

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5751610号
(P5751610)

(45) 発行日 平成27年7月22日 (2015. 7. 22)

(24) 登録日 平成27年5月29日 (2015. 5. 29)

(51) Int. Cl.			F I		
B 2 5 J	13/08	(2006. 01)	B 2 5 J	13/08	Z
B 2 5 J	5/00	(2006. 01)	B 2 5 J	5/00	A
A 6 3 H	11/00	(2006. 01)	A 6 3 H	11/00	Z

請求項の数 7 (全 15 頁)

(21) 出願番号	特願2010-221556 (P2010-221556)	(73) 特許権者	899000068
(22) 出願日	平成22年9月30日 (2010. 9. 30)		学校法人早稲田大学
(65) 公開番号	特開2012-76162 (P2012-76162A)		東京都新宿区戸塚町 1 丁目 1 0 4 番地
(43) 公開日	平成24年4月19日 (2012. 4. 19)	(74) 代理人	100080089
審査請求日	平成25年9月27日 (2013. 9. 27)		弁理士 牛木 護
		(74) 代理人	100137800
			弁理士 吉田 正義
		(74) 代理人	100148253
			弁理士 今枝 弘充
		(74) 代理人	100148079
			弁理士 梅村 裕明
		(74) 代理人	100125081
			弁理士 小合 宗一
		(74) 代理人	100161665
			弁理士 高橋 知之

最終頁に続く

(54) 【発明の名称】 会話ロボット

(57) 【特許請求の範囲】

【請求項 1】

胴体部に回動可能に設けられた上体部と、前記上体部に回動可能に設けられた頭部とを備え、外部センサからの出力結果に基づいて複数の対象物と自律的にグループ会話を行う会話ロボットにおいて、

前記外部センサから取得した出力結果に基づいて前複数の対象物の位置を検出する位置検出手段と、

前記外部センサから取得した出力結果に基づいて、前記複数の対象物のうち1つを主注目対象物として認識し、他の残りの前記対象物を従注目対象物として認識する役割識別手段と、

前記主注目対象物及び前記従注目対象物の位置から求めた重心方向線で示された重心方向に、前記上体部の正面を向けるように前記上体部の回動角度を制御する胴体部制御手段と、

前記主注目対象物の方向に前記頭部の正面を向けるように前記頭部の回動角度を制御する頭部制御手段とを備え、

前記胴体部制御手段と前記頭部制御手段は、前記役割識別手段からのデータに基づいて、前記上体部の正面を前記重心方向線で示された重心方向に向けた状態のまま前記頭部だけを回動させて前記注目対象物の方向へ前記頭部の正面を向けるように制御でき、かつ、前記頭部の正面を前記主注目対象物の方向に向けた状態のまま前記上体部だけを回動させて前記重心方向線で示された重心方向へ前記上体部の正面を向けるように制御できる構成と

されている

ことを特徴とする会話ロボット。

【請求項 2】

前記複数の対象物がユーザであり、

前記役割識別手段は、

前記外部センサから取得した画像及び又は音声に基づいて、複数の前記ユーザのうち、発話するユーザあるいは主聴者であるユーザを前記主注目対象物として認識し、他の残りの前記ユーザを前記従注目対象物として認識する

ことを特徴とする請求項 1 記載の会話ロボット。

【請求項 3】

前記役割識別手段は、

前記主注目対象物として認識した前記ユーザが注目する次注目対象推定物を認識し、該ユーザが発話終了後に前記次注目対象推定物を新たな主注目対象物とする

ことを特徴とする請求項 2 記載の会話ロボット。

【請求項 4】

前記重心方向線と所定の基準線のなす重心方向角度は、

前記主注目対象物の方向に延びる主注目対象方向線と前記基準線のなす主注目対象角度と、前記従注目対象物の方向に延びる従注目対象方向線と前記基準線のなす従注目対象角度を算出し、

前記主注目対象角度と前記従注目対象角度とを全て合算した値を前記複数の対象物の総数で除算することによって求める

ことを特徴とする請求項 1 ~ 3 のうちのいずれか 1 項記載の会話ロボット。

【請求項 5】

前記胴体部制御手段は、

前記複数の対象物から求めた前記重心方向へ前記上体部の正面を向けると、前記上体部の回動角度が所定回動角度以上になる場合、前記上体部の回動角度が前記所定回動角度範囲内となるように、前記複数の対象物の中から所定の対象物を除外して前記重心方向を求める

ことを特徴とする請求項 1 ~ 4 のうちいずれか 1 項記載の会話ロボット。

【請求項 6】

指示部の可動を制御する指示部制御手段を備えたことを特徴とする請求項 1 ~ 5 のうちいずれか 1 項記載の会話ロボット。

【請求項 7】

前記重心方向に前記上体部の正面を向けると、前記上体部の回動角度が所定回動角度以上になる場合、前記上体部の回動を前記回動角度の範囲内になるように、前記胴体部を移動手段によって移動させる移動制御手段を備える

ことを特徴とする請求項 1 ~ 4、6 のうちいずれか 1 項記載の会話ロボット。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、会話ロボットに関し、例えば話題を共有しながら複数人で行われる会話（以下、これをグループ会話と呼ぶ）を行う際に適用して好適なものである。

【背景技術】

【0002】

近年、電氣的若しくは磁氣的な作用を用いて人間や動物の動作に似せた運動を行うロボットが数多く商品化されている。このようなロボットに対し、例えば人間同士が日常的に行う会話と同様の会話をユーザとの間で行い得るような音声会話機能を搭載したロボットも知られている（例えば、非特許文献 1 参照）。

【0003】

實際上、これら様々な会話ロボットのなかには、C C D（ChargeCoupled Device）カメ

10

20

30

40

50

ラやマイクロホン等の各種外部センサを搭載し、これら外部センサの出力に基づいて外部状況を認識して、認識結果に基づいて自律的に行動し得るようになされたものなどもある。例えば、その一例としては、外部センサを基に、発話するユーザの方向を認識し、例えば胴体部を移動させて胴体部正面及び頭部正面を当該ユーザの方向に向けて、あたかもユーザの発話に反応してユーザに視線を向けて会話を行うような会話ロボットも知られている。

【先行技術文献】

【非特許文献】

【0004】

【非特許文献1】伊吹征太，木村憲次，武田夏佳：コミュニケーションロボットを用いた高齢者生活支援システム，日本機械学会誌，Vol.108，No.1038，pp.392-395(2005)

10

【発明の概要】

【発明が解決しようとする課題】

【0005】

しかしながら、このような会話ロボットでは、実際にグループ会話が行われている際に、発話しているユーザの方向に胴体部正面及び頭部正面を移動させ、発話しているユーザにだけ視線を向けるような動作が行われると、発話していない他のユーザもグループ会話の参加者であるにもかかわらず、あたかもグループ会話から外されたかのような不自然さを、発話しているユーザや他のユーザに感じさせてしまうという問題があった。

【0006】

20

本発明は以上の点を考慮してなされたもので、状況に応じた自然な会話を行い得る会話ロボットを提案することを目的とする。

【課題を解決するための手段】

【0007】

かかる課題を解決するため本発明の請求項1は、胴体部に回動可能に設けられた上体部と、前記上体部に回動可能に設けられた頭部とを備え、外部センサからの出力結果に基づいて複数の対象物と自律的にグループ会話を行う会話ロボットにおいて、前記外部センサから取得した出力結果に基づいて前複数の対象物の位置を検出する位置検出手段と、前記外部センサから取得した出力結果に基づいて、前記複数の対象物のうち1つを主注目対象物として認識し、他の残りの前記対象物を従注目対象物として認識する役割識別手段と、前記主注目対象物及び前記従注目対象物の位置から求めた重心方向線で示された重心方向に、前記上体部の正面を向けるように前記上体部の回動角度を制御する胴体部制御手段と、前記主注目対象物の方向に前記頭部の正面を向けるように前記頭部の回動角度を制御する頭部制御手段とを備え、前記胴体部制御手段と前記頭部制御手段は、前記役割識別手段からのデータに基づいて、前記上体部の正面を前記重心方向線で示された重心方向に向けた状態のまま前記頭部だけを回動させて前記注目対象物の方向へ前記頭部の正面を向けるように制御でき、かつ、前記頭部の正面を前記主注目対象物の方向に向けた状態のまま前記上体部だけを回動させて前記重心方向線で示された重心方向へ前記上体部の正面を向けるように制御できる構成とされていることを特徴とする。

30

【0008】

40

また、本発明の請求項2は、前記複数の対象物がユーザであり、前記役割識別手段は、前記外部センサから取得した画像及び又は音声に基づいて、複数の前記ユーザのうち、発話するユーザあるいは主聴者であるユーザを前記主注目対象物として認識し、他の残りの前記ユーザを前記従注目対象物として認識することを特徴とする。

【0009】

また、本発明の請求項3は、前記役割識別手段は、前記主注目対象物として認識した前記ユーザが注目する次注目対象推定物を認識し、該ユーザが発話終了後に前記次注目対象推定物を新たな主注目対象物とすることを特徴とする。

【0010】

また、本発明の請求項4は、前記重心方向線と所定の基準線のなす重心方向角度は、前

50

記注目対象物の方向に延びる注目対象方向線と前記基準線のなす注目対象角度と、前記従注目対象物の方向に延びる従注目対象方向線と前記基準線のなす従注目対象角度を算出し、前記注目対象角度と前記従注目対象角度とを全て合算した値を前記複数の対象物の総数で除算することによって求めることを特徴とする。

【0011】

また、本発明の請求項5は、前記胴体部制御手段は、前記複数の対象物から求めた前記重心方向へ前記上部部の正面を向けると、前記上部部の回動角度が所定回動角度以上になる場合、前記上部部の回動角度が前記所定回動角度範囲内となるように、前記複数の対象物の中から所定の対象物を除外して前記重心方向を求めることを特徴とする。

【0012】

また、本発明の請求項6は、指示部の可動を制御する指示部制御手段を備えたことを特徴とする。

【0013】

また、本発明の請求項7は、前記重心方向に前記上部部の正面を向けると、前記上部部の回動角度が所定回動角度以上になる場合、前記上部部の回動を前記回動角度の範囲内となるように、前記胴体部を移動手段によって移動させる移動制御手段を備えることを特徴とする。

【発明の効果】

【0014】

本発明によれば、注目対象物だけでなく従注目対象物に対しても、あたかも注目しているかのような印象を与えることができ、かくして状況に応じた自然な会話を行い得る。

【図面の簡単な説明】

【0015】

【図1】本発明の会話ロボットの外觀構成を示す概略図である。

【図2】ユーザBが新たに加わったときの会話ロボットの動作の様子を示す概略図である。

【図3】会話ロボットに搭載された会話装置の回路構成を示すブロック図である。

【図4】重心方向の求め方の説明に供する概略図である。

【図5】ユーザBがグループ会話から離脱したときの会話ロボットの動作の様子を示す概略図である。

【発明を実施するための形態】

【0016】

以下図面に基づいて本発明の実施の形態を詳述する。

【0017】

(1) 本願発明の概要

図1において、1は会話ロボットを示し、この会話ロボット1は、人に模して外觀が形成されており、胴体部2に対して回動自在に連結された頭部3と、当該胴体部2の左右に可動自在に連結された腕部4a,4bとを備えている。實際上、この胴体部2は、移動可能な車輪5を備えた基台6を有し、基台6に対して上部部8がz方向を回動軸として回動方向Caに回動自在に設けられた構成を有する。

【0018】

上部部8には、上部左右に肩関節部9a,9bを介して腕部4a,4bが設けられているとともに、上部に設けた首部8aに上下左右に可動する頭部3が連結されている。腕部4aには、肩関節部9aによって上部部8に対し上腕部11がx方向、y方向及びz方向に可動するように設けられているとともに、上腕部11に肘関節部12を介して前腕部13が回動自在に設けられ、さらにこの前腕部13に手首関節部14を介して手部15が回動自在に設けられている。これにより、腕部4a,4bは、これら肩関節部9a、肘関節部12及び手首関節部14を各アクチュエータ(図示せず)により駆動させ、上腕部11、前腕部13及び手部15を上部部8の正面(以下、単に胴体部正面と呼ぶ)8b側に向けて突き出す等、人の腕の動作に模した動作を行なえ得る。

10

20

30

40

50

【0019】

また、頭部3は、胴体部正面8bと同じ側を頭部3の正面（以下、単に頭部正面と呼ぶ）3aとして、人の目を模した目部21と、人の口を模した口部22とが当該頭部正面3aに形成されている。また、頭部3には、頭部正面3aの目部21に「目」として機能する一対のCCD（Charge Coupled Device）カメラ23が設けられているとともに、口部22の内部に発声装置として機能するスピーカ（図示せず）が配設されている。

【0020】

さらに、この会話ロボット1は、ユーザの音声をマイクロホン（図1では図示せず）により集音し、当該ユーザの発話した内容に応じて、当該ユーザに対し最適な返答や質問等をスピーカから発話したり、頭部3、上体部8、腕部4a,4b及び車輪5を動かして、ユーザとの間で状況に応じた自律的な行動を実行し得るようになされている。また、かかる構成に加えて、本発明による会話ロボット1は、会話するユーザの人数に応じて、頭部3及び上体部8を回動方向Caに回動させて、頭部正面3a及び胴体部正面8bの向く方向を最適な方向に適宜変更させ、会話するユーザが増えても、自然なグループ会話が行なえるようになされている。以下、複数のユーザとの間でグループ会話を行なう際の会話ロボット1の動作について説明する。

10

【0021】

ここで、図2（A）は、会話ロボット1とユーザAとの間で行われる1対1の会話の状況を、上方から見た様子を示し、これに対して、図2（B）は、新たにユーザBが加わり、会話ロボット1とユーザAとユーザBとの間で行われる複数人によるグループ会話の状況を、上方から見た様子を示している。なお、図2（A）及び（B）中「」マークは、ユーザAやユーザBの顔正面A1,B2や、会話ロボット1の頭部正面3aを示し、実線の矢印は、ユーザA及びユーザBの顔正面A1,B2、会話ロボット1の頭部正面3aが向けられた方向を示している。また、点線の矢印は、ユーザA及びユーザBの体正面A2,B2、会話ロボット1の胴体部正面8bが向けられた方向を示している。

20

【0022】

因みに、この実施の形態の場合では、ユーザA及びユーザBの音声を集音する手法の一例として、ユーザA及びユーザBにそれぞれ専用のマイクロホンを用意し、ユーザA及びユーザBの所定部位にそれぞれ付けたマイクロホンが、会話ロボット1に配線を介して接続されている。これにより、会話ロボット1は、マイクロホンから集音したユーザA及びユーザBの各音声を音声信号として取得し、当該音声信号を解析することによりユーザA及びユーザBの発話状況や発話内容を認識し得るようになされている。

30

【0023】

ここで、図2（A）に示すように、この会話ロボット1は、ユーザAとの間で会話を行うとき、頭部3を回動させて頭部正面3aをユーザAに向けるとともに、上体部8も回動させて胴体部正面8bをユーザAに向け、ユーザAを注視しているかのような動作を行なう。このように、会話ロボット1は、頭部正面3a及び胴体部正面8bをともにユーザAに向けて、ユーザAとの会話を行なうことを明確に意思表示しているかのような印象を、ユーザAに対し与えることができる。

【0024】

かかる構成に加えて、この会話ロボット1は、ユーザAとの間の会話中に、例えばユーザBが近づくと、CCDカメラ23により撮像した動画像に基づいて当該ユーザBの顔正面B1や体正面B2の向きを検出し、この検出結果から、ユーザBの顔正面B1の一部及び体正面B2の一部（後述する）が、会話ロボット1の方向に向いていると認識すると、ユーザBが新たに会話に加わりグループ会話が行われると判断し得るようになされている。

40

【0025】

また、会話ロボット1は、図2（B）に示すように、ユーザA及びユーザBを含めたグループ会話であると認識すると、CCDカメラ23で撮像した動画像を基に、グループ会話の参加者（図2（B）ではユーザA及びユーザB）全員の位置を特定しその位置関係から重心方向線CG1（後述する）を算出し、現在会話中のユーザAの方向に頭部正面3aを向け

50

た状態のまま、上部部 8 だけを回動させて、重心方向線CG1で示された重心方向へ胴体部正面8bを向けるように構成されている。これにより、会話ロボット 1 は、頭部正面3aをユーザ A に向けることで、ユーザ A との会話を継続しているかのような印象をユーザ A 及びユーザ B に対して与えることができる。また、これに加えて会話ロボット 1 は、ユーザ A 及びユーザ B の位置から求めた重心方向に胴体部正面8bを向けることで、ユーザ A だけでなく、あたかもユーザ B にも注視しているかのような印象をユーザ A 及びユーザ B に与え、ユーザ A 及びユーザ B との間で自然な会話を実現し得る。

【 0 0 2 6 】

(2) 会話ロボットの回路構成

次に、図 2 (A) 及び (B) に示すような会話ロボット 1 の行動を、図 3 に示す回路構成を用いて以下説明する。この実施の形態の場合、会話ロボット 1 には、図 3 に示すような会話装置30が内蔵されており、例えば頭部正面3aに設けられたCCDカメラ23は、頭部正面3a方向を撮像して得られた動画像を動画像データとして、顔向き・体向き検出部32と、顔認識部33とにそれぞれ送付する。顔認識部33には、動画像データから生成される動画像の中から、統計的手法によって予め定められた肌色尤度を基に、ほぼ楕円状の肌色領域を特定してこれを顔領域として抽出する。顔認識部33は、この肌色領域を正規化した後、この正規化した画像から両目の距離や鼻の幅等の特徴(顔特徴量)を算出する。ここで、顔認識部33には、ユーザ A やユーザ B の両目の距離や鼻の幅等の顔特徴量が顔データとして予め記憶されている。これにより顔認識部33は、動画像データを基に検出した顔特徴量と、登録されている顔データの特徴量とを比べることで、動画像中にユーザ A やユーザ B が存在していることを認識し得、これを顔識別結果データとして位置検出部35に送付する。

【 0 0 2 7 】

また、この実施の形態の場合、顔向き・体向き検出部32でも、CCDカメラ23から動画像データを受け取ると、動画像データから生成される動画像の中から、統計的手法によって予め定められた肌色尤度を基に、ほぼ楕円状の肌色領域を特定してこれを顔領域として抽出した後、この肌色領域を正規化し、この正規化した画像から両目の距離や鼻の幅等の特徴(顔特徴量)を算出する。また、顔向き・体向き検出部32は、動画像データから生成される動画像の中から、統計的手法によって予め定められた人の上半身の輪郭データを基に、ユーザ A 及びユーザ B の各上半身を特定してこれを上半身領域として抽出した後、この上半身領域を正規化し、この正規化した画像から肩幅等の特徴(上半身輪郭特徴量)を算出する。ここで、顔向き・体向き検出部32は、例えば顔向きテンプレート情報(後述する)と、体向きテンプレート情報(後述する)とを予め記憶しており、これら顔向きテンプレート情報と顔特徴量とを照らし合わせることで、両目距離の変化等から顔正面A1,B1の向きを識別するとともに、体向きテンプレート情報と上半身輪郭特徴量とを照らし合わせることで、肩幅の変化等から体正面A2,B2の向きを識別し得るようになされている。

【 0 0 2 8 】

例えば、顔向きテンプレート情報は、CCDカメラ23に対して人の顔が正面のとき、斜め前左右30度のとき、斜め前左右60度のとき、斜め前左右90度のときのそれぞれ統計的な顔特徴量のモデル(以下、統計的顔向き特徴量と呼ぶ)を示したものであり、これら統計的顔向き特徴量と、ユーザ A 及びユーザ B の顔特徴量とを比べることで、顔正面A1,B1の向きを識別し、これを顔向き識別データとして得るようになされている。また、体向きテンプレート情報は、例えばCCDカメラに対して人の上半身が正面のとき、斜め前左右30度のとき、斜め前左右60度のとき、斜め前左右90度のときのそれぞれ統計的な上半身輪郭特徴量のモデル(以下、統計的体向き特徴量と呼ぶ)を示したものであり、これら統計的体向き特徴量と、ユーザ A 及びユーザ B の上半身輪郭特徴量とを比べることで、体正面A2,B2の向きを識別し、これを体向き識別データとして得るようになされている。そして、顔向き・体向き検出部32は、これら動画像データを基に検出したユーザ A 及びユーザ B の顔向き識別データ及び体向き識別データをそれぞれ役割識別部36に送付する。

【 0 0 2 9 】

なお、ここで、ユーザ A 及びユーザ B の顔正面A1,B1及び体正面A2,B2の向きを識別する手法としては、例えば「顔と身体の外観及び形状の変動傾向を考慮した上体輪郭抽出・追跡手法」(俵直弘 藤江真也 小林哲則(「画像の認・理解シンポジウム(MIRU2010)」2010年7月))に記載された技術内容を適用するようによく、ユーザ A 及びユーザ B の顔正面A1,B1及び体正面A2,B2の向きを識別できれば、その他種々の手法を適用してもよい。

【0030】

このときマイクロホン37a,37bは、ユーザ A 及びユーザ B の各音声をそれぞれ集音すると、これらを音声信号としてそれぞれ音声処理部39に送出する。音声処理部39には、ユーザ A 及びユーザ B の音声の特徴(音声特徴量)が音声識別データとして予め記憶されており、マイクロホン37a,37bから受け取った各音声信号からそれぞれ特徴量を抽出し、この特徴量と音声識別データとを比べることで、どのマイクロホン37a,37bがユーザ A 又はユーザ B に用いられているか否かを認識し得るようになされている。

10

【0031】

そして、音声処理部39は、例えば一方のマイクロホン37aがユーザ A の音声を集音し、他方のマイクロホン37bがユーザ B の音声を集音していることを示す音声識別結果データを生成し、これを位置検出部35に送出する。また、この音声処理部39は、バイグラム言語モデル、HMM(Hidden Markov Model;隠れマルコフモデル)を用いた語彙量約七百のフレーム同期の連続音声認識を行ない、音声信号を単語の列へと変換し、これを単語列データとして行動選択部40に送出する。

20

【0032】

一方、位置検出部35は、顔認識部33から受け取った顔識別結果データと、音声処理部39から受け取った音声識別結果データとを対応付けることにより、動画像中のユーザ A 及びユーザ B がどの位置に存在し、かつユーザ A 及びユーザ B のいずれが発話しているかを認識し得るようになされている。

【0033】

實際上、位置検出部35は、動画像中において認識したユーザ A と、一方のマイクロホン37aで得られた音声信号とを対応付けるとともに、動画像中において認識したユーザ B と、他方のマイクロホン37bで得られた音声信号とを対応付け、動画像中のユーザ A 及びユーザ B のいずれかが発話しているかを認識し、これを位置検出結果データとして役割識別部36に送信する。

30

【0034】

役割識別部36は、顔向き・体向き検出部32から受け取った顔向き識別データ及び体向き識別データから、ユーザ A 及びユーザ B が会話に参加しているか否かを判断し得るようになされている。實際上、役割識別部36は、例えばユーザ A の動画像から得られたユーザ A の顔向き識別データ及び体向き識別データから、ユーザ A の顔正面A1の一部及び体正面A2の一部がともに、会話ロボット1側に向いているか否かを判断する。その結果、ユーザ A の顔向き識別データ及び体向き識別データから、ユーザ A の顔正面A1の一部及び体正面A2の一部が、会話ロボット1側に向いている場合(例えば、会話ロボット1に対してユーザ A の顔が正面のとき、斜め前左右30度のとき、斜め前左右60度のとき)、このことはユーザ A が会話ロボット1や他のユーザ B とグループ会話を行うために、会話ロボット1や他のユーザ B 側に顔正面A1及び体正面A2を向けていると判断し、役割識別部36は、ユーザ A がグループ会話の参加者であると認識する。

40

【0035】

また、役割識別部36は、ユーザ B についても同様に、ユーザ B の動画像から得られたユーザ B の顔向き識別データ及び体向き識別データから、ユーザ B の顔正面B1の一部及び体正面B2の一部が、会話ロボット1側に向いているか否かを判断する。その結果、ユーザ B の顔向き識別データ及び体向き識別データから、ユーザ B の顔正面B1の一部及び体正面B2の一部が、会話ロボット1側に向いている場合、このことはユーザ B が会話ロボット1や他のユーザ A とグループ会話を行うために、会話ロボット1や他のユーザ A 側に顔正面B1

50

及び体正面B2を向けていると判断し、役割識別部36は、ユーザBがグループ会話の参加者であると認識する。

【0036】

次に、役割識別部36は、位置検出部35から受け取った位置検出結果データに基づいて、グループ会話の参加者として判断したユーザA及びユーザBのうち、いずれかが発話者であるか否かを判断し得る。例えば、役割識別部36は、位置検出部35から受け取った位置検出結果データに基づいて、グループ会話の参加者であると判断したユーザAのマイクロホン37aから音声信号を取得すると、当該ユーザAを発話者（主注目対象物）とし、他方のユーザBを聴者（従注目対象物）とし、これらユーザA及びユーザBの各役割（この場合、ユーザAを発話者（主注目対象物）とし、他方のユーザBを聴者（従注目対象物）とする）を、位置検出結果データに対応付けた主従注目位置検出データを生成し、これを頭部駆動制御部41及び胴体部駆動制御部42にそれぞれ送出する。

10

【0037】

なお、このとき、役割識別部36は、例えば現在発話している発話者たるユーザAの顔向き識別データに基づいて、ユーザAの顔正面A1の向きがユーザB方向であると判断すると、当該ユーザAがユーザBに向けて現在発話しており、ユーザAの発話終了後にユーザBが何らかの返答をするため発話する可能性が高いと推定し、このユーザBを次発話推定者（ユーザAの発話終了後にこの次発話推定者（次注目対象推定物）を主注目対象物とする）として、これを主従注目位置検出データに対応付けて頭部駆動制御部41に送出する。

20

【0038】

さらに、役割識別部36は、後述する行動選択部40にて選択された発話内容を当該行動選択部40から受け取り、会話ロボット1自身が現在発話していると認識したとき、顔向き・体向き検出部32からの顔向き識別データ及び体向き識別データから、会話ロボット1に顔正面A1,B1及び体正面A2,B2を向けているユーザA又はユーザBが存在しているか否かを判断する。その結果、役割識別部36は、顔正面A1,B1及び体正面A2,B2を会話ロボット1に向いているユーザA又はユーザBを主聴者（主注目対象物）とし、これを主従注目位置検出データに対応付けて頭部駆動制御部41及び胴体部駆動制御部42にそれぞれ送出する。因みに、この実施の形態の場合、ユーザA及びユーザBともに顔正面A1,B1及び体正面A2,B2が会話ロボット1に向いているとき、ユーザA及びユーザBのいずれか一方をランダムに主聴者として選択し、これを主従注目位置検出データに対応付けて頭部駆動制御部41に送出する。

30

【0039】

頭部駆動制御部41は、役割識別部36から主従注目位置検出データを受け取ると、発話者や主聴者等の主注目対象物としたユーザA又はユーザBの方向（以下、これを主注目対象方向と呼ぶ）を特定した後、頭部正面3aと、主注目対象方向との角度差を算出する。實際上、この頭部駆動制御部41は、図4(A)に示すように、CCDカメラ23の撮像画枠51の中央線（以下、画枠中央線と呼ぶ）dが頭部正面3aとして予め設定されており、当該画枠中央線下端を角度中心点Oとして、この角度中心点Oから、主注目対象物とした例えばユーザAの胴体中心方向に延びる主注目対象方向線d1を算出する。また、頭部駆動制御部41は、角度中心点Oを基準に画枠中央線dから主注目対象方向線d1までの角度（以下、これを主注目対象角度と呼ぶ） θ_1 を算出し、この主注目対象角度 θ_1 を頭部回動命令として頭部アクチュエータ43に送出する。

40

【0040】

なお、図4(A)では、既に画枠中央線dと主注目対象方向線d1とが一致していることから、主注目対象角度 θ_1 は0度となる。仮に画枠中央線dと主注目対象方向線d1とがずれているときには、頭部アクチュエータ43が頭部回動命令に基づいて主注目対象角度 θ_1 だけ頭部3を回動させることにより、画枠中央線dと主注目対象方向線d1とを一致させ、主注目対象方向に頭部正面3aを向けさせ得る（図2(A)及び(B)）。

【0041】

因みに、胴体部駆動制御部42は、会話ロボット1がユーザAとだけ会話を行なっている

50

とき、上体部 8 を主注目対象角度 θ_1 まで回転させるのに必要な上体部回転角度を算出し、この上体部回転角度を胴体部回転命令として胴体部アクチュエータ 44 に送出する。胴体部アクチュエータ 44 は、胴体部回転命令に基づいて上体部回転角度だけ上体部 8 を回転させることにより、頭部正面 3a が向いている主注目対象方向に胴体部正面 8b も向けさせ得る（図 2 (A)）。

【 0 0 4 2 】

これに対して、胴体部駆動制御部 42 は、ユーザ A だけでなくユーザ B と会話を行なっているとき、役割識別部 36 から主注目位置検出データを受け取ると、図 4 (A) に示すように、角度中心点 O から、主注目対象物としたユーザ A の胴体中心方向に延びる主注目対象方向線 d1 を算出し、角度中心点 O を基準に画枠中央線 d から主注目対象方向線 d1 までの主注目対象角度 θ_1 を算出する。

10

【 0 0 4 3 】

また、このとき胴体部駆動制御部 42 は、角度中心点 O から、従注目対象物としたユーザ B の胴体中心方向に延びる従注目対象方向線 d2 を算出し、角度中心点 O を基準に画枠中央線 d から従注目対象方向線 d2 までの角度（以下、これを従注目対象角度 θ_2 と呼ぶ） θ_2 を算出する。次いで、胴体部駆動制御部 42 は、これら主注目対象角度 θ_1 と従注目対象角度 θ_2 とを全て合算して、認識したユーザ（ユーザ A 及びユーザ B）の総数である「2」で除算し、角度中心点 O を基準に画枠中央線 d から重心方向線 CG1 までの角度（以下、これを重心方向角度と呼ぶ） θ_{CG1} を算出して、これを胴体部回転命令として胴体部アクチュエータ 44 に送出する。

20

【 0 0 4 4 】

これにより、胴体部アクチュエータ 44 は、胴体部回転命令に基づいて重心方向角度 θ_{CG1} だけ上体部 8 を回転させることにより、胴体部正面 8b を重心方向線 CG1 側に向けさせ得る（図 2 (A) 及び (B)）。かくして、会話ロボット 1 は、上体部 8 だけを回転させて、胴体部正面 8b を重心方向へ向けることで、あたかも会話ロボット 1、ユーザ A 及びユーザ B の全員でグループ会話を行っているかのような意思表示を、ユーザ A 及びユーザ B に行ない得る。

【 0 0 4 5 】

因みに、上述した実施の形態においては、グループ会話として、ユーザ A 及びユーザ B の 2 人をグループ会話の参加者としたときの会話ロボット 1 の動作について述べたが、本発明はこれに限らず、3 人や 4 人等その他複数人をグループ会話の参加者としたときでも、この会話ロボット 1 は同様の動作を実行し得る。例えば、図 4 (B) に示すように、ユーザ A 及びユーザ B に加えて、新たにユーザ C が加わり、3 人をグループ会話の参加者としたときの会話ロボット 1 における動作について以下説明する。

30

【 0 0 4 6 】

この場合、役割識別部 36 は、ユーザ C についても同様に、ユーザ C の動画像から得られたユーザ C の顔向き識別データ及び体向き識別データから、ユーザ C の顔正面 C1 の一部及び体正面 C2 の一部が、会話ロボット 1 側に向いているか否かを判断する。その結果、ユーザ C の顔向き識別データ及び体向き識別データから、ユーザ C の顔正面 C1 の一部及び体正面 C2 の一部が、会話ロボット 1 側に向いている場合、このことはユーザ C が会話ロボット 1 や他のユーザ A、ユーザ B とグループ会話を行うために、会話ロボット 1 側に顔正面 C1 及び体正面 C2 を向けていると判断し、役割識別部 36 は、ユーザ C がグループ会話の参加者であると認識する。

40

【 0 0 4 7 】

また、胴体部駆動制御部 42 は、役割識別部 36 から主注目位置検出データを受け取ると、上述と同様にして、角度中心点 O を基準に画枠中央線 d から主注目対象方向線 d1 までの主注目対象角度 θ_1 を算出するとともに、角度中心点 O を基準に画枠中央線 d から従注目対象方向線 d2 までの従注目対象角度 θ_2 を算出する。また、ここでは、図 4 (B) に示すように、ユーザ A 及びユーザ B に加えて、新たにユーザ C が認識されている。これにより、胴体部駆動制御部 42 は、角度中心点 O から、従注目対象物としたユーザ C の胴体中心方向に

50

延びる従注目対象方向線d3を算出し、角度中心点0を基準に画枠中央線dから従注目対象方向線d3までの従注目対象角度 θ_3 を算出する。

【0048】

そして、胴体部駆動制御部42は、これら主注目対象角度 θ_1 と従注目対象角度 θ_2 と従注目対象角度 θ_3 を全て合算して、認識したユーザ（ユーザA、ユーザB及びユーザC）の総数である「3」で除算し、画枠中央線dから重心方向線CG1までの重心方向角度 θ_{CG2} を算出して、これを胴体部回動命令として胴体部アクチュエータ44に送出する。すなわち、この胴体部駆動制御部42は、 $\theta_1 + \theta_2 + \theta_3 + \dots + \theta_n / n$ の計算式に基づいて、角度中心点0を基準に画枠中央線dからの重心方向角度 θ_{CG} を算出し得るようになされている（但し、 θ_1 は主注目対象角度、 $\theta_2 \sim \theta_n$ は従注目対象角度を示し、nは認識したユーザ総数を示す）。

10

【0049】

因みに、図3に示すように、行動選択部40は、音声処理部39から単語列データを受け取ると、データベース47に予め記憶されているキーワードを読み出して、当該単語列データの中に含まれるキーワードを抽出し、予め定められたテンプレートの中からこれら抽出したキーワード列と対応するテンプレートをデータベース47から読み出す。これにより行動選択部40は、キーワード列に基づいて選択された所定のテンプレートから、キーワード列がどのような意味を示しているのかを判断し得るようになされている。ここでデータベース47には、各テンプレート毎に、会話ロボット1が発話する発話内容や、腕部4a,4bを動かす等の動作内容を示す行動パターンが対応付けられた行動パターンテーブルが予め記憶されている。

20

【0050】

これにより、行動選択部40は、文字列データから抽出したキーワード列に対応する行動パターンを、テンプレートを基に行動パターンテーブルの中から選択し、この選択した行動パターンに対応付けられた所定の発話内容及び動作内容をデータベース47から読み出して、発話内容を音声合成部48に送出するとともに、動作内容を腕部4a,4b等の各駆動部に送出する。音声合成部48は、行動選択部40から与えられる発話内容を音声信号に変換する機能を有し、かくして得られた音声信号をスピーカ49に送出するようになされている。これによりこの音声信号に基づく音声をスピーカ49から出力させることができるようになされている。また、腕部4a,4b等の駆動部は、行動選択部40から与えられる動作内容を基に、ユーザAやユーザBの発話に応じて手部15を上げる等、状況に応じた自律的な動作を実現し得る。

30

【0051】

因みに、会話ロボット1は、図2(B)に示すように、ユーザA及びユーザBとの間でグループ会話を行なっている際に、図5(A)に示すように、例えばユーザBが会話ロボット1側に顔正面B1の一部及び体正面B2の一部が向かないように向きを変えると、ユーザBがグループ会話の参加者ではなくなったと判断し、図5(B)に示すように、ユーザA側に胴体部正面8bを向け、ユーザAとの間だけで会話を行なうようになされている。

【0052】

實際上、役割識別部36は、ユーザBの動画像から得られたユーザBの顔向き識別データ及び体向き識別データから、ユーザBの顔正面B1の一部及び体正面B2の一部が、会話ロボット1側に向いていないと判断すると（例えば、会話ロボット1に対してユーザBの顔が、斜め前左右90度以上のとき）、ユーザBがグループ会話の参加者でなりことを示す主従注目位置検出データを生成し、これを胴体部駆動制御部42にそれぞれ送出する。

40

【0053】

胴体部駆動制御部42は、上体部8を主注目対象角度 θ_1 まで回動させるのに必要な上体部回動角度を算出し、この上体部回動角度を胴体部回動命令として胴体部アクチュエータ44に送出する。胴体部アクチュエータ44は、胴体部回動命令に基づいて上体部回動角度だけ上体部8を回動させることにより、頭部正面3aが向いている主注目対象方向に胴体部正面8bも向けさせ得る（図5(B)）。これにより、会話ロボット1は、ユーザBがグルー

50

プ会話から離脱しても、ユーザ A に対し胴体部正面8bを向かせることにより、ユーザ A との間で会話を続ける意思表示を行なえ、状況に応じた自然な会話を行い得る。

【 0 0 5 4 】

(3) 動作及び効果

以上の構成において、会話ロボット 1 では、ユーザ A やユーザ B の音声や動画像に基づいて、ユーザ A 及びユーザ B のうちいずれが発話者であるかを認識し、ユーザ A 又はユーザ B の発話内容に応じて、スピーカ49から発する発話内容や、腕部4a,4bの動作等の行動パターンを変化させ、ユーザ A やユーザ B との会話に応じた自律的行動を実現できる。

【 0 0 5 5 】

また、この会話ロボット 1 では、ユーザ A 及びユーザ B のうち、発話者や主聴者、或いは発話終了後の次発話推定者を注目対象物と認識し、当該注目対象物の方向に頭部正面3aが向くように頭部 3 を回動させる。これにより、会話ロボット 1 では、あたかも注目対象物となるユーザ A 又はユーザ B の動作に応じて、ユーザ A 又はユーザ B と会話するため注視しているかのような印象を与えることができ、かくしてユーザ A 又はユーザ B との間で状況に応じた自然な会話を行い得る。

10

【 0 0 5 6 】

これに加えて、この会話ロボット 1 では、注目対象角度 1 及び従注目対象角度 2 を算出し、これら注目対象角度 1 と従注目対象角度 2 とを全て合算した後、認識したユーザ (ユーザ A 及びユーザ B) の総数で除算して重心方向角度 CG_1 を算出し、胴体部正面8bをこの重心方向角度 CG_1 まで回動させる。このように、会話ロボット 1 では、胴体部正面8bを重心方向角度 CG_1 まで回動させることで、重心方向に胴体部正面8bを向けさせて、ユーザ A だけでなくユーザ B に対しても、あたかも注視しているかのような印象を与えることができ、かくしてユーザ A 又はユーザ B との間で状況に応じた自然な会話を実現し得る。

20

【 0 0 5 7 】

(4) 他の実施の形態

なお、本発明は、本実施形態に限定されるものではなく、本発明の要旨の範囲内で種々の変形実施が可能であり、例えば犬等の動物に似せた会話ロボットを適用してもよい。また、上述した実施の形態においては、対象物としてのユーザ A 及びユーザ B のうち、ユーザ A を注目対象物とし、ユーザ B やユーザ C を従注目対象物とした場合について述べたが、本発明はこれに限らず、例えば筆記具により文字を記載可能な掲示型のボードや、各種情報が表示された表示装置、掲示型の印刷物等を注目対象物又は従注目対象物 (対象物) としてもよい。

30

【 0 0 5 8 】

この場合、会話ロボット 1 は、注目対象物としてボードを認識し、従注目対象物としてユーザ B を認識したとき、腕部4a,4bを可動制御する指示部制御手段 (図示せず) によって、指示部としての腕部4a,4bを注目対象物たるボードの方向に向けるように、当該腕部4a,4bの可動を制御するようにしてもよい。

【 0 0 5 9 】

この際、会話ロボット 1 は、胴体部駆動制御部42によって、注目対象物としてのボードの位置から求めた注目対象角度 1 と、従注目対象物としてのユーザ B の位置から求めた従注目対象角度 2 とを全て合算して、認識した対象物 (ボード及びユーザ B) の総数である「 2 」で除算し、角度中心点 O を基準に画枠中央線 d から重心方向線 CG までの重心方向角度 CG を算出して、これを胴体部回動命令として胴体部アクチュエータ44に送出する。

40

【 0 0 6 0 】

これにより、会話ロボット 1 では、胴体部アクチュエータ44により、胴体部回動命令に基づいて重心方向角度 CG だけ上体部 8 を回動させ、胴体部正面8bを重心方向線 CG1 側に向けさせることができる。かくして、会話ロボット 1 の上体部 8 だけを回動させて、胴体部正面8bを重心方向へ向けることで、あたかも会話ロボット 1 が、ボードを腕部4a,4bで

50

指示しつつ、上体部 8 の一部をユーザ B 側に向けて、ボードに注目しつつユーザ B との間で会話を行なっているかのような印象を与えることができる。

【 0 0 6 1 】

また、上述した実施の形態においては、撮像画枠 51 内のユーザ A 及びユーザ B の位置を基に重心方向を求め、基台 6 を動かすことなく、上体部 8 のみを回動させて重心方向に胴体部正面 8b を向けさせるようにした場合について述べたが、本発明はこれに限らず、例えば撮像画枠 51 内のユーザ A 及びユーザ B の位置を基に重心方向を求め、この重心方向へ胴体部正面 8b を向けるための回動角度が、所定の回動角度以上であるとき、上体部 8 の回動を所定の回動角度範囲内になるように、移動制御手段によって基台 6 の車輪 5 を回動させ、胴体部 2 自体を回動角度方向に回動させたり、或いは胴体部 2 を回動角度方向側に平行移動させる等して、回動角度方向に胴体部 2 を自動的に移動させるようにしてもよい。

10

【 0 0 6 2 】

さらに、上述した実施の形態においては、撮像画枠 51 内の全てのユーザ A 及びユーザ B の位置から重心方向を求め、当該上体部 8 を重心方向へ回動させるようにした場合について述べたが、本発明はこれに限らず、撮像画枠 51 内の全てのユーザ A 及びユーザ B の位置から求めた重心方向に、胴体部正面 8b を向けると、所定回動角度以上に上体部 8 を回動させる必要があるとき、当該所定回動角度以内となるように、所定回動角度以上の位置に存在するユーザ A 又はユーザ B を除外する等し、ユーザ A 又はユーザ B の中から一部を除外して重心方向を求めるようにしてもよい。

20

【 0 0 6 3 】

さらに、上述した実施の形態においては、外部センサとして、会話ロボット 1 の目部 21 に撮像手段である CCD カメラ 23 を設け、この CCD カメラ 23 で撮像された撮像画枠 51 内のユーザ A 及びユーザ B の位置から重心方向を求めるようにした場合について述べたが、本発明はこれに限らず、例えば外部センサとして室内天井に撮像手段であるカメラを設置し、ユーザ A、ユーザ B 及び会話ロボット 1 の位置関係を上方からカメラで撮像し、この撮像画像内のユーザ A 及びユーザ B の位置から重心方向を求め、当該重心方向に会話ロボット 1 の胴体部正面 8b を向けるようにしてもよい。

【 0 0 6 4 】

また、上述した実施の形態のその他の形態として、例えば、音源方向の特定に関しては、例えば「ロボット頭部に設置した 2 系統のマイクによる音源定位（日本音響学会、春季研究発表 講演論文誌 p p 469-470 1999 小林哲則、宮田大介、松坂要佐）」に記載された技術内容を用いてもよく、この場合、頭部 3 に 2 つのマイクロホンを設け、各マイクロホンで受信した信号から音源定位を行い、おおよその到来方向を求めた後、CCD カメラ 23 で得られた動画像を画像処理してその方向にいるユーザを検索し、発話者として認定するようにしてもよい。

30

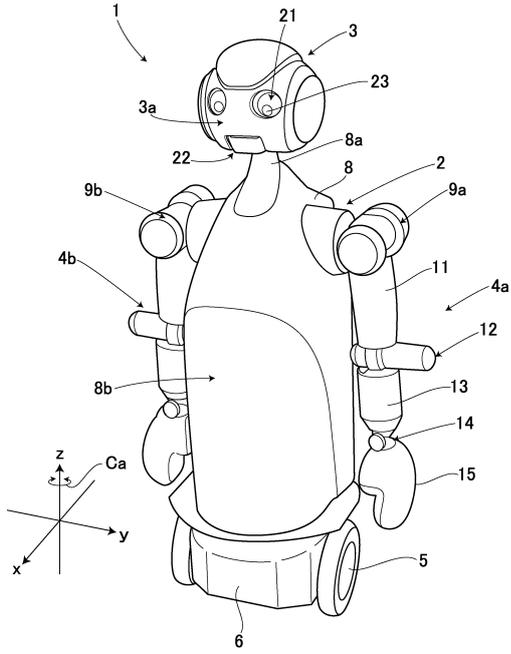
【符号の説明】

【 0 0 6 5 】

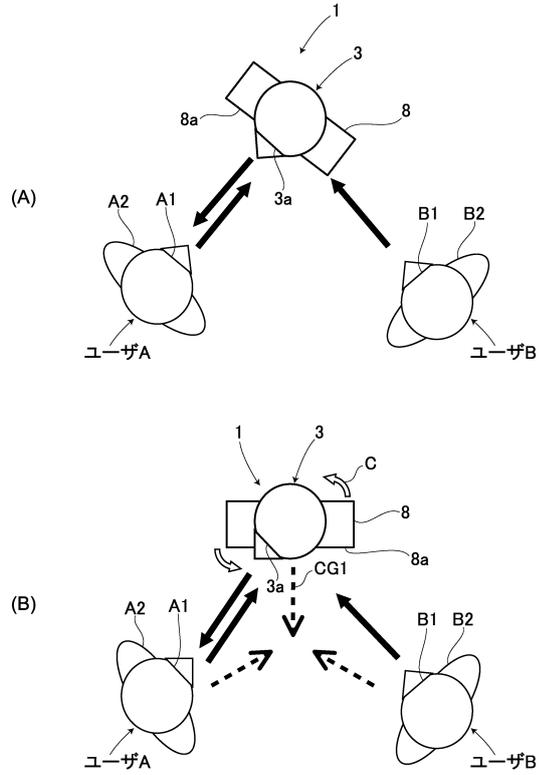
- 1 会話ロボット
- 2 胴体部
- 3 頭部
- 4a, 4b 腕部（指示部）
- 5 車輪（移動手段）
- 8 上体部
- 23 CCD カメラ（外部センサ）
- 37a, 37b マイクロホン（外部センサ）
- 35 位置検出部（位置検出手段）
- 36 役割識別部（役割識別手段）
- 41 頭部駆動制御部（頭部制御手段）
- 42 胴体部駆動制御部（胴体部制御手段）

40

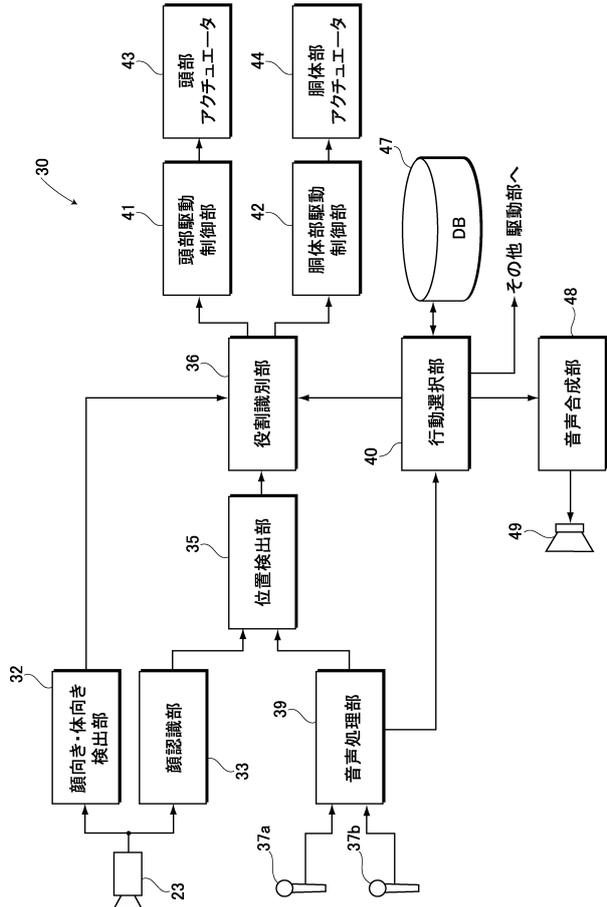
【図1】



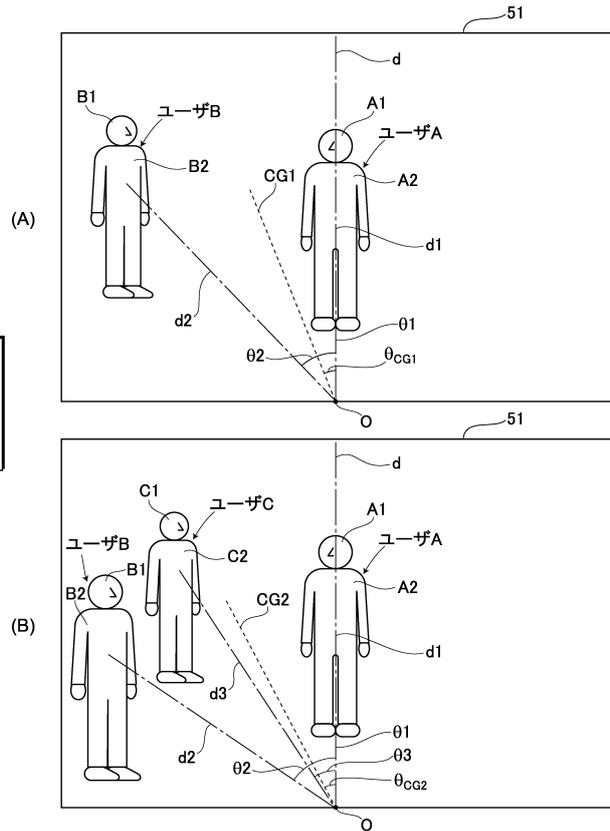
【図2】



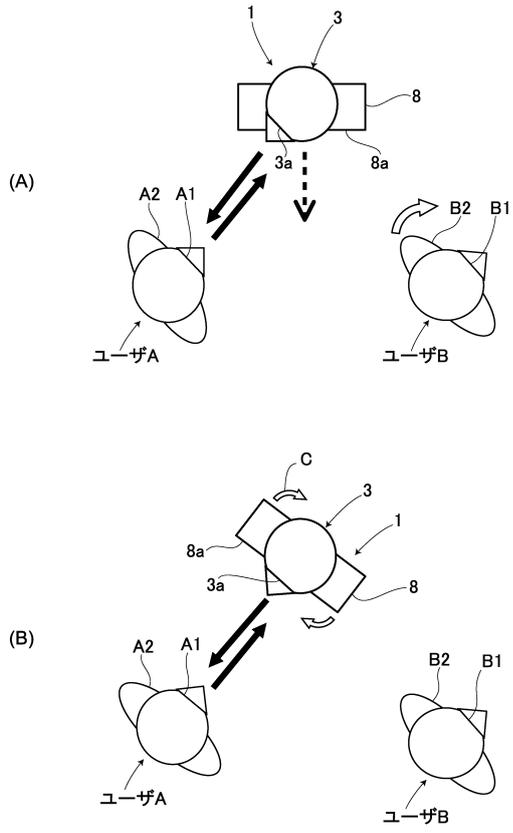
【図3】



【図4】



【 図 5 】



フロントページの続き

- (72)発明者 小林 哲則
東京都新宿区戸塚町1丁目104番地 学校法人早稲田大学内
- (72)発明者 藤江 真也
東京都新宿区戸塚町1丁目104番地 学校法人早稲田大学内
- (72)発明者 松山 洋一
東京都新宿区戸塚町1丁目104番地 学校法人早稲田大学内

審査官 鈴木 崇文

- (56)参考文献 特開2007-222968(JP,A)
特開2008-087140(JP,A)
特開2006-231447(JP,A)
特開2006-247780(JP,A)
特開2007-118129(JP,A)
松坂 要佐, 東條 剛史, 小林 哲則, グループ会話に参与する対話ロボットの構築, 電子情報通信学会論文誌. D-II, 情報・システム, II-パターン処理, 2001年 6月, pp.898-908
Mutlu, B. ; Shiwa, T. ; Kanda, T. ; Ishiguro, H. ; Hagita, N. , Footing in human-robot conversations: How robots might shape participant roles using gaze cues , Human-Robot Interaction (HRI), 2009 4th ACM/IEEE International Conference on , 2009年 3月
Shiomi, M. ; Kanda, T. ; Ishiguro, H. ; Hagita, N. , A larger audience, please! Encouraging people to listen to a guide robot , Human-Robot Interaction (HRI), 2010 5th ACM/IEEE International Conference on , 2010年 3月
Bennewitz, M. ; Faber, F. ; Joho, D. ; Schreiber, M. ; Behnke, S. , Integrating vision and speech for conversations with multiple persons , Intelligent Robots and Systems, 2005. (IROS 2005). 2005 IEEE/RSJ International Conference on , 2005年 8月

(58)調査した分野(Int.Cl. , DB名)

B25J 1/00 - 21/02

A63H 11/00