

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第4635486号
(P4635486)

(45) 発行日 平成23年2月23日 (2011.2.23)

(24) 登録日 平成22年12月3日 (2010.12.3)

(51) Int.Cl.

F I

G06N 5/04 (2006.01)

G06N 5/04 550N

G06N 5/04 580A

請求項の数 8 (全 15 頁)

(21) 出願番号 特願2004-191935 (P2004-191935)
 (22) 出願日 平成16年6月29日 (2004.6.29)
 (65) 公開番号 特開2006-12082 (P2006-12082A)
 (43) 公開日 平成18年1月12日 (2006.1.12)
 審査請求日 平成19年6月14日 (2007.6.14)

(73) 特許権者 000002185
 ソニー株式会社
 東京都港区港南1丁目7番1号
 (74) 代理人 100067736
 弁理士 小池 晃
 (74) 代理人 100086335
 弁理士 田村 榮一
 (74) 代理人 100096677
 弁理士 伊賀 誠司
 (72) 発明者 花形 理
 東京都品川区北品川6丁目7番35号 ソ
 ニー株式会社内
 (72) 発明者 小川 浩明
 東京都品川区北品川6丁目7番35号 ソ
 ニー株式会社内

最終頁に続く

(54) 【発明の名称】 概念獲得装置及びその方法、並びにロボット装置及びその行動制御方法

(57) 【特許請求の範囲】

【請求項 1】

画像から人の顔を認識し、人物IDを出力するとともに、顔の特徴を抽出し、顔の特徴を示す属性情報を出力する顔認識手段と、

音声から発話内容を認識し、テキスト文章を出力する音声認識手段と、

上記属性情報に基づいて顔の特徴の有無を示す正負例を上記人物ID毎に生成するとともに、上記テキスト文章に基づいて発話内容と同じ意味又は反対の意味を示す正負例を上記人物ID毎に生成する正負例生成手段と、

上記属性情報又は上記テキスト文章に基づいて目標概念を決定する目標概念決定手段と

、
 上記目標概念に対して上記正負例を組み合わせ、最も多くの人物IDに対して事例が成り立つものを概念として獲得する概念獲得手段と

を有する概念獲得装置。

【請求項 2】

上記目標概念決定手段は、目標概念が予め規定されたデータベースを有し、該データベースの目標概念と上記属性情報又は上記テキスト文章とを比較し、合致したものを目標概念として決定する請求項1記載の概念獲得装置。

【請求項 3】

上記概念獲得手段にて獲得した概念を記憶する知識記憶手段を有する請求項1記載の概念獲得装置。

【請求項 4】

顔認識手段により、画像から人の顔を認識し、人物IDを出力するとともに、顔の特徴を抽出し、顔の特徴を示す属性情報を出力する顔認識工程と、

音声認識手段により、音声から発話内容を認識し、テキスト文章を出力する音声認識工程と、

情報処理装置により、上記属性情報又は上記テキスト文章に基づいて目標概念を決定する目標概念決定工程と、

情報処理装置により、上記属性情報に基づいて顔の特徴の有無を示す正負例を上記人物ID毎に生成するとともに、上記テキスト文章に基づいて発話内容と同じ意味又は反対の意味を示す正負例を上記人物ID毎に生成する正負例生成工程と、

情報処理装置により、上記目標概念に対して上記正負例を組み合わせ、最も多くの人物IDに対して事例が成り立つものを概念として獲得する概念獲得工程と

を有する概念獲得方法。

【請求項 5】

自律的な動作が可能なロボット装置において、

画像から人の顔を認識し、人物IDを出力するとともに、顔の特徴を抽出し、顔の特徴を示す属性情報を出力する顔認識手段と、

音声から発話内容を認識し、テキスト文章を出力する音声認識手段と、

上記属性情報に基づいて顔の特徴の有無を示す正負例を上記人物ID毎に生成するとともに、上記テキスト文章に基づいて発話内容と同じ意味又は反対の意味を示す正負例を上記人物ID毎に生成する正負例生成手段と、

上記属性情報又は上記テキスト文章に基づいて目標概念を決定する目標概念決定手段と、

上記目標概念に対して上記正負例を組み合わせ、最も多くの人物IDに対して事例が成り立つものを概念として獲得する概念獲得手段と、

上記概念獲得手段にて獲得した概念を記憶する知識記憶手段と

を有するロボット装置。

【請求項 6】

上記属性情報に基づいて上記知識記憶手段に記憶された概念の真偽を識別する概念識別手段と、

上記概念識別手段にて真であると識別された概念を反映させる行動を選択する行動選択手段とを有する請求項5記載のロボット装置。

【請求項 7】

上記目標概念決定手段は、目標概念が予め規定されたデータベースを有し、該データベースの目標概念と上記属性情報又は上記テキスト文章とを比較し、合致したものを目標概念として決定する請求項5記載のロボット装置。

【請求項 8】

顔認識手段により、画像から人の顔を認識し、人物IDを出力するとともに、顔の特徴を抽出し、顔の特徴を示す属性情報を出力する顔認識工程と、

音声認識手段により、音声から発話内容を認識し、テキスト文章を出力する音声認識工程と、

情報処理装置により、上記属性情報又は上記テキスト文章に基づいて目標概念を決定する目標概念決定工程と、

情報処理装置により、上記属性情報に基づいて顔の特徴の有無を示す正負例を上記人物ID毎に生成するとともに、上記テキスト文章に基づいて発話内容と同じ意味又は反対の意味を示す正負例を上記人物ID毎に生成する正負例生成工程と、

情報処理装置により、上記目標概念に対して上記正負例を組み合わせ、最も多くの人物IDに対して事例が成り立つものを概念として獲得する概念獲得工程と、

情報処理装置により、上記概念獲得工程にて獲得した概念を記憶手段に記憶する知識記憶工程と

10

20

30

40

50

を有するロボット装置の行動制御方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、外界から得られた情報と装置内部に記憶されている知識とに基づいて新しい概念を獲得する概念獲得装置及びその方法、並びに概念に基づいて行動するロボット装置及びその行動制御方法に関する。

【背景技術】

【0002】

論理に基づく帰納推論を行うシステム（ILP: Inductive Logic Programing）が90年代後半にStephen Muggletonらによって開発された（例えば、非特許文献1参照。）。ILPは、多くの事例から共通するルールを導き出すシステムである。このILPは、事例として正しい例（正例）と間違っただけの例（負例）が与えられることにより、これらの正負例を事例に関連する背景知識に基づいて組み合わせ、多くの事例が成り立つ仮説を作り出す。作り出された仮説は、概念と呼ばれる。このようにして実現されたILPとして、Progolと呼ばれるシステムがある。

【0003】

【非特許文献1】Stephen Muggleton, Luc De Raedt, “Inductive Logic Programming : Theory And Methods”, Journal of Logic Programming, 1994, pp.629-679

【発明の開示】

【発明が解決しようとする課題】

【0004】

しかしながら、Progolでは、求めようとする概念（目標概念）、背景知識及び正負例はプログラムで与えることになっており、どのように目標概念、背景知識及び正負例を選択するか決まっておらず、新しい概念を自動的に獲得することができなかった。

【0005】

したがって、Progolがロボット装置に組み込まれた場合であっても、新しい概念を獲得することができず、概念に対応した動作を行うことができなかった。

【0006】

本発明は、このような従来の実情に鑑みて提案されたものであり、外界から得られた情報と装置内部に記憶されている知識とに基づいて新しい概念を獲得する概念獲得装置及びその方法、並びに概念に基づいて行動するロボット装置及びその行動制御方法を提供することを目的とする。

【課題を解決するための手段】

【0007】

そこで、上述した目的を達成するために、本発明に係る概念獲得装置は、画像から人の顔を認識し、人物IDを出力するとともに、顔の特徴を抽出し、顔の特徴を示す属性情報を出力する顔認識手段と、音声から発話内容を認識し、テキスト文章を出力する音声認識手段と、上記属性情報に基づいて顔の特徴の有無を示す正負例を上記人物ID毎に生成するとともに、上記テキスト文章に基づいて発話内容と同じ意味又は反対の意味を示す正負例を上記人物ID毎に生成する正負例生成手段と、上記属性情報又は上記テキスト文章に基づいて目標概念を決定する目標概念決定手段と、上記目標概念に対して上記正負例を組み合わせ、最も多くの人物IDに対して事例が成り立つものを概念として獲得する概念獲得手段とを有することを特徴としている。

【0008】

また、本発明に係る概念獲得方法は、顔認識手段により、画像から人の顔を認識し、人物IDを出力するとともに、顔の特徴を抽出し、顔の特徴を示す属性情報を出力する顔認識工程と、音声認識手段により、音声から発話内容を認識し、テキスト文章を出力する音声認識工程と、情報処理装置により、上記属性情報又は上記テキスト文章に基づいて目標概念を決定する目標概念決定工程と、情報処理装置により、上記属性情報に基づいて顔の

10

20

30

40

50

特徴の有無を示す正負例を上記人物ID毎に生成するとともに、上記テキスト文章に基づいて発話内容と同じ意味又は反対の意味を示す正負例を上記人物ID毎に生成する正負例生成工程と、情報処理装置により、上記目標概念に対して上記正負例を組み合わせ、最も多くの人物IDに対して事例が成り立つものを概念として獲得する概念獲得工程とを有することを特徴としている。

【0009】

また、本発明に係るロボット装置は、自律的な動作が可能なロボット装置において、画像から人の顔を認識し、人物IDを出力するとともに、顔の特徴を抽出し、顔の特徴を示す属性情報を出力する顔認識手段と、音声から発話内容を認識し、テキスト文章を出力する音声認識手段と、上記属性情報に基づいて顔の特徴の有無を示す正負例を上記人物ID毎に生成するとともに、上記テキスト文章に基づいて発話内容と同じ意味又は反対の意味を示す正負例を上記人物ID毎に生成する正負例生成手段と、上記属性情報又は上記テキスト文章に基づいて目標概念を決定する目標概念決定手段と、上記目標概念に対して上記正負例を組み合わせ、最も多くの人物IDに対して事例が成り立つものを概念として獲得する概念獲得手段と、上記概念獲得手段にて獲得した概念を記憶する知識記憶手段とを有することを特徴としている。

10

【0010】

また、本発明に係るロボット装置の行動制御方法は、顔認識手段により、画像から人の顔を認識し、人物IDを出力するとともに、顔の特徴を抽出し、顔の特徴を示す属性情報を出力する顔認識工程と、音声認識手段により、音声から発話内容を認識し、テキスト文章を出力する音声認識工程と、情報処理装置により、上記属性情報又は上記テキスト文章に基づいて目標概念を決定する目標概念決定工程と、情報処理装置により、上記属性情報に基づいて顔の特徴の有無を示す正負例を上記人物ID毎に生成するとともに、上記テキスト文章に基づいて発話内容と同じ意味又は反対の意味を示す正負例を上記人物ID毎に生成する正負例生成工程と、情報処理装置により、上記目標概念に対して上記正負例を組み合わせ、最も多くの人物IDに対して事例が成り立つものを概念として獲得する概念獲得工程と、情報処理装置により、上記概念獲得工程にて獲得した概念を記憶手段に記憶する知識記憶工程とを有することを特徴としている。

20

【発明の効果】

【0011】

本発明によれば、外界から得られた情報より目標概念を決定し、正負例を生成することができるため、正負例を組み合わせることで新しい概念を獲得することができる。

30

【0012】

また、外界から得られた情報よりロボット装置の行動に反映させる概念を識別し、行動を選択することができるため、どのような場面でも概念に対応した行動をとることができる。

【発明を実施するための最良の形態】

【0013】

以下、本発明を適用した具体的な実施の形態について、図面を参照しながら詳細に説明する。この実施の形態は、本発明を、外界から得られた知識と装置内部に記憶されている知識とに基づいて新しい概念を獲得し、行動に反映させる2足歩行タイプのロボット装置に適用したものである。このロボット装置は、住環境その他の日常生活上の様々な場面における人的活動を支援する実用ロボットであり、人間が行う基本的な動作を表出できるエンターテインメントロボットでもある。以下ではまず、このようなロボット装置の構成について説明し、次いで、このロボット装置の概念獲得方法及び行動制御方法について詳細に説明する。

40

【0014】

(1) ロボット装置の構成

まず、本実施の形態におけるロボット装置の構成について説明する。図1に示すように、本実施の形態におけるロボット装置1は、体幹部ユニット2の所定の位置に頭部ユニッ

50

ト 3 が連結されるとともに、左右 2 つの腕部ユニット 4 R / L と、左右 2 つの脚部ユニット 5 R / L が連結されて構成されている（但し、R 及び L の各々は、右及び左の各々を示す接尾辞である。以下において同じ。）。

【 0 0 1 5 】

このロボット装置 1 が具備する関節自由度構成を図 2 に模式的に示す。頭部ユニット 3 を支持する首関節は、首関節ヨー軸 1 0 1 と、首関節ピッチ軸 1 0 2 と、首関節ロール軸 1 0 3 という 3 自由度を有している。

【 0 0 1 6 】

また、上肢を構成する各々の腕部ユニット 4 R / L は、肩関節ピッチ軸 1 0 7 と、肩関節ロール軸 1 0 8 と、上腕ヨー軸 1 0 9 と、肘関節ピッチ軸 1 1 0 と、前腕ヨー軸 1 1 1 と、手首関節ピッチ軸 1 1 2 と、手首関節ロール軸 1 1 3 と、手部 1 1 4 とで構成される。手部 1 1 4 は、実際には、複数本の指を含む多関節・多自由度構造体である。但し、手部 1 1 4 の動作は、ロボット装置 1 の姿勢制御や歩行制御に対する寄与や影響が少ないので、本明細書ではゼロ自由度と仮定する。したがって、各腕部ユニット 4 R / L は、7 自由度で構成される。

【 0 0 1 7 】

また、体幹部ユニット 2 は、体幹ピッチ軸 1 0 4 と、体幹ロール軸 1 0 5 と、体幹ヨー軸 1 0 6 という 3 自由度を有する。

【 0 0 1 8 】

また、下肢を構成する各々の脚部ユニット 5 R / L は、股関節ヨー軸 1 1 5 と、股関節ピッチ軸 1 1 6 と、股関節ロール軸 1 1 7 と、膝関節ピッチ軸 1 1 8 と、足首関節ピッチ軸 1 1 9 と、足首関節ロール軸 1 2 0 と、足部 1 2 1 とで構成される。本明細書中では、股関節ピッチ軸 1 1 6 と股関節ロール軸 1 1 7 の交点は、ロボット装置 1 の股関節位置を定義する。人体の足部は、実際には多関節・多自由度の足底を含んだ構造体であるが、ロボット装置 1 の足部 1 2 1 は、ゼロ自由度とする。したがって、各脚部ユニット 5 R / L は、6 自由度で構成される。

【 0 0 1 9 】

以上を総括すれば、ロボット装置 1 全体としては、合計で $3 + 7 \times 2 + 3 + 6 \times 2 = 32$ 自由度を有することになる。但し、エンターテインメント向けのロボット装置 1 が必ずしも 32 自由度に限定される訳ではない。設計・制作上の制約条件や要求仕様等に応じて、自由度すなわち関節数を適宜増減することができることはいうまでもない。

【 0 0 2 0 】

上述したようなロボット装置 1 が持つ各自由度は、実際にはアクチュエータを用いて実装される。外観上で余分な膨らみを排してヒトの自然体形状に近似させること、2 足歩行という不安定構造体に対して姿勢制御を行うことなどの要請から、アクチュエータは小型且つ軽量であることが好ましい。

【 0 0 2 1 】

図 3 には、ロボット装置 1 の制御システム構成を模式的に示している。図 3 に示すように、制御システムは、ユーザ入力などに動的に反応して情緒判断や感情表現を司る思考制御モジュール 2 0 0 と、アクチュエータ 3 5 0 の駆動などロボット装置 1 の全身協調運動を制御する運動制御モジュール 3 0 0 とで構成される。

【 0 0 2 2 】

思考制御モジュール 2 0 0 は、情緒判断や感情表現に関する演算処理を実行する CPU (Central Processing Unit) 2 1 1 や、RAM (Random Access Memory) 2 1 2、ROM (Read Only Memory) 2 1 3、及び外部記憶装置（ハード・ディスク・ドライブなど）2 1 4 で構成され、モジュール内で自己完結した処理を行うことができる独立駆動型の情報処理装置である。

【 0 0 2 3 】

この思考制御モジュール 2 0 0 は、画像入力装置 2 5 1 から入力される画像データや音声入力装置 2 5 2 から入力される音声データなど、外界からの刺激などに従って、ロボッ

10

20

30

40

50

ト装置 1 の現在の感情や意思を決定する。ここで、画像入力装置 2 5 1 は、例えば C C D (Charge Coupled Device) カメラを左右に 1 つずつ備えており、また、音声入力装置 2 5 2 は、例えばマイクロホンを複数備えている。また、思考制御モジュール 2 0 0 は、スピーカを備える音声出力装置 2 5 3 を介して、音声を出力することができる。

【 0 0 2 4 】

また、思考制御モジュール 2 0 0 は、意思決定に基づいた動作又は行動シーケンス、すなわち四肢の運動を実行するように、運動制御モジュール 3 0 0 に対して指令を発行する。

【 0 0 2 5 】

一方の運動制御モジュール 3 0 0 は、ロボット装置 1 の全身協調運動を制御する C P U 3 1 1 や、R A M 3 1 2、R O M 3 1 3、及び外部記憶装置 (ハード・ディスク・ドライブなど) 3 1 4 で構成され、モジュール内で自己完結した処理を行うことができる独立駆動型の情報処理装置である。外部記憶装置 3 1 4 には、例えば、オフラインで算出された歩行パターンや目標とする Z M P 軌道、その他の行動計画を蓄積することができる。ここで、Z M P とは、歩行中の床反力によるモーメントがゼロとなる床面上の点のことであり、また、Z M P 軌道とは、例えばロボット装置 1 の歩行動作期間中に Z M P が動く軌跡を意味する。なお、Z M P の概念並びに Z M P を歩行ロボットの安定度判別規範に適用する点については、Miomir Vukobratovic 著 “LEGGED LOCOMOTION ROBOTS” (加藤一郎外著『歩行ロボットと人工の足』(日刊工業新聞社))に記載されている。

【 0 0 2 6 】

運動制御モジュール 3 0 0 には、図 2 に示したロボット装置 1 の全身に分散するそれぞれの関節自由度を実現するアクチュエータ 3 5 0、体幹部ユニット 2 の姿勢や傾斜を計測する姿勢センサ 3 5 1、左右の足底の離床又は着床を検出する接地確認センサ 3 5 2、3 5 3、バッテリーなどの電源を管理する電源制御装置 3 5 4 などの各種の装置が、バス・インターフェース (I / F) 3 0 1 経由で接続されている。ここで、姿勢センサ 3 5 1 は、例えば加速度センサとジャイロ・センサの組み合わせによって構成され、接地確認センサ 3 5 2、3 5 3 は、近接センサ又はマイクロ・スイッチなどで構成される。

【 0 0 2 7 】

思考制御モジュール 2 0 0 と運動制御モジュール 3 0 0 とは、共通のプラットフォーム上で構築され、両者間はバス・インターフェース 2 0 1、3 0 1 を介して相互接続されている。

【 0 0 2 8 】

運動制御モジュール 3 0 0 では、思考制御モジュール 2 0 0 から指示された行動を体現すべく、各アクチュエータ 3 5 0 による全身協調運動を制御する。すなわち、C P U 3 1 1 は、思考制御モジュール 2 0 0 から指示された行動に応じた動作パターンを外部記憶装置 3 1 4 から取り出し、又は内部的に動作パターンを生成する。そして、C P U 3 1 1 は、指定された動作パターンに従って、足部運動、Z M P 軌道、体幹運動、上肢運動、腰部水平位置及び高さなどを設定するとともに、これらの設定内容に従った動作を指示する指令値を各アクチュエータ 3 5 0 に転送する。

【 0 0 2 9 】

また、C P U 3 1 1 は、姿勢センサ 3 5 1 の出力信号によりロボット装置 1 の体幹部ユニット 2 の姿勢や傾斜を検出するとともに、各接地確認センサ 3 5 2、3 5 3 の出力信号により各脚部ユニット 5 R / L が遊脚又は立脚の何れの状態であるかを検出することによって、ロボット装置 1 の全身協調運動を適応的に制御することができる。

【 0 0 3 0 】

また、C P U 3 1 1 は、Z M P 位置が常に Z M P 安定領域の中心に向かうように、ロボット装置 1 の姿勢や動作を制御する。

【 0 0 3 1 】

さらに、運動制御モジュール 3 0 0 は、思考制御モジュール 2 0 0 において決定された意思通りの行動がどの程度発現されたか、すなわち処理の状況を、思考制御モジュール 2

10

20

30

40

50

00に返すようになっている。

【0032】

このようにしてロボット装置1は、制御プログラムに基づいて自己及び周囲の状況を判断し、自律的に行動することができる。

【0033】

(2) 概念獲得方法

次に、本実施の形態における概念獲得方法について説明する。図4は、本実施の形態における概念獲得装置40を示すブロック図である。概念獲得装置40を搭載したロボット装置1は、外界の情報を取得するセンサに図3に示す画像入力装置251に設けられたCCDカメラ41と音声入力装置252に設けられたマイクロホン43とを用いることができる。なお、ここではILP(Inductive Logic Programming)システムの1つであるP
r o g o lを用いて概念を獲得する方法を説明する。

【0034】

概念獲得装置40は、CCDカメラ41で撮像された画像から人の顔を認識し、その顔の特徴を示す属性情報とその顔の人物IDとを出力する顔認識器42と、マイクロホン43で集音した音声から発話内容を認識し、発話内容のテキスト文章を出力する音声認識器44と、顔認識器42から入力された属性情報又は音声認識器44から入力されたテキスト文章より正負例データベース46を参照して正例又は負例を生成するF a c t生成器45と、音声認識器44から入力されたテキスト文章より目標概念データベース48を参照して目標概念を決定する目標概念決定器47と、F a c t生成器45で生成された正負例及び目標概念決定器47で決定された目標概念から背景知識データベース50を参照して新しい概念を獲得する概念獲得器49と、概念獲得器49で獲得した概念を記憶する知識記憶器51とを備えている。

【0035】

顔認識器42は、例えば平均的な正面顔のテンプレート画像を使用し、CCDカメラ41から入力された画像と、テンプレート画像との相関を求めることにより、顔検出を行う。そして、検出した顔の人物IDを識別し、顔の特徴、例えば、髭が生えているか、眼鏡をかけているかなどを示す属性情報を人物IDとともに出力する。また、顔認識器42は、新しい顔を検出した場合、その顔に人物IDを付すとともに、その顔の特徴を抽出し、人物IDと属性情報とを出力する。顔認識器42から出力された人物ID及び属性情報は、F a c t生成器44に入力される。

【0036】

音声認識器44は、マイクロホン43で集音した音声から発話内容を認識し、発話内容のテキスト文章をF a c t生成器45及び目標概念決定器47に出力する。

【0037】

F a c t生成器45は、音声認識器44から入力されたテキスト文章より正負例を生成する。この際、F a c t生成器45は、正負例データベース46に格納された正負例決定ルールを参照し、正例又は負例を生成する。この正負例決定ルールは、入力されたテキスト文章に対応して正例又は負例を生成するように規定されている。例えば、「わたしはらーめんがすきです」と入力されたテキスト文章に対しては正例である記述like(person, r
amen)を生成するように規定され、その反対の意味を表す、例えば「わたしはらーめんがすきではありません」、「わたしはらーめんがきらいです」などと入力されたテキスト文章に対しては負例である記述not like(person, ramen)を生成するように規定されている。ここで、personには顔認識器42で認識された人のIDが入力される。

【0038】

また、F a c t生成器45は、顔認識器42から入力された属性情報に基づいて正負例を生成する。例えば、眼鏡をかけている人という属性情報が入力された場合には、正例である記述megane(person)を生成し、眼鏡をかけていない人という属性情報が入力された場合には、負例である記述not megane(person)を生成する。F a c t生成器45は、生成した正負例を概念獲得器49に出力する。

【 0 0 3 9 】

目標概念決定器 4 7 は、音声認識器 4 4 から入力されたテキスト文章を、目標概念データベース 4 8 に格納された目標概念決定ルールと比較して、合致したものを目標概念として決定する。目標概念決定ルールは、入力されたテキスト文章に対応して目標概念を決定するように規定されている。例えば、「わたしはらーめんがすきです」と入力されたテキストに対しては、like を目標概念とするように規定されている。目標概念決定器 4 7 は、決定した目標概念を概念獲得器 4 9 に出力する。

【 0 0 4 0 】

概念獲得器 4 9 は、背景知識データベース 5 0 に格納された背景ルールで規定された正負例を組み合わせて最も多くの事例が成り立つ仮説を作り出す。そして、作り出された仮説を概念として獲得する。背景ルールには、目標概念に関する概念を獲得するために利用される正負例が規定されている。例えば、目標概念決定器 4 7 から入力された目標概念 like に対しては、F a c t 生成器 4 5 から出力された髭が生えているか否か、眼鏡をかけているか否かなどの正負例を利用するように規定されている。概念獲得器 4 9 は、獲得した概念を知識記憶器 5 1 に出力する。

10

【 0 0 4 1 】

知識記憶器 5 1 は、概念獲得器 4 9 で獲得された概念を記憶する。記憶された概念は、後述するロボット装置 1 の行動に反映される。

【 0 0 4 2 】

次に、本実施の形態におけるロボット装置 1 の概念獲得方法について、具体的な例を示して説明する。図 5 は、本実施の形態におけるロボット装置 1 の概念獲得方法を説明するフローチャートである。

20

【 0 0 4 3 】

まず、人が、例えば「私はラーメンが好きです。」とロボット装置 1 の C C D カメラ 4 1 の前で話しかける（ステップ S 5 1）。顔認識器 4 2 は、C C D カメラ 4 1 で撮像された顔画像より話しかけた人物を特定する（ステップ S 5 2）。この際、顔認識器 4 2 は、人物 I D と髭が生えているか否か、眼鏡をかけているか否か等の属性情報とを次のようなテーブル型式で出力する。

人物 I D : p1

髭 : true

眼鏡 : false

30

ステップ S 5 1 にて、マイクロホン 4 3 から入力された音声は、音声認識器 4 4 にて発話内容が認識され、テキスト文章が出力される（ステップ S 5 3）。

【 0 0 4 4 】

目標概念決定器 4 7 は、ステップ S 5 3 にて認識されたテキスト文章と目標概念データベース 4 8 に格納された目標概念決定ルールとを比較し、目標概念として決定する（ステップ S 5 4）。例えば、目標概念データベース 4 8 には、次のような目標概念決定ルールが格納されているとする。

ルール：わたしは X がすきです , like

ルール：わたしは X がすきではありません , like

ルール：わたしは X がきらいです , like

40

そして、例えば「わたしはらーめんがすきです。」とステップ S 5 3 で認識された場合、目標概念決定器 4 7 は、その発話内容に対応する目標概念決定ルールに従って like を目標概念として決定する。

【 0 0 4 5 】

F a c t 生成器 4 5 は、ステップ S 5 3 で認識されたテキスト文章に対応する正負例決定ルールに従い、正負例を生成する（ステップ S 5 5）。例えば、正負例データベース 4 6 には、次のような正負例決定ルールが格納されているとする。

ルール：わたしは TARGET がすきです , like(person, TARGET)

ルール：わたしは TARGET がすきではありません , not(like(person, TARGET))

50

ルール：わたしはTARGETがきらいです , not(like(person, TARGET))

ここで、TARGETには、テキスト文章に応じて様々な要素が入力される。そして、Fact生成器45は、ステップS53にて、例えば「わたしはラーメンが大好きです。」と認識されたテキスト文章に対応する正負例決定ルールに従って、正例であるlike(person, ramen)を生成する。

【0046】

また、Fact生成器45は、顔認識器42から出力された属性情報に基づいて、正負例を生成する。例えば、髭が生えている場合、正例であるhige(person)を生成し、髭が生えていない場合、負例であるnot hige(person)を生成する。

【0047】

以上、ステップS51からステップS55までを順次繰り返し行い、正負例を生成する。例えば、p1からp9までのIDの人が、ラーメン及びうどんの好き嫌いについて話した場合、正負例は次のように生成される。

like(p1, ramen).	not like(p1, udon).	not hige(p1).	not megane(p1).
not like(p2, ramen).	like(p2, udon).	hige(p2).	not megane(p2).
like(p3, ramen).	not like(p3, udon).	not hige(p3).	megane(p3).
like(p4, ramen).	not like(p4, udon).	not hige(p4).	megane(p4).
not like(p5, ramen).	not like(p5, udon).	hige(p5).	megane(p5).
like(p6, ramen).	not like(p6, udon).	not hige(p6).	megane(p6).
not like(p7, ramen).	like(p7, udon).	hige(p7).	not megane(p7).
not like(p8, ramen).	like(p8, udon).	hige(p8).	not megane(p8).
not like(p9, ramen).	not like(p9, udon).	not hige(p9).	not megane(p9).

ここで、例えば1段目のp1の人について説明すれば、顔には髭も眼鏡もなく、ラーメンが好きで、うどんが嫌いな人であることを示している。

【0048】

概念獲得器49は、背景知識データベース50に格納された複数のルールの中からステップS54で決定された目標概念に関するルール、例えば次に示すようなlikeに関するルールを取り出す。

```

modeh(1, like(+person, #food))
modeb(1, hige (+person))
modeb(1, megane (+person))
modeb(1, not hige(+person)))
modeb(1, not (megane(+person,)))
modeb(*, +any, +any)

```

ここで、modeh(1, like(+person, #food))の記述は、食べ物#foodに対する目標概念likeについてmodebに記述されている正負例の分類規則を見つけるように規定したものである。また、modeb(*, +any, +any)の記述は、modebに記述されている正負例を組み合わせ分類できるように規定したものである。なお、ラーメン及びうどんについては、背景知識データベース50に食べ物であることが記述されているものとする。

【0049】

このようにステップS55で得られた正負例を組み合わせ最も多くの事例が成り立つ仮説を作り出し、概念を獲得する(ステップS56)。例えば、上述のp1からp9までのIDの人が、ラーメン及びうどんの好き嫌いについて話した場合、次のような概念を獲得する。

```

like(A, ramen) not hige(A)
like(A, udon) not(megane(A))and hige(A)

```

ここで、like(A, ramen) not hige(A)の記述は、「髭を生やしていない人はラーメンが好きである」という概念を表すものである。これは、髭を生やしていない人であるp1、p3、p4、p6及びp9のうち、ラーメンが好きではない人はp9のみであるという事実に基づくものである。また、like(A, udon) not(megane(A))and hige(A)の記述は、「眼鏡をかけ

10

20

30

40

50

ておらず、且つ髭を生やしている人は、うどんが好きである」という概念を表すものである。これは、眼鏡をかけておらず、且つ髭を生やしている人であるp2、p7及びp8は、皆うどんが好きであるという事実に基づくものである。

【0050】

ステップS56にて獲得した概念は、知識記憶器51に記憶される(ステップS57)

。

【0051】

なお、ステップS51において、人がロボット装置1のCCDカメラ41の前で話しかける際に、顔を撮像することとしたが、ロボット装置1が話しかけている人の方向を検知し、その方向にCCDカメラ41を向けて顔を撮像するようにしてもよい。

10

【0052】

また、ステップS54において、音声認識器44で認識されたテキスト文章に基づいて目標概念を決定することとしたが、顔認識器42で認識された顔の属性情報に基づいて目標概念を決定するようにしてもよい。例えば、眼鏡をかけている人という属性情報が入力された場合、視力について目標概念を決定するようにしておき、例えば、「視力が著しく悪い人は眼鏡をかけている」という概念を獲得してもよい。

【0053】

(3) 行動制御方法

本実施の形態における行動制御方法について説明する。図6は、本実施の形態におけるロボット装置1の行動制御に関わる部分を示すブロック図である。ロボット装置1の行動制御に関わる部分は、外界の情報を取得するセンサとして図3に示す画像入力装置251に設けられたCCDカメラ61を用いることができる。

20

【0054】

行動制御に関わる部分は、CCDカメラ61で撮像された画像から人の顔を識別し、その人のIDとその人の顔の特徴を示す属性情報とを出力する顔認識器62と、知識記憶器64に記憶された概念が真か偽かを顔認識器62から入力された属性情報に基づいて識別する概念識別器63と、概念識別器63にて真であると識別された概念を反映させるロボット装置1の行動を行動選択データベース66に基づいて選択する行動選択器65と、音声波形を合成する音声合成器67と、合成された音声波形を出力するスピーカ68とを備えている。

30

【0055】

顔認識器62は、例えば平均的な正面顔のテンプレート画像を使用し、CCDカメラ61から入力された画像と、テンプレート画像との相関を求めることにより、顔検出を行う。そして、検出した顔の人物IDを識別し、顔の特徴、例えば、髭が生えているか、眼鏡をかけているかなどを示す属性情報を人物IDとともに出力する。また、顔認識器62は、新しい顔を検出した場合、その顔に人物IDを付すとともに、その顔の特徴を抽出し、人物IDと属性情報とを出力する。顔認識器62から出力された人物ID及び属性情報は、概念識別器63に入力される。

【0056】

概念識別器63は、顔認識器62から入力された属性情報に基づいて知識記憶器64に記憶された概念が真か偽かを識別する。そして、真であると識別された概念をロボット装置1に反映させる概念として行動選択器65に出力する。

40

【0057】

知識記憶器64は、上述した概念獲得装置40で獲得された概念が記憶されている。

【0058】

行動選択器65は、概念識別器63にて真であると識別された概念を反映させるロボット装置1の行動を、行動選択データベース66に格納された行動選択ルールに基づいて選択する。行動選択ルールは、概念識別器63にて真であると識別された概念に対応する行動が規定されている。

【0059】

50

音声合成器 6 7 は、行動選択器 6 5 にて発話行動が選択された場合、テキスト文章などを音声波形に合成する。合成された音声波形はスピーカ 6 8 から出力される。

【 0 0 6 0 】

次に、本実施の形態におけるロボット装置 1 の行動制御方法について、図 7 のフローチャートを参照して説明する。なお、ここでは、ロボット装置 1 がこれまでに会ったことのない人物に遭遇した場合の行動制御例について説明する。

【 0 0 6 1 】

まず、顔認識器 6 2 は、新しい顔が検出された場合、CCDカメラ 6 1 で撮像された顔画像より話しかけた人物を特定する(ステップ S 7 1)。この際、顔認識器 6 2 は、人物 ID と髭が生えているか否か、眼鏡をかけているか否か等の属性情報とを次のようなテーブル型式で出力する。なお、顔認識器 6 2 は、新しい顔を検出した場合、人物 ID を付すこととしているが、本実施の形態のロボット装置 1 の行動制御方法においては、必ずしも必要なものではない。

人物 ID : p10

髭 : false

眼鏡 : false

概念識別器 6 3 は、顔認識器 6 2 から入力された属性情報に基づいて知識記憶器 6 4 に記憶された概念が真か偽かを識別する(ステップ S 7 2)。そして、真である概念をロボット装置 1 の行動に反映させる概念として行動選択器 6 5 に出力する。例えば、人物 ID p10 の属性情報に基づいて、次のような概念が知識記憶器 6 4 から取得される。

【 0 0 6 2 】

like(A, ramen) not hige(A)

like(A, udon) not(megane(A))and hige(A)

ここで、人物 ID p10 の属性情報を元に概念が真か偽かを識別する。人物 ID p10 の属性情報より like(p10, ramen) が真となることから、この概念をロボット装置 1 の行動に反映させる概念として行動選択器 6 5 に出力する。

【 0 0 6 3 】

行動選択器 6 5 は、ステップ S 7 2 で真と識別された概念を反映させるロボット装置 1 の行動を選択する(ステップ S 7 3)。この選択は、概念識別器 6 3 にて真と識別された概念に対応した行動が規定された行動選択ルールに基づいて行われる。例えば、like(A, ramen) not hige(A) が真であると識別された場合には、次のような行動選択ルールに基づいて行動が選択される。

like(P, OBJ), speak(Pさん " は " O B J " がすきなんだね ")

行動選択器 6 5 は、行動識別子 speak に基づいて発話行動を選択し、speak (" あなた は ラーめん がすきなんだね ") を音声合成器 6 7 に出力する。ここでは、顔認識器 6 2 が新しい顔を検出し、人物名を知らないため、" P さん " には " あなた " が挿入される。なお、人物名を知っている場合は、" P さん " に人物 ID で識別された人物名が挿入される。

【 0 0 6 4 】

音声合成器 6 7 では、行動選択器 6 5 で発話行動が選択された場合、行動選択器 6 5 から入力されたテキスト文章を音声波形に変換し、スピーカ 6 8 を介して出力する(ステップ S 7 4)。したがって、本例では、ロボット装置 1 が、人物 ID p10 の人に「あなたは、ラーメンが好きなんだね」と話す。

【 0 0 6 5 】

このように、ロボット装置 1 の行動制御方法によれば、CCDカメラ 6 1 で撮像された画像情報に基づいて知識記憶器 6 4 に記憶された概念から真となる概念を識別し、その概念に対応した行動を選択することができるため、ロボット装置 1 は、概念に基づいて自律的に行動することができる。

【 0 0 6 6 】

本実施の形態のロボット装置 1 によれば、外界から入力された画像、音声等の情報から

10

20

30

40

50

新しい概念を作り出すことができるだけでなく、その概念に基づいて行動することができる。したがって、ロボット装置 1 が、本実施の形態のように未知の場面に遭遇した場合、それまでの経験を元に概念を獲得していれば、例えば画像情報から真となる概念を識別し、その概念に基づいて適切な行動することができる。

【0067】

なお、本発明は上述した実施の形態のみに限定されるものではなく、本発明の要旨を逸脱しない範囲において種々の変更が可能であることは勿論である。本実施の形態では、獲得した概念を発話行動に反映させることとしたが、これに限定されるものではなく、例えば、「髭を生やしている人は、ダンスが好きである」という概念に対し、髭が生えている属性情報が入力された場合、ロボット装置 1 がダンスをするようにしてもよい。

10

【図面の簡単な説明】

【0068】

【図 1】本実施の形態におけるロボット装置の構成を示す模式図である。

【図 2】本実施の形態におけるロボット装置が具備する関節自由度の構成を示す模式図である。

【図 3】本実施の形態におけるロボット装置の制御システムの構成を示すブロック図である。

【図 4】本実施の形態における概念獲得装置を示すブロック図である。

【図 5】本実施の形態における概念獲得方法を説明するフローチャートである。

【図 6】本実施の形態におけるロボット装置の行動制御方法に関わる部分を示すブロック図である。

20

【図 7】本実施の形態におけるロボット装置の行動制御方法を説明するフローチャートである。

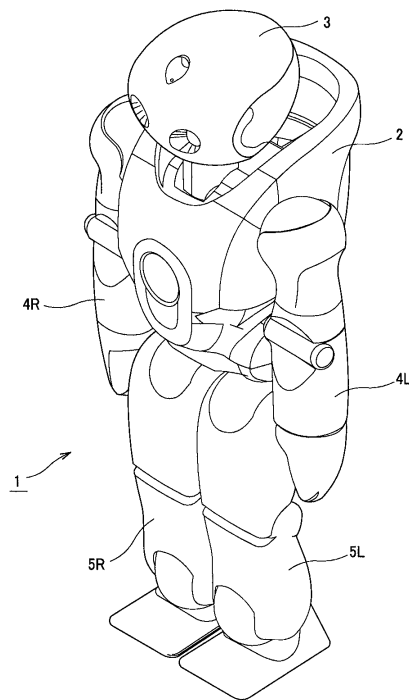
【符号の説明】

【0069】

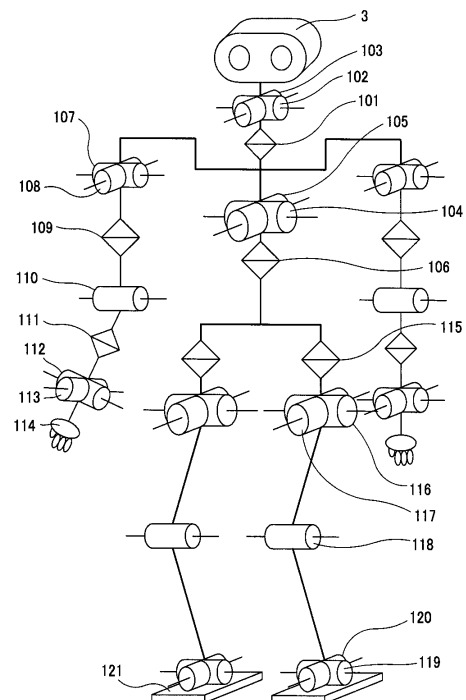
40 概念獲得装置、 41 CCDカメラ、 42 顔認識器、 43 マイクロホン、 44 音声認識器、 45 Fact生成器、 46 正負例データベース、 47 目標概念決定器、 48 目標概念データベース、 49 概念獲得器、 50 背景知識データベース、 51 知識記憶器、 61 CCDカメラ、 62 顔認識器、 63 概念識別器、 64 知識記憶器、 65 行動選択器、 66 行動選択データベース、 67 音声合成器、 68 スピーカ

30

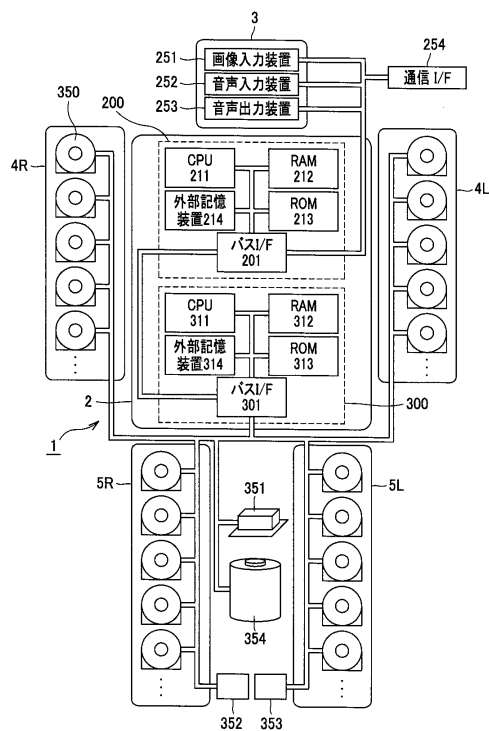
【図 1】



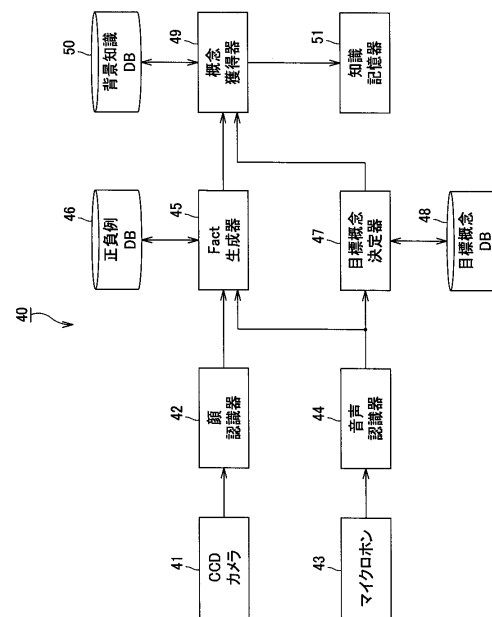
【図 2】



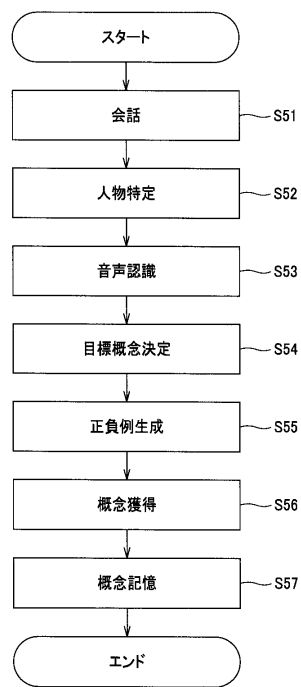
【図 3】



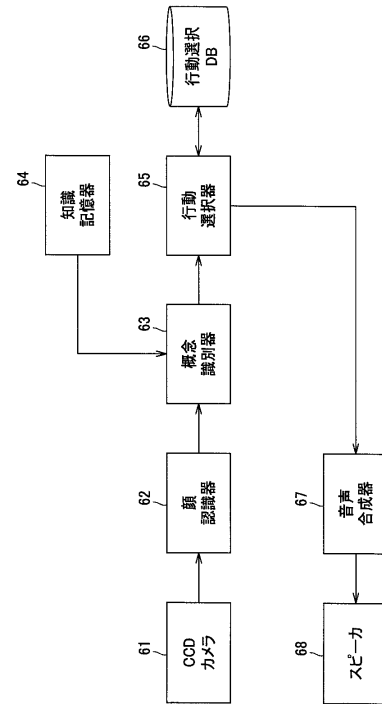
【図 4】



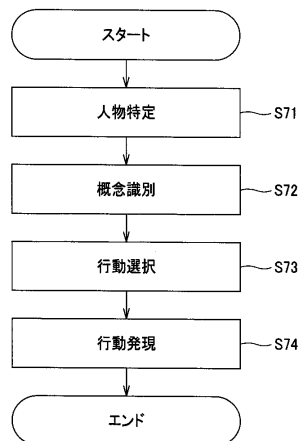
【図 5】



【図 6】



【図 7】



フロントページの続き

審査官 北元 健太

(56)参考文献 特表2002-533827(JP,A)

特開平7-141364(JP,A)

特開平6-222789(JP,A)

秋葉友良、外2名、時間関係と対話性を考慮したマルチモーダル対話記述用スクリプト、電子情報通信学会技術研究報告、日本、社団法人電子情報通信学会、1997年12月12日、第97巻、第440号、pp.71-78

麻生英樹、機械の学習と人間の学習 - 統計的学習の立場から - 、人工知能学会誌、日本、(社)人工知能学会、2003年9月1日、第18巻、第5号、pp.526-530

(58)調査した分野(Int.Cl., DB名)

G06N 3/00 - 3/12

G06N 5/04