

(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2021-163503

(P2021-163503A)

(43) 公開日 令和3年10月11日(2021.10.11)

(51) Int.Cl.	F I	テーマコード (参考)
G06T 7/00 (2017.01)	G06T 7/00 350C	5L096
G06N 3/08 (2006.01)	G06N 3/08	
G06N 20/00 (2019.01)	G06N 20/00	

審査請求 未請求 請求項の数 20 O L 外国語出願 (全 15 頁)

(21) 出願番号 特願2021-62699 (P2021-62699)
 (22) 出願日 令和3年4月1日(2021.4.1)
 (31) 優先権主張番号 16/839, 274
 (32) 優先日 令和2年4月3日(2020.4.3)
 (33) 優先権主張国・地域又は機関
 米国 (US)

(71) 出願人 390008235
 ファナック株式会社
 山梨県南部留郡忍野村忍草字古馬場358
 〇番地
 (74) 代理人 100099759
 弁理士 青木 篤
 (74) 代理人 100123582
 弁理士 三橋 真二
 (74) 代理人 100112357
 弁理士 廣瀬 繁樹
 (72) 発明者 トー タン
 アメリカ合衆国, カリフォルニア 945
 55, フリーモント, シー クリフ テラ
 ス 34913

最終頁に続く

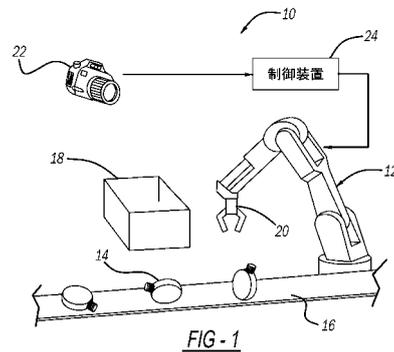
(54) 【発明の名称】 2次元カメラによる3次元の姿勢推定

(57) 【要約】

【課題】 2Dカメラからの2D画像及び学習ベースのニューラルネットワークを用いて、対象物の3Dポーズを取得するシステム及び方法の提供。

【解決手段】 ニューラルネットワークは、2D画像から対象物の複数の特徴を抽出し、抽出された特徴ごとに、色彩表現によって対象物の特徴点の位置の確率を識別するヒートマップを生成する。方法は、2D画像上のヒートマップからの特徴点を含む特徴点画像を提供し、特徴点画像を対象物の3D仮想CADモデルと比較することによって対象物の3Dポーズを推定する。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

現実の対象物の 3 次元ポーズを推定する方法であって、
2 次元カメラを用いて前記対象物の 2 次元画像を取得することと、
ニューラルネットワークを用いて、前記 2 次元画像から前記対象物の複数の特徴を抽出することと、
抽出された前記特徴ごとに、前記対象物の特徴点の位置の確率を識別する現実のヒートマップを生成することと、
前記ヒートマップからの特徴点と前記 2 次元画像とを結合する特徴点画像を提供することと、
前記特徴点画像を用いて、前記対象物の 3 次元ポーズを推定することと、
を含む方法。

10

【請求項 2】

前記対象物の 3 次元ポーズを推定することは、前記特徴点画像と前記対象物の 3 次元仮想モデルとを比較することを含む、請求項 1 に記載の方法。

【請求項 3】

前記対象物の 3 次元ポーズを推定することは、パースペクティブ n ポイントアルゴリズムを使用することを含む、請求項 2 に記載の方法。

【請求項 4】

前記ヒートマップにおける前記特徴点の位置の確率は、前記ヒートマップの色彩として表される、請求項 1 に記載の方法。

20

【請求項 5】

前記対象物の特徴点の位置の確率には、それが特徴点であるという信頼値が割り当てられる、請求項 1 に記載の方法。

【請求項 6】

マーカを有するマーカボード上において前記現実の対象物を表すトレーニング対象物を配置することを含むデータセットを取得することで前記ニューラルネットワークをトレーニングすることと、前記マーカボードの少なくとも 1 つのトレーニング 2 次元画像を得ることと、前記トレーニング 2 次元画像内のマーカの 3 次元ポーズを識別することと、前記マーカの前記 3 次元ポーズにオフセットを付与して、前記トレーニング対象物の特徴点を取得することと、各特徴点の位置の確率を特定する、個々のトレーニングヒートマップを生成することと、前記トレーニングヒートマップを使用して前記ニューラルネットワークをトレーニングすることと、をさらに含む、請求項 1 に記載の方法。

30

【請求項 7】

少なくとも 1 つのトレーニング 2 次元画像を得ることは、互いに異なる向きで複数のトレーニング 2 次元画像を得ることを含む、請求項 6 に記載の方法。

【請求項 8】

個々のトレーニングヒートマップを生成することは、ガウス分布プロセスを使用する、請求項 6 に記載の方法。

【請求項 9】

前記現実の対象物の複数の特徴を抽出することは、少なくとも 4 つの特徴を抽出することを含む、請求項 1 に記載の方法。

40

【請求項 10】

前記方法はロボットシステムにおいて実行され、前記現実の対象物はロボットにピックアップされる、請求項 1 に記載の方法。

【請求項 11】

ロボットによってピックアップされる対象物の 3 次元ポーズを推定する方法であって、
2 次元カメラを用いて前記対象物の 2 次元画像を取得することと、
収集されたトレーニングセットから、前記対象物の代表対象物を用いてトレーニングされた重みを備えたノードを含むニューラルネットワークを用いて、前記 2 次元画像から前

50

記対象物の複数の特徴を抽出することと、

抽出された前記特徴ごとに、前記対象物の特徴点の位置の確率を色彩で表示して識別するヒートマップを生成することと、

前記ヒートマップからの特徴点と前記２次元画像とを結合する特徴点画像を提供することと、

パースペクティブ n ポイントアルゴリズムを用いて、前記特徴点画像を前記対象物の３次元仮想モデルと比較することで現実の対象物の３次元ポーズを推定することと、を含む方法。

【請求項 1 2】

前記対象物の特徴点の位置の確率には、それが特徴点であるという信頼値が割り当てられる、請求項 1 1 に記載の方法。 10

【請求項 1 3】

前記対象物の複数の特徴を抽出することは、少なくとも４つの特徴を抽出することを含む、請求項 1 1 に記載の方法。

【請求項 1 4】

対象物の３次元ポーズを推定するシステムであって、
２次元カメラを用いて現実の対象物の２次元画像を取得する手段と、
ニューラルネットワークを用いて、前記２次元画像から前記対象物の複数の特徴を抽出する手段と、

抽出された前記特徴ごとに、前記対象物の特徴点の位置の確率を識別する現実のヒートマップを生成する手段と、 20

前記ヒートマップからの特徴点と前記２次元画像とを結合する特徴点画像を提供する手段と、

前記特徴点画像を用いて、前記対象物の３次元ポーズを推定する手段と、
を有するシステム。

【請求項 1 5】

前記対象物の３次元ポーズを推定する手段は、前記特徴点画像と前記対象物の３次元仮想モデルとを比較する、請求項 1 4 に記載のシステム。

【請求項 1 6】

前記対象物の３次元ポーズを推定する手段は、パースペクティブ n ポイントアルゴリズムを使用する、請求項 1 5 に記載のシステム。 30

【請求項 1 7】

前記ヒートマップにおける前記特徴点の位置の確率は、前記ヒートマップの色彩として表される、請求項 1 4 に記載のシステム。

【請求項 1 8】

前記対象物の特徴点の位置の確率には、それが特徴点であるという信頼値が割り当てられる、請求項 1 4 に記載のシステム。

【請求項 1 9】

収集されたトレーニングセットから、前記対象物の代表対象物を用いて前記ニューラルネットワークのノードをトレーニングする手段をさらに含む、請求項 1 4 に記載のシステム。 40

【請求項 2 0】

前記システムはロボット制御に使用され、前記対象物はロボットにピックアップされる、請求項 1 4 に記載のシステム。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本開示は、広くは、２次元カメラを使用して対象物の３次元の姿勢（３Ｄポーズ）を取得するためのシステム及び方法に関し、特に、２次元カメラ及び学習ベースのニューラルネットワークを使用して対象物の３次元の姿勢を取得するための、特定のロボット制御 50

用アプリケーションを含むシステム及び方法に関する。

【背景技術】

【0002】

ロボットは、ピックアンドブレース操作を含む多数のタスクを実行し、この操作では、ロボットが複数の対象物をピックアップして、コンベヤベルト等のある場所から、収集ビン等の他の場所に搬送するが、対象物の3Dポーズとして知られる位置及び向きは、僅かに異なる。そこで、ロボットが対象物を効率的にピックアップするために、多くの場合、ロボットが対象物の3Dポーズを知る必要がある。この目的のために、対象物の3Dポーズを決定するための種々の技術では、対象物の3D画像をリアルタイムで取得する3Dカメラが使用される。この場合、3D画像は、カメラの視野内の多くのポイントまでの距離測定値を識別し、これらの測定値はロボット制御装置に送信される。これらの技術では、構造化光を使用してシーンの3Dマップを作成し、対象物の3D特徴を識別して該マップ内の対象物を検索することができる。ロボットアプリケーションのために、対象物の3Dポーズを決定すべく3Dカメラを使用することは、効果的であることが証明されている。しかしながら、3Dカメラは高価なデバイスであり、多くの場合、必要とされるものよりも動作が遅い。

10

【0003】

ロボットアプリケーションのために対象物の3Dポーズを取得するための当技術分野で知られている技術の一つでは、代わりに2次元(2D)カメラを使用する。この技術は、シーンの単一の2Dカラー画像を取得し、完全畳み込み(fully-convolutional)アーキテクチャを使用して画像を処理する。アルゴリズムが、画像をS×Sセルを含む2Dグリッドに分割し、セル内の対象物を識別する。アルゴリズムは、画像内の各対象物の周囲に境界ボックスを投影し、該ボックスの8つのコーナの2D位置を予測する。しかしながら、境界ボックスのコーナを予測するこの技術では、精度に制限があり、オクルージョンに対するロバスト性(robustness to occlusion)が低い。

20

【0004】

対象物の3Dポーズを決定するための上記プロセスのいくつかは、人工知能(AI)を使用する。これは、アルゴリズムを使用するコンピュータサイエンスの一部であり、該アルゴリズムによって、ソフトウェアアプリケーションが環境から学習し、そこから決定を下して、特定の結果を達成することができる。機械学習は、ソフトウェアアプリケーションを使用するAIの一部であり、該ソフトウェアアプリケーションは、大量の生の入力データを繰り返し分析することで独自の知識を取得し、該データからパターンを抽出して、該ソフトウェアアプリケーションがあるタスクを実行するために特別にプログラムされていなくても、該タスクの実行を学習できる。深層学習(ディープラーニング)は、特定の現実世界の環境を、増加する複雑な概念の階層として表すことにより、より優れた学習パフォーマンスを提供する特定のタイプの機械学習である。

30

【0005】

深層学習は典型的には、非線形処理を実行するニューラルネットワークの複数の層(レイヤ)を含むソフトウェア構造を使用し、この場合、前の層からの出力を次の層の各々が受け取る。一般に層は、センサから生データを受信する入力層と、該データから抽象的な特徴を抽出する複数の隠れ層と、該隠れ層からの特徴抽出に基づいて特定のものを識別する出力層とを含む。ニューラルネットワークは、各々が「重み」を有するニューロン又はノードを含み、この重みにノードへの入力を乗算することで、あることが正しいか否かの確率を取得することができる。より具体的には、各ノードは浮動小数点数である重みを有し、この重みはノードへの入力と乗算され、該入力の割合である、そのノードの出力が生成される。この重みは最初に、監視された処理の下でニューラルネットワークに既知のデータのセットを分析させることと、コスト関数を最小化してネットワークが正しい出力を最高の確率で取得できるようにすることとにより、「トレーニングされ(trained)」又は設定されている。

40

【0006】

50

深層学習ニューラルネットワークは、画像内の対象物の視覚的な検出及び分類のための、画像特徴抽出及び変換を行うために、頻繁に使用される。画像のビデオ又はストリームは、該ネットワークによって分析可能であり、それにより対象物の識別・分類と、プロセスを通じて対象物をよりよく認識するための学習が可能となる。従って、これらのタイプのネットワークでは、システムは同じ処理構成を使用して特定の対象物を検出し、アルゴリズムが対象物を認識するためにどのように学習したかに基づいて、対象物を異なる方法で分類することができる。

【発明の概要】

【0007】

以下の説明は、2次元カメラからの2次元画像及び学習ベースのニューラルネットワークを使用して対象物の3次元の姿勢（ポーズ）を取得するためのシステム及び方法を開示・説明する。ニューラルネットワークは、2次元画像から対象物の複数の特徴を抽出し、抽出された特徴ごとに、対象物の特徴点の位置の確率を色彩表現によって特定するヒートマップを生成する。この方法は、2次元画像上のヒートマップの特徴点を含む特徴点画像を提供し、特徴点画像と対象物の3次元仮想CADモデルとを比較することによって、対象物の3次元の姿勢を推定する。

10

【0008】

本開示の追加の特徴は、添付の図面と併せて、以下の説明及び特許請求の範囲から明らかになるであろう。

【図面の簡単な説明】

20

【0009】

【図1】対象物をコンベヤベルトからピックアップしてピン内に配置するロボットを含むロボットシステムの図であり、該システムは対象物の2D画像を取得する2Dカメラと、2D画像を用いて対象物の3Dポーズを計算する制御装置とを使用する。

【図2】ニューラルネットワークをトレーニングするために、2Dカメラを使用して対象物からデータセットを取得するプロセスを示すフロータイプの図である。

【図3】調整可能な重みを各々が有する複数のノードを含むニューラルネットワークの図である。

【図4】対象物の2D画像及びニューラルネットワークを使用して対象物の3Dポーズを推定するために、トレーニングされたニューラルネットワークを使用する学習ベースのニューラルネットワークプロセスを示すフロータイプの図である。

30

【図5】図4に示すプロセスにおいて、対象物の3Dポーズ推定を決定するためのパーспекティブnポイント（PnP）プロセスを示す図である。

【図6】対象物を撮像する複数のカメラを使用するカメラシステムの図である。

【図7】複数の2Dカメラを使用して対象物の3Dポーズを推定するシステムのブロック図である。

【図8】図7に示すシステムにおいて、対象物の3Dポーズを推定するためのPnPプロセスを示す図である。

【図9】対象物上の特徴点を予測するためのベクトル場推定プロセスを示すフロータイプの図である。

40

【発明を実施するための形態】

【0010】

2Dカメラを使用して対象物の3Dポーズを取得するために学習ベースのニューラルネットワークを使用するシステム及び方法を対象とする本開示の実施形態の以下の説明は、本質的に単なる例示であり、本発明又はそのアプリケーション若しくは使用を限定することを意図するものではない。例えば、システム及び方法は、ロボットに把持された対象物の位置及び向きを求めるアプリケーションを有する。しかしシステム及び方法は、他のアプリケーションを有してもよい。

【0011】

図1は、対象物14をコンベヤベルト16からピックアップして収集ピン18内に配置

50

するロボット12を含むロボットシステム10の図である。システム10は、本明細書の記載から利益を享受できる任意のタイプのロボットシステムを表すことを意図しており、ロボット12は、その目的に適した任意のロボットであり得る。各対象物14は、コンベヤベルト16上を下流側に移動するとき、他の対象物14に対して、コンベヤベルト16上での向き及び位置が異なる。ロボット12が対象物14を効果的に把持及びピックアップするためには、ロボット12のエンドエフェクタ20が物体14を把持する前に、エンドエフェクタ20を適切な位置及び姿勢に配置できる必要がある。これを達成するために、システム10は、対象物14に対して所望の位置に設けられる2Dカメラ22を含み、2Dカメラ22は、エンドエフェクタ20を動かすべくロボット12を制御するロボット制御装置24に2D画像を提供する。以下で詳述するように、ロボット制御装置24は、エンドエフェクタ20が適切な向きになるように、アルゴリズムを使用して、カメラ22からの2D画像を対象物14の3Dポーズに変換する。

10

【0012】

後述するように、制御装置24が対象物14の3Dポーズを正確に決定するために、一実施形態では、学習ベースのニューラルネットワークが、制御装置24で実行されて対象物の特徴を抽出するアルゴリズム及びソフトウェアにおいて使用される。ニューラルネットワークは、代表対象物から取得したデータでトレーニングする必要がある。図2は、ニューラルネットワークをトレーニングするために使用される2Dカメラ40を使用して、システム10内の対象物14と同一の対象物32から画像データを取得するためのプロセスを示すフロータイプのダイアグラム30である。対象物32は、適切な数のマーカ36を有するマーカボード34上に配置され、マーカ36は、対象物32のグラウンドトゥールズ又は実際の位置を識別するために使用される特別な検出可能なデザインを備える。ボード34上のマーカ36のうちの一つのみが図示されており、マーカ34と対象物32との間の線38によって表されるオフセット距離が測定される。

20

【0013】

カメラ40は、マーカボード34の一つ又は複数の2D画像42、ここでは4つの画像を、互いに異なる角度で提供する。次にプロセスは、画像42上での投影を実行して、座標フレーム44によって識別されるマーカ36のうちの一つの、回転及び位置を求める。次にアルゴリズムは、マーカ36と対象物32との間の、以前に測定されたオフセットをマーカ36の回転及び位置に付加し、対象物32の回転及び位置を求める。対象物32の真の位置が決定されれば、対象物32のコーナ等の特徴点46の位置は、既知の特徴点抽出プロセスを使用して計算することができるが、ここでは少なくとも4つの特徴点が必要である。例えばガウス分布プロセスを使用可能なアルゴリズムは、画像42の識別された各特徴点46を見て、その点46の「ヒートマップ」48を生成し、ここでは、各ヒートマップ48は、色の強度によって、対象物32の特徴点の一つが特定の場所に存在する可能性を求める。例えば、領域50等の、特徴点があると考えられる画像42内の場所には、赤等の特定の色が割り当てられる。また特徴点領域50から離れた場所に存在する可能性は低いので、色は、例えば領域52のように青に変化する。

30

【0014】

次に、2D画像42と、ダイアグラム30で示すプロセスによって生成された対応するヒートマップ48とによって提供されるデータは、既知のAIトレーニングスキームを使用してニューラルネットワークをトレーニングするために使用され、そして対象物14の3Dポーズを識別するためにシステム10で使用される。図3は、各々が調整可能な重みWを有する複数のノード62を含むニューラルネットワーク60の図であり、ネットワーク60は、本明細書で論じられるニューラルネットワークを概略的に表すことを意図している。ニューラルネットワーク60は、2D画像から個々のピクセルデータを受信する入力層64と、2D画像のピクセルデータ内の特徴を識別する複数の残差ブロック層66及び68と、複数の畳み込み層70とを有し、畳み込み層70の各ノード62はヒートマップを提供する。畳み込み層70並びに残差ブロック層66及び68は、深層学習の領域で利用される標準モジュールであることに留意されたい。これらのタイプの畳み込み層は、

40

50

画像の入力又は前の畳み込み層に対して畳み込み操作を実行する複数のフィルタカーネルで構成されているため、エッジ、テクスチャ又は輪郭情報が入力から抽出可能である。残差ブロック層は、スキップ接続を持つ2つの畳み込み層を含む構造であり、第2の畳み込み層の出力が第1の畳み込み層の入力に付加されて、ブロック出力として機能する。

【0015】

図4は、制御装置24で動作するアルゴリズムを示すフロータイプのダイアグラム80であり、該アルゴリズムが採用する学習ベースのニューラルネットワーク78は、トレーニングされたニューラルネットワークを使用して、カメラ22で撮像された対象物14の2D画像82から、対象物14の3Dポーズを推定する。画像82は、入力層84と、制御装置24のAIソフトウェアで動作するニューラルネットワーク78内のフィードフォワードループを含む複数の連続する残差ブロック層86及び88とに提供され、フィルタリングプロセスを使用した、画像82内の対象物14の可能な特徴点の勾配、エッジ、輪郭等の特徴抽出が行われる。抽出された特徴を含む画像は、ニューラルネットワーク78内の複数の連続する畳み込み層90に提供され、抽出された特徴から得られる可能な特徴点は、特徴点ごとに1つずつ、一連のヒートマップ92として定義される。ヒートマップ92は、ヒートマップ92の色に基づいて、特徴点が対象物14上に存在する可能性を表している。画像94は、全てのヒートマップ92からの全ての特徴点のための特徴点96を含む、対象物14の画像82を使用して生成される。各特徴点96には、その特徴点のヒートマップ92の色に基づく信頼値が割り当てられ、特定の閾値を超える信頼値を持たない特徴点96は使用されない。

10

20

【0016】

次に画像94は、ポーズ推定プロセッサ98内に同じ特徴点を有する、対象物14の公称又は仮想の3DCADモデルと比較されて、対象物14の推定3Dポーズ100を提供する。画像94をCADモデルと比較するための適切なアルゴリズムの1つは、当技術分野ではパースペクティブnポイント(PnP)として周知である。一般に、PnPプロセスは、ワールド座標フレーム内の対象物のn個の3Dポイントのセットと、カメラからの画像内の対応する2D投影とが与えられてキャリブレーションされたカメラに対する対象物のポーズを推定する。ポーズには、カメラの座標フレームに対する対象物の回転(ロール、ピッチ、ヨー)と3D平行移動とで構成される6自由度(DOF)が含まれる。

30

【0017】

図5は、対象物14の3Dポーズを取得するために、本実施例ではPnPプロセスがどのように実行され得るかを示すイラスト104である。イラスト104は、グラウンドトゥールズ又は現実位置にある対象物14を表す3D対象物106を示す。対象物106は、カメラ22を表すカメラ112によって観測され、2D画像平面110上に2D対象物画像108として投影される。ここでは、対象物画像108は画像94を表し、画像108上の点102はニューラルネットワーク78によって予測され、点96を表す特徴点である。またイラスト104は、特徴点96と同じ位置に特徴点132を有する対象物14の仮想3DCADモデル114を示しており、モデル114は、カメラ112の前にランダムに配置されるとともに、投影された特徴点118も含む2Dモデル画像116として画像平面110上に投影される。CADモデル114は、カメラ112の前で回転及び並進させられ、それによりモデル画像116は、モデル画像116の各特徴点118と対象物画像108の対応する特徴点102との間の距離が最小化するように、すなわち画像116と画像108とが位置合わせするように、回転及び並進させられる。モデル画像116が対象物画像108に可能な限り位置合わせされると、カメラ112に対するCADモデル114のポーズは、対象物14の推定された3Dポーズ100となる。

40

【0018】

この分析は、画像108と画像116との間の対応する特徴点の1つについて式(1)で表され、式(1)は、画像108及び116の全ての特徴点について使用される。式(1)において、 V_i はCADモデル114の特徴点132の1つであり、 v_i はモデル画像116の対応する投影された特徴点102であり、 a_i は対象物画像108の特徴点1

50

図9は、カメラ22によって撮像された対象物14の2D画像82を使用して対象物14の3Dポーズを生成する制御装置24で動作するベクトル場推定プロセスを示すフロータイプのダイアグラム180であり、同様の要素は同じ参照符号で示される。画像入力は、3次元行列 $N \times N \times 3$ によって定義される3チャンネルRGB(赤-緑-青)画像であり、ここで N は、画像82の高さ及び幅である。画像82は、エンコーダ/デコーダニューラルネットワーク182に提供され、このネットワーク182は、画像82内の可能な主要特徴を抽出するエンコーダネットワーク184と、抽出された特徴から画像を生成して、以下で説明する対象物14の可能な特徴点を識別するデコーダネットワーク186とを含む。ここでネットワーク184及び186は、上述したタイプの複数のニューラルネットワーク層を含む。特に、ある非限定的な実施形態では、エンコーダネットワーク184のニューラルネットワーク層は、畳み込み層188、バッチ正規化(BN)層190及び正規化線形ユニット(ReLU)層192を含み、デコーダネットワーク186のニューラルネットワーク層は、プーリング層194、アップサンプリング層196及びソフトマックス層198を含む。

10

20

30

40

50

【0029】

ネットワーク182は、画像82の3次元行列を処理し、3つのチャンネルから、1次元($N \times N \times 1$)マスク画像200と、識別された各特徴点について2次元($N \times N \times 2$)ベクトル場画像202とを生成する。マスク画像200を生成するために、ネットワーク182は、対象物14の一部が特定のピクセルに存在するか否かに基づいて、画像82内の各ピクセルにデジタル1又は0を割り当てる。ここでマスク画像200の明るい領域のピクセルはデジタル1を有し、マスク画像200の暗い領域のピクセルはデジタル0を有する。各ベクトル場画像202は、画像82内の各ピクセルに与えられる x 方向ベクトル値及び y 方向ベクトル値によってそれぞれ定義される複数の矢印を含み、矢印の方向は、画像82の特徴点の1つに向かっている。故に画像82内の各ピクセルは、その特徴点への方向を予測する能力を有する。マスク画像200内のピクセルの値は、ベクトル場画像202内の対応するピクセルの値で乗算されるので、対象物14の一部を含むピクセルに関連付けられておらず、かつ特徴点に対して有用な方向を提供しない矢印は、ベクトル場抽出画像204に示されるように、破棄される。

【0030】

ベクトル場抽出画像204内の各矢印は、特徴点の1つへの方向を表すが、その特徴点までの距離は示さない。そこで、ベクトル場抽出画像204内の、値を有する2つの対象物ピクセル p_i 及び p_j の全ての組み合わせについて、プロセスは点 x_{ij} を識別する。点 x_{ij} のベクトル v_i 及び v_j は、ベクトル交差画像206に示すように、互いに交差する。各点 x_{ij} には、ピクセル p_i と点 x_{ij} との間の距離、ピクセル p_j と点 x_{ij} との間の距離、並びにピクセル p_i 及び p_j での勾配の大きさに応じて、スコアが割り当てられる。例えば、ピクセル p_i と点 x_{ij} との間の距離、及びピクセル p_j と点 x_{ij} との間の距離が小さいほどスコアは高くなり、ピクセル p_i 及び p_j での勾配の大きさが大きいほどスコアは高くなる。その結果、各々がスコアを有する多くの点 x_{ij} を含むグリッドが得られ、殆どの点 x_{ij} を含む画像内の領域は、点投票画像208に示されるように、可能性のある特徴点の位置を識別する。

【0031】

ベクトル場画像202は、ネットワーク182によって識別された可能な1つの特徴点のベクトル矢印を示す。ネットワーク182はまた、マスク200によって乗算された可能な特徴点ごとに別個のベクトル場画像202を生成する。全ての点投票画像208の全ての特徴点は、特徴点96を含む画像94のように、単一の画像内で結合され、個別の特徴点96の各々は、点投票画像208の1つによって提供される。次に画像94は、ポーズ推定プロセッサ98内でPnPプロセスによって処理されて、上述のやり方で3Dポーズ100を生成することができる。さらに、上述したような単一のカメラの光軸に沿った不確実性のために、複数のカメラを使用することができ、それらのカメラからの画像の各々が、ベクトル場推定プロセスによって処理される。

【0032】

前述のように、ベクトル場推定プロセスは、特徴点の抽出を必要とする他のプロセスに適用され得る。例えば、ロボット工学のアプリケーションの1つでは、ロボットによってピックアップされている対象物の中心を探索するが、そこでは対象物の向きは必要ない。別のアプリケーションは、品質保証の目的で対象物の測定値を決定することであり得る。

【0033】

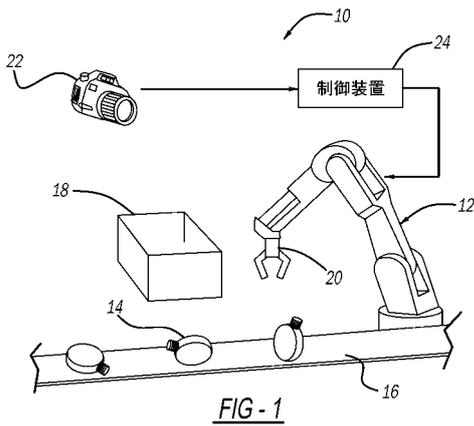
当業者によってよく理解されるように、本発明を説明するために本明細書で論じられるいくつかの様々なステップ及びプロセスは、電氣的現象を使用してデータを操作及び/又は変換するコンピュータ、プロセッサ又は他の電子計算デバイスによって実行される操作を意味し得る。それらのコンピュータ及び電子デバイスは、コンピュータ又はプロセッサによって実行可能な種々のコード又は実行可能命令を有する実行可能プログラムが格納された非一時的なコンピュータ可読媒体を含む、様々な揮発性及び/又は不揮発性メモリを使用することができる。メモリ及び/又はコンピュータ可読媒体は、あらゆる形態及びタイプの、メモリ及び他のコンピュータ可読媒体を含み得る。

10

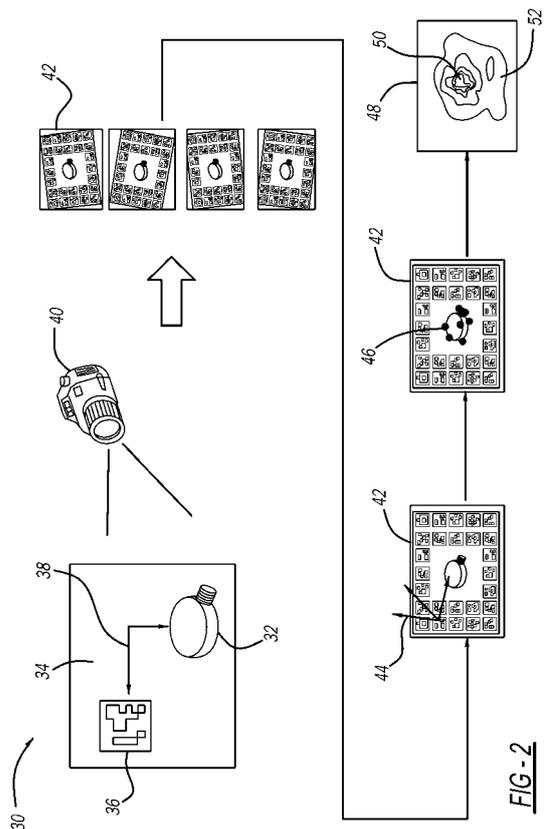
【0034】

上述の説明は、本開示の単なる例示的な実施形態を開示及び記載している。当業者は、そのような説明、並びに添付の図面及び特許請求の範囲から、以下の特許請求の範囲で規定される本開示の精神及び範囲から逸脱することなく、様々な変更、修正及び変形を行うことができることを容易に認識するであろう。

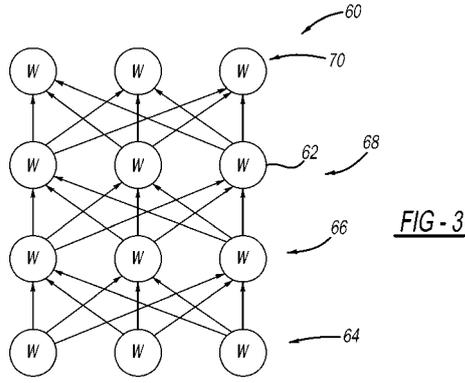
【図1】



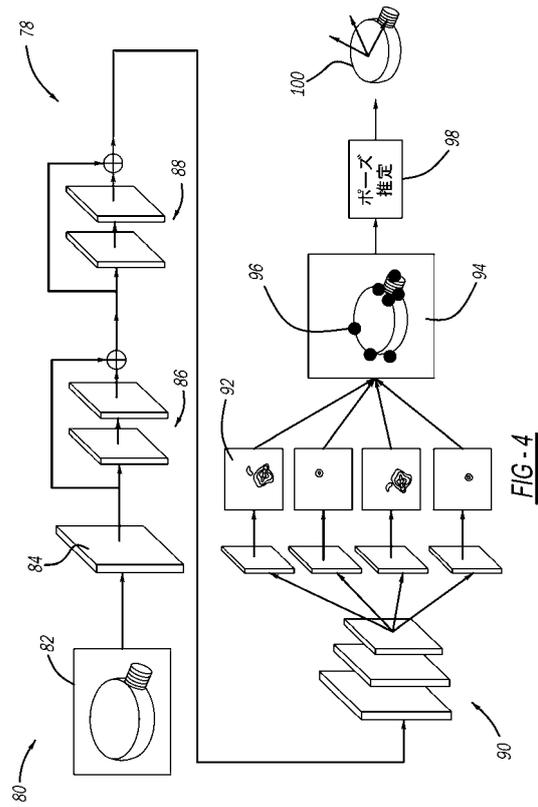
【図2】



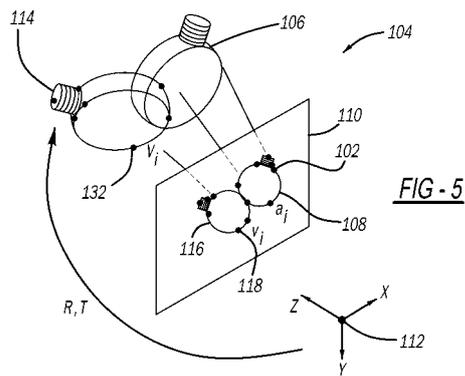
【 図 3 】



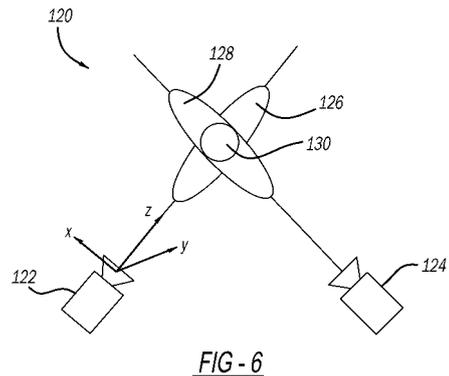
【 図 4 】



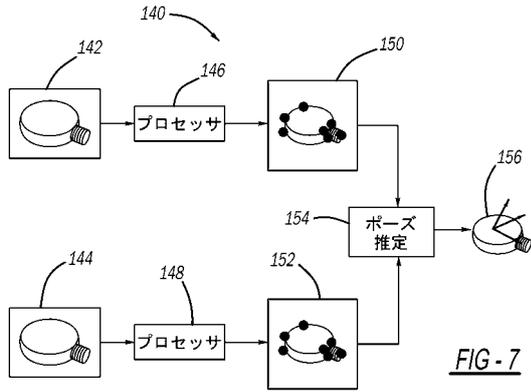
【 図 5 】



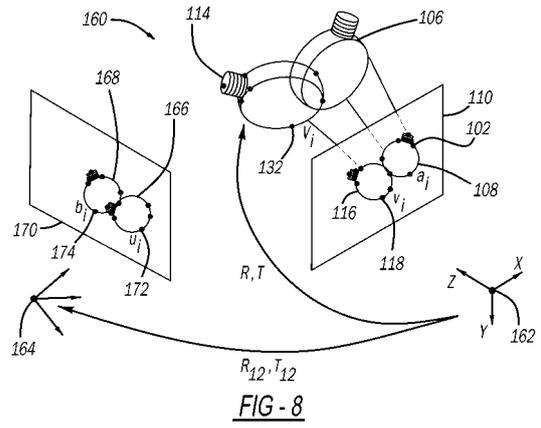
【 図 6 】



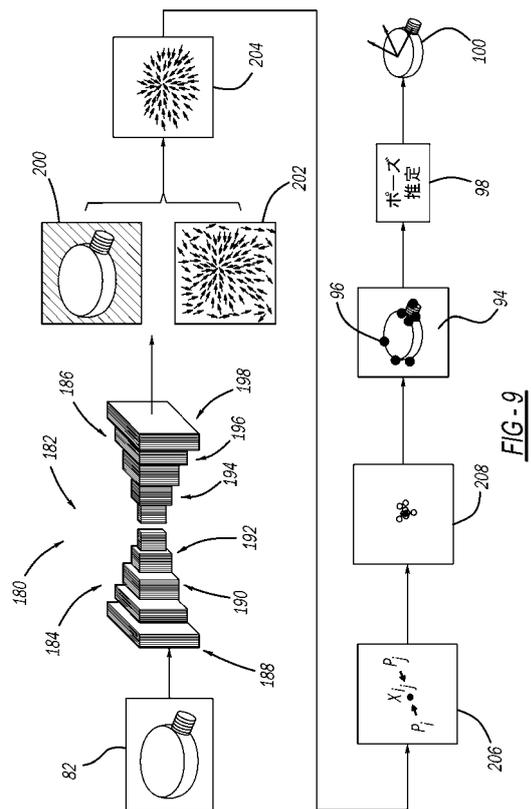
【 図 7 】



【 図 8 】



【 図 9 】



フロントページの続き

(72)発明者 加藤 哲朗

アメリカ合衆国, カリフォルニア 9 4 5 5 5, フリーモント, ランチョ デル サー ドライブ
5 3 8 5

Fターム(参考) 5L096 BA05 CA02 DA02 FA06 FA67 FA69 GA02 HA11 JA11 KA04

【外国語明細書】

2021163503000001.pdf