

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7520187号
(P7520187)

(45)発行日 令和6年7月22日(2024.7.22)

(24)登録日 令和6年7月11日(2024.7.11)

(51)国際特許分類

F I

G 0 6 T 7/00 (2017.01)

G 0 6 T 7/00 C

B 2 5 J 13/08 (2006.01)

B 2 5 J 13/08 A

請求項の数 18 外国語出願 (全22頁)

(21)出願番号	特願2023-98168(P2023-98168)	(73)特許権者	521105662
(22)出願日	令和5年6月15日(2023.6.15)		デクステリティ・インコーポレーテッド
(62)分割の表示	特願2021-549235(P2021-549235)の分割		D E X T E R I T Y , I N C .
原出願日	令和1年12月3日(2019.12.3)		アメリカ合衆国 カリフォルニア州 9 4
(65)公開番号	特開2023-120314(P2023-120314 A)		0 6 3 レッドウッド・シティ, ペテラ ンズ・ブルバード, 1 2 0 5
(43)公開日	令和5年8月29日(2023.8.29)	(74)代理人	110000028
審査請求日	令和5年6月22日(2023.6.22)		弁理士法人明成国際特許事務所
(31)優先権主張番号	16/667,661	(72)発明者	チャベス・ケヴィン・ホセ
(32)優先日	令和1年10月29日(2019.10.29)		アメリカ合衆国 カリフォルニア州 9 4
(33)優先権主張国・地域又は機関	米国(US)		3 0 6 バロ・アルト, ランバート・ア ベニュー, 2 9 2
(31)優先権主張番号	62/809,389	(72)発明者	ガオ・ユアン
(32)優先日	平成31年2月22日(2019.2.22)		アメリカ合衆国 カリフォルニア州 9 4
	最終頁に続く		3 0 6 バロ・アルト, ランバート・ア 最終頁に続く

(54)【発明の名称】 マルチカメラ画像処理

(57)【特許請求の範囲】

【請求項 1】

システムであって、
作業空間に関連付けられている複数のセンサの各々から画像データを受信するよう構成されている通信インターフェースと、
前記通信インターフェースに接続されているプロセッサと、
を備え、
前記プロセッサは、
前記複数のセンサによって取得された画像データに少なくとも一部に基づいて、前記複数のセンサの少なくとも1つが再校正されるべきであることを検出し、
前記作業空間における校正パターンまたは校正マーカーに関して前記複数のセンサの内の基準センサを校正し、
前記基準センサから1または複数の画像を取得し、
前記複数のセンサの少なくとも1つをクロス校正するために前記1または複数の画像を使用し、
前記複数のセンサが校正されたとの判定に応じて、
前記複数のセンサから画像データを取得し、
前記複数のセンサからの前記画像データをマージして、マージ済み点群データを生成し、
前記複数のセンサの内の一部からの画像データに基づいて分割を実行して、分割結果

を生成し、前記分割結果は一つまたは複数の物体の物体境界に関連しているものとして識別された複数のピクセルを含み、

前記マージ済み点群データおよび前記分割結果を用いて、前記作業空間のマージ済み三次元分割ビューを生成するよう構成され、前記マージ済み点群データおよび前記分割結果を用いて、前記作業空間のマージ済み三次元分割ビューを生成することは、

各物体の点の周りに分割された形状を備える点群に逆投影するために、前記分割結果を用い、

前記複数のセンサの各々について、各物体に関してセンサによって生成された前記点群をラベル付けし、対応する重心を計算し、

前記作業空間内において物体を分割するために、前記複数のセンサの対応する物体点群の重心の間で最近傍計算を用いる、ことを含む、
システム。

【請求項 2】

請求項 1 に記載のシステムであって、前記複数のセンサは、1 または複数の三次元 (3 D) カメラを含む、システム。

【請求項 3】

請求項 1 に記載のシステムであって、前記プロセッサは、前記マージ済み点群データおよび前記分割結果の一方または両方を用いて、前記作業空間内の物体に関するボックス適合を実行するよう構成されている、システム。

【請求項 4】

請求項 3 に記載のシステムであって、前記プロセッサは、さらに、前記マージ済み点群データおよび前記分割結果の一方または両方を用いて、ロボットアームで前記物体を把持するための戦略を合成するよう構成されている、システム。

【請求項 5】

請求項 4 に記載のシステムであって、前記プロセッサは、さらに、前記ロボットアームで前記物体を把持するための前記戦略を実施するよう構成されている、システム。

【請求項 6】

請求項 5 に記載のシステムであって、前記プロセッサは、前記物体を元の位置からピックアップして、前記物体を前記作業空間内の移動先位置にプレースするために、ロボット動作によって前記物体を把持するよう構成されている、システム。

【請求項 7】

請求項 1 に記載のシステムであって、前記プロセッサは、さらに、前記作業空間の前記マージ済み三次元分割ビューを用いて、前記作業空間の視覚化表現を表示するよう構成されている、システム。

【請求項 8】

請求項 7 に記載のシステムであって、前記表示された視覚化表現は、前記作業空間内に描写された前記物体を強調している、システム。

【請求項 9】

請求項 1 に記載のシステムであって、前記プロセッサは、さらに、前記マージ済み点群データをサブサンプリングするよう構成されている、システム。

【請求項 10】

請求項 9 に記載のシステムであって、前記プロセッサは、さらに、前記サブサンプリングされた点群データにクラスタ処理を実行するよう構成されている、システム。

【請求項 11】

請求項 10 に記載のシステムであって、前記プロセッサは、前記サブサンプリングおよびクラスタリングされた点群データと前記分割結果とを用いて、前記作業空間内の物体に関するボックス適合結果を生成するよう構成されている、システム。

【請求項 12】

請求項 1 に記載のシステムであって、前記プロセッサは、さらに、前記複数のセンサの内の前記一部のセンサに含まれないセンサに関連付けられている画像データに少なくとも

10

20

30

40

50

部分的に基づいて、前記作業空間の前記マージ済み三次元分割ビューを検証するよう構成されている、システム。

【請求項 13】

請求項 12 に記載のシステムであって、前記プロセッサは、少なくとも部分的には、前記作業空間の前記マージ済み三次元分割ビューを用いて、前記作業空間内の物体に関する第 1 ボックス適合結果を生成し、前記複数のセンサの内の前記一部のセンサに含まれないセンサに関連付けられている画像データを用いて、前記作業空間内の前記物体に関する第 2 ボックス適合結果を生成し、前記第 1 ボックス適合結果および前記第 2 ボックス適合結果を用いて、前記物体の検証済みボックス適合を決定することにより、前記複数のセンサの内の前記一部のセンサに含まれない前記センサに関連付けられている前記画像データに少なくとも部分的に基づいて、前記作業空間の前記マージ済み三次元分割ビューを検証するよう構成されている、システム。

10

【請求項 14】

請求項 1 に記載のシステムであって、前記プロセッサは、少なくとも部分的には、ユーザ定義された処理パイプラインを実装することにより、前記画像データをマージして処理するよう構成されている、システム。

【請求項 15】

請求項 14 に記載のシステムであって、前記プロセッサは、前記ユーザ定義された処理パイプラインの定義を受信してパースするよう構成されている、システム。

【請求項 16】

請求項 15 に記載のシステムであって、前記プロセッサは、前記定義を用いて、前記パイプラインを構成するモジュールのインスタンスを作成し、前記モジュールおよびパイプラインを実装するためのバイナリコードを自動生成するよう構成されている、システム。

20

【請求項 17】

コンピュータによって実行される方法であって、
作業空間に関連付けられている複数のセンサの各々から画像データを受信し、
前記複数のセンサによって取得された画像データに少なくとも一部に基づいて、前記複数のセンサの少なくとも 1 つが再校正されるべきであることを検出し、

前記作業空間における校正パターンまたは校正マーカーに関して前記複数のセンサの内の基準センサを校正し、

30

前記基準センサから 1 または複数の画像を取得し、
前記複数のセンサの少なくとも 1 つをクロス校正するために前記 1 または複数の画像を使用し、

前記複数のセンサが校正されたとの判定に応じて、
前記複数のセンサからの前記画像データをマージして、マージ済み点群データを生成し、
前記複数のセンサの一部からの画像データに基づいて分割を実行して、分割結果を生成し、前記分割結果は一つまたは複数の物体の物体境界に関連しているものとして識別された複数のピクセルを含み、

前記マージ済み点群データおよび前記分割結果を用いて、前記作業空間のマージ済み三次元分割ビューを生成することを備え、前記マージ済み点群データおよび前記分割結果を用いて、前記作業空間のマージ済み三次元分割ビューを生成することは、

40

各物体の点の周りに分割された形状を備える点群に逆投影するために、前記分割結果を用い、

前記複数のセンサの各々について、各物体に関してセンサによって生成された前記点群をラベル付けし、対応する重心を計算し、

前記作業空間内において物体を分割するために、前記複数のセンサの対応する物体点群の重心の間で最近傍計算を用いる、ことを含む、
方法。

【請求項 18】

コンピュータプログラムであって、コンピュータに、

50

作業空間に関連付けられている複数のセンサの各々から画像データを受信させるためのコンピュータ命令と、

前記複数のセンサによって取得された画像データに少なくとも一部に基づいて、前記複数のセンサの少なくとも1つが再校正されるべきであることを検出させるためのコンピュータ命令と、

前記作業空間における校正パターンまたは校正マーカーに関して前記複数のセンサの内の基準センサを校正させるためのコンピュータ命令と、

前記基準センサから1または複数の画像を取得させるためのコンピュータ命令と、

前記複数のセンサの少なくとも1つをクロス校正するために前記1または複数の画像を使用させるためのコンピュータ命令と、

前記複数のセンサが校正されたとの判定に応じて、

前記複数のセンサからの前記画像データをマージして、マージ済み点群データを生成させるためのコンピュータ命令と、

前記複数のセンサの内の一部からの画像データに基づいて分割を実行して、分割結果を生成させるためのコンピュータ命令と、前記分割結果は一つまたは複数の物体の物体境界に関連しているものとして識別された複数のピクセルを含み、

前記マージ済み点群データおよび前記分割結果を用いて、前記作業空間のマージ済み三次元分割ビューを生成させるためのコンピュータ命令であって、

各物体の点の周りに分割された形状を備える点群に逆投影するために、前記分割結果を用い、

前記複数のセンサの各々について、各物体に関してセンサによって生成された前記点群をラベル付けし、対応する重心を計算し、

前記作業空間内において物体を分割するために、前記複数のセンサの対応する物体点群の重心の間で最近傍計算を用いる、ことを含む、前記マージ済み点群データおよび前記分割結果を用いて、前記作業空間のマージ済み三次元分割ビューを生成させるためのコンピュータ命令と、

を備える、コンピュータプログラム。

【発明の詳細な説明】

【背景技術】

【0001】

他の出願への相互参照

本願は、2019年2月22日出願の名称を「ROBOTIC MULTI-ITEM TYPE PALLETIZING & DEPALLETIZING」とする米国仮特許出願第62/809,389号に基づく優先権を主張し、その仮特許出願は、すべての目的のために参照により本明細書に組み込まれる。

【0002】

本願は、2019年4月10日出願の同時係属の名称を「ROBOTIC MULTI-ITEM TYPE PALLETIZING & DEPALLETIZING」とする米国特許出願第16/380,859号の一部継続出願であり、その出願は、すべての目的のために参照により本明細書に組み込まれ、2019年2月22日出願の名称を「MULTI-ITEM TYPE PALLETIZING & DEPALLETIZING」とする米国仮特許出願第62/809,389号に基づく優先権を主張し、その仮特許出願は、すべての目的のために参照により本明細書に組み込まれる。

【0003】

例えば、アイテムをピック、移動、操作、および、プレースするために、多くの環境でロボットが利用されている。物理環境（本明細書では「作業空間」とも呼ばれる）においてタスクを実行するために、ロボットシステムは、典型的には、ロボットシステムによって操作される物体（ロボットアームを用いてピックアッププレースされるアイテムなど）を検出するため、ならびに、例えば、環境内の1または複数の物体を把持し、作業空間内の新たな位置へかかる物体を移動するなど、物体を操作するための計画を生成および実行

10

20

30

40

50

するために、カメラおよびその他のセンサを利用する。

【 0 0 0 4 】

センサは、複数のカメラを含んでよく、その内の 1 または複数は、三次元（「 3 D 」）カメラであってよく、 3 D カメラは、従来の（例えば、赤 - 青 - 緑、すなわち、「 R B G 」）画像データに加えて、画像内の点までの距離を示す「深度ピクセル」を生成する。しかしながら、単一のカメラでは、物体またはその一部が不明瞭であることなどにより、作業空間内のすべての物体に対して画像データおよび/または完全 3 D 画像データを生成することができない場合がある。操作を成功させるために、ロボットシステムは、変化する条件に応じることができなければならない、操作上意味のある時間枠内で操作を計画および実行できなければならない。

10

【図面の簡単な説明】

【 0 0 0 5 】

以下の詳細な説明と添付の図面において、本発明の様々な実施形態を開示する。

【 0 0 0 6 】

【図 1】ロボットシステムの一実施形態を示す図。

【 0 0 0 7 】

【図 2】複数のカメラからの画像データを用いてロボット動作を実行するための処理の一実施形態を示すフローチャート。

【 0 0 0 8 】

【図 3】分割された画像データを用いてロボット動作を実行するための処理の一実施形態を示すフローチャート。

20

【 0 0 0 9 】

【図 4】作業空間内に配備された複数のカメラを較正するための処理の一実施形態を示すフローチャート。

【 0 0 1 0 】

【図 5】作業空間からの画像データに物体インスタンス分割処理を実行するための処理の一実施形態を示すフローチャート。

【 0 0 1 1 】

【図 6】作業空間内のカメラの較正を維持するための処理の一実施形態を示すフローチャート。

30

【 0 0 1 2 】

【図 7】作業空間内のカメラを再較正するための処理の一実施形態を示すフローチャート。

【 0 0 1 3 】

【図 8】複数のカメラからの画像データを用いて、作業空間内でロボット動作を実行し、および/または、作業空間の可視化表現を提供するための処理の一実施形態を示すフローチャート。

【 0 0 1 4 】

【図 9 A】ロボット制御のためのマルチカメラ画像処理システムの一実施形態を示す図。

【 0 0 1 5 】

【図 9 B】ロボット制御のためのマルチカメラ画像処理システムの一実施形態を示す図。

40

【 0 0 1 6 】

【図 9 C】ロボット制御のためのマルチカメラ画像処理システムの一実施形態を示す図。

【 0 0 1 7 】

【図 1 0】マルチカメラ画像処理システムの一実施形態において生成および提供された視覚的表示の一例を示す図。

【 0 0 1 8 】

【図 1 1】ロボット制御のためのマルチカメラ画像処理システムにおいてセンサデータを処理するためのコードを生成する処理の一実施形態を示すフローチャート。

【発明を実施するための形態】

【 0 0 1 9 】

50

本発明は、処理、装置、システム、物質の組成、コンピュータ読み取り可能な格納媒体上に具現化されたコンピュータプログラム製品、および／または、プロセッサ（プロセッサに接続されたメモリに格納および／またはそのメモリによって提供される命令を実行するよう構成されているプロセッサ）を含め、様々な形態で実施されうる。本明細書では、これらの実施例または本発明が取りうる任意の他の形態が、技術と呼ばれうる。一般に、開示されている処理の工程の順序は、本発明の範囲内で変更されてもよい。特に言及しない限り、タスクを実行するよう構成されるものとして記載されたプロセッサまたはメモリなどの構成要素は、或る時間にタスクを実行するよう一時的に構成されている一般的な構成要素として、または、タスクを実行するよう製造された特定の構成要素として実装されてよい。本明細書では、「プロセッサ」という用語は、１または複数のデバイス、回路、および／または、コンピュータプログラム命令などのデータを処理するよう構成されている処理コアを指すものとする。

10

【 0 0 2 0 】

以下では、本発明の原理を示す図面を参照しつつ、本発明の１または複数の実施形態の詳細な説明を行う。本発明は、かかる実施形態に関連して説明されているが、どの実施形態にも限定されない。本発明の範囲は、特許請求の範囲によってのみ限定されるものであり、本発明は、多くの代替物、変形物、および、等価物を含む。以下の説明では、本発明の完全な理解を提供するために、多くの具体的な詳細事項が記載されている。これらの詳細事項は、例示を目的としたものであり、本発明は、これらの具体的な詳細事項の一部または全てがなくとも特許請求の範囲に従って実施可能である。簡単のために、本発明に関連する技術分野で周知の技術事項については、本発明が必要以上にわかりにくくならないように、詳細には説明していない。

20

【 0 0 2 1 】

複数のカメラまたはその他の画像センサを含むセンサのセットを用いて、作業空間の三次元ビューを生成するための技術が開示されている。いくつかの実施形態において、三次元ビューは、作業空間で作業を実行するために（例えば、不均質なアイテム（例えば、異なるサイズ、形状、重量、重量分布、剛性、壊れやすさ、など）の任意のセットをパレタイズ／デパレタイズならびに／もしくは他の方法で梱包および／または開梱するために）、１または複数のロボット（例えば、吸着器、グリッパ、および／または、その他のエンドエフェクタを作用端に備えたロボットアーム）を含むロボットシステムをプログラムの

30

【 0 0 2 2 】

様々な実施形態において、３Ｄカメラ、力センサ、および、その他のセンサが、ピックおよび／またはブレースされるアイテムの属性を検出および判定するため、ならびに／もしくは、最初の位置で１または複数のアイテムを把持し、作業空間内の対応する移動先位置に１または複数のアイテムの各々を移動させるために、用いられる。（例えば、例えばプログラムの決定された信頼スコアによって示されるように、十分な信頼性で）タイプを判定されたアイテムが、アイテムタイプ固有モデルから導出された戦略を用いて、把持およびブレースされてよい。識別できないアイテムは、所与のアイテムタイプに固有ではない戦略を用いてピックアンドブレースされる。例えば、サイズ、形状、および、重量の

40

【 0 0 2 3 】

いくつかの実施形態において、本明細書で開示されている技術は、作業空間の少なくとも一部の視覚的表現を生成および表示するために用いられてよい。様々な実施形態において、視覚的表現は、完全自動モードまたは部分自動モードで動作しているロボットを監視するため、および／または、遠隔操作でロボットアームまたはその他のロボットアクチュエータを制御するために、人間オペレータが利用するワークステーションを含むコンピュータまたはその他のディスプレイデバイスを介して表示されてよい。

【 0 0 2 4 】

例えば、いくつかの実施形態において、ロボットシステムがスタックした場合（例えば

50

、設定されたパラメータ（例えば、タイムアウト、信頼スコア、など）の範囲内で、次のタスクまたは操作を実行または完了できない場合に、人的介入が要請されてよい。いくつかの実施形態において、表示された作業空間の画像および／またはビデオが、遠隔操作を実行するために用いられてよい。人間オペレータは、表示された画像またはビデオを用いて、作業空間を見ながらロボットを制御することで、ロボットを手動で制御してよい。いくつかの実施形態において、ディスプレイは、インタラクティブな部分自動システムに組み込まれていてよい。例えば、人間オペレータが、ディスプレイを介して、ロボットが物体を把持すべき点をシーンの表示画像内で示してよい。

【 0 0 2 5 】

図 1 は、ロボットシステムの一実施形態を示す図である。図の例において、ロボットシステム 1 0 0 は、ロボットアーム 1 0 2 を備えており、異質のアイテムをパレタイズ（パレットに搭載）および／またはデパレタイズ（パレットから取り出し）するよう構成されている。この例において、ロボットアーム 1 0 2 は、固定されているが、様々な別の実施形態において、ロボットアーム 1 0 2 は、例えば、レール上に取り付けられている、モータ駆動シャーシ上で完全に移動可能である、など、完全にまたは部分的に移動可能であってもよい。図に示すように、ロボットアーム 1 0 2 は、コンベヤベルト（またはその他のソース） 1 0 4 から任意のおよび／または異種のアイテムをピックして、パレットまたはその他の荷役台 1 0 6 に積み重ねるために用いられる。図の例において、荷役台 1 0 6 は、四隅に車輪を備え、四方の側面の内の 3 つが少なくとも部分的に閉じているパレットまたはベースを備えており、三面の「ロールパレット」、「ロールケージ」、および／または

10

20

【 0 0 2 6 】

図の例において、ロボットアーム 1 0 2 は、吸着タイプのエンドエフェクタ 1 0 8 を備えている。エンドエフェクタ 1 0 8 は、複数の吸着カップ 1 1 0 を有する。ロボットアーム 1 0 2 は、図に示すように、ピックアップされるアイテムの上にエンドエフェクタ 1 0 8 の吸着カップ 1 1 0 を配置するために用いられ、真空源が、アイテムを把持し、アイテムをそのコンベヤ 1 0 4 から持ち上げ、アイテムを荷役台 1 0 6 上の移動先位置にプレースするための吸着力を提供する。

30

【 0 0 2 7 】

様々な実施形態において、エンドエフェクタ 1 0 8 に取り付けられたカメラ 1 1 2 と、ロボットシステム 1 0 0 が配備されている空間内に取り付けられたカメラ 1 1 4、1 1 6 と、の内の 1 または複数が、コンベヤ 1 0 4 上のアイテムを識別するため、および／または、アイテムを把持し、ピック／プレースし、荷役台 1 0 6 上に積み重ねるための計画を決定するために用いられる。様々な実施形態において、図示されていないさらなるセンサ（例えば、コンベヤ 1 0 4 および／またはロボットアーム 1 0 2 の中におよび／または隣接して具現化されている重量センサまたは力センサ、吸着カップ 1 1 0 の x - y 平面および／または z - 方向（垂直方向）の力センサ、など）が、例えばシステム 1 0 0 によって、アイテムが配置および／または再配置されうるコンベヤ 1 0 4 および／またはその他の移動元および／またはステージングエリア上のアイテムを識別し、その属性を決定し、把持し、ピックアップし、決定された軌道を通して移動させ、および／または、荷役台 1 0 6 の上または中の移動先位置にプレースするために用いられてよい。

40

【 0 0 2 8 】

図の例において、カメラ 1 1 2 は、エンドエフェクタ 1 0 8 の本体の側面に取り付けられているが、いくつかの実施形態において、カメラ 1 1 2 および／またはさらなるカメラが、他の位置に取り付けられてもよい（例えば、吸着カップ 1 1 0 の間の位置から下向きに、エンドエフェクタ 1 0 8 の本体の下面に取り付けられ、もしくは、ロボットアーム 1

50

02のセグメントまたはその他の構造に取り付けられ、もしくは、その他の位置に取り付けられる、など)。様々な実施形態において、112、114、および、116などのカメラが、コンベヤ104上のアイテム上に見えるおよび/またはアイテムを構成する、テキスト、ロゴ、写真、図、画像、マーク、バーコード、QRコード(登録商標)、または、その他の符号化および/またはグラフィック情報またはコンテンツを読み取るために用いられてよい。

【0029】

図1をさらに参照すると、図の例において、システム100は、この例では無線通信を介して(ただし、様々な実施形態において、有線通信および無線通信の一方または両方で)、ロボットアーム102、コンベヤ104、エフェクタ108、および、センサ(カメラ112、114、および、116、ならびに/もしくは、図1に示していない重量センサ、力センサ、および/または、その他のセンサなど)などの要素と通信するよう構成されている制御コンピュータ118を備える。様々な実施形態において、制御コンピュータ118は、センサ(カメラ112、114、および、116、ならびに/もしくは、図1に示していない重量センサ、力センサ、および/または、その他のセンサなど)からの入力を用いて、荷役台106へ荷積みおよび/または荷役台106から荷下ろしされるアイテムの1または複数の属性を観察、識別、および、決定するよう構成されている。様々な実施形態において、制御コンピュータ118は、例えば、画像および/またはその他のセンサデータに基づいて、制御コンピュータ118に格納されたライブラリ内および/または制御コンピュータ118にとってアクセス可能なライブラリ内のアイテムモデルデータを用いて、アイテムおよび/またはその属性を識別する。制御コンピュータ118は、アイテムに対応するモデルを用いて、移動先(荷役台106など)の中/上に、他のアイテムと共に、アイテムを積み重ねるための計画を決定および実施する。様々な実施形態において、アイテム属性および/またはモデルは、アイテムを把持、移動させて、移動先位置(例えば、荷役台106の中/上にアイテム積み重ねるための計画/再計画処理の一部として、アイテムがプレースされると決定された位置)にプレースするための戦略を決定するのに利用される。

【0030】

図の例において、制御コンピュータ118は、「オンデマンド」遠隔操作装置122に接続されている。いくつかの実施形態において、制御コンピュータ118が、完全自動モードで続行できない場合、例えば、制御コンピュータ118が完全自動モードでアイテムのピックアンドプレースを完了するための戦略を持たなくなるように、アイテムを把持、移動、および、プレースするための戦略が、決定できなくなり、および/または、失敗した場合、制御コンピュータ118は、例えば、アイテムを把持、移動、および、プレースするために、遠隔操作装置122を用いて、ロボットアーム102および/またはエンドエフェクタ108を操作することによって介入するように人間ユーザ124に指示する。

【0031】

様々な実施形態において、制御コンピュータ118は、画像データ(例えば、二次元RGB画像データまたはその他の画像データ、ビデオデータを構成する連続フレーム、3Dセンサによって生成された点群データ、2D画像データの対応するフレームに各々関連付けられている連続的な点群データセット、など)を受信して処理するよう構成されている。いくつかの実施形態において、制御コンピュータ118は、カメラ112、114、および、116、ならびに/もしくは、その他のセンサ(レーザセンサ、および、その他の光センサ、熱センサ、レーダセンサ、ソナーセンサ、もしくは、投写、反射、放射、および/または、その他の方法で受信された電磁放射および/または信号を用いて、画像を作成するために利用されるまたは利用されうる情報を検出および/または伝達するその他のセンサ)によって生成されてそこから受信された画像データに基づいて、別個のコンピュータ、アプリケーション、サービスなどによって生成された集約および/またはマージされた画像データを受信する。本明細書で利用される画像は、物理空間またはシーン(図1に示す例においてロボットシステム100が配置されている作業空間など)に存在する物

10

20

30

40

50

体および／または特徴の視覚的および／またはコンピュータまたはその他のマシンで認識可能な表現、描写、などを含む。

【 0 0 3 2 】

様々な実施形態において、カメラ 1 1 2、1 1 4、および、1 1 6、ならびに／もしくは、その他のセンサによって生成および提供された画像データは、ロボットシステム 1 0 0 が配置されている作業空間の少なくとも一部の三次元ビューを生成するために処理および利用される。いくつかの実施形態において、複数のカメラ（例えば、1 1 2、1 1 4、1 1 6）からの画像データは、作業空間の三次元ビューを生成するためにマージされる。マージされた画像データは、作業空間内の対象物体の境界を決定するために分割される。分割された画像データは、作業空間内の物体の把持、作業空間での物体の移動、移動先位置への物体のブレース、の内の 1 または複数を実行するための戦略または計画を、自動処理で決定するなどのタスクを実行するために用いられる。

10

【 0 0 3 3 】

様々な実施形態において、複数のカメラ（例えば、カメラ 1 1 2、1 1 4、1 1 6）によって生成された 3 D 点群データビューは、レジストレーションとして知られる処理によって、作業空間の完全なモデルまたはビューにマージされる。別個に取得されたビューにキャプチャされた作業空間の物体および特徴のそれぞれの位置および向きは、それらの交わる領域が可能な限り完全に重なるように、グローバル三次元座標フレームワークに変換される。異なるカメラまたはその他のセンサから取得された点群データセット（すなわち、異なるビュー）のあらゆるセットについて、様々な実施形態において、システムは、後続の処理工程（分割および物体再構築など）を適用できるように、本明細書に開示されているようにそれらを一緒に単一の点群モデルに整列させる。

20

【 0 0 3 4 】

様々な実施形態において、作業空間の三次元ビューが、少なくとも部分的には、カメラ（例えば、カメラ 1 1 2、1 1 4、および、1 1 6）をクロス較正し、データをマージして、利用可能できる限り多くの角度およびビューから作業空間および作業空間に存在するアイテム／物体のビューを生成することによって、カメラ 1 1 2、1 1 4、および／または、1 1 6 によって生成および提供された画像データを用いて生成される。例えば、図 1 に示す例において、カメラ 1 1 2 および 1 1 6 は、コンベヤ 1 0 4 上の物体が見える位置にあってよく、一方、カメラ 1 1 4 は、図の例において荷役台 1 0 6 に向けられている様子が示されており、コンベヤ 1 0 4 が配置されている作業空間の部分からの画像データを（現在は）全く持ちえない。同様に、アーム 1 0 2 は、カメラ 1 1 2 がもはやコンベヤ 1 0 4 のビューを持たないような位置に移動されてもよい。様々な実施形態において、作業空間内のカメラからの画像データ（例えば、RGB ピクセル、深度ピクセル、など）は、画像データが任意の所与の瞬間にカメラおよび／またはその他のセンサから受信されると仮定して、可能な限り完全かつ正確な作業空間の三次元ビューを動的に生成および連続更新するためにマージされる。カメラが、作業空間内の物体または領域のビューを遮られている場合、影響を受けた物体または領域への見通しを持ち続けているカメラからの画像データが、物体または領域の可能な限り完全かつ正確なビューを生成するために利用し続けられる。

30

40

【 0 0 3 5 】

様々な実施形態において、本明細書に開示されている技術は、カメラからの複数の画像データが、作業空間および作業空間内の物体のより完全なビューを生成および維持するために用いられることを可能にする。例えば、作業空間における異なる位置および／または向きの複数のカメラを用いると、或る視点からは大きい物体によって隠されうる小さい物体を、小さい物体が隠されない見通しのよい視点からその物体を見るように配置された 1 または複数のカメラの画像データを介して見ることができる。同様に、物体を多くの角度から見ることで、その物体のすべての隠されていない側面および特徴を判別することが可能になるため、把持戦略を決定および実施し、物体にぴったりと隣接するようにアイテムをブレースすることを決定し、人間の作業員またはロボットアクチュエータ（例えば、ロ

50

ボットアーム、コンベヤ、ロボット制御の移動可能な棚、など）が作業空間を移動する時に物体のビューを維持する、などの動作を円滑にしうる。

【 0 0 3 6 】

いくつかの実施形態において、分割された画像（例えば、ビデオ）データが、作業空間の可視化表現を生成して表示するために用いられる。いくつかの実施形態において、対象の物体は、表示された可視化表現内で強調されてよい。例えば、色付きの境界形状または輪郭が表示されてよい。いくつかの実施形態において、人間オペレータが、対象の物体の自動生成された境界を補正、微調整、または、他の方法で境界に関するフィードバックを提供することを可能にするために、人間の操作可能なインターフェースが提供される。例えば、ユーザが、自動生成された境界形状または輪郭の位置を移動または調整すること、もしくは、強調された領域が、1ではなく2（以上）の物体を実際には含むと示すことを可能にするために、インターフェースが提供されてよい。いくつかの実施形態において、表示された可視化表現は、人間オペレータが、遠隔操作モードで作業空間内のロボットを制御することを可能にするために用いられてよい。例えば、人間オペレータは、分割されたビデオを用いて、ロボットアーム（またはその他のアクチュエータ）を適所に移動させ、強調された物体を（例えば、コンベヤ 1 0 4 から）把持し、強調された物体を（例えば、荷役台 1 0 6 上の）移動先位置へ移動させることができる。

10

【 0 0 3 7 】

様々な実施形態において、本明細書に開示のタスクを実行する目的で複数のカメラからの画像データをマージすることを可能にするために、少なくともマスタまたは較正基準カメラが、較正パターン、物体、もしくは、固定のおよび/またはその他の方法で既知の位置、向きなどを有するその他の参照に対して較正される。図 1 に示す例において、例えば、カメラ 1 1 2、1 1 4、および、1 1 6 の内の 1 または複数の、作業空間内の固定位置（例えば、壁上の既知の位置）に取り付けられたマーカ 1 3 0 および/またはロボットアーム 1 0 2 上に取り付けられた（プリントなどされた）マーカ 1 3 2 の 1 または複数の画像に基づいて、較正されてよい。様々な実施形態において、較正されたマスタカメラによって生成された画像が、作業空間内の他のカメラの内の 1 または複数のクロス較正するために用いられる。

20

【 0 0 3 8 】

いくつかの実施形態において、例えば、カメラの誤差、カメラへの衝突もしくは意図的な再配置または再方向付け、画像データに基づいて試行された動作がカメラの誤差またはズレを示唆するように失敗したこと、或るカメラからの画像データに基づいて、別のカメラの位置、向き、などが期待と異なることを、システムが検出したこと、などに起因して、カメラを再較正および/またはクロス較正する必要性を検出するための処理が実行される。様々な実施形態において、システム 1 0 0（例えば、制御コンピュータ 1 1 8）は、本明細書で開示されているように、1 または複数のカメラを再較正する必要性を自動的に検出し、自動的および動的に再較正を行うよう構成されている。例えば、再較正は、様々な実施形態において、ロボットアクチュエータ上に取り付けられたカメラ（例えば、カメラ 1 1 2）を用いて、作業空間内の基準マーカ（例えば、マーカ 1 3 0）を再位置特定すること、基準マーカを用いてカメラから作業空間への変換を再評価すること、および、ロボット上のマーカ（例えば、マーカ 1 3 2）に合わせて再較正を行うこと、の内の 1 または複数のによって実行される。

30

40

【 0 0 3 9 】

図 2 は、複数のカメラからの画像データを用いてロボット動作を実行するための処理の一実施形態を示すフローチャートである。様々な実施形態において、図 2 の処理 2 0 0 は、図 1 の制御コンピュータ 1 1 8 など、コンピュータまたはその他のプロセッサによって実行される。図の例において、画像データが、作業空間内でビデオ（例えば、RGB ピクセルおよび深度ピクセルを含む 3 D ビデオ）をキャプチャするために配置された複数のカメラから受信される（工程 2 0 2）。受信された画像データは、作業空間の三次元ビュー（「シーン」とも呼ばれる）を生成するために処理およびマージされ、作業空間内の物体

50

を区別するために分割される（工程 2 0 4）。分割されたビデオは、物体の内の 1 または複数に関する操作（例えば、物体を把持して、作業空間内の新たな位置に移動させる操作）を実行するために用いられる（工程 2 0 6）。

【 0 0 4 0 】

図 3 は、分割された画像データを用いてロボット動作を実行するための処理の一実施形態を示すフローチャートである。様々な実施形態において、図 3 の処理は、図 2 の処理の工程 2 0 6 を実施するために用いられる。図の例において、分割されたビデオデータは、1 または複数の物体を把持、移動、および、ブレースするための戦略を決定および実施するために用いられる（工程 3 0 2）。分割されたビデオおよび関連付けられている境界ボックス（またはその他の形状）は、作業空間の可視化表現を生成および表示するために用いられる（工程 3 0 4）。例えば、分割データは、いくつかの実施形態において、1 または複数のマスキレイヤを生成して、作業空間内の物体の少なくとも一部の各々の上に、物体の輪郭と（できる限り近く）一致した半透明の色付きの形状を重ね合わせるために用いられる。

10

【 0 0 4 1 】

様々な実施形態において、可視画像は、自律モードのロボットシステムの動作を監視するため、および / または、遠隔操作でロボットアームまたはその他のロボットアクチュエータを操作するために、人間オペレータによって用いられてよい。

【 0 0 4 2 】

図 4 は、作業空間内に配備された複数のカメラを較正するための処理の一実施形態を示すフローチャートである。様々な実施形態において、図 4 の処理 4 0 0 は、作業空間内の複数のカメラ（図 1 のカメラ 1 1 2、1 1 4、および、1 1 6 など）またはその他のセンサから受信されたデータを制御および処理するよう構成されているコンピュータ（図 1 の制御コンピュータ 1 1 8 など）によって実行される。図の例において、較正基準が、1 または複数のカメラを用いて作業空間内の基準マーカまたはその他の基準の画像を生成することによって取得される（工程 4 0 2）。例えば、図 1 に示した例において、カメラ 1 1 2、1 1 4、および、1 1 6 の内の 1 または複数の、マーカ 1 3 0 および / またはマーカ 1 3 2 の 1 または複数の画像を生成するために用いられてよい。いくつかの実施形態において、ロボットアームまたはその他のアクチュエータが、キーもしくはその他のアイテムまたは付属物に対応する穴またはその他の受け部へ挿入し、既知の位置および向きにある間に画像を生成することなどによって、既知の固定位置へ移動されてよい。

20

30

【 0 0 4 3 】

図 4 をさらに参照すると、較正基準は、作業空間内のすべてのカメラをクロス較正するために用いられる（工程 4 0 4）。実行時に、複数のカメラからの点群をマージするために、反復最近接点（ICP）処理が実行される（工程 4 0 6）。インスタンス分割処理が、作業空間内の物体を判別、識別（例えば、タイプなどで）、および、ラベル付けするために実行される（工程 4 0 8）。

【 0 0 4 4 】

図 5 は、作業空間からの画像データに物体インスタンス分割処理を実行するための処理の一実施形態を示すフローチャートである。様々な実施形態において、図 5 の処理は、図 4 の処理の工程 4 0 8 を実施するために用いられる。図の例において、分割処理は、作業空間内のすべてのカメラよりも少ないカメラからの RGB（2D）画像データに実行される（工程 5 0 2）。いくつかの実施形態において、分割は、1 つのカメラからの RGB データを用いて実行される。物体境界に関連しているものとして分割処理で識別された RGB ピクセルが、対応する深度ピクセルに対してマッピングされる（工程 5 0 4）。分割データおよびマッピングされた深度ピクセル情報は、物体の点の周りに分割されたボックス（またはその他の形状）を備えた点群へ逆投影するために用いられる（工程 5 0 6）。各カメラの画像データについて、各物体の点群がラベル付けされ、重心が計算される（工程 5 0 8）。最近傍計算が、物体を分割するためにそれぞれのカメラに対応する物体点群の重心の間で実行される（工程 5 1 0）。

40

50

【 0 0 4 5 】

図 6 は、作業空間内のカメラの較正を維持するための処理の一実施形態を示すフローチャートである。様々な実施形態において、図 6 の処理 6 0 0 は、図 1 の制御コンピュータ 1 1 8 など、コンピュータまたはその他のプロセッサによって実行される。図の例において、作業空間内の 1 または複数のカメラを再較正する必要性が、検出される（工程 6 0 2）。様々な実施形態において、以下の内の 1 または複数が、再較正の必要性を示唆しうる。（例えば、ベース上の A r U c o またはその他のマーカの画像に基づいて）ロボットベース位置が移動したのをカメラが捕らえたこと、作業空間内に取り付けられたカメラが移動した（例えば、衝突した、人間またはロボット作業者によって意図的に再配置された、など）のをロボットアームまたはその他のアクチュエータ上のカメラが捕らえたこと、ならびに、システムが、連続して数回の（閾値回数よりも多い）把持の失敗を検出したこと。再較正は、人的介入なしに、例えば、ピックアップブレースまたはその他のロボット動作を中断せずにリアルタイムで、動的に実行される（工程 6 0 4）。様々な実施形態において、再較正は、以下の内の 1 または複数を含んでよい。ロボットアクチュエータ上に取り付けられたカメラ（例えば、カメラ 1 1 2）を用いて、作業空間内の基準マーカ（例えば、マーカ 1 3 0）を再位置特定すること、基準マーカを用いてカメラから作業空間への変換を再評価すること、および、ロボット上のマーカ（例えば、マーカ 1 3 2）に合わせて再較正を行うこと。

10

【 0 0 4 6 】

図 7 は、作業空間内のカメラを再較正するための処理の一実施形態を示すフローチャートである。様々な実施形態において、図 7 の処理は、図 6 の処理の工程 6 0 4 を実施するために用いられる。図の例において、ロボットシステム動作（図 1 に示した例のように、所望の移動先または移動先セット（パレットまたはその他の荷役台など）へ所定のセットのアイテムをピックアップブレースする動作など）が一時的に中断される（工程 7 0 2）。指定またはその他の方法でまだ確立されていない場合には、基準カメラが決定され、1 または複数の基準画像が生成される（工程 7 0 4）。例えば、固定カメラ、ロボットアームに取り付けられたカメラなどが、基準カメラとして指定される。あるいは、1 または複数の他のカメラと同期したままであると思われるセットに含まれるカメラが、それらの他のカメラなどを再較正するための基準カメラとして選択されてもよい。画像は、作業空間内の基準マーカまたはその他の基準の、基準カメラおよび 1 または複数のその他のカメラによる画像を含んでよい。基準画像は、カメラをクロス較正するために用いられる（工程 7 0 6）。

20

30

【 0 0 4 7 】

図 8 は、複数のカメラからの画像データを用いて、作業空間内でロボット動作を実行し、および / または、作業空間の可視化表現を提供するための処理の一実施形態を示すフローチャートである。様々な実施形態において、図 8 の処理 8 0 0 は、図 1 の制御コンピュータ 1 1 8 など、コンピュータまたはその他のプロセッサによって実行される。図の例において、画像データが、作業空間内の複数のカメラから受信され、作業空間フィルタが適用される（工程 8 0 2）。様々な実施形態において、作業空間フィルタは、実行に画像 / センサデータが利用されているロボット動作のために無視できる作業空間の部分、作業空間の特徴、作業空間内のアイテム、などに関連付けられている画像および / または点群データを除外してよい。余分な情報をフィルタアウトすることで、作業空間内で操作の対象となりおよび / または関係する操作をなされる要素（把持およびブレースされる物体、アイテムが把持またはブレースされるパレットまたはその他の移動先、アイテムを移動先まで移動させる経路で遭遇しうる障害物、など）のよりクリアおよび / または集中的なビューが実現される。いくつかの実施形態において、センサからのノイズを除去するために、統計的外れ値データが作業空間フィルタによって除去されてよい。それぞれのカメラからの点群データがマージされる（工程 8 0 4）。分割が、カメラの内の 1 または複数からの R G B 画像データを用いて実行される（工程 8 0 6）。例えば、最初に、R G B 分割が、1 つだけのカメラからの画像データに対して実行されてよい。点群データは、サブサン

40

50

リングされ、クラスタリングが実行される（工程 8 0 8）。「ボックス（またはその他の 3 D 幾何プリミティブ）適合」処理が、サブサンプリングおよびクラスタリングされた点群データと R G B 分割結果データとを用いて実行される（工程 8 1 0）。安定物体マッチングが実行される（工程 8 1 2）。様々な実施形態において、同じ物体の複数の 3 D 表現が、異なるカメラから 2 D 方法を用いることで生じうる。これらの表現は、部分的にのみ重複しうる。様々な実施形態において、「安定物体マッチング」は、どのセグメントが同じ物体に対応するかを調整して、表現をマージすることを含む。様々な実施形態において、（各点と共に R G B データを含む）点群の空間的、幾何学的、曲率の特性または特徴が、安定物体マッチングを実行するために用いられてよい。いくつかの実施形態において、安定物体マッチングは、2 つの異なるカメラの代わりに同じカメラからの複数フレームにわたって、経時的に実行されてもよい。処理された画像データは、把持合成を実行するために用いられる（工程 8 1 4）。例えば、物体の位置および境界情報が、ロボットグリッパまたはその他のエンドエフェクタを用いて物体を把持する戦略を決定するために用いられる（工程 8 1 6）。例えば、作業空間の生ビデオが、作業空間内の様々な物体に関連付けられている境界ボックス（またはその他の幾何プリミティブ）にカラーまたはその他の強調を適用して、表示されてよい。

【 0 0 4 8 】

ロボット動作を実行するために、複数のセンサ（複数の 3 D カメラなど）からのセンサデータを処理およびマージするようロボットシステムを構成する技術が開示されている。様々な実施形態において、管理者ユーザインターフェース、構成ファイル、アプリケーションプログラミングインターフェース（A P I）、または、その他のインターフェースが、センサ出力を処理および利用してロボット動作を実行するために、センサを識別し、1 または複数の処理パイプラインを定義するのに用いられてよい。様々な実施形態において、パイプラインは、処理モジュールと、かかるモジュールのそれぞれの入力および出力が処理パイプラインを形成するためにリンクされる方法とを識別することによって定義されてよい。いくつかの実施形態において、定義は、ロボット動作を実行するために、センサ入力を受信、処理、および、利用するためのバイナリコードを生成するのに用いられる。他の実施形態において、定義は、処理を実行するためのプラグインを動的にロードする単一の全体的なバイナリコードによって用いられる。

【 0 0 4 9 】

図 9 A、図 9 B、および、図 9 C は、それぞれ、様々な実施形態において、本明細書で開示されている技術を用いて構成および実装されたパイプラインの例を示す。

【 0 0 5 0 】

図 9 A は、ロボット制御のためのマルチカメラ画像処理システムの一実施形態を示す図である。図の例において、システム 9 0 0 は、センサノード 9 0 2 および 9 0 4 として図 9 A に示されている複数のセンサ（3 D またはその他のカメラなど）を備える。ノード 9 0 2 および 9 0 4 のセンサ出力は、それぞれ、作業空間フィルタ 9 0 6 および 9 0 8 によって処理される。結果として得られる点群データ（フィルタリングされたセンサ出力の 3 D 部分）は、マージされる（9 1 0）。マージされた点群データ 9 1 0 は、サブサンプリング（9 1 2）およびクラスタリング（9 1 4）され、この例においては、結果として得られる物体クラスタデータが、ターゲット物体（作業空間内に存在するアイテムが移動されるパレットまたはトレイなど）の位置および向きを推定するために、「ターゲット推定」モジュール / 処理 9 1 5 によって用いられる。

【 0 0 5 1 】

図 9 A に示す例では、分割処理 9 1 6、9 2 2 が、センサノード 9 0 2、9 0 4 からの R G B データに対して実行され、それぞれの分割結果が、「ボックス適合」処理 9 1 8、9 2 4 を実行して R G B 分割で識別された物体インスタンスの境界ボックス（またはその他の多面体）を決定するために用いられる。ボックス適合結果 9 1 8、9 2 4 は、マージされ（9 2 5）、安定物体マッチング 9 2 6 を実行するために用いられ、その結果は、把

持合成 9 2 7 を実行するため、ならびに、作業空間の可視化表現 9 2 8 を生成および表示するために用いられる。

【 0 0 5 2 】

図 9 A に示す例において、R G B 分割情報のみがマージされ (9 2 5)、安定物体マッチング 9 2 6 を実行し、把持合成 9 2 7 を実行し、作業空間内の物体の可視化表現 9 2 8 を生成するために用いられているが、いくつかの別の実施形態においては、点群データをマージし (9 1 0)、サブサンプリングし (9 1 2)、マージされた点群データに基づいて物体クラスを識別した (9 1 4) 結果が、安定物体マッチング 9 2 6 を実行し、把持合成 9 2 7 を実行し、作業空間内の物体の視覚化画像 9 2 8 を生成するために、R G B 分割およびボックス適合 (9 1 6、9 1 8、9 2 2、9 2 4) の結果と共にマージされてもよい (9 2 5)。

10

【 0 0 5 3 】

様々な実施形態において、パイプラインは、条件に基づいて、予め定義されてよく、および/または、リアルタイムで適合されてよい。例えば、作業空間内の物体が著しく散らかっている場合、R G B 分割結果は、3 D クラスタリング処理よりも良好な信号でありえ、いくつかの実施形態において、かかる条件下では、ボックス (ポリゴン) 適合は、R G B 分割出力のみに適用されてよい。他の条件においては、分割の両方のソース (R G B、点群データ) が、幾何プリミティブ適合の実施時などに適用されてよい。

【 0 0 5 4 】

図 9 A に示す例において、すべてのセンサ (例えば、カメラ) 9 0 2、9 0 4 からの R G B 分割 9 1 6、9 2 2 データに基づいた「ボックス適合」結果 9 1 8、9 2 4 が、下流タスク (安定物体マッチング 9 2 6、把持合成 9 2 7、および、可視化 9 2 8、など) を実行するためにマージされて (9 2 5) 利用されているが、様々な実施形態において、すべてのセンサよりも少ないセンサからの R G B 分割および/またはボックス適合データが、下流タスクの内の 1 または複数を実行するためにマージされて利用されてもよい。いくつかの実施形態において、所与のセンサからのデータの質および/または内容が、所与の下流タスクを実行する際に、信頼できないおよび/またはその他のセンサからのデータと同程度に有用である旨の判定が動的になされてよく、そのセンサからのデータは、そのタスクの実行時に省略されてよい (すなわち、破棄されて利用されなくてよい)。

20

【 0 0 5 5 】

いくつかの実施形態において、パイプラインが、所与のパイプラインパスおよび/またはタスクから所与のセンサを省略するために定義されてよい。例えば、本明細書で開示されているパイプラインを定義するユーザは、センサの性能、質、信頼性、および/または、センサの位置に基づいて、そのセンサが一部のタスクに有用であるが他のタスクには有用でありえないと判断してよく、そのセンサが適切および/または有用であるタスクにそのセンサの出力のみを利用するようにパイプラインを定義してよい。

30

【 0 0 5 6 】

様々な実施形態において、図 9 A のパイプライン 9 0 0 などのパイプラインは、例えば、特定のセンサの利用方法を変更するため、および/または、センサ 9 0 2、9 0 4 のセットにセンサを追加または除去するために、いつでも変更されてよい。様々な実施形態において、新しいまたは更新されたパイプラインが定義され、パイプラインを実装するためのコードが、本明細書で開示されているように、生成および配備され、新しいパイプライン、追加されたセンサ、などが、その後のロボット動作を実行するために配備および利用されることを可能にする。

40

【 0 0 5 7 】

図 9 B は、ロボット制御のためのマルチカメラ画像処理システムの一実施形態を示す図である。図 9 B に示す処理パイプライン 9 4 0 の例において、センサ 9 0 2、9 0 4 の出力 (例えば、カメラフレームデータ) が作業空間フィルタ 9 0 6、9 0 8 によって処理されることで、フィルタリングされたセンサ出力を生成し、その出力の 3 D 点群データ部分がマージされる (9 1 0) ことで、作業空間のマージ済み点群データを生成する。この例

50

においてマージされた点群データは、以下の３つの別個のモジュールに提供される。マージされた点群データに存在する情報を強化するための処理を実行する事前可視化モジュール 942、作業空間内の対象の物体である 3D 空間および / または対象の物体によって占められていない 3D 空間を識別する「ボクセル」処理モジュール 944、ならびに、マージされた点群データと事前可視化モジュール 942 およびボクセルモジュール 944 の出力とに基づいて、可視化表現を生成して、例えばディスプレイデバイスを介して、提供する可視化モジュール 946。様々な実施形態において、事前可視化モジュール 942 は、処理パイプラインの残り部分からのデータを再フォーマットすることで、効率的なレンダリングを可能にし、ユーザが、パン、ズーム、回転などを円滑に実行できる高度にインタラクティブな可視化を実現する。

10

【0058】

図 9C は、ロボット制御のためのマルチカメラ画像処理システムの一実施形態を示す図である。図の例において、パイプライン 960 は、単一のセンサノード 962 を備えており、その出力（例えば、3D またはその他のカメラフレームデータ）は、フィルタリングされたデータをクラスタリングモジュールまたは処理 966、事前可視化モジュール 970、および、可視化モジュール 972 へ提供するために、作業空間フィルタ 964 によって処理される。さらに、クラスタリングモジュール 966 の出力は、可視化モジュール 972 および把持合成モジュール 968 へ提供され、次に、把持合成モジュール 968 の出力は、事前可視化モジュール 970 および可視化モジュール 972 へ提供される。図 9C に示す例は、本明細書で開示されているように、処理パイプラインを定義するためにモジュールがどれだけ容易にリンクされるのかを示しており、中間モジュールの出力は、各処理モジュールがパイプライン内の次のモジュールへの出力として最良の情報を提供するために必要な情報を有することを保証するために、複数のパスに沿って伝えられる。図 9C に示す例において、例えば、可視化モジュールは、センサ 962 および作業空間フィルタ 964 による生フィルタリング済みセンサノード出力と、作業空間フィルタリングされたセンサ出力 962、964 および把持合成結果 968 に基づいて生成されたクラスタ情報 966、把持合成データ 968、および、事前可視化結果 970 とにアクセスすることで、ロボット動作を監視し、および / または、（例えば、遠隔操作で）介入する際に、人間オペレータにとって非常に有用でありうる情報で強化された高品質な可視化を可能にする。

20

30

【0059】

図 10 は、マルチカメラ画像処理システムの一実施形態において生成および提供された視覚的表示の一例を示す図である。様々な実施形態において、表示 1000 は、例えば、図 8 および図 9 に示した処理およびシステムによって、本明細書で開示されているように、作業空間内の複数のカメラからの画像データに基づいて生成および表示されてよい。図の例において、表示 1000 は、カメラが（図に示すように右側に）取り付けられている吸着カップ式のエンドエフェクタ 1004 を備えたロボットアーム 1002 を含む作業空間（またはその一部）を示している。この例におけるロボットシステムは、テーブル 1006 からアイテムを取ってきて、それらを移動先位置（例えば、図 10 に示されていないパレットまたは他の移動先）へ移動させるよう構成されてよい。テーブル 1006 上に示されている物体 1008、1010、および、1012 は各々、物体インスタンスとして識別されており、図に示すようなフィルパターンは、物体を強調し物体を区別するために用いられる異なる色を表す。様々な実施形態において、表示 1000 は、例えば、2 つの隣接する物体が、単色および / または境界ボックス（またはその他の形状）内に示されている場合に、人間オペレータが、ロボットシステムによって操作される物体の分割におけるエラーを検出すること、ならびに / もしくは、遠隔操作でロボットアーム 1002 を制御すること、を可能にするために提供されてよい。

40

【0060】

図 11 は、ロボット制御のためのマルチカメラ画像処理システムにおいてセンサデータを処理するためのコードを生成する処理の一実施形態を示すフローチャートである。様々

50

な実施形態において、処理 1100 は、コンピュータ（図 1 の制御コンピュータ 118 など）によって実行される。図の例において、パイプライン定義が、受信およびパースされる（工程 1102）。例えば、パイプライン定義は、ユーザインターフェース、構成ファイル、APIなどを介して受信されてよい。定義されている処理パイプラインに含まれる処理構成要素のインスタンスが、パイプライン定義で定義されているように、作成およびリンクされる（工程 1104）。構成要素およびパイプラインを実装するためのバイナリコードが、（必要に応じて）コンパイルされ、配備される（工程 1106）。

【0061】

様々な実施形態において、本明細書で開示されている技術は、いくつかの実施形態におけるロボットアームまたはその他のロボットアクチュエータに取り付けられた 1 または複数のカメラなど、作業空間内の複数のカメラによって生成された画像データに基づいて、完全または部分的に自律動作でおよび/または完全または部分的に遠隔操作で、ロボット動作を実行するために用いられてよい。

【0062】

上述の実施形態は、理解しやすいようにいくぶん詳しく説明されているが、本発明は、提供されている詳細事項に限定されるものではない。本発明を実施する多くの代替方法が存在する。開示されている実施形態は、例示であり、限定を意図するものではない。

〔適用例 1〕

システムであって、

作業空間に関連付けられている複数のセンサの各々から画像データを受信するよう構成されている通信インターフェースと、前記画像データは、前記複数のセンサの内の各センサについて、視覚画像情報および深度情報の一方または両方を含み、

前記通信インターフェースに接続されているプロセッサと、

を備え、

前記プロセッサは、

前記複数のセンサからの画像データをマージして、マージ済み点群データを生成し、

前記複数のセンサの内の少なくとも一部の前記センサからの視覚画像データに基づいて分割を実行して、分割結果を生成し、

前記マージ済み点群データおよび前記分割結果の一方または両方を用いて、前記作業空間のマージ済み三次元分割ビューを生成するよう構成されている、システム。

〔適用例 2〕適用例 1 に記載のシステムであって、前記複数の画像センサは、1 または複数の三次元（3D）カメラを含む、システム。

〔適用例 3〕適用例 1 に記載のシステムであって、前記視覚画像データは、RGB データを含む、システム。

〔適用例 4〕適用例 1 に記載のシステムであって、前記プロセッサは、前記マージ済み点群データおよび前記分割結果の一方または両方を用いて、前記作業空間内の物体に関するボックス適合を実行するよう構成されている、システム。

〔適用例 5〕適用例 4 に記載のシステムであって、前記プロセッサは、さらに、前記マージ済み点群データおよび前記分割結果の一方または両方を用いて、ロボットアームで前記物体を把持するための戦略を合成するよう構成されている、システム。

〔適用例 6〕適用例 5 に記載のシステムであって、前記プロセッサは、さらに、前記ロボットアームで前記物体を把持するための前記戦略を実施するよう構成されている、システム。

〔適用例 7〕適用例 6 に記載のシステムであって、前記プロセッサは、前記物体を元の位置からピックして、前記物体を前記作業空間内の移動先位置にブレースするために、ロボット動作によって前記物体を把持するよう構成されている、システム。

〔適用例 8〕適用例 1 に記載のシステムであって、前記プロセッサは、さらに、前記作業空間の前記マージ済み三次元分割ビューを用いて、前記作業空間の視覚化表現を表示するよう構成されている、システム。

〔適用例 9〕適用例 8 に記載のシステムであって、前記表示された視覚化表現は、前記作

10

20

30

40

50

業空間内に描写された前記物体を強調している、システム。

〔適用例 1 0〕適用例 1 に記載のシステムであって、前記プロセッサは、前記マージ済み点群データおよび前記分割結果を用いて、少なくとも部分的には、前記分割結果を含む 1 セットの点を前記マージ済み点群データに逆投影することによって、前記作業空間のマージ済み三次元分割ビューを生成するよう構成されている、システム。

〔適用例 1 1〕適用例 1 に記載のシステムであって、前記プロセッサは、さらに、前記マージ済み点群データをサブサンプリングするよう構成されている、システム。

〔適用例 1 2〕適用例 1 1 に記載のシステムであって、前記プロセッサは、さらに、前記サブサンプリングされた点群データにクラスタ処理を実行するよう構成されている、システム。

10

〔適用例 1 3〕適用例 1 2 に記載のシステムであって、前記プロセッサは、前記サブサンプリングおよびクラスタリングされた点群データと前記分割結果とを用いて、前記作業空間内の物体に関するボックス適合結果を生成するよう構成されている、システム。

〔適用例 1 4〕適用例 1 に記載のシステムであって、前記プロセッサは、さらに、前記一部のセンサに含まれないセンサに関連付けられている視覚画像データに少なくとも部分的に基づいて、前記作業空間の前記マージ済み三次元分割ビューを検証するよう構成されている、システム。

〔適用例 1 5〕適用例 1 4 に記載のシステムであって、前記プロセッサは、少なくとも部分的には、前記作業空間の前記マージ済み三次元分割ビューを用いて、前記作業空間内の物体に関する第 1 ボックス適合結果を生成し、前記一部のセンサに含まれないセンサに関連付けられている視覚画像データを用いて、前記作業空間内の前記物体に関する第 2 ボックス適合結果を生成し、前記第 1 ボックス適合および前記第 2 ボックス適合を用いて、前記物体の検証済みボックス適合を決定することにより、前記一部のセンサに含まれない前記センサに関連付けられている前記視覚画像データに少なくとも部分的に基づいて、前記作業空間の前記マージ済み三次元分割ビューを検証するよう構成されている、システム。

20

〔適用例 1 6〕適用例 1 に記載のシステムであって、前記プロセッサは、少なくとも部分的には、ユーザ定義された処理パイプラインを実装することにより、前記画像データをマージして処理するよう構成されている、システム。

〔適用例 1 7〕適用例 1 6 に記載のシステムであって、前記プロセッサは、前記ユーザ定義された処理パイプラインの定義を受信してパースするよう構成されている、システム。

30

〔適用例 1 8〕適用例 1 7 に記載のシステムであって、前記プロセッサは、前記定義を用いて、前記パイプラインを構成するモジュールのインスタンスを作成し、前記モジュールおよびパイプラインを実装するためのバイナリコードを自動生成するよう構成されている、システム。

〔適用例 1 9〕

方法であって、

作業空間に関連付けられている複数のセンサの各々から画像データを受信し、前記画像データは、前記複数のセンサの内の各センサについて、視覚画像情報および深度情報の一方または両方を含み、

前記複数のセンサからの画像データをマージして、マージ済み点群データを生成し、前記複数のセンサの内の少なくとも一部の前記センサからの視覚画像データに基づいて分割を実行して、分割結果を生成し、

40

前記マージ済み点群データおよび前記分割結果の一方または両方を用いて、前記作業空間のマージ済み三次元分割ビューを生成すること、を備える、方法。

〔適用例 2 0〕

コンピュータプログラム製品であって、非一時的なコンピュータ読み取り可能媒体内に具現化され、

作業空間に関連付けられている複数のセンサの各々から画像データを受信するためのコンピュータ命令と、前記画像データは、前記複数のセンサの内の各センサについて、視覚

50

画像情報および深度情報の一方または両方を含み、
前記複数のセンサからの画像データをマージして、マージ済み点群データを生成するためのコンピュータ命令と、
前記複数のセンサの内の少なくとも一部の前記センサからの視覚画像データに基づいて分割を実行して、分割結果を生成するためのコンピュータ命令と、
前記マージ済み点群データおよび前記分割結果の一方または両方を用いて、前記作業空間のマージ済み三次元分割ビューを生成するためのコンピュータ命令と、
を備える、コンピュータプログラム製品。

【図面】

【図 1】

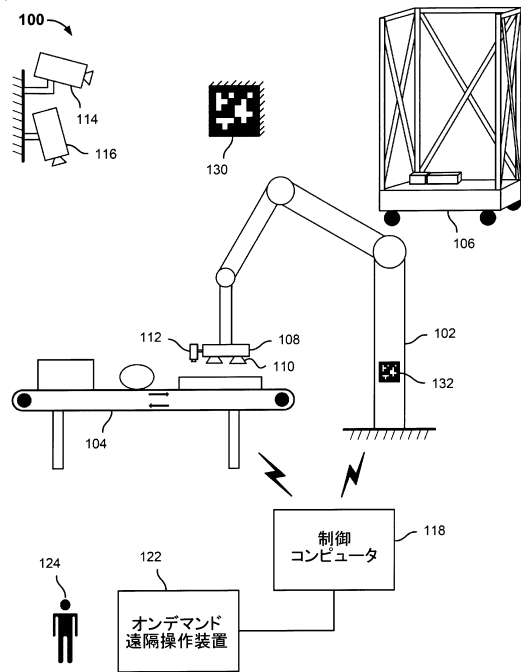


FIG. 1

【図 2】

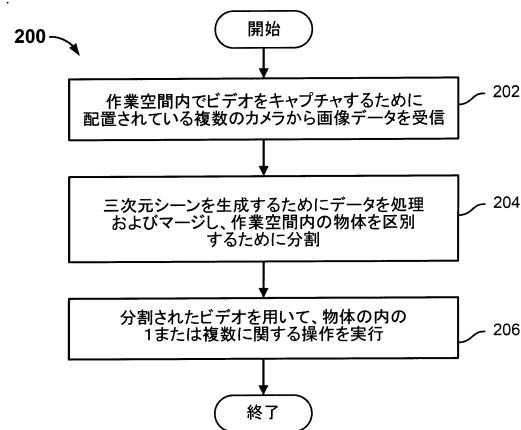


FIG. 2

10

20

30

40

50

【図 3】

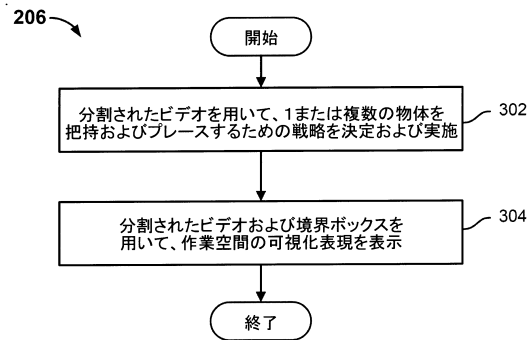


FIG. 3

【図 4】

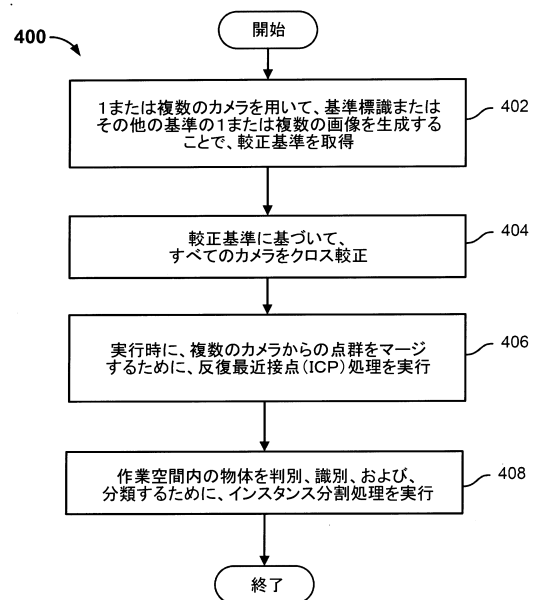


FIG. 4

【図 5】

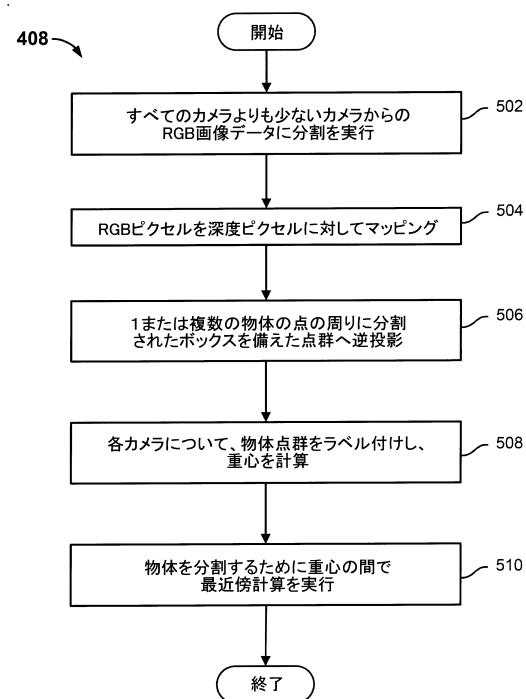


FIG. 5

【図 6】

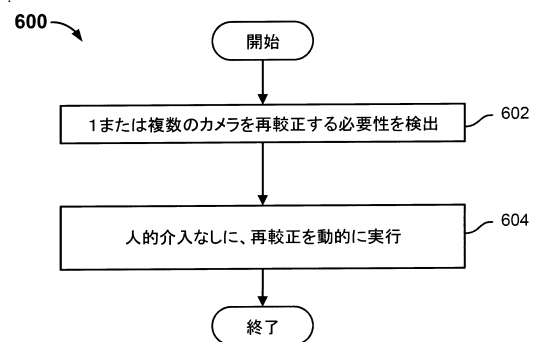


FIG. 6

10

20

30

40

50

【図 7】

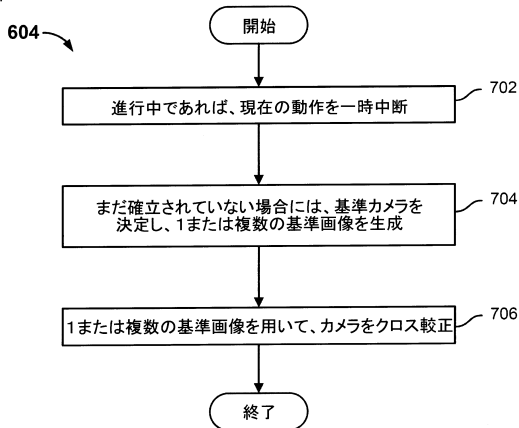


FIG. 7

【図 8】

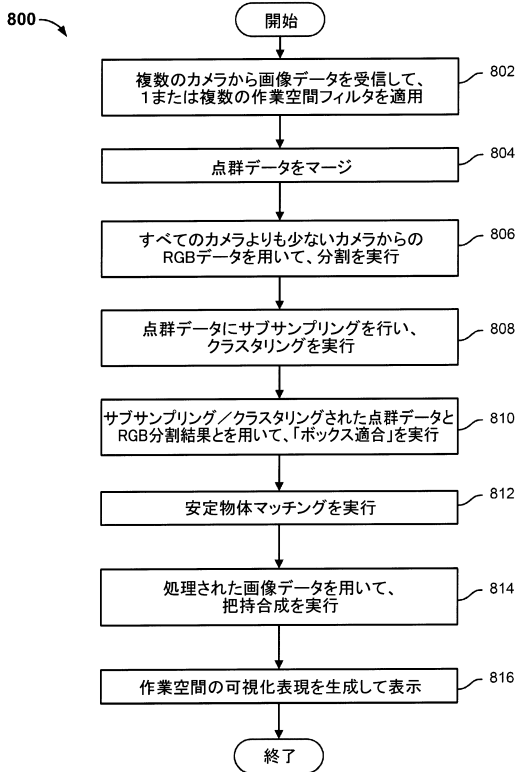


FIG. 8

【図 9 A】

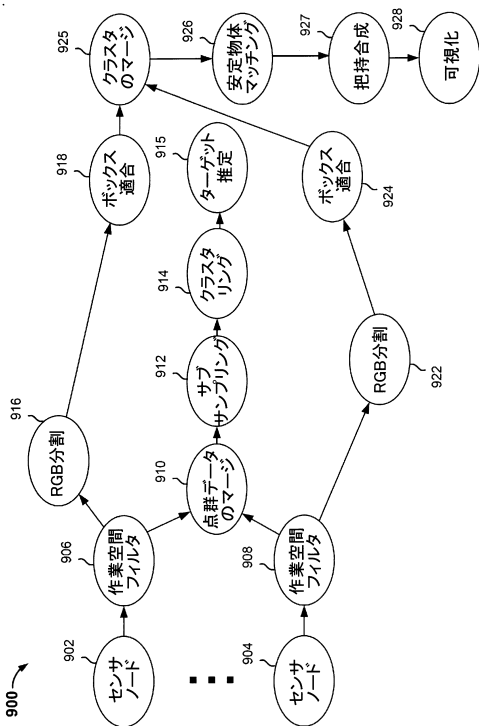


FIG. 9A

【図 9 B】

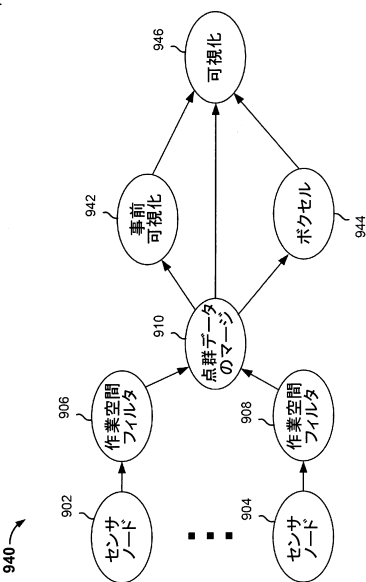


FIG. 9B

10

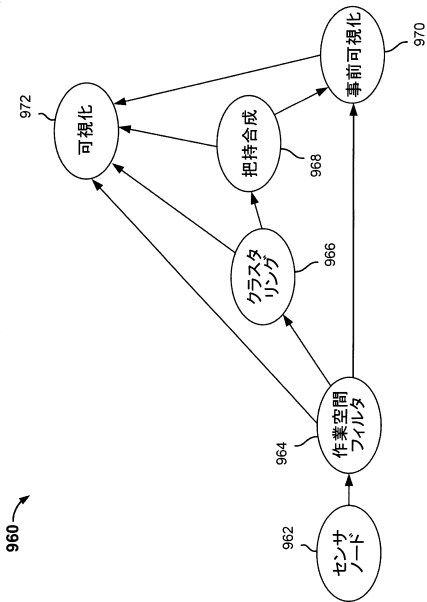
20

30

40

50

【図 9 C】



【図 1 0】

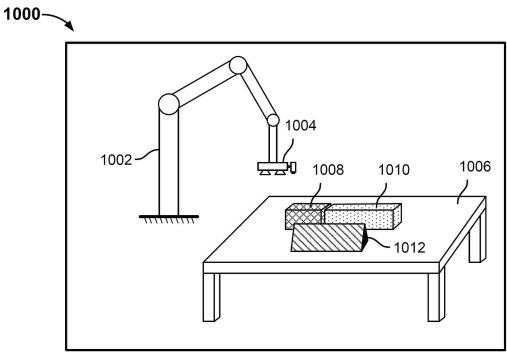


FIG. 9C

FIG. 10

【図 1 1】

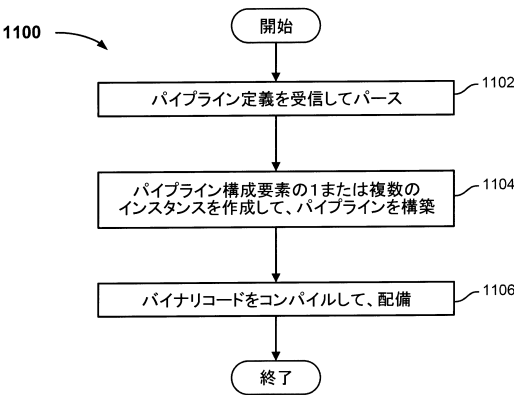


FIG. 11

10

20

30

40

50

フロントページの続き

(33)優先権主張国・地域又は機関

米国(US)

(31)優先権主張番号 16/380,859

(32)優先日 平成31年4月10日(2019.4.10)

(33)優先権主張国・地域又は機関

米国(US)

ベニユー, 292

(72)発明者 ビダパーシ・ロイト・アルカ

アメリカ合衆国 カリフォルニア州 94306 パロ・アルト, ランバート・アベニュー, 292

(72)発明者 モリス・ダウニング・タルボット

アメリカ合衆国 カリフォルニア州 94306 パロ・アルト, ランバート・アベニュー, 292

(72)発明者 スー・ハリー・ツェ

アメリカ合衆国 カリフォルニア州 94306 パロ・アルト, ランバート・アベニュー, 292

(72)発明者 メノン・サミア

アメリカ合衆国 カリフォルニア州 94306 パロ・アルト, ランバート・アベニュー, 292

審査官 佐田 宏史

(56)参考文献 特開2006-302195(JP, A)

特開2001-184500(JP, A)

特開2015-135331(JP, A)

米国特許出願公開第2018/0308254(US, A1)

米国特許出願公開第2015/0352721(US, A1)

米国特許第09802317(US, B1)

国際公開第2018/130491(WO, A1)

国際公開第2017/195801(WO, A1)

加藤 有祐、外5名, “三次元点群からの参照点候補の抽出と相対的位置概念の学習”, 第23回 画像センシングシンポジウム SSI 2017, 日本, 画像センシング技術研究会, 2017年06月07日, pp.1-6

西田 貴亮、外2名, “物体のパーツ分割と形状特徴に基づく移動可能性を考慮した種別分類”, 計測自動制御学会論文集, 日本, 公益社団法人計測自動制御学会, 2015年05月31日, Vol.51, No.5, pp.319-328

武口 智行、金子 俊一, “規則的ボクセル基準による距離アスペクト画像を用いたロバスト物体認識”, 電子情報通信学会論文誌, 日本, 社団法人電子情報通信学会, 2003年09月01日, Vol.J86-D-II, No.9, pp.1308-1319

Yu-Kai Chen et al., "Random Bin Picking with Multi-view Image Acquisition and CAD-Based Pose Estimation", 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 米国, IEEE, 2018年10月07日, pp.2218-2223

(58)調査した分野 (Int.Cl., DB名)

G06T 1/00, 7/00 - 7/90

G06V 10/00 - 10/98

G01B 11/00 - 11/30

B25J 13/08