

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6550643号
(P6550643)

(45) 発行日 令和1年7月31日 (2019.7.31)

(24) 登録日 令和1年7月12日 (2019.7.12)

(51) Int. Cl.		F I			
G 0 6 F	3/01	(2006.01)	G 0 6 F	3/01	5 7 0
B 2 5 J	13/08	(2006.01)	B 2 5 J	13/08	Z

請求項の数 11 (全 32 頁)

(21) 出願番号	特願2015-46241 (P2015-46241)	(73) 特許権者	000005326
(22) 出願日	平成27年3月9日 (2015.3.9)		本田技研工業株式会社
(65) 公開番号	特開2015-187856 (P2015-187856A)		東京都港区南青山二丁目1番1号
(43) 公開日	平成27年10月29日 (2015.10.29)	(74) 代理人	100165179
審査請求日	平成29年11月29日 (2017.11.29)		弁理士 田▲崎▼ 聡
(31) 優先権主張番号	特願2014-52295 (P2014-52295)	(74) 代理人	100126664
(32) 優先日	平成26年3月14日 (2014.3.14)		弁理士 鈴木 慎吾
(33) 優先権主張国	日本国 (JP)	(74) 代理人	100154852
			弁理士 酒井 太一
		(74) 代理人	100194087
			弁理士 渡辺 伸一
		(74) 代理人	100064908
			弁理士 志賀 正武
		(74) 代理人	100146835
			弁理士 佐伯 義文

最終頁に続く

(54) 【発明の名称】 動作推定装置、ロボット、及び動作推定方法

(57) 【特許請求の範囲】

【請求項1】

センサと対象物との距離を取得する取得部と、

人に対して挙手の終了を促した時刻に基づき設定される第1時刻における前記取得部によって取得された前記距離に基づいて人の反応を推定する状況推定部と、

前記取得部によって取得された前記距離に基づいて人を検出し、前記検出した人毎に識別子を付与する人処理部と、

人の反応を推定する所定の範囲を1つ以上の領域に分割し、前記取得部によって取得された前記距離に基づいて該領域毎の距離を求める領域距離処理部と、

を備え、

前記状況推定部は、

人に対して挙手の終了を促した時刻に基づき設定される第1期間における前記取得部によって取得された前記距離に基づいて人の反応を推定し、

前記人処理部によって前記識別子が付与された前記人の前記距離を用いて、前記第1時刻と前記第1期間それぞれにおける前記識別子毎の高さ方向の距離を抽出し、前記第1時刻と前記第1期間それぞれにおける前記抽出した前記識別子毎の高さ方向の距離同士を比較することで、人の反応を推定し、

前記人処理部によって前記識別子が付与された前記人の前記距離と、前記領域距離処理部によって求められた距離のうち、すくなくとも1つの距離によって人の反応を推定する動作推定装置。

【請求項 2】

前記状況推定部は、

前記第 1 時刻と前記第 1 期間それぞれにおける対象物の高さ方向の距離を抽出し、前記第 1 時刻に前記抽出した高さ方向の距離と、前記第 1 期間に前記抽出した高さ方向の距離とを比較することで、人の反応を推定する請求項 1 に記載の動作推定装置。

【請求項 3】

前記領域距離処理部は、

前記第 1 時刻と前記第 1 期間における前記領域毎の高さ方向の距離を抽出し、前記第 1 時刻と前記第 1 期間における前記抽出した前記領域毎の高さ方向の距離に基づく値を比較することで、人の反応を推定する請求項 1 または請求項 2 に記載の動作推定装置。

10

【請求項 4】

前記領域毎の高さ方向の距離に基づく値は、

前記第 1 時刻と前記第 1 期間における前記領域毎の高さ方向の距離が抽出された領域における前記高さ方向の距離の平均値である請求項 3 に記載の動作推定装置。

【請求項 5】

前記人処理部が付与した識別子の距離を追跡する距離追跡部、を備え、

前記状況推定部は、

前記距離追跡部が付与した前記識別子の数を参照し、所定の期間に追跡不能になった前記識別子の数が所定の値を超えたとき、前記領域距離処理部によって求められた距離を用いて人の反応を推定し、所定の期間に追跡不能になった前記識別子の数が所定の値より少ないとき、前記人処理部によって求められた距離を用いて、人の反応を推定する請求項 1 から請求項 4 のいずれか 1 項に記載の動作推定装置。

20

【請求項 6】

前記第 1 時刻または前記第 1 期間は、

前記領域距離処理部によって測定される前記距離の変化が所定の値より小さくなったときに対応して決定される請求項 1 から請求項 5 のいずれか 1 項に記載の動作推定装置。

【請求項 7】

前記状況推定部は、

人が挙手を行っていない第 2 期間における前記取得部によって取得された前記距離と、人の体格データとに基づいて算出される挙手をしたときに人の手が存在する高さ方向の範囲である挙手存在範囲を算出し、算出した前記挙手存在範囲にある前記第 1 時刻または前記第 1 期間における前記取得部によって取得された前記距離を抽出し、前記抽出した距離に基づいて人の反応を推定する請求項 1 から請求項 6 のいずれか 1 項に記載の動作推定装置。

30

【請求項 8】

前記人の反応は、人が挙手を行っている反応か挙手を行っていない反応かであり、

前記状況推定部は、

前記人の反応を複数回推定して、前記複数回推定した結果に基づいて、多数決の判定を行う請求項 1 から請求項 7 のいずれか 1 項に記載の動作推定装置。

【請求項 9】

所定の範囲に対して自センサと対象物との距離の測定を行うセンサ、を備える請求項 1 から請求項 8 のいずれか 1 項に記載の動作推定装置。

40

【請求項 10】

前記人に対して挙手の終了を促すこと示す信号を出力する出力部と、

請求項 1 から請求項 9 のいずれか 1 項に記載の動作推定装置を備えるロボット。

【請求項 11】

取得部が、センサと対象物との距離を取得する取得手順と、

人処理部が、前記取得手順によって取得された前記距離に基づいて人を検出し、前記検出した人毎に識別子を付与する人処理手順と、

領域距離処理部が、人の反応を推定する所定の範囲を 1 つ以上の領域に分割し、前記取

50

得部によって取得された前記距離に基づいて該領域毎の距離を求める領域距離処理手順と、

状況推定部が、人に対して挙手の終了を促した時刻に基づき設定される第1期間における前記取得部によって取得された前記距離に基づいて人の反応を推定する手順と、

前記状況推定部が、前記人処理手順によって前記識別子が付与された前記人の前記距離を用いて、人に対して挙手の終了を促した時刻に基づき設定される第1時刻と前記第1期間それぞれにおける前記識別子毎の高さ方向の距離を抽出し、前記第1時刻と前記第1期間それぞれにおける前記抽出した前記識別子毎の高さ方向の距離同士を比較することで、人の反応を推定する手順と、

前記状況推定部が、前記人処理手順によって前記識別子が付与された前記人の前記距離と、前記領域距離処理部によって求められた距離のうち、すくなくとも1つの距離によって人の反応を推定する手順と、

を含む動作推定方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、動作推定装置、ロボット、及び動作推定方法に関する。

【背景技術】

【0002】

腕、手、指、足、頭などの身体の各部分の動きによって意志や感情が表現されるジェスチャーを認識する装置（以下、ジェスチャー認識装置という）が提案されている。

【0003】

このようなジェスチャー認識装置では、画像センサによって生成された検知対象エリアに存在する物体までの距離値を画素値とする画像に基づいて、検知対象エリアに存在する人の特定部位を抽出する。そして、ジェスチャー認識装置では、抽出された特定部位の形状の時系列データに基づいて、特定空間での所定のジェスチャーを認識することが提案されている。なお、特定部位とは、例えば人の手の手首より先の部分であり、所定のジェスチャーとは、例えば、画像センサに向けて手がかざされる動作である（例えば、特許文献1参照）。

【先行技術文献】

【特許文献】

【0004】

【特許文献1】特開2006-99749号公報

【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、例えば画像センサに向けて手をかざすタイミングはジェスチャー認識装置にとって予測が困難であった。このため、特許文献1に記載の技術では、検知対象エリアにいる人のジェスチャーを短時間で精度良く認識することができない場合がある。

【0006】

本発明は、上記の問題点を鑑みてなされたものであって、検知対象エリアにいる人の動作を高精度に検出することができる動作推定装置、ロボット、及び動作推定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】

（1）上記目的を達成するため、本発明の一態様に係る動作推定装置は、センサと対象物との距離を取得する取得部と、人に対して挙手の終了を促した時刻に基づき設定される第1時刻における前記取得部によって取得された前記距離に基づいて人の反応を推定する状況推定部と、前記取得部によって取得された前記距離に基づいて人を検出し、前記検出した人毎に識別子を付与する人処理部と、人の反応を推定する所定の範囲を1つ以上の領域

10

20

30

40

50

に分割し、前記取得部によって取得された前記距離に基づいて該領域毎の距離を求める領域距離処理部と、を備え、前記状況推定部は、人に対して挙手の終了を促した時刻に基づき設定される第1期間における前記取得部によって取得された前記距離に基づいて人の反応を推定し、前記人処理部によって前記識別子が付与された前記人の前記距離を用いて、前記第1時刻と前記第1期間それぞれにおける前記識別子毎の高さ方向の距離を抽出し、前記第1時刻と前記第1期間それぞれにおける前記抽出した前記識別子毎の高さ方向の距離同士を比較することで、人の反応を推定し、前記人処理部によって前記識別子が付与された前記人の前記距離と、前記領域距離処理部によって求められた距離のうち、すくなくとも1つの距離によって人の反応を推定する。

(2) また、本発明の一態様に係る動作推定装置であって、前記状況推定部は、前記第1時刻と前記第1期間それぞれにおける対象物の高さ方向の距離を抽出し、前記第1時刻に前記抽出した高さ方向の距離と、前記第1期間に前記抽出した高さ方向の距離とを比較することで、人の反応を推定するようにしてもよい。

(3) また、本発明の一態様に係る動作推定装置であって、前記領域距離処理部は、前記第1時刻と前記第1期間における前記領域毎の高さ方向の距離を抽出し、前記第1時刻と前記第1期間における前記抽出した前記領域毎の高さ方向の距離に基づく値を比較することで、人の反応を推定するようにしてもよい。

(4) また、本発明の一態様に係る動作推定装置であって、前記領域毎の高さ方向の距離に基づく値は、前記第1時刻と前記第1期間における前記領域毎の高さ方向の距離が抽出された領域における前記高さ方向の距離の平均値であるようにしてもよい。

(5) また、本発明の一態様に係る動作推定装置は、前記人処理部が付与した識別子の距離を追跡する距離追跡部、を備え、前記状況推定部は、前記距離追跡部が付与した前記識別子の数を参照し、所定の期間に追跡不能になった前記識別子の数が所定の値を超えたとき、前記領域距離処理部によって求められた距離を用いて人の反応を推定し、所定の期間に追跡不能になった前記識別子の数が所定の値より少ないとき、前記人処理部によって求められた距離を用いて、人の反応を推定するようにしてもよい。

(6) また、本発明の一態様に係る動作推定装置であって、前記第1時刻または前記第1期間は、前記領域距離処理部によって測定される前記距離の変化が所定の値より小さくなったときに対応して決定されるようにしてもよい。

(7) また、本発明の一態様に係る動作推定装置であって、記状況推定部は、人が挙手を行っていない第2期間における前記取得部によって取得された前記距離と、人の体格データとに基づいて算出される挙手をしたときに人の手が存在する高さ方向の範囲である挙手存在範囲を算出し、算出した前記挙手存在範囲にある前記第1時刻または前記第1期間における前記取得部によって取得された前記距離を抽出し、前記抽出した距離に基づいて人の反応を推定するようにしてもよい。

(8) また、本発明の一態様に係る動作推定装置であって、前記人の反応は、人が挙手を行っている反応か挙手を行っていない反応かであり、状況推定部は、人の反応を複数回推定して、複数回推定した結果に基づいて、多数決の判定を行う。

(9) また、本発明の一態様に係る動作推定装置は、所定の範囲に対して自センサと対象物との距離の測定を行うセンサ、を備えるようにしてもよい。

【0008】

(10) 上記目的を達成するため、本発明の一態様に係るロボットは、前記人に対して挙手の終了を促すこと示す信号を出力する出力部と、上記(1)から(11)のいずれか1つに記載の動作推定装置を備える。

(11) 上記目的を達成するため、本発明の一態様に係る動作推定方法は、取得部が、センサと対象物との距離を取得する取得手順と、人処理部が、前記取得手順によって取得された前記距離に基づいて人を検出し、前記検出した人毎に識別子を付与する人処理手順と、領域距離処理部が、人の反応を推定する所定の範囲を1つ以上の領域に分割し、前記取得部によって取得された前記距離に基づいて該領域毎の距離を求める領域距離処理手順と、状況推定部が、人に対して挙手の終了を促した時刻に基づき設定される第1期間にお

10

20

30

40

50

る前記取得部によって取得された前記距離に基づいて人の反応を推定する手順と、前記状況推定部が、前記人処理手順によって前記識別子が付与された前記人の前記距離を用いて、人に対して挙手の終了を促した時刻に基づき設定される第1時刻と前記第1期間それぞれにおける前記識別子毎の高さ方向の距離を抽出し、前記第1時刻と前記第1期間それぞれにおける前記抽出した前記識別子毎の高さ方向の距離同士を比較することで、人の反応を推定する手順と、前記状況推定部が、前記人処理手順によって前記識別子が付与された前記人の前記距離と、前記領域距離処理部によって求められた距離のうち、すくなくとも1つの距離によって人の反応を推定する手順と、を含む。

【発明の効果】

【0009】

本発明の態様(1)、(9)、(10)、及び(11)によれば、手が挙げられる期間に比べて短時間に精度良く人の反応を推定することができる。

また、本発明の態様(1)、及び(2)によれば、人が手を下げる前の期間に人が手を挙げているときの高さ方向の値と、人が手を下げたときの高さ方向の値とに基づいて、手を挙げている人数を検出することができる。

また、本発明の態様(1)、(3)、及び(4)によれば、手を挙げている期間と、手を下げたときとの、領域における高さ方向の変化量を推定することができる。そして、動作推定装置3では、推定された結果に基づいて、多数決によってどの質問への挙手が最も多いかを推定することができる。

また、本発明の態様(5)によれば、検出結果に応じて、領域距離処理部によって求められた距離、または人処理部によって求められた距離を用いて、人の反応を推定することができる。

また、本発明の態様(6)によれば、人に対して挙手の終了を促すこと示すことにより設定されるときである手下げ時、及び人に対して挙手の終了を促すこと示すことにより設定される第1期間を、領域距離処理部によって検出された結果に基づいて決定することができる。

また、本発明の態様(7)によれば、個人を特定する必要がないので人が密集している領域が存在しても、人の反応を推定することができる。

また、本発明の態様(8)によれば、多数決の判定を行うことができる。

【図面の簡単な説明】

【0010】

【図1】第1実施形態に係るロボットの概略のブロック図である。

【図2】第1実施形態に係る音声信号の出力するタイミングと挙手を検出するタイミングを説明する図である。

【図3】第1実施形態に係るロボットによる挙手検出の処理手順のフローチャートである。

【図4】第1実施形態に係るLRFセンサによる測定とセンシング処理部と人統合処理部とによる処理を説明する図である。

【図5】第1実施形態に係るクラスタリング処理、人クラスタデータの生成の処理手順のフローチャートである。

【図6】第1実施形態に係る人統合処理部によって行われる統合処理手順のフローチャートである。

【図7】第1実施形態に係るトラッキングを行った結果の例を示す図である。

【図8】第1実施形態に係るLRFセンサによる測定とセンシング処理部とグリッド統合処理部による処理を説明する図である。

【図9】第1実施形態に係るグリッドデータの生成の処理手順のフローチャートである。

【図10】第1実施形態に係るグリッド統合処理部によって行われる統合処理手順のフローチャートである。

【図11】第1実施形態に係る第1処理及び第2処理を行うタイミングを説明する図である。

10

20

30

40

50

【図 1 2】第 2 実施形態に係るロボットの概略のブロック図である。

【図 1 3】第 2 実施形態に係る人の部位の高さを説明する図である。

【図 1 4 A】人が手を下げているイメージ図である。

【図 1 4 B】第 2 実施形態に係る人が手を下げている期間の x y z 値の第 1 マップの例を説明する図である。

【図 1 5 A】人が手を挙げているイメージ図である。

【図 1 5 B】第 2 実施形態に係る人が手を挙げている期間の x y z 値の第 2 マップの例を説明する図である。

【図 1 6】第 2 実施形態に係る状況推定部が第 1 マップ～第 5 マップを用いて行う処理を説明する図である。

【図 1 7】第 2 実施形態に係る状況推定部がクラスタリングを行って手を挙げている人数を推定する処理を説明する図である。

【図 1 8】第 2 実施形態に係る状況推定部の処理手順のフローチャートである。

【図 1 9】人に対して挙手の終了を促した時刻、第 1 時刻、及び第 1 期間を説明する図である。

【発明を実施するための形態】

【0 0 1 1】

まず、本発明の概要を説明する。

例えば、イベント会場において、人型のロボットが司会進行を行う場合がある。このような状況では、予めロボットに、シナリオ、動作、及び音声信号などを記憶させておき、記憶させた動作を行わせながら、シナリオに応じたタイミングで記憶されている音声信号を再生させる場合がある。このようなロボットの動作では、会場にいる人と、ロボットとのコミュニケーションが取りにくい場合がある。

このため、本実施形態では、会場にいる人達の反応に応じて、ロボットが次に行う動作や、出力する音声信号を選択することで、ロボットと人とのコミュニケーションを豊かにする。本実施形態では、ロボットが、会場にいる人による挙手を推定して、推定した結果に基づいて多数決を行うことで、提示した話題の中から話題の選択を行う。

【0 0 1 2】

以下、図面を参照しながら本発明の実施形態について説明する。

【0 0 1 3】

[第 1 実施形態]

図 1 は、本実施形態に係るロボット 1 の概略のブロック図である。

図 1 に示すように、ロボット 1 は、処理部 2、及び動作推定装置 3 を含んで構成される。ロボット 1 は、 n (n は 1 以上の整数) 個の LRF (Laser Range Finder; レーザレンジファインダー) センサ 10-1 ~ 10- n に無線または有線で接続される。なお、LRF センサ 10-1 ~ 10- n のうちいずれか 1 つを特定しない場合は、LRF センサ 10 という。なお、本実施形態では、動作推定装置 3 をロボット 1 が備える例を説明するが、これに限られない。例えば、イベント会場等で用いられるガイドシステム等が動作推定装置 3 を備えるようにしてもよい。

【0 0 1 4】

LRF センサ 10 は、レーザ光を照射しながら、検知対象の空間を所定のピッチでスキャンして、物体に反射して戻ってくるまでの時間を測定することで、物体との距離と方向を検出する 3 次元の測域センサである。LRF センサ 10-1 ~ 10- n それぞれは、検出した物体との距離と方向を含む検出結果である測距点の情報を、対応するセンシング処理部 20-1 ~ 20- n (取得部) に出力する。なお、測距点の情報には、x 軸成分、y 軸成分、及び z 軸成分が含まれる。LRF センサ 10-1 ~ 10- n それぞれは、検知対象の空間の一部が重なるように、人を俯瞰できる位置、例えば検知対象の空間の天井や壁などに設置される。レーザ光の波長は、例えば 905 nm、785 nm 等である。また、1 つの LRF センサ 10 のスキャン範囲は、例えば、水平面において 270 度であり、ステップ角度が 0.25 度、ステップ数が 1080 ステップである。なお、本実施形態では、

10

20

30

40

50

センサの例として、L R Fセンサ 1 0 を用いる例を説明するが、これに限られず、他のセンサであってもよい。

【0 0 1 5】

まず、処理部 2 の構成について説明する。

処理部 2 は、ロボット 1 の動作等、各種の制御を行う。処理部 2 は、記憶部 1 0 1、制御部 1 0 2、及び音声出力部 1 0 3 を備える。

記憶部 1 0 1 には、ロボット 1 の制御に関する情報、音声出力データ、シナリオ情報等が記憶されている。なお、シナリオ情報とは、所定の動作を所定の時間に行うことが記述されている情報である。

【0 0 1 6】

制御部 1 0 2 は、記憶部 1 0 1 に記憶されているシナリオ情報を読み出し、読み出したシナリオ情報に基づいて、音声信号を音声出力部 1 0 3 から出力させる。また、制御部 1 0 2 は、音声信号を出力させるタイミングでトリガー信号 d 5 を生成し、生成したトリガー信号 d 5 を状況推定部 5 0 に出力する。また、制御部 1 0 2 は、動作推定装置 3 から入力される検出結果を示す情報（以下、検出情報という）d 6 に基づいて、シナリオ情報の中から次に行う処理を選択する。

音声出力部 1 0 3 は、制御部 1 0 2 の処理に応じて、音声信号を出力する。音声出力部 1 0 3 は、例えばスピーカーである。

【0 0 1 7】

次に、動作推定装置 3 の構成について説明する。

動作推定装置 3 は、n 個のセンシング処理部 2 0 - 1 ~ 2 0 - n、人統合処理部（人処理部、距離追跡部）3 0、グリッド統合処理部（領域距離処理部）4 0、状況推定部 5 0、及び記憶部 6 0 を備える。センシング処理部 2 0 - 1 ~ 2 0 - n のうちいずれか 1 つを特定しない場合は、センシング処理部 2 0 という。また、L R Fセンサ 1 0 - 1 ~ 1 0 - n それぞれは、対応するセンシング処理部 2 0 - 1 ~ 2 0 - n に無線または有線を介して接続されている。また、動作推定装置 3 は、L R Fセンサ 1 0 を含んで構成されていてもよい。

【0 0 1 8】

センシング処理部 2 0 は、L R Fセンサ 1 0 が出力した検出結果を取得する。センシング処理部 2 0 は、L R Fセンサ 1 0 から入力された検出結果を用いて、例えば所定の周期毎に測距点を分類する（クラスタリング）。センシング処理部 2 0 は、分類したクラスタ毎の重心位置を算出し、算出した重心位置を、位置を示す情報とする。センシング処理部 2 0 は、分類したクラスタ毎に高さの最大値を抽出し、抽出した最大値を、高さを示す情報とする。そして、センシング処理部 2 0 は、クラスタ毎に、位置を示す情報と高さを示す情報とを関連づけて人クラスタデータ d 1 を生成し、生成した人クラスタデータ d 1 を人統合処理部 3 0 に出力する。

また、センシング処理部 2 0 は、L R Fセンサ 1 0 から入力された検出結果を用いて、例えば所定の周期毎に測距点を、予め定められているグリッドに分類する。センシング処理部 2 0 は、分類したグリッド毎に高さの最大値を抽出し、抽出した最大値を、高さを示す情報とする。そして、センシング処理部 2 0 は、グリッド毎に、グリッドの位置を示す情報と高さを示す情報とを関連づけてグリッドデータ d 2 を生成し、生成したグリッドデータ d 2 をグリッド統合処理部 4 0 に出力する。

【0 0 1 9】

センシング処理部 2 0 は、測距点を、例えば最短距離法などの階層的手法や、k - m e a n s 法などの分割最適化手法を用いて、分類を行う。センシング処理部 2 0 は、高さに対応する z 軸方向の値が所定の範囲内または所定の高さ以上であり、かつ x y 平面上に所定の密度ある測距点を 1 つのクラスタとして検出する。例えば、センシング処理部 2 0 は、x y 平面上に投影した測距点に対して、階層的クラスタリングを行うことで、得られた測距点を分類する。

【0 0 2 0】

人統合処理部30は、センシング処理部20-1~20-nそれぞれから入力された人クラスタデータd1を統合する。人統合処理部30は、統合した人クラスタデータを用いて、人毎に識別子(ID)を割り当てる。人統合処理部30は、IDを割り当てた人クラスタデータに対してトラッキング(以下、IDトラッキングという)を行う。なお、IDトラッキングは、例えば、後述するように、測距点をxy平面に投影し、投影した投影点が分類されたもの、クラスタの重心位置を用いて行う。人統合処理部30は、IDを割り当てた位置を示す情報(以下、人位置ともいう)と、IDを割り当てた人クラスタデータと、割り当てたIDとを対応づけて人データを生成する。人統合処理部30は、生成した人データd3を状況推定部50に出力する。

【0021】

グリッド統合処理部40は、センシング処理部20-1~20-nそれぞれから入力されたグリッドデータd2を統合する。グリッド統合処理部40は、統合したグリッドデータを用いて、グリッド毎の高さの最大値を抽出する。そして、グリッド統合処理部40は、グリッドの高さが抽出されたグリッドの数をカウントする。グリッド統合処理部40は、グリッドの高さが抽出されたグリッドにおけるグリッドの高さの平均値を算出し、算出した高さを、グリッド高さを示す情報d4として状況推定部50に出力する。

【0022】

状況推定部50には、制御部102からシナリオ情報に基づくトリガー信号d5が入力される。なお、トリガー信号d5には、後述する音声信号の種類を示す情報が含まれているようにしてもよい。状況推定部50は、入力されたトリガー信号d5に応じて、人統合処理部30から入力された人データd3を用いて、所定の期間(第1期間)、ID毎に人データに含まれる高さを示す情報の中から最大値を検出する。状況推定部50は、所定の期間に検出された最大値を、そのIDの第1クラスタ高さ値として、記憶部60に記憶させる。

また、状況推定部50は、入力されたトリガー信号d5に応じて、グリッド統合処理部40から入力されたグリッド高さを示す情報d4を用いて、所定の期間(第1期間)におけるグリッド高さを示す情報の最大値を検出し、所定の期間に検出された最大値を第1グリッド高さ値として、記憶部60に記憶させる。

【0023】

状況推定部50は、入力されたトリガー信号d5に応じて、人統合処理部30から入力された人データd3を用いて、所定の時刻(第1時刻)に、ID毎に人データに含まれる高さを示す情報を、そのIDの第2クラスタ高さ値として、記憶部60に記憶させる。

また、状況推定部50は、入力されたトリガー信号d5に応じて、所定の時刻のグリッド高さを示す情報の最大値を、第2グリッド高さ値として、記憶部60に記憶させる。

【0024】

状況推定部50は、ID毎に、記憶部60に記憶させた第1クラスタ高さ値と第2クラスタ高さ値とを読み出し、読み出した第1クラスタ高さ値と第2クラスタ高さ値とを比較することで、ID毎に手を挙げているか否かを判別する。

状況推定部50は、例えば、第1クラスタ高さ値から第2クラスタ高さ値を減算した値が所定の値以上の時、手を挙げていると判別する。または、状況推定部50は、例えば、第1クラスタ高さ値を第2クラスタ高さ値で除算した値が所定の値以上の時、手を挙げていると判別する。

【0025】

記憶部60には、第1クラスタ高さ値、第2クラスタ高さ値、第1グリッド高さ値、第2グリッド高さ値が記憶される。

【0026】

図2は、本実施形態に係る音声信号の出力するタイミングと挙手を検出するタイミングを説明する図である。図2において、横軸は時間である。なお、図2に示す例では、ロボット1が提示する話題がAとBの2つの場合を示している。

まず、制御部102が、音声信号を出力するタイミングについて説明する。

時刻 t_0 において、人は、ロボット 1 から出力される音声聞いている。この時点で人は挙手を行っていない。

時刻 t_1 のとき、制御部 102 は、記憶部 101 に記憶されているシナリオ情報に基づいて、音声信号「A について聞きたい人は、手を挙げてください」を、音声出力部 103 からの出力を開始する。

【0027】

また、時刻 t_{11} のとき、制御部 102 は、記憶部 101 に記憶されているシナリオ情報に基づいて、音声信号「B について聞きたい人は、手を挙げてください」を、音声出力部 103 からの出力を開始する。期間 p_1 は、時刻 t_1 または t_{11} に出力を開始した音声信号が出力されている期間であり、時刻 $t_1 \sim t_2$ 、または時刻 $t_{11} \sim t_{12}$ の期間である。

10

【0028】

時刻 t_1 から期間 p_2 経過後の時刻 t_4 のとき、制御部 102 は、記憶部 101 に記憶されているシナリオ情報に基づいて、音声信号「はい、わかりました」を、音声出力部 103 から出力する。また、時刻 t_{11} から期間 p_2 経過後の時刻 t_{14} のとき、制御部 102 は、記憶部 101 に記憶されているシナリオ情報に基づいて、音声信号「はい、わかりました」を、音声出力部 103 から出力する。期間 p_2 は、シナリオ情報に設定されている時刻 t_1 または t_{11} からの所定時間であり、例えば 10 秒間である。また、期間 p_5 は、時刻 t_4 または時刻 t_{14} に出力を開始した音声信号が出力されている期間であり、時刻 $t_4 \sim t_6$ 、または時刻 $t_{14} \sim t_{16}$ の期間である。時刻 t_4 及び t_{14} は、人

20

【0029】

時刻 t_3 は、時刻 t_4 から期間 p_4 （第 1 期間）だけ前の時刻である。また、時刻 t_{13} は、時刻 t_{14} から期間 p_4 だけ前の時刻である。期間 p_4 は、例えば、（所定の時間の） $1/2$ である。例えば、所定の時間は 4 秒間であり、期間 p_4 は 2 秒間である。この期間 p_4 の間、人統合処理部 30 は、ID 毎にクラスタ高さを検出する。また、この期間 p_4 の間、グリッド統合処理部 40 は、グリッド毎にグリッド高さを検出し、グリッド高さが検出された全てのグリッドにおいて、グリッド高さの平均値を算出する。

【0030】

時刻 t_5 は、時刻 t_4 から期間 p_6 （第 1 期間）だけ後の時刻である。また、時刻 t_{15} は、時刻 t_{14} から期間 p_6 だけ後の時刻である。期間 p_6 は、例えば、（所定の時間の） $1/2$ である。例えば、所定の時間は 4 秒間であり、期間 p_6 は 2 秒間である。時刻 t_5 及び t_{15} のとき、人統合処理部 30 は、ID 毎にクラスタ高さを検出する。また、時刻 t_5 及び t_{15} のとき、グリッド統合処理部 40 は、グリッド毎にグリッド高さを検出し、グリッド高さが検出された全てのグリッドにおいて、グリッド高さの平均値を算出する。

30

【0031】

本実施形態では、図 2 に示したように、音声信号が出力されるタイミングである時刻 t_4 から期間 p_4 （第 1 期間）だけ前の間のデータを用いて検出された ID 毎にクラスタ高さと、時刻 t_4 から期間 p_6 （第 1 期間）だけ後の時刻（第 1 時刻）のデータを用いて検出されたクラスタ高さとを状況推定部 50 が比較することで、挙手されている人数を推定する。同様に、時刻 t_{14} から期間 p_4 だけ前の間のデータを用いて検出された ID 毎にクラスタ高さと、時刻 t_{14} から期間 p_6 だけ後の時刻のデータを用いて検出された高さとを状況推定部 50 が比較することで、挙手されている人数を推定する。そして、時刻 t_{16} 後、制御部 102 は、質問 A 及び B に対して挙手されている人数を比較することで多数決を行う。

40

【0032】

または、時刻 t_4 から期間 p_4 だけ前の間のデータを用いて算出されたグリッド高さの平均値と、時刻 t_4 から期間 p_6 だけ後の時刻のデータを用いて算出されたグリッド高さの平均値とを状況推定部 50 が比較することで、グリッド高さを推定する。同様に、時刻

50

t 1 4から期間 p 4だけ前の間のデータを用いて算出されたグリッド高さの平均値と、時刻 t 1 4から期間 p 6だけ後の時刻のデータを用いて算出されたグリッド高さとを状況推定部 5 0が比較することで、グリッド高さを推定する。そして、時刻 t 1 6後、制御部 1 0 2は、質問 A 及び B に対するグリッド高さを比較することで多数決を行う。

【0033】

次に、ロボット 1 による挙手検出の処理手順を説明する。

図 3 は、本実施形態に係るロボット 1 による挙手検出の処理手順のフローチャートである。

(ステップ S 1) 制御部 1 0 2 は、記憶部 1 0 1 に記憶されているシナリオ情報を読み出す。制御部 1 0 2 は、処理をステップ S 2 に進める。

10

(ステップ S 2) 制御部 1 0 2 は、読み出したシナリオ情報に基づいて、音声信号を音声出力部 1 0 3 から出力させる。制御部 1 0 2 は、処理をステップ S 3 に進める。

【0034】

(ステップ S 3) センシング処理部 2 0 は、L R F センサ 1 0 から入力された検出結果の中から、z 軸成分の値が所定の値以上の測距点を抽出する。センシング処理部 2 0 は、処理をステップ S 4 に進める。

【0035】

(ステップ S 4) センシング処理部 2 0 は、ステップ S 3 で抽出した測距点の情報を用いて、人クラスタデータを生成する。次に、センシング処理部 2 0 は、ステップ S 3 で抽出した測距点の情報を用いて、グリッド毎にグリッドデータを生成する。なお、生成する順番は、グリッドデータが先であってもよく、人クラスタデータの生成とグリッドデータの生成とが平行して行われてもよい。センシング処理部 2 0 は、処理をステップ S 5 に進める。

20

【0036】

(ステップ S 5) 人統合処理部 3 0 は、ステップ S 4 で生成された人クラスタデータを用いて統合処理を行う。

次に、グリッド統合処理部 4 0 は、ステップ S 4 で生成されたグリッドデータを用いて統合処理を行う。なお、処理する順番は、人統合処理部 3 0 及びグリッド統合処理部 4 0 のどちらかが先であってもよく、平行して行うようにしてもよい。人統合処理部 3 0 及びグリッド統合処理部 4 0 は処理をステップ S 6 に進める。

30

【0037】

(ステップ S 6) 状況推定部 5 0 は、期間 (p 2 - p 4) (所定の期間 (第 1 期間)) が経過したか否か、期間 (p 2 + p 4) が経過したか否かを判別する。状況推定部 5 0 は、期間 (p 2 - p 4) が経過していないと判別した場合 (ステップ S 6 ; 経過していない)、ステップ S 3 に戻る。状況推定部 5 0 は、期間 (p 2 - p 4) が経過したと判別した場合 (ステップ S 6 ; 期間 (p 2 - p 4) が経過)、ステップ S 7 に進み、期間 (p 2 + p 4) (所定の時刻) が経過したと判別した場合 (ステップ S 6 ; 期間 (p 2 + p 4) が経過)、ステップ S 9 に進む。

【0038】

(ステップ S 7) 状況推定部 5 0 は、期間 p 2 が経過したか否かを判別する。状況推定部 5 0 は、期間 p 2 が経過していないと判別した場合 (ステップ S 7 ; NO)、ステップ S 8 に進み、期間 p 2 が経過したと判別した場合 (ステップ S 7 ; YES)、ステップ S 3 に戻る。

40

(ステップ S 8) 状況推定部 5 0 は、第 1 処理を行う。状況推定部 5 0 は、処理終了後、ステップ S 3 に戻る。

(ステップ S 9) 状況推定部 5 0 は、第 2 処理を行う。状況推定部 5 0 は、処理終了後、ステップ S 1 0 に進む。

【0039】

(ステップ S 1 0) 状況推定部 5 0 は、ステップ S 8 の第 1 処理で検出した結果と、ステップ S 9 の第 2 処理で検出した結果とを比較する。

50

(ステップS11) 状況推定部50は、ステップS10で比較された結果に基づいて、手を挙げた人数を推定する。次に、状況推定部50は、ステップS10で比較された結果に基づいて、グリッド高さの変化量を推定する。

動作推定装置3は、シナリオ情報に基づいて、所定の質問を所定回数提示するまで、ステップS2～S11を繰り返す。

【0040】

(ステップS12) 状況推定部50は、シナリオ情報に基づいて、所定の質問が所定回数提示された後、ステップS11で推定された質問毎の推定結果に基づいて、多数決判定を行う。具体的には、状況推定部50は、設定されている条件に応じて、人統合処理部30またはグリッド統合処理部40が処理した結果を用いるかを選択する。状況推定部50は、ステップS11で推定された質問毎の手を挙げた人数を比較し、最も人数が多い質問を選択する。または、状況推定部50は、ステップS11で推定された質問毎のグリッド高さの変化量を比較し、最も変化量が多い質問を選択する。

以上で、挙手検出の処理を終了する。

【0041】

<人クラスタデータの生成>

次に、ステップS4(図3)で行われる人クラスタデータの生成について説明する。

図4は、本実施形態に係るLRFセンサ10による測定とセンシング処理部20と人統合処理部30とによる処理を説明する図である。図4では、紙面に向かって左右方向をx軸方向、奥行き方向をy軸方向、高さ方向をz軸方向とする。

図4に示す例では、検出対象の空間の天井に2つのLRFセンサ10-1とLRFセンサ10-2が、距離L1の間隔で設置されている。符号111が示す領域は、LRFセンサ10-1の検出可能な領域(以下、検出領域という)を示し、符号112が示す領域は、LRFセンサ10-2の検出領域を示している。また、図4に示した例では、所定の高さ以上の測距点を示している。

【0042】

また、図4に示す例では、LRFセンサ10-1の検出領域111に人hu1がいて、LRFセンサ10-2の検出領域112に人hu1と人hu2がいる。

測距点121～126は、人hu1に対する測距点であり、測距点131～133は、人hu2に対する測距点である。測距点は、例えば、人の頭部、耳、服の襟、肩などに対応する点である。これらの測距点には、x、y、zの各成分が含まれている。すなわち、測距点には、xy平面上における位置を示す情報と、高さを示す情報とが含まれている。

そして、図4に示す例では、測距点122～126がLRFセンサ10-1によって測定された測距点であり、測距点121～126及び131～133がLRFセンサ10-2によって測定された測距点である。

また、図4において、投影点141～146は、測距点131～136をxy平面に投影した点であり、投影点151～153は、測距点131～133をxy平面に投影した点である。

【0043】

次に、クラスタリング処理、人クラスタデータの生成の処理手順を説明する。

図5は、本実施形態に係るクラスタリング処理、人クラスタデータの生成の処理手順のフローチャートである。

(ステップS101) センシング処理部20は、対応するLRFセンサ10から入力された測距点の情報のうち、z軸成分の値が所定の値以上の測距点を、各測距点の情報に含まれるx軸成分とy軸成分を用いて、xy平面に投影する。センシング処理部20は、処理をステップS102に進める。

【0044】

(ステップS102) センシング処理部20は、xy平面に投影した投影点を分類(クラスタ分け)する。センシング処理部20は、処理をステップS103に進める。

【0045】

(ステップS103) センシング処理部20は、クラスタ分けしたクラスタ毎に、クラスタ内に含まれる投影点の重心位置を算出し、算出した重心位置をそのクラスタの位置を示す情報とする。センシング処理部20は、処理をステップS104に進める。

(ステップS104) センシング処理部20は、クラスタ分けしたクラスタ毎に、クラスタ内に含まれる投影点に対応する測距点のうち、z軸成分の最大値を抽出し、抽出した最大値をそのクラスタの高さを示す情報とする。センシング処理部20は、処理をステップS105に進める。

(ステップS105) センシング処理部20は、クラスタ毎の位置を示す情報と、高さを示す情報とを関連づけて、クラスタ毎の人クラスタデータを生成する。

【0046】

次に、図4を用いて、クラスタリング処理、人クラスタデータの生成の具体例を、図5を参照して説明する。

センシング処理部20-1は、入力された測距点122~126をxy平面に投影し、投影点142~146の情報を生成する。センシング処理部20-2は、入力された測距点121~126及び131~133をxy平面に投影し、投影点141~146及び151~153の情報を生成する(ステップS101)。

【0047】

次に、センシング処理部20-1は、投影点142~146を1つのクラスタに分類する。また、センシング処理部20-2は、投影点141~146を1つのクラスタに分類し、投影点151~153を1つのクラスタに分類する(ステップS102)。

センシング処理部20-1は、クラスタリングされた投影点142~146の重心位置を計算する。センシング処理部20-1は、クラスタリングされた投影点141~146の重心位置、投影点151~153の重心位置を計算する(ステップS103)。

【0048】

次に、センシング処理部20-1は、分類したクラスタに含まれる投影点142~146に対応する測距点122~126のうち、z軸成分の最大値である測距点122のz軸成分を抽出する。センシング処理部20-2は、測距点121~126のうち、z軸成分の最大値である測距点121のz軸成分を抽出し、測距点131~133のうち、z軸成分の最大値である測距点132のz軸成分を抽出する(ステップS104)。

【0049】

次に、センシング処理部20-1は、投影点142~146の重心位置である位置を示す情報と、測距点122の高さを示す情報とを関連づけて人クラスタデータd1(11)を生成する。センシング処理部20-2は、投影点141~146の重心位置である位置を示す情報と、測距点121の高さを示す情報とを関連づけて人クラスタデータd1(21)を生成し、投影点151~153の重心位置である位置を示す情報と、測距点132の高さを示す情報とを関連づけて人クラスタデータd1(22)を生成する(ステップS105)。

【0050】

<人統合処理部30によって行われる第1統合処理>

次に、人統合処理部30によって行われるステップS5(図3)の第1統合処理について説明する。

図6は、本実施形態に係る人統合処理部30によって行われる第1統合処理手順のフローチャートである。

【0051】

(ステップS201) 人統合処理部30は、センシング処理部20-1~20-nから入力される人クラスタデータを統合する。人統合処理部30は、処理をステップS202に進める。

(ステップS202) 人統合処理部30は、統合された人クラスタデータを用いて、クラスタ毎にIDを割り当てる。人統合処理部30は、処理をステップS203に進める。

【0052】

10

20

30

40

50

(ステップS203) 人統合処理部30は、ID毎に、IDが割り当てられたクラスタの位置を示す情報と、IDが割り当てられたクラスタの高さを示す情報と、割り振られたIDとを関連づけて人データを生成する。人統合処理部30は、処理をステップS204に進める。

(ステップS204) 人統合処理部30は、生成した人データを用いて、ID毎にIDが割り振られた人の位置、及び高さに対してトラッキングを行う。なお、人統合処理部30は、トラッキングができなくなった場合、トラッキングできなくなったIDに対応する人データを削除する。

【0053】

次に、人統合処理部30によって行われる第1統合処理の具体例を、図4、図6を参照して説明する。

人統合処理部30は、センシング処理部20-1から入力される人クラスタデータd1(11)と、センシング処理部20-2から入力される人クラスタデータd1(21)を、位置を示す情報が所定の範囲内にあると判別して統合する(ステップS201)。次に、人統合処理部30は、人クラスタデータd1(11)とd1(21)が統合されたクラスタデータに対して、例えばID1を割り当てる(ステップS202)。また、人統合処理部30は、人クラスタデータd1(11)とd1(21)に含まれる高さを示す情報のうち、最大値として測距点121のz軸成分の値を抽出する。

また、人統合処理部30は、センシング処理部20-2から入力される人クラスタデータd1(22)と、位置を示す情報が所定の範囲内にある人クラスタデータが他にないと判別する。この場合、人統合処理部30は、人クラスタデータd1(22)に他の人クラスタデータを統合しない(ステップS201)。次に、人統合処理部30は、人クラスタデータd1(22)にID2を割り当てる(ステップS202)。また、人統合処理部30は、人クラスタデータd1(22)に含まれる高さを示す情報を抽出する。

【0054】

次に、人統合処理部30は、ID1の人データ、ID2の人データを生成する(ステップS203)。

次に、人統合処理部30は、各IDをトラッキングする。トラッキングは、例えば時刻t1(図2)のときに割り当てられた各IDに含まれる統合後の人クラスタデータが、サンプリング時間Δt毎に、所定の範囲内にある場合、同じIDであると判別する。所定の範囲とは、例えば、サンプリング時間Δtに人が歩いて移動したときに、人クラスタデータが移動可能な範囲等である。さらに、人統合処理部30は、同じIDであると判別された人データに含まれる高さを示す情報をトラッキングする(ステップS204)。

【0055】

図7は、本実施形態に係るトラッキングを行った結果の例を示す図である。なお、図7に示す例は、ID4~ID7の4人が検出された例である。また、図7に示したクラスタ高さとは、人統合処理部30によって抽出されたID毎の最も高さが高い測距点のz軸成分の値である。

【0056】

図7において、横軸は時間、縦軸はクラスタの高さである。また、符号g101の曲線は、人統合処理部30によってID4が割り振られた人のクラスタ高さの時間変化、符号g102の曲線は、ID5が割り振られた人のクラスタ高さの時間変化、符号g103の曲線は、ID6が割り振られた人のクラスタ高さの時間変化、符号g104の曲線は、ID7が割り振られた人のクラスタ高さの時間変化を表す曲線である。

【0057】

時刻t21は、制御部102が、記憶部101に記憶されているシナリオ情報に基づいて、音声信号「Aについて聞きたい人は、手を挙げてください」を、音声出力部103から出力した時刻である。

【0058】

曲線g101~g104それぞれは、ID4~ID7が割り振られたクラスタ高さの変

10

20

30

40

50

化である。曲線 g 1 0 1 が示すように、I D 4 のクラスタ高さは、約 1 6 2 0 [mm] であり、時刻 t 2 2 においてトラッキングが失敗している。なお、トラッキングが失敗するとは、トラッキングを行っていた人クラスタデータが消失した場合などに発生する。人クラスタデータが消失は、例えば、人が検出領域（所定の範囲）外に移動した場合などに発生する。

【0059】

曲線 g 1 0 2 が示すように、I D 5 のクラスタ高さは、時刻 t 2 1 ~ t 2 5 の期間、約 1 5 5 0 [mm] であり、時刻 t 2 5 においてクラスタ高さが約 1 8 0 0 [mm] に変化している。

また、曲線 g 1 0 3 に示すように、I D 6 のクラスタ高さは、時刻 t 2 1 ~ t 2 3 の期間、約 1 5 8 0 [mm] であり、時刻 t 2 3 においてクラスタ高さが約 1 9 0 0 [mm] に変化し、時刻 t 2 4 においてトラッキングが失敗している。

曲線 g 1 0 4 に示すように、I D 7 のクラスタ高さは、時刻 t 2 1 ~ t 2 6 の期間、約 1 4 8 0 [mm] のまま、変化していない。

【0060】

＜グリッドデータの生成＞

次に、ステップ S 4（図 3）で行われるグリッドデータの生成について説明する。

図 8 は、本実施形態に係る L R F センサ 1 0 による測定とセンシング処理部 2 0 とグリッド統合処理部 4 0 による処理を説明する図である。図 8 では、図 4 と同様に、紙面に向かって左右方向を x 軸方向、奥行き方向を y 軸方向、高さ方向を z 軸方向とする。また、検出対象の空間には、予めグリッド a（1、1）～a（3、5）を設定しておく。グリッドは、縦 L 2、横 L 2 であり、L 2 は例えば 1 0 c m である。なお、図 4 と同様のものについては、同じ符号を用いて説明を省略する。また、図 8 に示す例では、図 4 と同様に、所定の高さ以上の測距点を示している。

【0061】

グリッド a（1、1）～a（3、5）の各頂点の座標値は、予めセンシング処理部 2 0 に記憶されている。例えば、グリッド a（1、1）の x y 座標値それぞれは、左下の頂点 c 1（0、0）、左上の頂点 c 2 が（1 0、0）、右上の頂点 c 3 が（1 0、1 0）、右下の頂点 c 4 が（0、1 0）である。センシング処理部 2 0 は、この範囲にある投影点を、このグリッド a（1、1）の投影点として分類する。

【0062】

図 9 は、本実施形態に係るグリッドデータの生成の処理手順のフローチャートである。（ステップ S 3 0 1）センシング処理部 2 0 は、ステップ S 1 0 1（図 5）後、投影した投影点を、グリッド毎に分類する。センシング処理部 2 0 は、処理をステップ S 3 0 2 に進める。

（ステップ S 3 0 2）センシング処理部 2 0 は、グリッド毎に、グリッドに含まれる投影点に対応する測距点の z 軸成分の最大値を抽出し、抽出した z 軸成分を、そのグリッドの高さを示す情報とする。センシング処理部 2 0 は、処理をステップ S 3 0 3 に進める。

（ステップ S 3 0 3）センシング処理部 2 0 は、グリッド毎に、グリッドの位置を示す情報と、抽出したグリッドの高さを示す情報と関連づけてグリッドデータを生成する。

【0063】

図 8 に示した例では、センシング処理部 2 0 - 1 は、投影点 1 4 2 ~ 1 4 6 をグリッド a（1、4）に分類し、投影点 1 4 2 に対応する測距点 1 2 2 の z 軸成分を、高さを示す情報として抽出する。そして、センシング処理部 2 0 - 1 は、グリッド a（1、4）の位置を示す情報と、測距点 1 2 2 の z 軸成分である高さを示す情報とを関連づけてグリッドデータ d 2（1 1）を生成する。

【0064】

また、センシング処理部 2 0 - 2 は、投影点 1 4 1 ~ 1 4 6 をグリッド a（1、4）に分類し、投影点 1 4 1 に対応する測距点 1 2 1 の z 軸成分を、高さを示す情報として抽出する。さらに、センシング処理部 2 0 - 2 は、投影点 1 5 1 ~ 1 5 3 をグリッド a（1、

5) に分類し、投影点 1 5 2 に対応する測距点 1 3 2 の z 軸成分を、高さを示す情報として抽出する。そして、センシング処理部 2 0 - 2 は、グリッド a (1、4) の位置を示す情報と、測距点 1 2 1 の z 軸成分である高さを示す情報とを関連づけてグリッドデータ d 2 (2 1) を生成し、グリッド a (1、5) の位置を示す情報と、測距点 1 3 2 の z 軸成分である高さを示す情報とを関連づけてグリッドデータ d 2 (2 2) を生成する。

なお、センシング処理部 2 0 は、高さを示す情報が抽出されたグリッドのみグリッドデータを作成する。

【0065】

<グリッド統合処理部 4 0 によって行われる統合処理>

次に、ステップ S 5 (図 3) で行われる統合処理について説明する。

10

図 1 0 は、本実施形態に係るグリッド統合処理部 4 0 によって行われる統合処理手順のフローチャートである。

(ステップ S 4 0 1) グリッド統合処理部 4 0 は、グリッドデータ d 2 に含まれる位置を示す情報が一致する、センシング処理部 2 0 - 1 ~ 2 0 - n から入力されるグリッドデータ d 2 を統合する。グリッド統合処理部 4 0 は、処理をステップ S 4 0 2 に進める。

【0066】

(ステップ S 4 0 2) グリッド統合処理部 4 0 は、統合後、グリッドの高さを示す情報が複数ある場合、最大値を抽出し、抽出した高さを示す情報を、そのグリッドの高さ値に決定する。グリッド統合処理部 4 0 は、処理をステップ S 4 0 3 に進める。

20

(ステップ S 4 0 3) グリッド統合処理部 4 0 は、高さを示す情報を有している統合後のグリッド数をカウントし、グリッド数とする。次に、グリッド統合処理部 4 0 は、高さを示す情報を有する全てのグリッドの高さ値を加算し、グリッド数で除算することで、グリッドの高さの平均値を算出する。

【0067】

例えば、グリッド統合処理部 4 0 は、高さを示す情報を有するグリッドの高さである z 軸方向の値の平均値 Z_{ave} を、次式 (1) を用いて算出する。

【0068】

平均値 $Z_{ave} = \Sigma$ (各グリッドの z 軸方向の高さ) / (高さを示す情報が付加されているグリッド数) … (1)

【0069】

30

図 8 に示す例において、平均値 Z_{ave} は、 { グリッド a (1、4) の高さ (測距点 1 2 1 の z 軸方向の成分) + グリッド a (1、5) の高さ (測距点 1 3 2 の z 軸方向の成分) } / 2 である。

次に、グリッド統合処理部 4 0 は、時刻 t 5 (図 2) のときの高さを示す情報を有するグリッドの高さの平均値 Z_{ave} を、式 (1) を用いて算出する。

【0070】

例えば、図 8 に示した人 h u 1 の身長が 1 7 3 [c m]、人 h u 2 の身長が 1 6 8 [c m] であり、人 h u 1 が手を挙げたときの高さが 2 1 0 [c m]、人 h u 2 が手を挙げたときの高さが 1 9 5 [c m] であるとする。

期間 p 4 (図 2) のとき、人 h u 1 と人 h u 2 が 2 人とも手を挙げている場合、手を挙げている期間 p 4 におけるグリッドの高さの平均値 Z_{ave} は、2 0 2 . 5 [c m] (= (2 1 0 [c m] + 1 9 5 [c m]) / 2) である。この場合、期間 p 4 における最大値 Z_{max} は、平均値 Z_{ave} であるとする。このまた、手を下げている時刻 t 5 のときのグリッドの高さの平均値 Z_{ave} は、1 7 0 . 5 [c m] (= (1 7 3 [c m] + 1 6 8 [c m]) / 2) である。期間 p 4 の最大値 Z_{max} 2 0 2 . 5 [c m] と、時刻 t 5 のときの平均値 Z_{ave} 1 7 0 . 5 [c m] とを比較すると、差は 3 2 [c m] である。

40

【0071】

また、図 8 において、期間 p 4 のとき、人 h u 1 のみが手を挙げている場合、手を挙げている期間 p 4 におけるグリッドの高さの平均値 Z_{ave} は、1 8 9 [c m] (= (2 1 0 [c m] + 1 6 8 [c m]) / 2) である。同様に、期間 p 4 における最大値 Z_{max}

50

は、平均値 Z_{ave} であるとする。また、手を下げている時刻 t_5 のときのグリッドの高さの平均値 Z_{ave} は、 $170.5 [cm]$ ($= (173 [cm] + 168 [cm]) / 2$) である。期間 p_4 の最大値 $Z_{max} 189 [cm]$ と、時刻 t_5 のときの平均値 $Z_{ave} 170.5 [cm]$ とを比較すると、差は $18.5 [cm]$ である。

すなわち、期間 p_4 の最大値 Z_{max} と時刻 t_5 のときの平均値 Z_{ave} との差が大きい程、手を挙げている人数が多いことになる。

【0072】

例えば、ロボット 1 は、シナリオ情報に応じて、図 3 に示した処理を 3 回繰り返すことで、三択を検出領域にいる人たちに与えさせるとする。

状況推定部 50 は、1 回目の時刻 $t_{1(1)}$ において、音声信号「A について聞きたい人は、手を挙げてください」を、音声出力部 103 から出力させる。そして、1 回目の期間 $p_{4(1)}$ の最大値と時刻 $t_{5(1)}$ のときの平均値との差 Δg_1 を算出する。

次に、状況推定部 50 は、2 回目の時刻 $t_{1(2)}$ ($=$ 時刻 t_{11} 、図 2) において、音声信号「B について聞きたい人は、手を挙げてください」を、音声出力部 103 から出力させる。そして、2 回目の期間 $p_{4(2)}$ の平均値と時刻 $t_{5(2)}$ ($=$ 時刻 t_{15} 、図 2) のときの平均値との差 Δg_2 を算出する。

【0073】

次に、状況推定部 50 は、3 回目の時刻 $t_{1(3)}$ において、音声信号「C について聞きたい人は、手を挙げてください」を、音声出力部 103 から出力させる。そして、3 回目の期間 $p_{4(3)}$ の平均値と時刻 $t_{5(3)}$ のときの平均値との差 Δg_3 を算出する。

状況推定部 50 は、このように算出した差 $\Delta g_1 \sim \Delta g_3$ を比較することで、質問の A $\sim$ C のうち、どの質問に対する挙手が最も多かったかを推定する。

【0074】

以上のように、本実施形態に係る動作推定装置 3 において、人の反応は、人が挙手を行っている反応か挙手を行っていない反応かであり、状況推定部 (例えば状況推定部 50) は、人の反応を複数回推定して、複数回推定した結果に基づいて、多数決の判定を行う。

この構成によって、本実施形態の動作推定装置 3 は、多数決の判定を行うことができる。

【0075】

なお、上述した例では、センシング処理部 20 が、高さ情報が抽出できたグリッドのみグリッドデータを生成する例を説明したが、これに限られない。センシング処理部 20 は、高さを示す情報がないグリッドについては高さを示す情報として、例えば 0 を関連づけてグリッドデータを生成するようにしてもよい。この場合、グリッド統合処理部 40 は、ステップ S403 において、高さ情報が 0 より大きい値のグリッドの数をグリッド数とし、また、高さを示す情報が 0 より大きい値をグリッドの高さ値とするようにしてもよい。

【0076】

図 8 に示す例において、グリッド統合処理部 40 は、センシング処理部 20-1 から入力されたグリッドデータ $d_{2(11)}$ と、センシング処理部 20-2 から入力されたグリッドデータ $d_{2(21)}$ とにそれぞれ含まれる位置を示す情報が一致するため、1 つのグリッドのグリッドデータとして統合する。グリッド統合処理部 40 は、複数のセンシング処理部 20 から入力される高さを表す情報が一致しない場合、複数のセンシング処理部 20 から入力される高さを表す情報の z 軸成分同士を比較し、値が大きい z 軸成分を最大値として抽出する。

また、グリッド統合処理部 40 は、センシング処理部 20-2 から入力されたグリッドデータ $d_{2(22)}$ に含まれる位置を示す情報と一致する他のグリッドデータがないため、他のグリッドデータとの統合を行わない。そして、グリッド統合処理部 40 は、高さを表す情報である測距点 132 の z 軸成分を最大値として抽出する。

そして、グリッド統合処理部 40 は、測距点 121 と 132 との z 軸成分の加算値を、グリッド数の 2 で除算してグリッド高さを算出する。

【0077】

<第1処理及び第2処理を行うタイミング>

次に、第1処理及び第2処理を行うタイミングについて説明する。

図11は、本実施形態に係る第1処理及び第2処理を行うタイミングを説明する図である。

図11において、図7と同様に、横軸は時間、縦軸はクラスタの高さである。また、曲線g111～g114それぞれは、人統合処理部30によってID4～ID7が割り振られた人のクラスタ高さの時間変化である。図11に示した例では、ID4～ID7が割り振られた4人をトラッキングできたときの結果である。なお、制御部102は、時刻t31において時刻t1（図2）と同様の音声信号「Aについて聞きたい人は、手を挙げてください」の出力を開始させ、時刻t35において時刻t4（図2）と同様の音声信号「はい、わかりました」の出力を開始させる。

10

【0078】

時刻t32において、曲線g113が示すようにID6が割り振られた人のクラスタ高さが増え、時刻t33において、曲線g112が示すようにID5が割り振られた人のクラスタ高さが増え、時刻t34において、曲線g111が示すようにID4が割り振られた人のクラスタ高さが増える。曲線g114が示すようにID7が割り振られた人のクラスタ高さは、所定の閾値未満の変化である。

【0079】

時刻t36において、符号g113のようにID6が割り振られた人のクラスタ高さが減り、時刻t37において、符号g111のようにID4が割り振られた人のクラスタ高さが減り、時刻t38において、符号g112のようにID5が割り振られた人のクラスタ高さが減る。

20

【0080】

図11が示すように、検出領域に複数の人がいて、その中の何人かが手を挙げる場合、手を挙げるタイミングには、ばらつきが多い。例えば、周りにいる人が手を挙げたことを確認してから、手を挙げる場合もある。例えば、図11に示した時刻t32と時刻t34との時間差は、例えば3秒である。さらに、複数の人がいつ手を挙げ始めるかが不明である。このように、手を挙げたことを検出しようとする、図11において、例えば時刻t31～時刻t35の間、検出を行う必要がある。

【0081】

30

一方、図11に示した時刻t36と時刻t38との時間差は、例えば1秒である。すなわち、図11に示すように、音声信号「はい、わかりました」が出力された時刻t35から所定の時間が経過すると、手を挙げていた人たちは、短い時間内に手を下げていることになる。

このため、本実施形態では、時刻t35において音声信号「はい、わかりました」が出力されたことをトリガー信号とし、手を挙げてしていると推定される、トリガー信号が入力される時刻より所定の時間前の間に第1処理を行う。そして、本実施形態では、トリガー信号が入力される時刻より所定の時間後に第2処理を行い、第1処理と第2処理との結果を比較することで、手を挙げている人数を検出する。所定の時間は、例えば2秒間である。

【0082】

40

<第1処理>

次に、ステップS8（図3）で行われる第1処理について説明する。

まず、状況推定部50が人データd3を用いて行う第1処理について、図2を用いて説明する。

状況推定部50は、期間（p2～p4）経過後～期間p2経過後まで、すなわち時刻t3～t4の間、ID毎に人データd3に含まれる高さを示す情報の中から最大値を検出する。最大値の検出は、例えば、時刻t-Δt（サンプリング時間）と時刻tとの高さを示す情報を比較し、値が大きい方を最大値として保持することで検出する。

状況推定部50は、時刻t3～t4の間に検出された最大値を、そのIDの第1クラスタ高さ値として、記憶部60に記憶させる。

50

【0083】

次に、状況推定部50がグリッド高さを示す情報d4を用いて行う第1処理について、図2を用いて説明する。

状況推定部50は、時刻t3～t4の間、グリッド高さを示す情報d4の中から最大値を検出する。状況推定部50は、時刻t3～t4の間に検出された最大値を、第1グリッド高さ値として、記憶部60に記憶させる。

【0084】

<第2処理>

次に、ステップS9（図3）で行われる第2処理について説明する。

まず、状況推定部50が人データd3を用いて行う第2処理について、図2を用いて説明する。

10

状況推定部50は、期間(p2+p4)経過後、すなわち時刻t5のとき、ID毎に人データに含まれる高さを示す情報を検出する。状況推定部50は、時刻t5のときに抽出された値を、そのIDの第2クラスタ高さ値として、記憶部60に記憶させる。

【0085】

次に、状況推定部50がグリッド高さを示す情報d4を用いて行う第2処理について、図2を用いて説明する。

状況推定部50は、時刻t5のときのグリッド高さを示す情報d4を、第2グリッド高さ値として、記憶部60に記憶させる。

20

【0086】

<比較処理>

次に、ステップS10及びS11（図3）で行われる比較処理及び推定処理について説明する。

まず、状況推定部50が人データd3に対して行う比較処理及び推定処理について説明する。

状況推定部50は、ID毎に、記憶部60に記憶させた第1クラスタ高さ値と第2クラスタ高さ値とを読み出し、読み出した第1クラスタ高さ値と第2クラスタ高さ値とを比較することで、ID毎に手を挙げているか否かを判別する。

状況推定部50は、例えば、第1クラスタ高さ値から第2クラスタ高さ値を減算した値が所定の値以上の時、手を挙げていると判別する。または、状況推定部50は、例えば、第1クラスタ高さ値を第2クラスタ高さ値で除算した値が所定の値以上の時、手を挙げていると判別する。

30

【0087】

次に、状況推定部50がグリッド高さを示す情報d4に対して行う比較処理及び推定処理について説明する。

状況推定部50は、記憶部60に記憶させた第1グリッド高さ値と第2グリッド高さ値とを読み出し、読み出した第1グリッド高さ値と第2グリッド高さ値とを比較することでグリッド高さの変化量を算出する。

状況推定部50は、例えば、第1グリッド高さ値から第2グリッド高さ値を減算した値をグリッド高さ変化量としてもよく、または、第1グリッド高さ値を第2グリッド高さ値で除算した値をグリッド高さ変化量としてもよい。

40

【0088】

次に、状況推定部50が、上述した人統合処理部30が検出した検出結果と、グリッド統合処理部40が検出した検出結果のいずれか一方を用いる例を説明する。

状況推定部50は、以下のような条件に応じて、人統合処理部30が検出した検出結果か、グリッド統合処理部40が検出した検出結果かを選択する。

（条件1）人統合処理部30から入力されたトラッキング結果に基づいて選択

トラッキングを失敗した回数または失敗した率が所定の値未満の場合、人統合処理部30が検出した検出結果を選択し、トラッキングを失敗した回数または失敗した率が所定の値以上の場合、グリッド統合処理部40が検出した検出結果を選択する。

50

(条件2) 人の密集度に基づいて選択

グリッド毎の密集度を検出し、検出したグリッド毎の密集度が所定の値未満の場合、人統合処理部30が検出した検出結果を選択し、検出したグリッド毎の密集度が所定の値以上の場合、グリッド統合処理部40が検出した検出結果を選択する。なお、本実施形態における密集度とは、例えば、所定の面積内にいる人数に基づく指標である。

【0089】

以上のように、本実施形態の動作推定装置3は、センサ（例えばLRFセンサ10）と対象物（人）との距離（例えば人の高さ方向の値）を取得する取得部（例えばセンシング処理部20）と、人に対して挙手の終了を促した時刻（例えば時刻t4）に基づき設定される第1時刻（例えば時刻t3またはt5）における取得部によって取得された距離に基づいて人の反応を推定する状況推定部（例えば状況推定部50）と、を備える。

10

【0090】

この構成によって、本実施形態の動作推定装置3では、人が手を下げたタイミングで、LRFセンサ10によって測定された測距点に基づいて、人の反応である手が下がった人数を、手を挙げた人数として推定し、または手を下げた人の割合（＝手を挙げた人の割合）である高さの変化量を推定する。これにより、本実施形態の動作推定装置3は、手が挙げられる期間に比べて短時間に精度良く人の反応を推定することができる。

【0091】

また、本実施形態の動作推定装置3において、状況推定部（例えば状況推定部50）は、人に対して挙手の終了を促した時刻に基づき設定される第1期間（例えば時刻t3からt4の期間p4、または時刻t4～t5の期間p6）における取得部によって取得された距離（例えば人の高さ方向の値）に基づいて人の反応を推定する。

20

また、本実施形態の動作推定装置3において、状況推定部（例えば状況推定部50）は、第1時刻（例えば時刻t3）と第1期間（例えば時刻t3からt4の期間p4）それぞれにおける対象物（例えば人）の高さ方向の距離を抽出し、第1時刻に抽出した高さ方向の距離と、第1期間に抽出した高さ方向の距離とを比較することで、人の反応を推定する。

【0092】

また、本実施形態の動作推定装置3において、取得部（例えばセンシング処理部20）によって取得された距離（例えば人の高さ方向の値）に基づいて人を検出し、検出した人毎に識別子（例えばID）を付与する人処理部（例えば人統合処理部30）、を備え、状況推定部（例えば状況推定部50）は、人処理部によって識別子が付与された人の距離を用いて、第1時刻（例えば時刻t3）と第1期間（例えば時刻t3からt4の期間p4）それぞれにおける識別子毎の高さ方向の距離を抽出し、第1時刻と第1期間それぞれにおける抽出した識別子毎の高さ方向の距離同士を比較することで、人の反応を推定する。

30

【0093】

この構成によって、本実施形態の動作推定装置3では、人が手を下げる前の期間に人が手を挙げているときの高さ方向の値と、人が手を下げたときの高さ方向の値とに基づいて、手を挙げている人数を検出することができる。

【0094】

40

また、本実施形態の動作推定装置3において、人の反応を推定する所定の範囲を1つ以上の領域に分割し、取得部（例えばセンシング処理部20）によって取得された距離に基づいて該領域毎の距離（例えば人の高さ方向の値）を求める領域距離処理部（例えばグリッド統合処理部40）、を備え、状況推定部（例えば状況推定部50）は、人処理部（例えば人統合処理部30）によって識別子（例えばID）が付与された人の距離と、領域距離処理部によって求められた距離のうち、すくなくとも1つの距離によって人の反応を推定する。

【0095】

この構成によって、本実施形態の動作推定装置3では、手を挙げている期間と、手を下げたときとの、領域における高さ方向の変化量を推定することができる。そして、動作推

50

定装置 3 では、推定された結果に基づいて、多数決によってどの質問への挙手が最も多いかを推定することができる。

【0096】

なお、本実施形態では、動作推定装置 3 の状況推定部 50 が、人統合処理部 30 の処理結果またはグリッド統合処理部 40 の処理結果のいずれかを選択して、選択した結果を処理部 2 に出力する例を説明したが、これに限られない。状況推定部 50 は、人統合処理部 30 の処理結果、及びグリッド統合処理部 40 の処理結果の両方を処理部 2 に出力し、処理部 2 がどちらの処理結果を用いるか選択するようにしてもよい。

【0097】

また、本実施形態では、手を下げることに基づくタイミングに基づいて、トリガー信号 d 5 を生成する例を説明したが、これに限られない。動作推定装置 3 は、例えば手を挙げることを促すタイミングに基づいて、手を下げる音声信号を出力するまでの時刻を算出し、算出した時刻に基づいて、手を挙げている期間である期間 p 4、また手を下げるときの時刻 t 5 を算出するようにしてもよく、このタイミングで制御部 102 が動作推定装置 3 にトリガー信号 d 5 を出力するようにしてもよい。

【0098】

また、本実施形態では、シナリオ情報、手を下げることを促すタイミングに基づいて、第 1 処理及び第 2 処理（図 3）を行う例を説明したが、これに限られない。

状況推定部 50 は、グリッド統合処理部 40 によって算出される高さを示す情報が抽出されたグリッドに高さの平均値をサンプリングタイム毎に取得するようにしてもよい。そして、状況推定部 50 は、取得した高さを示す情報が抽出されたグリッドに高さの平均値が、予め定められている値より小さくなったか否かを判別し、予め定められている値より小さくなったと判別したとき、検出領域にいる人達の手が下がったときと判別するようにしてもよい。そして、状況推定部 50 は、検出領域にいる人達の手が下がったときに基づいて、手が上がっている期間、及び手が下がっているときを決定するようにしてもよい。この場合、動作推定装置 3 は、人統合処理部 30 によって処理された結果と、グリッド統合処理部 40 によって処理された結果それぞれを、時刻を示す情報と関連づけて記憶部 60 に記憶させ、決定した期間及び時刻の情報を記憶部 60 から読み出して第 1 処理及び第 2 処理を行うようにしてもよい。

【0099】

以上のように、本実施形態の動作推定装置 3 において、第 1 時刻（例えば時刻 t 3 または t 5）または第 1 期間（例えば時刻 t 3 から t 4 の期間 p 4）は、領域距離処理部（例えばグリッド統合処理部 40）によって測定される距離（例えば高さ方向の値）の変化が所定の値より小さくなったときに対応して決定される。

【0100】

この構成によって、本実施形態の動作推定装置 3 では、シナリオ情報によらず、LRF センサ 10 によって測定された測距点がグリッド統合処理部 40 によって処理された結果にもとづいて、手を挙げている期間および手を下げた期間を決定することができる。この結果、本実施形態では、トリガー信号 d 5 を用いなくても、人が手を挙げている人数の推定、または人の手の高さの変化量を推定することができる。

【0101】

[第 2 実施形態]

図 12 は、本実施形態に係るロボット 1 A の概略のブロック図である。

図 12 に示すように、ロボット 1 A は、処理部 2、及び動作推定装置 3 A を含んで構成される。ロボット 1 A は、n（n は 1 以上の整数）個の LRF センサ 10-1 ~ 10-n に無線または有線で接続される。なお、ロボット 1 と同じ機能を有する機能部については、同じ符号を用いて説明を省略する。また、LRF センサ 10-1 ~ 10-n のうちいずれか 1 つを特定しない場合は、LRF センサ 10 という。なお、イベント会場等で用いられるガイドシステム等が動作推定装置 3 A を備えるようにしてもよい。

【0102】

動作推定装置3Aは、n個のセンシング処理部20A-1~20A-n、人統合処理部30、グリッド統合処理部40、状況推定部50A、及び記憶部60Aを備える。センシング処理部20A-1~20A-nのうちいずれか1つを特定しない場合は、センシング処理部20Aという。また、ロボット1と同様に、LRFセンサ10-1~10-nそれぞれは、対応するセンシング処理部20A-1~20A-nに無線または有線を介して接続されている。また、動作推定装置3Aは、LRFセンサ10を含んで構成されていてもよい。

【0103】

センシング処理部20Aは、LRFセンサ10から入力された測距点の情報d7を、状況推定部50Aに出力する。

10

センシング処理部20Aは、LRFセンサ10から入力された検出結果を用いて、例えば所定の周期毎に測距点を分類する（クラスタリング）。センシング処理部20Aは、分類したクラスタ毎の重心位置を算出し、算出した重心位置を、位置を示す情報とする。センシング処理部20Aは、分類したクラスタ毎に高さの最大値を抽出し、抽出した最大値を、高さを示す情報とする。そして、センシング処理部20Aは、クラスタ毎に、位置を示す情報と高さを示す情報とを関連づけて人クラスタデータd1を生成し、生成した人クラスタデータd1を人統合処理部30に出力する。

【0104】

また、センシング処理部20Aは、LRFセンサ10から入力された検出結果を用いて、例えば所定の周期毎に測距点を、予め定められているグリッドに分類する。センシング処理部20Aは、分類したグリッド毎に高さの最大値を抽出し、抽出した最大値を、高さを示す情報とする。そして、センシング処理部20Aは、グリッド毎に、グリッドの位置を示す情報と高さを示す情報とを関連づけてグリッドデータd2を生成し、生成したグリッドデータd2をグリッド統合処理部40に出力する。

20

【0105】

状況推定部50Aは、制御部102から入力されるトリガー信号d5に基づいて、センシング処理部20Aから入力される測距点の情報d7を統合して、手が下げられている期間の平均の第1マップ(Az)と、手が挙げられているときの第2マップ(B)とを生成する。なお、手が下げられている期間とは、手を挙げることを促す音声信号が出力されるまでの期間、例えば、図2において、時刻t0~t1の間である。

30

【0106】

また、状況推定部50Aは、生成した平均の第1マップ(Az)に記憶部60Aに記憶されている体格データを乗じて、人の手が存在できるz値の最大値の第3マップ(AzMax)と、z値の最小値の第4マップ(AzMin)とを生成する。状況推定部50Aは、生成した第2マップと、第3マップ及び第4マップとを比較することで、手が挙げられていることによると推定されるxy平面における第5マップ(C)を生成する。状況推定部50Aは、生成した第5マップに対してクラスタの分類を行って手を挙げている人数を推定する。

【0107】

状況推定部50Aは、状況推定部50と同様に、人統合処理部30によって処理された結果を用いて、手を挙げている人数を推定する。さらに、状況推定部50Aは、状況推定部50と同様に、グリッド統合処理部40によって処理された結果を用いて、グリッド高さの変化量を推定する。

40

さらに、状況推定部50Aは、予め定められている条件、または人統合処理部30とグリッド統合処理部40とセンシング処理部20Aとのうち少なくとも1つから入力された結果に基づいて、手を挙げている人数、または手を挙げていることによるグリッド高さの変化量を推定し、推定した結果を制御部102に出力する。

【0108】

記憶部60Aには、第1クラスタ高さ値、第2クラスタ高さ値、第1グリッド高さ値、第2グリッド高さ値に加えて、人の体格データが記憶される。人の体格データとは、身長

50

と手首の橈骨茎突点との比、肩峰高と手首の橈骨茎突点との比である。なお、これらの人の体格データは、国や地域ごとに記憶されていてもよい。

【0109】

ここで、本実施形態における挙手判定における仮定と根拠とを説明する。

(仮定1) 人の身長と腕の長さとは、統計的に相関があり、また、年齢、性別を問わずにはば一定であることが知られている。

(仮定2) 一般的に、人は手をまっすぐ上に向けて上げる。

(仮定3) 人が密集している場合、人は、近接する人の挙げた手と接触を避けるように、手を挙げることが多い。すなわち、人は、暗黙的に手のパーソナル空間のようなものを設定している。

(仮定4) 人の体幹の厚さ(腹厚)以上には他人には近づけない傾向がある。

【0110】

図13は、本実施形態に係る人の部位の高さを説明する図である。図13において、肩峰高 H_s は、人が直立した時の床又は地面から肩峰付近までの高さである。身長 H_h は、人が直立した時の床又は地面から頭頂までの高さである。手首 H_w は、人が直立した時に、手を挙げた時の床又は地面から手首の橈骨茎突点までの高さである。

上述した仮定1のように、身長 H_h に対する手首 H_w の比率と、身長 H_h に対する肩峰高 H_s との比率とは、大きな差異がない。このため、人が挙手した時に手がある手首 H_w の高さは、子供も大人も関係なく身長 H_h または肩峰高 H_s から暗黙的な比率によりわかる。

このため、本実施形態では、測距点が、頭の位置の高さ(=身長 H_h)から手を挙げたときの高さ(=手首 H_w)の比率と、肩峰高 H_s から手を挙げたときの高さ(=手首 H_w)の比率との間にある場合、手であると判別する。

【0111】

図14Aは、人が手を下げているイメージ図である。図14Bは、本実施形態に係る人が手を下げている期間の $x y z$ 値の第1マップの例を説明する図である。図14Aに示す例では、代表として1個のLRFセンサ10のみを示しているが、LRFセンサ10は、第1実施形態と同様に n 個が例えば天井に設置されている。また、図14A及び図14Bに示す例では、5人(人 $h u 1 \sim$ 人 $h u 5$)に対応する測距点が、LRFセンサ10によって測定された場合の例である。

図14Bでは、紙面に向かって左右方向を x 軸方向、奥行き方向を y 軸方向、高さ方向を z 軸方向とする。図14Bにおいて、点は測距点を表している。また、図14Bに示す第1マップは、1回のサンプリングによって取得された測距点による第1マップ(A)である。

【0112】

図15Aは、人が手を挙げているイメージ図である。図15Bは、本実施形態に係る人が手を挙げている期間の $x y z$ 値の第2マップの例を説明する図である。また、図15A及び図15Bに示す例では、5人(人 $h u 1 \sim$ 人 $h u 5$)に対応する測距点が、LRFセンサ10によって測定された場合の例である。

図15Bでは、紙面に向かって左右方向を x 軸方向、奥行き方向を y 軸方向、高さ方向を z 軸方向とする。図15Bにおいて、点は測距点を表している。また、図15Bに示す第1マップは、所定の時刻に取得された測距点による第2マップ(B)である。

図15A及び図15Bに示した例では、人 $h u 2$ と人 $h u 4$ が手を挙げている例である。

【0113】

次に、状況推定部50Aが行う処理を説明する。

図16は、本実施形態に係る状況推定部50Aが第1マップ～第5マップを用いて行う処理を説明する図である。図17は、本実施形態に係る状況推定部50Aがクラスタリングを行って手を挙げている人数を推定する処理を説明する図である。図18は、本実施形態に係る状況推定部50Aの処理手順のフローチャートである。また、図16では、紙面

10

20

30

40

50

に向かって左右方向をx軸方向、奥行き方向をy軸方向、高さ方向をz軸方向とする。図17では、紙面に向かって左右方向をx軸方向、奥行き方向をy軸方向とする。

また、以下の処理の説明において、状況推定部50Aが第1マップ～第5マップを用いて行う処理のみを説明するが、状況推定部50Aは、第1実施形態と同様に、ステップS1～S11（図3）の処理も行う。

【0114】

（ステップS501）状況推定部50Aは、手が下げられている期間における測距点の情報をサンプリングタイミング毎にセンシング処理部20Aから取得して、取得した測距点を統合する。次に、状況推定部50Aは、統合された測距点を用いて図16の符号m301が示す領域の図のような第1マップ（A）を生成する。例えば、時刻t0（図2）から、制御部102から時刻t1における音声信号を出力することを示すトリガー信号d5が入力される時刻t1までの間のサンプリング数が10個の場合、第1マップ（A）が状況推定部50Aによって10個生成される。状況推定部50Aは、処理をステップS502に進める。

10

【0115】

（ステップS502）状況推定部50Aは、図16の符号m302が示す領域の図のような、手が下げられている期間におけるxyz値の平均値マップである平均の第1マップ（Av）を生成する。すなわち、ステップS502において、状況推定部50Aは、全ての測距点を手が下げられているときの値であるとみなす。状況推定部50Aは、処理をステップS503に進める。

20

（ステップS503）状況推定部50Aは、ステップS502で生成した平均の第1マップ（Av）のz軸成分の値に、記憶部60Aから読み出した手首Hwと肩峰高Hsとの比（ Hw/Hs ）を乗じて、図16の符号m303が示す領域の図のような第3マップ（AzMax）を生成する。 Hw/Hs は、例えば、1.1である。この意味合いは、手が下げられているときに測定された測距点が肩峰だったとき、手が挙げられた場合の手首のz軸成分の値の上限値を算出している。状況推定部50Aは、処理をステップS504に進める。

【0116】

（ステップS504）状況推定部50Aは、ステップS502で生成した平均の第1マップ（Av）のz軸成分の値に、記憶部60Aから読み出した手首Hwと身長Hhとの比（ Hw/Hh ）を乗じて、図16の符号m304が示す領域の図のような第4マップ（AzMin）を生成する。 Hw/Hh は、例えば、1.3である。この意味合いは、手が下げられているときに測定された測距点が頭頂だったとき、手が挙げられた場合の手首のz軸成分の値の上限値を算出している。状況推定部50Aは、処理をステップS505に進める。

30

【0117】

（ステップS505）状況推定部50Aは、手が挙げられているときの測距点の情報をセンシング処理部20Aから取得して、取得した測距点を統合する。次に、状況推定部50Aは、統合された測距点を用いて図16の符号m305が示す領域の図のような第2マップ（B）を生成する。ステップS505においても、状況推定部50Aは、全ての測距点を手が下げられているときの値であるとみなす。状況推定部50Aは、処理をステップS506に進める。

40

【0118】

（ステップS506）状況推定部50Aは、ステップS505で生成した第2マップ（B）のうち、ステップS503で生成した第3マップ（AzMax）より大きく、ステップS504で生成した第4マップ（AzMin）より小さいz軸成分の値を有する測距点を抽出する。状況推定部50Aは、この処理によって、手が挙げられたときに、手が存在すると思われるz軸成分の値を有する測距点を抽出している。状況推定部50Aは、処理をステップS507に進める。

【0119】

50

(ステップS507) 状況推定部50Aは、抽出した測距点に含まれるx軸成分とy軸成分とをxy平面に投影して、図16の符号m306が示す領域の図のような投影点による第5マップ(C)を生成する。すなわち、符号m306が示す領域の図における点は投影点を表し、各投影は、x軸成分及びy軸成分の各値を有している。状況推定部50Aは、処理をステップS508に進める。

【0120】

(ステップS508) 状況推定部50Aは、ステップS507で生成した第5マップ(C)の投影点に対して、図17の符号m307が示す領域の図のように分類(クラスタリング)を行う。図17において、符号c1～c5が示す領域の投影点は、各クラスタに分類された投影点の集まりを表している。状況推定部50Aは、処理をステップS509に進める。

10

【0121】

(ステップS509) 状況推定部50Aは、ステップS508で分類された結果に基づいて、手を挙げた人数を推定する。なお、状況推定部50Aは、LRFセンサ10の測定性能に基づいて、クラスタに含まれる投影点が所定の値以下の場合、分類されたクラスタをノイズと判別する。図17に示した例では、状況推定部50Aは、投影点が1つの場合、ノイズであると判別する。このため、符号c1、c4、及びc5に含まれる領域の投影点をノイズであると判別する。次に、状況推定部50Aは、ステップS509で判別された結果に基づいて、手を挙げている人数を推定する。図17に示す例では、手を挙げていると判別されるクラスタは、符号c2とc3に含まれる領域の投影点である。

20

動作推定装置3Aは、シナリオ情報に基づいて、所定の質問を所定回数提示するまで、ステップS501～S509を繰り返す。

【0122】

(ステップS510) 状況推定部50は、シナリオ情報に基づいて、所定の質問が所定回数提示された後、ステップS509で推定された質問毎の推定結果に基づいて、多数決判定を行う。具体的には、状況推定部50は、設定されている条件に応じて、人統合処理部30が処理した結果、グリッド統合処理部40が処理した結果、また第1マップ～第5マップを用いて処理した結果のうち、少なくとも1つを選択する。状況推定部50Aは、質問毎の手を挙げた人数に対する推定結果に応じて、対応する質問を選択する。

以上で、挙手検出の処理を終了する。

30

【0123】

以上のように、本実施形態の動作推定装置3Aにおいて、状況推定部(例えば状況推定部50)は、人が挙手を行っていない第2期間(例えば時刻t0～t1の期間)における取得部(例えばセンシング処理部20)によって取得された距離と、人の体格データとに基づいて算出される挙手をしたときに人の手が存在する高さ方向の範囲である挙手存在範囲(例えば第3マップ(AzMax)～第4マップ(AzMin))を算出し、算出した挙手存在範囲にある第1時刻(例えば時刻t3またはt5)または第1期間(例えば時刻t3からt4の期間p4、または時刻t4～t5の期間p6)における取得部によって取得された距離(例えば第5マップ(C))を抽出し、抽出した距離に基づいて人の反応を推定する。

40

【0124】

この構成によって、本実施形態の動作推定装置3Aは、個人を特定する必要がないので人が密集している領域が存在しても、挙手をした人数を推定することができる。また、本実施形態によれば、LRFセンサ10の測定結果のバラツキによって測定値にばたつきがあっても、手を下げている所定の時間において、同期を行って加算しているため、ロバスト性を向上することができる。

【0125】

なお、第1実施形態及び第2実施形態では、記憶部101に記憶されているシナリオ情報に基づいて、処理部2がトリガー信号d5を生成する例を説明したが、これに限られない。

50

例えば、プレゼンターが人であってもよい。この場合、プレゼンターは、検出領域にいる人に向けて質問を発するタイミングで、不図示の操作スイッチを押すことで、トリガー信号d 5を動作推定装置（3、または3 A）に出力するようにしてもよい。例えば、プレゼンターは、挙手を促す質問を発するときにAボタンを押し、手を挙げることを促す音声を発するときにBボタンを押し、全ての質問が終了したときにCボタンを押すようにしてもよい。操作スイッチは、押されたボタンに応じたトリガー信号d 5を動作推定装置（3、または3 A）に出力するようにしてもよい。この場合であっても、動作推定装置（3、または3 A）は、入力されたトリガー信号d 5に応じて、所定の期間、及び所定の時刻に上述した処理を行うことで、多数決処理を行うことができる。

【0126】

なお、第1実施形態及び第2実施形態で説明した検出領域は、例えば1つの会場に限られない。検出領域は複数であってもよい。この場合、動作推定装置（3、または3 A）は、例えば会場毎に手を挙げている人数を推定、またはグリッド高さの変化量を推定し、全ての会場で推定した結果を状況推定部（50、または50 A）、あるいは制御部102が判別して多数決処理を行うようにしてもよい。そして、これらの会場は、離れていてもよく、会場が離れている場合は、複数の動作推定装置（3、または3 A）が処理結果を、例えば無線回線を介して、1つの動作推定装置（3、または3 A）に送信し、全ての処理結果を受信した動作推定装置（3、または3 A）が処理を行うようにしてもよい。動作推定装置（3、または3 A）が受信する処理結果は、例えばLRFセンサ10によって測距点の情報であってもよく、各会場に設置される動作推定装置（3、または3 A）によって処理された人データd 3、グリッド高さ情報d 4、及び測距点の情報d 7であってもよい。

【0127】

このように、離れた複数の会場にロボット（1、1 A）を設置し、同期を行って処理を行う場合、ロボット（1、1 A）が出力する音声信号は、国や地域に合わせたものであってもよい。

【0128】

なお、第1実施形態及び第2実施形態では、人が質問に対して手を挙げたり下げたりする例を説明したがこれに限られない。例えば、検出領域にいる複数の、頭頂より高い高さにプラカード等を挙げたり下げたりする場合であっても、上述した処理によって多数決処理を行うことができる。

また、加速度センサやGPS（グローバル・ポジショニング・システム；Global Positioning System）など位置検出を行える携帯端末を人が上げ下げするようにしてもよい。そして、動作推定装置（3、または3 A）は、これらの携帯端末が送信する情報を受信し、受信した情報も用いて、人が手を挙げたか否かの判別、手を挙げた人の人数等の人の反応を推定するようにしてもよい。

【0129】

なお、第1実施形態及び第2実施形態では、例えば図2を用いて説明したタイミングで処理を行う例を説明したが、これに限られない。

図19は、人に対して挙手の終了を促した時刻、第1時刻、及び第1期間を説明する図である。

時刻t 101～t 102の期間は、音声出力部103が人に対して挙手の終了を促す音声信号「てをさげてください」を出力した期間である。

【0130】

時刻t 4は、人に対して挙手の終了を促した時刻（第1時刻）である。すなわち、人に対して挙手の終了を促した時刻は、人に対して挙手の終了を促す音声信号が出力されている時刻t 101～t 102の期間内であればよい。人に対して挙手の終了を促した時刻は、例えば、人に対して挙手の終了を促す音声信号が出力開始されたタイミングであってもよく、音声信号が終了したタイミングであってもよい。または、人に対して挙手の終了を促した時刻は、例えば、音声信号の「（を）さ（げ）」のタイミングであってもよく、音声信号の「だ」のタイミングであってもよい。

【0131】

時刻 t_3 及び t_5 は、人に対して挙手の終了を促した時刻に基づき設定される第1時刻である。

また、時刻 $t_3 \sim$ 時刻 t_4 の期間 p_4 と、時刻 $t_4 \sim$ 時刻 t_5 の期間 p_6 とは、人に対して挙手の終了を促した時刻に基づき設定される第1期間である。

【0132】

状況推定部50は、例えば、第1実施形態と同様に、時刻 t_4 から期間 p_4 だけ前の間のデータ（例えば平均値、最大値、統計処理された値）を用いて検出された高さ、時刻 t_4 から期間 p_6 だけ後の時刻 t_5 のデータを用いて検出された高さ、を比較して挙手されている人数を推定する。

10

または、状況推定部50は、時刻 t_3 のデータを用いて検出された高さ、時刻 t_5 のデータを用いて検出された高さ、を比較して挙手されている人数を推定するようにしてもよい。

または、状況推定部50は、時刻 t_3 のデータを用いて検出された高さ、期間 p_6 のデータ（例えば平均値、最小値、統計処理された値）を用いて検出された高さ、を比較して挙手されている人数を推定するようにしてもよい。

または、状況推定部50は、期間 p_4 のデータを用いて検出された高さ、期間 p_6 のデータ（例えば平均値、最小値、統計処理された値）を用いて検出された高さ、を比較して挙手されている人数を推定するようにしてもよい。

以上のように、状況推定部50は、人が手を挙げていると推定される時刻または期間のいずれかの高さ、人が手を下げていると推定される時刻または期間のいずれかの高さ、を比較して挙手されている人数を推定するようにしてもよい。

20

【0133】

なお、本発明における動作推定装置（3、または3A）の機能を実現するためのプログラムをコンピュータ読み取り可能な記録媒体に記録して、この記録媒体に記録されたプログラムをコンピュータシステムに読み込ませ、実行することにより挙手の推定を行ってもよい。なお、ここでいう「コンピュータシステム」とは、OSや周辺機器等のハードウェアを含むものとする。また、「コンピュータシステム」は、ホームページ提供環境（あるいは表示環境）を備えたWWWシステムも含むものとする。また、「コンピュータ読み取り可能な記録媒体」とは、フレキシブルディスク、光磁気ディスク、ROM、CD-ROM等の可搬媒体、コンピュータシステムに内蔵されるハードディスク等の記憶装置のことをいう。さらに「コンピュータ読み取り可能な記録媒体」とは、インターネット等のネットワークや電話回線等の通信回線を介してプログラムが送信された場合のサーバやクライアントとなるコンピュータシステム内部の揮発性メモリ（RAM）のように、一定時間プログラムを保持しているものも含むものとする。

30

【0134】

また、上記プログラムは、このプログラムを記憶装置等に格納したコンピュータシステムから、伝送媒体を介して、あるいは、伝送媒体中の伝送波により他のコンピュータシステムに伝送されてもよい。ここで、プログラムを伝送する「伝送媒体」は、インターネット等のネットワーク（通信網）や電話回線等の通信回線（通信線）のように情報を伝送する機能を有する媒体のことをいう。また、上記プログラムは、前述した機能の一部を実現するためのものであってもよい。さらに、前述した機能をコンピュータシステムにすでに記録されているプログラムとの組み合わせで実現できるもの、いわゆる差分ファイル（差分プログラム）であってもよい。

40

【符号の説明】

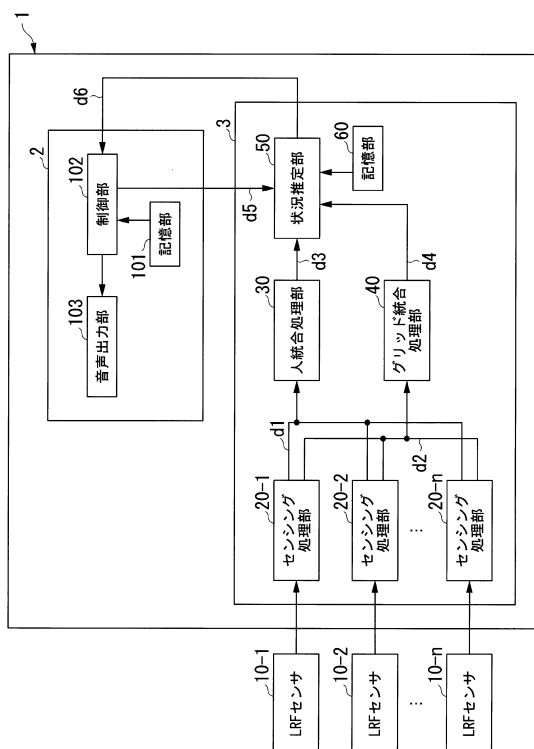
【0135】

1、1A…ロボット、2…処理部、3、3A…動作推定装置、10、10-1～10-n…LRFセンサ、20、20A、20-1～20-n、20A-1～20A-n…センシング処理部、30…人統合処理部、40…グリッド統合処理部、50、50A…状況推定部、60、60A…記憶部、101…記憶部、102…制御部、103…音声出力部、d

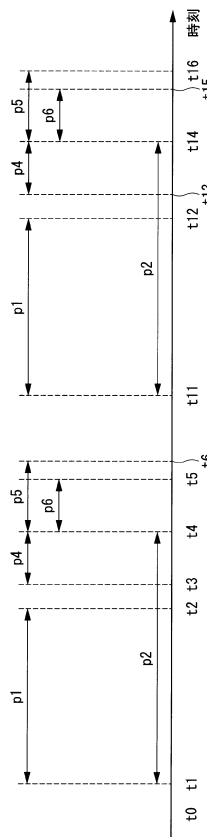
50

1…人クラスタデータ、d 2…グリッドデータ、d 3…人データ、d 4…グリッド高さを示す情報、d 5…トリガー信号、d 6…検出情報、d 7…測距点の情報、A z…第 1 マップ、B…第 2 マップ、A z M a x…第 3 マップ、A z M i n…第 4 マップ、C…第 5 マップ

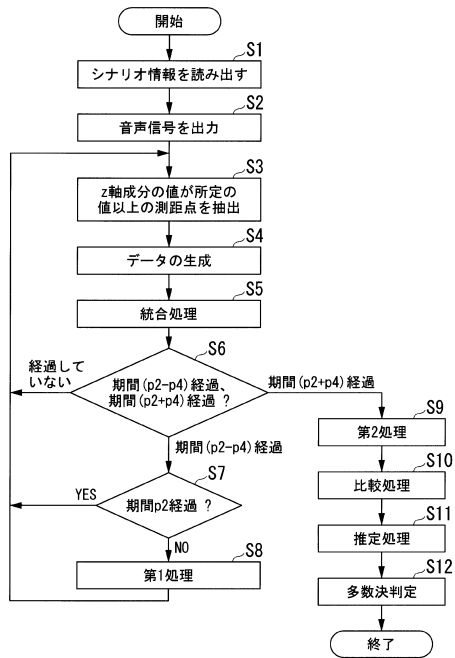
【図 1】



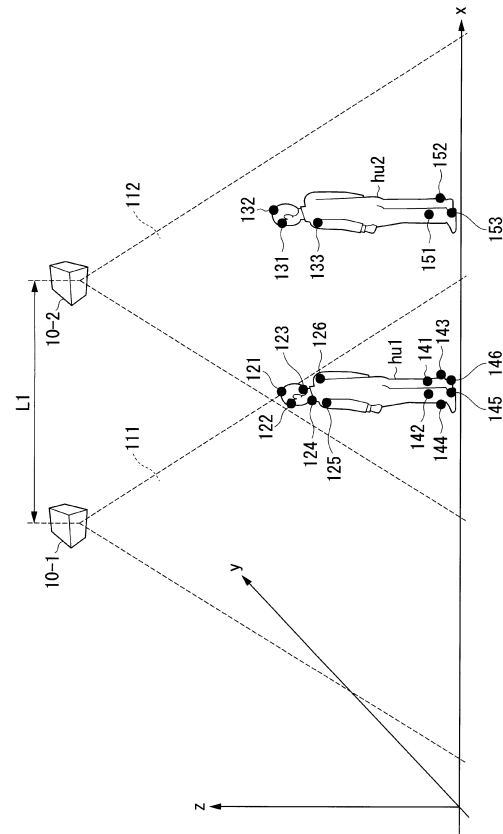
【図 2】



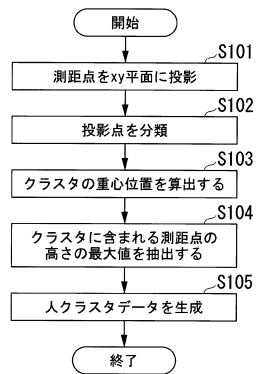
【図 3】



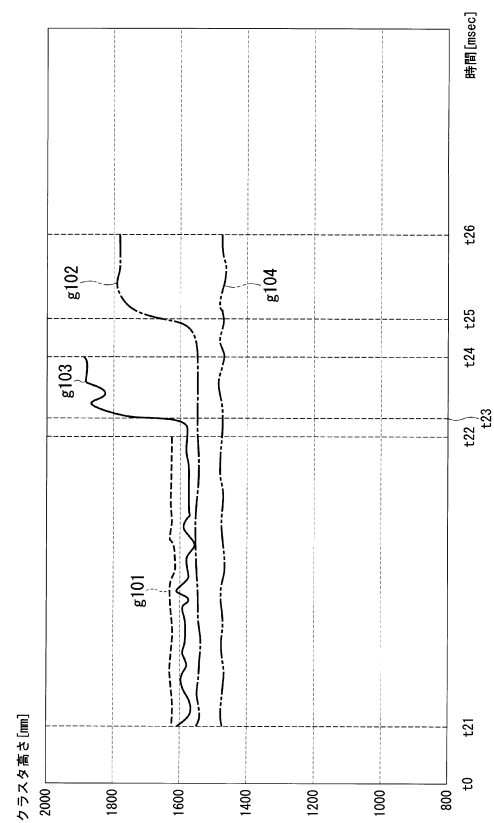
【図 4】



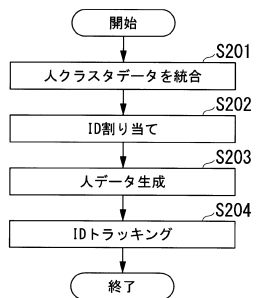
【図 5】



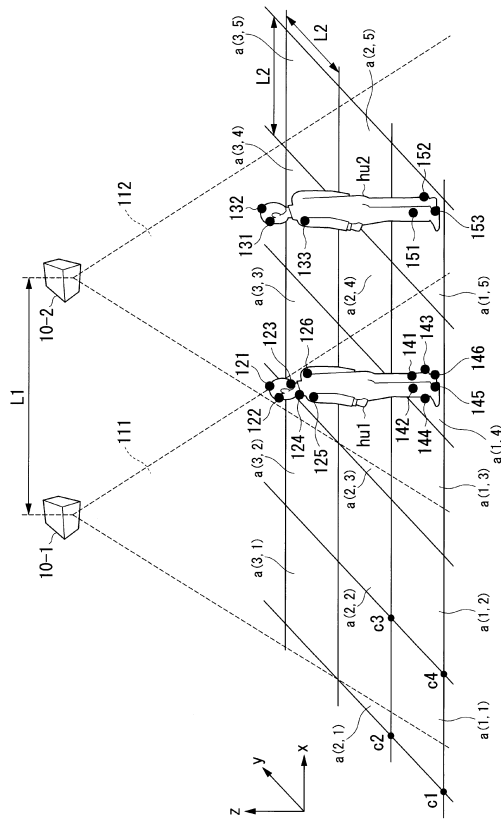
【図 7】



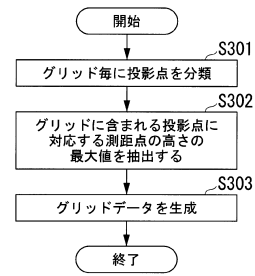
【図 6】



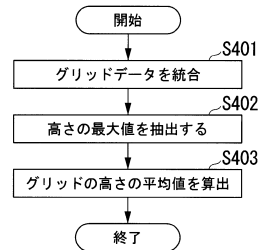
【図 8】



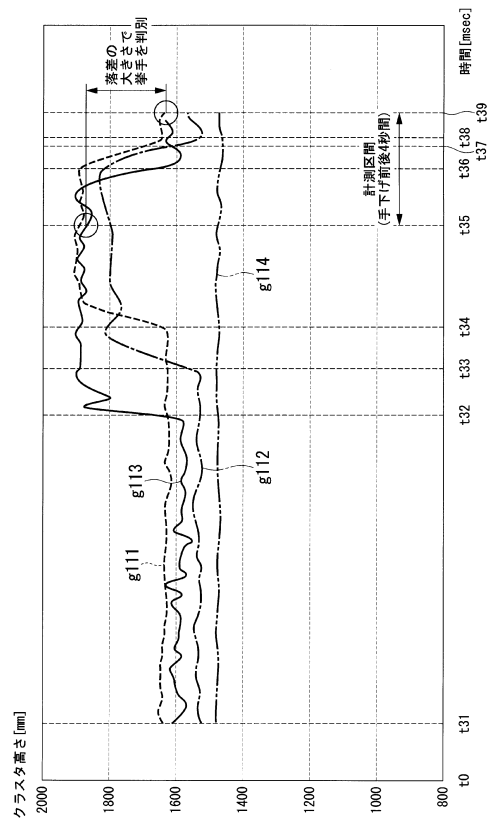
【图 9】



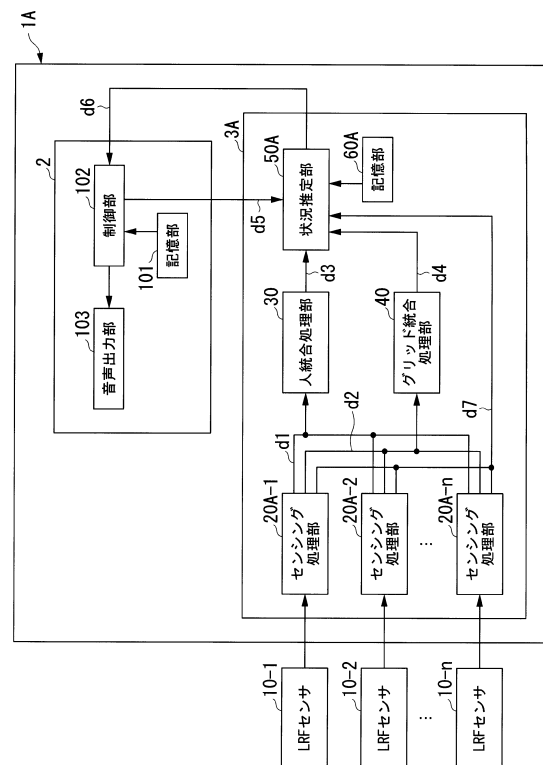
【図 10】



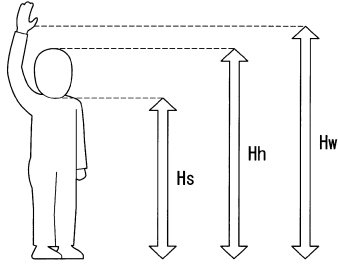
【図 1 1】



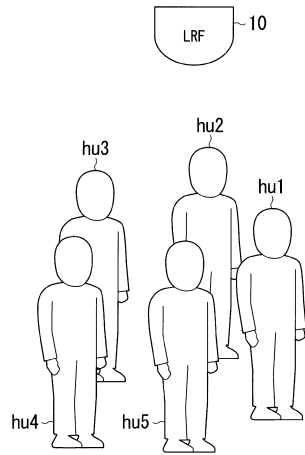
【图 1 2】



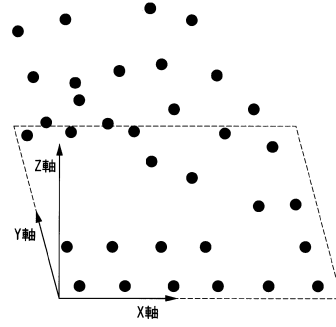
【図 1 3】



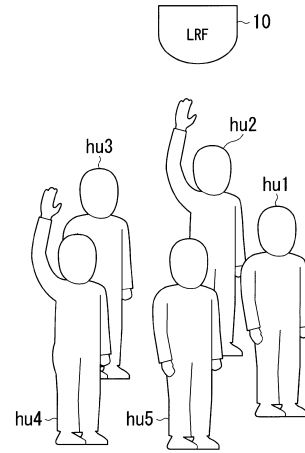
【図 1 4 A】



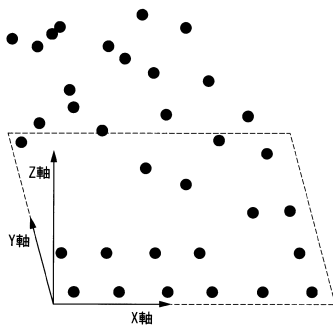
【図 1 4 B】



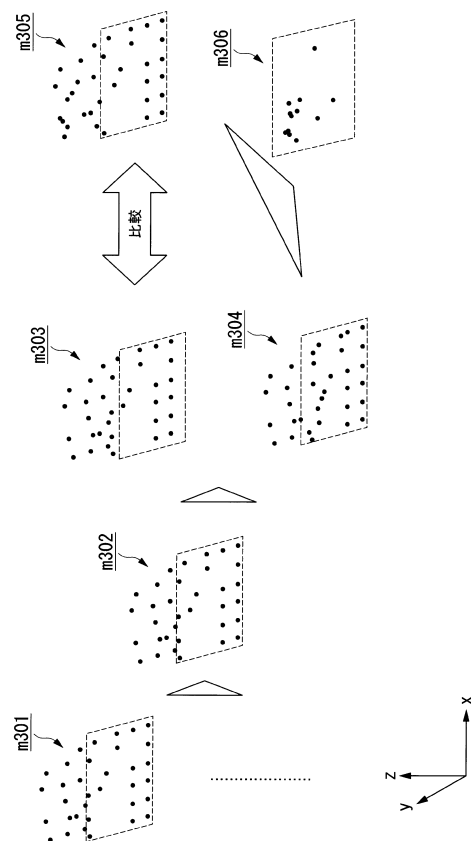
【図 1 5 A】



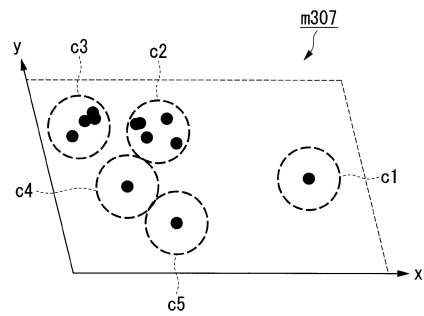
【図 1 5 B】



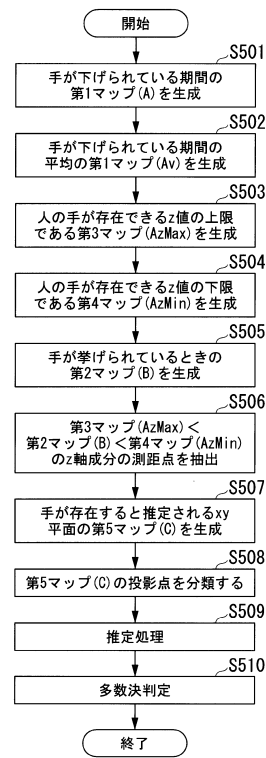
【図 1 6】



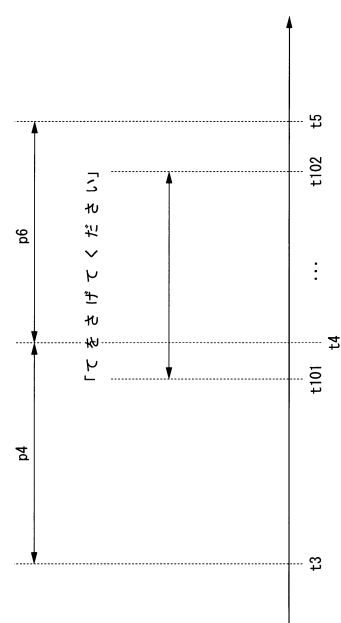
【図 17】



【図 18】



【図 19】



フロントページの続き

- (74)代理人 100175802
弁理士 寺本 光生
- (74)代理人 100094400
弁理士 鈴木 三義
- (72)発明者 鳥羽 薫
埼玉県和光市中央1丁目4番1号 株式会社本田技術研究所内
- (72)発明者 大原 慶久
埼玉県和光市中央1丁目4番1号 株式会社本田技術研究所内

審査官 佐田 宏史

- (56)参考文献 特開2012-065833 (JP, A)
特開2010-239327 (JP, A)
特表2013-514585 (JP, A)
特開2011-194073 (JP, A)
特開2001-128134 (JP, A)
特開2013-231520 (JP, A)
特開2011-086286 (JP, A)
特開平11-190816 (JP, A)
特開2002-197463 (JP, A)
特開2005-215927 (JP, A)
土屋 友寛、外3名, “レンジイメージを用いた空間内の時系列人物認識とモデル化”, 情報処理学会研究報告, 日本, 社団法人情報処理学会, 2007年 7月 2日, Vol.2007, No.65, p
p.151-156

- (58)調査した分野(Int.Cl., DB名)
- | | |
|------|--------------|
| G06F | 3/01 |
| G06T | 7/00-7/90 |
| B25J | 13/08, 19/02 |