

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4661074号  
(P4661074)

(45) 発行日 平成23年3月30日(2011.3.30)

(24) 登録日 平成23年1月14日(2011.1.14)

(51) Int.Cl.	F I	
<b>G 1 0 L 13/00 (2006.01)</b>	G 1 0 L 13/00	1 0 0 N
<b>G 1 0 L 13/08 (2006.01)</b>	G 1 0 L 13/00	1 0 0 A
<b>G 1 0 L 11/00 (2006.01)</b>	G 1 0 L 13/08	1 3 1 Z
<b>G 1 0 L 15/00 (2006.01)</b>	G 1 0 L 13/08	1 3 2
	G 1 0 L 11/00	1 0 1 G
請求項の数 6 (全 30 頁) 最終頁に続く		

(21) 出願番号	特願2004-113589 (P2004-113589)	(73) 特許権者	000002185
(22) 出願日	平成16年4月7日(2004.4.7)		ソニー株式会社
(65) 公開番号	特開2005-300692 (P2005-300692A)		東京都港区港南1丁目7番1号
(43) 公開日	平成17年10月27日(2005.10.27)	(74) 代理人	100093241
審査請求日	平成19年4月4日(2007.4.4)		弁理士 宮田 正昭
		(74) 代理人	100101801
			弁理士 山田 英治
		(74) 代理人	100086531
			弁理士 澤田 俊夫
		(72) 発明者	藤田 雅博
			東京都品川区北品川6丁目7番35号 ソニー株式会社内
		審査官	菊池 智紀
最終頁に続く			

(54) 【発明の名称】 情報処理システム、情報処理方法、並びにロボット装置

(57) 【特許請求の範囲】

【請求項1】

音韻系列の学習を行なう情報処理システムにおいて、  
フィルタ関数  $K$  を用いて話者毎の音韻系列 ( $s$ 、 $s'$ ) を基準周波数の通倍の中心周波数を持つ複数の成分で構成される音韻系列  $u$  に変換する蝸牛特性フィルタと、

学習関数  $H$  を用いて音韻系列 ( $u$ ) を音声合成パラメータ ( $y$ ) に変換するとともに、音韻系列 ( $u$ ) から変換した音声合成パラメータ ( $y$ ) と音韻系列 ( $u$ ) の聞き真似をして発生する音声合成パラメータ ( $y_0$ ) との関係を利用する学習又は評価する音声信号学習手段と

、  
音声合成関数  $G$  を用いて音声合成パラメータ ( $y$ ) から音韻系列 ( $s$ ) を生成する音声合成手段と、  
を具備し、

前記音声合成手段が前記音声合成関数  $G$  を用いて当該システムに入力された音声信号の音声合成パラメータ ( $y$ ) から音韻系列 ( $s$ ) を生成し、

前記蝸牛特性フィルタが前記フィルタ関数  $K$  を用いて前記の生成した音韻系列 ( $s$ ) を音韻系列 ( $u$ ) に変換し、

前記音声信号学習手段が前記学習関数  $H$  を用いて音韻系列 ( $u$ ) から聞き真似をするための音声合成パラメータ ( $y_0$ ) を生成するとともに、前記音声合成手段が前記音声合成関数  $G$  を用いて聞き真似をするための音声合成パラメータ ( $y_0$ ) から聞き真似をした音韻系列を合成し、

前記音声信号学習手段が、当該システムへの入力としての音声合成パラメータ ( $y$ ) と聞き真似をするための音声合成パラメータ ( $y_0$ ) との相違を評価して前記学習関数  $H$  を学習する、

ことを特徴とする情報処理システム。

【請求項 2】

前記音声信号学習手段は、フィードフォワード型ニューラル・ネットワークで構成される、

ことを特徴とする請求項 1 に記載の情報処理システム。

【請求項 3】

フィルタ関数  $K$  を用いて話者毎の音韻系列 ( $s$ 、 $s'$ ) を基準周波数の通倍の中心周波数を持つ複数の成分で構成される音韻系列  $u$  に変換する蝸牛特性フィルタと、

学習関数  $H$  を用いて音韻系列 ( $u$ ) を音声合成パラメータ ( $y$ ) に変換するとともに、音韻系列 ( $u$ ) から変換した音声合成パラメータ ( $y$ ) と音韻系列 ( $u$ ) の聞き真似をして発生する音声合成パラメータ ( $y_0$ ) との関係进行学习又は評価する音声信号学習手段と

、  
内部状態又は身体的制約 ( $x$ ) を生成する内部状態生成手段と、

写像関数  $F$  を用いて内部状態又は身体的制約 ( $x$ ) から内部状態又は身体的制約 ( $x$ ) を表す音声合成パラメータ ( $y$ ) に変換する写像手段と、

音声合成関数  $G$  を用いて音声合成パラメータ ( $y$ ) から音韻系列 ( $s$ ) を生成する音声合成手段と、

を具備し、

前記写像手段が、前記写像関数  $F$  を用いて、前記内部状態生成手段が生成した現在の内部状態又は身体的制約 ( $x$ ) から内部状態又は身体的制約 ( $x$ ) を表す音声合成パラメータ ( $y$ ) に変換し、

前記音声合成手段が、前記音声合成関数  $G$  を用いて、前記現在の内部状態又は身体的制約 ( $x$ ) から内部状態又は身体的制約 ( $x$ ) を表す音声合成パラメータ ( $y$ ) から音韻系列 ( $s$ ) を生成し、

前記蝸牛特性フィルタが前記フィルタ関数  $K$  を用いて前記の生成した音韻系列 ( $s$ ) を音韻系列 ( $u$ ) に変換し、

前記音声信号学習手段が前記学習関数  $H$  を用いて音韻系列 ( $u$ ) から聞き真似をするための音声合成パラメータ ( $y_0$ ) を生成するとともに、前記音声合成手段が前記音声合成関数  $G$  を用いて聞き真似をするための音声合成パラメータ ( $y_0$ ) から聞き真似をした音韻系列を合成し、

前記音声信号学習手段が、前記現在の内部状態又は身体的制約 ( $x$ ) から内部状態又は身体的制約 ( $x$ ) を表す音声合成パラメータ ( $y$ ) と聞き真似をするための音声合成パラメータ ( $y_0$ ) との相違を評価して前記学習関数  $H$  を学習する、

ことを特徴とする情報処理システム。

【請求項 4】

ケアギバが発する音韻系列 ( $s'$ ) を前記蝸牛特性フィルタが前記フィルタ関数  $K$  を用いて音韻系列 ( $u'$ ) に変換し、

前記音声信号学習手段が前記学習関数  $H$  を用いて前記ケアギバの音韻系列 ( $u'$ ) から聞き真似をするための音声合成パラメータ ( $y'$ ) を生成し、

前記写像手段は、前記写像関数  $F$  の逆関数  $F^{-1}$  を用いて、前記ケアギバの音韻系列 ( $u'$ ) に基づく音声合成パラメータ ( $y'$ ) に相当する内部状態 ( $x'$ ) を逆推定する、

ことを特徴とする請求項 3 に記載の情報処理システム。

【請求項 5】

情報処理装置上で音韻系列の学習を行なう情報処理方法において、

前記情報処理装置が備える蝸牛特性フィルタが、フィルタ関数  $K$  を用いて話者毎の音韻系列 ( $s$ 、 $s'$ ) を基準周波数の通倍の中心周波数を持つ複数の成分で構成される音韻系列  $u$  に変換するフィルタリング・ステップと、

10

20

30

40

50

前記情報処理装置が備える音声信号学習手段が、学習関数 H を用いて音韻系列 ( u ) を音声合成パラメータ ( y ) に変換するとともに、音韻系列 ( u ) から変換した音声合成パラメータ ( y ) と音韻系列 ( u ) の聞き真似をして発生する音声合成パラメータ ( y<sub>o</sub> ) との関係を学習又は評価する音声信号学習ステップと、

前記情報処理装置が備える音声合成手段が、音声合成関数 G を用いて音声合成パラメータ ( y ) から音韻系列 ( s ) を生成する音声合成ステップと、  
を有し、

前記音声合成ステップで前記音声合成関数 G を用いて当該システムに入力された音声信号の音声合成パラメータ ( y ) から音韻系列 ( s ) を生成し、

前記フィルタリング・ステップで前記フィルタ関数 K を用いて前記の生成した音韻系列 ( s ) を音韻系列 ( u ) に変換し、

前記音声信号学習ステップで前記学習関数 H を用いて音韻系列 ( u ) から聞き真似をするための音声合成パラメータ ( y<sub>o</sub> ) を生成するとともに、前記音声合成ステップで前記音声合成関数 G を用いて聞き真似をするための音声合成パラメータ ( y<sub>o</sub> ) から聞き真似をした音韻系列を合成し、

前記音声信号学習ステップで当該システムへの入力としての音声合成パラメータ ( y ) と聞き真似をするための音声合成パラメータ ( y<sub>o</sub> ) との相違を評価して前記学習関数 H を学習する、

ことを特徴とする情報処理方法。

【請求項 6】

音韻系列の学習を行なうロボット装置において、

複数の可動部と、

前記可動部を駆動する駆動手段と、

前記駆動手段を制御する駆動制御手段と、

フィルタ関数 K を用いて話者毎の音韻系列 ( s 、 s' ) を基準周波数の逡倍の中心周波数を持つ複数の成分で構成される音韻系列 u に変換する蝸牛特性フィルタと、

学習関数 H を用いて音韻系列 ( u ) を音声合成パラメータ ( y ) に変換するとともに、音韻系列 ( u ) から変換した音声合成パラメータ ( y ) と音韻系列 ( u ) の聞き真似をして発生する音声合成パラメータ ( y<sub>o</sub> ) との関係を学習又は評価する音声信号学習手段と

音声合成関数 G を用いて音声合成パラメータ ( y ) から音韻系列 ( s ) を生成する音声合成手段と、

を具備し、

前記音声合成手段が前記音声合成関数 G を用いて当該システムに入力された音声信号の音声合成パラメータ ( y ) から音韻系列 ( s ) を生成し、

前記蝸牛特性フィルタが前記フィルタ関数 K を用いて前記の生成した音韻系列 ( s ) を音韻系列 ( u ) に変換し、

前記音声信号学習手段が前記学習関数 H を用いて音韻系列 ( u ) から聞き真似をするための音声合成パラメータ ( y<sub>o</sub> ) を生成するとともに、前記音声合成手段が前記音声合成関数 G を用いて聞き真似をするための音声合成パラメータ ( y<sub>o</sub> ) から聞き真似をした音韻系列を合成し、

前記音声信号学習手段が、当該システムへの入力としての音声合成パラメータ ( y ) と聞き真似をするための音声合成パラメータ ( y<sub>o</sub> ) との相違を評価して前記学習関数 H を学習する、

ことを特徴とするロボット装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、自律的な動作を行ないユーザとのリアリスティックなコミュニケーションを

10

20

30

40

50

実現する情報処理システム、ロボットの行動制御システム及び行動制御方法、並びにロボット装置に係り、特に、視覚や聴覚などの外部環境の認識結果や本能や感情などの内部状態などのロボットが置かれている状況を統合的に判断して適当な行動を選択する情報処理システム、ロボットのための行動制御システム及び行動制御方法、並びにロボット装置に関する。

【0002】

さらに詳しくは、本発明は、外部環境や内部状態に基づいて本能的、反射的な状況依存型の行動、又はその他の形態により発現される行動を選択する情報処理システム、ロボットの行動制御システム及び行動制御方法、並びにロボット装置に係り、特に、模倣や学習などのメカニズムを取り入れ、自己発達により進化を遂げる情報処理システム、ロボットの行動制御システム及び行動制御方法、並びにロボット装置に関する。

10

【背景技術】

【0003】

電氣的若しくは磁氣的な作用を用いて人間の動作に似せた運動を行う機械装置のことを「ロボット」という。ロボットの語源は、スラブ語のROBOTA（奴隷機械）に由来すると言われている。わが国では、ロボットが普及し始めたのは1960年代末からであるが、その多くは、工場における生産作業の自動化・無人化などを目的としたマニピュレータや搬送ロボットなどの産業用ロボット(industrial robot)であった。

【0004】

20

最近では、イヌやネコのように4足歩行の動物の身体メカニズムやその動作を模したペット型ロボット、あるいは、ヒトやサルなどの2足直立歩行を行う動物の身体メカニズムや動作を模した「人間形」若しくは「人間型」のロボット(humanoid robot)など、脚式移動ロボットの構造やその安定歩行制御に関する研究開発が進展し、実用化への期待も高まってきている。これら脚式移動ロボットは、クローラ式ロボットに比し不安定で姿勢制御や歩行制御が難しくなるが、階段の昇降や障害物の乗り越えなど、柔軟な歩行・走行動作を実現できるという点で優れている。

【0005】

また、最近では、ロボット装置は、産業上の用途以外に、生活密着型、すなわち人間との「共生」あるいは「エンターテインメント」という目的で使用されることが多くなってきている。従来の玩具機械では、ユーザ操作と応答動作との関係が固定的で、玩具の動作をユーザの好みに合わせて変更することはできないため、ユーザは同じ動作の繰り返しにやがては飽きてしまう。これに対し、知能型のロボット装置は、動作に起因する行動モデルや学習モデルを備え、外部からの音声や画像、触覚などの入力情報に基づいてモデルを変化させて動作を決定し、自律的な思考及び動作制御を実現することができる。また、人間とのインターフェースとして画像認識処理や音声認識処理などの認識技術を導入し、より高度でリアリスティックなコミュニケーションが可能となる。

30

【0006】

ロボットあるいはその他のリアリスティックな対話システムでは、視覚や聴覚など外部環境の変化に応じて逐次的に行動を選択していくのが一般的である。また、行動選択メカニズムの他の例として、本能や感情といった情動をモデル化してシステムの内部状態を管理して、内部状態の変化に応じて行動を選択するものを挙げることができる。

40

【0007】

ここで、内部状態には、例えば生体で言えば大脳辺縁系へのアクセスに相当する本能のような要素や、大脳新皮質へのアクセスに相当する内発的欲求や社会的欲求などのように動物行動学的モデルで捉えられる要素、さらには喜びや悲しみ、怒り、驚きなどのような感情と呼ばれる要素などで構成される。

【0008】

勿論、システムの内部状態は、外部環境の変化によっても変化するし、選択された行動を発現することによっても変化する。

50

## 【 0 0 0 9 】

例えば、視覚や聴覚などの外部環境の認識結果や本能や感情などの内部状態など、ロボットが置かれている状況を総合的に判断して行動を選択するロボット装置について提案されている（例えば、特許文献1を参照のこと）。

## 【 0 0 1 0 】

この場合のロボット装置は、内部状態又は外部入力に応じた前記ロボット装置の行動評価を出力する行動評価手段と、前記ロボット装置の行動命令を出力する行動命令出力手段とをそれぞれ備えた複数の行動モジュールを備えている。そして、各行動モジュールの行動評価手段から得られる評価を基にして、外部刺激と内部状態の変化に応じて最適な行動モジュールを統合的な判断すること、すなわちConcurrentな評価を行なうことができる。また、前記階層構造の下位層の行動モジュールから上位層の行動モジュールへ出力される行動評価及び前記ロボット装置の使用リソースに基いて行動モジュールを選択するという、リソースを考慮した行動選択、すなわちConcurrentな実行を実現することができる。

10

## 【 0 0 1 1 】

ここで、ロボット装置のさらなるインテリジェント化を考察した場合、上述したような行動の発現方法にとどまらず、ロボットが自律的に学習し、さらに自己発達により進化を遂げることが期待される。

## 【 0 0 1 2 】

ロボットやその他の自動機械における学習メカニズムは、一般に、ニューラル・ネットワークやその他の数学モデルを用いて実現される。

20

## 【 0 0 1 3 】

ここで、ニューラル・ネットワークとは、人間の脳における神経回路網を簡略化したモデルであり、神経細胞ニューロンが一方向にのみ信号を通過するシナプスを介して結合されているネットワークを意味する。ニューロン間の信号の伝達はシナプスを介して行なわれ、シナプスの抵抗、すなわち重みを適当に調整することによりさまざまな情報処理が可能となる。各ニューロンは、他の1以上のニューロンからの出力をシナプスによる重み付けをして入力し、それら入力値の総和を非線型応答関数の変形を加え、再度他のニューロンへ出力する。ニューラル・ネットワークによる制御では、摩擦や粘性などの非線型問題にそのまま対応することができるとともに、学習機能を備えているので、パラメータの設定変更が不要になる。

30

## 【 0 0 1 4 】

例えば、ニューラル・ネットワークを用い、教示された動作パターンを音声などの入力シンボルと連想して記憶することができる（例えば、特許文献2を参照のこと）。この場合、ロボットは、比較的長い教示動作やその他の行動を分節化するとともに、分節化された各動作を、音声認識などにより認識された入力シンボルと連想して記憶することができる。動作再生時においては、比較的長い動作パターンを容易に指示することができる。また、直接的には同一のものを学習していない行動であっても、音声指示により入力されたシンボルから連想していくことにより、指示により近い行動を発現することができる。また、リカレント・ニューラル・ネットワークを適用することで、内部にフィードバック結合を備え、1周期前の情報をネットワーク内に持ち、これによって時系列データの履歴を把握することができる。

40

## 【 0 0 1 5 】

他方、ロボットのモデルである人間やその他の動物自体は、模倣、すなわち親（若しくはケアギバ（care giver））や周囲の環境の物まねを通じて学習するのが一般的である。

## 【 0 0 1 6 】

【特許文献1】特開2003-334785号公報

【特許文献2】特開2002-337075号公報

【発明の開示】

50

## 【発明が解決しようとする課題】

## 【0017】

本発明の目的は、自律的な動作を行ないユーザとのリアリスティックなコミュニケーションを実現することができる、優れた情報処理システム、ロボットの行動制御システム及び行動制御方法、並びにロボット装置を提供することにある。

## 【0018】

本発明のさらなる目的は、視覚や聴覚などの外部環境の認識結果や本能や感情などの内部状態などのロボットが置かれている状況を統合的に判断して適当な行動を選択することができる、優れた情報処理システム、ロボットの行動制御システム及び行動制御方法、並びにロボット装置を提供することにある。

10

## 【0019】

本発明のさらなる目的は、外部環境や内部状態に基づいて本能的、反射的な状況依存型の行動、又はその他の形態により発現される行動を選択することができる、優れた情報処理システム、ロボットの行動制御システム及び行動制御方法、並びにロボット装置を提供することにある。

## 【0020】

本発明のさらなる目的は、模倣や学習などのメカニズムを取り入れ、自律的に学習し、自己発達により進化を遂げることができる、優れた情報処理システム、ロボットの行動制御システム及び行動制御方法、並びにロボット装置を提供することにある。

## 【課題を解決するための手段】

20

## 【0021】

本発明は、上記課題を参酌してなされたものであり、その第1の側面は、音韻系列の学習を行なうロボットの行動制御システムにおいて、

音韻系列を入力する音声入力手段と、

音韻系列を音声合成パラメータに変換するとともに、自己が生成した音韻系列に関する音声合成パラメータと聞き真似をする音声合成パラメータとの関係を学習又は評価する音声信号学習手段と、

前記音声信号学習手段による音声合成パラメータに基づいて音韻系列を生成する音声合成手段と、

を具備することを特徴とするロボットの行動制御システムである。ここで、前記音声入力手段は蝸牛特性フィルタ・バンクを備えていてもよい。

30

## 【0022】

但し、ここで言う「システム」とは、複数の装置（又は特定の機能を実現する機能モジュール）が論理的に集合した物のことを言い、各装置や機能モジュールが単一の筐体内にあるか否かは特に問わない。

## 【0023】

本発明の第1の側面に係るロボットの行動制御システムによれば、音韻系列の素となる音声合成パラメータが音声合成手段に投入されると、音声信号からなる音韻系列 $s$ が出力される。この音韻系列 $s$ は蝸牛特性フィルタ・バンクにより話者間の相違が吸収された音韻系列 $u$ に変換される。そして、音声信号学習手段は、この音韻系列 $u$ から、聞き真似をするための音声合成パラメータ $y_0$ を生成する。さらに、この音声合成パラメータ $y_0$ が音声合成手段に投入されると、聞き真似をした音韻系列が発されることになる。

40

## 【0024】

そして、音声信号学習手段は、当該システムへの入力としての音声合成パラメータ $y_0$ と、聞き真似するための音声合成パラメータ $y_0$ との相違を評価して、学習を行なうことができる。

## 【0025】

本発明の第1の実施形態に係るロボットの行動制御システムは、内部状態又は身体的制約を生成する内部状態生成手段と、内部状態を音声合成パラメータに変換する写像手段とをさらに備えることにより、内部状態又は身体的制約を考慮し、これらにより特性の異なる

50

る発音内容を生成することができる。

【0026】

この場合、内部状態生成手段により内部状態  $x$  が生成されると、写像手段により、内部状態を表した音韻系列を生成するための音声合成パラメータ  $y$  に変換される。そして、音声合成パラメータ  $y$  が音声合成手段に投入されると、音声信号  $s$  からなる音韻系列が出力される。そして、この音韻系列  $s$  が蝸牛特性フィルタ・バンクに入力され、話者間の相違が吸収された音韻系列  $u$  に変換される。音声信号学習手段は、この音韻系列  $u$  から、聞き真似をするための音声合成パラメータ  $y_0$  を生成する。さらに、この音声合成パラメータ  $y_0$  が音声合成手段に投入されると、聞き真似をした音韻系列が発されることになる。

【0027】

このような行動制御システムによれば、ロボット装置は、ランダムな発生に対して、自分で入力を分析して学習（予測誤差、あるいは相関学習）することができる。また、人間が意味を感じた音に対して同じ音を出すと、それを分析して同じく学習（個人性を吸収する学習 + 発生の強化学習）することができる。すなわち、ロボット装置は、聞いた声が自分の発生でどのパラメータに相当するかを覚えることができ、意味のある発生を覚えることができるようになる。

【0028】

また、このような行動制御システムによれば、写像手段が写像時の逆関数を用いて、入力された音声合成パラメータに相当する内部状態を逆推定することができる。すなわち、人の声を聞いたときに、自分の内部状態が逆写像で求まるといふ、身体的理解を実現することができる。

【0029】

また、本発明の第2の側面は、内部状態と外部環境に基づいて発現すべき行動を決定するとともに、予測学習を行なうロボットの行動制御システムにおいて、

前記ロボットは、駆動指令に基づいて動作する駆動手段と、内部状態を管理する内部状態管理手段と、外部刺激又は物理環境を計測する計測手段を備えており、

前記行動制御システムは、前記ロボットの駆動指令、内部状態、物理環境に関する現在の観測値と現在の目標値を入力し、次の時刻におけるこれらの推定値を推定する予測学習手段を備える、

ことを特徴とするロボットの行動制御システムである。

【0030】

ロボット装置は、現在観測されている世界、現在目標とする世界、次の時刻において予測される世界、並びに実ロボットが動作した結果観測される次の時刻における世界という4つの世界に取り囲まれている。ここでは、各世界をロボットの駆動指令、内部状態、物理環境というパラメータで記述することになっている。

【0031】

この予測学習手段は、例えば、フィードフォワード型ニューラル・ネットワークや相関学習器により構成することができる。

【0032】

そして、予測学習手段は、次の時刻における前記ロボットの駆動指令、内部状態、物理環境の推定値と、次の時刻におけるこれらの観測値との偏差や相関などに基づいて予測学習を行なうことができる。

【0033】

また、本発明の第2の側面に係る行動制御システムによれば、ロボット装置は、ロボット単体の身体的ダイナミクス以外に、道具を使った道具ダイナミクスを実現することもできる。後者の場合、ロボット装置は、外部環境に対し道具を用いたインタラクションを行なう。

【0034】

このような場合、予測学習手段は、物理環境を操作するために使用した道具と、道具を介した物理環境の操作結果としてのボディー・イメージと、物理環境を取り扱うことで、

10

20

30

40

50

上述とは同じ枠組みで、道具の力学系を学習し、それを“道具”として使うことが可能になる。

【0035】

また、学習のメカニズムを利用してロボット装置の自己発達を実現する場合、過大なテーマが与えられると、学習の結果がなかなか収束せず、学習効率が上がらないという問題がある。そこで、本発明では、学習メカニズムに発達のシナリオの概念を取り入れることができる。

【0036】

発達のシナリオは、複数の学習フェーズで構成される。各学習フェーズでは、ロボット装置が実現する自由度構成（関節の可動角やトルク限界など）や取るべき行動リスト（本能的行動、反射行動など）を、学習の進行に従って定義している。学習フェーズ毎にファイルで管理することにより、学習の進行に応じたファイルをアクセスし、発達のシナリオを実現することができる。

10

【0037】

学習フェーズの遷移は、時間の経過でコントロールしてもよいが、予測学習手段における推定値と現実の観測値との偏差や相関に基づいてコントロールすることができる。すなわち、推定値通りに実ロボットが行動できるようになれば、現在の学習フェーズをクリアしたとみなし、次の学習フェーズへ遷移させるようにすればよい。

【発明の効果】

【0038】

本発明によれば、自律的な動作を行ないユーザとのリアリスティックなコミュニケーションを実現することができる、優れた情報処理システム、ロボットの行動制御システム及び行動制御方法、並びにロボット装置を提供することができる。

20

【0039】

また、本発明によれば、視覚や聴覚などの外部環境の認識結果や本能や感情などの内部状態などのロボットが置かれている状況を統合的に判断して適当な行動を選択することができる、優れた情報処理システム、ロボットの行動制御システム及び行動制御方法、並びにロボット装置を提供することができる。

【0040】

また、本発明によれば、外部環境や内部状態に基づいて本能的、反射的な状況依存型の行動、又はその他の形態により発現される行動を選択することができる、優れた情報処理システム、ロボットの行動制御システム及び行動制御方法、並びにロボット装置を提供することができる。

30

【0041】

また、本発明によれば、模倣や学習などのメカニズムを取り入れ、自律的に学習し、自己発達により進化を遂げることができる、優れた情報処理システム、ロボットの行動制御システム及び行動制御方法、並びにロボット装置を提供することができる。

【0042】

本発明のさらに他の目的、特徴や利点は、後述する本発明の実施形態や添付する図面に基づくより詳細な説明によって明らかになるであろう。

40

【発明を実施するための最良の形態】

【0043】

以下、図面を参照しながら本発明の実施形態について詳解する。

【0044】

#### A. システム構成

図1には、本発明に実施に供されるロボット装置1の機能構成を模式的に示している。同図に示すように、ロボット装置1は、全体の動作の統括的制御やその他のデータ処理を行なう制御ユニット20と、入出力部40と、駆動部50と、電源部60とで構成される。以下、各部について説明する。

【0045】

50

入出力部 40 は、入力部としてロボット装置 1 の目に相当する CCD カメラ 15 や、耳に相当するマイクロフォン 16、頭部や背中などの部位に配設されてユーザの接触を感知するタッチ・センサ 18、あるいは五感に相当するその他の各種のセンサを含む。また、出力部として、口に相当するスピーカ 17、あるいは点滅の組み合わせや点灯のタイミングにより顔の表情を形成する LED インジケータ（目ランプ）19などを装備している。これら出力部は、音声やランプの点滅など、脚などによる機械運動パターン以外の形式でもロボット装置 1 からのユーザ・フィードバックを表現することができる。

#### 【0046】

ここで、カメラ 15 や、マイクロフォン 16、タッチ・センサ 18 などの入力機器は、検出信号をデジタル変換並びにデジタル信号処理を行なうものとする。例えば、マイクロフォン 16 は、蝸牛特性フィルタ・バンクによるフィルタリングやその他のデジタル信号処理を行なう DSP (Digital Signal Processor) を含んでいる。

10

#### 【0047】

ロボット装置 1 は、カメラ 15 を含むことで、作業空間上に存在する任意の物体の形状や色彩を認識することができる。また、ロボット装置 1 は、カメラのような視覚手段の他に、赤外線、音波、超音波、電波などの発信波を受信する受信装置をさらに備えていてもよい。この場合、各伝送波を検知するセンサ出力に基づいて発信源からの位置や向きを計測することができる。

#### 【0048】

20

駆動部 50 は、制御部 20 が指令する所定の運動パターンに従ってロボット装置 1 の機体動作を実現する機能ブロックであり、行動制御による制御対象である。駆動部 50 は、ロボット装置 1 の各関節における自由度を実現するための機能モジュールであり、それぞれの関節におけるロール、ピッチ、ヨーなど各軸毎に設けられた複数の駆動ユニットで構成される。各駆動ユニットは、所定軸回りの回転動作を行なうモータ 51 と、モータ 51 の回転位置を検出するエンコーダ 52 と、エンコーダ 52 の出力に基づいてモータ 51 の回転位置や回転速度を適応的に制御するドライバ 53 の組み合わせで構成される。

#### 【0049】

駆動ユニットの組み合わせ方によって、ロボット装置 1 を例えば 2 足歩行又は 4 足歩行などの脚式移動ロボットとして構成することができる。

30

#### 【0050】

電源部 60 は、その字義通り、ロボット装置 1 内の各電気回路などに対して給電を行なう機能モジュールである。本実施形態に係るロボット装置 1 は、バッテリーを用いた自律駆動式であり、電源部 60 は、充電バッテリー 61 と、充電バッテリー 61 の充放電状態を管理する充放電制御部 62 とで構成される。

#### 【0051】

充電バッテリー 61 は、例えば、複数本のリチウムイオン 2 次電池セルをカートリッジ式にパッケージ化した「バッテリー・パック」の形態で構成される。また、充放電制御部 62 は、バッテリー 61 の端子電圧や充電/放電電流量、バッテリー 61 の周囲温度などを測定することでバッテリー 61 の残存容量を把握し、充電の開始時期や終了時期などを決定する。充放電制御部 62 が決定する充電の開始及び終了時期は制御ユニット 20 に通知され、ロボット装置 1 が充電オペレーションを開始及び終了するためのトリガとなる。

40

#### 【0052】

制御部 20 は、「頭脳」に相当し、例えばロボット装置 1 の頭部ユニットあるいは胴体部ユニットに搭載される。

#### 【0053】

図 2 には、制御ユニット 20 の構成をさらに詳細に図解している。同図に示すように、制御ユニット 20 は、メイン・コントローラとしての CPU (Central Processing Unit) 121 が、メモリその他の各回路コンポーネントや周辺機器とバス接続された構成となっている。バス 28 は、データ・バス、アドレス・バス、コント

50

ロール・バスなどを含む共通信号伝送路である。

【0054】

RAM (Random Access Memory) 22は、DRAM (Dynamic RAM)などの揮発性メモリで構成された書き込み可能メモリであり、CPU 21が実行するプログラム・コードをロードしたり、実行プログラムによる作業データの一時的に保存したりするために使用される。

【0055】

ROM (Read Only Memory) 23は、プログラムやデータを恒久的に格納する読み出し専用メモリである。ROM 23に格納されるプログラム・コードには、ロボット装置1の電源投入時に実行する自己診断テスト・プログラムや、ロボット装置1の動作を規定する動作制御プログラムなどが挙げられる。

10

【0056】

本実施形態に係るロボット装置1は、動作制御プログラムによって、外部刺激の認識結果や内部状態の変化に応じて行動制御を行なうことができるが、行動制御のシステム構成については後述に譲る。

【0057】

不揮発性メモリ24は、例えばEEPROM (Electrically Erasable and Programmable ROM)のように、電氣的に消去再書き込みが可能なメモリ素子で構成され、逐次更新すべきデータを不揮発的に保持するために使用される。逐次更新すべきデータには、例えば、製造番号や暗号鍵などのセキュリティ情報などが挙げられる。

20

【0058】

インターフェース25は、制御ユニット20外の機器と相互接続し、データ交換を可能にするための装置である。インターフェース25は、例えば、カメラ15やマイクロフォン16、スピーカ17との間でデータ入出力を行なう。また、インターフェース25は、駆動部50内の各駆動回路53-1...との間でデータやコマンドの入出力を行なう。

【0059】

また、インターフェース25は、RS (Recommended Standard) - 232Cなどのシリアル・インターフェース、IEEE (Institute of Electrical and Electronics Engineers) 1284などのパラレル・インターフェース、USB (Universal Serial Bus) インターフェース、i-Link (IEEE 1394) インターフェース、SCSI (Small Computer System Interface) インターフェース、メモリ・カード・インターフェースなどのような、コンピュータの周辺機器接続用の汎用インターフェースを備え、ローカル接続された外部機器との間でプログラムやデータの移動を行なうようにしてもよい。

30

【0060】

また、インターフェース25の他の例として、赤外線通信 (IrDA) インターフェースを備え、外部機器と無線通信を行なうようにしてもよい。

【0061】

さらに、制御ユニット20は、無線通信インターフェース26を含み、Bluetoothのような近距離無線データ通信やIEEE 802.11bのような無線ネットワークにより機体の数メートル以内に配設されたアクセスポイント (図示しない) との間で無線データ通信を行なうことができる。アクセスポイントは、さらにLANなどの小規模ネットワークやインターネットなどの広域ネットワークに相互接続され、ロボット装置1をネットワーク上の情報提供空間へと導くことができる。

40

【0062】

B. 行動制御システム

ロボットの行動は、本能的行動、反射的行動、条件反応的行動、学習行動、学習済み行動などに分類することができる。学習行動は、行動の順序を覚える (すなわち運動計画を

50

立てる)、あるいはダイナミクス(ロボット単体の身体的ダイナミクス、道具を使った道具ダイナミクス(後述)を含む)を覚えることを含み、探索学習や、物まね学習などが含まれる。また、条件反動的行動、学習行動、学習済み行動は、本能行動と反射行動に基づいて実装することが可能である。

【0063】

ロボットの行動制御システムを実装する上で、これらの行動を統一的に扱う仕組みが重要であると、本発明者らは思料する。例えば、どういうときに「ものまね」を行ない、どういうときに「因果行動」をとるべきかである。

【0064】

図3には、本発明の実施形態に係るロボット装置1の行動制御システム100の機能構成を模式的に示している。ロボット装置1は、外部刺激の認識結果や内部状態の変化に応じて行動制御を行なうことができる。さらには、長期記憶機能を備え、外部刺激から内部状態の変化を連想記憶することにより、外部刺激の認識結果や内部状態の変化に応じて行動制御を行なうことができる。

【0065】

図示の行動制御システム100にはオブジェクト指向プログラミングを採り入れて実装することができる。この場合、各ソフトウェアは、データとそのデータに対する処理手続きとを一体化させた「オブジェクト」というモジュール単位で扱われる。また、各オブジェクトは、メッセージ通信と共有メモリを使ったオブジェクト間通信方法によりデータの受け渡しとInvokeを行なうことができる。

【0066】

行動制御システム100は、外部環境(Environments)を認識するために、視覚認識機能部101と、聴覚認識機能部102と、接触認識機能部103を備えている。

【0067】

視覚認識機能部(Video)51は、例えば、CCDカメラのような画像入力装置を介して入力された撮影画像を基に、顔認識や色認識などの画像認識処理や特徴抽出を行なう。視覚認識機能部51は、後述する"MultiColorTracker", "FaceDetector", "FaceIdentify"といった複数のオブジェクトで構成される。

【0068】

聴覚認識機能部(Audio)52は、マイクなどの音声入力装置を介して入力される音声データを音声認識して、特徴抽出したり、単語セット(テキスト)認識を行ったりする。聴覚認識機能部52は、後述する"AudioRecog", "AuthurDecoder"といった複数のオブジェクトで構成される。

接触認識機能部(Tactile)53は、例えば機体の頭部などに内蔵された接触センサによるセンサ信号を認識して、「なでられた」とか「叩かれた」という外部刺激を認識する。

【0069】

内部状態管理部(ISM: Internal Status Manager)104は、本能や感情といった数種類の情動を数式モデル化して管理しており、上述の視覚認識機能部101と、聴覚認識機能部102と、接触認識機能部103によって認識された外部刺激(ES: External Stimula)に応じてロボット装置1の本能や情動といった内部状態を管理する。

【0070】

感情モデルと本能モデルは、それぞれ認識結果と行動履歴を入力に持ち、感情値と本能値を管理している。行動モデルは、これら感情値や本能値を参照することができる。

【0071】

本実施形態では、情動についてその存在意義による複数階層で構成され、それぞれの階層で動作する。決定された複数の動作から、そのときの外部環境や内部状態によってどの

10

20

30

40

50

動作を行なうかを決定するようになっていいる。また、それぞれの階層で行動は選択されるが、より低次の行動から優先的に動作を発現していくことにより、反射などの本能的行動や、記憶を用いた動作選択などの高次の行動を1つの個体上で矛盾なく発現することができる。

【0072】

ロボット装置1は、外部刺激の認識結果や内部状態の変化に応じて行動制御を行なうために、時間の経過とともに失われる短期的な記憶を行なう短期記憶部105と、情報を比較的長期間保持するための長期記憶部106を備えている。短期記憶と長期記憶という記憶メカニズムの分類は神経心理学に依拠する。

【0073】

短期記憶部(Short Term Memory)105は、上述の視覚認識機能部101と、聴覚認識機能部102と、接触認識機能部103によって外部環境から認識されたターゲットやイベントを短期間保持する機能モジュールである。例えば、カメラ15からの入力画像を約15秒程度の短い期間だけ記憶する。

【0074】

長期記憶部(Long Term Memory)106は、物の名前など学習により得られた情報を長期間保持するために使用される。長期記憶部106は、例えば、ある行動モジュールにおいて外部刺激から内部状態の変化を連想記憶することができる。

【0075】

また、本実施形態に係るロボット装置1の行動制御は、反射行動部109によって実現される「反射行動」と、状況依存行動階層108によって実現される「状況依存行動」と、熟考行動階層107によって実現される「熟考行動」と、行動の学習と評価を行なう学習行動部110によって実現される「学習行動」に大別される。

【0076】

反射的行動部(Reflexive Situated Behaviors Layer)109は、上述の視覚認識機能部101と、聴覚認識機能部102と、接触認識機能部103によって認識された外部刺激に応じて反射的な機体動作を実現する機能モジュールである。

【0077】

反射行動とは、基本的に、センサ入力された外部情報の認識結果を直接受けて、これを分類して、出力行動を直接決定する行動のことである。例えば、人間の顔を追いかけたり、うなずいたりといった振る舞いは反射行動として実装することが好ましい。

【0078】

状況依存行動階層(Situated Behaviors Layer)108は、短期記憶部105並びに長期記憶部106の記憶されている外部環境や、内部状態管理部104によって管理される内部状態を基に、ロボット装置1が現在置かれている状況に即応した本能的行動又はその他の形態の状況依存型の行動を制御する。

【0079】

状況依存行動階層108は、各行動毎にステートマシン(又は状態遷移モデル)を用意しており、それ以前の行動や状況に依存して、センサ入力された外部情報の認識結果を分類して、行動を機体上で発現する。また、状況依存行動階層108は、内部状態をある範囲に保つための行動(「ホメオスタシス行動」とも呼ぶ)も実現し、内部状態が指定した範囲内を越えた場合には、その内部状態を当該範囲内に戻すための行動が出現し易くなるようにその行動を活性化させる(実際には、内部状態と外部環境の両方を考慮した形で行動が選択される)。状況依存行動は、反射行動に比し、反応時間が遅い。

【0080】

熟考行動階層(Deliberative Layer)107は、短期記憶部105並びに長期記憶部106の記憶内容に基づいて、ロボット装置1の比較的長期にわたる行動計画などを行なう。

【0081】

10

20

30

40

50

熟考行動とは、与えられた状況あるいは人間からの命令により、推論やそれを実現するための計画を立てて行なわれる行動のことである。例えば、ロボットの位置と目標の位置から経路を探索することは熟考行動に相当する。このような推論や計画は、ロボット装置1がインタラクションを保つための反応時間よりも処理時間や計算負荷を要する（すなわち処理時間がかかる）可能性があるので、上記の反射行動や状況依存行動がリアルタイムで反応を返しながらか、熟考行動は推論や計画を行なう。

#### 【0082】

熟考行動階層107や状況依存行動階層108、反射行動部109は、ロボット装置1のハードウェア構成に非依存の上位のアプリケーション・プログラムとして記述することができる。これに対し、ハードウェア依存層制御部（ConfigurationDependentActionsAndReactions）111は、これら上位アプリケーション（「スキーマ」と呼ばれる行動モジュール）からの命令に応じて、関節アクチュエータの駆動などの機体のハードウェア（外部環境）を直接操作する。

#### 【0083】

学習行動部110は、行動の順序を覚える（すなわち運動計画を立てる）、あるいはダイナミクス（ロボット単体の身体的ダイナミクス、道具を使った道具ダイナミクス（後述）を含む）を覚えることを含み、探索学習や、物まね学習などを行なう。

#### 【0084】

図3に示したようなロボット1の行動制御システム100における各機能モジュールは、オブジェクトとして構成される。各オブジェクトは、メッセージ通信と共有メモリを使ったオブジェクト間通信方法によりデータの受け渡しとInvokeを行なうことができる。図4には、本実施形態に係る行動制御システム100のオブジェクト構成を模式的に示している。

#### 【0085】

視覚認識機能部101は、“FaceDetector”、“MultiColorTracker”、“FaceIdentify”という3つのオブジェクトで構成される。

#### 【0086】

FaceDetectorは、画像フレーム中から顔領域を検出するオブジェクトであり、検出結果をFaceIdentifyに出力する。MultiColorTrackerは、色認識を行なうオブジェクトであり、認識結果をFaceIdentify及びShortTermMemory（短期記憶部105を構成するオブジェクト）に出力する。また、FaceIdentifyは、検出された顔画像を手持ちの人物辞書で検索するなどして人物の識別を行ない、顔画像領域の位置、大きさ情報とともに人物のID情報をShortTermMemoryに出力する。

#### 【0087】

聴覚認識機能部102は、“AudioRecog”と“SpeechRecog”という2つのオブジェクトで構成される。AudioRecogは、マイクなどの音声入力装置からの音声データを受け取って、特徴抽出と音声区間検出を行なうオブジェクトであり、音声区間の音声データの特徴量及び音源方向をSpeechRecogやShortTermMemoryに出力する。SpeechRecogは、AudioRecogから受け取った音声特徴量と音声辞書及び構文辞書を使って音声認識を行なうオブジェクトであり、認識された単語のセットをShortTermMemoryに出力する。

#### 【0088】

触覚認識記憶部103は、接触センサからのセンサ入力を認識する“TactileSensor”というオブジェクトで構成され、認識結果はShortTermMemoryや内部状態を管理するオブジェクトであるInternalStateModel（ISM）に出力する。

#### 【0089】

ShortTermMemory（STM）は、短期記憶部105を構成するオブジェ

10

20

30

40

50

クトであり、上述の認識系の各オブジェクトによって外部環境から認識されたターゲットやイベントを短期間保持（例えばカメラ15からの入力画像を約15秒程度の短い期間だけ記憶する）する機能モジュールであり、STMクライアントであるSituatingBehaviorLayerに対して外部刺激の通知（Notify）を定期的に行なう。

【0090】

LongTermMemory（LTM）は、長期記憶部106を構成するオブジェクトであり、物の名前など学習により得られた情報を長期間保持するために使用される。LongTermMemoryは、例えば、ある行動モジュールにおいて外部刺激から内部状態の変化を連想記憶することができる。

10

【0091】

InternalStatusManager（ISM）は、内部状態管理部104を構成するオブジェクトであり、本能や感情といった数種類の情動を数式モデル化して管理しており、上述の認識系の各オブジェクトによって認識された外部刺激（ES：ExternalStimula）に応じてロボット装置1の本能や情動といった内部状態を管理する。

【0092】

SituatingBehaviorLayer（SBL）は状況依存型行動階層108を構成するオブジェクトである。SBLは、ShortTermMemoryのクライアント（STMクライアント）となるオブジェクトであり、ShortTermMemoryからは定期的に外部刺激（ターゲットやイベント）に関する情報の通知（Notify）を受け取ると、スキーマ（schema）すなわち実行すべき行動モジュールを決定する（後述）。

20

【0093】

ReflexiveSituatingBehaviorLayerは、反射的行動部109を構成するオブジェクトであり、上述した認識系の各オブジェクトによって認識された外部刺激に応じて反射的・直接的な機体動作を実行する。例えば、人間の顔を追いかけたり、うなずいたり、障害物の検出により咄嗟に避けたりといった振る舞いを行なう。

【0094】

SituatingBehaviorLayerは、外部刺激や内部状態の変化などの状況に応じて行動を選択する。これに対し、ReflexiveSituatingBehaviorLayerは、外部刺激に応じて反射的を行動する。これら2つのオブジェクトによる行動選択は独立して行なわれるため、互いに選択された行動モジュール（スキーマ）を機体上で実行する場合に、ロボット1のハードウェア・リソースが競合して実現不可能なこともある。ResourceManagerというオブジェクトは、SituatingBehaviorLayerとReflexiveSituatingBehaviorLayerによる行動選択時のハードウェアの競合を調停する。そして、調停結果に基づいて機体動作を実現する各オブジェクトに通知することにより機体が駆動する。

30

【0095】

SoundPerformer、MotionController、LedControllerは、機体動作を実現するオブジェクトである。SoundPerformerは、音声出力を行なうためのオブジェクトであり、ResourceManager経由でSituatingBehaviorLayerから与えられたテキスト・コマンドに応じて音声合成を行ない、ロボット1の機体上のスピーカから音声出力を行なう。また、MotionControllerは、機体上の各関節アクチュエータの動作を行なうためのオブジェクトであり、ResourceManager経由でSituatingBehaviorLayerから手や脚などを動かすコマンドを受けたことに応答して、該当する関節角を計算する。また、LedControllerは、LED19の点滅動作を行なうためのオブジェクトであり、ResourceManager経由でSituating

40

50

ed Behavior Layer からコマンドを受けたことに応答して LED 19 の点滅駆動を行なう。

【0096】

### C. 状況依存行動

本実施形態に係るロボット装置は、外部環境や内部状態に基づいて、状況に即応した本能的行動又はその他の形態の状況依存型の行動を制御することができる。

【0097】

図5には、状況依存行動階層(SBL)108(但し、反射行動部109を含む)による状況依存行動制御の形態を模式的に示している。認識系101~103による外部環境の認識結果は、外部刺激として状況依存行動階層108(反射行動部109を含む)に与えられる。また、認識系による外部環境の認識結果に応じた内部状態の変化も状況依存行動階層108に与えられる。そして、状況依存行動階層108では、外部刺激や内部状態の変化に応じて状況を判断して、行動選択を実現することができる。

10

【0098】

状況依存行動階層108は、行動モジュール毎にステートマシンを用意しており、それ以前の行動や状況に依存して、センサ入力された外部情報の認識結果を分類して、行動を機体上で発現する。行動モジュールは、機体動作を記述し行動実行に伴う状態遷移(ステートマシン)を実現するAction機能と、Action機能において記述された行動の実行を外部刺激や内部状態に応じて評価して状況判断を行なうMonitor機能とを備えたスキーマ(schema)として記述される。図6には、状況依存行動階層108が複数のスキーマによって構成されている様子を模式的に示している。

20

【0099】

状況依存行動階層108(より厳密には、状況依存行動階層108のうち、通常の状態依存行動を制御する階層)は、複数のスキーマが階層的に連結されたツリー構造として構成され、外部刺激や内部状態の変化に応じてより最適なスキーマを統合的に判断して行動制御を行なうようになっている。ツリーは、例えば動物行動学的(Ethological)な状況依存行動を数式化した行動モデルや、感情表現を実行するためのサブツリーなど、複数のサブツリー(又は枝)を含んでいる。

【0100】

各スキーマは外部刺激と内部状態を入力している。また、各スキーマは、少なくともMonitor関数とAction関数を備えている。図7には、スキーマの内部構成を模式的に示している。同図に示すように、スキーマは、所定の事象の発生に従って状態(又はステート)が移り変わっていく状態遷移モデル(ステートマシン)の形式で機体動作を記述したAction関数と、外部刺激や内部状態に応じてAction関数の各状態を評価して活動度レベル値として返すMonitor関数と、Action関数のステートマシンをREADY(準備完了)、ACTIVE(活動中)、SLEEP(待機中)いずれかの状態としてスキーマの状態を記憶管理する状態管理部で構成されている。

30

【0101】

Monitor関数は、外部刺激と内部状態に応じて当該スキーマの活動度レベル(Activation Level: AL値)を算出する関数である。図6に示すようなツリー構造を構成する場合、上位(親)のスキーマは外部刺激と内部状態を引数として下位(子供)のスキーマのMonitor関数をコールすることができ、子供のスキーマはAL値を返り値とする。また、スキーマは自分のAL値を算出するために、さらに子供のスキーマのMonitor関数をコールすることができる。そして、ルートのスキーマには各サブツリーからのAL値が返されるので、外部刺激と内部状態の変化に応じた最適なスキーマすなわち行動を統合的に判断することができる。

40

【0102】

例えばAL値が最も高いスキーマを選択したり、AL値が所定の閾値を越えた2以上のスキーマを選択して並列的に行動実行したりするようにしてもよい(但し、並列実行するときは各スキーマ同士でハードウェア・リソースの競合がないことを前提とする)。

50

## 【0103】

図7で示す例では、スキーマすなわち行動モジュールの管理を行なう行動状態制御部（仮称）は、Monitor関数からの戻り値に基づいて、実行すべき行動を選択し、該当するスキーマのAction関数をコールし、あるいは状態管理部に記憶されているスキーマの状態の移行を指示する。例えば行動誘発評価値としての活動度レベルが最も高いスキーマを選択したり、リソースが競合しないように優先順位に従って複数のスキーマを選択したりする。また、行動状態制御部は、より優先順位の高いスキーマが起動し、リソースの競合が生じた場合、優先順位が下位のスキーマの状態をACTIVEからSLEEPに退避させ、競合状態が解かれるとACTIVEに回復するなど、スキーマの状態を制御する。

10

## 【0104】

D. ロボットの内部状態管理

本実施形態に係るロボットの行動制御システム100では、状況依存行動階層108は内部状態と外部環境によって行動を決定する。

## 【0105】

ロボット装置1の内部状態は、本能や感情といった数種類の情動で構成され、数式モデル化して扱われる。内部状態管理部（ISM: Internal Status Manager）104は、上述した各認識機能部101~103によって認識された外部刺激（ES: External Stimula）と、時間的経過に基づいて、内部状態を管理する。

20

## 【0106】

図8には、内部状態管理部104と他の機能モジュールとの通信経路を模式的に示している。

## 【0107】

短期記憶部105は、外部環境の変化を認識する各認識機能部101~103からの認識結果を、内部状態管理部104と状況依存行動階層108に出力する。内部状態管理部104は、状況依存行動階層108に内部状態を通知する。これに対し、状況依存行動階層108は、連想又は決定した本能や感情の情報を返す。

## 【0108】

また、状況依存行動階層108は、内部状態と外部環境から算出される活動度レベルを基に行動を選択するとともに、選択した行動の実行と完了を短期記憶部105経由で内部状態管理部104に通知する。

30

## 【0109】

内部状態管理部104は、行動毎に内部状態を長期記憶部106に出力する。これに対し、長期記憶部106は、記憶情報を返す。

## 【0110】

バイオリズム管理部は、バイオリズム情報を内部状態管理部104に供給する。

## 【0111】

内部状態の指標は時間経過により変化する。例えば、1次情動すなわち本能であるHunger（空腹感）、Fatigue（疲労）、Sleep（眠気）は、時間経過によりそれぞれ以下のように変化する。

40

## 【0112】

Hunger：おなかが減る（仮想値又はバッテリー残量）

Fatigue：疲れがたまる

Sleep：眠気がたまる

## 【0113】

また、本実施形態では、ロボットの2次情動すなわち感情（Emotion）の要素としてPleasantness（満足度）、Activation（活動度）、Certainty（確信度）を定義しているが、時間経過によりそれぞれ以下のように変化する。

50

## 【0114】

Pleasantness: Neutral (中立) に向かって変化する

Activation: バイオリズムやSleep (眠気) に依存する

Certainty: Attention に依存する

## 【0115】

また、内部状態は、ロボットが動作を実行することによっても変化する。

## 【0116】

例えば、「眠る」という行動を行なうスキーマは、下位の1次情動としてのSleep (眠気) を満たす行動が最終目標としている。状況依存行動階層108では、1次情動としてのSleepと2次情動としてのActivationを基に各スキーマの活動度レベルを算出・比較して、「眠る」スキーマを選択し、この結果、眠るという行動が実現される。

10

## 【0117】

一方、状況依存行動階層108は、眠るという行動の実行完了を短期記憶部105経由で内部状態管理部104に伝達する。これに対し、内部状態管理部104は、「眠る」行動の実行により、1次情動であるSleepの指標値を変更する。

## 【0118】

そして、状況依存行動階層108では、Sleepが満たされた度合いと2次情動としてのActivationを基に各スキーマの活動度レベルを改めて算出・比較する。この結果、優先度が高くなった他のスキーマを選択し、眠るというスキーマから抜ける。

20

## 【0119】

また、内部状態は、ロボットが動作を実行したときに得られるセンサ情報によっても変化する。

## 【0120】

ロボットが動作を実行したときのその動作程度は、各認識機能部101~103によって認識され、短期記憶部105経由で内部状態管理部104に通知される。内部状態管理部104は、この動作程度を例えばFatigue (疲労) として1次情動の変化に反映させることができる。また、この1次情動の変化に応答して、2次情動も変化させることができる。

## 【0121】

30

### E. 聞き真似をするシステム

ロボット装置の自己発達、行動創発を実現する上で、模倣のメカニズムを採り入れることは効果的であると本発明者らは思料する。

## 【0122】

これまで、物まねやジェスチャ認識を行なうロボット装置の研究開発はなされてきた。ロボット装置同士で物まねやジェスチャ認識を行なうこともできるが、人間すなわちケアギバを直接真似ることできれば、ロボット装置の学習や自己発達はより効率的となり、進歩の速度も向上するであろう。

## 【0123】

ここで、人間の模倣メカニズムを実現するには、なぜロボットが人間の模倣を行なうか (動機付け)、人体が持つ自由度の相違のため、適する動きをどのように選択するか、どうやって対応をとっていくのか、といった問題を解決しなければならない。

40

## 【0124】

また、聞き真似を考えた場合、同じ音韻を発する場合であっても、ケアギバとロボット装置における周波数特性の相違や、ケアギバの個体差の問題がある。さらに、人間の場合、同じ音韻であっても、そのときの感情などの内部状態又は身体的制約により特性の異なる発音内容となる。人間の肉声の聞き真似を行なうロボット装置に関する事例は皆無に等しい。

## 【0125】

以下では、ロボットが人間の肉声を基に音韻系列を学習するメカニズムについて提案す

50

る。この聞き真似メカニズムによれば、身体的制約を考慮して、意味のある音韻系列のみを学習することができる。

【0126】

図9には、本発明の実施形態に係る聞き真似システムの構成を模式的に示している。図示の聞き真似システム1000は、蝸牛特性フィルタ・バンク1001と、音声信号学習器1002と、音声合成器1003で構成される。

【0127】

蝸牛特性フィルタ・バンク1001は、音韻が基準周波数の通倍の中心周波数を持つ複数の成分で構成されるという性質を利用して、話者間での同じ音韻系列の相違を見え易くする機能モジュールである。以下では、蝸牛特性フィルタ・バンク1001を、話者毎の音韻系列  $s$  並びに  $s'$  を入力とし、 $s$  と  $s'$  の相違を見え易い形式の系列  $u$  に変換する関数  $K$  として扱う。蝸牛特性フィルタ・バンク1001は、例えば、マイクロフォン16に付随するDSP（前述）内に実装することができる。

【0128】

なお、蝸牛特性フィルタ・バンクは、人間の聴覚末梢系である内耳の蝸牛特性を擬似化したもので、J. M. Kates の提案による (IEEE TRANSACTIONS ON SIGNAL PROCESSING, VOL. 39, NO. 12, DECEMBER 1991 及び IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING VOL. 1, NO. 4, OCTOBER 1993)。一般に、蝸牛には周波数解析機能があると言われており、本実施形態ではデジタル蝸牛モデル中のフィルタ・バンク部分を使用して、音声の周波数解析を行なう。

【0129】

音声信号学習器1002は、入力された音声信号の音声合成パラメータ  $y$  と、ロボット装置自身がこれを聞き真似して発生する音声合成パラメータ  $y_0$  との関係を評価・学習する機能モジュールである。以下では、音声信号学習器1002を、 $y$  を  $y_0$  に変換する学習の評価関数  $H$  として扱う。この評価関数  $H$  は、音韻系列  $u$  を入力とし、これを音声合成のパラメータとなる  $y$  に変換する。なお、学習メカニズムとして、例えばフィードフォワード型ニューラル・ネットワークを適用することができる。音声信号学習器1002は、例えば、認識系のオブジェクトである  $Speech\ Recog$ （前述）内に組み込むことができる。

【0130】

音声合成器1003は、入力された音声合成パラメータ  $y$  に基づいて、出力する音声信号  $s$  を合成する機能モジュールであり、以下では、 $y$  を入力とし  $s$  を出力する関数  $G$  として扱う。音声合成器1003は、例えば出力系のオブジェクト  $Sound\ Performer$ （前述）内に組み込むことができる。

【0131】

図9に示したシステムによる聞き真似のメカニズムは、下式のように記述することができる。

【0132】

- (1)  $y \quad G \quad s$   
 (2)  $s \quad K \quad u$   
 (3)  $u \quad H \quad y_0$

【0133】

すなわち、音韻系列の素となる音声合成パラメータ  $y$  が音声合成関数  $G$  に投入されると、音声信号  $s$  からなる音韻系列が出力される。そして、この音韻系列  $s$  が蝸牛特性フィルタ・バンク1001に入力されると、そのフィルタ関数  $K$  により話者間の相違が吸収された音韻系列  $u$  に変換される。音声信号学習器1002としての学習関数  $H$  は、この音韻系列  $u$  から、聞き真似をするための音声合成パラメータ  $y_0$  を生成する。さらに、この音声合成パラメータ  $y_0$  が音声合成関数  $G$  に投入されると、聞き真似をした音韻系列が発されることになる。

## 【 0 1 3 4 】

ここで、音声信号学習器 1 0 0 2 としての学習関数 H は、当該システムへの入力としての音声合成パラメータ  $y$  と、聞き真似するための音声合成パラメータ  $y_0$  との相違を評価して、学習を行なう。

## 【 0 1 3 5 】

人間の場合、同じ音韻であっても、そのときの感情などの内部状態又は身体的制約により特性の異なる発音内容となるが、図 9 に示したシステム構成では、このような身体的制約は考慮されない。

## 【 0 1 3 6 】

図 1 0 には、このような身体的制約を考慮し、内部状態を理解することができる聞き真似システムの構成例を模式的に示している。図示の聞き真似システム 1 0 0 0 は、蝸牛特性フィルタ・バンク 1 0 0 1 と、音声信号学習器 1 0 0 2 と、音声合成器 1 0 0 3 に加えて、内部状態生成器 1 0 0 4 と、写像器 1 0 0 5 を備えている。

10

## 【 0 1 3 7 】

蝸牛特性フィルタ・バンク 1 0 0 1 は、音韻が基準周波数の低倍の中心周波数を持つ複数の成分で構成されるという性質を利用して、話者間での同じ音韻系列の相違を見え易くする機能モジュールであり、関数 K として扱う。

## 【 0 1 3 8 】

音声信号学習器 1 0 0 2 は、入力された音声信号の音声合成パラメータ  $y$  と、ロボット装置自身がこれを聞き真似して発生する音声合成パラメータ  $y_0$  との関係を評価・学習する機能モジュールであり、評価関数 H として扱う。

20

## 【 0 1 3 9 】

音声合成器 1 0 0 3 は、入力された音声合成パラメータ  $y$  に基づいて、出力する音声信号  $s$  を合成する機能モジュールであり、関数 G として扱う。

## 【 0 1 4 0 】

内部状態生成器 1 0 0 4 は、息や身体伸ばしなどの信号からなる現在の内部状態  $x$  を生成する。

## 【 0 1 4 1 】

写像器 1 0 0 5 は、現在の内部状態  $x$  を音声合成パラメータ  $y$  に射影する。以下では、写像器 1 0 0 5 を、入力  $x$  を  $y$  に変換する関数 F として扱う。

30

## 【 0 1 4 2 】

図 1 0 に示したシステムにおいて、ロボット自身の声から内部状態を理解するメカニズムは、下式のように記述することができる。

## 【 0 1 4 3 】

( 1 )	$x$	F	$y$
( 2 )	$y$	G	$s$
( 3 )	$s$	K	$u$
( 4 )	$u$	H	$y_0$

## 【 0 1 4 4 】

内部状態生成器 1 0 0 4 により内部状態  $x$  が生成されると、写像器 1 0 0 5 により、内部状態を表した音韻系列を生成するための音声合成パラメータ  $y$  に変換される。そして、音声合成パラメータ  $y$  が音声合成関数 G に投入されると、音声信号  $s$  からなる音韻系列が出力される。そして、この音韻系列  $s$  が蝸牛特性フィルタ・バンク 1 0 0 1 に入力されると、そのフィルタ関数 K により話者間の相違が吸収された音韻系列  $u$  に変換される。音声信号学習器 1 0 0 2 としての学習関数 H は、この音韻系列  $u$  から、聞き真似をするための音声合成パラメータ  $y_0$  を生成する。さらに、この音声合成パラメータ  $y_0$  が音声合成関数 G に投入されると、聞き真似をした音韻系列が発されることになる。上式のうち、( 2 ) ~ ( 4 ) は上述した聞き真似のメカニズムに相当する。

40

## 【 0 1 4 5 】

図 1 0 に示したシステムでは、ランダムな発声に対して、自分で入力を分析して学習 (

50

予測誤差、あるいは相関学習)することができる。また、人間が意味を感じた音に対して同じ音を出すと、それを分析して同じく学習(個人性を吸収する学習+発生の強化学習)することができる。

【0146】

したがって、このようなメカニズムによれば、聞いた声が自分の発声でどのパラメータに相当するかを覚えることができる。すなわち、ロボット装置は、意味のある発声<sup>1</sup>を覚えることができる。

【0147】

また、図10に示したシステムにおいて、ケアギバの声からその内部状態を理解するメカニズムは、下式のように記述することができる。

【0148】

(1)  $s'$              $K$              $u'$   
 (2)  $u'$              $H$              $y_o'$   
 (3)  $y_o'$             $y'$   
 (4)  $y'$              $G$              $s$   
 (5)  $y'$              $F^{-1}$          $x'$

【0149】

ケアギバが発する音韻系列  $s'$  が蝸牛特性フィルタ・バンク 1001 に入力されると、そのフィルタ関数  $K$  により話者間の相違が吸収された音韻系列  $u'$  に変換される。音声信号学習器 1002 としての学習関数  $H$  は、この音韻系列  $u'$  から、聞き真似をするための音声合成パラメータ  $y_o'$  を生成する。

【0150】

ここで、音声信号学習器 1002 としての学習関数  $H$  は、ケアギバの音韻系列に基づく音声合成パラメータ  $y'$  と、聞き真似するための音声合成パラメータ  $y_o'$  との相違を評価して、学習を行なう。

【0151】

そして、音声合成パラメータ  $y_o'$  が音声合成関数  $G$  に投入されると、ケアギバの聞き真似としての音韻系列  $s$  が出力される。

【0152】

また、写像器 1005 では、逆関数  $F^{-1}$  を用いて、音声合成パラメータ  $y'$  に相当する内部状態  $x'$  を導き出す。これは、内部状態を逆推定することに相当する。すなわち、人の声を聞いたときに、自分の内部状態が逆写像で求まるという、身体的理解を実現することができる訳である。

【0153】

#### F. 学習行動

学習とは、作業上における知識の増大と技能の習熟の過程を表すものであり、ロボット装置の自己発達を実現する上で欠くことができない機能と言える。

【0154】

ロボット装置の学習行動は、行動の順序を覚える(すなわち運動計画を立てる)、あるいはダイナミクス(ロボット単体の身体的ダイナミクス、道具を使った道具ダイナミクス(後述)を含む)を覚えることを含み、探索学習や、物まね学習などが含まれる。また、条件反動的行動、学習行動、学習済み行動は、本能行動と反射行動に基づいて実装することが可能である。

【0155】

本発明に係るロボット装置は、環境(すなわち外部刺激)とロボット自身の内部状態を含む力学構造を学習するとともに、これを予測に用い、コマンド・シーケンスを決定するという予測学習型の学習行動システムを実現する。

【0156】

10

20

30

40

50

### G - 1 . 予測学習システム

図 1 1 には、本発明に係るロボット装置が備える予測学習器のシステム構成を模式的に示している。

【 0 1 5 7 】

ここで、ロボット装置自身すなわち実ロボットは、人やその他のモデルとなる動物の筋肉や骨格に相当するアクチュエータ及びフレーム、並びに外部刺激を検出する感覚に相当するセンサを備えている。また、ロボット装置の行動制御システムは、現在の内部状態を保持するとともに、センサ入力により外部環境を認識し、内部状態と外部環境に基づいて発現すべき行動を決定することができる（例えば、図 3、図 5 を参照のこと）。そして、行動制御は、実ロボットに対するモータ・コマンドという形態で外部に表出される。

10

【 0 1 5 8 】

一方、予測学習器は、現在観測されている世界、現在目標とする世界、次の時刻において予測される世界、並びに実ロボットが動作した結果観測される次の時刻における世界という 4 つの世界に取り囲まれている。

【 0 1 5 9 】

ここで言う世界とは、ロボット装置の内部状態、ロボット装置の行動を制御するモータ・コマンド、及び、行動実行時にセンサで検出される外部環境を表す各パラメータでそれぞれ記述される。言い換えれば、予測学習器は、上述の各世界を記述する（内部状態、モータ・コマンド、センサ出力）を入力し、それぞれの世界を理解することができる。

【 0 1 6 0 】

20

予測学習は一般に、現在の観測値に基づいて次の時刻における観測値（推定値）を推定するとともに、次の時刻における実際の観測値を測定し、次の時刻における予測された推定値と実際の観測値との偏差や相関などに基づいて、予測を学習し、コマンド・シーケンスの決定に利用していくものである。本実施形態に係る予測学習器は、現在の観測値以外に、現在の目標値を入力に持つという点で特徴がある。すなわち、予測学習器は、目標とする世界すなわち目標値と、現在観測されている世界すなわち観測値を入力として学習と評価を行なう。予測学習器自体は、フォワード型ニューラル・ネットワークやリカレントニューラルネットワーク、又は相関学習器を用いて構成することができる。

【 0 1 6 1 】

本実施形態に係る予測学習器による学習過程について、図 1 2 を参照しながら説明する。

30

【 0 1 6 2 】

同図において、現実のロボット装置に相当する実ロボットは、全身の筋肉や骨格など駆動部分の総称としてのモータと、感情や情動などの内部状態と、周囲の物理環境を外部刺激として計測するセンサを備えている。モータは、行動制御システムなどから供給されるモータ・コマンドによって動作する。また、センサは、モータ・コマンド実行時すなわち実ロボットが動作したときの物理環境を検出する。そして、内部状態は、モータによる動作結果と、センサの検出結果に基づいて変化していく（図 8 を参照のこと）。

【 0 1 6 3 】

現在時刻  $T$  において、実ロボットは、投入されたモータ・コマンドを実行し、そのときのモータの実行結果（関節角度やモータ・トルクなど）と、物理環境若しくは外部刺激を計測したセンサ出力と、モータの実行結果とセンサ出力に基づいて変化した内部状態を、現在の観測値  $O_b(T)$  として、予測学習器に入力する。

40

【 0 1 6 4 】

また、目標値生成器は、現在時刻  $T$  における目標となるモータの実行結果、センサ出力、及び内部状態からなる目標値  $T_a(T)$  を生成し、予測学習器に入力する。

【 0 1 6 5 】

そして、予測学習器は、次の時刻  $T + \Delta T$  におけるモータの実行結果、センサ出力、及び内部状態の推定値  $P_r(T + \Delta T)$  を推定する。この推定は、例えばフォワード型ニューラル・ネットワーク、リカレントニューラルネットワーク又は相関学習に基づく。

50

## 【0166】

一方、実ロボットでは、外部環境と内部状態に応じた行動制御（前述）に基づいて、次の時刻  $T + \Delta T$  における行動すなわちモータ・コマンドが生成され、この結果、次の時刻  $T + \Delta T$  におけるモータの実行結果と、物理環境若しくは外部刺激を計測したセンサ出力と、モータの実行結果とセンサ出力に基づいて変化した内部状態を、次の時刻の観測値  $O_b(T + \Delta T)$  として、予測学習器に入力する。

## 【0167】

そして、予測学習器では、次の時刻  $T + \Delta T$  における推定値  $P_r(T + \Delta T)$  と現実の観測値  $O_b(T + \Delta T)$  に基づいて、予測学習を行なう。予測学習器は、例えば、推定値  $P_r(T + \Delta T)$  と現実の観測値  $O_b(T + \Delta T)$  との偏差や相関に基づいて学習を行なう。

10

## 【0168】

なお、目標生成器は、ユーザ入力に基づいて目標値  $T_a(T)$  を生成してもよいし、ロボット装置の内部状態と外部環境に基づいて状況依存により目標値  $T_a(T)$  を生成するようにしてもよい。

## 【0169】

G-2. 道具を用いたインタラクションの学習

本実施形態に係るロボット装置は、ロボット単体の身体的ダイナミクス以外に、道具を使った道具ダイナミクスを実現することもできる。後者の場合、ロボット装置は、外部環境に対し道具を用いたインタラクションを行なう。

20

## 【0170】

図13には、ロボット装置が道具を使った道具ダイナミクスを学習する予測学習器のシステム構成を模式的に示している。図示の予測学習器は、図11に示したシステム構成と同じ枠組みで、道具の力学系を学習し、それを“道具”として使う。

## 【0171】

図11に示した例では、予測学習器は、実ロボットがモータ・コマンドを実行したときのモータの実行結果（関節角度やモータ・トルクなど）と、物理環境を計測したセンサ出力と、内部状態を取り扱っていた。これに対し、図13に示すシステム構成では、物理環境を操作するために使用した道具と、道具を介した物理環境の操作結果としてのボディー・イメージと、物理環境を取り扱う。これによって、予測学習器には、道具のモデルが学習される。ここで、実ロボットがどの道具を用いたかは、モータ・コマンド及びセンサ出力により検出することができる。

30

## 【0172】

現在時刻  $T$  において、実ロボットは、投入されたモータ・コマンドを実行し、そのときの道具を使ったインタラクションの結果（関節角度やモータ・トルクなど）をボディー・イメージとして、使った道具、及びセンサ出力からなる物理環境とともに、現在の観測値  $O_b(T)$  として、予測学習器に入力する。

## 【0173】

また、目標値生成器は、現在時刻  $T$  における目標となるボディー・イメージと、使う道具、及び物理環境に関する目標値  $T_a(T)$  を生成し、予測学習器に入力する。

40

## 【0174】

そして、予測学習器は、次の時刻  $T + \Delta T$  におけるボディー・イメージと、使う道具、及び物理環境に関する推定値  $P_r(T + \Delta T)$  を推定する。この推定は、例えばフォワード型ニューラル・ネットワーク、又は相関学習に基づく。

## 【0175】

一方、実ロボットでは、外部環境と内部状態に応じた行動制御（前述）に基づいて、次の時刻  $T + \Delta T$  における行動が生成され、この結果、次の時刻  $T + \Delta T$  におけるボディー・イメージと、使う道具、及び物理環境に関する観測値  $O_b(T + \Delta T)$  として、予測学習器に入力する。

## 【0176】

50

そして、予測学習器では、次の時刻  $T + \Delta T$  における推定値  $P_r(T + \Delta T)$  と現実の観測値  $O_b(T + \Delta T)$  に基づいて、予測学習を行なう。予測学習器は、例えば、推定値  $P_r(T + \Delta T)$  と現実の観測値  $O_b(T + \Delta T)$  との偏差や相関に基づいて学習を行なう。この結果、予測学習器には、道具のモデルが学習される。

【0177】

### G-3. 発達のシナリオを持った学習

上述したように、本発明では、学習のメカニズムを取り入れることにより、ロボット装置の自己発達を実現する。

【0178】

ここで、過大なテーマが与えられると、学習の結果がなかなか収束せず、学習効率が上がらないという問題がある。そこで、本発明では、学習メカニズムに発達のシナリオの概念を取り入れることにした。

【0179】

発達のシナリオは、複数の学習フェーズで構成される。各学習フェーズでは、ロボット装置が実現する自由度構成（関節の可動角やトルク限界など）や取るべき行動リスト（本能的行動、反射行動など）を、学習の進行に従って定義している。

【0180】

学習フェーズ毎にファイルで管理することにより、学習の進行に応じたファイルをアクセスし、発達のシナリオを実現することができる。

【0181】

また、学習フェーズの遷移は、時間の経過でコントロールしてもよいが、予測学習器における推定値  $P_r(T + \Delta T)$  と現実の観測値  $O_b(T + \Delta T)$  との偏差や相関に基づいてコントロールすることができる。すなわち、推定値通りに実ロボットが行動できるようになれば、現在の学習フェーズをクリアしたとみなし、次の学習フェーズへ遷移させる。

【0182】

図14には、発達のシナリオを持つ予測学習器のシステム構成を模式的に示している。

【0183】

ロボット装置自身すなわち実ロボットは、人やその他のモデルとなる動物の筋肉や骨格に相当するアクチュエータ及びフレーム、並びに外部刺激を検出する感覚に相当するセンサを備えている。また、ロボット装置の行動制御システムは、現在の内部状態を保持するとともに、センサ入力により外部環境を認識し、内部状態と外部環境に基づいて発現すべき行動を決定する。

【0184】

予測学習器は、目標とする世界すなわち目標値と、現在観測されている世界すなわち観測値を入力として学習と評価を行なう。予測学習器自体は、フォワード型ニューラル・ネットワーク、又は相関学習器を用いて構成することができる。

【0185】

現在時刻  $T$  において、実ロボットは、投入されたモータ・コマンドを実行し、そのときのモータの実行結果（関節角度やモータ・トルクなど）と、物理環境若しくは外部刺激を計測したセンサ出力と、モータの実行結果とセンサ出力に基づいて変化した内部状態を、現在の観測値  $O_b(T)$  として、予測学習器に入力する。

【0186】

ここで、発達のシナリオには、学習フェーズに応じた自由度構成（関節の可動角やトルク限界など）や取るべき行動リスト（本能的行動、反射行動など）を記述した学習フェーズ・ファイルが格納されている。選択器は、実ロボットの現在の学習の度合いに適合した学習フェーズ・ファイルを取り出す。そして、実ロボットは、このファイルの記述内容に従って、モータ・コマンドの実行を行なう。

【0187】

また、目標値生成器は、現在時刻  $T$  における目標となるモータの実行結果、センサ出力、及び内部状態からなる目標値  $T_a(T)$  を生成し、予測学習器に入力する。

10

20

30

40

50

## 【0188】

そして、予測学習器は、次の時刻  $T + \Delta T$  におけるモータの実行結果、センサ出力、及び内部状態の推定値  $P_r(T + \Delta T)$  を推定する。この推定は、例えばフォワード型ニューラル・ネットワーク、又は相関学習に基づく。

## 【0189】

一方、実ロボットでは、外部環境と内部状態に応じた行動制御（前述）に基づいて、次の時刻  $T + \Delta T$  における行動すなわちモータ・コマンドが生成され、この結果、次の時刻  $T + \Delta T$  におけるモータの実行結果と、物理環境若しくは外部刺激を計測したセンサ出力と、モータの実行結果とセンサ出力に基づいて変化した内部状態を、次の時刻の観測値  $O_b(T + \Delta T)$  として、予測学習器に入力する。但し、モータ・コマンドの実行は、学習フェーズの定義に従う（同上）。

10

## 【0190】

そして、予測学習器では、次の時刻  $T + \Delta T$  における推定値  $P_r(T + \Delta T)$  と現実の観測値  $O_b(T + \Delta T)$  に基づいて、予測学習を行なう。予測学習器は、例えば、推定値  $P_r(T + \Delta T)$  と現実の観測値  $O_b(T + \Delta T)$  との偏差や相関に基づいて学習を行なう。

## 【0191】

また、選択器は、時間の経過、又は、推定値  $P_r(T + \Delta T)$  と現実の観測値  $O_b(T + \Delta T)$  との偏差や相関に基づいて、学習フェーズの遷移をコントロールする。すなわち、推定値通りに実ロボットが行動できるようになれば、現在の学習フェーズをクリアしたとみなし、発達のシナリオから、次の学習フェーズに相当するファイルを取り出し、学習を進行させる。

20

## 【0192】

G - 4 : 報酬評価と目標生成

報酬評価は、内部状態及び内部状態をどのような状態にすることが評価が高いか、に基づいて行なうことができる。本能行動や反射行動は、この設計された報酬評価に基づいて行動が選択されていく。一方、目標生成をこの報酬評価に基づいて行なうことができる。高い評価を得た状態を目標として記憶することで、本能反射行動で探索を行ないながら高い評価の状態を目標とし、以降探索を行なわずにこの状態を目標とする行動生成を予測学習器に学習させることができる。

30

## 【0193】

G - 5 : リハースとプランニング

目標に対して学習された予測学習器を用いて、現在の状態から目標の状態までを実際に行動を生成せずに、予測学習器の出力を入力にフィードバックして時刻  $T$ 、 $T + 1$ 、 $T + 2$ 、... と生成し、リハースすることが可能である。この結果を用いて実際の行動を選択し、効率的に目標達成の行動を行なうことができる（図15を参照のこと）。

## 【0194】

G - 6 : 行動出力の調停

本能反射行動と予測学習器の生成する行動のどちらを実際の行動として出力するかを調停することにより、探索や模倣などが本能反射行動として実装されている場合には、本能反射行動の出力を用いて学習しながら行動し、学習が進んだ段階では、偏差を利用することで予測学習器の出力を利用するように調停を行なうことができる。これにより、学習を行ないながら、学習が進んだ状態では予測学習期の出力を利用する自律行動を実現することができる。

40

## 【産業上の利用可能性】

## 【0195】

以上、特定の実施例を参照しながら、本発明について詳解してきた。しかしながら、本発明の要旨を逸脱しない範囲で当業者が該実施例の修正や代用を成し得ることは自明である。

## 【0196】

50

本発明の要旨は、必ずしも「ロボット」と称される製品には限定されない。すなわち、電氣的若しくは磁氣的な作用を用いて人間の動作に似せた運動を行なう機械装置であるならば、例えば玩具等のような他の産業分野に属する製品であっても、同様に本発明を適用することができる。

【0197】

要するに、例示という形態で本発明を開示してきたのであり、本明細書の記載内容を限定的に解釈するべきではない。本発明の要旨を判断するためには、冒頭に記載した特許請求の範囲の欄を参酌すべきである。

【図面の簡単な説明】

【0198】

【図1】図1は、本発明に実施に供されるロボット装置1の機能構成を模式的に示した図である。

【図2】図2は、制御ユニット20の構成を示した図である。

【図3】図3は、本発明の実施形態に係るロボット装置1の行動制御システム100の機能構成を模式的に示した図である。

【図4】図4は、行動制御システム100のオブジェクト構成を模式的に示した図である。

【図5】図5は、状況依存行動階層(SBL)108(但し、反射行動部109を含む)による状況依存行動制御の形態を模式的に示した図である。

【図6】図6は、状況依存行動階層108が複数のスキーマによって構成されている様子を模式的に示した図である。

【図7】図7は、スキーマの内部構成を模式的に示した図である。

【図8】図8は、内部状態管理部104と他の機能モジュールとの通信経路を模式的に示した図である。

【図9】図9は、本発明に係る聞き真似システムの構成を模式的に示した図である。

【図10】図10は、本発明に係る聞き真似システムについての他の構成を模式的に示した図である。

【図11】図11は、本発明に係るロボット装置が備える予測学習器のシステム構成を模式的に示した図である。

【図12】図12は、予測学習器による学習過程を説明するための図である。

【図13】図13は、ロボット装置が道具を使った道具ダイナミクスを学習する予測学習器のシステム構成を模式的に示した図である。

【図14】図14は、発達のシナリオを持つ予測学習器のシステム構成を模式的に示した図である。

【図15】図15は、リハースとプランニングを説明するための図である。

【符号の説明】

【0199】

1 ... ロボット装置

15 ... CCDカメラ

16 ... マイクロフォン

17 ... スピーカ

18 ... タッチ・センサ

19 ... LEDインジケータ

20 ... 制御部ユニット

21 ... CPU

22 ... RAM

23 ... ROM

24 ... 不揮発メモリ

25 ... インターフェース

26 ... 無線通信インターフェース

10

20

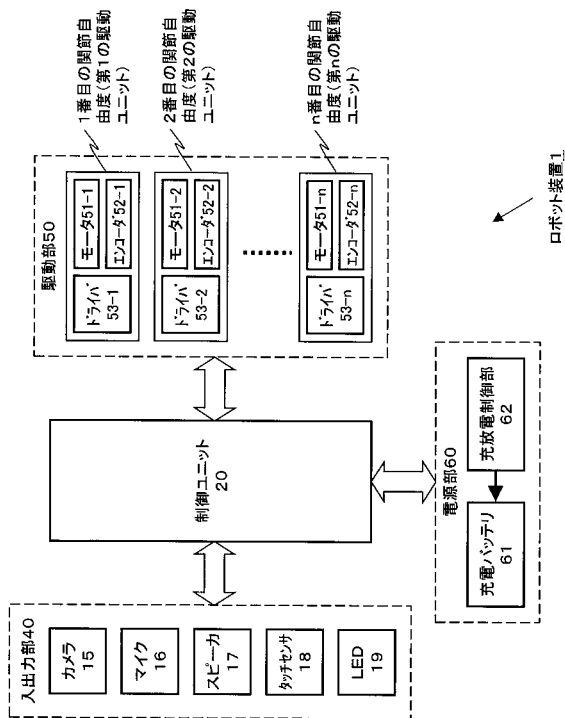
30

40

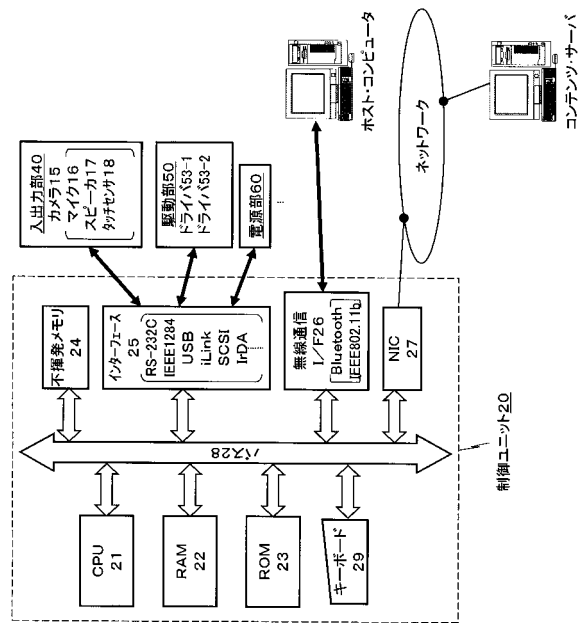
50

- 2 8 ... バス
- 4 0 ... 入出力部
- 5 0 ... 駆動部
- 5 1 ... モータ
- 5 2 ... エンコーダ
- 5 3 ... ドライバ
- 1 0 0 0 ... 聞き真似システム
- 1 0 0 1 ... 蝸牛特性フィルタ・バンク
- 1 0 0 2 ... 音声信号学習器
- 1 0 0 3 ... 音声合成器
- 1 0 0 4 ... 内部状態生成器
- 1 0 0 5 ... 写像器

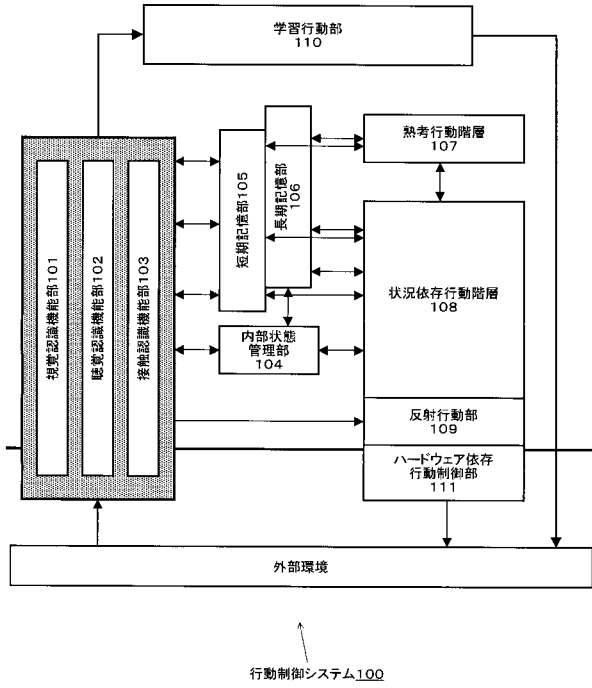
【 図 1 】



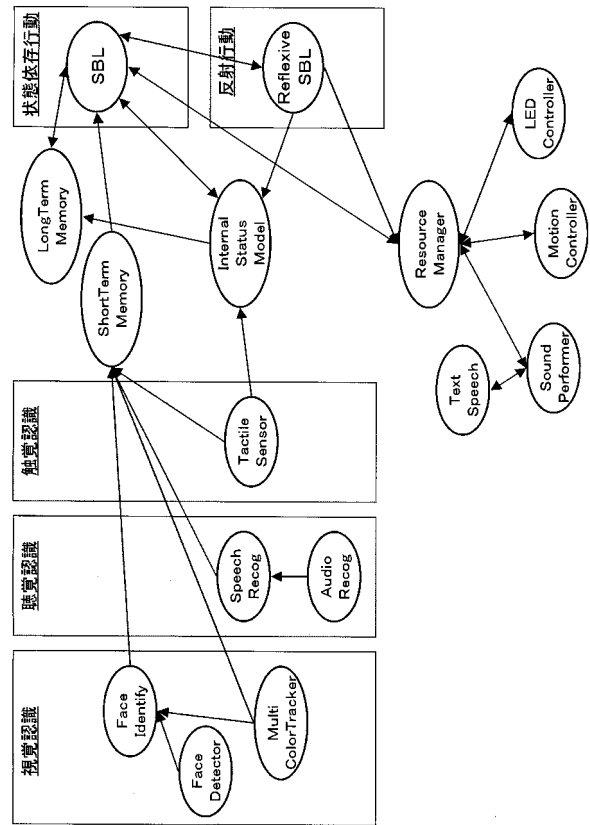
【 図 2 】



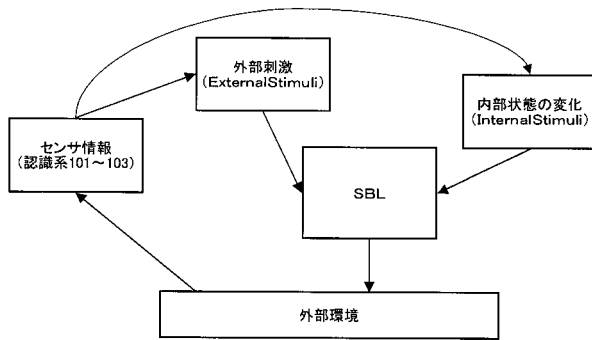
【図3】



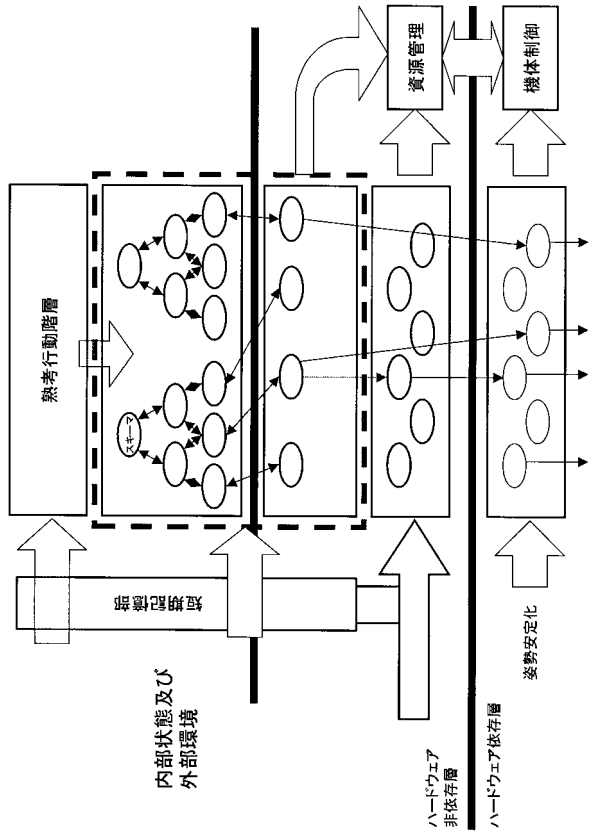
【図4】



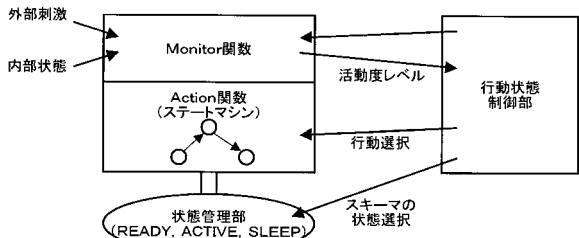
【図5】



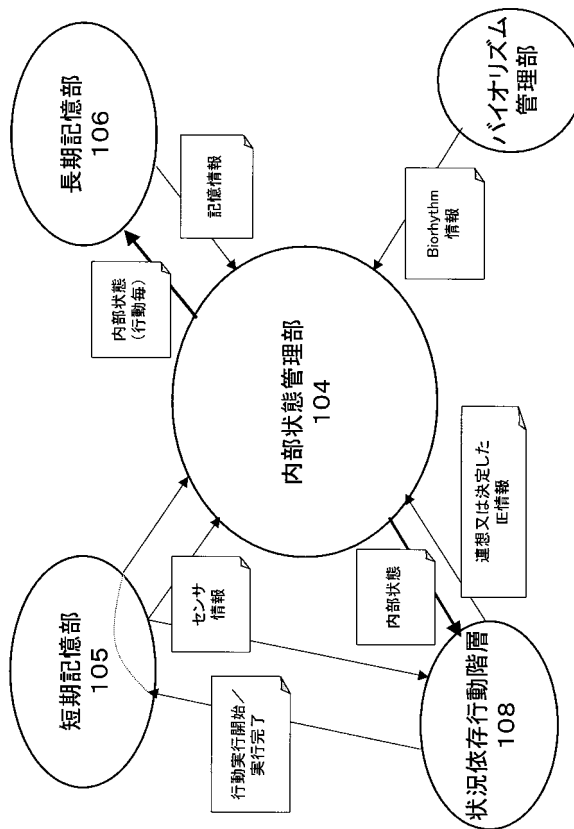
【図6】



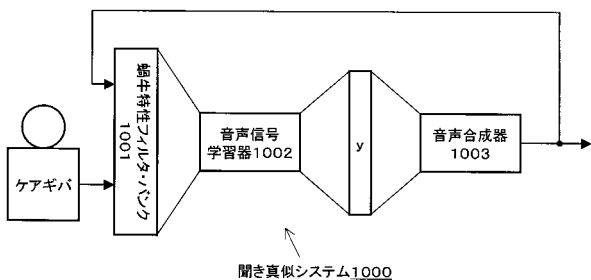
【図7】



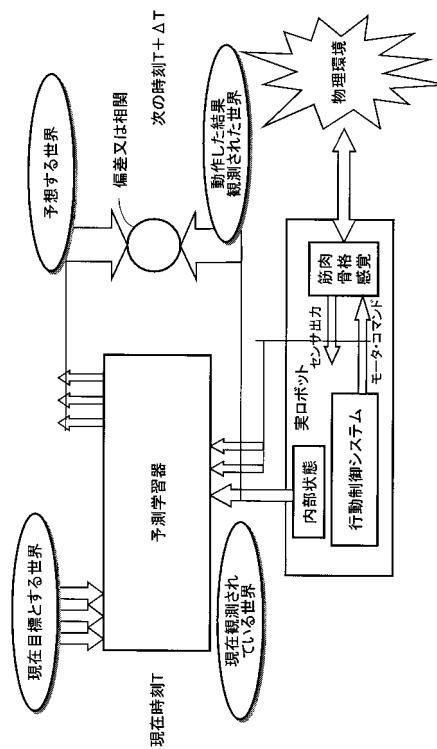
【図8】



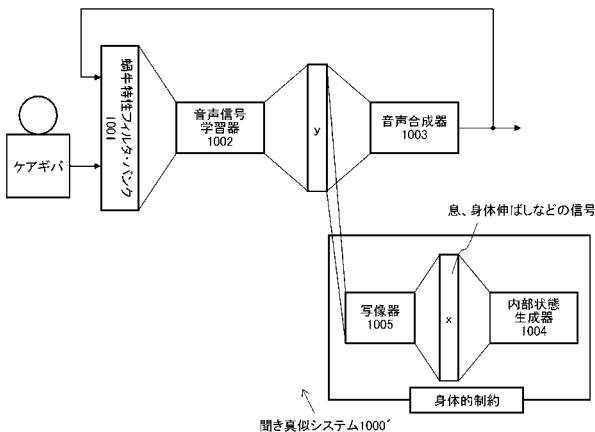
【図9】



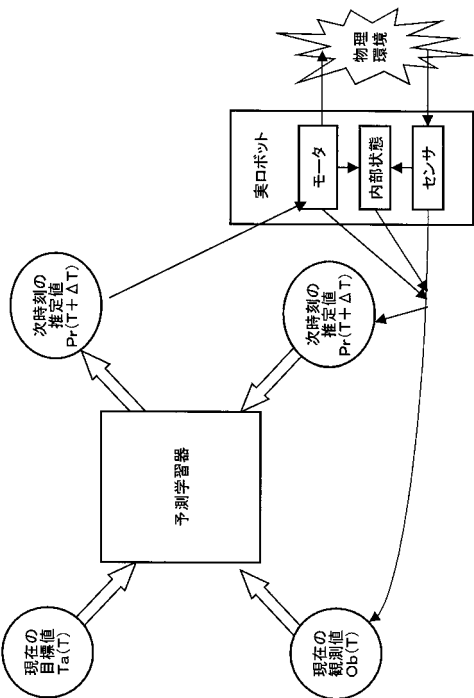
【図11】



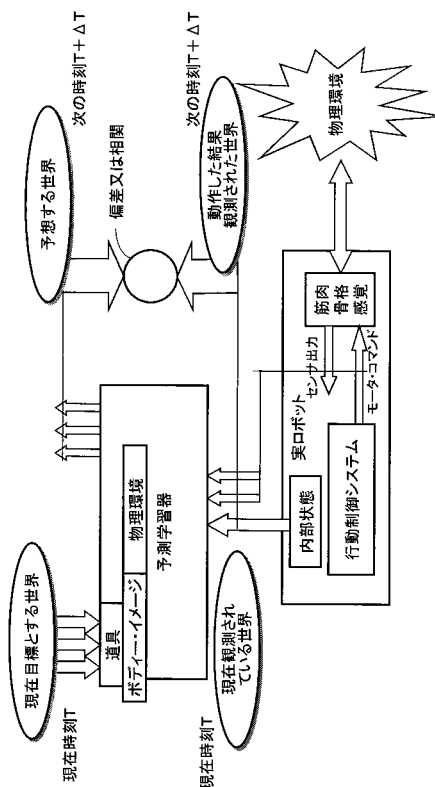
【図10】



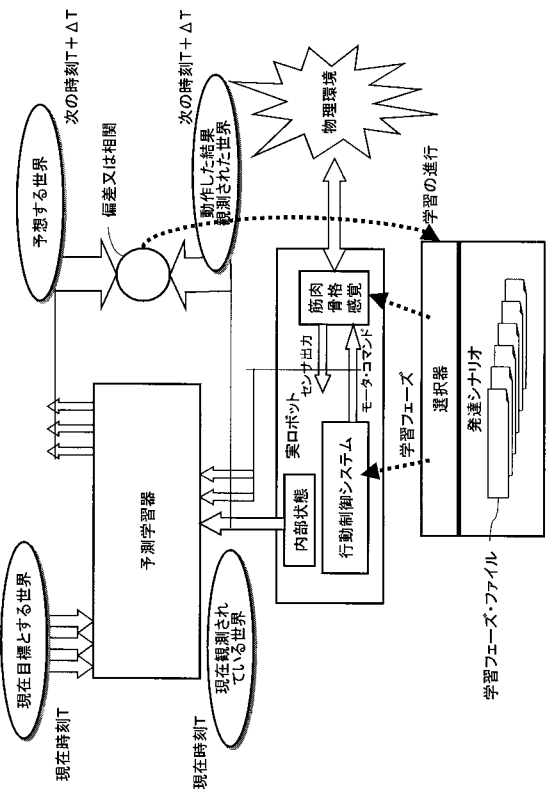
【 図 1 2 】



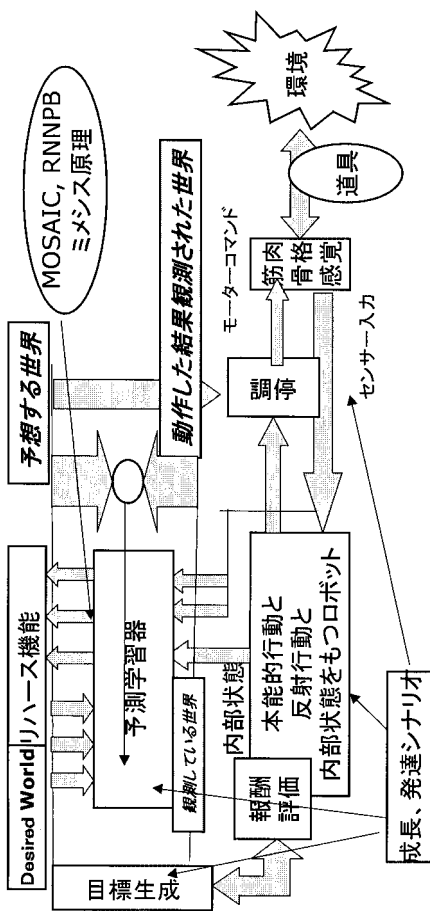
【 図 1 3 】



【 図 1 4 】



【 図 1 5 】



---

フロントページの続き

(51)Int.Cl. F I  
G 1 0 L 15/00 2 0 0 H

(56)参考文献 特開2002-189488(JP,A)  
特開2002-268656(JP,A)  
特開2002-049385(JP,A)  
特開2002-264057(JP,A)  
特開平10-049188(JP,A)  
特開2002-239952(JP,A)  
特開平10-097267(JP,A)  
特開平11-311999(JP,A)  
特開平11-085194(JP,A)  
藤田雅博 他, "Entertainment Robotにおける音によるインタラクション", AIチャレンジ研究会  
(第13回) 人工知能学会研究会資料, 2001年 6月15日, p.29-34

(58)調査した分野(Int.Cl., DB名)  
G 1 0 L 1 1 / 0 0  
1 3 / 0 0 - 1 3 / 0 8  
1 5 / 0 0  
J S T P l u s ( J D r e a m I I )