

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2003年11月27日 (27.11.2003)

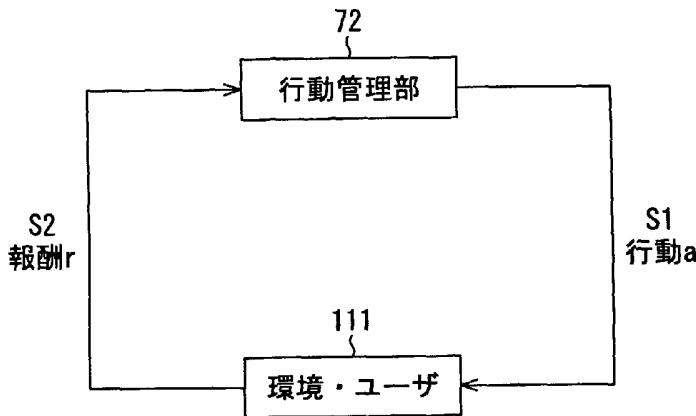
PCT

(10) 国際公開番号
WO 03/097308 A1

- (51) 国際特許分類: B25J 13/00, 5/00
- (21) 国際出願番号: PCT/JP03/06179
- (22) 国際出願日: 2003年5月19日 (19.05.2003)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ: 特願2002-145334 2002年5月20日 (20.05.2002) JP
- (71) 出願人 (米国を除く全ての指定国について): ソニー株式会社 (SONY CORPORATION) [JP/JP]; 〒141-0001 東京都品川区北品川6丁目7番35号 Tokyo (JP).
- (72) 発明者; および
- (75) 発明者/出願人 (米国についてのみ): 澤田 務 (SAWADA, Tsutomu) [JP/JP]; 〒141-0001 東京都品川区北品川6丁目7番35号 ソニー株式会社内 Tokyo (JP). 藤田 雅博 (FUJITA, Masahiro) [JP/JP]; 〒141-0001 東京都品川区北品川6丁目7番35号 ソニー株
- (54) 式会社内 Tokyo (JP). 花形 理 (HANAGATA, Osamu) [JP/JP]; 〒141-0001 東京都品川区北品川6丁目7番35号 ソニー株式会社内 Tokyo (JP). 高木 剛 (TAKAGI, Tsuyoshi) [JP/JP]; 〒141-0001 東京都品川区北品川6丁目7番35号 ソニー株式会社内 Tokyo (JP).
- (74) 代理人: 稲本 義雄 (INAMOTO, Yoshio); 〒160-0023 東京都新宿区西新宿7丁目11番18号 711ビルディング4階 Tokyo (JP).
- (81) 指定国 (国内): US.
- (84) 指定国 (広域): ヨーロッパ特許 (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR).
- 添付公開書類:
— 国際調査報告書
- 2文字コード及び他の略語については、定期発行される各PCTガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

(54) Title: ROBOT DEVICE AND ROBOT CONTROL METHOD

(54) 発明の名称: ロボット装置およびロボット制御方法



(57) Abstract: A robot device and a robot control method which enable a robot to behave without users feeling bored with. A behavior control unit (72) selects between a searching behavior and a knowledge-utilizing behavior based on a behavior selection probability. When a behavior (a) selected by the behavior control unit (72) is executed, a reward (r) is given by a environment/user (111). The unit (72) updates a behavior value based on the reward (r), and applies the behavior value to Boltzmann distribution to determine a behavior selection probability. The unit (72) also changes Boltzmann temperature based on an input information. This invention is applicable to a robot reinforced learning system.

S2...REWARD r 72...BEHAVIOR CONTROL UNIT

S1...BEHAVIOR a 111...ENVIRONMENT/USER

WO 03/097308 A1



(57) 要約:

本発明は、ロボットに、ユーザを飽きさせないように行動させることができるようにしたロボット装置およびロボット制御方法に関する。行動管理部(72)は、行動選択確率に基づいて、探索的行動と知識利用的行動を選択する。行動管理部(72)が、選択した行動(a)を実行すると、環境・ユーザ(111)から報酬(r)が与えられる。行動(72)は、報酬(r)に基づき、行動価値を更新し、その行動価値を、ボルツマン分布に適用して、行動選択確率を決定する。また、行動管理部(72)は、ボルツマン温度を入力情報に基づいて変化させる。本発明は、ロボットの強化学習システムに適用できる。

明細書

ロボット装置およびロボット制御方法

技術分野

- 5 本発明は、ロボット装置およびロボット制御方法に関し、特に、ロボットに、人間と同じように、かつ、ユーザを飽きさせないように、行動させることができるようにしたロボット装置およびロボット制御方法に関する。

背景技術

- 10 生命体を模擬するロボット装置とインタラクションする場合、ユーザは、ロボット装置に対して「人と同じような」振る舞いを期待している。

ロボット装置にこのような行動を実行させるには、ロボット装置に行動獲得のための強化学習を行わせればよい。ロボット装置は、行動価値から算出される行動選択確率に基づいて、行動を選択する。選択する行動としては、探査を主とした行動（以下、探査的行動と称する）と知識利用を主とした行動（以下、知識利

15 用的行動と称する）がある。探査的行動が選択されると、ロボット装置は今までと異なる新たな行動をとるようになる。これに対して、知識的行動が選択されると、ロボット装置は、一般的に期待される通りの行動をとる。

- 20 しかしながら、このような従来の強化学習では、行動選択確率は、ロボット装置の外部状態や内部状態によらず一定であった。

したがって、ロボット装置に、人と同じように、かつ、ユーザを飽きさせないように、振る舞わせることは困難であった。

発明の開示

- 25 本発明はこのような状況に鑑みてなされたものであり、ロボット装置に、人と同じように行動させつつ、ユーザを飽きさせないように行動させることができるようにすることを目的とする。

本発明のロボット装置は、過去の行動に関する知識に基づく行動と、過去の行動に関する知識に基づかない行動の、いずれかの行動を選択する行動管理手段と、ロボット装置を制御する制御手段とを備え、制御手段は、行動管理手段により選択された過去の行動に関する知識に基づく行動、または過去の行動に関する知識

5 に基づかない行動に基づいて、ロボット装置を制御することにより、所定の行動を出力することを特徴とする。

過去の行動に関する知識に基づく行動は、知識利用的行動であり、過去の行動に関する知識に基づかない行動は、探査的行動であるようにすることができる。

行動管理手段は、知識利用的行動または探査的行動のいずれかを行動選択確率

10 に基づいて選択するようすることができる。

行動選択確率は、ロボット装置が実行した行動に対するユーザからの報酬により変動するようすることができる。

行動選択確率は、ボルツマン温度に基づいて変動するようすることができる。

ロボット装置は、ロボット装置の感情を決定する感情パラメータを保持し、ボルツマン温度は、感情パラメータに基づいて変動するようすることができる。

15

本発明のロボット制御方法は、過去の行動に関する知識に基づく行動と、過去の行動に関する知識に基づかない行動の、いずれかの行動を選択する行動管理ステップと、ロボット装置を制御する制御ステップとを含み、制御ステップの処理は、行動管理ステップの処理により選択された過去の行動に関する知識に基づく

20 行動、または過去の行動に関する知識に基づかない行動に基づいて、ロボット装置を制御することにより、所定の行動を出力することを特徴とする。

本発明のロボット装置およびロボット制御方法においては、過去の行動に関する知識に基づく行動と、過去の行動に関する知識に基づかない行動の、いずれかの行動が選択され、その選択された過去の行動に関する知識に基づく行動、または過去の行動に関する知識に基づかない行動に基づいて、ロボット装置が制御されることにより、所定の行動が出力される。

25

図面の簡単な説明

図 1 は、本発明を適用したペットロボットの外観の構成例を示す斜視図である。

図 2 は、図 1 のペットロボットの内部の構成例を示すブロック図である。

図 3 は、ペットロボットの機能モジュールの構成例を示す図である。

5 図 4 は、図 3 の本能・感情管理部の機能の例を模式的に示す図である。

図 5 は、本発明を適用した強化学習システムの構成を示す図である。

図 6 は、感情パラメータに基づいてボルツマン温度を演算する処理を説明するためのフローチャートである。

図 7 は、感情パラメータに基づくボルツマン温度の変化の例を示す図である。

10 図 8 は、本能パラメータに基づいてボルツマン温度を演算する処理を説明するためのフローチャートである。

図 9 は、本能パラメータに基づくボルツマン温度の変化の例を示す図である。

図 10 は、場所に基づいてボルツマン温度を決定する処理を説明するためのフローチャートである。

15 図 11 は、場所に基づくボルツマン温度の変化の例を示す図である。

図 12 は、時刻に基づいてボルツマン温度を演算する処理を説明するためのフローチャートである。

図 13 は、時刻に基づくボルツマン温度の変化の例を示す図である。

20 図 14 は、成長時刻に基づいてボルツマン温度を演算する処理を説明するためのフローチャートである。

図 15 は、成長時刻に基づくボルツマン温度の変化の例を示す図である。

発明を実施するための最良の形態

以下、本発明の実施の形態について、図面を参照して説明する。図 1 は、本発明を適用したペットロボット 1 の例を示す斜視図である。

例えば、ペットロボット 1 は、4 本足の小熊形状のものとされており、胴体部ユニット 2 の前後左右に、それぞれ脚部ユニット 3 A、3 B、3 C、3 D が連結

されるとともに、胴体部ユニット 2 の前端部と後端部に、それぞれ頭部ユニット 4 と尻尾部ユニット 5 が連結されている。

図 2 は、図 1 のペットロボット 1 の内部構成の例を示すブロック図である。胴体部ユニット 2 には、ペットロボット 1 の全体を制御するコントローラ 10、ペットロボット 1 の各部に電力を供給するバッテリー 11、並びにバッテリーセンサ 12、および熱センサ 13 からなる内部センサ 14 が格納されている。このコントローラ 10 には、CPU (Central Processing Unit) 10A、CPU 10A が各部を制御するためのプログラムが記憶されているメモリ 10B、および計時動作を行い、現在の日時、起動後の経過時間等を計測する時計 10C が設けられている。

10 また、CPU 10A には、インターネットに代表されるネットワークを介してデータを通信する通信部 63、プログラムなどの各種データを格納する半導体メモリなどよりなる記憶部 62 が接続されている。さらに、リムーバブルメモリ 61 などの記録媒体に対してデータを読み書きするドライブ 60 が必要に応じて接続される。

15 このペットロボット 1 に本発明を適用したロボット装置としての動作を実行させるロボット制御プログラムは、リムーバブルメモリ 61 に格納された状態でペットロボット 1 に供給され、ドライブ 60 によって読み出されて、記憶部 62 に内蔵されるハードディスクドライブにインストールされる。記憶部 62 にインストールされたロボット制御プログラムは、ユーザから入力されるコマンドに対応する CPU 10A の指令によって、記憶部 62 からメモリ 10B にロードされて実行される。

25 頭部ユニット 4 には、外部からの刺激を感知するセンサとして、音を感知する「耳のような聴覚器官」に相当するマイクロフォン 15、CCD (Charge Coupled Device)、CMOS (Complementary Metal Oxide Semiconductor)、およびイメージセンサなどから構成され、外部の画像信号を取得する「目のような視覚器官」に相当するビデオカメラ 16、およびユーザが接触することによる圧力等を感知する「肌等のような触覚器官」に相当するタッチセンサ 17 が、それぞれ所定の

位置に設けられている。また、頭部ユニット4には、対象物までの距離を測定する位置検出センサ18、および所定の音階を出力するペットロボット1の「口のような発声器官」に相当するスピーカ19が、それぞれ所定の位置に設置されている。

5 脚部ユニット3A乃至3Dのそれぞれの関節部分、脚部ユニット3A乃至3Dのそれぞれと胴体部ユニット2の連結部分、頭部ユニット4と胴体部ユニット2の連結部分、並びに尻尾部ユニット5と胴体部ユニット2の連結部分などには、アクチュエータが設置されている。アクチュエータは、コントローラ10からの指示に基づいて各部を動作させる。

10 図2の例においては、脚部ユニット3Aには、アクチュエータ $3A A_1$ 乃至 $3A A_k$ が設けられ、脚部ユニット3Bには、アクチュエータ $3B A_1$ 乃至 $3B A_k$ が設けられている。また、脚部ユニット3Cには、アクチュエータ $3C A_1$ 乃至 $3C A_k$ が設けられ、脚部ユニット3Dには、アクチュエータ $3D A_1$ 乃至 $3D A_k$ が設けられている。さらに、頭部ユニット4には、アクチュエータ $4 A_1$ 乃至 $4 A_l$ が設けられており、尻尾部ユニット5には、アクチュエータ $5 A_1$ および $5 A_2$ がそれぞれ設けられている。

以下、脚部ユニット3A乃至3Dに設けられているアクチュエータ $3A A_1$ 乃至 $3D A_k$ 、頭部ユニット4に設けられているアクチュエータ $4 A_1$ 乃至 $4 A_l$ 、および尻尾部ユニットに設けられているアクチュエータ $5 A_1$ および $5 A_2$ のそれぞれを個々に区別する必要がない場合、適宜、まとめて、アクチュエータ $3A A_1$ 乃至 $5 A_2$ と称する。

さらに、脚部ユニット3A乃至3Dには、アクチュエータの他にスイッチ3AB乃至3DBが、ペットロボット1の足の裏に相当する場所に設置されている。そして、ペットロボット1が歩行したとき、スイッチ3AB乃至3DBが押下され、それを表す信号がコントローラ10に入力されるようになっている。

25 頭部ユニット4に設置されるマイクロフォン15は、ユーザの発話を含む周囲の音声（音）を集音し、得られた音声信号をコントローラ10に出力する。ビデ

オカメラ16は、周囲の状況を撮像し、得られた画像信号を、コントローラ10に出力する。タッチセンサ17は、例えば、頭部ユニット4の上部に設けられており、ユーザからの「撫でる」や「叩く」といった物理的な働きかけにより受けた圧力を検出し、その検出結果を圧力検出信号としてコントローラ10に出力する。位置検出センサ18は、例えば、赤外線を出射し、その反射光を受光したタイミングにおいての検出結果をコントローラ10に出力する。

コントローラ10は、マイクロフォン15、ビデオカメラ16、タッチセンサ17、および位置検出センサ18から与えられる音声信号、画像信号、圧力検出信号等に基づいて、周囲の状況や、ユーザからの指令、ユーザからの働きかけなどの有無を判断し、その判断結果に基づいて、ペットロボット1が次に実行する動作を決定する。そして、コントローラ10は、その決定に基づいて、必要なアクチュエータを駆動させ、これにより、頭部ユニット4を上下左右に振らせたり、尻尾部ユニット5を動かしたり、脚部ユニット3A乃至3Dのそれぞれを駆動して、ペットロボット1を歩行させるなどの動作を実行させる。

その他にも、コントローラ10は、ペットロボット1の頭部ユニット4などに設けられた、図示しないLED (Light Emitting Diode) を点灯、消灯または点滅させるなどの処理を行う。

図3は、図2のコントローラ10の機能的構成例を示すブロック図である。なお、図3に示す各機能は、CPU10Aがメモリ10Bに記憶されている制御プログラムを実行することによって実現される。

コントローラ10は、外部からの刺激を検知するセンサ（マイクロフォン15乃至位置検出センサ18、およびスイッチ3AB乃至3DB）からの各種信号を検出するセンサ入力処理部31と、センサ入力処理部31により検出された情報等に基づいて、ペットロボット1を動作させる情報処理部32から構成されている。

センサ入力処理部31を構成する角度検出部41は、アクチュエータ3AA₁乃至5A₂のそれぞれに設けられるモータが駆動されたとき、アクチュエータ3

A A₁ 乃至 5 A₂ のそれぞれから通知される情報に基づいて、その角度を検出する。角度検出部 4 1 により検出された角度情報は、情報処理部 3 2 の行動管理部 7 2、および音データ生成部 7 5 に出力される。

5 音量検出部 4 2 は、マイクロフォン 1 5 から供給される信号に基づいて、その音量を検出し、検出した音量情報を行動管理部 7 2、および音データ生成部 7 5 に出力する。

10 音声認識部 4 3 は、マイクロフォン 1 5 から供給される音声信号について音声認識を行う。音声認識部 4 3 は、その音声認識結果としての、例えば、「お話しよう」、「歩け」、「伏せ」、「ボールを追いかける」等の指令その他を、音声認識情報として、本能・感情管理部 7 1、行動管理部 7 2、および音データ生成部 7 5 に通知する。

15 画像認識部 4 4 は、ビデオカメラ 1 6 から供給される画像信号を用いて、画像認識を行う。画像認識部 4 4 は、その処理の結果、例えば、「赤い丸いもの」、「地面に対して垂直な、かつ、所定の高さ以上の平面」、「広い開放的な場所」、「家族がいる」、「家族の中の子供の友人がいる」等を検出したときには、「ボールがある」、「壁がある」、「畑である」、「家である」、「学校である」等の画像認識結果を、画像認識情報として、本能・感情管理部 7 1、行動管理部 7 2、および音データ生成部 7 5 に通知する。

20 圧力検出部 4 5 は、タッチセンサ 1 7 から与えられる圧力検出信号を処理する。例えば、圧力検出部 4 5 は、その処理の結果、所定の閾値以上で、かつ、短時間の圧力を検出したときには、「叩かれた（しかられた）」と認識し、所定の閾値未満で、かつ、長時間の圧力を検出したときには、「なでられた（ほめられた）」と認識して、その認識結果を、状態認識情報として、本能・感情管理部 7 1、行動管理部 7 2、および音データ生成部 7 5 に通知する。

25 位置検出部 4 6 は、位置検出センサ 1 8 から供給される信号に基づいて、所定の対象物までの距離を測定し、その距離情報を行動管理部 7 2、および音データ生成部 7 5 に通知する。例えば、位置検出部 4 6 は、目の前にユーザの手などが

差し出されたとき、その手までの距離や、画像認識部 4 4 により認識されたボールまでの距離を検出する。

5 スイッチ入力検出部 4 7 は、ペットロボット 1 の足の裏に相当する部分に設けられているスイッチ 3 A B 乃至 3 D B から供給される信号に基づき、例えば、ペットロボット 1 が歩行している状態において、その歩行タイミングや、ユーザにより足の裏が触れられたことを本能・感情管理部 7 1、および行動管理部 7 2 に通知する。

10 一方、情報処理部 3 2 を構成する本能・感情管理部 7 1 は、ペットロボット 1 の本能、および感情を管理し、所定のタイミングで、ペットロボット 1 の本能を表すパラメータや、感情を表すパラメータを行動管理部 7 2、および音データ生成部 7 5 に出力する。

15 ペットロボット 1 の本能を表すパラメータと感情を表すパラメータについて、図 4 を参照して説明する。図 4 に示すように、本能・感情管理部 7 1 は、ペットロボット 1 の感情を表現する感情モデル 1 0 1 と、本能を表現する本能モデル 1 0 2 を記憶し、管理している。

20 感情モデル 1 0 1 は、例えば、「うれしさ」、「悲しさ」、「怒り」、「驚き」、「恐れ」、「苛立ち」等の感情の状態（度合い）を、所定の範囲（例えば、0 乃至 1 0 0 等）の感情パラメータによってそれぞれ表し、センサ入力処理部 3 1 の音声認識部 4 3、画像認識部 4 4、および圧力検出部 4 5 からの出力や時間経過等に基づいて、その値を変化させる。

この例において、感情モデル 1 0 1 は、「うれしさ」を表す感情ユニット 1 0 1 A、「悲しさ」を表す感情ユニット 1 0 1 B、「怒り」を表す感情ユニット 1 0 1 C、「驚き」を表す感情ユニット 1 0 1 D、「恐れ」を表す感情ユニット 1 0 1 E、および「苛立ち」を表す感情ユニット 1 0 1 F から構成されている。

25 本能モデル 1 0 2 は、例えば、「運動欲」、「愛情欲」、「食欲」、「好奇心」、「睡眠欲」等の本能による欲求の状態（度合い）を、所定の範囲（例えば、0 乃至 1 0 0 等）の本能のパラメータによってそれぞれ表し、音声認識部 4 3、

画像認識部 4 4、および圧力検出部 4 5等からの出力や時間経過等に基づいて、その値を変化させる。また、本能モデル 1 0 2 は、行動履歴に基づいて、「運動欲」を表すパラメータを高めたり、バッテリー 1 1 の残量（電圧）に基づいて、「食欲」を表すパラメータを高めたりする。

- 5 この例において、本能モデル 1 0 2 は、「運動欲」を表す本能ユニット 1 0 2 A、「愛情欲」を表す本能ユニット 1 0 2 B、「食欲」を表す本能ユニット 1 0 2 C、「好奇心」を表す本能ユニット 1 0 2 D、および「睡眠欲」を表すユニット 1 0 2 E から構成されている。

- 10 このような感情ユニット 1 0 1 A 乃至 1 0 1 F と本能ユニット 1 0 2 A 乃至 1 0 2 E のパラメータは、外部からの入力だけでなく、図の矢印で示されるように、それぞれのユニット同士が相互に影響し合うことによっても変化される。

- 例えば、「うれしさ」を表現する感情ユニット 1 0 1 A と「悲しさ」を表現する感情ユニット 1 0 1 B が相互抑制的に結合することにより、本能・感情管理部 7 1 は、ユーザにほめてもらったときには「うれしさ」を表現する感情ユニット 1 0 1 A のパラメータを大きくするとともに、「悲しさ」を表現する感情ユニット 1 0 1 B のパラメータを小さくするなどして、表現する感情を変化させる。

また、感情モデル 1 0 1 を構成する各ユニット同士、および本能モデル 1 0 2 を構成する各ユニット同士だけでなく、双方のモデルを超えて、それぞれのユニットのパラメータが変化される。

- 20 例えば、図 4 に示されるように、本能モデル 1 0 2 の「愛情欲」を表す本能ユニット 1 0 2 B や、「食欲」を表す本能ユニット 1 0 2 C のパラメータの変化に応じて、感情モデル 1 0 1 の「悲しさ」を表現する感情ユニット 1 0 2 B や「怒り」を表現する感情ユニット 1 0 2 C のパラメータが変化される。

- 25 具体的には、「愛情欲」を表す本能ユニット 1 0 2 B のパラメータ、または「食欲」を表す本能ユニット 1 0 2 C のパラメータが大きくなったとき、感情モデル 1 0 1 の「悲しさ」を表現する感情ユニット 1 0 1 B のパラメータと「怒り」を表現する感情ユニット 1 0 1 C のパラメータが小さくなる。

このようにして、本能・感情管理部 7 1 により、管理される感情のパラメータ、または本能のパラメータは、所定の周期で計測され、行動管理部 7 2、および音データ生成部 7 5 に出力される。

5 なお、本能・感情管理部 7 1 には、音声認識部 4 3、画像認識部 4 4、および圧力検出部 4 5 等から認識情報が供給される他に、行動管理部 7 2 から、ペットロボット 1 の現在、または過去の行動、具体的には、例えば、「長時間歩いた」などの行動の内容を示す行動情報が供給されるようになされている。そして、本能・感情管理部 7 1 は、同一の認識情報等が与えられた場合であっても、行動情報により示されるペットロボット 1 の行動に応じて、異なる内部情報を生成する。

10 例えば、ペットロボット 1 がユーザに挨拶をし、ユーザに頭をなでられた場合には、ユーザに挨拶をしたという行動情報と、頭をなでられたという認識情報が本能・感情管理部 7 1 に供給される。このとき、本能・感情管理部 7 1 においては、「うれしさ」を表す感情ユニット 1 0 1 A の値が増加される。

図 3 の説明に戻り、行動管理部 7 2 は、音声認識部 4 3、および画像認識部 4 4 等から供給されてきた情報と、本能・感情管理部 7 1 から供給されてきたパラメータ、および時間経過等に基づいて次の行動を決定し、決定した行動の実行を指示するコマンドを姿勢遷移管理部 7 3 に出力する。姿勢遷移管理部 7 3 は、行動管理部 7 2 から指示された行動に基づいて、姿勢の遷移を決定し、制御部 7 4 に出力する。制御部 7 4 は、姿勢遷移管理部 7 3 からの出力に基づき、アクチュエータ 3 A A 1 乃至 5 A 2 を制御して、行動管理部 7 2 が決定した動作を行う。

20 また、音データ生成部 7 5 は、音声認識部 4 3、および画像認識部 4 4 等から供給されてきた情報と、本能・感情管理部 7 1 から供給されてきたパラメータ、および時間経過等に基づいて音データを生成する。そして、行動管理部 7 2 は、ペットロボット 1 に発話をさせるとき、あるいは所定の動作に対応する音をスピーカ 1 9 から出力させるとき、音の出力を指示するコマンドを音声合成部 7 6 に出力し、音声合成部 7 6 は、音データ生成部 7 5 から出力された音データに基づいて、スピーカ 1 9 に音を出力させる。

図5は、本発明を適用した強化学習システムの原理的構成を示している。ステップS1において、行動管理部72は、行動選択確率（後述する式(3)）に基づいて、ランダムに動き、新しい行動をする探査的行動と、過去の行動の中から最も良いと考えられる行動を選択する知識利用的行動のいずれかを選択し、選択された行動aをペットロボット1に実行させる。ステップS2において、環境・ユーザ111は、その行動aに対して、ペットロボット1に報酬rを与える。ここで、報酬rとは、行動aが正しかったときに環境・ユーザ111により行われる、例えば、「頭をなでる」行為や、行動aが誤っていた（ユーザが期待していない行動であった）ときに環境・ユーザ111により行われる、例えば、「頭をたたく」行為である。

行動管理部72は、獲得した報酬rに基づき、以下のような式(1)に従って、この行動aに対する行動価値Q(a)を、新しい（報酬rに基づく）行動価値 $Q_1(a)$ に更新する。

$$Q_1(a) = (1 - \alpha)Q(a) + \alpha r \quad \dots(1)$$

なお、式(1)において、 α は、0と1の間の値の係数であって、獲得した報酬をどの程度行動価値に反映させるかを決定するパラメータ（学習率）であり、報酬rは、「頭をなでられる」、「叩かれる」、「与えられたタスクを達成する」、「与えられたタスクを失敗する」等のセンサ入力処理部31からの検出信号に応じて決定される値である。

このようにして求められた行動価値Q(a)は、行動aを実行して、報酬rを獲得するごとに更新される。したがって、新しく獲得した報酬 r_2 によって、更新された行動価値 $Q_2(a)$ は、前の行動価値 $Q_1(a)$ とその前の行動価値 $Q_0(a)$ を使って表すと、

$$Q_2(a) = (1 - \alpha)Q_1(a) + \alpha r_2 = (1 - \alpha)^2 Q_0(a) + (1 - \alpha)\alpha r_1 + \alpha r_2 \quad \dots(2)$$

となり、学習率 α は0と1の間の値なので、新しく獲得した報酬 r_2 の係数 α は、前の報酬 r_1 の係数 $(1 - \alpha)\alpha$ よりも必ず大きくなる。したがって、行動

価値 $Q_2(a)$ は、過去の報酬 r_1 よりも新しく受け取った報酬 r_2 のほうに、より重み付けされる。即ち、式 (2) により、行動価値 $Q(a)$ は、遠い過去の報酬よりも、最近受け取った報酬が反映される。

さらに、行動管理部 72 は、このようにして求められた行動価値 $Q(a)$ に基づき、行動 a の行動選択確率 $P(a)$ を、以下に示すボルツマン分布の式 (3) から演算する。

$$P(a) = \frac{\exp\left(\frac{Q(a)}{T}\right)}{\sum_{a \in A} \exp\left(\frac{Q(a)}{T}\right)} \quad \dots (3)$$

なお、式 (3) において、 A は選択可能な全ての行動を表す。 T はボルツマン温度を表し、行動価値 $Q(a)$ を、どの程度行動選択確率 P に反映させるかを決定するパラメータである。

式 (3) より明らかなように、ボルツマン温度 T の値が小さいほど、行動価値 $Q(a)$ が、より行動選択確率 P に反映され、行動管理部 72 は、より知識利用的な行動を選択する。一方、ボルツマン温度 T の値が大きいほど、行動価値 $Q(a)$ が、行動選択確率 P に反映され難くなり、行動管理部 72 は、より探査的な行動を選択する。

本発明においては、探査的行動と知識利用的行動の選択が、動的に変更される。このため、行動管理部 72 は、本能・感情管理部 71 からの感情のパラメータの出力に応じて、ボルツマン温度 T を適宜変化させる。この場合の処理を図 6 を参照して説明する。

ステップ S11 において、行動管理部 72 は、本能・感情管理部 71 から、感情のパラメータを読み出す。ステップ S12 において、行動管理部 72 は、読み出した感情のパラメータに基づいて、メモリ 10B に記憶されている式またはテーブル (例えば、図 7 の例の場合、「苛立ち」とボルツマン温度 T との関係を表す式またはテーブル) に基づき、ボルツマン温度 T を演算する。

「苛立ち」を表すパラメータに基づいて演算されたボルツマン温度 T の例を図7に示す。図7に示されるように、「苛立ち」を表すパラメータの値が大きいほど、ボルツマン温度 T の値は大きくなる。したがって、「苛立ち」を表すパラメータの値が大きいほど、行動価値 $Q(a)$ は、行動選択確率 P に反映され難くなり、

5 ペットロボット1は、探査的な行動をする。即ち、ペットロボット1は、人間と同様に、苛立っているときは、理性的な行動をとりにくくなる。

また、本能・感情管理部71からの本能のパラメータの出力に応じて、ボルツマン温度 T を変化させるようにしてもよい。この場合の処理を図8を参照して説明する。

10 ステップS21において、行動管理部72は、本能・感情管理部71から、本能のパラメータを読み出す。ステップS22において、行動管理部72は、読み出した本能のパラメータに基づいて、メモリ10Bに記憶されている式またはテーブル（例えば、図9の例の場合、「好奇心」とボルツマン温度 T との関係を表す式またはテーブル）に基づき、ボルツマン温度 T を演算する。

15 「好奇心」を表すパラメータに基づいて演算されたボルツマン温度 T の例を図9に示す。図9の例の場合、「好奇心」を表すパラメータの値が大きいほど、ボルツマン温度 T の値は小さくなる。したがって、「好奇心」を表すパラメータの値が大きいほど、行動価値 $Q(a)$ は、行動選択確率 P に反映されやすくなり、

20 ペットロボット1は、知識利用的な行動をする。即ち、ペットロボット1は、人間と同様に、好奇心があるときは、その行動を集中して(熱中して)行う。

さらに、行動管理部72は、画像認識部44からの場所に関する情報に応じて、ボルツマン温度 T を変化させることができる。この場合の処理を図10を参照して説明する。

25 ステップS31において、行動管理部72は、画像認識部44の出力から、ペットロボット1が位置する場所を検出する。ステップS32において、行動管理部72は、ペットロボット1が位置する場所が学校であるか否かを判定する。ステップS32において、ペットロボット1の位置する場所が学校であると判定さ

れた場合、行動管理部 7 2 は、処理をステップ S 3 3 に進め、学校におけるボルツマン温度 T をメモリ 1 0 B から読み出す。

5 ステップ S 3 2 において、ペットロボット 1 が位置する場所が学校ではないと判定された場合、行動管理部 7 2 は、処理をステップ S 3 4 に進め、ペットロボット 1 が位置する場所が家であるか否かを判定する。ステップ S 3 4 において、ペットロボット 1 が位置する場所が家であると判定された場合、行動管理部 7 2 は、処理をステップ S 3 5 に進め、家におけるボルツマン温度 T をメモリ 1 0 B から読み出す。

10 ステップ S 3 4 において、ペットロボット 1 が位置する場所が家ではないと判定された場合、行動管理部 7 2 は、処理をステップ S 3 6 に進め、ペットロボット 1 が位置する場所が畑であるか否かを判定する。ステップ S 3 6 において、ペットロボット 1 が位置する場所が畑であると判定された場合、行動管理部 7 2 は、処理をステップ S 3 7 に進め、畑におけるボルツマンの温度 T をメモリ 1 0 B から読み出す。

15 ステップ S 3 3, S 3 5, または S 3 7 の処理の後、行動管理部 7 2 は、ステップ S 3 8 において、読み出したボルツマン温度 T を、新しいボルツマン温度 T として設定する。そして、処理は終了する。

20 ステップ S 3 6 において、ペットロボット 1 が位置する場所が畑ではないと判定された場合、行動管理部 7 2 は、処理をステップ S 3 9 に進め、エラー処理を行ない、処理を終了する。

25 場所によるボルツマン温度 T の値を示す例を、図 1 1 に示す。ペットロボット 1 が学校に位置するとき、ボルツマン温度 T の値は一番小さく、行動価値 $Q(a)$ は行動選択確率 P に反映され、ペットロボット 1 は、知識利用的な行動をする。即ち、ペットロボット 1 は、人間と同様に、学校にいるときは、決まった行動をする。

ペットロボット 1 が家に位置するときは、ボルツマン温度 T の値は平均的な値とされ、ペットロボット 1 は、知識利用的行動と探査的行動の両方をバランスよ

く行う。そして、ペットロボット1が畑に位置するとき、ボルツマン温度 T の値は、一番大きくなり、行動価値 $Q(a)$ は行動選択確率 P に反映され難くなる。したがって、ペットロボット1は、より探査的な行動をする。即ち、ペットロボット1は、人間と同様に、畑のような開放的な場所にいるときは、新しい行動を多
5 くする。

また、時刻に応じて、ボルツマン温度 T を変化させることもできる。この場合の処理を、図12を参照して説明する。

ステップS51において、行動管理部72は、時計10Cから、現在の時刻を読み出す。ステップS52において、行動管理部72は、その読み出された時刻
10 に基づいて、メモリ10Bに記憶されている式またはテーブル（例えば、図13の例の場合、「一日の流れの中の時刻」とボルツマン温度 T との関係を表す式またはテーブル）に基づき、ボルツマン温度 T を演算する。

行動管理部72が、時刻に基づいて、ボルツマン温度 T を変化させる例を図13に示す。朝起きたばかりの時刻 t_1 のとき、ボルツマン温度 T は、最も小さい
15 値とされ、ペットロボット1は、知識利用的な行動（決まった行動）をとり、新しい行動を行なわない。そして、時刻の経過とともに、徐々にボルツマン温度 T は上がっていき、朝食を食べた後の時刻 t_2 で、ボルツマン温度 T は、最大となる。したがって、朝食を食べた後の時間帯では、ペットロボット1は、探査的な行動（新しい行動）を行い、活発に動く。

朝食を食べて、活発に動いた後、眠くなる時刻 t_3 まで、ボルツマン温度 T は徐々に下がっていく。したがって、時刻 t_3 を中心とする眠くなる時間帯のとき、
20 ペットロボット1は、知識利用的な行動（決まった行動）をとり、活発な動きをしない（ただし、時刻 t_1 を中心とする時間帯よりは活発に活動する）。その後、夕方から夜にかけての時刻 t_4 まで、ボルツマン温度 T は、徐々に上がっていき、
25 ペットロボット1は、また、探査的な行動（新しい行動）を行う（時刻 t_2 の時間帯と時刻 t_3 の時間帯の中間のレベルの活動を行う）。

そして、寝る前の時刻 t_5 になるにつれて、ボルツマン温度 T は下がっていき、ペットロボット 1 は、徐々に知識利用的な行動（決まった行動）をとるようになっていく。そして、ペットロボット 1 は就寝する。このように、ペットロボット 1 は、人間と同様な行動を、1 日の時間の流れの中で行なう。

5 さらに、起動後の経過時刻（ペットロボット 1 の成長時刻）に応じて、ボルツマン温度 T を変化させてもよい。この場合の処理を、図 1 4 を参照して説明する。

ステップ S 6 1 において、行動管理部 7 2 は、時計 1 0 C から、ペットロボット 1 が生まれて（起動されて）からの時刻（成長時刻）を読み出す。ステップ S 6 2 において、行動管理部 7 2 は、その読み出された成長時刻に基づいて、メモリ 1 0 B に記憶されている式またはテーブル（例えば、図 1 5 の例の場合、「成長時刻」とボルツマン温度 T との関係を表す式またはテーブル）ボルツマン温度 T を演算する。

行動管理部 7 2 は、成長時刻 t に基づいて、例えば、以下に示す式（4）のようにボルツマン温度 T を変化させる。

$$T(t) = T_{\min} + (T_{\max} - T_{\min}) \times \text{sigmoid}(t - \tau) = T_{\min} + \frac{T_{\max} - T_{\min}}{1 + e^{\beta(t - \tau)}} \dots (4)$$

15 なお、式（4）において、 τ は基準時刻、 β は基準時刻付近での変化率を、それぞれ表す。また、 T_{\min} と T_{\max} は、それぞれ最小と最大のボルツマン温度を表す。

この式（4）を用いて計算されたボルツマン温度 T の変化を示す例を図 1 5 に示す。図 1 5 に示されるように、ペットロボット 1 が生まれたばかり（起動当初）（ $t = 0$ ）のときのボルツマン温度 T は、 T_{\max} で表される最大値とされ、探査的な行動（新しい行動）を多くする。その後、時刻（年齢）が経過すると、徐々にボルツマン温度 T は小さくなる。そして、ボルツマン温度 T は、最終的には T_{\min} で表される最小値に収束し、ペットロボット 1 は、知識利用的な行動（過去に学んだ行動）を行うようになる。即ち、人間と同様に、年をとるにつれて、

探査的な行動（新しい行動）をとりにくくなり、知識利用的行動（決まった行動）を行うようになる。予め設定してある基準時刻（基準年齢） t において、ボルツマン温度 T は、最大値と最小値の中間の値となる。

5 なお、上述した一連の処理は、図1に示したような動物型のペットロボットに実行させるだけでなく、例えば、2足歩行が可能な人間型ロボットや、コンピュータ内で活動する仮想ロボット等に実行させるようにしてもよい。また、本明細書において、ロボットには、人工エージェントも含まれる。

10 また、上述した一連の処理は、ハードウェアにより実行させることもできるが、ソフトウェアにより実行させることもできる。一連の処理をソフトウェアにより実行させる場合には、そのソフトウェアを構成するプログラムが、専用のハードウェアに組み込まれているロボット装置、または、各種のプログラムをインストールすることで、各種の機能を実行することが可能な、例えば汎用のロボット装置などに、ネットワークや記録媒体からインストールされる。

15 この記録媒体は、図2に示されるように、装置本体とは別に、ユーザにプログラムを提供するために配布される、プログラムが記録されているリムーバブルメモリ61などよりなるパッケージメディアにより構成されるだけでなく、装置本体に予め組み込まれた状態でユーザに提供される、プログラムが記録されているメモリ10Bに含まれるハードディスクなどで構成される。

20 なお、本明細書において、記録媒体に記録されるプログラムを記述するステップは、記載された順序に沿って時系列的に行われる処理はもちろん、必ずしも時系列的に処理されなくとも、並列的あるいは個別に実行される処理をも含むものである。

産業上の利用可能性

25 以上のように、本発明によれば、ロボット装置を動作させることができる。また、その行動を通して、ユーザに、より生命らしい擬似感覚を与えることができる。従って、ユーザが飽きない、ロボット装置を実現することができる。

請求の範囲

1. 所定の行動を出力するロボット装置において、
過去の行動に関する知識に基づく行動と、過去の行動に関する知識に基づかない行動の、いずれかの行動を選択する行動管理手段と、
- 5 前記ロボット装置を制御する制御手段とを備え、
前記制御手段は、前記行動管理手段により選択された前記過去の行動に関する知識に基づく行動、または前記過去の行動に関する知識に基づかない行動に基づいて、前記ロボット装置を制御することにより、前記所定の行動を出力することを特徴とするロボット装置。
- 10 2. 前記過去の行動に関する知識に基づく行動は、知識利用的行動であり、
前記過去の行動に関する知識に基づかない行動は、探査的行動であることを特徴とする請求の範囲第1項に記載のロボット装置。
3. 前記行動管理手段は、前記知識利用的行動または前記探査的行動のいずれかを行動選択確率に基づいて選択することを特徴とする請求の範囲第2項に記載のロボット装置。
- 15 4. 前記行動選択確率は、前記ロボット装置が実行した行動に対するユーザからの報酬により変動することを特徴とする請求の範囲第3項に記載のロボット装置。
5. 前記行動選択確率は、ボルツマン温度に基づいて変動することを特徴とする請求の範囲第3項に記載のロボット装置。
- 20 6. 前記ロボット装置は、前記ロボット装置の感情を決定する感情パラメータを保持し、
前記ボルツマン温度は、前記感情パラメータに基づいて変動することを特徴とする請求の範囲第5項に記載のロボット装置。
- 25 7. 所定の行動を出力するロボット装置のロボット制御方法において、
過去の行動に関する知識に基づく行動と、過去の行動に関する知識に基づかない行動の、いずれかの行動を選択する行動管理ステップと、

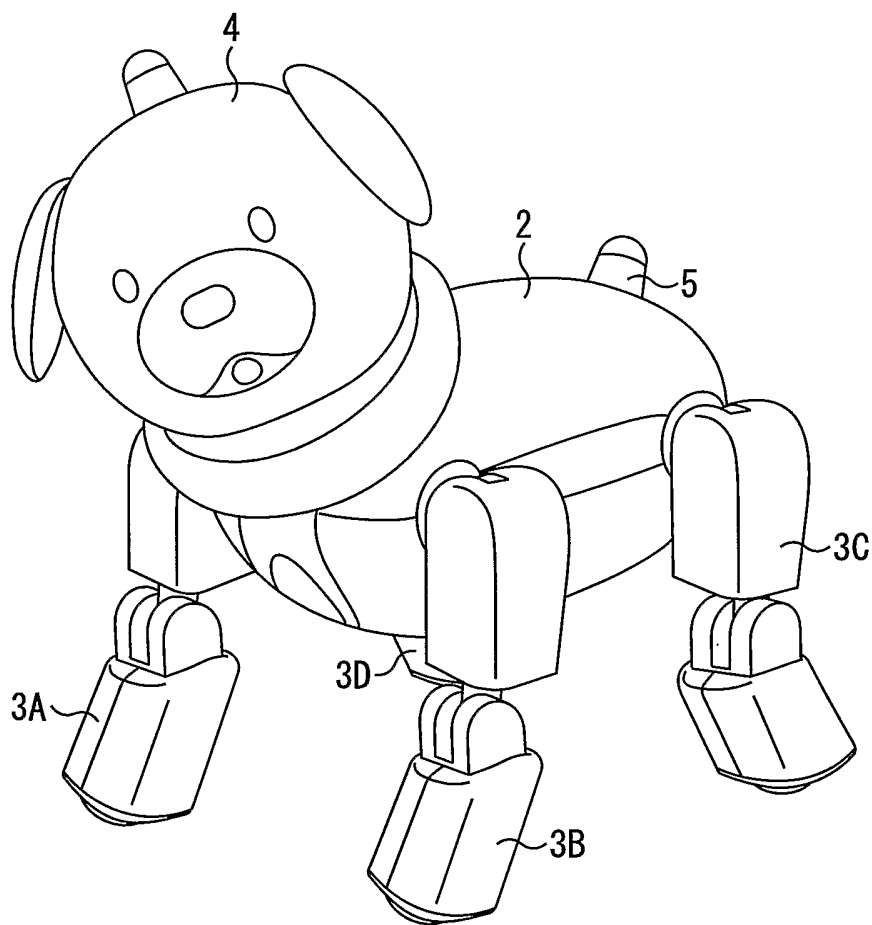
前記ロボット装置を制御する制御ステップとを含み、

前記制御ステップの処理は、前記行動管理ステップの処理により選択された前記過去の行動に関する知識に基づく行動、または前記過去の行動に関する知識に基づかない行動に基づいて、前記ロボット装置を制御することにより、前記所定

5 の行動を出力する

ことを特徴とするロボット制御方法。

図 1



1

図 2

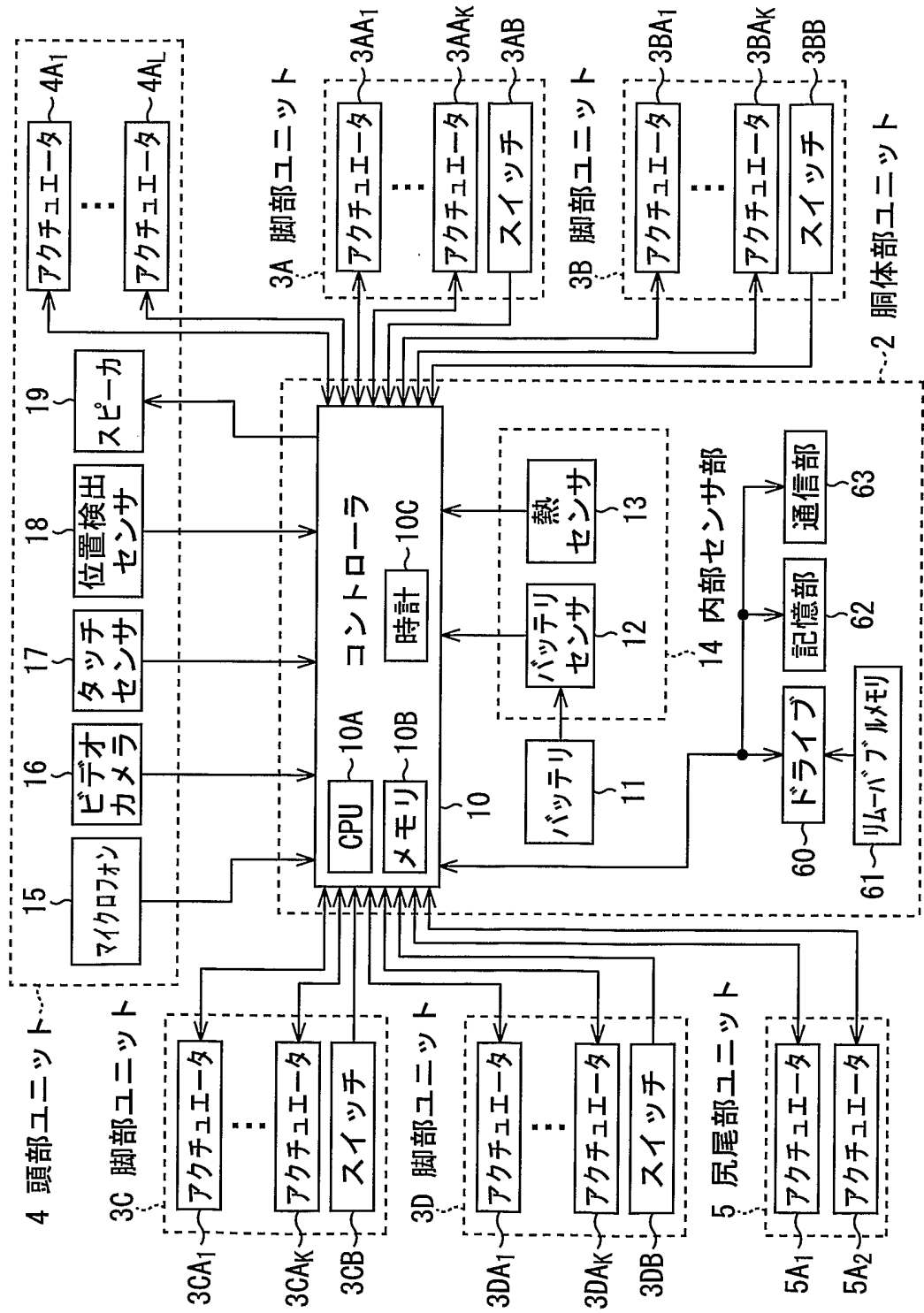


図 3

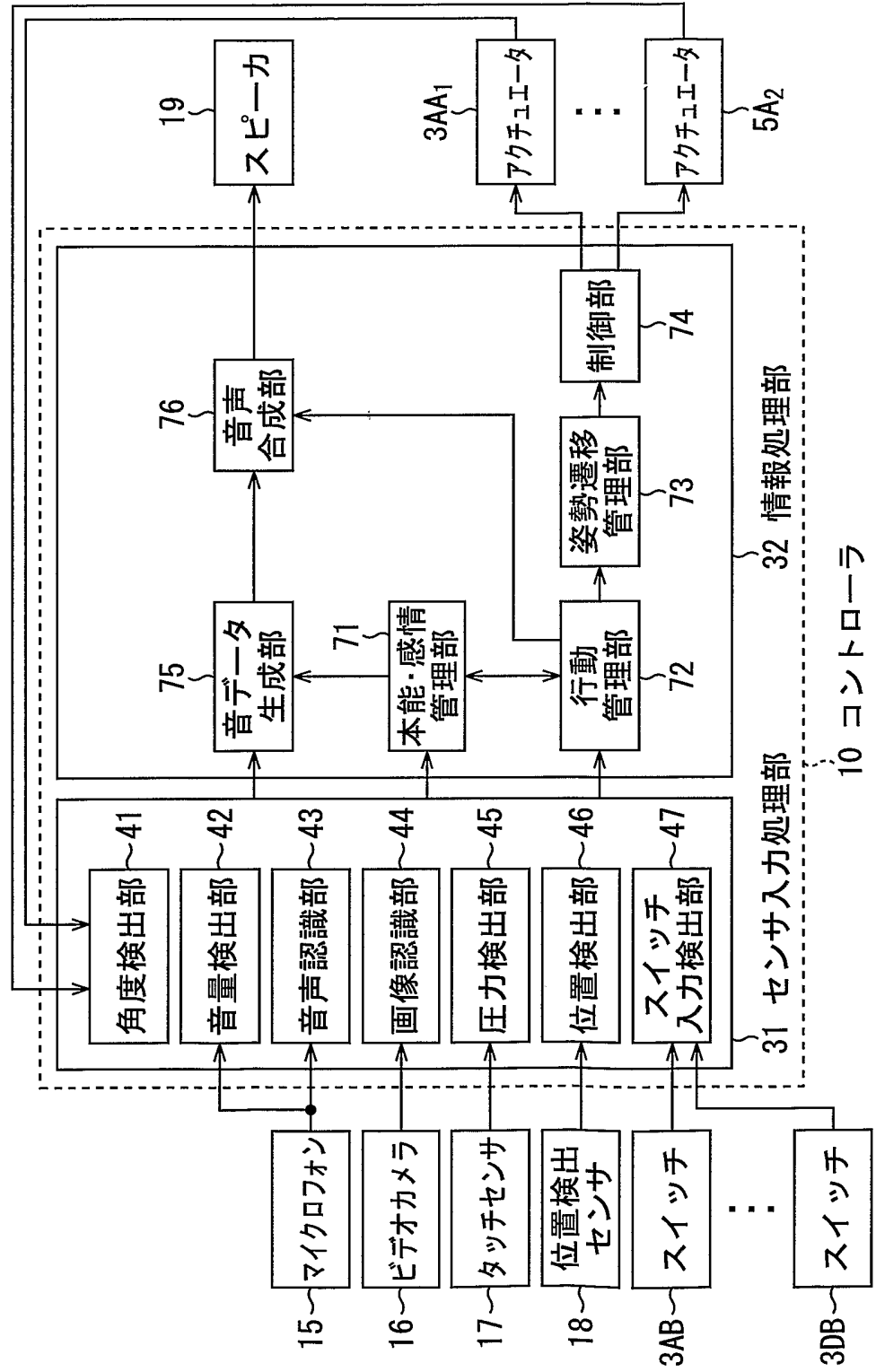


図 4

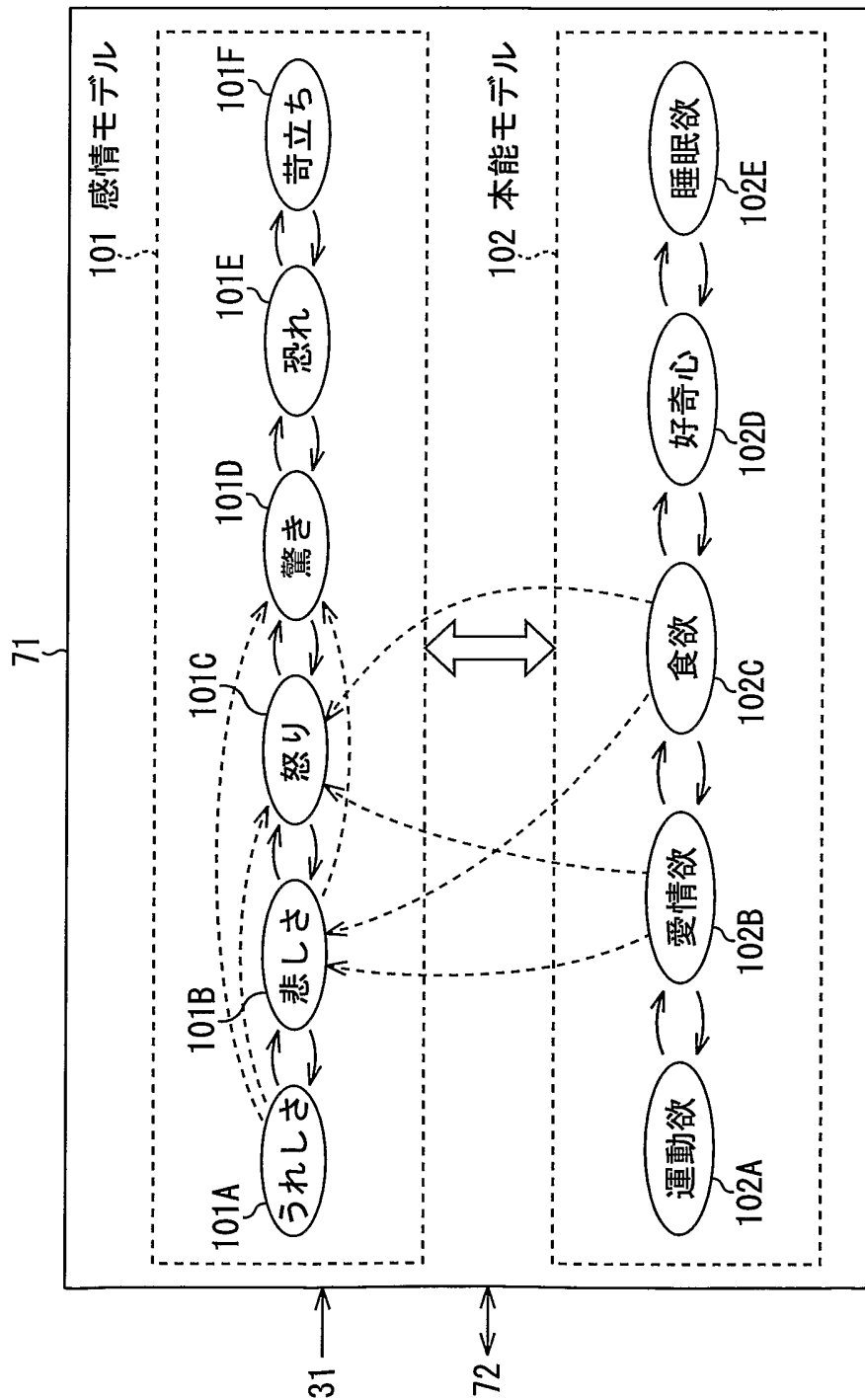


図 5

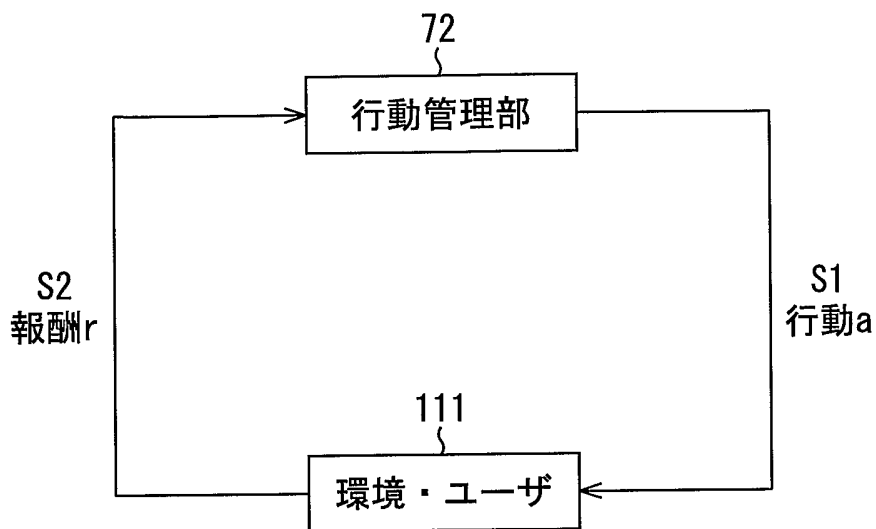


図 6

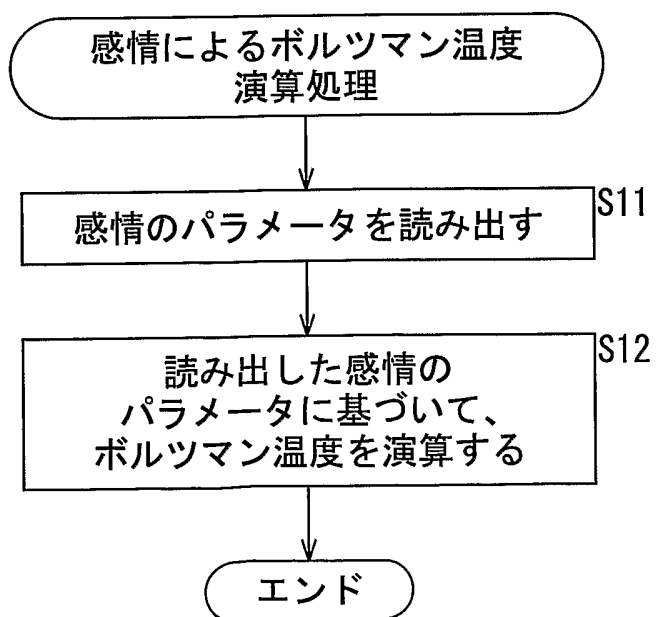


図 7

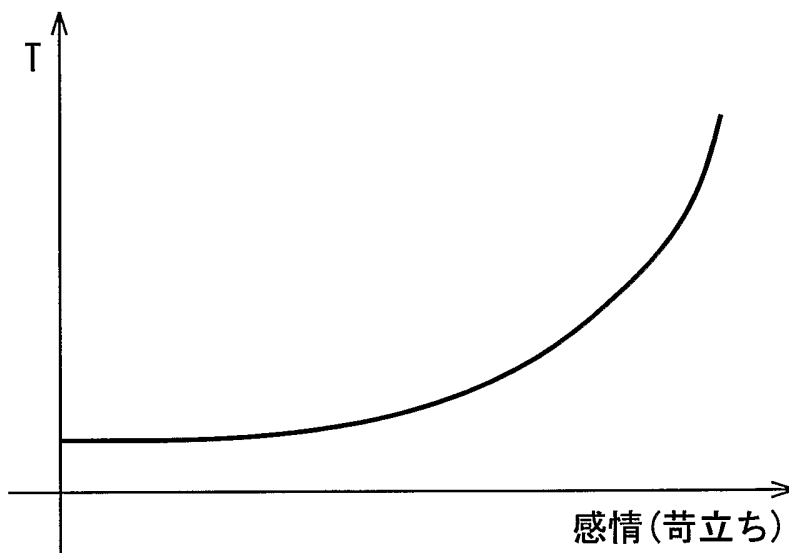


図 8

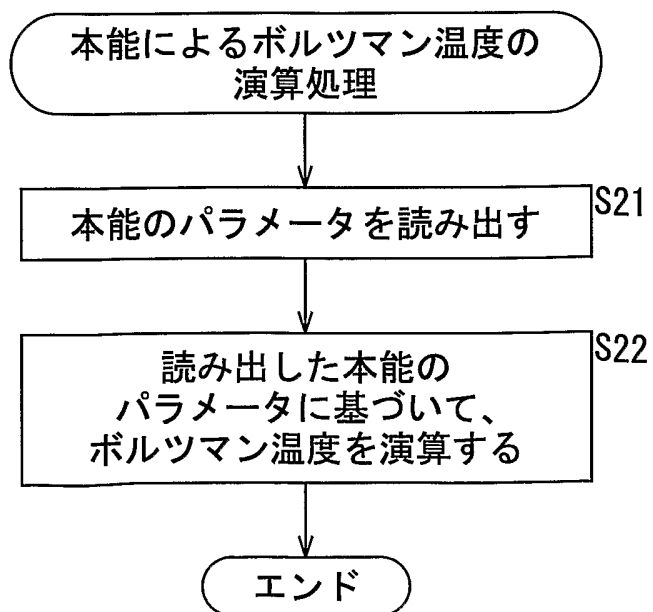


図 9

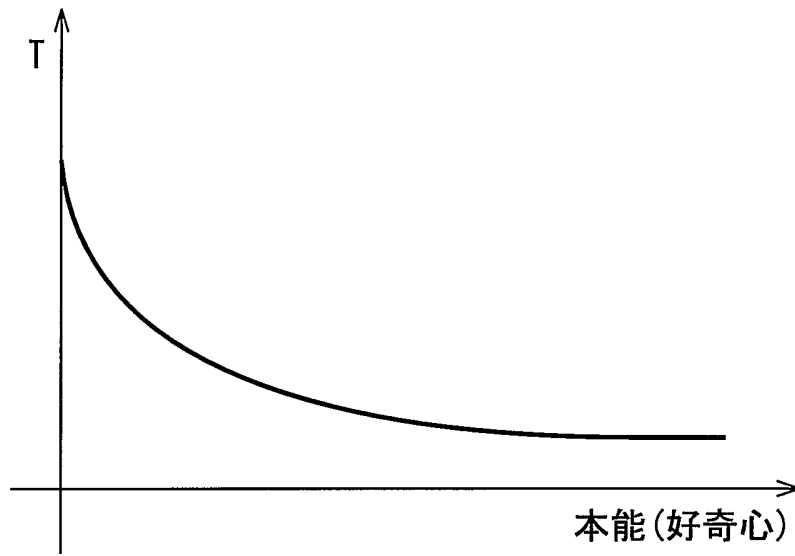


図10

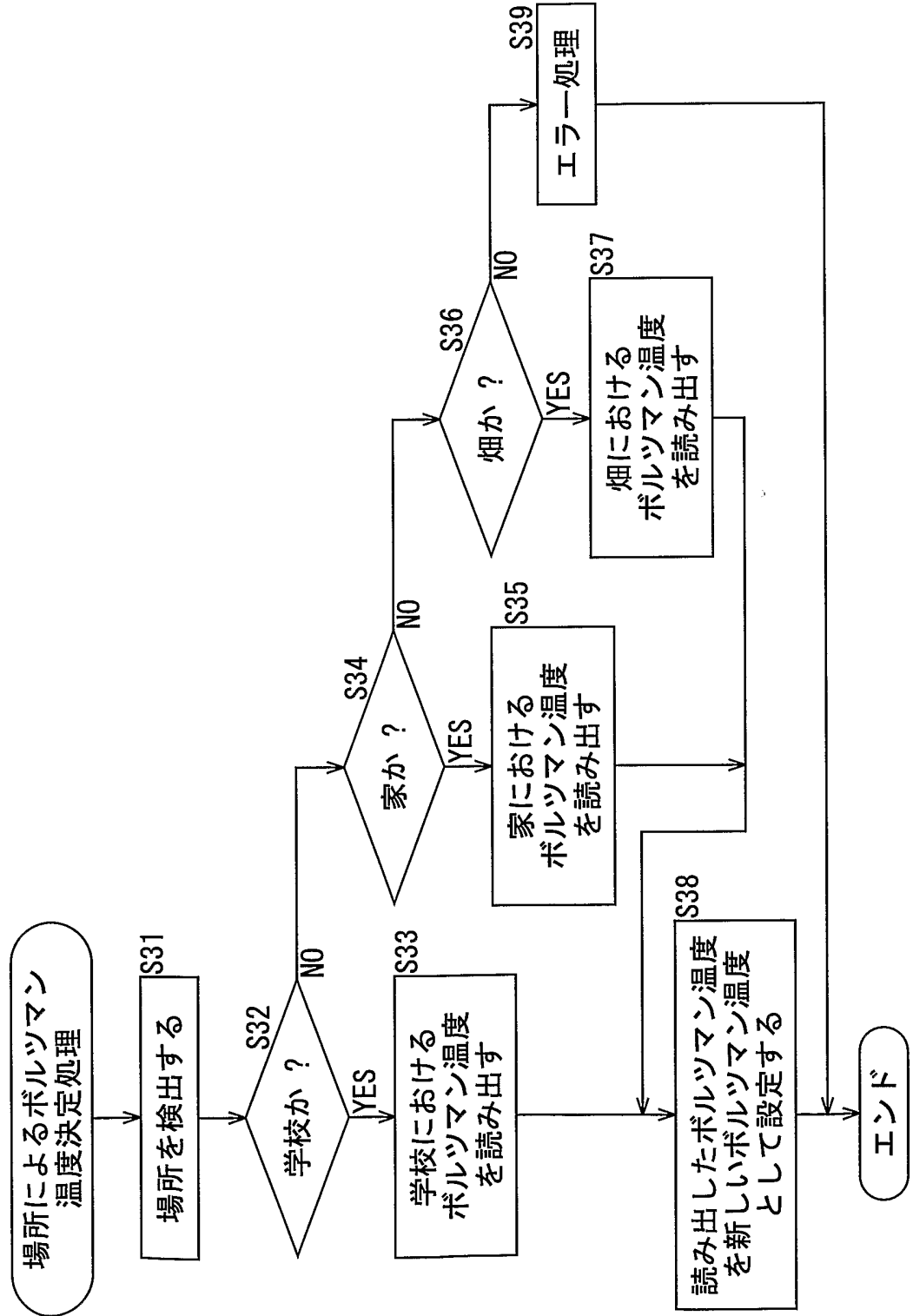


図11

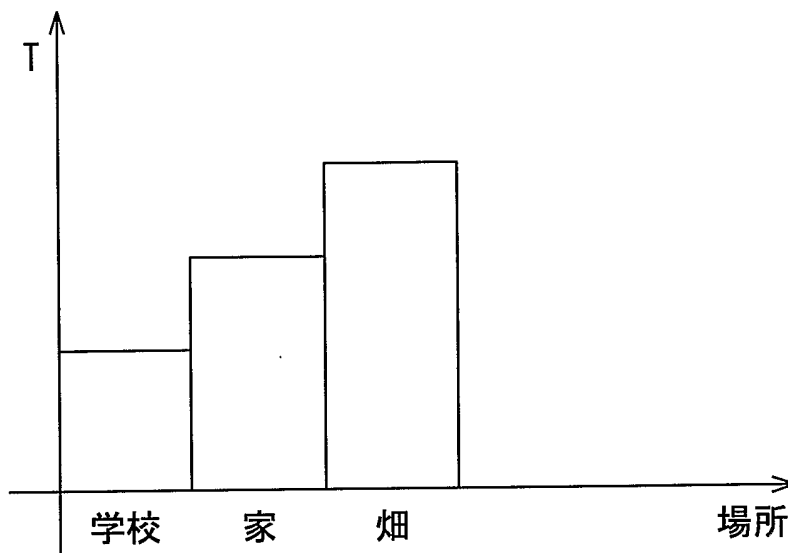


図12

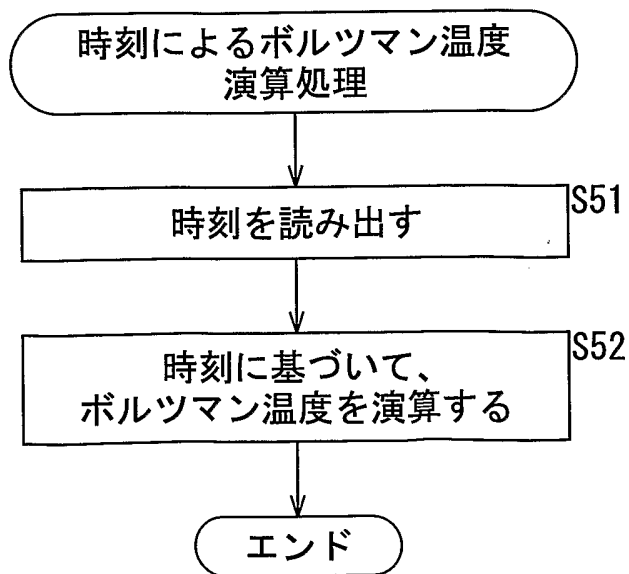


図13

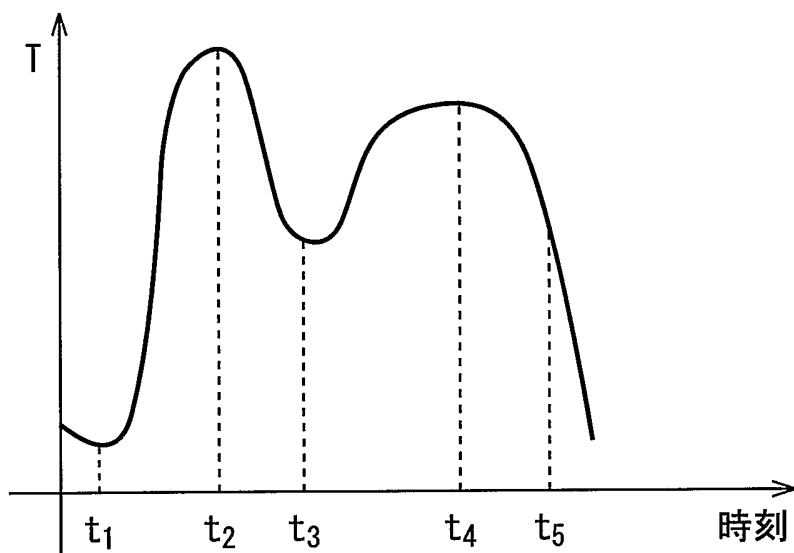


図14

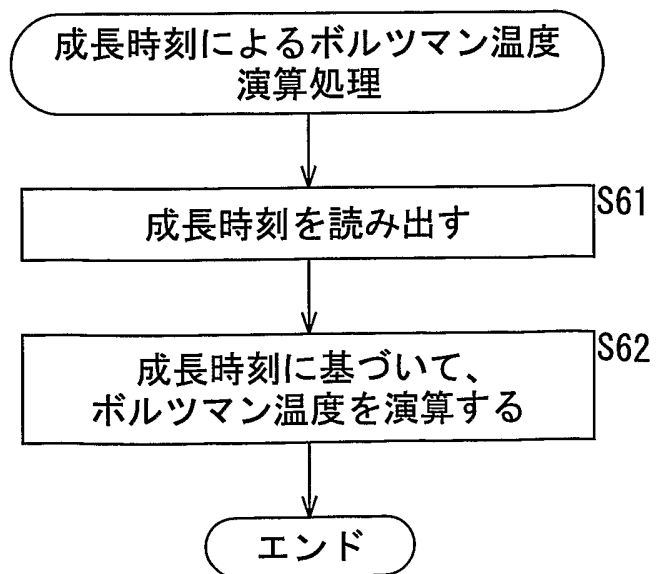
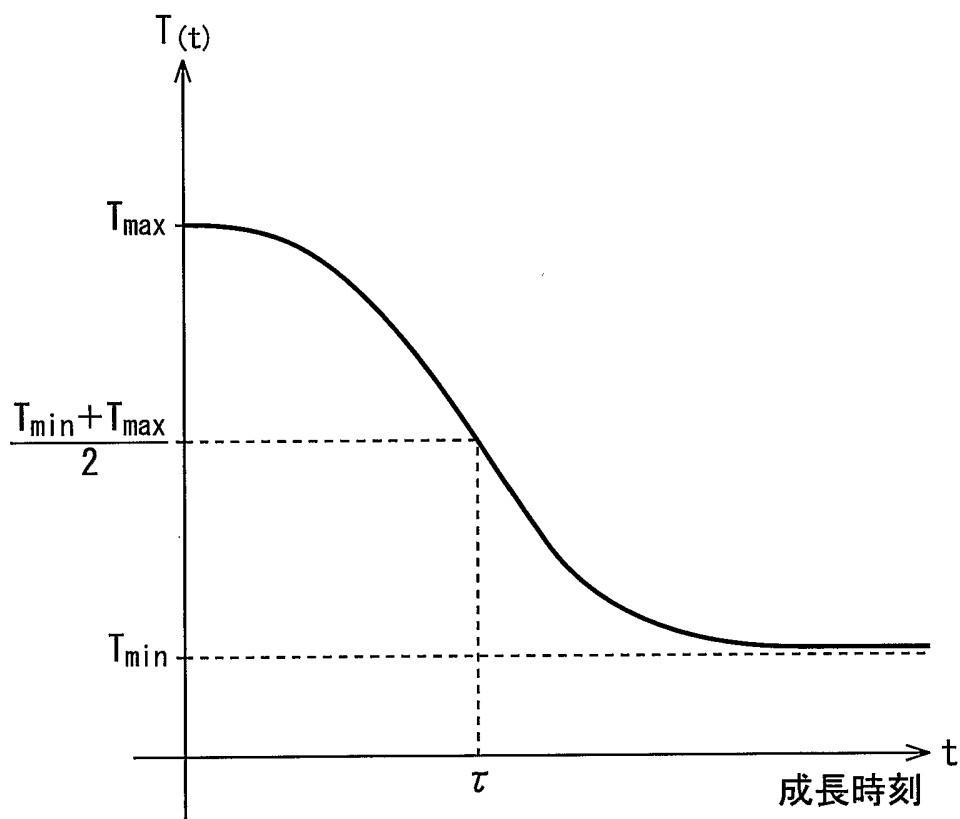


図15



INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP03/06179

<p>A. CLASSIFICATION OF SUBJECT MATTER Int.Cl.⁷ B25J13/00, 5/00</p> <p>According to International Patent Classification (IPC) or to both national classification and IPC</p>														
<p>B. FIELDS SEARCHED</p> <p>Minimum documentation searched (classification system followed by classification symbols) Int.Cl.⁷ B25J1/00-21/02</p> <p>Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Jitsuyo Shinan Koho 1926-1996 Kokai Jitsuyo Shinan Koho 1971-2003 Toroku Jitsuyo Shinan Koho 1994-2003</p> <p>Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) JOIS (JICST)</p>														
<p>C. DOCUMENTS CONSIDERED TO BE RELEVANT</p> <table border="1"> <thead> <tr> <th>Category*</th> <th>Citation of document, with indication, where appropriate, of the relevant passages</th> <th>Relevant to claim No.</th> </tr> </thead> <tbody> <tr> <td>A</td> <td>JP 2001-157981 A (Sony Corp.), 12 June, 2001 (12.06.01), Full text; all drawings (Family: none)</td> <td>1-7</td> </tr> <tr> <td>A</td> <td>JP 11-143849 A (Omron Corp.), 28 May, 1999 (28.05.99), Full text; all drawings (Family: none)</td> <td>1-7</td> </tr> <tr> <td>A</td> <td>US 5400244 A (KABUSHIKI KAISHA TOSHIBA), 21 March, 1995 (21.03.95), Full text; all drawings & JP 5-61539 A</td> <td>1-7</td> </tr> </tbody> </table>			Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.	A	JP 2001-157981 A (Sony Corp.), 12 June, 2001 (12.06.01), Full text; all drawings (Family: none)	1-7	A	JP 11-143849 A (Omron Corp.), 28 May, 1999 (28.05.99), Full text; all drawings (Family: none)	1-7	A	US 5400244 A (KABUSHIKI KAISHA TOSHIBA), 21 March, 1995 (21.03.95), Full text; all drawings & JP 5-61539 A	1-7
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.												
A	JP 2001-157981 A (Sony Corp.), 12 June, 2001 (12.06.01), Full text; all drawings (Family: none)	1-7												
A	JP 11-143849 A (Omron Corp.), 28 May, 1999 (28.05.99), Full text; all drawings (Family: none)	1-7												
A	US 5400244 A (KABUSHIKI KAISHA TOSHIBA), 21 March, 1995 (21.03.95), Full text; all drawings & JP 5-61539 A	1-7												
<p><input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> See patent family annex.</p>														
<p>* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed</p>		<p>"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family</p>												
<p>Date of the actual completion of the international search 19 August, 2003 (19.08.03)</p>		<p>Date of mailing of the international search report 09 September, 2003 (09.09.03)</p>												
<p>Name and mailing address of the ISA/ Japanese Patent Office</p> <p>Facsimile No.</p>		<p>Authorized officer</p> <p>Telephone No.</p>												

A. 発明の属する分野の分類 (国際特許分類 (IPC)) Int Cl ⁷ B25J13/00, 5/00		
B. 調査を行った分野 調査を行った最小限資料 (国際特許分類 (IPC)) Int Cl ⁷ B25J1/00-21/02		
最小限資料以外の資料で調査を行った分野に含まれるもの 日本国実用新案公報 1926-1996年 日本国登録実用新案公報 1994-2003年 日本国公開実用新案公報 1971-2003年		
国際調査で使用した電子データベース (データベースの名称、調査に使用した用語) JOIS (JICST)		
C. 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
A	JP 2001-157981 A (ソニー株式会社) 2001.06.12, 全文, 全図 (ファミリーなし)	1-7
A	JP 11-143849 A (オムロン株式会社) 1999.05.28, 全文, 全図 (ファミリーなし)	1-7
A	US 5400244 A (KABUSHIKI KAISHA TOSHIBA) 1995.03.21, 全文, 全図& JP 5-61539 A	1-7
<input type="checkbox"/> C欄の続きにも文献が列挙されている。 <input type="checkbox"/> パテントファミリーに関する別紙を参照。		
* 引用文献のカテゴリー 「A」 特に関連のある文献ではなく、一般的技術水準を示すもの 「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの 「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す) 「O」 口頭による開示、使用、展示等に言及する文献 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願		
の日の後に公表された文献 「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの 「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの 「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの 「&」 同一パテントファミリー文献		
国際調査を完了した日 19.08.03	国際調査報告の発送日 09.09.03	
国際調査機関の名称及びあて先 日本国特許庁 (ISA/JP) 郵便番号100-8915 東京都千代田区霞が関三丁目4番3号	特許庁審査官 (権限のある職員) 八木 誠	3C 9348 電話番号 03-3581-1101 内線 3324