

(19)日本国特許庁(JP)

(12)公開特許公報(A)

(11)公開番号

特開2024-23193

(P2024-23193A)

(43)公開日 令和6年2月21日(2024.2.21)

(51)国際特許分類		F I			
G 0 6 F	3/01 (2006.01)	G 0 6 F	3/01	5 1 0	
A 6 3 H	11/00 (2006.01)	A 6 3 H	11/00		Z
B 2 5 J	13/00 (2006.01)	B 2 5 J	13/00		Z

審査請求 有 請求項の数 20 O L (全26頁)

(21)出願番号	特願2023-184256(P2023-184256)	(71)出願人	000002185 ソニーグループ株式会社 東京都港区港南1丁目7番1号
(22)出願日	令和5年10月26日(2023.10.26)	(74)代理人	110002147 弁理士法人酒井国際特許事務所
(62)分割の表示	特願2019-549861(P2019-549861))の分割	(72)発明者	小山 裕一郎 東京都港区港南1丁目7番1号 ソニーグループ株式会社内
原出願日	平成30年8月1日(2018.8.1)	(72)発明者	下江 健晶 東京都港区港南1丁目7番1号 ソニーグループ株式会社内
(31)優先権主張番号	特願2017-209311(P2017-209311)	(72)発明者	小原 一太郎 東京都港区港南1丁目7番1号 ソニーグループ株式会社内
(32)優先日	平成29年10月30日(2017.10.30)		
(33)優先権主張国・地域又は機関	日本国(JP)		

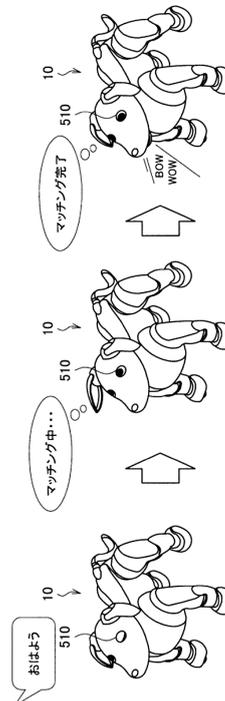
(54)【発明の名称】 情報処理装置及び情報処理方法

(57)【要約】

【課題】認識処理の実行に係るフィードバックをより自然な動作で実現する。

【解決手段】認識処理に基づいて動作体の動作制御を行う動作制御部、を備え、動作制御部は、認識対象情報の入力開始が検知されたことに基づいて第1応答を実行し、認識対象情報の入力完了が検知されたことに基づいて第2応答を実行し、センサ情報から推定される状況に基づいて、第1応答、および第2応答に係る動作種別を動的に決定し、動作体が他の欲求もしくは周囲の環境情報を優先する場合は、第1応答及び第2応答を実行しない、情報処理装置が提供される。

【選択図】図9



10

20

【特許請求の範囲】

【請求項 1】

認識処理に基づいて動作体の動作制御を行う動作制御部、
を備え、

前記動作制御部は、認識対象情報の入力開始が検知されたことに基づいて第 1 応答を実行し、前記認識対象情報の入力完了が検知されたことに基づいて第 2 応答を実行し、

センサ情報から推定される状況に基づいて、前記第 1 応答、および前記第 2 応答に係る動作種別を動的に決定し、

前記動作体が他の欲求もしくは周囲の環境情報を優先する場合は、前記第 1 応答及び前記第 2 応答を実行しない、

情報処理装置。

10

【請求項 2】

前記周囲の環境情報を優先する場合は、前記周囲の環境の安全性が損なわれる場合を含む

請求項 1 に記載の情報処理装置。

【請求項 3】

前記他の欲求を優先する場合は、前記動作体の充電処理を含む

請求項 1 に記載の情報処理装置。

【請求項 4】

前記他の欲求を優先する場合は、前記動作体がユーザに愛されたいという欲求を含む

請求項 1 に記載の情報処理装置。

20

【請求項 5】

前記動作体が他の欲求もしくは周囲の環境情報を優先するかは、前記他の欲求及び前記周囲の環境情報を総合的に判断して決定する

請求項 1 に記載の情報処理装置。

【請求項 6】

前記動作制御部は、前記認識対象情報の入力に基づいて前記動作体に応答動作を実行させ、

前記第 1 応答に係る前記応答動作は、前記認識処理の実行に係る暗示的なフィードバックである、

請求項 1 に記載の情報処理装置。

30

【請求項 7】

前記動作制御部は、認識対象情報の入力開始が検知されたことに基づいて前記動作体に前記第 1 応答を実行させ、認識対象情報の入力完了が検知されたことに基づいて前記動作体に前記第 2 応答を実行させ、

前記第 1 応答に係る動作および前記第 2 応答に係る動作は、前記認識処理の実行に係る暗示的なフィードバックである、

請求項 1 に記載の情報処理装置。

【請求項 8】

前記動作制御部は、前記認識処理が完了したことに基づいて、前記動作体に前記認識処理の実行に係るフィードバックである第 3 応答を実行させる、

請求項 1 に記載の情報処理装置。

40

【請求項 9】

前記動作制御部は、前記第 3 応答の実行後、前記動作体に前記認識処理に基づく行動を実行させる、

請求項 8 に記載の情報処理装置。

【請求項 10】

前記動作体は、生物を模した形態および能力を有し、

前記第 1 応答に係る動作、前記第 2 応答に係る動作、および前記第 3 応答に係る動作は、身体動作、眼球動作、または音を用いた情動表出動作のいずれかを含む、

50

請求項 8 に記載の情報処理装置。

【請求項 1 1】

前記動作制御部は、センサ情報から推定される状況に基づいて、前記第 1 応答、前記第 2 応答、および前記第 3 応答に係る動作種別を動的に決定する、
請求項 8 に記載の情報処理装置。

【請求項 1 2】

前記動作制御部は、ユーザ状態に基づいて、前記第 1 応答、前記第 2 応答、および前記第 3 応答に係る動作種別を動的に決定する、
請求項 8 に記載の情報処理装置。

【請求項 1 3】

前記動作制御部は、前記認識処理に係る信頼度に基づいて、前記第 3 応答に係る動作を動的に決定する、
請求項 8 に記載の情報処理装置。

【請求項 1 4】

前記認識処理は、音声認識処理である、
請求項 8 に記載の情報処理装置。

【請求項 1 5】

前記第 1 応答に係る動作は、眼球動作である、
請求項 1 に記載の情報処理装置。

【請求項 1 6】

前記動作体は、駆動部を有する装置である、
請求項 1 に記載の情報処理装置。

【請求項 1 7】

前記動作制御部は、ユーザの発話対象が前記動作体ではないと判定されたことに基づいて、前記動作体に前記第 1 応答、前記第 2 応答、または前記第 3 応答に係る動作のうち少なくともいずれかを実行させない、
請求項 8 に記載の情報処理装置。

【請求項 1 8】

前記動作制御部は、前記動作体の行動中にユーザの発話開始が検出された場合、前記行動を緩やかに停止させる、
請求項 1 に記載の情報処理装置。

【請求項 1 9】

前記動作制御部は、前記行動の停止後、前記動作体が音を発生しないよう制御する、
請求項 9 に記載の情報処理装置。

【請求項 2 0】

コンピュータが、
認識処理に基づいて動作体の動作制御を行う動作制御部
を含み、

前記動作制御部は、認識対象情報の入力開始が検知されたことに基づいて第 1 応答を実行し、前記認識対象情報の入力完了が検知されたことに基づいて第 2 応答を実行し、

センサ情報から推定される状況に基づいて、前記第 1 応答、および前記第 2 応答に係る動作種別を動的に決定し、

前記動作体が他の欲求もしくは周囲の環境情報を優先する場合は、前記第 1 応答及び前記第 2 応答を実行しない、
情報処理方法。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本開示は、情報処理装置、情報処理方法、およびプログラムに関する。

【背景技術】

10

20

30

40

50

【0002】

近年、認識処理に基づく動作を行う種々の装置が開発されている。上記の装置には、認識した状況に基づいて自律動作を行うロボットなどの動作体が含まれる。例えば、特許文献1には、状況に応じた自律動作や感情表出を行う脚式移動ロボットが開示されている。

【先行技術文献】

【特許文献】

【0003】

【特許文献1】特開2003-71763号公報

【発明の概要】

【発明が解決しようとする課題】

【0004】

ところで、認識処理に基づく動作を実行する装置では、ユーザに対し、当該認識処理が実行されていることを示すことが重要となる。しかし、特許文献1に記載される脚式移動ロボットのようにランプを用いたフィードバックを行う場合、動作体の趣旨に反する不自然な表現となる場合も想定される。

【0005】

そこで、本開示では、認識処理の実行に係るフィードバックをより自然な動作で実現することが可能な、新規かつ改良された情報処理装置及び情報処理方法を提案する。

【課題を解決するための手段】

【0006】

本開示によれば、認識処理に基づいて動作体の動作制御を行う動作制御部、を備え、前記動作制御部は、認識対象情報の入力開始が検知されたことに基づいて第1応答を実行し、前記認識対象情報の入力完了が検知されたことに基づいて第2応答を実行し、センサ情報から推定される状況に基づいて、前記第1応答、および前記第2応答に係る動作種別を動的に決定し、前記動作体が他の欲求もしくは周囲の環境情報を優先する場合は、前記第1応答及び前記第2応答を実行しない、情報処理装置が提供される。

【0007】

また、本開示によれば、コンピュータが、認識処理に基づいて動作体の動作制御を行う動作制御部を含み、前記動作制御部は、認識対象情報の入力開始が検知されたことに基づいて第1応答を実行し、前記認識対象情報の入力完了が検知されたことに基づいて第2応答を実行し、センサ情報から推定される状況に基づいて、前記第1応答、および前記第2応答に係る動作種別を動的に決定し、前記動作体が他の欲求もしくは周囲の環境情報を優先する場合は、前記第1応答及び前記第2応答を実行しない、情報処理方法が提供される。

【発明の効果】

【0008】

以上説明したように本開示によれば、認識処理の実行に係るフィードバックをより自然な動作で実現することが可能となる。

【0009】

なお、上記の効果は必ずしも限定的なものではなく、上記の効果とともに、または上記の効果に代えて、本明細書に示されたいずれかの効果、または本明細書から把握され得る他の効果が奏されてもよい。

【図面の簡単な説明】

【0010】

【図1】本開示の一実施形態に係る自律移動体のハードウェア構成例を示す図である。

【図2】本開示の一実施形態に係る自律移動体が備えるアクチュエータの構成例である。

【図3】本開示の一実施形態に係る自律移動体が備えるアクチュエータの動作について説明するための図である。

【図4】本開示の一実施形態に係る自律移動体が備えるアクチュエータの動作について説明するための図である。

10

20

30

40

50

【図 5】本開示の一実施形態に係る自律移動体が備えるディスプレイの機能について説明するための図である。

【図 6】本開示の一実施形態に係る自律移動体の動作例を示す図である。

【図 7】本開示の第 1 の実施形態に係る自律移動体 10 の機能構成例を示す機能ブロック図である。

【図 8】同実施形態に係る比較手法による動作制御の一例を示す図である。

【図 9】同実施形態に係る動作制御の概要を示す図である。

【図 10】同実施形態に係る情報処理方法と比較手法との差異について説明するための図である。

【図 11】同実施形態に係る第 1 応答の一例を示す図である。

10

【図 12】同実施形態に係る第 2 応答の一例を示す図である。

【図 13】同実施形態に係る第 3 応答の一例を示す図である。

【図 14】同実施形態に係る状況に基づく動作種別の動的制御について説明するための図である。

【図 15】同実施形態に係る発話対象の認識に基づく動作制御について説明するための図である。

【図 16】同実施形態に係る応答動作への移行制御について説明するための図である。

【図 17】同実施形態に係る仮想空間における動作体の制御について説明するための図である。

【図 18】同実施形態に係る動作制御の流れを示すフローチャートである。

20

【図 19】本開示の一実施形態に係る動作制御装置のハードウェア構成例を示す図である。

【発明を実施するための形態】

【0011】

以下に添付図面を参照しながら、本開示の好適な実施の形態について詳細に説明する。なお、本明細書及び図面において、実質的に同一の機能構成を有する構成要素については、同一の符号を付することにより重複説明を省略する。

【0012】

なお、説明は以下の順序で行うものとする。

1. 自律移動体 10 の概要
2. 自律移動体 10 のハードウェア構成例
3. 第 1 の実施形態
 - 3.1. 概要
 - 3.2. 機能構成例
 - 3.3. 動作制御の具体例
 - 3.4. 制御の流れ
4. 動作制御装置のハードウェア構成例
5. まとめ

30

【0013】

< 1. 自律移動体 10 の概要 >

まず、本開示の一実施形態に係る自律移動体 10 の概要について説明する。本開示の一実施形態に係る自律移動体 10 は、収集したセンサ情報に基づく状況推定を実行し、状況に応じた種々の動作を自律的に選択し実行する情報処理装置である。自律移動体 10 は、単にユーザの指示コマンドに従った動作を行うロボットとは異なり、状況ごとに最適であると推測した動作を自律的に実行することを特徴の一つとする。

40

【0014】

このため、本開示の一実施形態に係る自律移動体 10 は、状況によっては、ユーザの指示に対応する動作を敢えて実行しない場合や、当該動作とは異なる他の挙動を実行する場合もある。上記の状況には、例えば、ユーザの指示に対応する動作を行った場合、ユーザや自律移動体 10、または周囲環境の安全性が損なわれる場合や、自律移動体 10 が例え

50

ば、充電処理などの他の欲求（本能）を優先する場合などが相当する。

【0015】

また、自律移動体10は、ユーザの指示に敢えて従わないことで、当該ユーザの興味の誘因を試みることや、自身の感情やハードウェアの状態をユーザに伝達しようと試みる場合もある。

【0016】

一方で、自律移動体10は、ユーザに愛されたいという強い欲求（本能）を有する。このため、自律移動体10は、ユーザを喜ばせるためにユーザの指示に対応する動作を繰り返し実行したり、ユーザが気に入った動作を学習し、指示がない場合でも自発的に当該動作を実行したりする。

10

【0017】

このように、本開示の一実施形態に係る自律移動体10は、ヒトを含む動物と同様に、欲求や感情、また周囲の環境などを総合的に判断して自律動作を決定、実行する。上記の点において、自律移動体10は、指示に基づいて対応する動作や処理を実行する受動的な装置とは明確に相違する。

【0018】

本開示の一実施形態に係る自律移動体10は、空間内を自律的に移動し、種々の動作を実行する自律移動型ロボットであってよい。自律移動体10は、例えば、ヒトやイヌなどの動物を模した形状や、動作能力を有する自律移動型ロボットであってよい。また、自律移動体10は、例えば、ユーザとのコミュニケーション能力を有する車両やその他の装置であってよい。本開示の一実施形態に係る自律移動体10の形状、能力、また欲求などのレベルは、目的や役割に応じて適宜設計され得る。

20

【0019】

< 2 . 自律移動体10のハードウェア構成例 >

次に、本開示の一実施形態に係る自律移動体10のハードウェア構成例について説明する。なお、以下では、自律移動体10がイヌ型の四足歩行ロボットである場合を例に説明する。

【0020】

図1は、本開示の一実施形態に係る自律移動体10のハードウェア構成例を示す図である。図1に示すように、自律移動体10は、頭部、胴部、4つの脚部、および尾部を有するイヌ型の四足歩行ロボットである。また、自律移動体10は、頭部に2つのディスプレイ510を備える。

30

【0021】

また、自律移動体10は、種々のセンサを備える。自律移動体10は、例えば、マイクロフォン515、カメラ520、ToF (Time of Flight) センサ525、人感センサ530、測距センサ535、タッチセンサ540、照度センサ545、足裏ボタン550、慣性センサ555を備える。

【0022】

(マイクロフォン515)

マイクロフォン515は、周囲の音を収集する機能を有する。上記の音には、例えば、ユーザの発話や、周囲の環境音が含まれる。自律移動体10は、例えば、頭部に4つのマイクロフォンを備えてもよい。複数のマイクロフォン515を備えることで、周囲で発生する音を感度高く収集すると共に、音源の定位を実現することが可能となる。

40

【0023】

(カメラ520)

カメラ520は、ユーザや周囲環境を撮像する機能を有する。自律移動体10は、例えば、鼻先と腰部に2つの広角カメラを備えてもよい。この場合、鼻先に配置される広角カメラは、自律移動体の前方視野（すなわち、イヌの視野）に対応した画像を撮像し、腰部の広角カメラは、上方を中心とする周囲領域の画像を撮像する。自律移動体10は、例えば、腰部に配置される広角カメラにより撮像された画像に基づいて、天井の特徴点などを

50

抽出し、SLAM (Simultaneous Localization and Mapping) を実現することができる。

【 0 0 2 4 】

(T o F センサ 5 2 5)

T o F センサ 5 2 5 は、頭部前方に存在する物体との距離を検出する機能を有する。T o F センサ 5 2 5 は、頭部の鼻先に備えられる。T o F センサ 5 2 5 によれば、種々の物体との距離を精度高く検出することができ、ユーザを含む対象物や障害物などとの相対位置に応じた動作を実現することが可能となる。

【 0 0 2 5 】

(人感センサ 5 3 0)

人感センサ 5 3 0 は、ユーザやユーザが飼育するペットなどの所在を検知する機能を有する。人感センサ 5 3 0 は、例えば、胸部に配置される。人感センサ 5 3 0 によれば、前方に存在する動物体を検知することで、当該動物体に対する種々の動作、例えば、興味、恐怖、驚きなどの感情に応じた動作を実現することが可能となる。

【 0 0 2 6 】

(測距センサ 5 3 5)

測距センサ 5 3 5 は、自律移動体 1 0 の前方床面の状況を取得する機能を有する。測距センサ 5 3 5 は、例えば、胸部に配置される。測距センサ 5 3 5 によれば、自律移動体 1 0 の前方床面に存在する物体との距離を精度高く検出することができ、当該物体との相対位置に応じた動作を実現することができる。

【 0 0 2 7 】

(タッチセンサ 5 4 0)

タッチセンサ 5 4 0 は、ユーザによる接触を検知する機能を有する。タッチセンサ 5 4 0 は、例えば、頭頂、あご下、背中など、ユーザが自律移動体 1 0 に対し触れる可能性が高い部位に配置される。タッチセンサ 5 4 0 は、例えば、静電容量式や感圧式のタッチセンサであってよい。タッチセンサ 5 4 0 によれば、ユーザによる触れる、撫でる、叩く、押すなどの接触行為を検知することができ、当該接触行為に応じた動作を行うことが可能となる。

【 0 0 2 8 】

(照度センサ 5 4 5)

照度センサ 5 4 5 は、自律移動体 1 0 が位置する空間の照度を検出する。照度センサ 5 4 5 は、例えば、頭部背面において尾部の付け根などに配置されてもよい。照度センサ 5 4 5 によれば、周囲の明るさを検出し、当該明るさに応じた動作を実行することが可能となる。

【 0 0 2 9 】

(足裏ボタン 5 5 0)

足裏ボタン 5 5 0 は、自律移動体 1 0 の脚部底面が床と接触しているか否かを検知する機能を有する。このために、足裏ボタン 5 5 0 は、4 つの脚部の肉球に該当する部位にそれぞれ配置される。足裏ボタン 5 5 0 によれば、自律移動体 1 0 と床面との接触または非接触を検知することができ、例えば、自律移動体 1 0 がユーザにより抱き上げられたことなどを把握することが可能となる。

【 0 0 3 0 】

(慣性センサ 5 5 5)

慣性センサ 5 5 5 は、頭部や胸部の速度や加速度、回転などの物理量を検出する 6 軸センサである。すなわち、慣性センサ 5 5 5 は、X 軸、Y 軸、Z 軸の加速度および角速度を検出する。慣性センサ 5 5 5 は、頭部および胸部にそれぞれ配置される。慣性センサ 5 5 5 によれば、自律移動体 1 0 の頭部および胸部の運動を精度高く検出し、状況に応じた動作制御を実現することが可能となる。

【 0 0 3 1 】

以上、本開示の一実施形態に係る自律移動体 1 0 が備えるセンサの一例について説明し

10

20

30

40

50

た。なお、図 1 を用いて説明した上記の構成はあくまで一例であり、自律移動体 10 が備え得るセンサの構成は係る例に限定されない。自律移動体 10 は、上記の構成のほか、例えば、温度センサ、地磁気センサ、GNSS (Global Navigation Satellite System) 信号受信機を含む各種の通信装置などをさらに備えてよい。自律移動体 10 が備えるセンサの構成は、仕様や運用に応じて柔軟に変形され得る。

【0032】

続いて、本開示の一実施形態に係る自律移動体 10 の関節部の構成例について説明する。図 2 は、本開示の一実施形態に係る自律移動体 10 が備えるアクチュエータ 570 の構成例である。本開示の一実施形態に係る自律移動体 10 は、図 2 に示す回転箇所に加え、
10

【0033】

例えば、自律移動体 10 は、頭部に 3 自由度を有することで、頷きや首を傾げる動作を両立することができる。また、自律移動体 10 は、腰部に備えるアクチュエータ 570 により、腰のスイング動作を再現することで、より現実のイヌに近い自然かつ柔軟な動作を実現することが可能である。

【0034】

なお、本開示の一実施形態に係る自律移動体 10 は、例えば、1 軸アクチュエータと 2 軸アクチュエータを組み合わせることで、上記の 22 の回転自由度を実現してもよい。例えば、脚部における肘や膝部分においては 1 軸アクチュエータを、肩や大腿の付け根には
20

【0035】

図 3 および図 4 は、本開示の一実施形態に係る自律移動体 10 が備えるアクチュエータ 570 の動作について説明するための図である。図 3 を参照すると、アクチュエータ 570 は、モータ 575 により出力ギアを回転させることで、可動アーム 590 を任意の回転位置および回転速度で駆動させることができる。

【0036】

図 4 を参照すると、本開示の一実施形態に係るアクチュエータ 570 は、リアカバー 571、ギア BOX カバー 572、制御基板 573、ギア BOX ベース 574、モータ 575、第 1 ギア 576、第 2 ギア 577、出力ギア 578、検出用マグネット 579、2 個
30

【0037】

本開示の一実施形態に係るアクチュエータ 570 は、例えば、磁気式 s v G M R (spin-valve Giant Magnetoresistive) であってもよい。制御基板 573 が、メインプロセッサによる制御に基づいて、モータ 575 を回転させることで、第 1 ギア 576 および第 2 ギア 577 を介して出力ギア 578 に動力が伝達され、可動アーム 590 を駆動させることが可能である。

【0038】

また、制御基板 573 に備えられる位置センサが、出力ギア 578 に同期して回転する検出用マグネット 579 の回転角を検出することで、可動アーム 590 の回転角度、すな
40

【0039】

なお、磁気式 s v G M R は、非接触方式であるため耐久性に優れるとともに、G M R 飽和領域において使用することで、検出用マグネット 579 や位置センサの距離変動による信号変動の影響が少ないという利点を有する。

【0040】

以上、本開示の一実施形態に係る自律移動体 10 が備えるアクチュエータ 570 の構成例について説明した。上記の構成によれば、自律移動体 10 が備える関節部の屈伸動作を精度高く制御し、また関節部の回転位置を正確に検出することが可能となる。

【0041】

続いて、図5を参照して、本開示の一実施形態に係る自律移動体10が備えるディスプレイ510の機能について説明する。図5は、本開示の一実施形態に係る自律移動体10が備えるディスプレイ510の機能について説明するための図である。

【0042】

(ディスプレイ510)

ディスプレイ510は、自律移動体10の目の動きや感情を視覚的に表現する機能を有する。図5に示すように、ディスプレイ510は、感情や動作に応じた眼球、瞳孔、瞼の動作を表現することができる。ディスプレイ510は、文字や記号、また眼球運動とは関連しない画像などを取って表示しないことで、実在するイヌなどの動物に近い自然な動作を演出する。

10

【0043】

また、図5に示すように、自律移動体10は、右眼および左眼にそれぞれ相当する2つのディスプレイ510rおよび510lを備える。ディスプレイ510rおよび510lは、例えば、独立した2つのOLED(Organic Light Emitting Diode)により実現される。OLEDによれば、眼球の曲面を再現することが可能となり、1枚の平面ディスプレイにより一对の眼球を表現する場合や、2枚の独立した平面ディスプレイにより2つの眼球をそれぞれ表現する場合と比較して、より自然な外装を実現することができる。

【0044】

以上述べたように、ディスプレイ510rおよび510lによれば、図5に示すような自律移動体10の視線や感情を高精度かつ柔軟に表現することが可能となる。また、ユーザはディスプレイ510に表示される眼球の動作から、自律移動体10の状態を直観的に把握することが可能となる。

20

【0045】

以上、本開示の一実施形態に係る自律移動体10のハードウェア構成例について説明した。上記の構成によれば、図6に示すように、自律移動体10の関節部や眼球の動作を精度高くまた柔軟に制御することで、より実在の生物に近い動作および感情表現を実現することが可能となる。なお、図6は、本開示の一実施形態に係る自律移動体10の動作例を示す図であるが、図6では、自律移動体10の関節部および眼球の動作について着目して説明を行うため、自律移動体10の外部構造を簡略化して示している。本開示の一実施形態に係る自律移動体10のハードウェア構成および外装は、図面により示される例に限定されず、適宜設計され得る。

30

【0046】

< 3 . 第1の実施形態 >

< < 3 . 1 . 概要 > >

次に、本開示の第1の実施形態について説明する。上述したように、本開示の一実施形態に係る自律移動体10(動作体、とも称する)は、イヌ型の情報処理装置であってよい。本開示の一実施形態に係る自律移動体10は、眼球運動による感情表現を除いた視覚情報の出力手段や、音声による言語伝達手段を有しないことを特徴の一つとする。当該特徴によれば、実際のイヌに近いより自然な動作を実現するとともに、自律移動体10が有する機能や外装に対するユーザの違和感を低減することが可能となる。

40

【0047】

しかし、自律移動体10のようにユーザに対する明示的な情報伝達手段を有しない装置の場合、ユーザが装置の状態を明確に把握することが困難な場合がある。例えば、自律移動体10は、ユーザの発話を認識し、当該認識の結果に基づく行動を実行する機能を有する。しかし、スマートフォンなどに搭載される音声認識機能とは異なり、自律移動体10による音声認識では、ユーザが明示的に認識の開始をボタンなどにより指示することがない。このため、ユーザは、認識結果に基づく行動が表出されるまでは、認識処理が実行されているのか否かを判断することが困難である。

【0048】

50

さらには、本開示の一実施形態に係る自律移動体 10 は、上述したように、状況によっては、ユーザの指示に対応する動作を敢えて実行しない場合や、当該動作とは異なる他の挙動を実行する場合もある。このため、認識処理が正常に実行し、かつ自律移動体 10 がユーザの意図に沿わない行動を行った場合、ユーザが、認識処理が失敗した、あるいは認識処理自体が行われなかった、と誤認する可能性も想定される。

【0049】

一方、上記のような可能性を排除するためには、例えば、「認識中です」などの文言を音声または視覚情報により出力することや、ランプの点灯などにより、認識処理の実行に係るフィードバックを明示的に行うことも考えられる。

【0050】

しかし、上述したように、上記のような明示的なフィードバックは、自律移動体 10 の挙動を不自然なものとし、自律移動体 10 に対するユーザの興味や熱中度を低下させる可能性がある。

【0051】

本実施形態に係る技術思想は上記の点に着目して発想されたものであり、認識処理の実行に係るより自然なフィードバックを実現することを可能とする。このために、本実施形態に係る情報処理方法を実現する自律移動体 10 は、認識対象情報の入力に基づいて、認識処理の実行に係る暗示的なフィードバックである応答動作を実行すること、を特徴の一つとする。

【0052】

以下、本実施形態に係る自律移動体 10 が有する上記の特徴と、当該特徴が奏する効果について詳細に説明する。

【0053】

<< 3. 2. 機能構成例 >>

まず、本実施形態に係る自律移動体 10 の機能構成例について説明する。図 7 は、本実施形態に係る自律移動体 10 の機能構成例を示す機能ブロック図である。図 7 を参照すると、本実施形態に係る自律移動体 10 は、入力部 110、認識部 120、学習部 130、行動計画部 140、動作制御部 150、駆動部 160、および出力部 170 を備える。

(入力部 110)

入力部 110 は、ユーザや周囲環境に係る種々の情報を収集する機能を有する。入力部 110 は、例えば、ユーザの発話や周囲で発生する環境音、ユーザや周囲環境に係る画像情報、および種々のセンサ情報を収集する。このために、入力部 110 は、図 1 に示す各種のセンサを備える。

【0054】

(認識部 120)

認識部 120 は、入力部 110 が収集した種々の情報に基づいて、ユーザや周囲環境、また自律移動体 10 の状態に係る種々の認識を行う機能を有する。一例としては、認識部 120 は、人識別、表情や視線の認識、物体認識、色認識、形認識、マーカー認識、障害物認識、段差認識、明るさ認識などを行ってよい。

【0055】

また、認識部 120 は、ユーザの発話に基づく音声認識、単語理解、感情認識、音源定位などを行う。また、認識部 120 は、ユーザなどによる接触や、周囲の温度、動物体の存在、自律移動体 10 の姿勢などを認識することができる。

【0056】

さらには、認識部 120 は、認識した上記の情報に基づいて、自律移動体 10 が置かれた周囲環境や状況を推定し、理解する機能を有する。この際、認識部 120 は、事前に記憶される環境知識を用いて総合的に状況推定を行ってもよい。

【0057】

(学習部 130)

学習部 130 は、環境(状況)と行動、また当該行動による環境への作用を学習する機

10

20

30

40

50

能を有する。学習部 130 は、例えば、深層学習 (Deep Learning) などの機械学習アルゴリズムを用いて、上記の学習を実現する。なお、学習部 130 が採用する学習アルゴリズムは、上記の例に限定されず、適宜設計可能である。

【0058】

(行動計画部 140)

行動計画部 140 は、認識部 120 が推定した状況と学習部 130 が学習した知識に基づいて、自律移動体 10 が行う行動を計画する機能を有する。本実施形態に係る行動計画部 140 は、例えば、認識部 120 が認識したユーザの発話に基づいて、ユーザの発話意図に応じた行動や、ユーザの発話意図に取って代わらない行動を実行することを決定する。

【0059】

(動作制御部 150)

動作制御部 150 は、認識部 120 による認識処理および行動計画部 140 による行動計画に基づいて、駆動部 160 および出力部 170 の動作を制御する機能を有する。動作制御部 150 は、例えば、上記の行動計画に基づいて、アクチュエータ 570 の回転制御や、ディスプレイ 510 の表示制御、スピーカによる音声出力制御などを行う。

【0060】

また、本実施形態に係る動作制御部 150 は、認識対象情報の入力に基づいて、認識処理の実行に係る暗示的なフィードバックである応答動作の実行を制御すること、を特徴の一つとする。本実施形態に係る動作制御部 150 の機能詳細については別途詳細に説明する。

【0061】

(駆動部 160)

駆動部 160 は、動作制御部 150 による制御に基づいて、自律移動体 10 が有する複数の関節部を屈伸させる機能を有する。より具体的には、駆動部 160 は、動作制御部 150 による制御に基づき、各関節部が備えるアクチュエータ 570 を駆動させる。

【0062】

(出力部 170)

出力部 170 は、動作制御部 150 による制御に基づいて、視覚情報や音情報の出力を行う機能を有する。このために、出力部 170 は、ディスプレイ 510 やスピーカを備える。なお、本実施形態に係る出力部 170 は、上述したように、明示的な言語伝達情報を出力しないことを特徴の一つとする。

【0063】

以上、本実施形態に係る自律移動体 10 の機能構成について説明した。なお、図 7 に示す構成はあくまで一例であり、本実施形態に係る自律移動体 10 の機能構成は係る例に限定されない。本実施形態に係る自律移動体 10 は、例えば、情報処理サーバや他の自律移動体と通信を行う通信部などを備えてよい。

【0064】

また、本実施形態に係る認識部 120、学習部 130、行動計画部 140、および動作制御部 150 などは、上記の情報処理サーバ (動作制御装置) の機能として実現されてもよい。この場合、動作制御部 150 は、自律移動体 10 の入力部 110 が収集したセンサ情報に基づいて決定された行動計画に基づいて、自律移動体 10 の駆動部 160 と出力部 170 の制御を行うことが可能である。本実施形態に係る自律移動体 10 の機能構成は仕様や運用に応じて柔軟に変形可能である。

【0065】

<< 3.3. 動作制御の具体例 >>

続いて、本実施形態に係る動作制御の具体例について詳細に説明する。上述したように、本実施形態に係る動作制御部 150 は、認識対象情報の入力に基づいて、認識処理の実行に係る暗示的なフィードバックである応答動作の実行を制御すること、を特徴の一つとする。当該特徴によれば、自律移動体 10 による認識処理の進捗をユーザが直観的に把握することが可能となる。

10

20

30

40

50

【 0 0 6 6 】

なお、以下においては、本実施形態に係る自律移動体 1 0 が音声認識を実行する場合を例に動作制御部 1 5 0 の機能について説明する。しかし、本実施形態に係る認識処理は係る例に限定されず、本実施形態に係る技術思想は、種々の認識処理や推定処理に適用可能である。本実施形態に係る動作制御部 1 5 0 は、例えば、物体認識、話者認識もしくは声紋認識、マーカー認識、感情推定などに係る暗示的なフィードバックを制御することが可能である。

【 0 0 6 7 】

ここで、まず、本実施形態に係る情報処理方法に対する比較手法による動作制御について説明する。上述したように、自律移動体 1 0 は、ユーザの発話を正しく認識した場合でも、ユーザの意図に沿わない行動を行う場合がある。このような場合、ユーザは、音声認識が実行されたか否かを判断することが困難であり、当該行動を自律移動体の不具合と誤認する可能性がある。

10

【 0 0 6 8 】

上記の可能性を排除するため、認識に基づく行動とは別途に、音声認識の完了を示すフィードバックを行うことも想定される。

【 0 0 6 9 】

図 8 は、比較手法による動作制御の一例を示す図である。図 8 には、比較手法に係る動作体 9 0 がユーザの発話に係る音声認識処理を実行する際の状態変化が時系列に示されている。

20

【 0 0 7 0 】

なお、本実施形態では、音声認識処理が、信号処理、発話検出、パターン認識、発話理解により実現され、発話理解については、取得パターンの辞書マッチングが採用される場合を例に説明する。しかし、上記はあくまで一例であり、本実施形態に係る情報処理方法は、種々の音声認識手法に適用され得る。

【 0 0 7 1 】

図 8 の左側には、ユーザの発話の開始を検出した際の動作体 9 0 の状態が、図 8 の中央には、ユーザの発話の終了を検出しマッチングを開始した際の動作体 9 0 の状態が、図 8 の右側には、マッチングが完了した際の動作体 9 0 の状態がそれぞれ示されている。

【 0 0 7 2 】

図示するように、比較手法では、マッチングが完了した際に、動作体 9 0 に耳を動かす動作を実行させることで、ユーザに対し音声認識処理が完了したことをフィードバックしている。当該制御によれば、この後に動作体 9 0 が意図に沿わない行動を行った場合であっても、音声認識処理が実行されたことをユーザが把握できる。

30

【 0 0 7 3 】

しかし、ユーザは、マッチング完了時における耳の動きを確認するまでは、音声認識処理が開始されたことや、実行中であることを把握することが困難である。

【 0 0 7 4 】

そこで、本実施形態に係る動作制御部 1 5 0 は、認識対象情報の入力開始が検知されたことに基づいて自律移動体 1 0 に第 1 応答を実行させ、認識対象情報の入力終了が検知されたことに基づいて自律移動体 1 0 に第 2 応答を実行させることで、上記の点を解決する。なお、上記の認識対象情報とは、本例の場合、ユーザの発話を指す。

40

【 0 0 7 5 】

図 9 は、本実施形態に係る動作制御の概要を示す図である。図 9 には、図 8 と同様に、発話開始検出時、発話完了検出時、マッチング完了時における自律移動体 1 0 の状態が時系列に示されている。

【 0 0 7 6 】

まず、本実施形態に係る動作制御部 1 5 0 は、認識部 1 2 0 により発話の開始が検出された際、眼球動作を用いた第 1 応答を出力部 1 7 0 に実行させてよい。上記の眼球動作は、ディスプレイ 5 1 0 により実現される。当該第 1 応答によれば、ユーザが、自律移動体

50

10 がユーザの発話に反応したことを、より少ない遅延で把握することができる。また、当該第1応答によれば、無音のフィードバックをユーザに対し行うことができ、アクチュエータ570の駆動音やスピーカからの音出力により音声認識の精度が低下することを効果的に防止することが可能となる。このように、眼球動作に係る視覚情報の出力による暗示的なフィードバックは、駆動部を備える音声認識装置に対し高い効果を奏する。

【0077】

次に、動作制御部150は、認識部120により発話の終了が検出され、マッチングが開始される際、駆動部160に耳を上げる動作を実行させてよい。当該第2応答によれば、自律移動体10がユーザの発話に反応し耳を傾けている動作を演出することができ、音声認識処理が実行されていることをユーザが直観的に把握することが可能となる。

10

【0078】

また、本実施形態に係る動作制御部150は、マッチング、すなわち認識処理が完了したことに基づいて、当該認識処理の完了を示すフィードバックである第3応答を自律移動体10に実行させる。動作制御部150は、例えば、耳を下げる動作と口を開く動作を駆動部160に実行させ、鳴き声に相当する音を出力部170に出力させてもよい。

【0079】

上記の第3応答によれば、音声認識処理が実行されたことをユーザが明確に把握することができる。なお、動作制御部150は、第3応答の実行後、行動計画部140が音声認識結果に基づいて計画した行動に相当する動作を自律移動体10に実行させてよい。なお、上述したように上記の行動は、ユーザの発話意図に沿わない行動である場合もある。

20

【0080】

図10は、音声認識処理および応答動作における本実施形態に係る情報処理方法と比較手法との差異について説明するための図である。図10には、本実施形態に係る情報処理方法および比較手法のそれぞれにおける声認識処理および応答動作の対応関係が時系列に示されている。なお、図中におけるMethod 1は上記比較手法に、Method 2は本実施形態に係る情報処理方法にそれぞれ対応する。

【0081】

図10を参照すると、比較手法では、ユーザに対する応答動作が行われるまで、約800msの時間を要することがわかる。このため、ユーザは、音声認識処理の完了を示す応答動作により当該処理が実行されたことを把握できるものの、約800msの間は、動作体が無反応であることに違和感を覚える可能性がある。

30

【0082】

一方、本実施形態に係る情報処理方法では、ユーザの「おはよう」という発話が始まってから程なく第1応答が行われ、発話の終了検出に伴うマッチングの開始時に第2応答が遅滞なく実行される。このように、本実施形態に係る情報処理方法によれば、ユーザの発話開始直後から、段階的に複数のフィードバックを行うことができる。当該手法によれば、ユーザが、マッチングが完了する前の段階であっても、自律移動体10がユーザの発話を理解しようとしていることを把握することが可能となる。

【0083】

次に、本実施形態に係る第1応答の具体例について詳細に説明する。上述したように、本実施形態に係る第1応答は、眼球動作であってもよい。

40

【0084】

図11は、本実施形態に係る第1応答の一例を示す図である。図11には、動作制御部150により制御されるディスプレイ510の時系列における変化が示されている。具体的には、本実施形態に係る動作制御部150は、認識部120によりユーザの発話開始が検出された場合、ディスプレイ510に瞬きに相当する画像を表示させてもよい。

【0085】

なお、動作制御部150は、図示するような瞬きのほか、ユーザに視線を合わせる表現や、ウィンクなどの表現をディスプレイ510に出力させてもよい。

【0086】

50

このように、本実施形態に係る動作制御部 150 は、第 1 応答として眼球動作に係る表示を出力部 170 に実行させることで、音声認識処理を妨げることなく、またユーザの発話に対するより遅滞の少ないフィードバックを実現することができる。

【0087】

なお、本実施形態に係る動作制御部 150 は、眼球動作のほか、アクチュエータ 570 の駆動を伴う身体動作や、音を用いた情動表出動作を第 1 応答として自律移動体 10 に実行させてもよい。なお、上記の音を用いた情動表出動作には、鳴き声、羽音、拍手などのノンバーバル動作が広く含まれる。

【0088】

この場合、アクチュエータ 570 の駆動音やスピーカから出力される音により音声認識の精度が低下する可能性も想定されるが、例えば、スピーカとマイクロフォンの位置関係が一定である場合には、リファレンス信号を用いたエコーキャンセルなどを行うことで、認識精度の低下を抑えることが可能である。また、後述するように、第 1 応答に眼球動作を採用しないことで、ユーザの利便性が高まる状況も存在する。

【0089】

次に、本実施形態に係る第 2 応答の具体例について詳細に説明する。本実施形態に係る第 2 応答は、眼球動作、身体動作、音を用いた情動表出動作のいずれか、または組み合わせであってもよい。図 12 は、本実施形態に係る第 2 応答の一例を示す図である。

【0090】

本実施形態に係る動作制御部 150 は、例えば、図 12 の左側に示すように、耳を持ち上げるなどの身体動作を制御してもよい。なお、動作制御部 150 は、耳の他、尾部や脚部などの動作を制御してもよい。

【0091】

一方、動作制御部 150 は、図 12 の右側に示すように、視線を斜め上方に向かせるなどの眼球動作を制御してもよい。また、動作制御部 150 は、軽くうなる、などの情動表出動作を制御することもできる。本実施形態に係る第 2 応答は、例えば、自律移動体 10 のモデルとなる生物種類に応じた、より自然な動作が採用されてよい。

【0092】

次に、本実施形態に係る第 3 応答の具体例について詳細に説明する。本実施形態に係る第 3 応答は、眼球動作、身体動作、音を用いた情動表出動作のいずれか、または組み合わせであってもよい。また、本実施形態に係る動作制御部 150 は、認識処理に係る信頼度に基づいて、第 3 応答の動作を動的に決定することができる。図 13 は、本実施形態に係る第 3 応答の一例を示す図である。

【0093】

認識処理に係る信頼度が高い場合、本実施形態に係る動作制御部 150 は、例えば、図 13 の左側に示すように、ユーザの発話を理解したことを示す肯定的な第 3 応答を自律移動体 10 に実行させてもよい。上記肯定的な挙動には、例えば、喜びや興奮、興味などに相当する情動の表出動作が含まれる。

【0094】

一方、認識処理に係る信頼度が低い場合、本実施形態に係る動作制御部 150 は、例えば、図 13 の右側に示すように、ユーザの再発話を促すための第 3 応答を自律移動体 10 に実行させてもよい。ユーザの再発話を促すための第 3 応答には、例えば、疑問、不安などに相当する情動の表出動作が含まれる。動作制御部 150 は、例えば、首を傾げながら耳を上げる動作を駆動部 160 に実行させてもよい。

【0095】

動作制御部 150 が有する上記の機能によれば、音声認識処理の結果が芳しくないことをユーザが直観的に把握し、再度発話を行うことが可能となる。

【0096】

以上、本実施形態に係る第 1 応答、第 2 応答、および第 3 応答について具体例を挙げて説明した。上述したように、本実施形態に係る第 1 応答、第 2 応答、および第 3 応答は、

10

20

30

40

50

眼球動作、身体動作、音を用いた情動表出動作のいずれか、または組み合わせにより実現されてよい。

【0097】

また、本実施形態に係る動作制御部150は、センサ情報から推定される状況に基づいて、第1応答、第2応答、および第3応答に係る動作種別を動的に決定することも可能である。なお、上記のセンサ情報から推定される状況には、ユーザ、自律移動体10、周囲環境に係る種々の状態・状況が含まれる。

【0098】

図14は、本実施形態に係る状況に基づく動作種別の動的制御について説明するための図である。図14には、ユーザU1が自律移動体10の背後から発話を行っている状況が示されている。この場合、ユーザU1の位置からは、自律移動体10のディスプレイ510を視認できない可能性が高い。

10

【0099】

このため、本実施形態に係る動作制御部150は、自律移動体10の背後から発話が発出された場合には、例えば、尾部を振る身体動作など、眼球動作を用いない応答動作を自律移動体10に実行させてよい。

【0100】

また、例えば、動作制御部150は、周囲の環境音が大きい場合には、眼球動作または身体動作を優先してもよいし、周囲が暗い状況では、身体動作が確認しづらいことから眼球動作や音を用いた情動表出動作を優先してもよい。

20

【0101】

また、動作制御部150は、特に、ユーザ状態に基づいて、第1応答、第2応答、および第3応答に係る動作種別を動的に決定してもよい。例えば、普段は視力矯正具を着用しているユーザが、視力矯正具を着用していないことが検出された場合、動作制御部150は、眼球動作による応答動作を採用せず、音を用いた情動表出動作を優先してよい。

【0102】

また、ユーザが視覚に障害を有すると推定された場合も同様である。認識部120は、例えば、ユーザが所持する白杖などから上記の推定を行うことが可能である。また、認識部120は、自律移動体10の行動に対するユーザの反応から、上記の推定を行ってもよい。なお、聴覚矯正器具や聴覚障がいについても同様のことが言える。

30

【0103】

このように、本実施形態に係る動作制御部150によれば、種々の状況に対応した、より利便性の高いフィードバックを実現することが可能である。

【0104】

また、本実施形態に係る動作制御部150は、ユーザの発話対象に基づく動作制御を行ってよい。図15は、本実施形態に係る発話対象の認識に基づく動作制御について説明するための図である。

【0105】

図15には、電話で会話を行っているユーザU1と、自律的な行動を行っている最中の自律移動体10とが示されている。この際、本実施形態に係る動作制御部150は、ユーザU1の発話対象が自律移動体10ではないと判定されたことに基づいて、第1応答、第2応答、および第3応答のいずれか、またはすべての実行が行われないう制御を行ってよい。

40

【0106】

本実施形態に係る動作制御部150が有する上記の機能によれば、ユーザの意図に沿う場合にのみ応答動作を行うことが可能となり、自律移動体10に対するユーザの評価を向上させる効果が期待される。また、上記機能によれば、不必要な動作制御による電力消費を抑えることも可能である。

【0107】

なお、認識部120は、ユーザが電話を把持していることや、ユーザの視線が自律移動

50

体 10 に向いていないことなどから、ユーザの発話対象が自律移動体 10 ではないと判定することができる。

【0108】

また、動作制御部 150 は、上記判定の確度が所定以上となるまでは、応答動作を自律移動体 10 に実行させてよい。例えば、第 2 応答を実行させた後に、発話対象が自律移動体 10 ではないと判定された場合には、動作制御部 150 は、第 3 応答を実行させずに、自律行動の制御に復帰してよい。

【0109】

次に、自律移動体が何らかの行動を行っている際にユーザの発話が検出された場合の動作制御例について説明する。図 16 は、本実施形態に係る応答動作への移行制御について説明するための図である。図 16 には、自律移動体 10 がボールで遊んでいる際に、ユーザの発話の開始が検出された場合の一例が示されている。

10

【0110】

この際、本実施形態に係る動作制御部 150 は、自律移動体 10 の行動、すなわちボールを追う行動を緩やかに停止させてよい。また、動作制御部 150 は、当該行動の停止後、自律移動体 10 が音を発生しないように制御を行う。

【0111】

動作制御部 150 による上記の制御によれば、突然に行動を停止することでユーザに違和感を与えることなく、かつ停止後にはアクチュエータ 570 を動作させないことで、駆動音による音声認識精度の低下を防ぐことができる。

20

【0112】

なお、行動の停止が間に合わず、アクチュエータ 570 の駆動音の影響により音声認識結果の確信度が低くなった場合には、動作制御部 150 は、図 16 の右側に示すように、再発話を促す第 3 応答を自律移動体 10 に実行させ、第 3 応答の完了後には、自律移動体 10 が音を発生しないように制御してよい。

【0113】

動作制御部 150 による上記の制御によれば、自律移動体 10 に、より自然な動作を実行させながらも、再度の音声認識処理に係る精度を高めることができる。

【0114】

以上、本実施形態に係る動作制御について具体例を挙げながら説明した。上記で述べた動作制御部 150 の機能によれば、実際の生物に近いより自然な動作を行いながら、自律移動体 10 による認識処理の進捗をユーザが直観的に把握することが可能となる。

30

【0115】

なお、上記では、自律移動体 10 が眼球動作、身体動作、音を用いた情動表出動作のいずれか、または組み合わせを制御する場合を例に述べたが、本実施形態に係る動作制御は、認識処理や自律移動体 10 の特性に応じて適宜変形され得る。例えば、認識部 120 がタッチセンサ 540 が収集したセンサ情報に基づいてユーザの接触パターンの認識などを行う場合には、動作制御部 150 は、圧電素子などによる振動を用いた応答動作を自律移動体 10 に実行させてもよい。

【0116】

また、本実施形態に係る自律移動体 10 は、仮想空間上における動作体（仮想動作体、とも称する）であってもよい。図 17 は、本実施形態に係る仮想動作体の制御について説明するための図である。

40

【0117】

図 17 には、情報処理端末 30 を装着するユーザ U2 の視野 FV と、視野 FV に表示される仮想動作体 VO が示されている。情報処理端末 30 は、例えば、ヘッドマウントディスプレイや眼鏡型のウェアラブル装置であってもよい。この際、動作制御部 150 は、情報処理端末 30 や、情報処理端末 30 と通信する情報処理サーバの機能として実現される。情報処理端末 30 や上記の情報処理サーバは、後述する動作制御装置に相当する。

【0118】

50

この場合、動作制御部 150 は、例えば、AR (Augmented Reality)、VR (Virtual Reality)、MR (Mixed Reality) などの技術を用いて仮想動作体 VO の表示を制御する。

【0119】

ここで、仮想動作体 VO は、イヌなどの言語伝達手段を有しない生物に相当する視覚情報であってよい。制御対象が上記のような仮想物体である場合であっても、動作制御部 150 による上述の制御によれば、実際の生物に近いより自然な挙動を実現しながら、認識処理の進捗に係るフィードバックをユーザに提示することが可能である。

【0120】

また、仮想動作体 VO が言語伝達手段を有するキャラクターなどに相当する視覚情報である場合であっても、ユーザの発話開始が検出された際には頷く動作を、マッチングが開始された際には考える動作をさせることなどにより、より現実に近い動作を実現し、没入感を高めることが可能である。

【0121】

<< 3.4. 制御の流れ >>

次に、本実施形態に係る動作制御の流れについて詳細に説明する。図 18 は、本実施形態に係る動作制御の流れを示すフローチャートである。

【0122】

図 18 を参照すると、まず、入力部 110 がセンサ情報を収集する (S1101)。ステップ S1101 におけるセンサ情報の収集は、図 1 で説明した各種のセンサにより実現される。

【0123】

次に、認識部 120 が、ステップ S1102 において収集されたセンサ情報に基づいて状況推定を行う (S1102)。なお、ステップ S1101 におけるセンサ情報の収集と、ステップ S1102 における状況推定は、常時継続的に実行されてよい。

【0124】

次に、認識部 120 によりユーザの発話の開始が検出され (S1103)、動作制御部 150 が第 1 応答の実行を制御する (S1104)。

【0125】

次に、認識部 120 によりユーザの発話の終了が検出され (S1105)、動作制御部 150 が第 2 応答の実行を制御する (S1106)。

【0126】

次に、認識部 120 によりマッチング処理が実行される (S1107)。

【0127】

ここで、マッチング処理に係る確信度が高い場合 (S1108: 高)、動作制御部 150 は、発話に対する理解を示す第 3 応答の実行を制御し (S1109)、また、マッチング結果に基づく動作の実行を制御する (S1110)。

【0128】

一方、マッチング処理に係る確信度が低い場合 (S1108: 低)、認識部 120 は、発話の対象が自律移動体 10 であるか否かを判定してよい (S1111)。

【0129】

ここで、認識部 120 が発話の対象が自律移動体 10 ではないと判定した場合 (S1111: NO)、動作制御部 150 は、応答動作に係る制御を終了する。

【0130】

一方、認識部 120 が発話の対象が自律移動体 10 であると判定した場合 (S1111: YES)、動作制御部 150 は、ユーザに対し再発話を促す第 3 応答の実行を制御し (S1112)、当該再発話にむけて自律移動体 10 が音を発生させないように待機させる (S1113)。

【0131】

< 4. 動作制御装置のハードウェア構成例 >

10

20

30

40

50

次に、本開示の一実施形態に係る動作制御部 150 の機能が自律移動体 10 とは別途の動作制御装置として実現される場合のハードウェア構成例について説明する。図 19 は、本開示の一実施形態に係る動作制御装置 20 のハードウェア構成例を示すブロック図である。図 19 を参照すると、動作制御装置 20 は、例えば、CPU 871 と、ROM 872 と、RAM 873 と、ホストバス 874 と、ブリッジ 875 と、外部バス 876 と、インターフェース 877 と、入力装置 878 と、出力装置 879 と、ストレージ 880 と、ドライブ 881 と、接続ポート 882 と、通信装置 883 と、を有する。なお、ここで示すハードウェア構成は一例であり、構成要素の一部が省略されてもよい。また、ここで示される構成要素以外の構成要素をさらに含んでもよい。

【0132】

(CPU 871)

CPU 871 は、例えば、演算処理装置又は制御装置として機能し、ROM 872、RAM 873、ストレージ 880、又はリムーバブル記録媒体 901 に記録された各種プログラムに基づいて各構成要素の動作全般又はその一部を制御する。

【0133】

(ROM 872、RAM 873)

ROM 872 は、CPU 871 に読み込まれるプログラムや演算に用いるデータ等を格納する手段である。RAM 873 には、例えば、CPU 871 に読み込まれるプログラムや、そのプログラムを実行する際に適宜変化する各種パラメータ等が一時的又は永続的に格納される。

【0134】

(ホストバス 874、ブリッジ 875、外部バス 876、インターフェース 877)

CPU 871、ROM 872、RAM 873 は、例えば、高速なデータ伝送が可能なホストバス 874 を介して相互に接続される。一方、ホストバス 874 は、例えば、ブリッジ 875 を介して比較的データ伝送速度が低速な外部バス 876 に接続される。また、外部バス 876 は、インターフェース 877 を介して種々の構成要素と接続される。

【0135】

(入力装置 878)

入力装置 878 には、例えば、マウス、キーボード、タッチパネル、ボタン、スイッチ、及びレバー等が用いられる。さらに、入力装置 878 としては、赤外線やその他の電波を利用して制御信号を送信することが可能なリモートコントローラ(以下、リモコン)が用いられることもある。また、入力装置 878 には、マイクロフォンなどの音声入力装置が含まれる。

【0136】

(出力装置 879)

出力装置 879 は、例えば、CRT (Cathode Ray Tube)、LCD、又は有機 EL 等のディスプレイ装置、スピーカ、ヘッドホン等のオーディオ出力装置、プリンタ、携帯電話、又はファクシミリ等、取得した情報を利用者に対して視覚的又は聴覚的に通知することが可能な装置である。また、本開示に係る出力装置 879 は、触覚刺激を出力することが可能な種々の振動デバイスを含む。

【0137】

(ストレージ 880)

ストレージ 880 は、各種のデータを格納するための装置である。ストレージ 880 としては、例えば、ハードディスクドライブ (HDD) 等の磁気記憶デバイス、半導体記憶デバイス、光記憶デバイス、又は光磁気記憶デバイス等が用いられる。

【0138】

(ドライブ 881)

ドライブ 881 は、例えば、磁気ディスク、光ディスク、光磁気ディスク、又は半導体メモリ等のリムーバブル記録媒体 901 に記録された情報を読み出し、又はリムーバブル記録媒体 901 に情報を書き込む装置である。

10

20

30

40

50

【0139】

(リムーバブル記録媒体901)

リムーバブル記録媒体901は、例えば、DVDメディア、Blu-ray(登録商標)メディア、HDDVDメディア、各種の半導体記憶メディア等である。もちろん、リムーバブル記録媒体901は、例えば、非接触型ICチップを搭載したICカード、又は電子機器等であってもよい。

【0140】

(接続ポート882)

接続ポート882は、例えば、USB(Universal Serial Bus)ポート、IEEE1394ポート、SCSI(Small Computer System Interface)、RS-232Cポート、又は光オーディオ端子等のような外部接続機器902を接続するためのポートである。

【0141】

(外部接続機器902)

外部接続機器902は、例えば、プリンタ、携帯音楽プレーヤ、デジタルカメラ、デジタルビデオカメラ、又はICレコーダ等である。

【0142】

(通信装置883)

通信装置883は、ネットワークに接続するための通信デバイスであり、例えば、有線又は無線LAN、Bluetooth(登録商標)、又はWUSB(Wireless USB)用の通信カード、光通信用のルータ、ADSL(Asymmetric Digital Subscriber Line)用のルータ、又は各種通信用のモデム等である。

【0143】

<5.まとめ>

以上説明したように、本開示の一実施系形態に係る情報処理方法を実現する自律移動体10は、認識対象情報の入力に基づいて、認識処理の実行に係る暗示的なフィードバックである応答動作を実行すること、を特徴の一つとする。係る構成によれば、認識処理の実行に係るフィードバックをより自然な動作で実現することが可能となる。

【0144】

以上、添付図面を参照しながら本開示の好適な実施形態について詳細に説明したが、本開示の技術的範囲はかかる例に限定されない。本開示の技術分野における通常の知識を有する者であれば、請求の範囲に記載された技術的思想の範疇内において、各種の変更例または修正例に想到し得ることは明らかであり、これらについても、当然に本開示の技術的範囲に属するものと了解される。

【0145】

また、本明細書に記載された効果は、あくまで説明的または例示的なものであって限定的ではない。つまり、本開示に係る技術は、上記の効果とともに、または上記の効果に代えて、本明細書の記載から当業者には明らかな他の効果を奏しうる。

【0146】

また、本明細書における自律移動体10の処理に係る各ステップは、必ずしもフローチャートに記載された順序に沿って時系列に処理される必要はない。例えば、自律移動体10の処理に係る各ステップは、フローチャートに記載された順序と異なる順序で処理されても、並列的に処理されてもよい。

【0147】

なお、以下のような構成も本開示の技術的範囲に属する。

(1)

認識処理に基づいて行動を行う動作体の動作制御を行う動作制御部、
を備え、

前記動作制御部は、認識対象情報の入力に基づいて前記動作体に応答動作を実行させ、

10

20

30

40

50

前記応答動作は、前記認識処理の実行に係る暗示的なフィードバックである、
情報処理装置。

(2)

前記動作制御部は、認識対象情報の入力開始が検知されたことに基づいて前記動作体に第 1 応答を実行させ、認識対象情報の入力完了が検知されたことに基づいて前記動作体に第 2 応答を実行させ、

前記第 1 応答および前記第 2 応答は、前記認識処理の実行に係る暗示的なフィードバックである、

前記 (1) に記載の情報処理装置。

(3)

前記動作制御部は、前記認識処理が完了したことに基づいて、前記動作体に前記認識処理の実行に係るフィードバックである第 3 応答を実行させる、

前記 (2) に記載の情報処理装置。

(4)

前記動作制御部は、前記第 3 応答の実行後、前記動作体に前記認識処理に基づく行動を実行させる、

前記 (3) に記載の情報処理装置。

(5)

前記動作体は、生物を模した形態および能力を有し、

前記第 1 応答、前記第 2 応答、および前記第 3 応答は、身体動作、眼球動作、または音を用いた情動表出動作のいずれかを含む、

前記 (3) または (4) に記載の情報処理装置。

(6)

前記動作制御部は、センサ情報から推定される状況に基づいて、前記第 1 応答、前記第 2 応答、および前記第 3 応答に係る動作種別を動的に決定する、

前記 (3) ~ (5) のいずれかに記載の情報処理装置。

(7)

前記動作制御部は、ユーザ状態に基づいて、前記第 1 応答、前記第 2 応答、および前記第 3 応答に係る動作種別を動的に決定する、

前記 (3) ~ (6) のいずれかに記載の情報処理装置。

(8)

前記動作制御部は、前記認識処理に係る信頼度に基づいて、前記第 3 応答の動作を動的に決定する、

前記 (3) ~ (7) のいずれかに記載の情報処理装置。

(9)

前記認識処理は、音声認識処理である、

前記 (3) ~ (8) のいずれかに記載の情報処理装置。

(10)

前記第 1 応答は、眼球動作である、

前記 (9) に記載の情報処理装置。

(11)

前記動作体は、駆動部を有する装置である、

前記 (9) または (10) に記載の情報処理装置。

(12)

前記動作制御部は、ユーザの発話対象が前記動作体ではないと判定されたことに基づいて、前記動作体に前記第 1 応答、前記第 2 応答、または第 3 応答のうち少なくともいずれかを実行させない、

前記 (9) ~ (11) のいずれかに記載の情報処理装置。

(13)

前記動作制御部は、前記動作体の行動中にユーザの発話開始が検出された場合、前記行

10

20

30

40

50

動を緩やかに停止させる、

前記(9)～(12)のいずれかに記載の情報処理装置。

(14)

前記動作制御部は、前記行動の停止後、前記動作体が音を発生しないよう制御する、
前記(13)に記載の情報処理装置。

(15)

前記動作制御部は、前記音声認識処理に係る信頼度が低い場合、前記動作体にユーザの再発話を促す前記第3応答を実行させる、

前記(9)～(14)のいずれかに記載の情報処理装置。

(16)

前記動作制御部は、前記再発話を促す前記第3応答の完了後、前記動作体が音を発生しないよう制御する、

前記(15)に記載の情報処理装置。

(17)

前記動作体は、言語伝達手段を持たない自律移動体である、

前記(1)～(6)のいずれかに記載の情報処理装置。

(18)

前記動作体である、

前記(1)～(17)のいずれかに記載の情報処理装置。

(19)

プロセッサが、認識処理に基づいて行動を行う動作体の動作制御を行うこと、
を含み、

前記動作制御を行うことは、認識対象情報の入力に基づいて前記動作体に応答動作を実行させること、

をさらに含み、

前記応答動作は、前記認識処理の実行に係る暗示的なフィードバックである、
情報処理方法。

(20)

コンピュータを、

認識処理に基づいて行動を行う動作体の動作制御を行う動作制御部、

を備え、

前記動作制御部は、認識対象情報の入力に基づいて前記動作体に応答動作を実行させ、

前記応答動作は、前記認識処理の実行に係る暗示的なフィードバックである、

情報処理装置、

として機能させるためのプログラム。

【符号の説明】

【0148】

10 自律移動体

110 入力部

120 認識部

130 学習部

140 行動計画部

150 動作制御部

160 駆動部

170 出力部

510 ディスプレイ

570 アクチュエータ

10

20

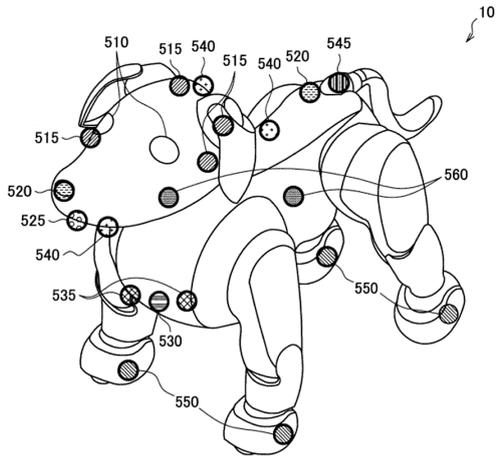
30

40

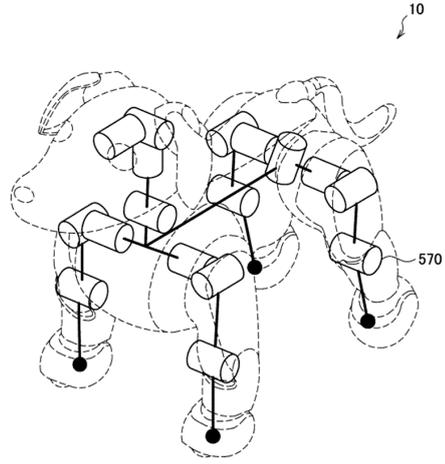
50

【 図面 】

【 図 1 】



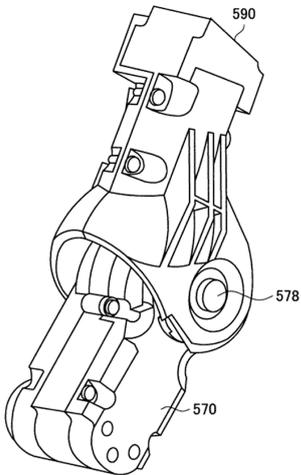
【 図 2 】



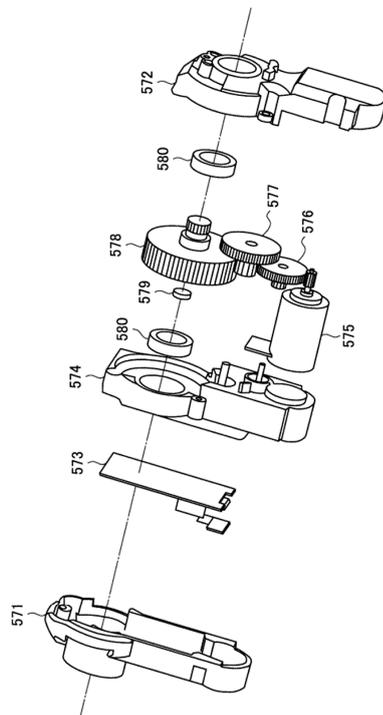
10

20

【 図 3 】



【 図 4 】

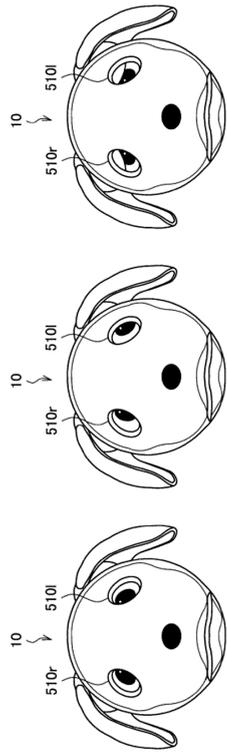


30

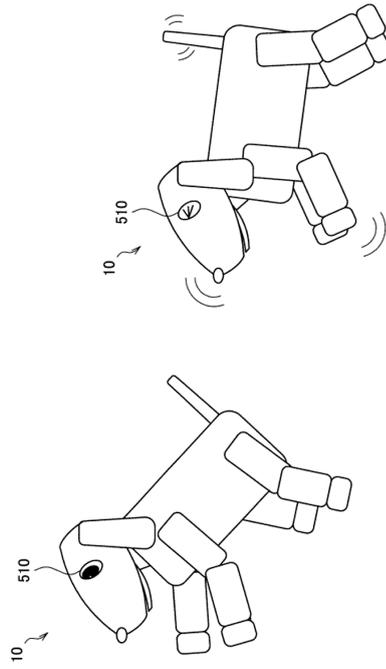
40

50

【 図 5 】



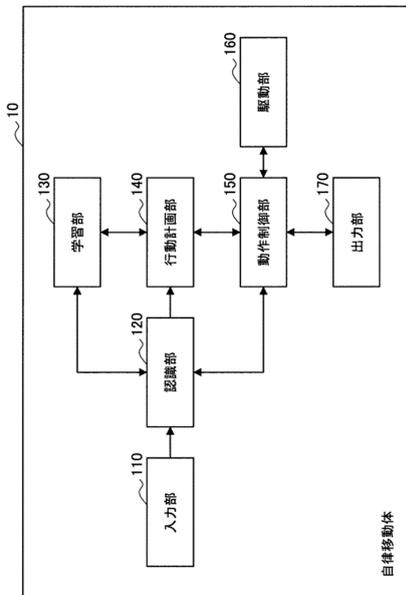
【 図 6 】



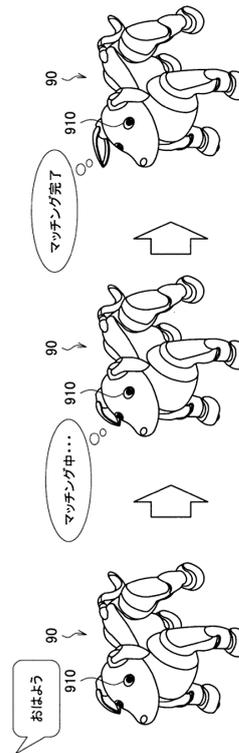
10

20

【 図 7 】



【 図 8 】

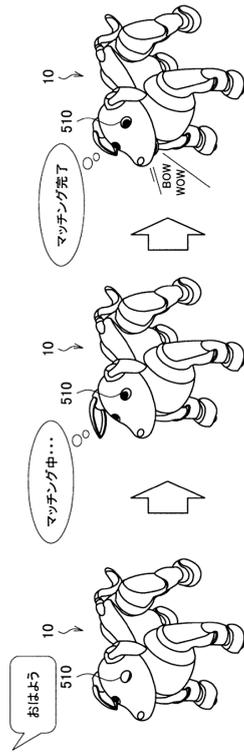


30

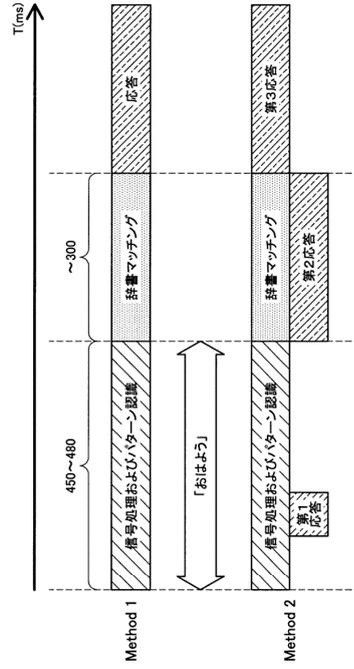
40

50

【 図 9 】



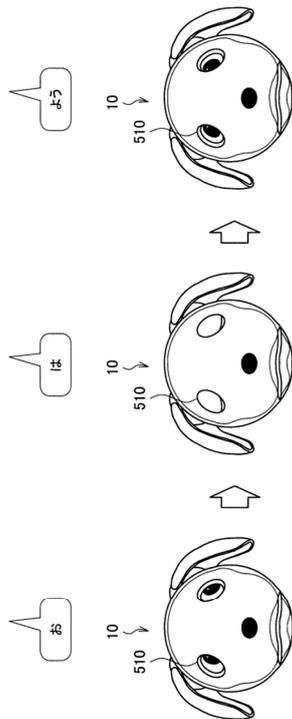
【 図 10 】



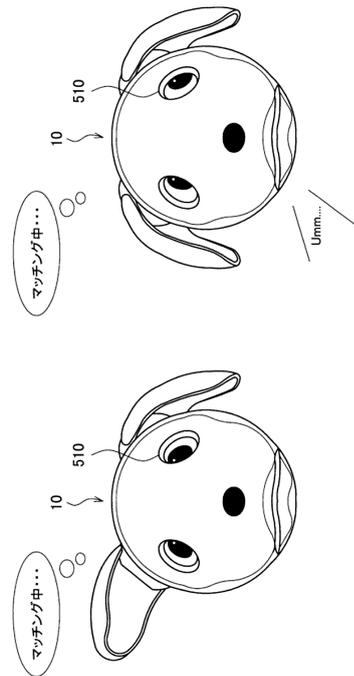
10

20

【 図 11 】



【 図 12 】

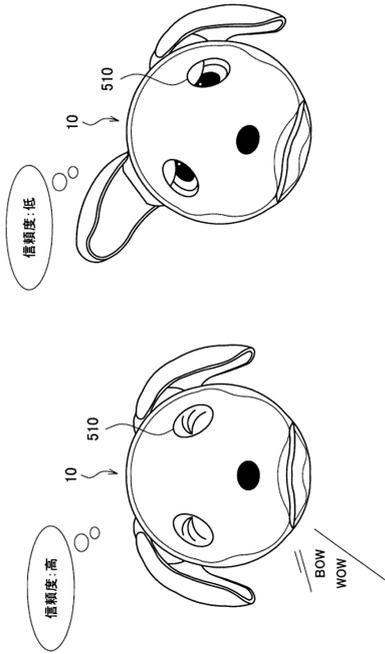


30

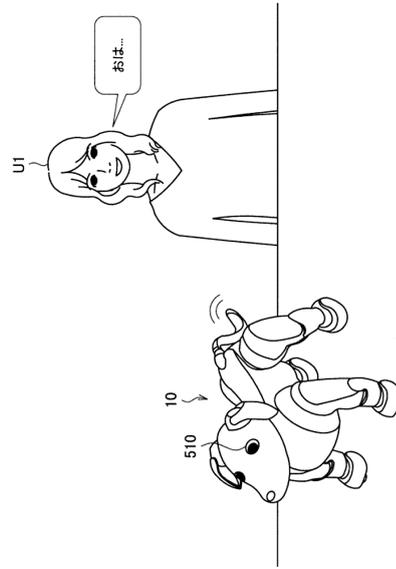
40

50

【 図 1 3 】



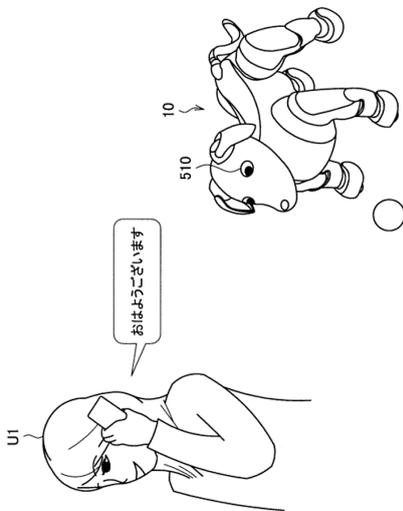
【 図 1 4 】



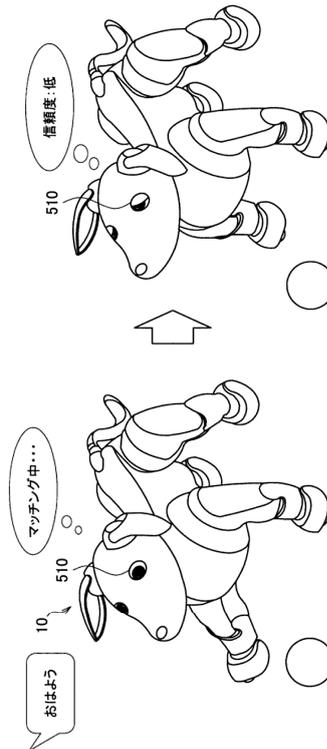
10

20

【 図 1 5 】



【 図 1 6 】

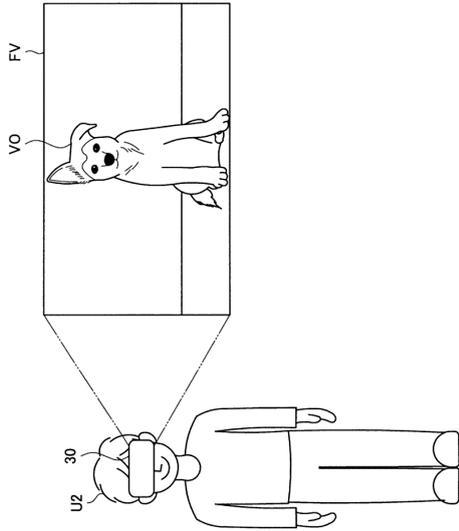


30

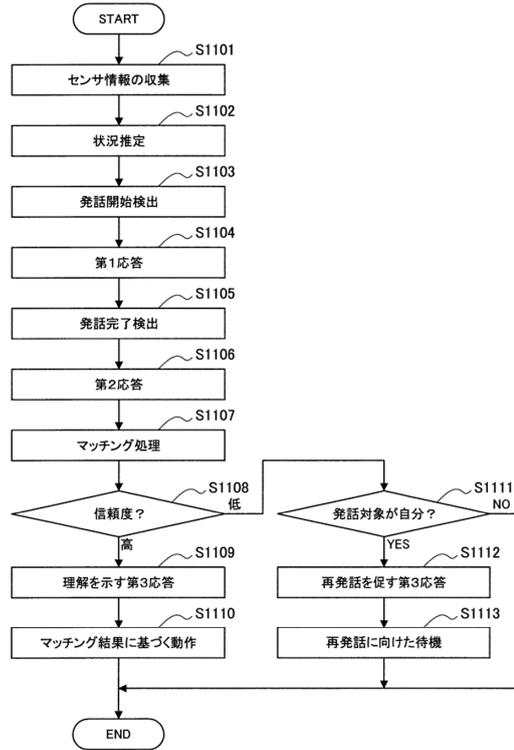
40

50

【 図 1 7 】



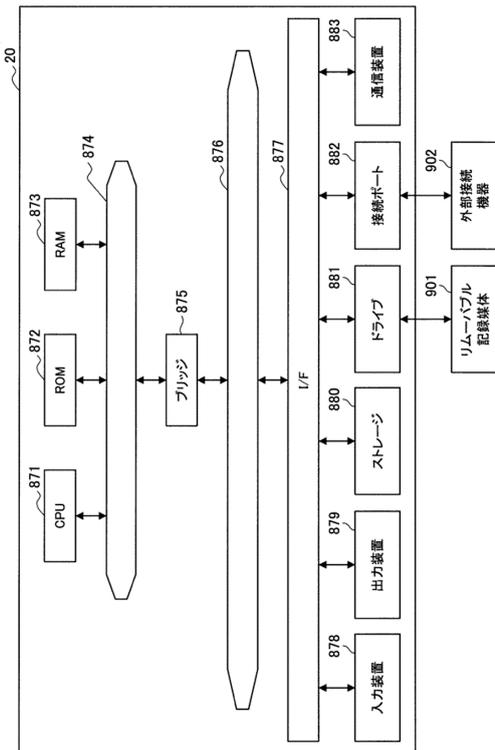
【 図 1 8 】



10

20

【 図 1 9 】



30

40

50