

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6872128号
(P6872128)

(45) 発行日 令和3年5月19日 (2021.5.19)

(24) 登録日 令和3年4月21日 (2021.4.21)

(51) Int. Cl.	F I
GO 1 B 11/00 (2006.01)	GO 1 B 11/00 H
GO 1 B 11/26 (2006.01)	GO 1 B 11/26 H
GO 6 T 7/70 (2017.01)	GO 6 T 7/70 Z

請求項の数 6 (全 17 頁)

(21) 出願番号	特願2017-543143 (P2017-543143)	(73) 特許権者	000002185
(86) (22) 出願日	平成28年9月16日 (2016.9.16)		ソニーグループ株式会社
(86) 国際出願番号	PCT/JP2016/077427		東京都港区港南1丁目7番1号
(87) 国際公開番号	W02017/057054	(74) 代理人	100121131
(87) 国際公開日	平成29年4月6日 (2017.4.6)		弁理士 西川 孝
審査請求日	令和1年8月22日 (2019.8.22)	(74) 代理人	100082131
(31) 優先権主張番号	特願2015-194554 (P2015-194554)		弁理士 稲本 義雄
(32) 優先日	平成27年9月30日 (2015.9.30)	(72) 発明者	貝野 彰彦
(33) 優先権主張国・地域又は機関	日本国 (JP)		東京都港区港南1丁目7番1号 ソニー株 式会社内
		(72) 発明者	鶴見 辰吾
			東京都港区港南1丁目7番1号 ソニー株 式会社内

最終頁に続く

(54) 【発明の名称】 情報処理装置、情報処理方法、およびプログラム

(57) 【特許請求の範囲】

【請求項 1】

移動体に搭載されたステレオカメラから構成され、所定のフレームレートで動画像を撮像することにより時系列のステレオ主画像対を出力する主撮像部と、

前記主撮像部とは撮像方向が異なるように前記移動体に搭載され、所定のフレームレートで動画像を撮像することにより時系列の副画像を出力する1または複数の副撮像部と、

前記主撮像部によって撮像された前記ステレオ主画像対から検出した特徴点の3D座標と、前記特徴点を前記副画像の複数フレームに亘ってトラッキングすることによって得られた、前記特徴点の前記ステレオ主画像対における2D座標とに基づいて前記移動体のポーズを推定する最終ポーズ決定部と

を備える情報処理装置。

【請求項 2】

前記副撮像部は、広角カメラ、魚眼カメラ、または全方位カメラから構成され、所定のフレームレートで動画像を撮像することにより、時系列の前記副画像を出力し、

前記最終ポーズ決定部は、

前記主撮像部によって撮像された前記ステレオ主画像対から検出した特徴点の3D座標を推定する3D推定部と、

前記特徴点を前記副画像の複数フレームに亘ってトラッキングし、トラッキング結果を前記ステレオ主画像対における2D座標に変換する2Dトラッキング部と、

前記3D推定部によって推定された前記特徴点の3D座標と、前記2Dトラッキング

10

20

部によって変換された前記特徴点の前記ステレオ主画像対における2D座標とに基づいて最終ポーズを推定するポーズ推定部とを有する

請求項1に記載の情報処理装置。

【請求項3】

前記3D推定部は、前記主撮像部によって撮像された前記ステレオ主画像対から検出した特徴点の3D座標を推定するとともに、前記特徴点の前記副画像における2D座標を推定し、

前記2Dトラッキング部は、前記3D推定部によって推定された前記特徴点の前記副画像における2D座標に基づき、前記特徴点を前記副画像の複数フレームに亘ってトラッキングし、トラッキング結果を前記ステレオ主画像対における2D座標に変換する

10

請求項2に記載の情報処理装置。

【請求項4】

複数の前記副撮像部は、前記3D推定部による前記特徴点の前記副画像における2D座標の推定結果に基づき、選択的に切り替えられて撮像を行う

請求項3に記載の情報処理装置。

【請求項5】

情報処理装置による情報処理方法において、

前記情報処理装置による、

移動体に搭載されたステレオカメラから構成され、所定のフレームレートで動画を撮像することにより時系列のステレオ主画像対を出力する主撮像部によって撮像された前記ステレオ主画像対から検出した特徴点の3D座標と、前記主撮像部とは撮像方向が異なるように前記移動体に搭載され、所定のフレームレートで動画を撮像することにより時系列の副画像を出力する1または複数の副撮像部によって撮像された前記副画像の複数フレームに亘って前記特徴点をトラッキングすることによって得られた、前記特徴点の前記ステレオ主画像対における2D座標とに基づいて前記移動体のポーズを推定する最終ポーズ決定ステップを

20

含む情報処理方法。

【請求項6】

情報処理装置の制御用のプログラムにおいて、

移動体に搭載されたステレオカメラから構成され、所定のフレームレートで動画を撮像することにより時系列のステレオ主画像対を出力する主撮像部によって撮像された前記ステレオ主画像対から検出した特徴点の3D座標と、前記主撮像部とは撮像方向が異なるように前記移動体に搭載され、所定のフレームレートで動画を撮像することにより時系列の副画像を出力する1または複数の副撮像部によって撮像された前記副画像の複数フレームに亘って前記特徴点をトラッキングすることによって得られた、前記特徴点の前記ステレオ主画像対における2D座標とに基づいて前記移動体のポーズを推定する最終ポーズ決定ステップを

30

含む処理を情報処理装置のコンピュータに実行させるプログラム。

【発明の詳細な説明】

【技術分野】

40

【0001】

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し、特に、例えば、自動車などの移動体に搭載して自己位置を推定する場合に用いて好適な情報処理装置、情報処理方法、およびプログラムに関する。

【背景技術】

【0002】

自動車などの移動体に搭載して自己位置を推定する技術としてSLAM(Simultaneous Localization and Mapping)が知られている。

【0003】

ステレオカメラ等を用いたSLAMによる自己位置推定技術は以下の処理から構成される。

50

すなわち、初めに、ステレオカメラにより撮像された視差の異なるステレオ画像対から特徴があるテクスチャ領域を検出し、テクスチャ領域から特徴点を検出して奥行推定を行う処理、すなわち、特徴点の3D座標を推定する処理が行われる。次に、得られた特徴点を複数フレームに亘ってトラッキングする処理、すなわち、特徴点の2D座標トラッキング処理が行われる。最後に、推定された特徴点の3D座標と、追跡された特徴点の2D座標の組み合わせを尤も適切に表現できるポーズ（自己位置と向き）を推定する処理、すなわち、6DOF（自由度6）の自己位置を推定する処理が行われる（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2008-304268号公報

【発明の概要】

【発明が解決しようとする課題】

【0005】

SLAMによる自己位置推定技術が上述した処理から構成されることにより、一般的には、ステレオカメラにより撮像された画像の大部分がテクスチャのない壁や床によって占められる場合、周辺環境（他の自動車や歩行者など）が動く場合、特徴点の3D座標の空間分布が偏る場合には自己位置推定が困難となる。

【0006】

特に、屋外で使用される自動車などの移動体にSLAMによる自己位置推定技術を適用する場合には、以下のような問題が発生して自己位置の推定が困難となる事態が起こり得る。

【0007】

すなわち、例えば、近くに特徴のある被写体が存在せず、遠方の建物がテクスチャとして検出される場合、テクスチャの時間変動が得られない。画像上において周辺環境が移動する場合、自己が動いているのか、周辺環境が動いているのかを区別することができない。自己（自動車）が高速に移動する場合、画面上から特徴点が消失してしまい、トラッキングが困難となる。

【0008】

本開示はこのような状況に鑑みてなされたものであり、継続的に自己位置を推定できるようにするものである。

【課題を解決するための手段】

【0009】

本開示の一側面である情報処理装置は、移動体に搭載されたステレオカメラから構成され、所定のフレームレートで動画像を撮像することにより時系列のステレオ主画像対を出力する主撮像部と、前記主撮像部とは撮像方向が異なるように前記移動体に搭載され、所定のフレームレートで動画像を撮像することにより時系列の副画像を出力する1または複数の副撮像部と、前記主撮像部によって撮像されたステレオ主画像対から検出した特徴点の3D座標と、前記特徴点を前記副画像の複数フレームに亘ってトラッキングすることによって得られた、前記特徴点の前記ステレオ主画像対における2D座標とに基づいて前記移動体のポーズを推定する最終ポーズ決定部とを備える。

【0013】

前記副撮像部は、広角カメラ、魚眼カメラ、または全方位カメラから構成され、所定のフレームレートで動画像を撮像することにより、時系列の前記副画像を出力することができ、前記最終ポーズ決定部は、前記主撮像部によって撮像された前記ステレオ主画像対から検出した特徴点の3D座標を推定する3D推定部と、前記特徴点を前記副画像の複数フレームに亘ってトラッキングし、トラッキング結果を前記ステレオ主画像対における2D座標に変換する2Dトラッキング部と、前記3D推定部によって推定された前記特徴点の3D座標と、前記2Dトラッキング部によって変換された前記特徴点の前記ステレオ主画像対における2D座標とに基づいて最終ポーズを推定するポーズ推定部とを有することが

10

20

30

40

50

できる。

【0014】

前記3D推定部は、前記主撮像部によって撮像された前記ステレオ主画像対から検出した特徴点の3D座標を推定するとともに、前記特徴点の前記副画像における2D座標を推定することができ、前記2Dトラッキング部は、前記3D推定部によって推定された前記特徴点の前記副画像における2D座標に基づき、前記特徴点を前記副画像の複数フレームに亘ってトラッキングし、トラッキング結果を前記ステレオ主画像対における2D座標に変換することができる。

【0015】

複数の前記副撮像部は、前記3D推定部による前記特徴点の前記副画像における2D座標の推定結果に基づき、選択的に切り替えられて撮像を行うことができる。

10

【0016】

本開示の一側面である情報処理方法は、情報処理装置による情報処理方法において、前記情報処理装置による、移動体に搭載されたステレオカメラから構成され、所定のフレームレートで動画像を撮像することにより時系列のステレオ主画像対を出力する主撮像部によって撮像された前記ステレオ主画像対から検出した特徴点の3D座標と、前記主撮像部とは撮像方向が異なるように前記移動体に搭載され、所定のフレームレートで動画像を撮像することにより時系列の副画像を出力する1または複数の副撮像部によって撮像された前記副画像の複数フレームに亘って前記特徴点をトラッキングすることによって得られた、前記特徴点の前記ステレオ主画像対における2D座標とに基づいて前記移動体のポーズを推定する最終ポーズ決定ステップを含む。

20

【0017】

本開示の一側面であるプログラムは、情報処理装置の制御用のプログラムにおいて、移動体に搭載されたステレオカメラから構成され、所定のフレームレートで動画像を撮像することにより時系列のステレオ主画像対を出力する主撮像部によって撮像された前記ステレオ主画像対から検出した特徴点の3D座標と、前記主撮像部とは撮像方向が異なるように前記移動体に搭載され、所定のフレームレートで動画像を撮像することにより時系列の副画像を出力する1または複数の副撮像部によって撮像された前記副画像の複数フレームに亘って前記特徴点をトラッキングすることによって得られた、前記特徴点の前記ステレオ主画像対における2D座標とに基づいて前記移動体のポーズを推定する最終ポーズ決定ステップを含む処理を情報処理装置のコンピュータに実行させる。

30

【0018】

本開示の一側面においては、移動体に搭載されたステレオカメラから構成され、所定のフレームレートで動画像を撮像することにより時系列のステレオ主画像対を出力する主撮像部によって撮像された前記ステレオ主画像対から検出した特徴点の3D座標と、前記主撮像部とは撮像方向が異なるように前記移動体に搭載され、所定のフレームレートで動画像を撮像することにより時系列の副画像を出力する1または複数の副撮像部によって撮像された前記副画像の複数フレームに亘って前記特徴点をトラッキングすることによって得られた、前記特徴点の前記ステレオ主画像対における2D座標とに基づいて前記移動体のポーズが推定される。

40

【発明の効果】

【0019】

本開示の一側面によれば、継続的に自己位置を推定することができる。

【図面の簡単な説明】

【0020】

【図1】本開示を適用した情報処理装置の第1の構成例を示すブロック図である。

【図2】図1の第1乃至第4の撮像部の自動車における配置例を示す図である。

【図3】情報処理装置の第1の構成例による最終ポーズ決定処理を説明するフローチャートである。

【図4】本開示を適用した情報処理装置の第2の構成例を示すブロック図である。

50

【図 5】図 4 の第 1 乃至第 3 の撮像部の自動車における配置例を示す図である。

【図 6】第 1 乃至第 3 の撮像部の代わりに全方位カメラを配置した例を示す図である。

【図 7】第 1 乃至第 3 の撮像部の代わりに下向きカメラを配置した例を示す図である。

【図 8】情報処理装置の第 2 の構成例による最終ポーズ決定処理を説明するフローチャートである。

【図 9】本開示を適用した情報処理装置をヘッドマウントディスプレイに搭載した場合の概念図である。

【図 10】本開示を適用した情報処理装置を自立型掃除ロボットに搭載した場合の概念図である。

【図 11】本開示を適用した情報処理装置をドローン（自律型無人飛行機）に搭載した場合の概念図である。 10

【図 12】本開示を適用した情報処理装置を 3D モデルや任意視点映像などの作成に応用する場合の概念図である。

【図 13】汎用のコンピュータの構成例を示すブロック図である。

【発明を実施するための形態】

【0021】

以下、本開示を実施するための最良の形態（以下、実施の形態と称する）について、図面を参照しながら詳細に説明する。以下においては、本開示の実施の形態である情報処理装置は、屋外で使用される自動車に搭載して使用することを想定して説明する。

【0022】 20

< 本開示を適用した情報処理装置の第 1 の構成例 >

図 1 は、本開示を適用した情報処理装置の第 1 の構成例を示している。

【0023】

該第 1 の構成例は、第 1 の撮像部 11 - 1 乃至第 4 の撮像部 11 - 4、SLAM 処理部 12 - 1 乃至 12 - 4、および統合部 13 から構成される。

【0024】

第 1 の撮像部 11 - 1 は、高解像度ステレオカメラから構成されており、所定のフレームレートで動画を撮像し、その結果得られる視差が異なるステレオ画像対を SLAM 処理部 12 - 1 に供給する。第 2 の撮像部 11 - 2 乃至第 4 の撮像部 11 - 4 についても第 1 の撮像部 11 - 1 と同様に構成される。なお、ステレオカメラの代わりに、1 台のカメラと被写体までの距離（3D 座標）を測定できる偏光センサ、アクティブ系センサ（ToF センサ、SL センサなど）を用いてもよい。 30

【0025】

図 2 は、第 1 の撮像部 11 - 1 乃至第 4 の撮像部 11 - 4 の自動車における配置例を示しており、同図 A は側面図を示しており、同図 B は上面図を示している。

【0026】

第 1 の撮像部 11 - 1 乃至第 4 の撮像部 11 - 4 は、それぞれ異なる方向を撮像範囲とするように配置される。同図の場合、第 1 の撮像部 11 - 1 は、前方を撮像範囲とするように、自動車の前側に配置されている。第 2 の撮像部 11 - 2 は、右側方を撮像範囲とするように、自動車の右側に配置されている。第 3 の撮像部 11 - 3 は、左側方を撮像範囲とするように、自動車の左側面に配置されている。第 4 の撮像部 11 - 4 は、後方を撮像範囲とするように、自動車の後側に配置されている。 40

【0027】

なお、第 1 の撮像部 11 - 1 乃至第 4 の撮像部 11 - 4 の配置は、図 2 の配置例に限定されるものではなく、複数の高解像度ステレオカメラによって自動車の周囲を撮像できるようにすればよい。また、図示は省略するが、路面を撮像範囲とするように、自動車の下部に第 2 の撮像部 11 - 2 乃至第 4 の撮像部 11 - 4 のいずれかを配置するようにしてもよい。

【0028】

図 1 に戻る。SLAM 処理部 12 - 1 は、第 1 の撮像部 11 - 1 から順次所定のフレームレ 50

ートで供給される、視差が異なるステレオ画像対を対象としてSLAM処理を行い、その結果えられるポーズとその信頼度を統合部13に出力する。ここで、ポーズとは、6DOFの自己位置である。信頼度は、有効な特徴点の数、特徴点の3次元空間分布、直前に得られた最終ポーズとの差分、または、推定における最適化演算における残差のうちの少なくとも一つを含むものとする。

【0029】

SLAM処理部12-2乃至12-4についてもSLAM処理部12-1と同様、前段から順次所定のフレームレートで供給される、視差が異なるステレオ画像対を対象としてSLAM処理を行い、その結果得られるポーズとその信頼度を統合部13に出力する。

【0030】

なお、SLAM処理部12-1乃至12-4が行うSLAM処理の一部または全部を、いわゆるクラウド上のサーバに実行させるようにしてもよい。

【0031】

統合部13は、SLAM処理部12-1乃至12-4から入力されるポーズを、その信頼度に基づいて統合して最終的なポーズを決定する。具体的には、SLAM処理部12-1乃至12-4から入力されるポーズをその信頼度に基づいて重み付け平均したり、信頼度が最も高いポーズを採用したり、SLAM処理部12-1乃至12-4から入力されるポーズを、その信頼度に基づくカルマンフィルタ(Kalman Filter)によって統合したりして最終的なポーズを決定する。

【0032】

<情報処理装置の第1の構成例による最終ポーズ決定処理>

次に、図3は、情報処理装置の第1の構成例による最終ポーズ決定処理を説明するフローチャートである。

【0033】

以下、第1の撮像部11-1乃至第4の撮像部11-4のそれぞれに対応するSLAM処理部12-1乃至12-4の組み合わせをそれぞれステレオカメラ系と称する。また、第1の撮像部11-1乃至第4の撮像部11-4を個々に区別する必要がある場合には撮像部11と称し、SLAM処理部12-1乃至12-4を個々に区別する必要がある場合にはSLAM処理部12と称する。

【0034】

ステップS1において、各ステレオカメラ系の撮像部11によって所定のフレームレートで動画の撮像を開始し、その結果得られる視差が異なるステレオ画像対を対象としてSLAM処理部12に供給する。

【0035】

ステップS2において、各ステレオカメラ系のSLAM処理部12は、前段の撮像部11から順次所定のフレームレートで供給される、視差が異なるステレオ画像対を対象としてSLAM処理を行い、ステップS3において、その結果えられるポーズとその信頼度を統合部13に出力する。

【0036】

ステップS4において、統合部13は、各ステレオカメラ系のSLAM処理部12から入力されるポーズを、その信頼度に基づいて統合して最終的なポーズを決定する。具体的には、上述した3種類の統合方法のいずれを採用してもよい。

【0037】

以上に説明した最終ポーズ決定処理によれば、全てのステレオカメラ系で自己位置が推定され、その結果が統合されるので、仮に1つのステレオカメラ系で自己位置推定ができなかったり、誤差が有ったりしても、継続的に最終的なポーズを推定することができる。

【0038】

なお、ステレオカメラ系の数は4に限らず、2以上であればよい。当然ながら、その数が増えれば、最終的なポーズの正確性を高めることができる。

【0039】

10

20

30

40

50

< 本開示を適用した情報処理装置の第 2 の構成例 >

図 4 は、本開示を適用した情報処理装置の第 2 の構成例を示している。

【 0 0 4 0 】

該第 2 の構成例は、大別して 3 D 推定部 2 1、2 D トラッキング部 2 7、およびポーズ推定部 3 1 から構成される。

【 0 0 4 1 】

3 D 推定部 2 1 は、ステレオ撮像部 2 2、特徴点検出部 2 3、奥行推定部 2 4、座標変換部 2 5 および保持部 2 6 を有する。

【 0 0 4 2 】

ステレオ撮像部 2 2 は、自動車の前方を撮像する高解像度ステレオカメラから構成されており、所定のフレームレートで動画像を撮像し、その結果得られる視差が異なるステレオ画像対を特徴点検出部 2 3 に供給する。

【 0 0 4 3 】

特徴点検出部 2 3 は、ステレオ撮像部 2 2 から供給されるステレオ画像対から特徴があるテクスチャ領域を検出し、テクスチャ領域から 1 以上の特徴点を検出して、検出結果を示す 2 D 特徴点群@cam0 を奥行推定部 2 4 に出力する。

【 0 0 4 4 】

奥行推定部 2 4 は、ステレオ画像対における特徴点の奥行推定を行う処理、すなわち、特徴点の 3 D 座標を推定する処理を行い、処理結果を示す 3 D 特徴点群@cam0 を座標変換部 2 5 および保持部 2 6 に出力する。

【 0 0 4 5 】

座標変換部 2 5 は、奥行推定部 2 4 から得られる 3 D 特徴点群@cam0 に基づき、ステレオ画像対における特徴点の、第 1 の撮像部 2 8 - 1 (後述) によって撮像される第 1 の画像における 2 D 座標を推定し、その結果得られる推定 2 D 特徴点群@cam[1] を保持部 2 6 に出力する。また、座標変換部 2 5 は、奥行推定部 2 4 から得られる 3 D 特徴点群@cam0 に基づき、ステレオ画像対における特徴点の、第 2 の撮像部 2 8 - 2 (後述) によって撮像される第 2 の画像における 2 D 座標を推定し、その結果得られる推定 2 D 特徴点群@cam[2] を保持部 2 6 に出力する。さらに、座標変換部 2 5 は、奥行推定部 2 4 から得られる 3 D 特徴点群@cam0 に基づき、ステレオ画像対における特徴点の、第 3 の撮像部 2 8 - 3 (後述) によって撮像される第 3 の画像における 2 D 座標を推定し、その結果得られる推定 2 D 特徴点群@cam[3] を保持部 2 6 に出力する。

【 0 0 4 6 】

保持部 2 6 は、奥行推定部 2 4 から入力される 3 D 特徴点群@cam0 と、座標変換部 2 5 から入力される推定 2 D 特徴点群@cam[i] (i=1,2,3) を保持する。

【 0 0 4 7 】

2 D トラッキング部 2 7 は、第 1 の撮像部 2 8 - 1 乃至第 3 の撮像部 2 8 - 3、トラッキング部 2 9 - 1 乃至 2 9 - 3、および座標変換部 3 0 - 1 乃至 3 0 - 3 を有する。

【 0 0 4 8 】

第 1 の撮像部 2 8 - 1 は、広角カメラまたは魚眼カメラなどの撮像範囲が広いカメラから成り、所定のフレームレートで動画像を撮像し、その結果得られる第 1 の画像を トラッキング部 2 9 - 1 に供給する。同様に、第 2 の撮像部 2 8 - 2 は、第 2 の画像を トラッキング部 2 9 - 2 に供給し、第 3 の撮像部 2 8 - 3 は、第 3 の画像を トラッキング部 2 9 - 3 に供給する。上述したように、第 1 の撮像部 2 8 - 1 乃至第 3 の撮像部 2 8 - 3 は、視差に基づくキャリブレーションなどが必須となるステレオカメラではなく、それぞれが 1 台の広角カメラなどで構成することができるので、ステレオカメラを用いた場合に比較して、低コスト化を実現でき、キャリブレーションの手間を省くことができる。

【 0 0 4 9 】

なお、第 1 の撮像部 2 8 - 1 乃至第 3 の撮像部 2 8 - 3 から得られる第 1 乃至第 3 の画像から有効なテクスチャ範囲を選定し、ステレオ撮像部 2 2 を構成する高解像度ステレオカメラの向きやズーム率を制御するようにしてもよい。

10

20

30

40

50

【 0 0 5 0 】

図 5 は、第 1 の撮像部 2 8 - 1 乃至第 3 の撮像部 2 8 - 3 の自動車における配置例を示している。

【 0 0 5 1 】

同図に示されるように、第 1 の撮像部 2 8 - 1 は、例えば、自動車の左側を撮像範囲とする位置に取り付けられる。第 2 の撮像部 2 8 - 2 は、例えば、自動車の右側を撮像範囲とする位置に取り付けられる。第 3 の撮像部 2 8 - 3 は、例えば、自動車の後方を撮像範囲とする位置に取り付けられる。

【 0 0 5 2 】

なお、第 1 の撮像部 2 8 - 1 乃至第 3 の撮像部 2 8 - 3 の全てに対して同時に撮像を行わせるのではなく、第 1 の撮像部 2 8 - 1 乃至第 3 の撮像部 2 8 - 3 を選択的に切り替えて撮像させてもよい。具体的には、ステレオ画像対から検出された特徴点の空間分布（特に奥行方向の分布）が偏らないように選択したり、動きが有る周辺環境を撮像していないものを選択したり、特徴点として有効なテクスチャを撮像しているものを選択したりして撮像させるようにしてもよい。これにより、省電力化を図ることができる。

10

【 0 0 5 3 】

また、第 1 の撮像部 2 8 - 1 乃至第 3 の撮像部 2 8 - 3 を設ける代わりに、図 6 に示されるように、周囲 3 6 0 度を撮像範囲とする全方位カメラ 2 8 を自動車の屋根などに配置してもよい。または、図 7 に示されるように、第 1 の撮像部 2 8 - 1 乃至第 3 の撮像部 2 8 - 3 を設ける代わりに、路面を撮像範囲とする下向きカメラ 2 8 を自動車の底部などに配置してもよい。

20

【 0 0 5 4 】

図 4 に戻る。トラッキング部 2 9 - 1 は、保持部 2 6 から推定 2 D 特徴点群@cam[1]を取得し、取得した推定 2 D 特徴点群@cam[1]に基づいて、ステレオ画像対から検出された特徴点を、第 1 の撮像部 2 8 - 1 から得られる第 1 の画像上でトラッキングし、その結果得られる 2 D 特徴点群@cam[1]を座標変換部 3 0 - 1 に出力する。なお、ステレオ画像対から検出された特徴点のうち、第 1 の画像上に存在しない特徴点については、2 D 特徴点群@cam[1]から削除する。同様に、トラッキング部 2 9 - 2 は、2 D 特徴点群@cam[2]を座標変換部 3 0 - 2 に出力し、トラッキング部 2 9 - 3 は、2 D 特徴点群@cam[3]を座標変換部 3 0 - 3 に出力する。

30

【 0 0 5 5 】

座標変換部 3 0 - 1 は、トラッキング部 2 9 - 1 から入力された 2 D 特徴点群@cam[1]に基づき、第 1 の画像上でトラッキングした特徴点の 2 D 座標を、ステレオ画像対における 2 D 座標に変換し、その結果得られる 2 D 特徴点群@cam0-1をポーズ推定部 3 1 に出力する。同様に、座標変換部 3 0 - 2 は 2 D 特徴点群@cam0-2をポーズ推定部 3 1 に出力し、座標変換部 3 0 - 3 は 2 D 特徴点群@cam0-3をポーズ推定部 3 1 に出力する。

【 0 0 5 6 】

ポーズ推定部 3 1 は、保持部 2 6 から 3 D 特徴点群@cam0を取得し、取得した 3 D 特徴点群@cam0と、2 D 特徴点群@cam0-1乃至 2 D 特徴点群@cam0-3に基づいて、3 D 座標と 2 D 座標との組み合わせが尤も適切に表現できる最終ポーズ（6 DOF）を推定する。

40

【 0 0 5 7 】

< 情報処理装置の第 2 の構成例による最終ポーズ決定処理 >

次に、図 8 は、情報処理装置の第 2 の構成例による最終ポーズ決定処理を説明するフローチャートである。

【 0 0 5 8 】

前提として、既にステレオ撮像部 2 2 は所定のフレームレートで動画像の撮像を開始しており、その結果得られるステレオ画像対を後段に出力しているものとする。

【 0 0 5 9 】

ステップ S 1 1 において、3 D 推定部 2 1 の特徴点検出部 2 3 は、ステレオ画像対から特徴があるテクスチャ領域を検出し、テクスチャ領域から 1 以上の特徴点を検出して、検

50

出結果を示す2D特徴点群@cam0を奥行推定部24に出力する。ステップS12において、奥行推定部24は、ステレオ画像対における特徴点の3D座標を推定し、その結果を示す3D特徴点群@cam0を座標変換部25および保持部26に出力する。

【0060】

ステップS13において、座標変換部25は、3D特徴点群@cam0に基づき、ステレオ画像対における特徴点の、第1の乃至第3の画像における2D座標を推定し、その結果得られる推定2D特徴点群@cam[i] ($i=1,2,3$)を保持部26に出力する。

【0061】

この後、2Dトラッキング部27の第1の撮像部28-1乃至第3の撮像部28-3は、選択的に切り替えられて動画像の撮像を開始する。

10

【0062】

ステップS14において、トラッキング部29-i ($i=1,2,3$)は、保持部26から推定2D特徴点群@cam[i]を取得し、取得した推定2D特徴点群@cam[i]に基づいて、ステレオ画像対から検出された特徴点を、前段から得られる第iの画像上でトラッキングすることによって2D特徴点群@cam[i]を得る。次に、ステップS15において、トラッキング部29-iは、第iの画像上でトラッキングできなかった特徴点の記録を2D特徴点群@cam[i]から削除して座標変換部30-iに出力する。

【0063】

ステップS16において、座標変換部30-i ($i=1,2,3$)は、トラッキング部29-iから入力された2D特徴点群@cam[i]に基づき、第iの画像上にトラッキングした特徴点の2D座標を、ステレオ画像対における2D座標に変換し、その結果得られる2D特徴点群@cam0-iをポーズ推定部31に出力する。

20

【0064】

ステップS17において、ポーズ推定部31は、保持部26から3D特徴点群@cam0を取得し、取得した3D特徴点群@cam0と、2D特徴点群@cam0-1乃至2D特徴点群@cam0-3に基づいて、3D座標と2D座標との組み合わせが尤も適切に表現できる最終ポーズ(6DOF)を推定する。

【0065】

以上説明した情報処理装置の第2の構成例による最終ポーズ決定処理によれば、第1の撮像部28-1乃至第3の撮像部28-3のうちのいずれかが、具体的には、ステレオ画像対から検出された特徴点を継続的にトラッキングできるように、また、特徴点の空間分布が偏らないように、周辺環境に影響されない特徴点を検出できるように選択的に切り替えられるので、最終的なポーズを継続的に推定することが可能となる。

30

【0066】

<本開示を適用した情報処理装置の応用例>

上述した実施の形態では、車載用途を想定していたが、情報処理装置の第1および第2の構成例は車載用途以外への応用も可能である。

【0067】

図9は、情報処理装置の第2の構成例をHMD(ヘッドマウントディスプレイ)に搭載した場合の概念図を示している。

40

【0068】

同図に示されるHMDを用いれば、ユーザの頭部の位置と姿勢を継続的に推定することができるので、動画再生やAR/VRアプリケーションにおいて、ユーザの状態に応じた映像を提示することが可能となる。

【0069】

図10は、情報処理装置の第2の構成例を自律型掃除ロボットに搭載した場合の概念図を示している。図11は、情報処理装置の第2の構成例をドローン(自律型無人飛行機)に搭載した場合の概念図を示している。

【0070】

図10に示される自律型掃除ロボットや図11に示されるドローンによれば、継続的に

50

安定して自己のポーズを推定することができるので、安定した移動や飛行が可能となる。

【 0 0 7 1 】

図示は省略するが、該情報処理装置を小型して保護すべき野生動物などに取り付ければ、位置を継続的に推定できるので、それらの行動解析に資することが可能となる。

【 0 0 7 2 】

図 1 2 は、情報処理装置の第 2 の構成例を 3 D モデルや任意視点映像の作成に応用する場合の概念図を示している。同図に示されるように、不動の被写体（同図の場合、建物）の周囲に各撮像部を配置すれば、各撮像部の位置と姿勢を推定できるので、各撮像部によって撮像された画像を用いて 3 D モデルや任意視点映像を作成することができる。

【 0 0 7 3 】

ところで、上述した情報処理装置の第 1 および第 2 の構成例による一連の処理は、ハードウェアにより実行することもできるし、ソフトウェアにより実行することもできる。一連の処理をソフトウェアにより実行する場合には、そのソフトウェアを構成するプログラムが、コンピュータにインストールされる。ここで、コンピュータには、専用のハードウェアに組み込まれているコンピュータや、各種のプログラムをインストールすることで、各種の機能を実行することが可能な、例えば汎用のパーソナルコンピュータなどが含まれる。

【 0 0 7 4 】

図 1 3 は、上述した一連の処理をプログラムにより実行するコンピュータのハードウェアの構成例を示すブロック図である。

【 0 0 7 5 】

このコンピュータ 1 0 0 において、CPU (Central Processing Unit) 1 0 1 , ROM (Read Only Memory) 1 0 2 , RAM (Random Access Memory) 1 0 3 は、バス 1 0 4 により相互に接続されている。

【 0 0 7 6 】

バス 1 0 4 には、さらに、入出力インタフェース 1 0 5 が接続されている。入出力インタフェース 1 0 5 には、入力部 1 0 6 、出力部 1 0 7 、記憶部 1 0 8 、通信部 1 0 9 、およびドライブ 1 1 0 が接続されている。

【 0 0 7 7 】

入力部 1 0 6 は、キーボード、マウス、マイクロフォンなどよりなる。出力部 1 0 7 は、ディスプレイ、スピーカなどよりなる。記憶部 1 0 8 は、ハードディスクや不揮発性のメモリなどよりなる。通信部 1 0 9 は、ネットワークインタフェースなどよりなる。ドライブ 1 1 0 は、磁気ディスク、光ディスク、光磁気ディスク、又は半導体メモリなどのリムーバブルメディア 1 1 1 を駆動する。

【 0 0 7 8 】

以上のように構成されるコンピュータ 1 0 0 では、CPU 1 0 1 が、例えば、記憶部 1 0 8 に記憶されているプログラムを、入出力インタフェース 1 0 5 およびバス 1 0 4 を介して、RAM 1 0 3 にロードして実行することにより、上述した一連の処理が行われる。

【 0 0 7 9 】

なお、コンピュータ 1 0 0 が実行するプログラムは、本明細書で説明する順序に沿って時系列に処理が行われるプログラムであってもよいし、並列に、あるいは呼び出しが行われたとき等の必要なタイミングで処理が行われるプログラムであってもよい。

【 0 0 8 0 】

なお、本開示の実施の形態は、上述した実施の形態に限定されるものではなく、本開示の要旨を逸脱しない範囲において種々の変更が可能である。

【 0 0 8 1 】

本開示は以下のような構成も取ることができる。

(1)

移動体に搭載されたステレオカメラから構成され、所定のフレームレートで動画像を撮像することにより時系列のステレオ主画像対を出力する主撮像部と、

10

20

30

40

50

前記主撮像部とは撮像方向が異なるように前記移動体に搭載され、所定のフレームレートで動画像を撮像する１または複数の副撮像部と、

前記主撮像部によって撮像されたステレオ主画像対と、前記副撮像部によって撮像された副画像とに基づいて前記移動体のポーズを推定する最終ポーズ決定部とを備える情報処理装置。

(2)

前記副撮像部は、ステレオカメラから構成され、所定のフレームレートで動画像を撮像することにより時系列のステレオ副画像対を出力し、

前記最終ポーズ決定部は、

前記主撮像部によって撮像された時系列のステレオ主画像対に基づいて前記移動体のポーズを推定する主推定部と、

前記副撮像部によって撮像された時系列のステレオ副画像対に基づいて前記移動体のポーズを推定する副推定部と、

前記主推定部によって推定されたポーズと、前記副推定部によって推定されたポーズとを統合することにより最終ポーズを決定する統合部とを有する

前記(1)に記載の情報処理装置。

(3)

前記主推定部は、前記ステレオ主画像対から検出した特徴点の３Ｄ座標と推定するとともに、前記特徴点を前記ステレオ主画像対の一方の複数フレームに亘ってトラッキングすることによって前記移動体のポーズを推定し、推定したポーズとその信頼度を前記統合部に出力し、

前記副推定部は、前記ステレオ副画像対から検出した特徴点の３Ｄ座標と推定するとともに、前記特徴点を前記ステレオ副画像対の一方の複数フレームに亘ってトラッキングすることによって前記移動体のポーズを推定し、推定したポーズとその信頼度を前記統合部に出力し、

前記統合部は、前記信頼度に基づいて、前記主推定部によって推定されたポーズと、前記副推定部によって推定されたポーズとを統合することにより最終ポーズを決定する

前記(2)に記載の情報処理装置。

(4)

前記信頼度は、前記ポーズを推定するに際して有効な前記特徴点の数、前記特徴点の３Ｄ空間分布、推定された前記ポーズと直近の最終ポーズとの差分、または、推定における最適化演算における残差のうちの少なくとも一つを含む

前記(3)に記載の情報処理装置。

(5)

前記副撮像部は、広角カメラ、魚眼カメラ、または全方位カメラから構成され、所定のフレームレートで動画像を撮像することにより、時系列の副画像を出力し、

前記最終ポーズ決定部は、

前記主撮像部によって撮像された前記ステレオ主画像対から検出した特徴点の３Ｄ座標を推定する３Ｄ推定部と、

前記特徴点を前記副画像の複数フレームに亘ってトラッキングし、トラッキング結果を前記ステレオ主画像対における２Ｄ座標に変換する２Ｄトラッキング部と、

３Ｄ推定部によって推定された前記特徴点の前記３Ｄ座標と、前記２Ｄトラッキング部によって変換された前記特徴点の前記ステレオ主画像対における２Ｄ座標とに基づいて最終ポーズを推定するポーズ推定部とを有する

前記(1)に記載の情報処理装置。

(6)

前記３Ｄ推定部は、前記主撮像部によって撮像された前記ステレオ主画像対から検出した特徴点の３Ｄ座標を推定するとともに、前記特徴点の前記副画像における２Ｄ座標を推定し、

前記２Ｄトラッキング部は、前記３Ｄ推定部によって推定された前記特徴点の前記副画

10

20

30

40

50

像における前記 2 D 座標に基づき、前記特徴点を前記副画像の複数フレームに亘ってトラッキングし、トラッキング結果を前記ステレオ主画像対における 2 D 座標に変換する

前記 (5) に記載の情報処理装置。

(7)

複数の前記副撮像部は、前記 3 D 推定部による前記特徴点の前記副画像における前記 2 D 座標に推定結果に基づき、選択的に切り替えられて撮像を行う

前記 (5) または (6) に記載の情報処理装置。

(8)

情報処理装置による情報処理方法において、

前記情報処理装置による、

移動体に搭載されたステレオカメラから構成され、所定のフレームレートで動画を撮像することにより時系列のステレオ主画像対を出力する主撮像部によって撮像されたステレオ主画像対と、前記主撮像部とは撮像方向が異なるように前記移動体に搭載され、所定のフレームレートで動画を撮像する 1 または複数の副撮像部によって撮像された副画像とに基づいて前記移動体のポーズを推定する最終ポーズ決定ステップを

含む情報処理方法。

(9)

情報処理装置の制御用のプログラムにおいて、

移動体に搭載されたステレオカメラから構成され、所定のフレームレートで動画を撮像することにより時系列のステレオ主画像対を出力する主撮像部によって撮像されたステレオ主画像対と、前記主撮像部とは撮像方向が異なるように前記移動体に搭載され、所定のフレームレートで動画を撮像する 1 または複数の副撮像部によって撮像された副画像とに基づいて前記移動体のポーズを推定する最終ポーズ決定ステップを

含む処理を情報処理装置のコンピュータに実行させるプログラム。

【符号の説明】

【 0 0 8 2 】

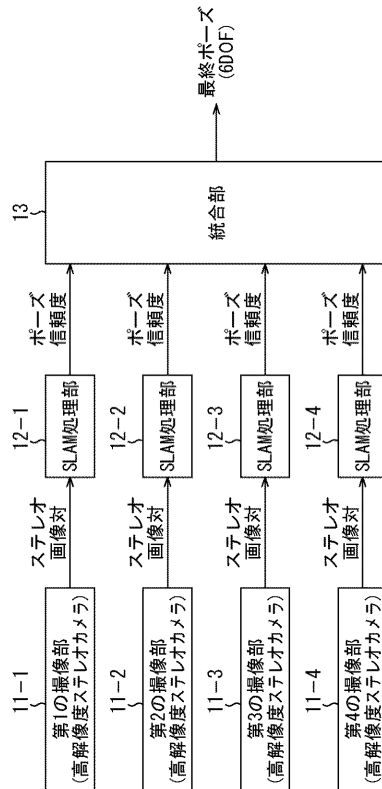
1 1 - 1 第 1 の撮像部, 1 1 - 2 第 2 の撮像部, 1 1 - 3 第 3 の撮像部,
1 1 - 4 第 4 の撮像部, 1 2 - 1 乃至 1 2 - 4 SLAM 処理部, 1 3 統合部, 2
1 3 D 推定部, 2 2 ステレオ撮像部, 2 3 特徴点検出部, 2 4 奥行推定部
, 2 5 座標変換部, 2 6 保持部, 2 7 2 D トラッキング部, 2 8 - 1 第
1 の撮像部, 2 8 - 2 第 2 の撮像部, 2 8 - 3 第 3 の撮像部, 2 9 - 1 乃至 2
9 - 3 トラッキング部, 3 0 - 1 乃至 3 0 - 3 座標変換部, 3 1 ポーズ推定部
, 1 0 0 コンピュータ, 1 0 1 CPU

10

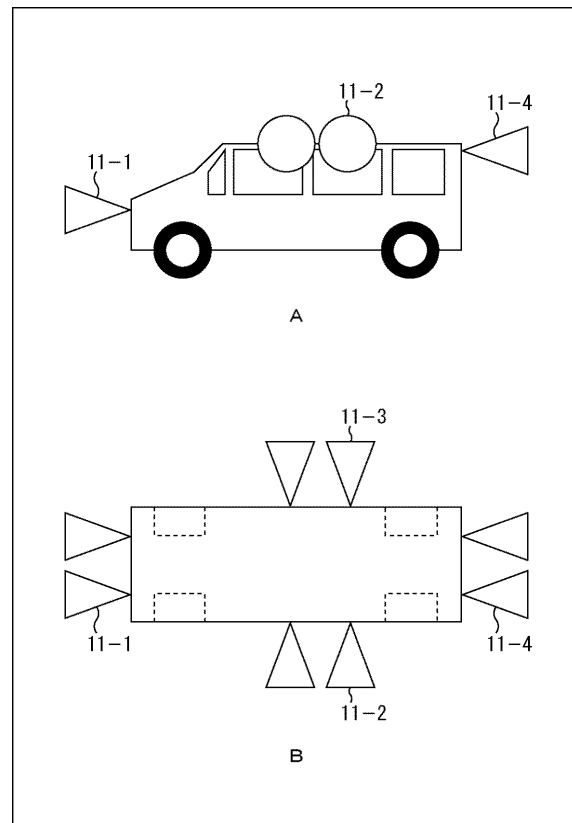
20

30

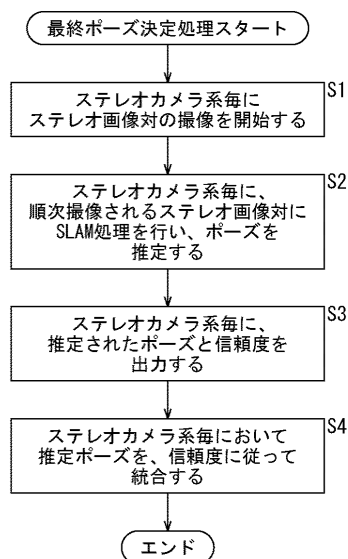
【 図 1 】
FIG. 1



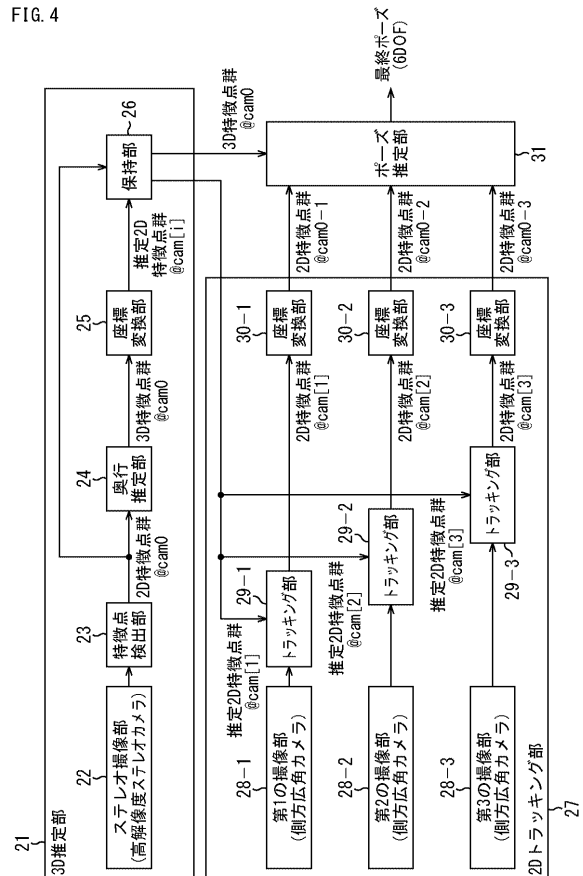
【 図 2 】
FIG. 2



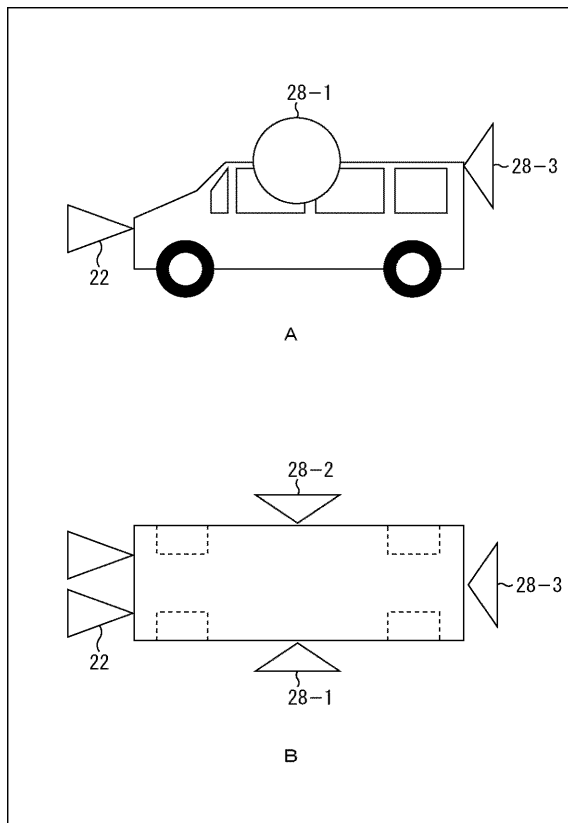
【 図 3 】
FIG. 3



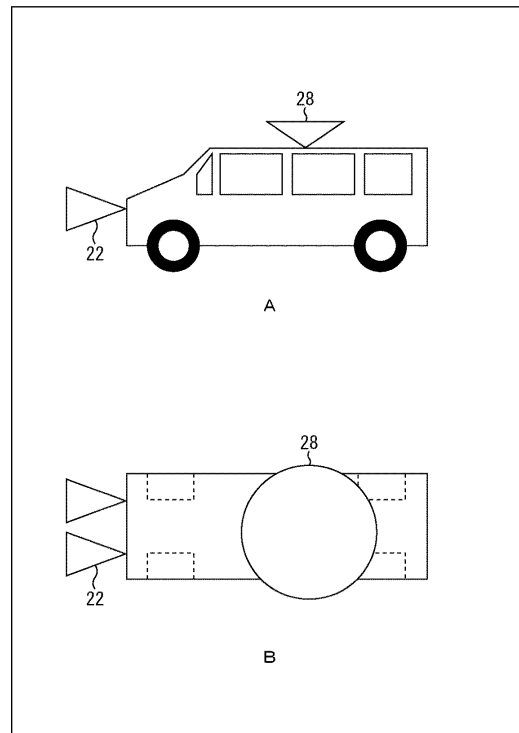
【 図 4 】
FIG. 4



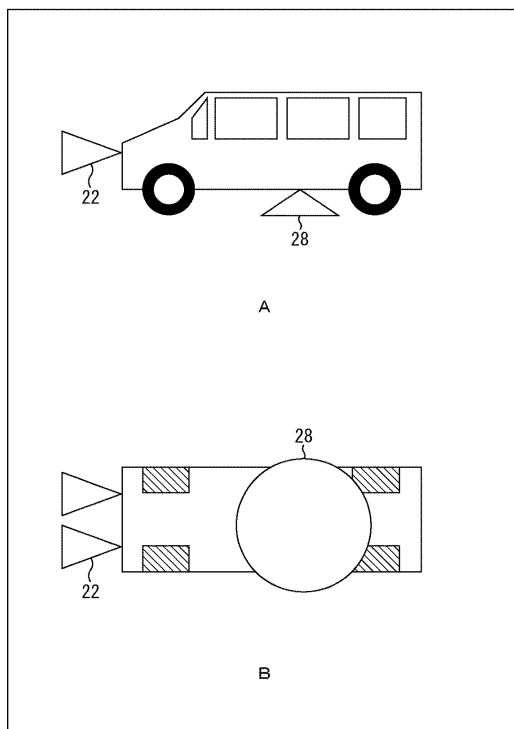
【図 5】
FIG. 5



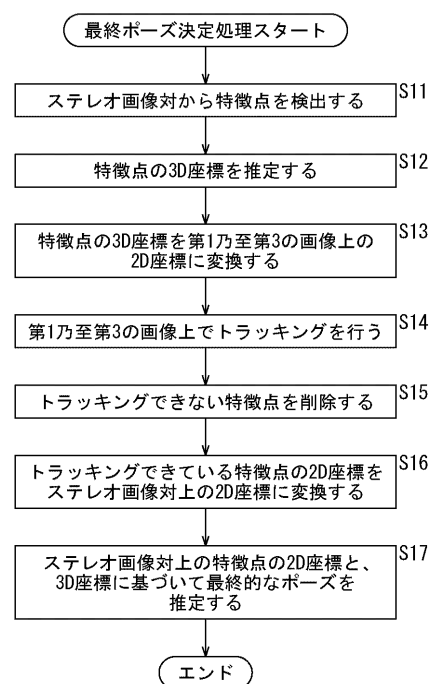
【図 6】
FIG. 6



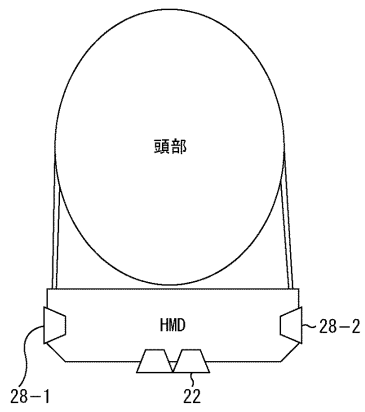
【図 7】
FIG. 7



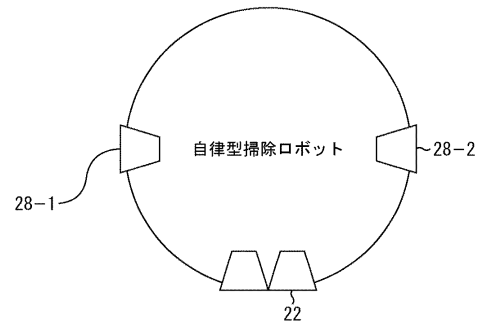
【図 8】
FIG. 8



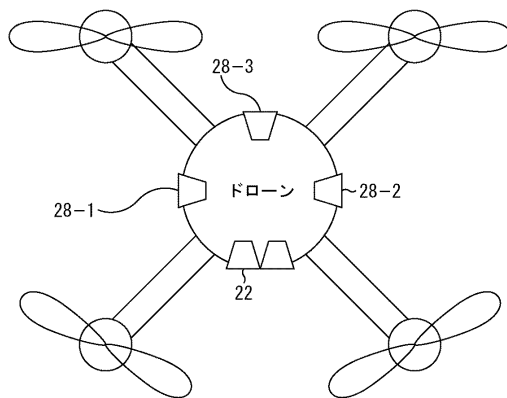
【図 9】
FIG. 9



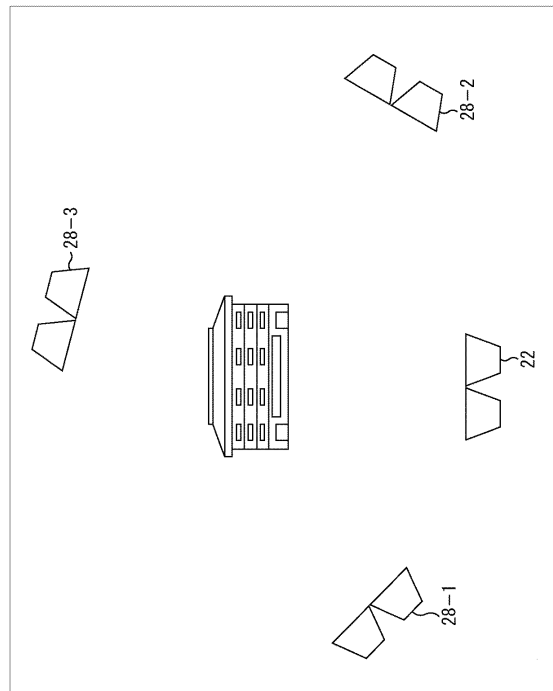
【図 10】
FIG. 10



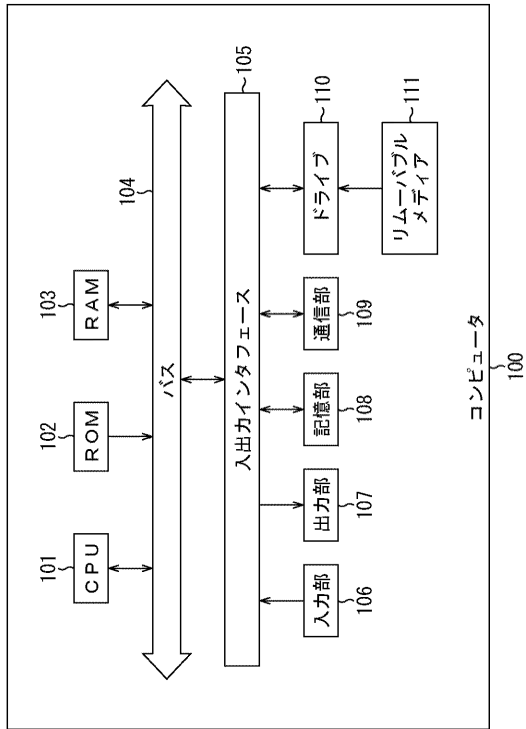
【図 11】
FIG. 11



【図 12】
FIG. 12



【図 13】
FIG. 13



フロントページの続き

- (72)発明者 江島 公志
東京都港区港南1丁目7番1号 ソニー株式会社内
- (72)発明者 加藤 高明
東京都港区港南1丁目7番1号 ソニー株式会社内
- (72)発明者 福地 正樹
東京都港区港南1丁目7番1号 ソニー株式会社内
- (72)発明者 柏谷 辰起
東京都港区港南1丁目7番1号 ソニー株式会社内
- (72)発明者 本間 俊一
東京都港区港南1丁目7番1号 ソニー株式会社内

審査官 三好 貴大

- (56)参考文献 特開2007-278871(JP, A)
国際公開第2012/176249(WO, A1)
特開2014-216813(JP, A)
特開2010-014450(JP, A)
米国特許出願公開第2013/0293714(US, A1)

(58)調査した分野(Int.Cl., DB名)

G01B 11/00 - 11/30
G06T 1/00
7/20
7/70