

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2017-191426

(P2017-191426A)

(43) 公開日 平成29年10月19日(2017.10.19)

(51) Int.Cl.			F I			テーマコード (参考)	
<b>G06F</b>	<b>3/01</b>	<b>(2006.01)</b>	G06F	3/01	570	5B087	
<b>G06T</b>	<b>7/20</b>	<b>(2017.01)</b>	G06T	7/20	300A	5E555	
<b>G06T</b>	<b>7/60</b>	<b>(2017.01)</b>	G06T	7/60	150P	5L096	
<b>G06F</b>	<b>3/038</b>	<b>(2013.01)</b>	G06F	3/038	310A		
<b>G06F</b>	<b>3/0346</b>	<b>(2013.01)</b>	G06F	3/0346	426		

審査請求 未請求 請求項の数 13 O L (全 17 頁)

(21) 出願番号 特願2016-80071 (P2016-80071)  
 (22) 出願日 平成28年4月13日 (2016.4.13)

(71) 出願人 000001007  
 キヤノン株式会社  
 東京都大田区下丸子3丁目30番2号  
 (74) 代理人 100099324  
 弁理士 鈴木 正剛  
 (72) 発明者 釜森 勇樹  
 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内  
 Fターム(参考) 5B087 AB11 AC02 AD02 BC32 DD03  
 DD09 DD10 DJ03  
 5E555 AA11 AA54 BA38 BB38 BC01  
 BE17 CA42 CB45 CB55 CB66  
 DA08 DA31 EA14 EA19 EA22  
 EA25 FA00

最終頁に続く

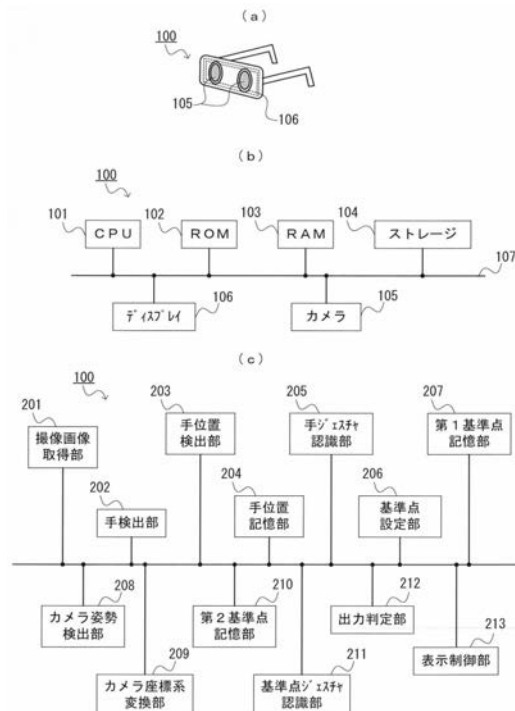
(54) 【発明の名称】 入力装置、入力制御方法、コンピュータプログラム、及び記憶媒体

(57) 【要約】

【課題】ジェスチャ操作の誤認識を抑制する入力装置を提供する。

【解決手段】HMD 100は、現実空間に固定して設定される基準点及びユーザの手を含む撮像画像を取得するカメラ105を備える。HMD 100は、カメラ105の動きから基準点の撮像画像内における位置を検出するカメラ座標系変換部209と、手の撮像画像内における位置の所定時間内の変化により、該手の撮像画像内におけるジェスチャ操作を認識する手ジェスチャ認識部205と、基準点の位置の撮像画像内における所定時間内の変化により、基準点の撮像画像内におけるジェスチャを認識する基準点ジェスチャ認識部211と、手の撮像画像内におけるジェスチャ操作と、基準点の撮像画像内におけるジェスチャとを比較し、比較結果に基づいて手によるジェスチャ操作に応じた結果を出力するか否かを判定する出力判定部212と、を備える。

【選択図】 図1



**【特許請求の範囲】****【請求項 1】**

現実空間に固定して設定される所定の基準点及び入力操作を行う指示体を含む範囲を連続して撮像する撮像手段と、

前記撮像手段で撮像された撮像画像内における前記指示体の位置を検出する第 1 位置検出手段と、

前記撮像手段の位置及び姿勢を検出する姿勢検出手段と、

前記撮像手段の位置及び姿勢から前記基準点の前記撮像画像内における位置を検出する第 2 位置検出手段と、

前記第 1 位置検出手段で連続する撮像画像から検出した前記指示体の位置の所定時間内の変化により、該指示体の前記撮像画像内における動きによる入力操作を認識する第 1 認識手段と、

前記第 2 位置検出手段で検出した前記基準点の位置の前記所定時間内の変化により、前記基準点の前記撮像画像内における動きを認識する第 2 認識手段と、

前記指示体の前記撮像画像内における動きと、前記基準点の前記撮像画像内における動きとを比較し、比較結果に基づいて前記指示体による前記入力操作に応じた結果を出力するか否かを判定する出力判定手段と、を備えることを特徴とする、

入力装置。

**【請求項 2】**

前記出力判定手段は、前記指示体の前記撮像画像内における動きと、前記基準点の前記撮像画像内における動きとが一致するときに、前記指示体による前記入力操作に応じた結果の出力を停止し、前記指示体の前記撮像画像内における動きと、前記基準点の前記撮像画像内における動きとが一致しないときに、前記指示体による前記入力操作に応じた結果の出力を許可することを特徴とする、

請求項 1 記載の入力装置。

**【請求項 3】**

前記指示体が動く軌跡の近傍に前記基準点を設定する基準点設定手段をさらに備えることを特徴とする、

請求項 1 又は 2 記載の入力装置。

**【請求項 4】**

前記基準点設定手段は、前記基準点を前記撮像手段から 50 [ c m ] の位置に設定することを特徴とする、

請求項 3 記載の入力装置。

**【請求項 5】**

前記姿勢検出手段により検出された前記撮像手段の位置及び姿勢の変化に基づいて、当該撮像手段が静止しているか否かを検出する静止検出手段をさらに備えており、

前記基準点設定手段は、前記撮像手段の静止が検出されると、前記基準点を再設定することを特徴とする、

請求項 3 又は 4 記載の入力装置。

**【請求項 6】**

前記撮像画像内における前記指示体の位置を、前記姿勢検出手段により検出された前記撮像手段の位置及び姿勢に基づいて前記現実空間の位置に変換する変換手段をさらに備えており、

前記基準点設定手段は、前記現実空間の位置に変換された前記指示体の位置に基づいて、前記基準点を設定することを特徴とする、

請求項 3 又は 4 記載の入力装置。

**【請求項 7】**

前記第 2 位置検出手段は、前記現実空間を表す第 1 座標系により設定される前記基準点の位置を、前記撮像画像に基づく第 2 座標系における位置に座標変換することで、前記基準点の前記撮像画像内における位置を検出し、

10

20

30

40

50

前記第 1 位置検出手段は、前記撮像画像に基づいて、前記指示体の前記第 2 座標系における位置を検出することを特徴とする、

請求項 1 ~ 6 のいずれか 1 項記載の入力装置。

【請求項 8】

前記第 2 認識手段は、前記基準点の前記撮像画像内における動きを、前記第 1 認識手段と同様の処理により認識することを特徴とする、

請求項 1 ~ 7 のいずれか 1 項記載の入力装置。

【請求項 9】

前記第 1 認識手段は、前記第 1 位置検出手段で連続する撮像画像から検出した前記指示体の位置の所定時間内の変化から特徴量を算出し、予め学習しておいた認識対象ジェスチャモデルとのマッチングを行うことで、該指示体の前記撮像画像内における動きによる入力操作を認識することを特徴とする、

請求項 8 記載の入力装置。

【請求項 10】

前記撮像手段はユーザの身体に装着されることを特徴とする、

請求項 1 ~ 9 のいずれか 1 項記載の入力装置。

【請求項 11】

現実空間に固定して設定される所定の基準点及び入力操作を行う指示体を含む範囲を連続して撮像する撮像手段から、撮像画像を取得する入力装置により実行される方法であって、

前記撮像手段の位置及び姿勢を検出する工程と、

検出した前記撮像手段の位置及び姿勢に応じて前記基準点の撮像画像内における位置を検出する工程と、

前記撮像手段で撮像された前記撮像画像内における前記指示体の位置を検出する工程と

、連続する撮像画像から検出した前記指示体の位置の所定時間内の変化により、該指示体の前記撮像画像内における動きによる入力操作を認識する工程と、

前記撮像画像内における前記基準点の位置の前記所定時間内の変化により、前記基準点の前記撮像画像内における動きを認識する工程と、

前記指示体の前記撮像画像内における動きと、前記基準点の前記撮像画像内における動きとを比較し、比較結果に基づいて前記指示体による前記入力操作に応じた結果を出力するか否かを判定する工程と、を含むことを特徴とする、

入力制御方法。

【請求項 12】

現実空間に固定して設定される所定の基準点及び入力操作を行う指示体を含む範囲を連続して撮像する撮像手段から撮像画像を取得するコンピュータを、

前記撮像手段で撮像された撮像画像内における前記指示体の位置を検出する第 1 位置検出手段、

前記撮像手段の位置及び姿勢を検出する姿勢検出手段、

前記撮像手段の位置及び姿勢から前記基準点の前記撮像画像内における位置を検出する第 2 位置検出手段、

前記第 1 位置検出手段で連続する撮像画像から検出した前記指示体の位置の所定時間内の変化により、該指示体の前記撮像画像内における動きによる入力操作を認識する第 1 認識手段、

前記第 2 位置検出手段で検出した前記基準点の位置の前記所定時間内の変化により、前記基準点の前記撮像画像内における動きを認識する第 2 認識手段、

前記指示体の前記撮像画像内における動きと、前記基準点の前記撮像画像内における動きとを比較し、比較結果に基づいて前記指示体による前記入力操作に応じた結果を出力するか否かを判定する出力判定手段、

として機能させるためのコンピュータプログラム。

10

20

30

40

50

**【請求項 1 3】**

請求項 1 2 記載のコンピュータプログラムを記憶するコンピュータにより読み取り可能な記憶媒体。

**【発明の詳細な説明】****【技術分野】****【0001】**

本発明は、ユーザのジェスチャを入力操作として認識する入力装置に関する。

**【背景技術】****【0002】**

ユーザを撮像装置により連続して撮像し、撮像画像からユーザのジェスチャを入力操作として認識することで、該入力操作を受け付ける入力装置がある。ジェスチャによる入力操作を「ジェスチャ操作」という。ジェスチャは、ユーザの手、指等の身体の一部、或いは指示棒等の指示体の動きである。ジェスチャ操作を用いた入力装置は、ユーザのジェスチャを正確に認識することが重要である。撮像装置がユーザの身体に装着可能なものである場合、撮像装置がユーザの動きに応じて動いてしまう。例えば、HMD (Head Mount Display) のように撮像装置と表示装置とが一体となってユーザの頭部に装着される装置では、ユーザの首の動きに応じて撮像装置の撮像範囲が動いてしまう。入力装置は、ユーザがジェスチャ操作を行っていない場合であっても、ユーザの動きに応じて撮像装置が動くことで、撮像画像内で指示体の位置が移動してしまい、これをジェスチャ操作に誤認識することがある。これはユーザの意図しないジェスチャ操作となり、操作性の低下につながる。

**【0003】**

特許文献 1 のジェスチャ入力装置は、撮像画像によりジェスチャを認識し、ユーザの身体の動きに伴う撮像装置の姿勢の変動量を求め、変動量に基づいてジェスチャに応じたコマンドを出力する技術を開示する。このジェスチャ入力装置は、撮像装置の変動量が多い場合、ジェスチャ操作に応じたコマンドの出力を制限することで、ジェスチャ操作の誤認識を抑制する。

**【先行技術文献】****【特許文献】****【0004】**

【特許文献 1】特開 2012 - 146220 号公報

**【発明の概要】****【発明が解決しようとする課題】****【0005】**

撮像装置の動きによるジェスチャ操作の誤認識の本質的な要因は、撮像装置の動き量の大小ではなく、撮像画像内の指示体が、撮像装置の動きに応じてジェスチャ操作のように動くことである。撮像画像内の指示体がジェスチャ操作のように動いているか否かは、撮像装置の動き量の大小から判断することは困難である。

**【0006】**

本発明は、上記の問題を解決するために、撮像装置の動きに起因するジェスチャ操作の誤認識を抑制する入力装置を提供することを主たる課題とする。

**【課題を解決するための手段】****【0007】**

本発明の入力装置は、現実空間に固定して設定される所定の基準点及び入力操作を行う指示体を含む範囲を連続して撮像する撮像手段と、前記撮像手段で撮像された撮像画像内における前記指示体の位置を検出する第 1 位置検出手段と、前記撮像手段の位置及び姿勢を検出する姿勢検出手段と、前記撮像手段の位置及び姿勢から前記基準点の前記撮像画像内における位置を検出する第 2 位置検出手段と、前記第 1 位置検出手段で連続する撮像画像から検出した前記指示体の位置の所定時間内の変化により、該指示体の前記撮像画像内における動きによる入力操作を認識する第 1 認識手段と、前記第 2 位置検出手段で検出し

た前記基準点の位置の前記所定時間内の変化により、前記基準点の前記撮像画像内における動きを認識する第2認識手段と、前記指示体の前記撮像画像内における動きと、前記基準点の前記撮像画像内における動きとを比較し、比較結果に基づいて前記指示体による前記入力操作に応じた結果を出力するか否かを判定する出力判定手段と、を備えることを特徴とする。

【発明の効果】

【0008】

本発明によれば、撮像画像内の基準点の動きと指示体の動きとを比較し、比較結果に応じてジェスチャ操作の誤認識を抑制することができる。

【図面の簡単な説明】

10

【0009】

【図1】(a)、(b)、(c)はHMDの構成の説明図。

【図2】(a)、(b)はジェスチャ操作の説明図。

【図3】(a)、(b)はジェスチャ操作の説明図。

【図4】(a)、(b)は基準点の説明図。

【図5】(a)、(b)は基準点の説明図。

【図6】入力制御処理を表すフローチャート。

【図7】(a)、(b)はジェスチャ認識処理の説明図。

【図8】(a)、(b)はジェスチャ認識処理の説明図。

【図9】HMDの機能構成図。

20

【図10】入力制御処理を表すフローチャート。

【図11】HMDの機能構成図。

【図12】入力制御処理を表すフローチャート。

【発明を実施するための形態】

【0010】

以下、図面を参照して実施形態を詳細に説明する。ただし、本実施形態に記載されている構成要素はあくまでも例示であり、本発明の範囲をそれらのみに限定する趣旨のものではない。

【0011】

[第1実施形態]

30

本実施形態では、入力装置としてユーザの頭部に装着されるHMDを用いる例を説明する。HMDは撮像装置及び表示装置を備えており、撮像装置で撮像した撮像画像を表示装置に表示する。入力装置は、ユーザの手、指等の身体の一部、或いは指示棒等の指示体の動き(ジェスチャ)による入力操作(ジェスチャ操作)を受け付けて、入力操作の結果を表示装置に表示する。或いは、入力装置は、受け付けた入力操作に応じたコマンドを所定の情報処理装置に送信して、該情報処理装置にコマンドに応じた処理を実行させる。なお、本実施形態の入力装置は、ユーザの身体に装着可能な撮像装置を備えた構成であれば、表示装置を備える必要はない。この場合、表示装置は撮像装置と分離して設けられる。撮像装置に分離して設けられる表示装置は、例えばプロジェクタである。

【0012】

40

(構成)

図1は、HMDの構成の説明図である。図1(a)は、HMD100の外観図である。HMD100は眼鏡型であり、ユーザの頭部に装着されて指示体によるジェスチャ操作の認識を行う。本実施形態では、指示体としてユーザの手を用いる例を説明する。ジェスチャ操作の認識のために、HMD100は、装着時にユーザの視界を撮像範囲に含むような位置に、撮像装置であるカメラ105を備える。カメラ105は、ステレオカメラであるが、赤外線距離カメラであってもよい。またHMD100は、カメラ105で撮像した画像を表示するディスプレイ106を備える。HMD100は、例えばジェスチャ操作に応じた操作結果を、ディスプレイ106に表示する。

【0013】

50

図1(b)は、HMD100のハードウェア構成図である。HMD100は、CPU(Central Processing Unit)101、ROM(Read Only Memory)102、RAM(Random Access Memory)103、及びストレージ104を備える情報処理装置である。CPU101、ROM102、RAM103、ストレージ104、カメラ105、及びディスプレイ106は、システムバス107を介して相互に通信可能に接続される。CPU101は、ROM102及びストレージ104に格納されたコンピュータプログラムを読み込み、RAM103を作業領域に用いて実行することで、HMD100の動作を制御する。ROM102は、例えば起動時にHMD100の各ハードウェアの起動処理を行うための起動プログラムを格納する。ストレージ104は、例えばHMD100による各種処理を行うための制御プログラムを格納する。ストレージ104は、ハードディスクや外部記憶媒体である。CPU101は、コンピュータプログラムの実行により、カメラ105及びディスプレイ106の動作を制御する。カメラ105は連続して撮像を行い、撮像した連続する撮像画像、例えば動画をCPU101に送信する。

10

#### 【0014】

本実施形態のHMD100は、カメラ105で撮像した撮像画像に基づくジェスチャ操作に応じた認識結果をディスプレイ106に表示する場合について説明する。図1(c)は、このような処理を行うためのHMD100の機能構成図である。HMD100は、CPU101がコンピュータプログラムを実行することで、各機能を実現するが、少なくとも一部をハードウェアにより実現してもよい。HMD100は、撮像画像取得部201、出力判定部212、及び表示制御部213として機能する。HMD100は、ジェスチャ操作を認識するための手検出部202、手位置検出部203、手位置記憶部204、及び手ジェスチャ認識部205として機能する。HMD100は、現実空間を表す世界座標系におけるカメラ105の動きの判断基準となる基準点を定めるための基準点設定部206及び第1基準点記憶部207として機能する。HMD100は、カメラ105の現実空間における動きを検出するためのカメラ姿勢検出部208、カメラ座標系変換部209、第2基準点記憶部210、及び基準点ジェスチャ認識部211として機能する。

20

#### 【0015】

撮像画像取得部201は、カメラ105により撮像された、指示体となるユーザの手及び基準点を含む範囲の撮像画像を取得する。カメラ105が連続して撮像するため、撮像画像取得部201は連続して撮像画像を取得する。

30

手検出部202は、撮像画像取得部201により取得された撮像画像から、ユーザの手領域を検出する。手検出部202は、撮像画像中の肌色領域を検出することで、手領域の検出を行う。手検出部202は、肌色領域、且つHMD100(カメラ105)からの距離が閾値以内にある領域のみを手領域として検出してもよい。閾値は、人間の手の長さに応じて決められる。これにより、ユーザの手以外の肌色領域を手領域として誤検出することが抑制される。

#### 【0016】

手位置検出部203は、手検出部202より検出された手領域の、撮像画像内における位置を検出する。撮像画像内における位置は、撮像画像に基づく座標系(以下、「カメラ座標系」という。)により表される。手領域の位置は、一般的に、手領域を構成する各ピクセルの重心座標で表される。この他に、例えば手領域を構成する各ピクセルの内、手領域の輪郭への最短距離が最長のピクセルを手領域の位置としてもよい。この場合、手の平の中心付近が手の位置となる。手位置記憶部204は、手位置検出部203より検出された手領域のカメラ座標系における座標(位置)を記憶する。撮像画像が連続して取得されるために、手位置記憶部204は、手領域の位置を該手領域の検出時刻に関連付けて記憶する。

40

#### 【0017】

手ジェスチャ認識部205は、手位置記憶部204に記憶された手領域の位置の所定時間内の時系列データに基づいて、手の動きがジェスチャ操作であるか否かを認識する。手ジェスチャ認識部205は、連続する撮像画像内の手の位置の変化により、手の動きを認

50

識する。手ジェスチャ認識部 205 は、例えば手領域の位置及び時間経過による手領域の位置の変化から手の速度、加速度等の特徴量を算出する。手ジェスチャ認識部 205 は、算出した特徴量と、予め学習しておいた認識対象ジェスチャモデルとのマッチングを行い、マッチングの尤度が大きい場合に該ジェスチャ操作を認識する。手ジェスチャ認識部 205 は、認識対象ジェスチャを HMM (Hidden Markov Model) でモデル化し、マッチング時の最尤経路探索にビタビアルゴリズムを用いる。この他に手ジェスチャ認識部 205 は、決定木等のルールベースによる認識アルゴリズムを用いてマッチングを行ってもよい。

#### 【0018】

基準点設定部 206 は、世界座標系内の所定の位置を基準点として設定する。基準点は、世界座標系において固定されており、移動しない。第 1 基準点記憶部 207 は、基準点設定部 206 より設定された基準点の世界座標系における座標 (位置) を記憶する。「基準点」の詳細は後述する。

10

#### 【0019】

カメラ姿勢検出部 208 は、カメラ 105 の位置、姿勢を検出する。カメラ姿勢検出部 208 は、例えば、撮像画像取得部 201 より取得された連続する撮像画像中の動きベクトルや、ユーザの周辺の世界座標系に設けられる位置姿勢検出用マーカを利用して、センサを用いずにカメラ 105 の位置及び姿勢を検出可能である。また、カメラ姿勢検出部 208 は、HMD 100 に赤外線ビーコンや加速度センサ、ジャイロセンサ等のセンサを別途追加することにより、カメラ 105 の位置及び姿勢を高精度に検出することも可能である。

20

#### 【0020】

カメラ座標系変換部 209 は、第 1 基準点記憶部 207 に記憶された基準点の世界座標系における座標をカメラ座標系に座標変換する。つまりカメラ座標系変換部 209 は、世界座標系に設定された基準点の位置をカメラ座標系内の位置に変換することで、基準点の撮像画像内の位置を検出する。カメラ座標系は、カメラ 105 により撮像された撮像画像内の座標系であり、カメラ姿勢検出部 208 により検出されたカメラ 105 の位置及び姿勢に基づいて決定される。カメラ座標系変換部 209 は、例えばカメラ 105 の焦点距離、主点座標、歪み係数に加え、カメラ 105 の動きに基づく平行移動パラメータと回転パラメータとに基づいて、基準点の世界座標系における座標のカメラ座標系への座標変換を行う。第 2 基準点記憶部 210 は、カメラ座標系に変換された基準点の座標 (位置) を記憶する。撮像画像が連続して取得されるために、カメラ姿勢検出部 208 が連続してカメラ 105 の位置及び姿勢を検出する。そのためにカメラ座標系変換部 209 は、カメラ座標系を連続して決定し、基準点の世界座標系における座標をカメラ座標系に座標変換する。第 2 基準点記憶部 210 は、カメラ座標系に座標変換された基準点の位置を該座標の決定時刻に関連付けて記憶する。

30

#### 【0021】

基準点ジェスチャ認識部 211 は、第 2 基準点記憶部 210 に記憶された基準点のカメラ座標系における位置の所定時間内の時系列データに基づいて、基準点の動きをジェスチャ操作として認識する。基準点ジェスチャ認識部 211 は、連続する撮像画像内の基準点の位置の変化により、基準点の動きを認識する。基準点ジェスチャ認識部 211 は、手ジェスチャ認識部 205 のジェスチャ認識処理においてジェスチャ実行中であるとみなされた時間区間 (以下、「ジェスチャ実行区間」という。) と同じ時間区間の基準点の位置の時系列データを用いて処理を行う。また、基準点ジェスチャ認識部 211 は、基準点のカメラ座標系における位置の時系列データを、手の位置の時系列データであると仮定した上でマッチングを行い、基準点のジェスチャを認識する。これは、基準点と同様の動きが手によって行われた場合に、その動きがジェスチャとして認識されるか否かを判定するためである。ただし、基準点の動きは、カメラ 105 の動きの変化に応じた撮像画像中の動きである。基準点ジェスチャ認識部 211 の認識処理は、手ジェスチャ認識部 205 の認識処理と同様のアルゴリズムにより可能である。基準点ジェスチャ認識部 211 は、基準点

40

50

のカメラ座標系における位置及びそれにかかる速度、加速度以外の特徴量が認識処理に必要となる場合、手ジェスチャ認識部 205 が認識処理で用いた手の特徴量を用いる。或いは基準点ジェスチャ認識部 211 は、基準点から取得できない特徴量に基づく尤度計算部分を除外したアルゴリズムを用いて認識処理を行う。

#### 【0022】

出力判定部 212 は、手ジェスチャ認識部 205 による手のジェスチャ操作の認識結果と、基準点ジェスチャ認識部 211 による基準点のジェスチャの認識結果とを比較し、比較結果に基づいて手のジェスチャ操作の認識結果を出力するか否かを判定する。表示制御部 213 は、ディスプレイ 106 に、手のジェスチャ操作の認識結果を表示する。

#### 【0023】

(ジェスチャ操作)

図 2、図 3 は、HMD 100 を装着したユーザによるジェスチャ操作の説明図である。図 2 では、ユーザが右手を右から左へ動かすジェスチャ操作を行っている。図 3 では、ユーザがジェスチャ操作を行っておらず、首を左から右へ振っている。図 2 (a)、図 3 (a) は HMD 100 を装着したユーザの俯瞰図であり、図 2 (b)、図 3 (b) はカメラ 105 による撮像画像の例示図である。

#### 【0024】

図 2 に示す通り、ユーザが右手を右から左へ移動させることで、HMD 100 は、撮像画像中の手の動き(位置の時系列データ)により、ユーザが手を右から左へ動かすジェスチャ操作を行ったことを認識する。図 3 に示す通り、ユーザが手を動かさずに首を左から右へ振った場合、撮像画像中の手が右から左へ移動する(図 3 (b))。これは、図 2 (b)の手の動きと同じである。そのために HMD 100 は、このままではユーザがジェスチャ操作を行っていないにも関わらず、ユーザが手を右から左へ動かすジェスチャ操作を行ったと誤認識してしまう。本実施形態の HMD 100 は、このような誤認識を抑制する。

#### 【0025】

(基準点)

図 4、図 5 は、ジェスチャ操作の誤認識を抑制するために用いる基準点の説明図である。本実施形態では、基準点のカメラ座標系内の動きにより、手によるジェスチャ操作の認識結果を出力するか否かを判定する。そのために基準点の適切な設定は、正確なジェスチャ操作の認識のために重要である。図 4 では、図 2 と同様にユーザが右手を右から左へ動かすジェスチャ操作を行っている。図 5 では、図 3 と同様に、ユーザがジェスチャ操作を行わず、首を左から右へ振っている。図 4 (a)、図 5 (a) は HMD 100 を装着したユーザの俯瞰図であり、図 4 (b)、図 5 (b) はカメラ 105 による撮像画像の例示図である。

#### 【0026】

図 4 (a) では、世界座標系におけるユーザの正面に基準点 401 が設定される。ユーザが首を動かさずに手を右から左へ振った場合、カメラ 105 は、手が右から左へ移動する画像を撮像する(図 4 (b))。HMD 100 の動きに起因する撮像画像中(カメラ座標系)の基準点 401 の動きは、カメラ 105 が動かないために、手の動きと大きく異なる。そのための基準点 401 の動きは、右から左へ振るジェスチャとは認識されない。これは、撮像画像中における手の動きが、HMD 100 の動きに起因するものではないことを表す。よって、この場合、HMD 100 は、手が右から左へ振られたというジェスチャ操作の認識結果の出力を許可する。

#### 【0027】

図 5 (a) では、図 4 (a) と同様の位置に基準点 401 が設定される。ユーザが手を動かさずに首を左から右へ振った場合、カメラ 105 は、手が右から左へ移動する画像を撮像する(図 5 (b))。HMD 100 の動きに起因する撮像画像中(カメラ座標系)の基準点 401 の動きは、手の動きに類似しており、撮像画像中を右から左へ振るジェスチャとして認識される。これは、撮像画像中における手のジェスチャが、HMD 100 の動

10

20

30

40

50

きに起因するものであることを表す。よって、この場合、HMD 100は、手が右から左へ振られたというジェスチャ操作の認識結果の出力を停止する。

#### 【0028】

基準点401は、ユーザがジェスチャ操作を行う際に手が動く軌跡の近傍に設定されることで、ジェスチャ認識の判定精度が向上する。世界座標系におけるユーザの手の位置と基準点401の位置とが近いほど、カメラ105の動きに起因する手の動きと基準点401の動きとが類似するためである。例えば基準点401は、ジェスチャ認識の処理開始時におけるHMD 100（カメラ105）の撮像範囲内で、HMD 100（カメラ105）からの距離が50 [cm]の位置に設定されることが好ましい。逆に、基準点401がHMD 100（カメラ105）から遠く離れた位置やユーザの背後に設定された場合、カメラ105の動きに起因する基準点401の動きと手の動きとが乖離し、ジェスチャ操作の認識結果の出力の判定精度が下がることになる。基準点401は、カメラ105による撮像画像中に実際には表示されない、世界座標系中の仮想の点である。また、基準点401として体積を有する仮想的な物体を設定してもよい。

10

#### 【0029】

（ジェスチャ認識処理）

図6は、このようなHMD 100によるユーザのジェスチャ操作による入力制御処理を表すフローチャートである。この処理は、ユーザがHMD 100を頭部に装着してジェスチャ操作を行う際に行われる。処理の開始時に、カメラ105は撮像を開始しており、CPU 101に連続して撮像画像を入力している。

20

#### 【0030】

HMD 100は、基準点設定部206により基準点401の世界座標系における位置を設定し、この位置の世界座標系における座標を第1基準点記憶部207に記憶させる（S301）。カメラ姿勢検出部208は、撮像画像取得部201により取得された撮像画像に基づいて、カメラ105の世界座標系における位置、姿勢を検出する（S302）。カメラ座標系変換部209は、第1基準点記憶部207に記憶された基準点401の世界座標系における座標を、カメラ105の位置、姿勢に基づくカメラ座標系に座標変換し、第2基準点記憶部210に記憶する（S303）。カメラ105が連続して撮像を行うため、第2基準点記憶部210は、時系列に応じて基準点401のカメラ座標系における位置を表す時系列データを記憶する。

30

#### 【0031】

手検出部202は、撮像画像取得部201により取得された撮像画像に基づいて、ユーザの手領域を検出する（S304）。手位置検出部203は、手検出部202により検出された手領域のカメラ座標系における位置を検出して、手位置記憶部204に記憶させる（S305）。カメラ105が連続して撮像を行うため、手位置記憶部204は、時系列に応じて手領域のカメラ座標系における位置を表す時系列データを記憶する。手ジェスチャ認識部205は、手位置記憶部204に記憶された手領域の位置の時系列データに基づいて、ユーザの手の動きがジェスチャ操作を行っているか否か認識する（S306）。手ジェスチャ認識部205がジェスチャ操作が行われていないと判定した場合（S307：N）、HMD 100は、次の撮像画像を取得したか否かの判定を行う（S312）。

40

#### 【0032】

ジェスチャ操作が行われたと認識された場合（S307：Y）、基準点ジェスチャ認識部211は、第2基準点記憶部210に記憶されたカメラ座標系の基準点の位置を表す時系列データに基づいて、基準点の動きによるジェスチャを認識する（S308）。基準点ジェスチャ認識部211は、S306の処理におけるジェスチャ実行区間と同じ時間区間のカメラ座標系の基準点の位置の時系列データを使用して、基準点によるジェスチャを認識する。なお、基準点401として体積のある物体を設定した場合、基準点ジェスチャ認識部211は、物体の大きさを特徴量としてジェスチャ認識処理に用いてもよい。

#### 【0033】

基準点によるジェスチャと、手によるジェスチャ操作との認識結果が一致する場合（S

50

309:Y)、出力判定部212は、手によるジェスチャ操作の認識結果の出力を停止する(S310)。基準点によるジェスチャと手によるジェスチャ操作とが一致するという事は、カメラ105の動きに起因する基準点401のカメラ座標系における動きが、ジェスチャ操作時の手の動きに類似することを意味する(図5(b)参照)。つまり、ユーザが手を動かさなくてもジェスチャ操作として認識されてしまう撮像画像が、カメラ105の動きにより撮像されたと解釈できる。よって、HMD100は、ジェスチャ操作の認識が誤りであると判定して、表示制御部213へのジェスチャ操作の認識結果の出力を停止する。

#### 【0034】

基準点によるジェスチャと、手によるジェスチャ操作との認識結果が一致しない場合(S309:N)、出力判定部212は、手によるジェスチャ操作の認識結果の出力を許可する(S311)。表示制御部213は、ジェスチャ操作の認識結果に応じた画像を表示する。基準点によるジェスチャと手のジェスチャ操作とが一致しないということは、カメラ105の動きに起因する基準点のカメラ座標系における動きが、ジェスチャ操作時の手の動きとは大きく異なることを表す(図4(b)参照)。つまり、カメラ105の動きが、ユーザが手を動かさなければジェスチャ操作として認識されないものであったと解釈できる。よって、HMD100は、ジェスチャ操作の認識結果がユーザの意図通りであると判定し、表示部123へのジェスチャ操作の認識結果の出力を許可する。

#### 【0035】

HMD100のCPU101は、カメラ105から次の撮像画像を取得したか否かの判定を行う(S312)。次の撮像画像を取得した場合(S312:Y)、CPU101はS302以降の処理を繰り返し行う。次の撮像画像を取得しない場合(S312:N)、CPU101は処理を終了する。

#### 【0036】

以上のような本実施形態のHMD100は、世界座標系に設定される基準点により、カメラ105の動きに起因する、手によるジェスチャ操作の誤認識を抑制する。そのためにHMD100は、ユーザの身体に装着されるカメラ105が動いた際の、ユーザの意図しないジェスチャ操作を防止することができる。特にHMD100は、ジェスチャ操作の誤認識の本質的な要因である、カメラ105の動きに起因する撮像画像内の手の動きが如何にジェスチャ操作に類似しているかを、既存のジェスチャ操作の認識アルゴリズムに基づいて判定することが可能となる。

#### 【0037】

##### [第2実施形態]

基準点は、カメラ105の撮像範囲に応じて適切に設定されることが好ましいが、カメラ105を装着したユーザの動きによっては、一度設定した基準点が適切な位置から外れることがある。図7、図8は、基準点401が適切な位置から外れた場合のジェスチャ認識処理の説明図である。図7は、基準点401がHMD100(カメラ105)から大きく離れた位置に設定される。図8では、ユーザの移動により基準点401がユーザの后背に位置し、カメラ105の撮像範囲から外れている。図7(a)、図8(a)はHMD100を装着したユーザの俯瞰図であり、図7(b)、図8(b)はカメラ105による撮像画像の例示図である。いずれの場合も、基準点401が適切な位置から外れているために、カメラ105の動きに応じた基準点401の動きとユーザの手の動きとが乖離し、正確なジェスチャ操作の認識が困難である。

#### 【0038】

そのために第2実施形態のHMD100は、撮像装置(カメラ105)を装着したユーザの位置、姿勢等の動きに伴い、基準点を再設定する。具体的には、HMD100は、カメラ105が静止したタイミングで、基準点の再設定を行う。HMD100の外観、ハードウェア構成は、第1実施形態のHMD100と同様であるため説明を省略する。

#### 【0039】

図9は、第2実施形態のHMD100の機能構成図である。第2実施形態のHMD100

10

20

30

40

50

0の機能構成は、図1(c)に示す第1実施形態のHMD100の機能構成に、カメラ静止検出部501を追加した構成である。同様の機能については説明を省略する。カメラ静止検出部501は、カメラ姿勢検出部208により検出されたカメラ105の位置及び姿勢の変化に基づいて、カメラ105が静止しているか否かを検出する。

【0040】

図10は、第2実施形態のHMD100によるユーザのジェスチャ操作による入力制御処理を表すフローチャートである。図6に示す第1実施形態のフローチャートと同じ処理には同じステップ番号を付してある。

【0041】

基準点の設定後にカメラ105の動きを検出したHMD100は、カメラ静止検出部501により、検出されたカメラ105の動きが静止しているか否かを判定する(S301、S302、S1001)。静止判定は、カメラ105の世界座標系における動き量、例えばカメラ105の位置や角度の変動量と所定の閾値との比較により行われる。カメラ静止検出部501は、動き量が閾値以下の場合にカメラ105が静止していると判定する。静止していない場合(S1001:N)、HMD100は、S303以降の処理を行う。静止している場合(S1001:Y)、基準点設定部206は世界座標系における新たな基準点を設定して、第1基準点記憶部207に記憶させる(S1002)。新たな基準点の設定が終了すると、HMD100は、S303以降の処理を行う。

【0042】

このように第2実施形態のHMD100は、カメラ105の動きが静止する度に新たな基準点を設定し、基準点が最適な位置から外れることを防止する。これにより基準点は、ユーザがジェスチャを行う位置の近傍に常に設定される。そのために第2実施形態のHMD100は、第1実施形態のHMD100の効果の他に、ユーザの動きが大きい場合であっても、HMD100は、カメラ105の動きに起因する、手によるジェスチャ操作の誤認識を抑制するという効果を持つ。

【0043】

[第3実施形態]

第1、第2実施形態では、手によるジェスチャ操作の認識前に基準点を設定しているが、基準点は、ジェスチャ操作の認識後に設定されてもよい。第3実施形態のHMD100は、手によるジェスチャ操作の実行中に、基準点を設定する。ジェスチャ操作の実行中であるために、基準点が、手の位置に基づく適切な位置に設定可能である。

【0044】

HMD100の外観、ハードウェア構成は、第1実施形態のHMD100と同様であるため説明を省略する。図11は、第3実施形態のHMD100の機能構成図である。第3実施形態のHMD100の機能構成は、図1(c)に示す第1実施形態のHMD100の機能構成に、世界座標系変換部601及びカメラ姿勢記憶部602を追加した構成である。同様の機能については説明を省略する。

【0045】

世界座標系変換部601は、カメラ姿勢検出部208より検出されたカメラ105の位置及び姿勢に基づいて、手位置検出部203より検出されたカメラ座標系における手の位置を、世界座標系に座標変換する。カメラ姿勢記憶部602は、カメラ姿勢検出部208より検出されたカメラ105の位置、姿勢等の動きを記憶する。

【0046】

図12は、第3実施形態のHMD100によるユーザのジェスチャ操作による入力制御処理を表すフローチャートである。図6に示す第1実施形態のフローチャートと同じ処理には同じステップ番号を付してある。この処理は、ユーザがHMD100を頭部に装着してジェスチャ操作を行う際に行われる。処理の開始時に、カメラ105は撮像を開始しており、CPU101に連続して撮像画像を入力している。

【0047】

カメラ姿勢検出部208は、撮像画像取得部201により取得された撮像画像に基づい

10

20

30

40

50

て、カメラ105の位置、姿勢等の動きを検出して、カメラ姿勢記憶部602に記憶させる(S1201)。手検出部202は、撮像画像取得部201により取得された撮像画像に基づいて、ユーザの手領域を検出する(S304)。手位置検出部203は、手検出部202により検出された手領域のカメラ座標系における位置を検出して、手位置記憶部204に記憶させる(S305)。カメラ105が連続して撮像を行うため、手位置記憶部204は、時系列に応じて手領域のカメラ座標系における位置を表す時系列データを記憶する。

#### 【0048】

世界座標系変換部601は、手位置記憶部204に記憶された手領域のカメラ座標系における位置を、カメラ姿勢記憶部602に記憶されたカメラ105の位置、姿勢(動き)に基づいて世界座標系に座標変換して、手位置記憶部204に記憶させる(S1202)。

10

#### 【0049】

手ジェスチャ認識部205は、手位置記憶部204に記憶された手領域のカメラ座標系における位置の時系列データに基づいて、ユーザの手の動きがジェスチャ操作を行っているか否か認識する(S306)。手ジェスチャ認識部205がジェスチャ操作が行われていないと判定した場合(S307:N)、HMD100は、次の撮像画像を取得したか否かの判定を行う(S312)。

#### 【0050】

ジェスチャ操作が行われたと認識された場合(S307:Y)、基準点設定部206は、手位置記憶部204に記憶された手領域の世界座標系における位置に基づいて、基準点の世界座標系における位置を設定する。基準点設定部206は、設定した基準点の世界座標系における位置を、第2基準点記憶部210に記憶させる(S1203)。ここで基準点設定部206は、直前に行われるS306のジェスチャ操作の認識処理において認識されたジェスチャ実行区間における手領域の平均位置を、基準点の位置とする。ジェスチャ操作の認識後に基準点の位置を設定することで、基準点をジェスチャ操作時の手の近くに配置することができる。

20

#### 【0051】

カメラ座標系変換部209は、カメラ姿勢記憶部602に記憶されたカメラ105の位置、姿勢に基づいて、S1203で設定された基準点の世界座標系における座標(位置)をジェスチャ実行区間におけるカメラ座標系の座標に変換する(S1204)。カメラ座標系変換部209は、カメラ座標系の座標に変換した基準点の座標(位置)を第2基準点記憶部210に記憶させる。カメラ座標系における基準点の座標(位置)の第2基準点記憶部210への記憶後に、HMD100は、S308以降の処理を行う。

30

#### 【0052】

このように第3実施形態のHMD100は、手によるジェスチャ操作を認識した後に基準点を設定し、カメラ座標系における基準点の位置がジェスチャ実行区間においてどのように変化したかを過去に遡って逆算する。これにより第3実施形態のHMD100は、第1実施形態のHMD100の効果の他に、基準点と手の位置との乖離を抑制して、カメラ105の動きに起因する手によるジェスチャ操作の誤認識を抑制するという効果を持つ。

40

#### 【0053】

第1~第3実施形態では、基準点を1つ設定する場合について説明したが、基準点は複数設けられてもよい。基準点を複数設ける場合、手の位置に最も近接する1以上の基準点のカメラ座標系における動きに応じてジェスチャ操作の判定を行うことで、基準点と手の位置の乖離を抑制することができる。また、撮像装置(カメラ105)は、ユーザの身体に装着するものの他に、ハンディカメラや車載カメラ等に対する適応も可能である。

#### 【0054】

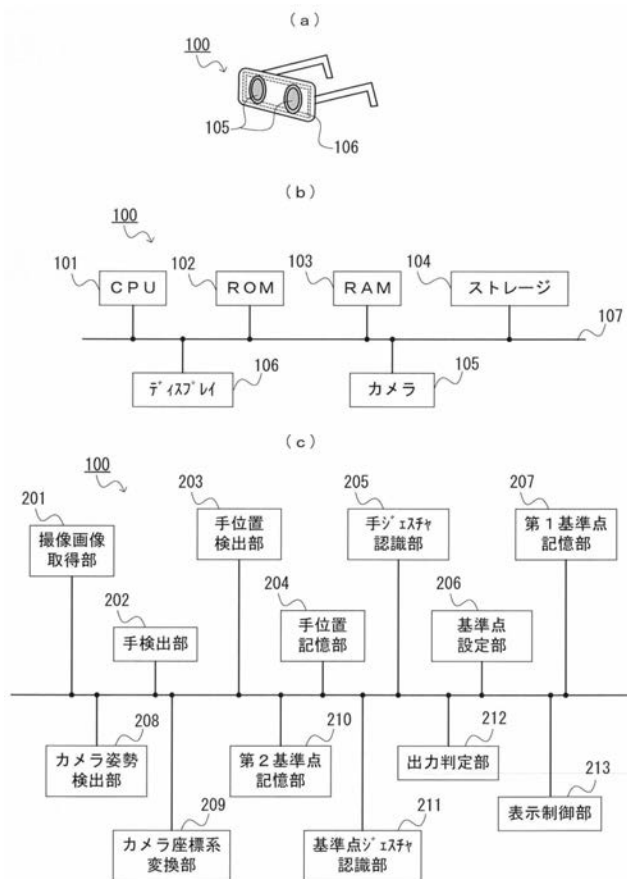
##### [その他の実施形態]

また、本発明は、以下の処理を実行することによっても実現される。即ち、上述した実施形態の機能を実現するソフトウェア(プログラム)を、ネットワーク又は各種記憶媒体を

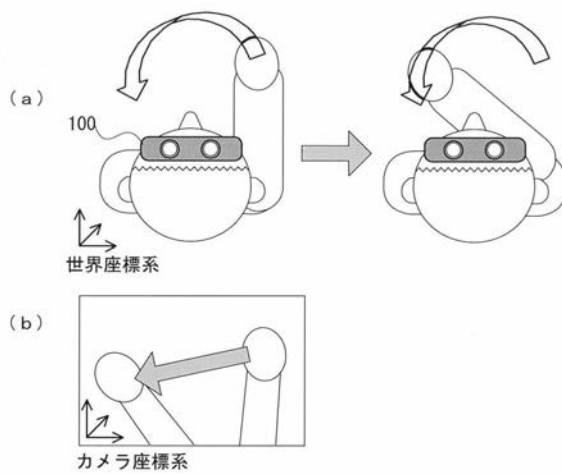
50

介してシステム或いは装置に供給し、そのシステム或いは装置のコンピュータ（又はCPU、MPU等）がプログラムを読み出して実行する処理である。

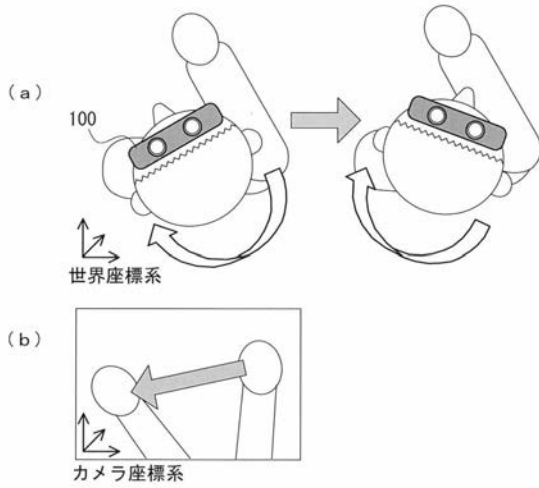
【図1】



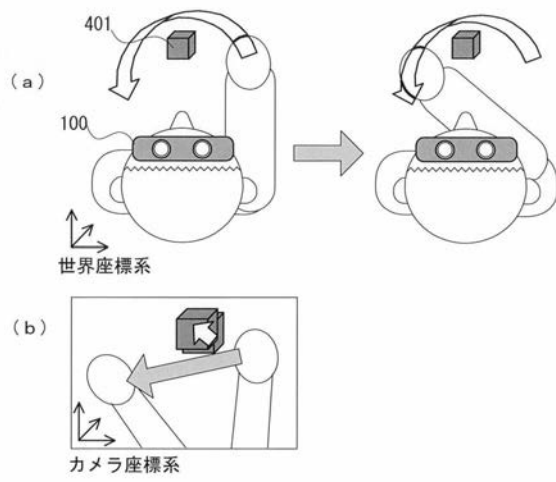
【図2】



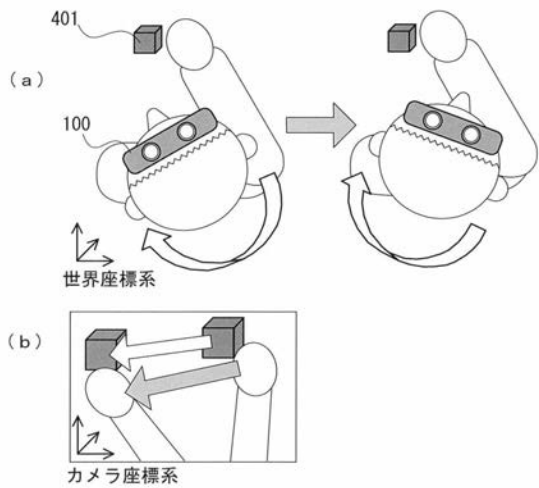
【 図 3 】



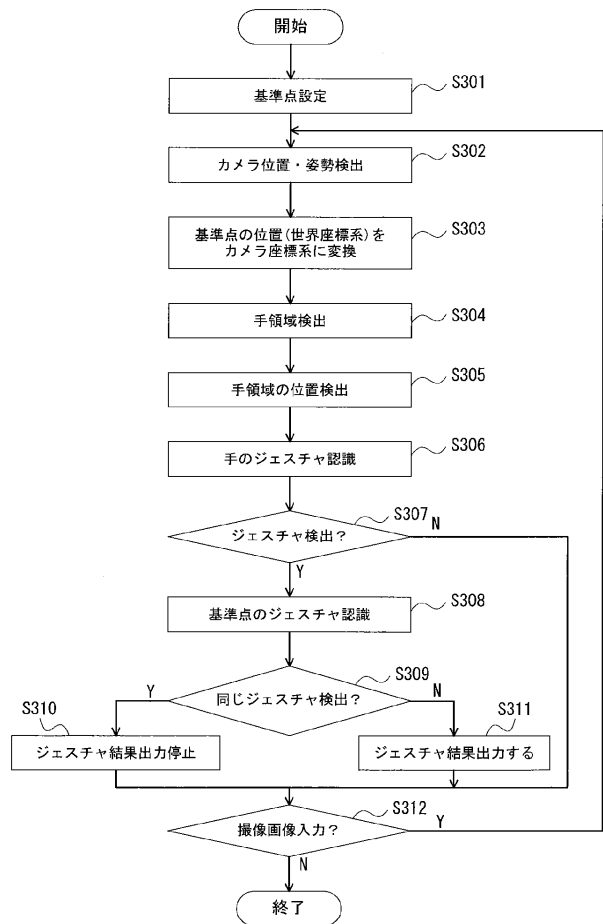
【 図 4 】



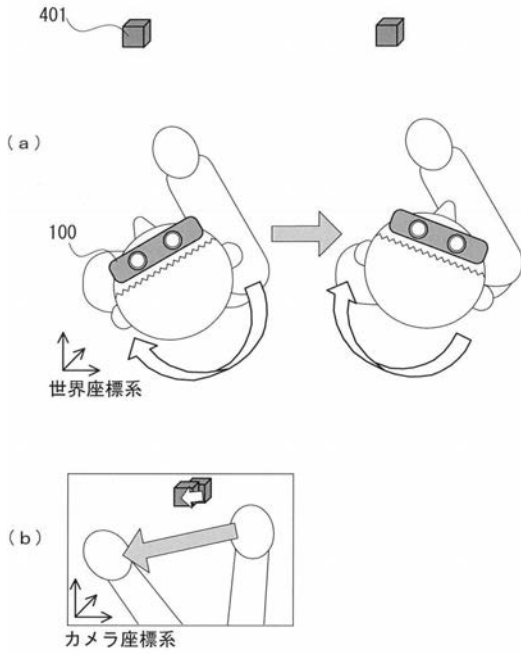
【 図 5 】



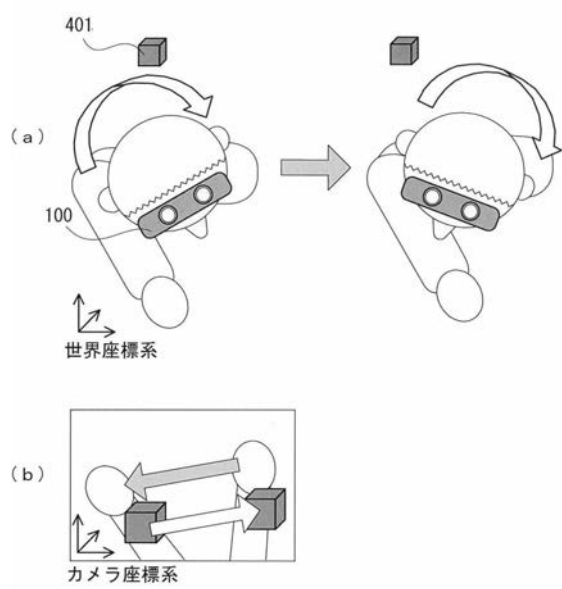
【 図 6 】



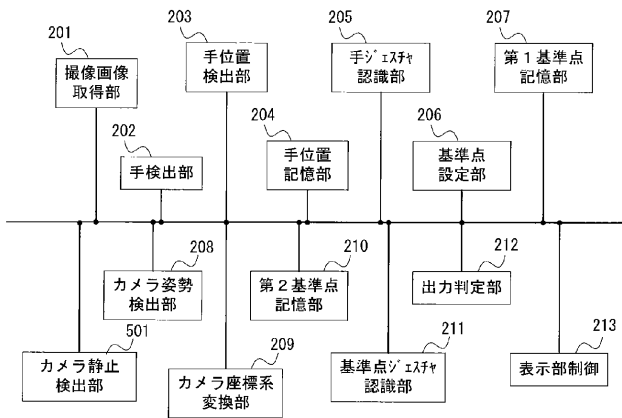
【図 7】



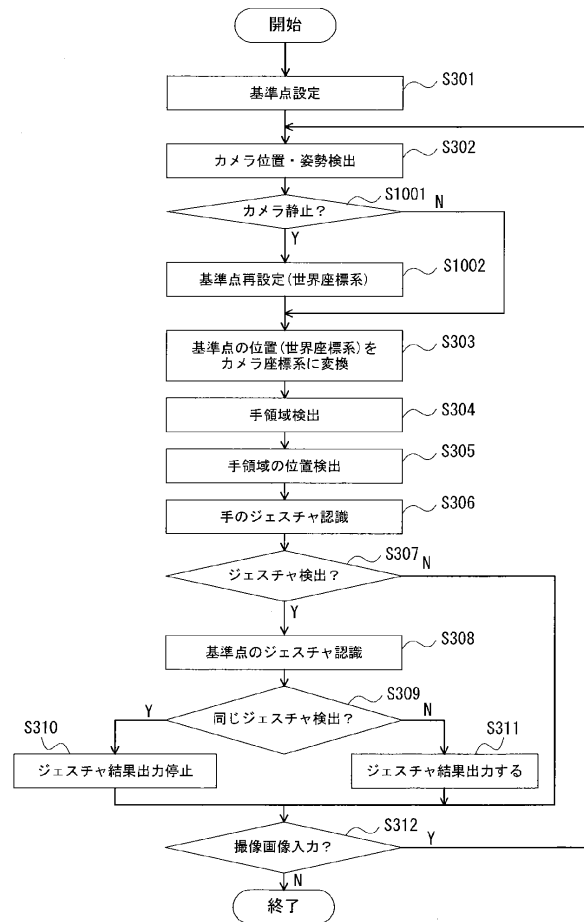
【図 8】



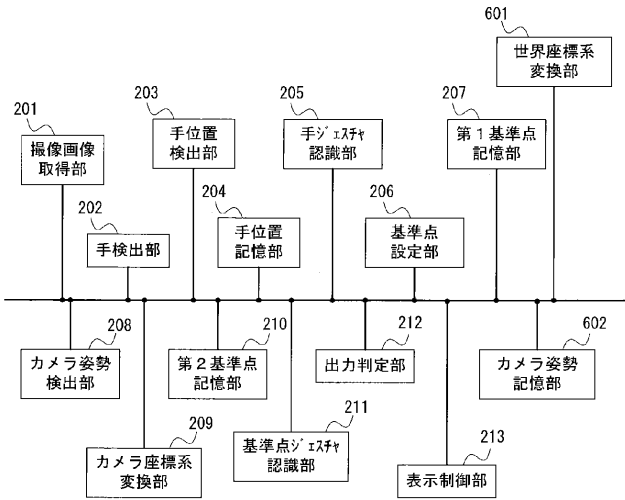
【図 9】



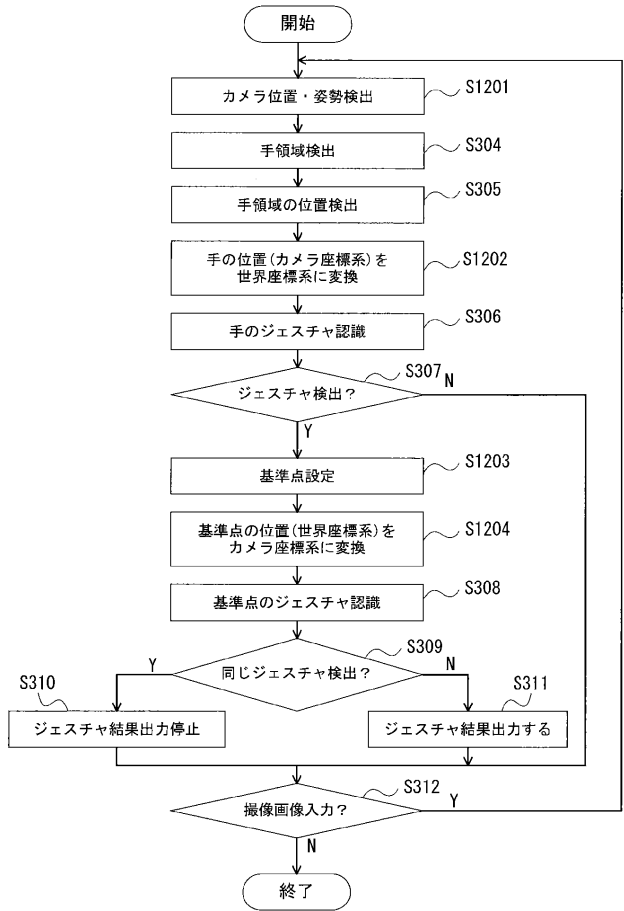
【図 10】



【図 1 1】



【図 1 2】



フロントページの続き

Fターム(参考) 5L096 AA02 AA06 CA04 CA25 DA02 FA06 FA15 FA60 FA66 FA67  
FA69 GA38 GA51 HA04 JA03 JA11 JA25 KA04