

(19)日本国特許庁(JP)

(12)公開特許公報(A)

(11)公開番号

特開2023-55217

(P2023-55217A)

(43)公開日 令和5年4月17日(2023.4.17)

(51)国際特許分類	F I	テーマコード(参考)
G 0 6 T 7/73 (2017.01)	G 0 6 T 7/73	3 C 7 0 7
G 0 6 T 7/00 (2017.01)	G 0 6 T 7/00	5 L 0 9 6
B 2 5 J 13/08 (2006.01)	B 2 5 J 13/08	A

審査請求 未請求 請求項の数 11 O L 外国語出願 (全23頁)

(21)出願番号	特願2022-160018(P2022-160018)	(71)出願人	390023711
(22)出願日	令和4年10月4日(2022.10.4)		ローベルト ボツシュ ゲゼルシャフト
(31)優先権主張番号	10 2021 211 185.8		ミット ベシユレンクテル ハフツング
(32)優先日	令和3年10月5日(2021.10.5)		ROBERT BOSCH GMBH
(33)優先権主張国・地域又は機関	ドイツ(DE)		ドイツ連邦共和国 シュツツトガルト (番地なし)
			Stuttgart, Germany
		(74)代理人	100114890
			弁理士 アインゼル・フェリックス=ライ
			インハルト
		(74)代理人	100098501
			弁理士 森田 拓
		(74)代理人	100116403
			弁理士 前川 純一
		(74)代理人	100134315

最終頁に続く

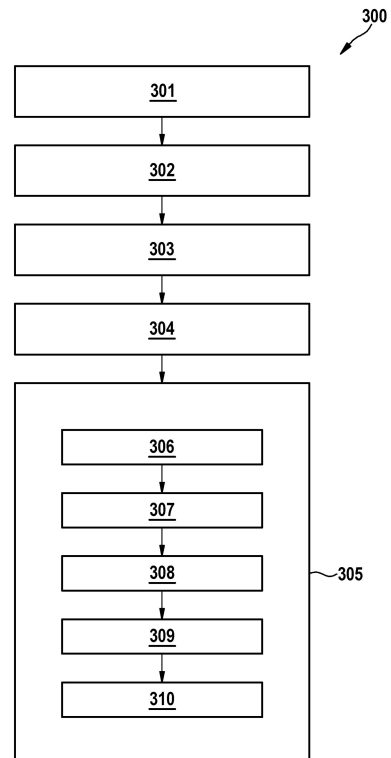
(54)【発明の名称】 ロボットを制御するための装置及び方法

(57)【要約】 (修正有)

【課題】ロボットを制御する装置及び方法を提供する。

【解決手段】ロボットを制御するための方法は、デモンストレーションを実施し301、物体に対するロボットの視点から、このデモンストレーションのために記述子画像を供給し302、記述子画像内において可視である選択された特徴点に報酬を与える目的関数の最適条件を探索することによって選択される、複数の特徴点から成る集合を選択し303、デモンストレーションを用いてロボット制御モデルをトレーニングする304。さらに、物体の記述子画像を求め306、複数の特徴点から成る選択された集合を物体の記述子画像内に配置し307、配置された特徴点のユークリッド座標を求め308、求められたユークリッド座標から姿勢を推定し309、推定した姿勢と共にロボット制御モデルを用いて、物体を処理するようにロボットを制御する310ことで物体を含む制御シーンに対してロボットを制御する305。

【選択図】図3



【特許請求の範囲】**【請求項 1】**

ロボットを制御するための方法であって、

デモンストレーションを実施するステップであって、各デモンストレーションは、物体の処理を実際に行う、ステップと、

デモンストレーションごとに、前記物体に対する前記ロボットの視点から、少なくとも 1 つの記述子画像を供給するステップであって、前記記述子画像によって、前記物体上のロケーションに関する特徴点が指定される、ステップと、

指定された前記特徴点から、複数の特徴点から成る集合を選択するステップであって、前記特徴点は、前記記述子画像内において可視である選択された特徴点に報酬を与える目的関数の最適条件を探索することによって選択される、ステップと、 10

デモンストレーションを用いてロボット制御モデルをトレーニングするステップであって、前記ロボット制御モデルは、入力された物体姿勢に依存して制御情報を出力するように構成されている、ステップと、

物体を含む制御シーンに対しロボットを制御するステップであって、

前記ロボットの視点から前記物体の記述子画像を求め、

複数の特徴点から成る選択された前記集合を前記物体の前記記述子画像内に配置し、

前記制御シーンに関して配置された前記特徴点のユークリッド座標を求め、

求められた前記ユークリッド座標から姿勢を推定し、

推定された前記姿勢が入力として前記ロボット制御モデルに供給され、前記ロボット制御モデルを用いて、前記物体を処理するように前記ロボットを制御することによって、ロボットを制御するステップと、 20
を含む方法。

【請求項 2】

前記目的関数は、選択された特徴点が記述子空間内において離間されていること、選択された前記特徴点に対応する前記物体上のロケーションがユークリッド空間内において離間されていること、及び、前記物体に対し選択された特徴点に関する検出誤差が小さいことのうちの 1 つ又は複数に、さらに報酬を与える、

請求項 1 に記載の方法。

【請求項 3】

前記物体に 1 つの平面をマッチングさせるステップと、
前記特徴点が前記平面上の座標フレームを規定するように前記特徴点を選択するステップと、 30

配置された前記特徴点の求められた前記ユークリッド座標と、マッチングされた前記平面の前記姿勢に関する情報とから、前記姿勢を推定するステップと、
をさらに含む、

請求項 1 又は 2 に記載の方法。

【請求項 4】

求められた前記ユークリッド座標から前記姿勢を推定するステップは、配置された前記特徴点の前記ユークリッド座標をマッチングされた超平面に投影するステップを含む、 40

請求項 3 に記載の方法。

【請求項 5】

前記平面が前記物体にマッチングされ、前記特徴点は、前記特徴点が前記平面上の 1 つの座標フレームを規定するように選択され、前記姿勢は、1 つの空間方向における前記物体の変化が予め定められた閾値を下回っている場合には、配置された前記特徴点の求められた前記ユークリッド座標と、前記平面の前記姿勢に関する情報とから推定される、

請求項 3 又は 4 に記載の方法。

【請求項 6】

選択された前記特徴点のユークリッド座標から座標フレームの導出ルールを求めるステップを含み、

求められた前記ユークリッド座標から前記姿勢を推定するステップは、選択された前記特徴点に前記導出ルールを適用するステップを含み、前記導出ルールは、前記ユークリッド座標におけるノイズから前記座標フレームの依存性の最小値を探索することによって求められる、

請求項 1 乃至 5 のいずれか一項に記載の方法。

【請求項 7】

前記デモンストレーションを用いて前記ロボット制御モデルをトレーニングするステップは、

デモンストレーションごとに、複数の特徴点から成る選択された前記集合を前記物体の前記記述子画像内に配置するステップと、

前記デモンストレーションのために配置された前記特徴点のユークリッド座標を求めるステップと、

前記デモンストレーションのために求められた前記ユークリッド座標から姿勢を推定するステップと、

を含む、

請求項 1 乃至 6 のいずれか一項に記載の方法。

【請求項 8】

前記物体の前記記述子画像を、高密度オブジェクトネットにより前記物体のカメラ画像から求めるステップを含む、

請求項 1 乃至 7 のいずれか一項に記載の方法。

【請求項 9】

請求項 1 乃至 8 のいずれか一項に記載の方法を実施するように構成されたロボットコントローラ。

【請求項 10】

コンピュータにより実行されるときに、前記コンピュータに請求項 1 乃至 8 のいずれか一項に記載の方法を実施させるための命令を含むコンピュータプログラム。

【請求項 11】

コンピュータにより実行されるときに、前記コンピュータに請求項 1 乃至 8 のいずれか一項に記載の方法を実施させるための命令を含むコンピュータ可読媒体。

【発明の詳細な説明】

【技術分野】

【0001】

従来技術

本開示は、ロボットを制御するための装置及び方法に関する。

【背景技術】

【0002】

デモンストレーションに基づく学習 (LfD) というアプローチを通して、ロボットのスキルをプログラミングすることができ、その際にスキルの公称プランが、デモンストレーションに基づきロボットによって学習される。LfDの主要な着想は、ロボットのワークスペースを監視するカメラ及びロボットにより処理されるべきターゲット物体の姿勢によって、ガウス分布をパラメータ化することである。この目的で、また、結果として生じるロボット制御モデルを利用するためにも、所与のシーンにおける物体の姿勢の推定を実施する必要がある。1つの可能性は、RGB画像に基づく姿勢推定技術である。しかしながら、これらは信頼性に欠ける場合がある。

【発明の概要】

【発明が解決しようとする課題】

【0003】

従って、LfDという状況において信頼性のある姿勢推定を可能にするアプローチが望まれている。

【課題を解決するための手段】

【 0 0 0 4 】

発明の開示

種々の実施形態によれば、ロボットを制御するための方法が提供され、この方法は、デモンストレーションを実施するステップであって、各デモンストレーションは物体の処理を実際に行うステップと、デモンストレーションごとに、物体に対するロボットの視点から、少なくとも1つの記述子画像を供給するステップであって、この記述子画像によって、物体上のロケーションに関する特徴点が指定される、ステップと、指定された特徴点から、複数の特徴点から成る集合を選択するステップであって、これらの特徴点は、記述子画像内において可視である選択された特徴点に報酬を与える目的関数の最適条件を探索することによって選択される、ステップと、デモンストレーションを用いてロボット制御モデルをトレーニングするステップであって、このロボット制御モデルは、入力された物体姿勢に依存して制御情報を出力するように構成されている、ステップと、物体を含む制御シーンに対しロボットを制御するステップであって、ロボットの視点から物体の記述子画像を求め、複数の特徴点から成る選択された集合を物体の記述子画像内に配置し、制御シーンに関して配置された特徴点のユークリッド座標を求め、求められたユークリッド座標から姿勢を推定し、推定された姿勢が入力としてロボット制御モデルに供給され、ロボット制御モデルを用いて、物体を処理するようにロボットを制御する、ことによって、ロボットを制御するステップと、を含む。

10

【 0 0 0 5 】

かくして、選択された特徴点が（例えば、記述子画像の大部分において）良好に可視であるように、ある1つの物体の姿勢推定に使用される特徴点の自動選択が実施される。従って、特徴点の選択を、以下の前提に基づくものであるとみなすことができる。即ち、この前提とは、デモンストレーション中に見られる特徴（例えば、デモンストレーションを記録するために撮影された画像）がこの意味で良好なものであり、それらは大抵の場合、個々の操作スキルがロボットによって実施されるべきであるシーンにおいて見られる、というものである。

20

【 0 0 0 6 】

ロボットの視点は、例えば、ロボットに取り付けられたカメラの視点であり、例えば、ロボットのエンドエフェクタの視点である。

【 0 0 0 7 】

このため上述の方法によって、ある1つの物体の姿勢（例えば、6D姿勢）の正確かつ確実な姿勢推定が可能となる。

30

【 0 0 0 8 】

以下、種々の実施例を挙げる。

【 0 0 0 9 】

実施例1は、上述のロボットを制御するための方法である。

【 0 0 1 0 】

実施例2は、実施例1に記載の方法において、目的関数が、選択された特徴点が記述子空間内において離間されていること、選択された特徴点に対応する物体上のロケーションがユークリッド空間内において離間されていること、及び、物体に対し選択された特徴点に関する検出誤差が小さいこと、のうちの1つ又は複数に、さらに報酬を与える、方法である。

40

【 0 0 1 1 】

これらの条件によって保証されることは、物体に関して特徴点を確実に検出することができること、及び、それらが検出されたときに物体の姿勢を確実に求めることができることである。

【 0 0 1 2 】

実施例3は、実施例1又は2に記載の方法において、物体に1つの平面をマッチングさせるステップと、特徴点がこの平面上の座標フレームを規定するように特徴点を選択するステップと、配置された特徴点の求められたユークリッド座標と、マッチングされた平面

50

の姿勢に関する情報とから、姿勢を推定するステップと、をさらに含む方法である。

【0013】

これによって、たとえば、物体が平坦であり、及び/又は、1つの空間方向において変化がほとんどなくても（即ち、「リアル」3D物体ではなくても）、信頼性のある姿勢推定が保証される。

【0014】

実施例4は、実施例3に記載の方法において、求められたユークリッド座標から姿勢を推定するステップが、配置された特徴点のユークリッド座標をマッチングされた超平面に投影するステップを含む、方法である。

【0015】

従って、たとえば、特徴点のユークリッド座標が平面上に位置しないように特徴点を検出されたとしても、（この平面内の）平面座標フレームを導出することができる。

【0016】

実施例5は、実施例3又は4に記載の方法において、平面が物体にマッチングされ、特徴点は、それらが平面上の1つの座標フレームを規定するように選択され、姿勢は、1つの空間方向における物体の変化が予め定められた閾値を下回っている場合には、配置された特徴点の求められたユークリッド座標と、平面の姿勢に関する情報とから推定される、方法である。

【0017】

かくして信頼性の高い姿勢決定スキームを、リアル3D物体と非リアル3D物体の双方のためにトレーニングすることができる。特に、（超）平面フィッティングは、必要なときにしか適用されないので、リアル3D物体に対する姿勢検出は、それでもやはり効率的である。

【0018】

実施例6は、実施例1乃至5のいずれか1つに記載の方法において、選択された特徴点のユークリッド座標から座標フレームの導出ルールを求めるステップを含み、求められたユークリッド座標から姿勢を推定するステップは、選択された特徴点に導出ルールを適用するステップを含み、この導出ルールは、ユークリッド座標におけるノイズから座標フレームの依存性の最小値を探索することによって求められる、方法である。

【0019】

これにより、たとえば、特徴点の検出誤差のケースであっても、確実な姿勢検出が保証される。

【0020】

実施例7は、実施例1乃至6のいずれか1つに記載の方法において、デモンストレーションを用いてロボット制御モデルをトレーニングするステップは、デモンストレーションごとに、複数の特徴点から成る選択された集合を、デモンストレーションにおける物体の記述子画像内に配置するステップと、デモンストレーションのために配置された特徴点のユークリッド座標を求めるステップと、デモンストレーションのために求められたユークリッド座標から姿勢を推定するステップと、を含む、方法である。

【0021】

このことが意味することは、選択された特徴点を、ロボット制御モデル自体のトレーニングに対し姿勢推定のために使用することができる、ということである。

【0022】

実施例8は、実施例1乃至7のいずれか1つに記載の方法において、物体の記述子画像を、高密度オブジェクトネットにより物体のカメラ画像から求めるステップを含む、方法である。

【0023】

ある1つの物体上の点を一意に識別するように、高密度オブジェクトネット（Dense Object Net, DON）をトレーニングすることができる。物体のコンフィギュレーションがそれぞれ異なっていたとしても、かかる点を確実に追跡することができ

10

20

30

40

50

る。

【0024】

実施例9は、実施例1乃至8のいずれか1つに記載の方法を実施するように構成されたロボットコントローラである。

【0025】

実施例10は、コンピュータにより実行されるときに、このコンピュータに実施例1乃至8のいずれか1つに記載の方法を実施させるための命令を含むコンピュータプログラムである。

【0026】

実施例11は、コンピュータにより実行されるときに、このコンピュータに実施例1乃至8のいずれか1つに記載の方法を実施させるための命令を含むコンピュータ可読媒体である。

10

【0027】

図面において、同様の参照符号は一般に、それぞれ異なる図面にわたり同様の部分を指す。図面は、必ずしも実寸の比率どおりではなく、その代わりに一般に本発明の原理を説明することに重点が置かれている。以下の記載においては、以下の図面を参照しながら種々の態様について説明する。

【図面の簡単な説明】

【0028】

【図1】ロボットを示す図である。

20

【図2】特徴点を選択し、1つの方向においてほとんど変化していないトレーニングデータセット中の物体について座標フレームを学習し、さらに配備中にその座標フレームを適用するための方法を説明する図である。

【図3】ロボットを制御するための方法を説明するフローチャートを示す図である。

【発明を実施するための形態】

【0029】

以下の詳細な説明は添付の図面を参照し、それらの図面には、本発明を実施可能な本開示の具体的な詳細及び態様が例示として示されている。他の態様を使用してもよく、本発明の範囲から逸脱することなく、構造的、論理的又は電氣的な変更を加えることができる。本開示のいくつかの態様を本開示の1つ又は複数の他の態様と組み合わせ、新しい態様を形成することができることから、本開示の種々の態様は、必ずしも相互に排他的なものではない。

30

【0030】

以下、種々の実施例についてさらに詳しく説明する。

【0031】

図1には、ロボット100が示されている。

【0032】

ロボット100は、ロボットアーム101を含み、例えば、ワークピース（あるいは、1つ又は複数の他の物体）を処理するための又は組み立てるための産業用ロボットアームを含む。ロボットアーム101は、マニピュレータ102、103、104、及び、これらのマニピュレータ102、103、104を支持する基台（又は支持体）105を含む。「マニピュレータ」という用語は、ロボットアーム101の可動部材を指し、この可動部材を作動させることによって、例えば、ある1つのタスクを実行するために、周囲との物理的インタラクション（相互作用）が可能となる。制御のために、ロボット100は、（ロボット）コントローラ106を含み、このコントローラ106は、制御プログラムに従って、周囲とのインタラクションを実施するように構成されている。マニピュレータ102、103、104のうち（支持体105から最も遠い）最後の部材104は、エンドエフェクタ104とも称され、溶接トーチ、把持器具、塗装設備など、1つ又は複数のツールを含み得る。

40

【0033】

50

(支持体105の比較的近くにある)他のマニピュレータ102、103は、エンドエフェクタ104と共に、終端にこのエンドエフェクタ104を備えたロボットアーム101が提供されるように、ポジショニング装置を成すことができる。ロボットアーム101は、(場合によっては、その端部に工具を有する)人間の腕と同様の機能を提供可能な機械式アームである。

【0034】

ロボットアーム101は、マニピュレータ102、103、104を互いにかつ支持体105と相互接続するジョイント部材107、108、109を含み得る。ジョイント部材107、108、109は、1つ又は複数のジョイントを有することができ、それらのジョイント各々によって、関連づけられたマニピュレータに対し互いに相対的に回転可能な運動(即ち、回転運動)及び/又は並進運動(即ち、変位)をもたらすことができる。マニピュレータ102、103、104の運動を、コントローラ106により制御されるアクチュエータを用いて開始することができる。

10

【0035】

「アクチュエータ」という用語は、駆動されることに応答して、メカニズム又はプロセスに作用を及ぼすように適応化された構成要素であると解することができる。アクチュエータは、コントローラ106によって発せられた命令(いわゆるアクティベーション)を実施して、機械的運動を生じさせることができる。アクチュエータを、例えば電気機械変換器を、駆動に応答して電気エネルギーを機械エネルギーに変換するように、構成することができる。

20

【0036】

「コントローラ」という用語は、任意のタイプのロジック実装エンティティのことであると解することができる。このロジック実装エンティティは、例えば、記憶媒体に格納されたソフトウェア、ファームウェア又はこれらの組合せを実行可能な回路及び/又はプロセッサを含み得るものであり、例えば、本実施例においては、アクチュエータに命令を発することができる。例えば、プログラムコード(例えば、ソフトウェア)によって、システム、本実施例においてはロボットの動作を制御するように、コントローラを構成することができる。

【0037】

本実施例によれば、コントローラ106は、1つ又は複数のプロセッサ110と、コード及びデータを格納するメモリ111とを含み、このコード及びデータに基づき、プロセッサ110はロボットアーム101を制御する。種々の実施形態によれば、コントローラ106は、メモリ111に格納された統計モデル112に基づき、ロボットアーム101を制御する。

30

【0038】

ロボット100は、デモンストレーションに基づく学習(LfD)というアプローチを活用して、ある1つのタスクの実行又は人間のパートナーとの協働を学習することができる。人間のデモンストレーションを、ロボットのためのタスクの公称プランを表現する確率モデル(統計モデルとも称される)によって、符号化することができる。次いで、コントローラ106は、ロボット軌跡モデルとも称される統計モデルを使用し、場合によっては、人間のパートナーと周囲の双方の状態の関数として、所望のロボット運動を生成することができる。

40

【0039】

LfDの基本的な着想は、GMMなどのような指定されたスキルモデルを一握りのデモンストレーションにフィットさせる、ということである。M個のデモンストレーションが行われるものとし、それらのデモンストレーション各々は、 $N = \sum_{t=1}^m T_m$ 個のすべての観測のデータセット

【数1】

$$\xi = \{\xi_t\}_{t=1}^N$$

50

について、 T 個のデータ点を含み、ここで、

【数 2】

$$\xi_t \in \mathbb{R}^d$$

である。また、（着目対象物体の局所座標系又は局所座標フレームといったタスクパラメータによって定められた）それぞれ異なる P 個の座標系の視点から、複数の同様のデモンストレーションが記録される、ということも前提とする。かかるデータを取得するための 1 つの一般的な方法は、デモンストレーションを静的なグローバルフレームから

【数 3】

$$\xi_t^{(p)} = \mathbf{A}^{(p)-1}(\xi_t - \mathbf{b}^{(p)})$$

によるフレーム p に変換することである。ここで、

【数 4】

$$\{(\mathbf{b}^{(p)}, \mathbf{A}^{(p)})\}_{p=1}^P$$

は、ワールド（即ち、グローバル）フレームに対する（局所的な）フレーム p の並進及び回転である。次いで、 TP -GMM が、モデルパラメータ

【数 5】

$$\{\pi_k, \{\boldsymbol{\mu}_k^{(p)}, \boldsymbol{\Sigma}_k^{(p)}\}_{p=1}^P\}_{k=1}^K$$

によって記述され、ここで、 K は、混合モデルにおけるガウス成分の個数を表し、 π_k は、各成分の事前確率であり、さらに

【数 6】

$$\{\boldsymbol{\mu}_k^{(p)}, \boldsymbol{\Sigma}_k^{(p)}\}_{p=1}^P$$

は、フレーム p 内の k 番目のガウス成分のパラメータである。

【0040】

標準的な GMM とは異なり、上述の混合モデルを、フレームごとに独立して学習させることができない。実際、混合係数 π_k は、すべてのフレームによって共有され、フレーム p 内の k 番目の成分は、グローバルフレーム内の対応する k 番目の成分にマッピングされなければならない。期待値最大化 (EM) は、かかるモデルを学習するために十分に確立された方法である。

【0041】

一旦学習された場合には、学習されたスキルについて軌跡を再現するために、実行中に、 TP -GMM を使用することができる。

【0042】

隠れセミマルコフモデル (HSMM) は、基礎を成す確率過程の時間情報を埋め込むことによって、標準隠れマルコフモデル (HMM) を拡張する。つまり、HMM の場合には基礎を成す隠れ過程がマルコフであるとみなされ、即ち、次の状態への遷移の確率は現在の状態のみに依存するのに対し、HSMM の場合には状態過程がセミマルコフであるとみなされる。このことが意味することは、次の状態への遷移が、現在の状態にも、その状態に入ってから経過した時間にも依存する、ということである。それらのパラメータを、 TP -GMM と組み合わせて、デモンストレーションの時空間特徴を学習するためのロボットスキル符号化に適用することができ、その結果、タスクパラメータ化された HSMM (TP -HSMM) モデルが得られる。

10

20

30

40

50

【 0 0 4 3 】

かかるロボット制御モデルを適用するためには、処理されるべき物体 1 1 3 の姿勢の推定が必要とされる。このことを、TP - GMM のパラメータを適正にセットできるように、処理されるべき物体 1 1 3 の初期姿勢に関係するものとして行うことができる。ただし、観測のために姿勢を推定しなければならない場合もある。例えば、デモンストレーションは、複数の画像から成るシーケンスを撮影することによって記録され、観測は、画像から抽出された処理される物体（又はロボットの一部分）の姿勢に対応する（又はこれを含む）。種々の実施形態によれば、選択された特徴点における姿勢を推定するアプローチが提供され、このアプローチによって、配備中に、例えば、RGBD（RGB + 深度）画像がある場合には、ある 1 つの制御シナリオ（シーン）において、ある 1 つの物体の姿勢（例えば、6 D 姿勢）を最初に推定することが可能となる。次いで、この姿勢が TP - HSMM に供給されて、この制御シナリオのための所望の軌跡が計算される。TP - HSMM のトレーニングの前に、これらのキー特徴点の選択が実施される。この選択がなされた場合には、次いで、デモンストレーションごとに、6 D 物体姿勢を導出することもできる。TP - HSMM のトレーニング中に、物体フレーム内のタスクパラメータを計算するためにこれらの 6 D 姿勢が使用される。

10

【 0 0 4 4 】

従って、種々の実施形態によれば、デモンストレーションに基づく学習が、画像を記述子画像にマッピングすることに基づく姿勢推定アプローチと組み合わせられる。かかるマッピングを実施するようにトレーニング可能なモデルの一例は、高密度オブジェクトネット

20

【 0 0 4 5 】

高密度オブジェクトネットにより、物体固有の特徴を追跡する手段が提供される。物体固有又は物体クラス固有のいずれかであるように、これをトレーニングすることができる。従って、これにより、単一のニューラルネットワークを用いて、幅広いクラスの物体にわたり物体クラス固有の特徴を検出する手法が提供される。

【 0 0 4 6 】

高密度オブジェクトネット（Dense Object Net, DON）は、

【 数 7 】

$$h \times w \times D, D \in \mathbb{N}^+$$

30

記述子空間における

【 数 8 】

$$h \times w \times 3$$

RGB 画像を表す。

【 0 0 4 7 】

DON をトレーニングするために、ピクセル単位の対照損失に基づく損失関数を、全結合ネットワークにおいて最小化することができる。ピクセル単位の対照損失は、2 つの所与の画像から損失を生成する技術である。この損失に従い、2 つの画像内のセマンティック的に類似するピクセルのすべてのペアは、記述子空間内の類似点ペアにマッピングされなければならない。しかも、セマンティック的に類似していないピクセルのすべてのペアは、記述子空間の異なる領域にマッピングされなければならない。ピクセル単位の対照損失を適用することによって、物体（クラス）固有の特徴を検出可能なネットワークがもたらされる。種々の実施形態によれば、この能力が活用される。特に、（トレーニングデータにおいて）まだ見たことのない画像（これは特定のコンフィギュレーション、例えば、ロボットアーム 1 0 1 を制御すべきときの出発点となる初期コンフィギュレーションを示す）において、物体固有の特徴点が発見される（記述子空間の 1 つの要素である場合には

40

50

、この要素が1つの特徴点であるとされる)。所与の(ターゲット)記述子にできるかぎり近い記述子により、記述子空間において表される画像のピクセルを見出すことによって、所与の記述子(従って、所与の特徴、例えば、物体113)のユークリッド座標が見出される。

【数9】

$$(u^*, v^*)$$

によって表されるこの「最良」のピクセルを見つけるために、いくつかのアプローチを使用することができる。1つのアプローチは、ターゲット記述子までの距離を最小化する記述子に対応するピクセルを選ぶことである。即ち、

【数10】

$$(u^*, v^*) = \operatorname{argmax}_{(u,v) \in h \times w} \|d^* - I(u, v)\|^2 \quad (1)$$

であり、ここで、

【数11】

$$d^*$$

は、特徴記述子を表し、Iは、記述子画像を表す。

【0048】

上述のようにこれは一例にすぎず、所与の記述子を用いて(新しい)画像においてマッチを見出すために、他のアプローチを使用することができる。例えば、より高いロバスト性のために、指数関数によって重み付けされた二乗誤差の期待値が取得される実施形態によれば、空間期待値と称される修正バージョンが使用される。

【0049】

画像座標(ピクセル座標)からワールド座標(例えば、ロボットの基準フレーム座標)への変換は、ロボット及びコンピュータビジョンにおける標準技術である。見出された「最良の」ピクセルへのこの変換を適用した結果、記述子空間に生じた物体固有の特徴点についてユークリッド座標を見出すためのルーチンが得られる。これらのユークリッド座標は、特徴点のユークリッド座標として表される。これらは、シーン(即ち、コンフィギュレーション)に依存する。

【0050】

ここで前提とすることは、高密度オブジェクトネットの出力を利用可能であり、(例えば、ロボットアーム101を制御すべき対象である初期コンフィギュレーションを示す)現在の画像の記述子画像内の各記述子について、それがターゲット物体(例えば、物体113)上の点に対応するのかが既知であり、もし対応する場合には、それが属する物体上の点のユークリッド座標も既知である、ということである。以下においては、デモンストレーションによる学習に適用可能な姿勢推定を見出すことができるようにするアプローチについて説明する。姿勢推定は、物体が動かされると姿勢も同様に变换されるという意味において、物体の動きに関して不変でなければならない。ここで留意されたいことは、このためにRGB画像が使用されない、ということである。特に、トレーニング(又は学習)データセット中のすべてのデータ点

【数12】

$$\{I_i, \text{mask}_i, \text{transformation}_i\}, i=1 \dots N$$

10

20

30

40

50

は、D次元の記述子画像Iと、画像マスク（これはある1つの点が物体に属するの否かの情報を含む）と、物体上に位置しているある1つの点を表す記述子ごとにユークリッド座標を計算するための手法（“transformation”として表される）とから成るものとする事ができる。ここで留意されたいことは、このトレーニングデータセットは、特徴点を選択するためのトレーニングデータセットである、ということである。DONについて前提とすることは、（教師あり又は教師なしの方式で）別の適切なトレーニングデータセットを用いて物体に対しすでにトレーニング済みであるか又は事前にトレーニングされる、ということである。

【0051】

種々の実施形態によれば、デモンストレーションによる学習に適用可能な姿勢推定を見出すことができるようにするアプローチは、（特徴点と称される）K個の記述子を見出すための最適化に基づく方法と、（任意選択肢として）部分集合

10

【数13】

$$X \subseteq \mathbb{R}^3$$

のための座標フレームの最適化に基づくパラメータ化と、（任意選択肢として）適切な超平面への直交投影とを含む。

【0052】

記述子の選択

20

これには、この場合には、座標フレームのパラメータ化に使用可能な特徴点を見出すことが含まれる。

【0053】

種々の実施形態によれば、特徴点の集合が種々の条件を満たすように選択される。以下においては4つの条件を挙げるが、ここで留意されたいことは、それらのすべてを考慮しなくてもよく、また、他の条件に対し1つの条件を重要視するために重みを定めることができる、ということである。

【0054】

第1の条件は、選択された特徴点は小さな検出誤差で検出されなければならない、ということである（このことは、それらが符号化する物体の特徴の可視性を含む）。この判定基準を定量化するために、生じ得る記述子ごとに、学習データセット内のすべてのデータ点にわたりマッチング誤差が計算され、この誤差が、未知のデータ点に対するマッチング誤差の期待値として計算される（検出誤差は正規化されるので、平均誤差の代わりに累積誤差を使用することができる）。

30

【0055】

マッチング誤差を計算することができる2つのドメインが存在し、これは記述子ドメイン内（記述子と記述子画像との間の距離）と、ユークリッドドメイン内（特徴点のユークリッド座標における変化）とである。記述子ドメイン内においてマッチング誤差を計算する利点は、どのピクセルが（セマンティック的に）他の画像内のどのピクセルに属しているのにかに関するいかなるグラウンドトゥルス情報も必要としない、ということである。かかるデータセット（高密度オブジェクトネットのトレーニングデータセット）の可用性にもかかわらず、これを使用するには一般性を損なうことになってしまう。しかしながら、記述子空間内においては非常に良好に追跡できるが、ユークリッド空間内においては不十分にしか追跡できない特徴点があるといったような物体の存在が認められるので、ユークリッドドメインを使用することもできる。

40

【0056】

特徴点の集合を選択するために、特徴点の集合を満たさなければならない第2の条件は、記述子空間における空間的变化である。この条件によって保証されることは、選択された特徴点が、ユークリッド空間における物体について高密度オブジェクトネットの視点で種々の特徴を記述する、ということである。記述子空間における空間分離条件と同様に、

50

第3の条件を、ユークリッド空間における空間分離とすることができる。特徴点がユークリッド空間内において互いに近すぎると（即ち、それらのユークリッド座標がある1つの物体について類似しすぎると）、マッチング誤差の影響が増加する。ユークリッド空間条件におけるマッチング誤差とは異なり、空間分離条件によっても、種々の画像相互の関係に関する知識を必要としないことから、一般性が損なわれてしまうことはない。

【0057】

第4の条件は、学習データセットのすべてのデータ点（又はできるかぎり多くのデータ点）に含まれる特徴が取得されなければならない、ということである。トレーニングデータセット中の記述子画像各々が、カメラ画像に基づきトレーニングされたDONによって生成される、ということを前提とする場合には、このことが意味することは、すべての記述子がトレーニングデータセットのすべての（又はできるかぎり多くの）画像において可視でなければならない、ということである。配備中にも存在するであろうコンフィギュレーションがデモンストレーションによって表された、ということを前提とすることができるので、これらの特徴点を有する物体部分が、この場合には配備中に取得された画像内においても可視となる可能性が高い、と予期することができる。

10

【0058】

式(2)によって、上述の4つの条件を考慮する特徴点を選択するための最適化問題が与えられる。

【数14】

$$\min_{\substack{p_i \in \mathbb{R}^D \\ i \in 1, \dots, K}} \left[-\alpha_1 (\text{NO}_{(\mathbb{R}^D)^K} \beta_1)(p_1, \dots, p_D) - \alpha_2 (\text{NO}_{(\mathbb{R}^D)^K} \beta_2)(p_1, \dots, p_D) \right. \\ \left. + \alpha_3 (\text{NO}_{(\mathbb{R}^D)^K} \beta_3)(p_1, \dots, p_D) + \alpha_4 (\text{NO}_{(\mathbb{R}^D)^K} \beta_4)(p_1, \dots, p_D) \right] \quad (2)$$

20

ただし、

$$\beta_1(p_1, \dots, p_K) = \sum_{\substack{j, k=1 \\ i \neq j}}^K \|p_j - p_k\|$$

30

$$\beta_2(p_1, \dots, p_K)$$

$$= \sum_{\substack{j, k=1 \\ i \neq j}}^K \frac{1}{N} \sum_{l=1}^N \|\text{transformation}_l(p_j) - \text{transformation}_l(p_k)\|$$

40

$$\beta_3(p_1, \dots, p_K) = \frac{1}{4} \sum_{j=1}^K e(p_j)$$

$$\beta_4(p_1, \dots, p_K) = \sum_{i=k}^K \sum_{l=1}^N \text{detected}_l(p_k)$$

50

【 0 0 5 9 】

この場合、 K は、選択された記述子の個数であり、

【 数 1 5 】

$$\alpha_1, \alpha_2, \alpha_3, \alpha_4$$

は、条件の結合パラメータとして用いられる正の実数である。正規化演算子

【 数 1 6 】

$$NO_I$$

10

は、関数

【 数 1 7 】

$$f : I \subseteq X \rightarrow \mathbb{R}^+$$

を関数

【 数 1 8 】

$$\tilde{f} : I \subseteq X \rightarrow [0, 1]$$

20

にマッピングする演算子であり、 I に制限された種々の

【 数 1 9 】

$$\tilde{f}$$

を、費用関数を規定するために追加することができるようにする（すべての

【 数 2 0 】

$$\tilde{f}$$

30

は、これに対し同様の影響を及ぼす）。

【 0 0 6 0 】

検出された関数 $l(p_k)$ は、画像 l において特徴点 p_k が検出された場合には、0 に等しく、そうでなければ、である指示関数である。

【 0 0 6 1 】

任意の有限集合 I について、正規化演算子を

【 数 2 1 】

$$NO_I f(x) = \frac{f(x) - \min_{\tilde{x} \in I} f(\tilde{x})}{\max_{\tilde{x} \in I} f(\tilde{x}) - \min_{\tilde{x} \in I} f(\tilde{x})}$$

40

(3)

によって規定することができる。

【 0 0 6 2 】

式 (3) において、関数 $e(\cdot)$ は、ユークリッド空間における対応するマッチング誤差に記述子をマッピングする。高密度オブジェクトネットのためのトレーニングデータセットを利用できる場合には（又は、物体のポジション及び姿勢が静的である複数のシーンから成る他の任意のデータセットを利用できる場合には）、 $e(x)$ のために以下の推定プロセスを使用することができる。即ち、シーン i ごとに（ランダムに） k 個の画像を選

50

択し、各画像における x についてユークリッド座標 p_{ij} を推論する。1つのシーン内において平均値

【数 2 2】

$$\bar{p}_i = \frac{1}{j} \sum_{l=1}^k p_{ij}$$

を計算する。従って、

【数 2 3】

$$e(x) \approx \sum_{i,j} p_{ij} - \bar{p}_i$$

10

となる。上述のように、累積マッチング誤差は、十分である。

【0 0 6 3】

座標フレームのパラメータ化

(例えば、上述のように学習された) 複数の特徴点から成る集合がある場合には、学習された特徴点に依存する空間 X に対する座標系 (又は座標フレーム) が導出される (学習される)。ここで、 D は、学習された特徴点のユークリッド座標である列を有する行列を表し、 $B(D)$ は、望ましい基底行列を表す (即ち、その列は、 X の基底である)。以下

20

【数 2 4】

$$X = \mathbb{R}^3$$

であることを前提とする。それ以外の場合には、学習された記述子が X に投影され、以下の演算が、(適切な n の場合に

【数 2 5】

$$\mathbb{R}^n$$

30

として識別される) X に対し、すべてのステップに必要な変更を加えて実行される。さらに前提とすることは、適切な次元を有する行列 B について

【数 2 6】

$$B(D) = DB$$

である、ということである。学習された座標系は、特徴点のユークリッド座標におけるノイズが基底ベクトルに重大な影響を及ぼしてはならないという意味において、信頼性が高くなければならない。数学的に表現すれば、このことを

40

【数 2 7】

$$\|B(D) - B(\tilde{D})\| \tag{4}$$

の最小化として定式化することができる。ただし、

【数 2 8】

$$\tilde{D}$$

50

は測定された記述子行列を表す。このことは、ノイズ項 E を伴う

【数 2 9】

$$\tilde{D} = D + E$$

であることを意味する。このノイズ項を、以下のように記述することができる。即ち、

【数 3 0】

$$E = (e_1 \ e_2 \ \dots \ e_K) \begin{pmatrix} w_1 & & & \\ & w_2 & & \\ & & \ddots & \\ & & & w_K \end{pmatrix} \quad (5) \quad 10$$

いくつかについて $w_i \in \mathbb{R} \quad i = 1, 2, \dots, K, e_i \in \mathbb{R}^3 \quad i = 1, 2, \dots, K$

ただし $\|(e_1 \ e_2 \ \dots \ e_K)\| \leq 1$

よって、

20

$$\begin{aligned} & \|B(D) - B(\tilde{D})\| \\ &= \left\| (e_1 \ e_2 \ \dots \ e_K) \begin{pmatrix} w_1 & & & \\ & w_2 & & \\ & & \ddots & \\ & & & w_K \end{pmatrix} B \right\| \\ &\leq \left\| \begin{pmatrix} w_1 & & & \\ & w_2 & & \\ & & \ddots & \\ & & & w_K \end{pmatrix} B \right\| \\ &= \|WB\| \end{aligned} \quad 30$$

【0 0 6 4】

ここで留意されたいことは、

【数 3 1】

$$\|WB\|$$

40

は、 WB の最大特異値である、ということである。

【0 0 6 5】

上述の累積マッチング誤差を使用して、重み

【数 3 2】

$$w_i, \quad i = 1, 2, \dots, K$$

を推定することができる。

【0 0 6 6】

50

簡単にするため、(4)の最適化を、以下の座標系を規定するすべてのBにわたって行うことができ、即ち、この座標系の軸は、2点間のベクトル、又は、1つの中心点から他の点を指すベクトルのいずれかである。直交基底を取得するために、直交化プロセスをB(D)に適用することができる。

【0067】

特徴点のユークリッド座標の平均値を、原点として使用することができる。

【0068】

支持超平面

トレーニング中の物体が「リアル」3Dである場合には、特徴点の選択及び座標フレームの学習のための上述のアプローチを、その物体に適用可能である。しかしながら、多くの工業用途においては、物体がわずかな変化しかもたらさない方向が存在するという意味においては、考察される物体はリアル3Dではない。かかるケースにおいては、上述のアプローチによって、少なくとも1つの信頼できない座標軸が生じることになる。従って、かかる物体(例えば、平坦な物体)のケースにおいては、この方向は、選択された適切な超平面への直交投影を用いることで除去される。

【0069】

図2には、特徴点を選択し、1つの方向においてほとんど変化していないトレーニングデータセット202中の物体201について座標フレーム204を学習し、さらに配備中にその座標フレームを適用するための方法が説明されている。

【0070】

この方法は、超平面203を選択(学習)するステップを含む。超平面は、物体の(マッチングされた)記述子に対応する点(物体の点群と称される)が位置している表面を近似しなければならない。表面がほぼ平坦である場合には、標準的な最小二乗法又は特異値分解のアプローチを使用することができる。しかしながら、常にこのことが当てはまるわけではない。例えば、古典的なPCB(プリント回路基板)の物体は、変化が少ない1つの方向(基板に対して垂直な方向)を有し、最小二乗法又は特異値分解といったアプローチは、電子部品のサイズ及び高密度オブジェクトネットのマッチング誤差に起因して、うまくいかない可能性がある。

【0071】

この問題を克服するために、種々の実施形態によれば、支持平面203を計算するために修正(又は適応化)されたTheil-Sen推定器が使用される。

【0072】

すべての平面は、2つの方向ベクトルと1つの支持点とによって一意に規定される。しかも、超平面を、関数

【数33】

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}$$

のグラフとして解釈することができる。

【0073】

fの傾きを、

【数34】

$$\mathbb{R}^2$$

の2つの標準ユークリッド軸上において推定することができる。

【0074】

fの画像内に生じた点yを、

【数35】

10

20

30

40

50

$$y = c_1 + c_2 = \alpha_1 \cdot m_1 \cdot e_1 + \alpha_2 \cdot m_2 \cdot e_2$$

として記述することができ、ただし、 $e_1 ; e_2$ は、標準ユークリッド軸の単位ベクトルであり、 $m_1 ; m_2$ は、 e_1 / e_2 の方向における f の傾きであり、 α_1, α_2 は、
 【数 3 6】

$$y = f(\alpha_1 e_1 + \alpha_2 e_2)$$

であるようないくつかのスカラである。

10

【0075】

y 及び c_2 / c_1 から、比 m_1 / m_2 を計算することができる。超平面を学習する（即ち、フィッティングさせる）ためには、双方の傾きを同時に求めなければならない。従って、1つの実施形態によれば、 m_1 / m_2 の決定と c_1, c_2 の推定とを交互に行う反復アルゴリズムが使用される。以下のアルゴリズム 1 は、点群 P （ここで、各点は、3つの空間座標を有する）において動作する 1つの例を示す。

【数 3 7】

入力 : $P = \{p_1, p_2, p_3\} = \{\{p_{1,i}\}, \{p_{2,i}\}, \{p_{3,i}\}\}$

出力 : n 及び b

20

/* 初期推定: */

1 $m_1, m_2 = 1;$

2 P における x_1, x_2 及び y の選択に対して **実行**

3 m_1, m_2 と先行のそれらとの偏差が大きすぎ、
 かつ、最大反復回数に達していない場合 **実行**

4 $c_1 = y - m_2 \cdot x_2;$

5 $c_2 = y - m_1 \cdot x_1;$

6 m_1, m_2 と、データ (x_1, c_1) 及び (x_2, c_2)
 に対する両方の切片とについての新たな推定を
 計算するために Theil-Sen を使用

30

7 フィッティング誤差を計算し、
 それを超平面パラメータと共に記憶

8 最小フィッティング誤差を有するフィットを選択

アルゴリズム 1 : 超平面フィッティングのために適応化された Theil-Sen 推定器

40

【0076】

トレーニングデータセットのデータポイントのいずれか 1つ（即ち、対応する点群を生成するための変換が付随する記述子画像のいずれか）から、超平面 2 0 3 を推定することができる。

【0077】

超平面 2 0 3 がある場合には、特徴点のユークリッド座標をその超平面に投影することができる。上述の座標フレームをパラメータ化するための方法を適用するために、（

【数 3 8】

50

\mathbb{R}^2

によって識別可能な)超平面203は、1つの座標系を備えている。1つの考えられる手法は、原点が記述子のすべてのユークリッド座標の平均に位置し、変換されたz軸が超平面の法線ベクトルと整列されるように、ワールド座標系を移動させることである。変換されたこのワールド座標系によって、超平面上に1つの平面座標系が誘導され(超平面上へのx軸及びy軸の投影)、この平面座標系においてパラメータ化を実施することができる。さらに、計算された変換によって、パラメータ化された平面的な基底を

【数39】

10

 \mathbb{R}^3

の基底に拡張することができる(これを、z軸/法線ベクトルを中心とした一般化された回転と解釈することができる)。このアプローチは、平面的な基底を見出す他の方法のためにも有効である。特に、点群の比較に基づく任意の方法を使用することができる。

【0078】

こうしてトレーニング中に、複数の(ユークリッド空間の)点から成る1つの点群が物体201から生成され、これは、物体201上のロケーションに相応する記述子に対応する。記述子画像ごとに、記述子画像に対する個々の変換によって、記述子のユークリッド座標が求められる。

20

【0079】

物体がリアル3Dでなければ、物体の支持超平面203は上述のようにして求められる。超平面(及びこの超平面への投影)が必要とされるのかを判定するための1つの考えられる手法は、1つの行列として解釈される点群の特異値によってもたらされる。この判定を自動的に行うことができる。

【0080】

点群の各点について、そのユークリッド座標と超平面との間のユークリッド距離が計算される。超平面まで最短のユークリッド距離を有する点を選択し、その他の点が除去されて、点群から外れ値が除かれる。

30

【0081】

次いで、複数の特徴点から成る選択された集合のユークリッド座標に依存する平面座標フレーム204(即ち、超平面に関する座標フレーム)を求めるために、上述のアプローチが使用される。物体がリアル3Dでない場合、複数の特徴点から成る選択された集合のこれらのユークリッド座標は、超平面に投影されたユークリッド座標である。従って、特徴点のユークリッド座標に関する条件を考慮して特徴点を選択され、かつ、物体がリアル3Dでなければ、この条件は、超平面に投影されたユークリッド座標に適用される。

【0082】

配備中(即ち、オンライン推定中)、観測209に基づき点群(例えば、トレーニングではまだ見たことがなかった制御シナリオにおける物体113の現在の画像)が得られた場合、物体がリアル3Dでなければ、この点群に対する超平面205及び選択された特徴点に対応する3D点のポジション206が、(選択された特徴点と現在の画像の記述子とのマッチングによって)求められる。このことが意味することは、(物体113の画像についてDONにより生成された)記述子画像と、カメラの内因性及び外因性のパラメータとがある場合には、点群の、特に(可能なマッチング誤差まで)選択された特徴点の、ユークリッド座標を求めることができる、ということである。上述のアプローチを用いてトレーニング中と同様に、超平面204を求める(即ち、点群にマッチングさせる)ことができる。

40

【0083】

選択された特徴点(即ち、平面フレーム204を規定する点)のユークリッド座標が、

50

207において超平面に投影されて、これらの投影によって生じた平面座標フレームが定められる。

【0084】

総合(3D)座標フレーム208は、上述のように、点群のユークリッド座標の平均値と、平面座標フレームと、変換されたワールド座標系との組合せである。投影が不要なケースにおいては(即ち、リアル3D物体のケースにおいては)、総合座標フレーム208は、(特徴点と現在の画像の記述子画像との間の特徴マッチングにより見出された)選択された特徴点のポジションによって定められるものにすぎない。

【0085】

式(2)による特徴点の最適化に基づく選択を、例えば、以下のヒューリスティックアプローチによって解くことができる。1つのデータセットがある場合には、1つのデータ点を選択され、物体上に位置する記述子の集合が推測される。この集合において所与の閾値よりも互いに近いすべての記述子が、元の記述子の座標単位の平均によって置き換えられる/近似される。第2のステップにおいて、トレーニングデータセットにおけるすべてのデータ点にわたり、圧縮された集合におけるすべての記述子のマッチング誤差が計算される。1つのデータ点について所与の閾値よりも大きいマッチング誤差を有する記述子が除去される。この集合において(記述子ごとに個別に)最も小さい累積マッチング誤差を有する記述子の部分集合が選択される。この部分集合におけるK個の記述子の考えられる組合せごとに、式(2)の最適化問題において使用された費用が評価される。ここで留意されたいことは、この費用は有限の多数の点に基づき評価されるので、式(3)を適用可能である、ということである。(記述子空間内又はユークリッド空間内において)所与の閾値よりも小さいユークリッド距離を有する特徴点のペアが存在する、というすべての特徴点ペアが拒否される。最小の総費用を有する記述子ペアが使用される。

【0086】

要約すると、種々の実施形態によれば、図3において説明されているような方法が提供される。

【0087】

図3には、ロボットを制御するための方法を説明するフローチャート300が示されている。

【0088】

ステップ301において、デモンストレーションが(例えば、人間ユーザによって)実施され、その際に各デモンストレーションによって物体の処理が実際に行われる。

【0089】

ステップ302において、デモンストレーションごとに、物体に対するロボットの視点から、少なくとも1つの記述子画像が供給され(又は取得され)、その際にこの記述子画像によって、物体上のロケーションに関する特徴点が指定される。記述子画像を、例えば、デモンストレーション中に、ロボットに取り付けられたカメラによって撮影された画像から、DONによって生成することができる。

【0090】

ステップ303において、指定された特徴点から、複数の特徴点から成る集合が(自動的に)選択され、その際にそれらの特徴点は、記述子画像内において可視である選択された特徴点に報酬を与える目的関数の最適条件を探索することによって選択される。

【0091】

ステップ304において、デモンストレーションを用いてロボット制御モデル(即ち、TP-HSMM)がトレーニングされ、この場合、ロボット制御モデルは、入力された物体姿勢に依存して制御情報を出力するように構成されている。

【0092】

ステップ305において、物体を含む制御シーンに対しロボットが、以下のステップにより制御される。即ち、

- ・ステップ306において、ロボットの視点から物体の記述子画像を求め、

10

20

30

40

50

- ・ステップ 307 において、複数の特徴点から成る選択された集合を物体の記述子画像内に配置し、
- ・ステップ 308 において、この制御シーンに関して配置された特徴点のユークリッド座標を求め、
- ・ステップ 309 において、求められたユークリッド座標から姿勢を推定し、
- ・ステップ 310 において、推定された姿勢が入力としてロボット制御モデルに供給され、ロボット制御モデルを用いて、物体を処理するようにロボットを制御する。

【0093】

図 3 のアプローチを使用して、例えば、ロボット、車両、家電製品、電動工具、製造機械、パーソナルアシスタント又はアクセスコントロールシステムといったコンピュータ制御機械などの物理システムを制御するための制御信号を計算することができる。種々の実施形態によれば、物理システムを制御するためのポリシーを学習することができ、その後、それに応じて物理システムを動作させることができる。

10

【0094】

種々の実施形態は、ビデオ、レーダ、LiDAR、超音波、サーマルイメージング、モーション、ソナーなど、種々の視覚センサ（カメラ）からの画像データ（即ち、デジタル画像）を、例えば、記述子画像の基礎として、受信し使用することができる。

【0095】

1 つの実施形態によれば、この方法はコンピュータによって実装されている。

【0096】

これまで特定の実施形態について例示し説明してきたが、当業者には自明であるとおり、本発明の範囲を逸脱することなく、ここで示し説明してきた特定の実施形態の代わりに、種々の選択的実施形態及び/又は等価の実施形態を用いることができる。本願は、ここで論じた特定の実施形態のどのような適応化又は変形であってもカバーすることを意図している。従って、本発明は、特許請求の範囲及びその等価物によってのみ限定されることを意図している。

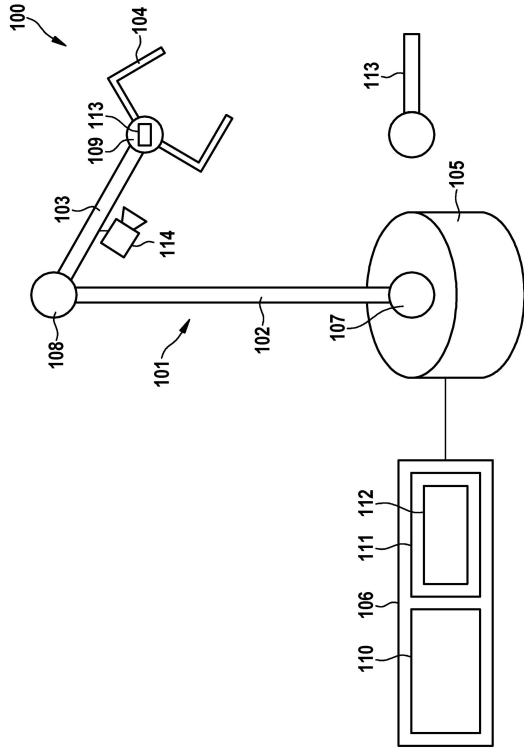
20

30

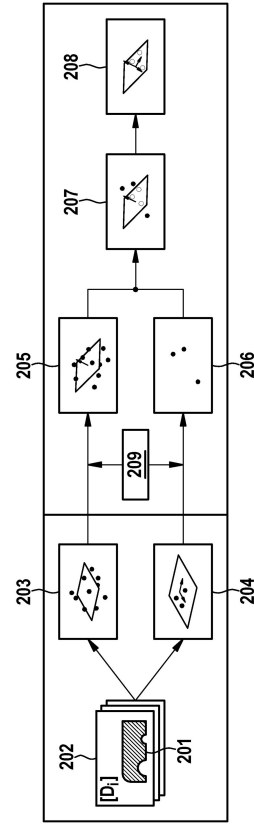
40

50

【 図 面 】
【 図 1 】



【 図 2 】



10

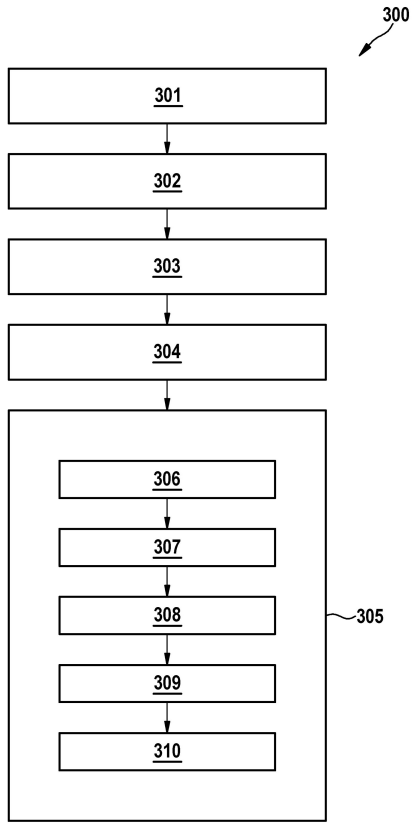
20

30

40

50

【 図 3 】



10

20

30

40

50

【 外国語明細書 】

2023055217000044.pdf

フロントページの続き

弁理士 永島 秀郎
(74)代理人 100162880
弁理士 上島 類
(72)発明者 アンドラーシュ ガーボル クブチク
ドイツ連邦共和国 ベープリングェン カールシュトラッセ 2 3
(72)発明者 メン グオ
中華人民共和国 베이ジン ハイディエン ドン シェン ユエン ゴン イー 1 2 ロウ 5 0 4
F ターム (参考) 3C707 AS06 AS11 AS13 BS10 HS23 KS04 KT01 KT05 KT11 LS15
LW12 LW15
5L096 AA02 AA06 AA09 BA05 BA18 CA02 DA02 EA15 EA16 EA31
EA39 FA32 GA32 GA51 HA11 JA03 KA04