

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7581320号
(P7581320)

(45)発行日 令和6年11月12日(2024.11.12)

(24)登録日 令和6年11月1日(2024.11.1)

(51)国際特許分類

| | | | | |
|---------|-----------------|-----|---------|-------------|
| G 0 5 D | 1/00 (2024.01) | F I | G 0 5 D | 1/00 |
| G 0 6 F | 3/01 (2006.01) | | G 0 6 F | 3/01 5 1 0 |
| G 0 6 F | 3/0481(2022.01) | | G 0 6 F | 3/0481 |
| G 0 6 T | 19/00 (2011.01) | | G 0 6 T | 19/00 6 0 0 |

請求項の数 20 (全30頁)

(21)出願番号 特願2022-503979(P2022-503979)
 (86)(22)出願日 令和2年6月5日(2020.6.5)
 (65)公表番号 特表2022-542241(P2022-542241
 A)
 (43)公表日 令和4年9月30日(2022.9.30)
 (86)国際出願番号 PCT/US2020/036428
 (87)国際公開番号 WO2021/015868
 (87)国際公開日 令和3年1月28日(2021.1.28)
 審査請求日 令和5年6月2日(2023.6.2)
 (31)優先権主張番号 62/877,792
 (32)優先日 令和1年7月23日(2019.7.23)
 (33)優先権主張国・地域又は機関
 米国(US)
 (31)優先権主張番号 62/877,791
 (32)優先日 令和1年7月23日(2019.7.23)

最終頁に続く

(73)特許権者 518102470
 トヨタ リサーチ インスティテュート,
 インコーポレイティド
 アメリカ合衆国, カリフォルニア 94
 022, ロスアルトス, エル カミノ リ
 アル 4440
 (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 100092624
 弁理士 鶴田 準一
 (74)代理人 100147555
 弁理士 伊藤 公一
 (74)代理人 100123593

最終頁に続く

(54)【発明の名称】 口ボット装置からの視覚的出力を補強するシステム及び方法

(57)【特許請求の範囲】**【請求項1】**

口ボット装置により生成されたデータを可視化する方法であって、
 第1の場所における現在の環境において、前記第1の場所とは異なる第2の場所における訓練環境内での口ボット装置の運転可能領域の学習に関連するニューラルネットワークに基づいて前記口ボット装置により識別された第1の運転可能領域を、ユーザ装置のディスプレイに表示することであって、前記運転可能領域は、前記訓練環境内で前記運転可能領域を走行する前記口ボット装置に関連するセンサデータおよび奥行データに基づいて学習される、ことと、

前記第1の運転可能領域内での前記口ボット装置の意図する経路であって、前記運転可能領域を学習する前記ニューラルネットワークに基づいて決定される前記口ボット装置の意図する経路を、前記ディスプレイに表示することと、

前記現在の環境における第2の運転可能領域を識別するための入力を前記ユーザ装置で受信することと、

前記第2の運転可能領域を示すメッセージを前記口ボット装置に送信し、前記メッセージの送信に応じて前記ニューラルネットワークが更新されることと、

を含む、方法。

【請求項2】

前記訓練環境の運転可能領域内で過去に走行した一つ以上の経路に基づいて前記意図する経路を判断するように前記口ボット装置を訓練することを更に含む、請求項1に記載の

方法。

【請求項 3】

前記意図する経路に重ね合わせて前記第1の運転可能領域を表示することを更に含む、請求項1に記載の方法。

【請求項 4】

前記入力は前記第2の運転可能領域に対応するピクセルにマークを付することを含む、請求項1に記載の方法。

【請求項 5】

前記ロボット装置は、前記第2の運転可能領域の受信に基づいて第3の運転可能領域を運転可能として判断する、請求項1に記載の方法。

10

【請求項 6】

現在の画像とキーフレーム画像との比較を表示することを更に含み、前記比較は前記現在の画像と前記キーフレーム画像との一致ピクセルを識別する可視化されたマークを有する、請求項1に記載の方法。

【請求項 7】

現在の画像とキーフレームのセットの中の各キーフレームとの間のマッチングの確信度を表示することを更に含む、請求項1に記載の方法。

【請求項 8】

前記意図する経路及び前記第1の運転可能領域を前記ロボット装置から離れた位置にて表示する、請求項1に記載の方法。

20

【請求項 9】

ロボット装置により生成されたデータを可視化する装置であって、
プロセッサと、

前記プロセッサに接続され命令を記憶するメモリであって、該命令は、前記プロセッサにより実行されると、前記装置に、

第1の場所における現在の環境において、前記第1の場所とは異なる第2の場所における訓練環境内でのロボット装置の運転可能領域の学習に関連するニューラルネットワークに基づいて前記ロボット装置により識別された第1の運転可能領域を表示させ、前記運転可能領域は、前記訓練環境内で前記運転可能領域を走行する前記ロボット装置に関連するセンサデータおよび奥行データに基づいて学習され、

30

前記第1の運転可能領域内での前記ロボット装置の意図する経路であって、前記運転可能領域を学習する前記ニューラルネットワークに基づいて決定される前記ロボット装置の意図する経路を表示させ、

前記現在の環境における第2の運転可能領域を識別するための入力を受信させ、

前記第2の運転可能領域を示すメッセージを前記ロボット装置に送信させ、前記メッセージの送信に応じて前記ニューラルネットワークが更新されるようにする、

ように動作可能である命令を記憶する、メモリと、
を具備する、装置。

【請求項 10】

前記命令の実行は更に、前記装置に、前記訓練環境の運転可能領域内で過去に走行した一つ以上の経路に基づいて前記意図する経路を判断するように前記ロボット装置を訓練させる、請求項9に記載の装置。

40

【請求項 11】

前記命令の実行は更に、前記装置に、前記意図する経路に重ね合わせて前記第1の運転可能領域を表示させる、請求項9に記載の装置。

【請求項 12】

前記入力は前記第2の運転可能領域に対応するピクセルにマークを付することを含む、請求項9に記載の装置。

【請求項 13】

前記ロボット装置は、前記第2の運転可能領域の受信に基づいて第3の運転可能領域を

50

運転可能として判断する、請求項 9 に記載の装置。

【請求項 14】

前記プロセッサは更に、現在の画像とキーフレーム画像との比較を表示するように構成され、前記比較は前記現在の画像と前記キーフレーム画像との一致ピクセルを識別する可視化されたマークを含む、請求項 9 に記載の装置。

【請求項 15】

前記プロセッサは更に、現在の画像とキーフレームのセットの中の各キーフレームとの間のマッチング確信度を表示するように構成されている、請求項 9 に記載の装置。

【請求項 16】

前記意図する経路及び前記第 1 の運転可能領域を前記ロボット装置から離れた位置にて表示する、請求項 9 に記載の装置。

10

【請求項 17】

ロボット装置により生成されたデータを可視化するためのプログラムコードを記録した非一時的なコンピュータ可読媒体であって、

前記プログラムコードはプロセッサにより実行され、

第 1 の場所における現在の環境において、前記第 1 の場所とは異なる第 2 の場所における訓練環境内でのロボット装置の運転可能領域の学習に関連するニューラルネットワークに基づいて前記ロボット装置により識別された第 1 の運転可能領域を、ユーザ装置のディスプレイに表示し、前記運転可能領域は、前記訓練環境内で前記運転可能領域を走行する前記ロボット装置に関連するセンサデータおよび奥行データに基づいて学習されるプログラムコードと、

20

前記第 1 の運転可能領域内での前記ロボット装置の意図する経路であって、前記運転可能領域を学習する前記ニューラルネットワークに基づいて決定される前記ロボット装置の意図する経路を、前記ディスプレイに表示するプログラムコードと、

前記現在の環境における第 2 の運転可能領域を識別するための入力を前記ユーザ装置で受信するプログラムコードと、

前記第 2 の運転可能領域を示すメッセージを前記ロボット装置に送信し、前記メッセージの送信に応じて前記ニューラルネットワークが更新されるようするプログラムコードと、を含む、非一時的なコンピュータ可読媒体。

【請求項 18】

30

前記プログラムコードは更に、前記訓練環境の運転可能領域内で過去に走行した一つ以上の経路に基づいて前記意図する経路を判断するよう前記ロボット装置を訓練するプログラムコードを有する、請求項 17 に記載の非一時的なコンピュータ可読媒体。

【請求項 19】

前記プログラムコードは更に、前記意図する経路に重ね合わせて前記第 1 の運転可能領域を表示するプログラムコードを有する、請求項 17 に記載の非一時的なコンピュータ可読媒体。

【請求項 20】

前記入力は前記第 2 の運転可能領域に対応するピクセルにマークを付することを含む、請求項 17 に記載の非一時的なコンピュータ可読媒体。

40

【発明の詳細な説明】

【技術分野】

【0001】

本出願は、2019年6月23日に出願され「キーフレームマッチャー」と題された米国特許仮出願第62/877,792、2019年6月23日に出願され「操作のための視覚的教示及び繰り返し 教示VR」と題された米国特許仮出願第62/877,791、及び2019年6月23日に出願され「視覚化」と題された米国特許仮出願第62/877,793の利益を主張する、2019年9月13日に出願され「ロボット装置からの視覚的出力を補強するシステム及び方法」と題された米国特許出願第16/570,540の利益を主張し、当該出願の全内容はここに参照により組み込まれるものとする。

50

【 0 0 0 2 】

本開示の特定の態様は概してロボット装置に関し、特にロボット装置からの視覚的出力を表示して編集するためのユーザーインターフェースを提供するシステム及び方法に関する。

【背景技術】**【 0 0 0 3 】**

自律エージェント（例えば車両、ロボット、ドローン等）及び半自律エージェントは周囲の環境における関心領域を解析するためにマシンビジョンを使用する。

運転中は、自律エージェントは周囲の環境の画像中の関心領域中に存在する物体を識別するために訓練されたニューラルネットワークに頼ることがある。例えば、ニューラルネットワークは、光検出と測距（L I D A R）センサ、ソナーセンサ、R G B カメラ、R G B - D カメラ等のような1つ以上のセンサにより捉えられた物体を識別し追従するよう訓練されていてもよい。センサは、自律エージェントのような装置と接続されるか、又は通信中であってもよい。自律エージェントのための物体検知アプリケーションは、自律エージェントの周囲の風景から物体（例えば歩行者、自転車に乗っている人、他の車等）を検知するために、センサ画像データを解析してもよい。

10

【 0 0 0 4 】

従来のシステムでは、エージェントの視覚的出力がオペレーターに提供され得た。例えば、意図する経路のような、エージェントが意図する経路がオペレーターに出力されるかもしれない。出力が正しくなければ、エージェントは正しい出力を得るために再度訓練されてもよい。関連するニューラルネットワークを再度訓練する必要なしにエージェントの出力を修正するようにシステムを改善することが望ましい。

20

【発明の概要】**【 0 0 0 5 】**

本開示のある態様では、ロボット装置により生成されたデータを可視化する方法が開示される。方法はロボット装置の意図する経路を環境中に表示することを含む。方法はまた環境中にロボット装置により運転可能と識別された(identified as drivable for the robotic device)第1の領域を表示することを含む。方法は更に環境中の第2の領域を運転可能として識別するための入力を受信することを含む。方法は更に前記第2の領域を前記ロボット装置へ送信することを含む。

30

【 0 0 0 6 】

本開示の別の態様では、非一時的なプログラムコードを記憶した非一時的なコンピュータ可読媒体が開示される。プログラムコードはロボット装置が生成したデータを可視化するためのものである。プログラムコードはプロセッサにより実行され、環境中にロボット装置の意図する経路を表示するプログラムコードを含む。プログラムコードはまた環境中にロボット装置が運転可能と識別された第1の領域を表示するプログラムコードを含む。プログラムコードは更に環境中の第2の領域を運転可能として識別するための入力を受信することを含む。プログラムコードは更に第2の領域をロボット装置へ送信するプログラムコードを含む。

30

【 0 0 0 7 】

本開示の別の態様は、ロボット装置により生成されたデータを可視化する装置に関する。装置はメモリ及びメモリに接続された1つ以上のプロセッサを有する。プロセッサはロボット装置の意図する経路を環境中に表示する。プロセッサはまたロボット装置にとって運転可能と識別された環境中の第1の領域を表示する。プロセッサは更に環境中の第2の領域を運転可能として識別するための入力を受信する。プロセッサは更に第2の領域をロボット装置へ送信する。

40

【 0 0 0 8 】

以上、この後の詳細な説明がより良く理解されるために本開示の特徴及び技術的利点を広く大まかに説明した。本開示の追加の機能及び利点を以下に説明する。本開示は、本開示と同じ目的を実行するためのその他の構造を変更又は設計するための基礎として容易に

50

使用され得るものであることが、当業者により理解されるはずである。そのような同等の構成は添付する特許請求の範囲により規定される本開示の教示から逸脱しないことも、当業者により認識されるはずである。本開示の特徴であると考えられる新しい機能は、その構成及び操作方法に関して、更なる目的及び利点と共に、添付する図と併せて考慮されたときに以下の説明からよりよく理解されるであろう。しかしながら、各図面は例示及び説明のみを目的として提供され、本開示の限界を定義することは意図しないことは、明白に理解されるべきである。

【図面の簡単な説明】

【0009】

本開示の機能、性質、及び利点は、類似の参照文字が全体に渡って対応する図面と組み合わせて考慮した場合に、以下に行う詳細な説明からより明らかになる。

10

【0010】

【図1】本開示の態様によりオペレーターがロボット装置を制御する例を示す。

【図2A】本開示の態様により環境中にて動作するロボット装置の例を示す。

【図2B】本開示の態様により環境中にて動作するロボット装置の例を示す。

【図2C】本開示の態様により環境中にて動作するロボット装置の例を示す。

【図3A】本開示の態様により環境中にて動作するロボット装置により提供された情報に基づいて生成された視覚的出力の例を示す。

【図3B】本開示の態様により環境中にて動作するロボット装置により提供された情報に基づいて生成された視覚的出力の例を示す。

20

【図4】本開示の態様によるロボット装置のハードウェア実装の例を示す。

【図5】本開示の態様による教示された動作のグラフシーケンスを示す。

【図6】本開示の態様によるロボットシステムのためのソフトウェアモジュールを示す。

【図7】本開示の態様によるロボット装置を制御する方法を示す。

【発明を実施するための形態】

【0011】

以下に行う添付図面に関連した詳細な説明は、様々な構成の説明を意図し、本明細書内で説明する概念を実施する単一の構成を提示することを意図しない。詳細な説明は、様々な概念の完全な理解を提供する目的により、特定の詳細を含む。しかしながら、これらの概念がこれらの特定の詳細なしに実施され得ることが当業者には明らかであろう。幾つかの事例では、そのような概念を不明瞭にすることを避けるために、周知の構造及び構成要素がブロック図にて示される。

30

【0012】

教示に基づき、本開示の範囲は、独立して実装されるか本開示のその他の態様と組み合わせるかに関わらず、本開示の任意の態様を含めることを意図することが、当業者により理解されるべきである。例えば、明らかにされる任意の数の態様を使用して装置を実装してもよく、又は方法を実施してもよい。加えて、本開示の範囲は、本開示で明らかにする様々な態様に加えて、又はその他の構造及び機能、又は構造及び機能を使用して実施されるそのような装置又は方法を含むことを意図する。本開示の任意の態様は特許請求の範囲の1つ以上の要素により具現化され得ることが理解されるべきである。

40

【0013】

本明細書において「例示的な」という語は「例、実例、又は例証の役割を果たす」という意味で使用される。「例示的」として説明される本明細書の任意の態様は必ずしも他の態様に比べて好ましい又は有利であるとして理解されるべきものではない。

【0014】

本明細書にて特定の態様を説明するが、本開示の範囲にはこれらの態様に対する多数の変形及び置換が含まれる。好ましい態様の幾つかの利益及び利点が記載されるが、本開示の範囲は特定の利益、使用又は目的に限定されることを意図しない。寧ろ、本開示の態様は、一部を図及び好ましい態様の説明に例示を目的として示す異なる技術、システム構成、ネットワーク、及びプロトコルへ広く適用可能であることが意図される。詳細な説明及

50

び図面は限定することよりも寧ろ本開示の説明のみを目的とし、本開示の範囲は添付する特許請求の範囲及び同等物によって定義される。

【 0 0 1 5 】

自律エージェント及び半自律エージェントは検出された物体に対してタスクを実行する。例えば、エージェントは環境中を走行し、環境中の物体を識別し、そして／又は環境中の物体と相互作用し得る。本願では、ロボット又はロボット装置は自律エージェント又は半自律エージェントのことである。簡略化のために、ロボット装置はロボットと称される。

【 0 0 1 6 】

本開示の態様は特定の型のロボットに限定されない。視覚／視覚化出力を提供する1つ以上の視覚システムを有する様々な型のロボット装置が考えられる。視覚出力は供給ビデオとしてオペレーターへ提供されてもよい。

10

【 0 0 1 7 】

本開示の態様によれば、ロボットは移動操作が可能である。すなわち、ロボットは関節構成の機能としてエンドエフェクタの位置を変更する能力を有する。ある構成では、ロボットは自動的に全身を制御し計画を立てる機能を備えている。そのことにより人間のオペレーターが運動学的制約又はロボットの姿勢をほとんど又は全く気にすることなく、タスク空間における継ぎ目ないエンドエフェクタ動作を仮想現実（VR）で実演することができる。

【 0 0 1 8 】

ロボットをVR環境にて訓練することに加えて、テスト時間中に訓練をアップデートすること（例えば実世界環境）が理想的である。本開示の態様はロボットのためにニューラルネットワークにより予測された経路を表示して編集するためのユーザーインターフェースを提供することに関する。本開示の態様はまたキーフレームマッチングの可視化及びロボットにより実行される処理を可視化する1つ以上のチャートを出力することに関する。

20

【 0 0 1 9 】

ある構成では、ロボットは、パンチルト雲台上の赤・緑・青及び距離（RGB-D）視野センサのような1つ以上のセンサを含む。RGB-D画像はRGB画像及び対応する奥行画像の組み合わせである。奥行画像とは、各ピクセルが画像平面とRGB画像上の対応する物体との距離に関連付けられている画像チャンネルである。

【 0 0 2 0 】

ロボットは環境の供給ビデオ（例えば供給画像）から得られた画像に基づいて環境中を走行してもよい。ロボットはまた供給ビデオからの画像に基づいてタスクを実行してもよい。従来の視覚システムでは、ロボットは特徴点に基づいたアプローチにより画像（例えば物体の画像）を識別する。

30

【 0 0 2 1 】

すなわち、タスクはロボットの現在の視覚の画像から抽出された特徴にもとづいて実行される。特徴点に基づくアプローチでは、テスト画像と参照画像の類似性に基づいてテスト画像（例えば現在の画像）を識別する。特に、従来の視覚システムではテスト画像の特徴と参照画像の特徴を比較する。テスト画像は、特徴が参照画像の特徴と一致したときに識別される。

40

【 0 0 2 2 】

特徴とは、画像の明白な（例えばユニークな）特色（例えば角、端、高コントラスト領域、低コントラスト領域等）である。画像は複数の特徴を含んでもよい。ディスクリプタが画像中の特徴をエンコードする。特徴ベクトルはディスクリプタの例であってもよい。従来の視覚システムでは、ディスクリプタは画像のユニークな特徴（例えば画像中の物体）のエンコードに限られていた。

【 0 0 2 3 】

ほとんどの場合、ディスクリプタは画像変換（例えば局在化、スケール、明るさ等）に対してロバストである。すなわち、従来の視覚システムは、様々な照明条件などの様々な条件において、物体を識別し得る。それにも関わらず、ディスクリプタが画像にユニーク

50

な特徴のエンコードに限られているため、ロバスト性は限定的である。ロボットが画像を識別してタスクを実行する能力を改善するためには、画像の視覚システムのロバスト性を改善することが理想的である。

【0024】

本開示の態様は画像のピクセルにディスクリプタを割り当てることに関する。すなわち、従来の視覚システムとは対照的に、本開示の態様ではディスクリプタはユニークな特徴に限られない。したがって、視覚システムの正確性が進歩し得る。特に、従来の視覚システムはテスト画像の明白な特徴（例えば特色）を参照画像の明白な特徴と比較することに限られていた。画像は特徴よりも多くのピクセルを含む。従って、特徴の代わりにピクセルを比較することにより、比較の正確性が増大する。

10

【0025】

すなわち、ある構成では、視覚システムはテスト画像のピクセルを参照画像のピクセルと比較する。参照画像はキーフレームと称されてもよい。キーフレームは訓練中に取得されてもよい。キーフレームの使用により姿勢や画像変換に対する不变性が得られる。視覚システムは一致したピクセルの数が閾値よりも大きいときに、テスト画像がキーフレームに一致したことを判断する。

【0026】

ある構成では、視覚システムは、テスト画像とキーフレームとの一致ピクセルを識別するために、ピクセルのディスクリプタを比較する。ピクセルディスクリプタはピクセルレベル情報と奥行情報を含む。ピクセルレベル情報はピクセルのRGB値や、画像／周囲のピクセル中におけるピクセルのコンテキストのような情報を含む。奥行情報は、ピクセルに対応し、面からその面の画像を取得するために使用されたセンサへの距離を示す。

20

【0027】

訓練中、ロボットがタスクを実行すると、1つ以上のセンサが風景の画像データを取得する。キーフレームは取得された画像データから生成される。キーフレームはロボットが特定のタスクを実行していた時の環境の記憶として概念化されてもよい。すなわち、キーフレームは環境中のタスク又は位置へのアンカーとして設定されてもよい。テスト後、ロボットにより現在の画像とキーフレームが一致したとき、キーフレームに対応する訓練された機能が実行されてもよい。

【0028】

30

キーフレームのピクセルにディスクリプタが割り当てられてもよい。ディスクリプタは値のベクトル又は配列であってもよい。例えば、ディスクリプタは16要素を持つベクトルであってもよい。ディスクリプタ（例えばピクセルディスクリプタ）は一致する画像を識別するために使用されてもよい。

【0029】

訓練後、ロボットは1つ以上の訓練されたタスクを実行するために初期化されてもよい。ロボットは命令されたタスクを実行する際に生の画像データを収集してもよい。生の画像データはロボットが通過及び／又は環境中の物体を操作する際にキーフレーム内のピクセルにピクセルレベルにてマッピングされる。現在の画像とキーフレームの対応関係はピクセルディスクリプタの類似性に基づいて判断される。

40

【0030】

すなわち、ロボットは現在の画像のピクセルディスクリプタの値をキーフレームのピクセルディスクリプタと比較する。比較の確信度は現在の画像のピクセルがキーフレームのピクセルへ一致する可能性を示す。現在の画像データ及びキーフレームの複数のピクセル間の斜め／並行の対応関係（例えば非交差相対関係）は、一致を示し得る。それは、交差関係は現在の画像データとキーフレームの間の非一致を示すということである。

【0031】

ある構成では、ロボットは環境中に位置し、風景の画像データが収集される。その後ロボットは（例えば仮想現実インターフェースを通して）制御され、タスクを実行する。仮想現実にて人間のオペレーターの視界をロボットの視界へと制限することにより、訓練中

50

にロボットが1人でタスクを実行する能力が改善される。

【0032】

ロボットは、仮想現実インターフェースを通してオペレーターが行う動作をパラメータ化することによりアクション／タスクを実行するよう教示されてもよい。例えば、仮想現実インターフェースは、オペレーターがVR環境を描画するヘッドセットを着用して操作するパドル、手持ちのコントローラ、ペイントブラシツール、拭き取り用具、及び／又は配置ツールの使用を含んでもよい。従って、人間のオペレーターは、直接的にタスク空間での動作を教示するよりも、パラメータ化されたプリミティブ（又は動作）のセットを教示する。パラメータ化されたプリミティブは、教示されるパラメーターを削減し、実行中のロバスト性を提供するために、衝突のない動作計画とハイブリッド（位置及び力）のデカルト制御を組み合わせる。

10

【0033】

パラメータ化された動作とは、そのタスクを小さい数の分離した動作の塊に分けることにより学習することである。各動作は、関節角度変化、回転角度等のパラメーターのセットにより定義される。これらのパラメーターの値はタスクを実行するときのロボットの状況に基づいて構成され更新されてもよい。

【0034】

パラメータ化された動作は、1つの学習済みタスクから学習及び抽出され、より大きいタスクを形成するためにその他の対応タスクと組み合わされてもよい。回転するハンドルを有するドアを開けるようなパラメータ化された動作は、任意のドアハンドルを開けることを実行することとして実装されてもよい（例えば30度の回転を要するドア、又は60度の回転を要するドア、又はより多くの回転）。例えば、回転角度は回転するドアハンドルを有するドアを開けることに対するパラメータ化された動作を定義する1つのパラメーターであり得る。

20

【0035】

タスクを実行するには、ロボットは（訓練中の初期位置から相対的に）同じ又は類似した環境中に位置してもよい。ロボットは異なるスタート地点に位置してもよく、相対的に異なる初期姿勢をとってもよい（例えば関節角度が異なるスタート位置へ調整される）。ロボットは瓶を取り上げる、キャビネットを開ける、そして瓶をキャビネットに入れるといった、（人間のオペレーターによる制御なしに）同じタスクの実行（例えばパラメータ化された動作のセット）を課され得る。例えば、ロボットは仮想現実にて制御されたシーケンスの中で教示された動作のパラメーターを更新することにより同じタスクを実行してもよい。パラメーターは、訓練中に使用された姿勢及び／又は位置に比較して、ロボットの現在の姿勢及び／又は位置に基づいて更新されてもよい。

30

【0036】

議論したように、パラメーターを更新するには、ロボットは風景の初期の画像を撮影し、新しい画像からキーフレームへピクセル及び／ディスクリプタをマップする。マッピングにより、初期画像とキーフレームとの間ににおける相対的な変換が定義される。相対的な変換によりキーフレームを新たな画像にマッピングすることができる。相対的な変換はロボットのx軸上の位置、y軸上の位置、z軸上の位置、ロール、ピッチ、及びヨーの変化により定義されてもよい。相対的な変換は、パラメータ化された動作のパラメーターを、教示されたパラメーターから観察された状況へアップデートするために使用されてもよい。

40

【0037】

相対的な変換はパラメータ化された動作に適用されてもよい。パラメータ化された動作に相対的な変換を適用することにより、ロボットは、初期位置及び／又は姿勢が変わっても、同じタスクを以前教示されたように実行し得る。ロボットシステムは、ピクセル及び／又は密集したニューラルネットワークのディスクリプタを、現在の風景からのものからキーフレームからのものへ、パラメータ化された動作から間隔を空けて（例えば連続して）調整が行われるように、継続してマッピングしてもよい。例えば、引き出しを引

50

いて開ける、ドアを開ける、カップ又は瓶を取り上げる等のような、パラメーター化された動作のセットにより定義された教示されたアクションへ、相対的な変換が適用されてもよい。

【0038】

幾つかの態様では、アクションは風景全体に関連するか、及び／又は物体特有であってもよい。例えば、瓶を取り上げるアクションを行うには、瓶のところまで走行するために風景全体に関連したキーフレームの使用が必要になり得る。瓶に接近すれば、瓶に特有のキーフレームが環境から独立して解析されてもよい。

【0039】

走行の動作はロボットのある地点から別の地点へ移動させるために使用されてもよい。このことにより、「取り上げる」アクションの訓練中に、ロボットが環境中の任意の場所に位置し得る物体の位置を特定し、その瓶の位置に関わらず「取り上げる」といったようなタスクを実行することができるようになり得る。操作する動作は、所望の物体と接触するためにロボットの部品（例えば胴体及び／又は腕）を動かすことに使用され得る。

10

【0040】

図1は本開示の態様によりオペレーター100がロボット106を制御する例を示す。図1に示されるように、オペレーター100は、ロボット106を制御するための視覚システム102及び動作制御器104（例えばジェスチャ追従システム）を備える。この例では、オペレーター100はロボット106を1つ以上のタスクを実行するために制御してもよい。図1の例では、ロボット106はキッチン108にて訓練される。本開示の態様はロボット106をキッチン108で訓練することに限定されず、その他の環境も考慮されている。

20

【0041】

視覚システム102はオペレーター100の視覚を捕捉するだけでなく、供給ビデオを提供してもよい。オペレーター100はロボット106から離れた位置にいてもよい。本例では、ロボット106はキッチン108に位置し、オペレーター100はロボット制御センター114のようなキッチン108と異なる場所に位置する。

【0042】

視覚システム102はロボット106の供給ビデオを提供してもよい。例えば、視覚システム102はロボット106の前方視点に基づくキッチン108の光景を提供してもよい。360°の風景のようなその他の視点が提供されてもよい。視点はロボット106のビデオカメラのような1つ以上の視覚センサを使用して提供される。視覚システム102は図1に示されるようなヘッドセットに限定されない。視覚システム102はモニタ110、画像プロジェクタ、又はロボット106からの供給ビデオを表示可能なその他の装置であってもよい。

30

【0043】

ロボット106の一つ以上のアクションが動作制御器104を介して制御されてもよい。例えば、動作制御器104はオペレーター100のジェスチャを捉え、捉えたジェスチャをロボット106が真似してもよい。オペレーター100はロボット106の運動、手足の動作、及びその他のアクションを、動作制御器104を介して制御してもよい。本開示の態様は動作制御器104を介してオペレーター100のジェスチャを捉えることに限定されない。その他の型のジェスチャ捕捉システムも考えられる。オペレーター100は無線接続112を介してロボット106を制御してもよい。加えて、ロボット106は無線接続112を介してオペレーター100へ供給ビデオのようなフィードバックを提供してもよい。

40

【0044】

ロボット106は特定の環境（例えばキッチン108）及び／又は類似した環境を走行するよう訓練されていてもよい。加えて、ロボット106は特定の環境中の物体、及び／又は任意の環境中の類似した物体に対してタスクを実行するように訓練されていてもよい。例えば、ロボット106はキッチン108内の引き出しを開け閉めするよう訓練されて

50

いてもよい。訓練はキッチン 108 内の引き出しに対してのみ、及び / 又は別のキッチンの引き出しなどの、任意の環境内の類似した引き出しに対して実装されていてもよい。

【 0 0 4 5 】

議論したように、テスト時間中に訓練をアップデートすること（例えば実世界環境）が理想的である。本開示の態様はロボットのためにニューラルネットワークにより決定された経路を表示して編集するためのユーザーインターフェースを提供することに関する。ある構成では、ロボットは意図する経路を提供し、ユーザーはピクセルを運転可能又は運転不可能としてマークして意図する経路を更新してもよい。

【 0 0 4 6 】

訓練中、オペレーターはロボットを環境内で走行させてもよい。加えて、又は代わりに、ロボットは訓練中に自動的に環境中を走行してもよい。ロボットは走行した経路をメモリに記憶する。記憶された走行経路は、運転可能な経路と関係する画像中のピクセルと運転不可能な経路と対応する画像中のピクセルとを判断するよう、ネットワークを訓練するために使用される。新しい環境又は訓練に使用された環境と同じ環境で運転可能な経路を判断するように訓練されるロボット。

【 0 0 4 7 】

運転可能な経路はロボットが走行し得る経路である。ロボットは運転不可能な経路を走行しないことがあり得る。例えば、経路上の物体により、経路が運転不可能なものとなり得る。ある構成では、運転可能な経路を判断するために奥行データが使用される。物体の高さが閾値以上の場合、物体のある経路の一部が運転不可能としてマークされる。

【 0 0 4 8 】

図 2 A は本開示の態様による環境 202 中のロボット 200 の例を示す。明確にするために、図 2 A は環境 202 の上面図である。図 2 A に示されるように、環境 202 はダイニングテーブル 204、シンク 206、及びカウンター 210 を含む。この例では、訓練中、ロボット 200 は環境 202 中を経路 212 に沿って走行した。ロボット 200 は経路 212 をメモリに記憶し、経路 212 をロボット 200 のニューラルネットワークを訓練するために使用する。訓練は、現在の環境 202 に加えて、現在の環境 202 と異なる環境で運転可能な経路を判断するために使用される。

【 0 0 4 9 】

図 2 B は本開示の態様による訓練されたニューラルネットワークに基づいて判断された、ロボット 200 の意図する経路 226 の例を示す。ニューラルネットワークは、訓練環境 202 からの経路 212 のような 1 つ以上の過去の経路に基づいて訓練される。図 2 B の現在の環境 250 は図 2 A の訓練環境 202 と異なる。それにも関わらずロボット 200 は、訓練中又は訓練後に訓練環境 202 中の意図する経路を決定してもよい。

【 0 0 5 0 】

図 2 B の例はロボット 200 の視点に基づいている。図 2 B の画像は遠隔のオペレーターへ出力され、遠隔地にある表示装置へ表示されてもよい。図 2 B に示すように、意図する経路 226 ではテーブル 220 の周りを走行する。意図する経路 226 に加えて、ニューラルネットワークは運転可能領域 228 を判断する。意図する経路 226 は運転可能領域 228 の境界内にある。ある構成では、運転可能領域 228 は意図する経路 226 に重なる。画像中のピクセルが運転可能領域 228 に属すると識別される。

【 0 0 5 1 】

図 2 B に示されるように、ニューラルネットワークは床の上の物体 224（例えば冷蔵庫 222 のコード）を識別する。訓練、物体 224 の奥行（例えば高さ）及び / 又はその他のファクタに基づき、ニューラルネットワークは物体 224 の領域は運転不可能であると判断する。かくして運転可能領域 228 は物体 224 の領域と交差しない。

【 0 0 5 2 】

ある構成では、ロボット 200 の視点がオペレーターへ出力されるとき、オペレーターはニューラルネットワークが運転可能領域 228 を正しく識別したかどうかを判断する。オペレーターはニューラルネットワークの決断を修正するために運転可能領域 228 を更

10

20

30

40

50

新してもよい。更新は、ディスプレイ中の領域を運転可能又は運転不可能としてマークするようなユーザー入力を通して提供されてもよい。

【0053】

図2Cは本開示の態様によりロボット200の運転可能領域228を更新する例を示す。図2Cに示されるように、人間のオペレーターは環境250の画像を、新しい運転可能領域230をマークすることにより更新する。加えて、又は代わりに、人間のオペレーターは運転可能領域228を運転不可能としてマークしてもよい(図2Cには示されない)。現在の環境における運転可能領域228を編集した後は、ニューラルネットワークは後続の経路を判断する際に人間のオペレーターによる編集を考慮する。

【0054】

ニューラルネットワークのコアの重みは、新しいユーザーにより提供された編集済み情報によって変わることはない。ある構成では、サポートベクタマシン(SVM)がネットワーク出力の末端に付加される。SVMの重みはユーザーにより編集された情報に基づいて更新される。SVMの重みの更新はリアルタイムで行われてもよい。従って、SVMの更新はネットワーク全体を再訓練するよりも速い。更に、ユーザーインターフェースを使用して新たに教示された運転可能/運転不可能領域の重要性を調整することが可能である。例えば、スライダーによりユーザーが新たな運転可能/運転不可能領域の重みを調整できてもよい。従って、SVMは適切な分類をリアルタイムで行わされ得る。

【0055】

議論したように、訓練後、ロボットは1つ以上の訓練されたタスクを実行するために初期化されてもよい。例えば、ロボットは環境中を走行し、そして/又は環境中の物体を操作するようなタスクが課されてもよい。ロボットは命令されたタスクを実行する際に生の画像データを収集する。生の画像データはキーフレームと比較される。

【0056】

一つ以上のキーフレームが現在の画像データと一致する場合は、合致したキーフレームと関連した動作が実行されてもよい。例えば、ロボットは引き出しを開け、又は閉めてもよい。別の例としては、ロボットは蛇口を開けるか、又は閉めてもよい。加えて、又は代わりに、前の動作が完了すると、行動の実行がキューに追加されてもよい。別の例として、ロボットはキーフレームに基づいて現在の位置を判断してもよく、現在の位置から別の位置へ経路を走行するために使用してもよい。ロボットが環境中を移動するか、又は環境中の物体の位置が変更されるに従って、新しいキーフレームが参照されてもよい。

【0057】

ある構成では、ピクセルディスクリプタを比較する前に1つ以上のピクセルがフィルタされる。外れ値をフィルタするために関数が実行されてもよい。例えば、観察されたデータをランダムにサンプルすることでピクセルの外れ値をフィルタするためにrandom sample consensus(RANSAC)関数を使用してもよい。ピクセルをフィルタすることによりノイズが削減され正確性が向上する。

【0058】

議論したように、現在の画像の奥行及び/又は角度は一致したキーフレームの奥行及び/又は角度から異なり得る。従って、ある構成では、現在の画像データの一致したピクセルとキーフレームのピクセルとの間の相対変換が判断される。すなわち、現在の画像とキーフレームとの比較に基づいて姿勢デルタが判断され得る。デルタはキーフレームから現在の画像へのロボットの姿勢のことである。

【0059】

姿勢デルタはピクセルディスクリプタの奥行情報に基づいて判断されてもよい。奥行情報はセンサからピクセルに対応する物体への距離である。姿勢デルタは現在の画像とキーフレームとの間の(x, y, z)座標、ロール、ピッチ、及びヨーの変化である。すなわち、姿勢デルタにより、キーフレームから現在の画像へと画像がどのように変化したかがセンサの移動に基づいて説明される。

【0060】

10

20

30

40

50

一致したピクセルの数の増加に関連して姿勢デルタの正確性が増加する。本開示では、一致したピクセルの数は一致した特徴の数よりも大きくなり得る。かくして、本開示の態様は画像を特徴のみに基づいて比較する従来のシステムと比較して姿勢デルタの判断の正確性を向上させる。姿勢デルタを判断するために最小二乗関数を使用してもよい。

【0061】

実行される動作のパラメーターの更新に、ピクセルのピクセルディスクリプタの1つ以上の数値の間の姿勢デルタ（例えば相対変換）を使用してもよい。例えば、ロボットがキーフレームにて定義された位置と比較して物体から1フィート（約30.5センチメートル）離れていることを姿勢デルタが示している場合、ロボットの現在の位置（例えば姿勢）を考慮して動作が更新されてもよい。キーフレームにて定義された位置は、タスクを実行するためにロボットを訓練する際に使用する位置である。

10

【0062】

図3Aは本開示の態様による視覚システムからの出力300の別の例を示す。図3Aに示されるように、出力300は現在の画像302、キーフレーム308、ヒストグラム306、及び姿勢確信度ウインドウ304を含む。出力300はロボットの現在の位置から離れた位置に表示されてもよい。例えば、出力300はロボットの動作を制御する制御センター（例えば図1のロボット制御センター114）に表示されてもよい。

【0063】

現在の画像302はロボットの現在の視界からの画像である。キーフレーム308は、現在の画像302と比較されるキーフレームのセットからの1つのキーフレームである。ヒストグラム306は現在の画像302からキーフレームのセット中の各キーフレームへの一致に関する確信度スコアを示す。

20

【0064】

訓練中には、走行タスクにおいて、ロボットは経路に沿ったスタートからゴールまでの特定の地点にて、キーフレームを撮影する。経路上で撮影されたキーフレームは走行タスクにおいてキーフレームのセットとして使用される。テスト中、ロボットは現在の位置を識別するために現在の画像をキーフレームのセットと比較する。

【0065】

一致したピクセル（例えばピクセルディスクリプタ）の数に基づいて、ロボットは現在の画像から特定のキーフレームへの一致の確信度を判断する。一致の確信度が閾値よりも高ければ、ロボットは一致したキーフレームに関連したタスクを実行する。タスクは位置判定（例えばロボットの現在の位置を判断すること）を含んでもよい。図3Aの例にて、ロボットは経路上の、一致の確信度310が最も高いキーフレームの位置に対応する位置にいると判断する。

30

【0066】

ヒストグラム306はタスクに応じて異なってもよい。図3Aのヒストグラム306は走行タスクのためのものである。ヒストグラム306はロボットが操作（例えば物体の操作）タスクを実行している場合には異なるものとなり得る。操作タスクのためのキーフレームのセットは、走行タスクのためのキーフレームのセットと比較して少ないキーフレームを有していてもよい。

40

【0067】

例えばキャビネットを開ける場合、ロボットはキャビネットが閉じていたり少し開いていたりする場合に開けるように訓練されていてもよい。この例では、ヒストグラムは、棒グラフ中に2本のキーフレーム一致棒グラフを有し得る。1本の棒グラフは現在の画像が閉じたキャビネットのキーフレームに一致する確信度を示し、もう1本の棒グラフは現在の画像が部分的に開いたキャビネットのキーフレームに一致する確信度を示し得る。

【0068】

姿勢確信度ウインドウ304は姿勢一致器から得た姿勢一致の確信度を示す。ある構成では、姿勢デルタを判断する前に、ロボットは数々の基準が1つ以上の閾値を満足するかどうかを判断する。基準は奥行、姿勢、クリーク、エラー、及びその他の要素に基づいて

50

いてもよい。

【 0 0 6 9 】

ある構成では、ロボットは奥行値を有するピクセルの数を判断する。ガラス又は艶のある面のような面では、ロボットは奥行を判断できないことがある。したがって、これらの面に対応するピクセルは奥行値を有さない場合がある。奥行値を有するピクセルの数はピクセルをフィルタする前に判断されてもよい。奥行値を有するピクセルの数が閾値よりも大きければ、奥行の基準は満足される。姿勢確信度ウインドウ 3 0 4 は、奥行値を有するピクセルの数の棒グラフ 3 2 0 と閾値を示す棒グラフ 3 2 2 とを比較する棒グラフを含んでいてもよい。棒グラフは色分けされていてもよい。

【 0 0 7 0 】

加えて、又は代わりに、姿勢確信度ウインドウ 3 0 4 の棒グラフは現在の画像とキーフレームの間で揃う特徴の数の棒グラフとアライメント閾値の棒グラフとの比較を含んでもよい。キーフレームと現在の画像の一一致の正確性を判断するために、ロボットは姿勢変換を適用し、現在の画像における物体の特徴がキーフレーム中の物体の特徴と幾つ揃うか、及びその反対を判断してもよい。もしも特徴が静的でない物体と紐づけられていた場合、特徴は揃わなくてもよい。例えば、キーフレームは現在の画像には既に存在しないカップを含み得る。従って、カップの特徴は画像間で揃わなくなり得る。揃った特徴の数はクリークインライアと呼ばれてもよい。クリークインライアが閾値よりも大きい場合に姿勢デルタが生成され得る。

【 0 0 7 1 】

加えて、又は代わりに、姿勢確信度ウインドウ 3 0 4 の棒グラフは姿勢を判断するために使用される特徴の数の棒グラフと姿勢閾値の棒グラフの比較を含んでもよい。すなわち、姿勢を判断するために一定数の特徴が使用されるべきである。姿勢を判断するために使用される特徴の数が姿勢閾値よりも小さい場合は、姿勢デルタは計算されなくてもよい。姿勢を判断するために使用される特徴の数は姿勢インライアと呼ばれてもよい。

【 0 0 7 2 】

姿勢確信度ウインドウ 3 0 4 の棒グラフは姿勢デルタの二乗平均平方根 (R M S) エラーの棒グラフと閾値の比較をも含み得る。R M S エラーが R M S エラー閾値よりも小さければ、姿勢デルタは申し分がないものであり得る。R M S エラーが R M S エラー閾値よりも大きければ、ユーザー又はロボットはタスクを実行するために姿勢デルタを使用しなくてもよい。

【 0 0 7 3 】

姿勢確信度ウインドウ 3 0 4 は棒グラフに限定されず、他のグラフ又は画像が使用されてもよい。姿勢確信度ウインドウ 3 0 4 の基準は議論された基準に限定されず、他の基準が用いられてもよい。

【 0 0 7 4 】

図 3 B は本開示の態様による視覚システムからの出力 3 5 0 の、更に別の例を示す。図 3 B に示されるように、出力 3 5 0 は現在の画像 3 5 2 、キーフレーム 3 5 8 、ヒストグラム 3 5 4 、及び姿勢確信度ウインドウ 3 5 6 を含む。出力 3 5 0 はロボットの現在の位置から離れた位置に表示されてもよい。

【 0 0 7 5 】

姿勢確信度ウインドウ 3 5 6 は一致経路に渡る特徴の数及び R M S エラーを可視化する。各特徴（例えば基準）に対して、現在の値が閾値と比較される。例えば、奥行に関して、現在の奥行（ 3 8 9 ）が閾値（ 6 5 ）と比較される。

各棒グラフの棒は現在の値と閾値に関する。

【 0 0 7 6 】

ヒストグラム 3 5 4 は現在の画像と一致が行われる各キーフレームの一致スコアを可視化する。閾値よりも高く最も高い一致スコアを有するキーフレームが一致するキーフレームとして選択され得る。一致したキーフレームは出力 3 5 0 の中にキーフレーム 3 5 8 として表示されてもよい。

10

20

30

40

50

【 0 0 7 7 】

議論したように、本開示の態様は、仮想現実内の人間からタスクをデモンストレーションにより教示された後に、実世界環境にて自律的に人間のレベルのタスクを実行することが可能な移動操作ハードウェア及びソフトウェアシステムに関する。ある構成では、移動操作ロボットが使用される。ロボットは全身タスク空間ハイブリッド位置／力制御を含んでもよい。加えて、議論したように、ロバストに学習された風景の密な視覚的エンベディング表現にリンクされた、パラメーター化されたプリミティブがロボットに教示される。最後に、教示された動作のタスクグラフが生成されてもよい。

【 0 0 7 8 】

固定された物体のセットを認識したり予め定義されたタスクを実行したりするようにロボットにプログラミングや訓練を行うよりも、本開示の態様によりロボットは新しい物体やタスクを人間によるデモンストレーションから学習することができる。学習されたタスクは自然に変化する条件下でロボットにより自律的に実行されてもよい。ロボットは以前の物体モデル又はマップを使用せず、1つの例から、与えられた動作のセットを任意の風景及び物体へ関連付けるよう教示されることがある。視覚システムは既存の教師付き及び教師なしデータセットを使用してオフラインで訓練されてもよく、システムの残りは追加の訓練データなしに機能してもよい。

10

【 0 0 7 9 】

タスク空間の動作を直接教示する従来のシステムとは対照的に、本開示の態様はパラメーター化された動作のセットを教示する。これらの動作は、教示されたパラメーターを最小化し、実行中のロバスト性を提供するために、衝突のない動作計画とエンドエフェクタのハイブリッド（位置及び力）デカルト制御を組み合わせる。

20

【 0 0 8 0 】

ある構成では、タスクに特化した、訓練された密な視覚に関するエンベディングが計算される。このピクセルに関するエンベディングはパラメーター化された動作を風景にリンクさせる。リンクにより、システムは、新しい状況への一般化と引き換えにロバスト性の高い様々な環境を取り扱い得る。

【 0 0 8 1 】

タスクの動作は、視覚的な入力条件と、成功に基づく終了条件を用いて独立に教示されてもよい。動作は動的タスクグラフ内にて互いに連結されてもよい。動作が連結されているため、ロボットはタスクシーケンスを実行するため行動を再利用してもよい。

30

【 0 0 8 2 】

ロボットは多自由度（D O F）であってもよい。例えば、ロボットは車台、胴体、左腕、右腕、及び頭の5つのサブシステムに分けられた31自由度（D O F）であってもよい。ある構成では、車台は、「疑似ホロノミック」な可動性を実現する4つの駆操舵可能な動輪（例えば計8自由度）を含む。駆動／操舵アクチュエータパッケージは様々なモーター及びギアヘッドを含んでもよい。胴体は5自由度（ヨー、ピッチ、ピッチ、ピッチ、ヨー）であってもよい。それぞれの腕は7自由度であってもよい。頭はパン／チルトの2自由度であってもよい。それぞれの腕は劣駆動の指を有する1自由度のグリッパーを含んでもよい。本開示の態様は上に議論したロボットに限定されない。その他の構成が考えられる。一例では、ロボットはスポンジ又はモップのようなカスタムツールを含んでもよい。

40

【 0 0 8 3 】

ある構成では、ロボットには環境との相互作用力を計測するための力／トルクセンサが一体化されている。例えば、力／トルクセンサはそれぞれの腕の手首に配されていてもよい。頭には、広い視野を提供し、人間やロボットがタスクを実行するためのVRコンテキストを提供するための知覚センサが統合されていてもよい。

【 0 0 8 4 】

本開示の態様はロボット制御のための数段階の抽象化を提供する。ある構成では、最も低い制御レベルにより、ロボットの全ての自由度のリアルタイムな協調制御が提供される。リアルタイム制御は関節制御及び部品の制御を含み得る。関節制御により、低レベルの

50

装置通信が実装され、装置のコマンドや状態を一般的な形で公開する。加えて、関節制御はアクチュエータ、力センサ、及び慣性計測装置をサポートする。関節制御は異なるロボットをサポートするためにランタイムにて構成されていてもよい。

【 0 0 8 5 】

部品制御により、ロボットを部品（例えば右腕、頭等）に分け、各部品に対してパラメーター化された制御器のセットを提供することで、ロボットのより高レベルの協働作用を扱い得る。部品制御により、関節位置及び速度、関節アドミッタンス、カメラ目視、車台位置及び速度、並びにハイブリッドタスク空間における姿勢、速度、及びアドミッタンス制御の制御器が提供されてもよい。

【 0 0 8 6 】

エンドエフェクタのタスク空間制御により、ロボット制御を別次元で抽象化することが可能になる。このレベルの抽象化により、所望の動作を達成するためのロボットの姿勢の問題が解決される。ハイブリッドデカルト制御のための全身の逆運動学（IK）が二次のプログラムとして形成され、解かれる。部品には関節位置、速度、加速度、及び重力トルクに関して線形拘束が存在し得る。

【 0 0 8 7 】

全身のIKはデカルト座標における姿勢のゴールに到達するための動作計画に使用されてもよい。ある構成では、環境の占有されたボクセルは球体やカプセル型が当てはめられる。ロボットと世界の衝突を避けるために、ボクセルの衝突拘束がIKの二次プログラムに加えられる。IKの二次プログラムにて、ノード間の操舵関数としてデカルト空間におけるサンプリングを行い、`rapidly-exploring random tree (RRT)`を用いて動作計画が行われてもよい。

【 0 0 8 8 】

デカルト空間における計画により、自然で直接的な動作になる。操舵機能としてIKの二次プログラムを使用することにより、計画の信頼性を向上させることができ、計画と実行に同じ制御器を、両者の食い違いを削減するために用いてもよい。同様に、関節位置のゴールに向けた動作計画と、操舵機能として働く部品制御による関節位置制御器とを組み合わせてRRTを使用する。

【 0 0 8 9 】

次の抽象化レベルによりパラメーター化された動作が定義される。ある構成では、パラメーター化された動作は、パラメーター化可能で、組み合わせることでタスクを成し遂げることが可能な、プリミティブなアクションである。動作は、握る、持ち上げる、置く、引く、ひっこめる、拭く、直接制御、のような操作アクション、関節を動かす、速度命令により運転する、位置命令により運転する、能動的な障害物回避を行いながらの経路追従、のような走行アクション、及び目視して停止するといった予備的アクションを限定せず含んでもよい。

【 0 0 9 0 】

各動作は、ロボットの部品の一つ以上の関節又はデカルト座標における動作のような、单一又は複数の異なる型のアクションを有することができる。各アクションは位置、速度、又はアドミッタンス制御のような異なる制御手法を使用することができ、外部の障害物を回避するために動作計画を使用することを選択することができる。動作計画を使用するかに関わらないロボットの動作は、自己衝突を回避し、動作制御拘束を満たす。

【 0 0 9 1 】

各動作は異なるアクションによりパラメーター化され、その代わりにアクションは独自のパラメーターを有してもよい。例えば、把持動作はグリッパ角度、6Dアプローチ、把持、及びグリッパの（随意的な）持ち上げ時の姿勢の4つのパラメーターから成り立っていてもよい。本例では、これらのパラメーターにより、次の予め定義されたアクションのシーケンスが定義される。（1）所望のグリッパ角度までグリッパを開く（2）6Dアプローチ姿勢までの衝突のない経路を計画し実行する（3）6D把持姿勢までグリッパを動かし接触したら停止する（4）グリッパを閉じる（5）6D持ち上げポーズまでグリッパ

10

20

30

40

50

を動かす。

【 0 0 9 2 】

最終レベルの制御の抽象化がタスクである。ある構成では、タスクは、ロボットが操作を行い人間の環境を走行することを可能にする動作のシーケンスとして定義される。タスクグラフ（図5参照）は、異なるタスクをノードとし、異なる移動状況をエッジとし、異常検出及び異常からの回復を含む、有効であり、周期的又は非周期的なグラフである。エッジ状況は異なる物体や環境を取り扱うための各動作の実行状況、力／トルクセンサを使用した手中の物体の検査、音声コマンド、及びキーフレームとの一致を含む。

【 0 0 9 3 】

本開示の態様により、ロボットが周囲の環境を理解するための知覚パイプラインが設計される。知覚パイプラインによりロボットはまた、教示されたタスクを踏まえてどのアクションを取るべきかを認識する能力が得られる。ある構成では、一つの視界画像（例えばワイドな視界の左画像）に高解像度カラーステレオペアの複数の奥行画像を投影することにより、融合RGB-D画像が作成される。システムは様々なピクセルレベルの分類及び特徴ベクトル（例えばエンベディング）を提供するためにディープニューラルネットワークのセットを実行する。教示されたシーケンスから呼び出された視覚的特徴に基づき、ピクセルレベルの分類及び特徴ベクトルは一時的な3Dボクセル表現へと累積される。

ピクセルレベルの分類及び特徴ベクトルは実行するべきアクションを呼び出すために使用されてもよい。

【 0 0 9 4 】

ある構成では、物体のカテゴリは定義されない。加えて、若しくは物体のモデル又は環境は想定されない。物体を明示的に検出してセグメント分けし、また明示的に6自由度の姿勢を推定するよりも、多用なタスクのために密なピクセルレベルのエンベディングを生成してもよい。教示されたシーケンスからの参照エンベディングが動作分類又は姿勢推定を行うために使用されてもよい。

【 0 0 9 5 】

訓練されたモデルは完全な畳み込み型であってもよい。ある構成では、入力画像のピクセルはエンベディング空間の中のある点へそれぞれマップされる。エンベディング空間はモデルの出力により定義される損失関数と訓練手順によって暗黙のうちに定義されるメトリックを与えられる。訓練されたモデルは様々なタスクに使用されてもよい。

【 0 0 9 6 】

ある構成では、1つの注釈が付された例が与えられれば、訓練されたモデルはセマンティッククラス中の全てのオブジェクトを検出する。セマンティッククラス中の物体は注釈中のエンベディングとその他の領域中のエンベディングを比較することにより検出されてもよい。モデルは識別損失関数（discriminative loss function）により訓練されてもよい。

【 0 0 9 7 】

モデルはオブジェクトインスタンスを判断するために訓練されてもよい。このモデルは独立したオブジェクトを識別し、そして／又は数える。モデルは各ピクセルのベクトル（2Dエンベディング）を予測するために訓練されてもよい。ベクトルはそのピクセルを含むオブジェクトの重心を指示してもよい。ランタイムにおいて、同じ重心を指すピクセルはその風景のセグメントとしてグループ分けされてもよい。ランタイムにおける実行は3Dで行われてもよい。

【 0 0 9 8 】

モデルは3Dの対応関係について訓練されてもよい。このモデルは風景内の任意の3D地点のビューが同じエンベディングに対してマッピングされるように、ビューや照明に対して不变なエンベディングをピクセルごとに提供する。このモデルは損失関数を使用して訓練されてもよい。

【 0 0 9 9 】

各RGB-Dフレームに対するピクセルに関するエンベディング（及び奥行データ）は

10

20

30

40

50

動的 3 D ボクセルマップへと融合される。各ボクセルは第 1 及び第 2 の順序の位置、色、エンベディングの統計を累積する。動的オブジェクトの有効期限はボクセルの奥行画像への逆投影に基づく。ボクセルマップは、セマンティック及びインスタンスレベル、並びに幾何学的近似性に基づいて、標準的なグラフセグメンテーションを使用して分割される。ボクセルマップは標高及び走行可能性の分類統計を有する 2 . 5 D マップへと次元が削減される。

【 0 1 0 0 】

2 . 5 D マップは衝突のない車台の動作に使用される一方、ボクセルマップは衝突のない全身の動作計画に使用される。3 D における衝突の検査のために、貪欲法を使用してマップ中のボクセルがカプセルへとグループ化されてもよい。セグメント化されたオブジェクトは、物体が把持された際に手に付属させるための動作に使用されてもよい。

10

【 0 1 0 1 】

ロボットは過去に教示されたタスクにて記録された特徴に高度に関連する風景中の（又は特定の操作オブジェクトの）特徴を認識するように、ワンショット学習アプローチにより訓練されてもよい。タスクがユーザーによって実演されると、特徴がタスク全体にわたってキーフレームの形態で保存される。キーフレームは多次元エンベディング及びピクセルごとの奥行（有効であれば）を有する R G B 画像であってもよい。

【 0 1 0 2 】

エンベディングは、現在の画像が教示時に存在した参照画像と充分類似しているという想定の下でランタイムにおけるピクセルごとの対応関係を確立し得る特徴ディスクリプタとして機能する。奥行が（ほとんど）全てのピクセルに存在するため、現在の画像と参照画像の姿勢のデルタを解くために対応関係を使用することができる。ユークリッド制約を使用してインライアが検出されてもよく、6 自由度の姿勢を解くために R A N S A C と共にレーベンバーグ・マルカート最小二乗関数を適用する。

20

【 0 1 0 3 】

姿勢のデルタは、教示された動作のシーケンスを現在の風景へ適合させるために適用可能な補正の役割を果たす。エンベディングが各ピクセルに対して定義されてもよいことから、キーフレームは画像中の全てのピクセルを含む程広くても良く、又はユーザーが定義したマスク内のピクセルのみを使用する程狭くてもよい。議論したように、ユーザーは、画像中の領域をタスクと関連するものとして、又は物体上にあるものとして選択的に注釈を付することにより、マスクを定義してもよい。

30

【 0 1 0 4 】

視覚センシングに加えて、ある構成では、ロボットは音声入力を収集して処理する。音声は、ロボットを教示するための入力として、別のエンベディングのセットを提供する。例として、ロボットは質問を行い、人間からの応答の音声言語を理解することにより音声入力を得る。音声による応答はカスタムのキーワード検出モジュールを使用して理解されてもよい。

【 0 1 0 5 】

ロボットは、完全畳み込みキーワードスポットティングモデルを利用して、カスタムのウェイクワード、物体のセット（例えば「マグ」又は「瓶」）、及び場所のセット（例えば「キャビネット」又は「冷蔵庫」）を理解してもよい。ある構成では、モデルは、例えば 3 2 m s のようなある間隔でウェイクワードを聞いている。ウェイクワードが検出されたら、ロボットは物体又は場所のキーワードが検出されるかを注意する。訓練の間は、認識をよりロバストにするために人工的にノイズが加えられる。

40

【 0 1 0 6 】

議論したように、ロボットにタスクを教示するためには、オペレーターは V R モードのセットを使用する。各動作は、その動作に特有のパラメーターを設定し命令するために、対応する V R モードを有してもよい。各動作モードは、各パラメーターの設定を支援するために、パラメーターの型に応じて、カスタマイズされた視覚化を含んでもよい。例えば、ドアを引く動きのパラメーターを設定する場合、蝶番の軸がラベルされ、線として視覚

50

化され、グリッパを引くための姿勢候補は蝶番を中心とする円弧上に制限される。教示プロセスを支援するために、動作の復元、関連物体による環境への注釈付け、仮想ロボットの再位置決め、カメラ画像、及びVR世界のメニューのような幾つかのユーティリティVRモードが使用される。

【0107】

実行中、ロボットの姿勢及び環境中の部品は訓練中に使用されたものと異なってもよい。教示されたものと類似した環境中の特徴を発見するために特徴一致が使用されてもよい。姿勢デルタは一致した特徴の対応関係から確立されてもよい。ユーザーにより教示された動作が計算された姿勢デルタにより変化してもよい。ある構成では、複数のキーフレームがマッチング問題へ渡される。対応関係の数に基づいて、最も良く一致したキーフレームが選択される。

10

【0108】

図4は本開示の態様によるロボットシステム400のハードウェア実装の例を示す図である。ロボットシステム400は車両、ロボット装置428、又はその他の装置のような、自律又は半自律システムの構成要素であってもよい。図4の例では、ロボットシステム400はロボット装置428の構成要素である。ロボットシステム400は、ロボット装置428のオペレーターの意図を推察することによってロボット装置428のアクションを制御するために使用されてもよい。

20

【0109】

ロボットシステム400はバス430として概ね表されるバスアーキテクチャにより実装されてもよい。バス430はロボットシステム400の特定の用途及び全体的な設計制約に応じて任意の数の相互接続バス及びブリッジを含んでもよい。バス430は、プロセッサ420として表される1つ以上のプロセッサ及び/又はハードウェアモジュール、通信モジュール422、位置モジュール418、センサモジュール402、移動モジュール426、メモリ424、視覚化モジュール408、及びコンピュータ可読媒体414のような様々な回路を接続する。バス430はタイミングソース、周辺機器、電圧制御器、電源管理回路のような、当業者に既知であり、故にこれ以上の説明は行わないその他の様々な回路も接続してもよい。

20

【0110】

ロボットシステム400は、プロセッサ420に接続された送受信機416、センサモジュール402、視覚化モジュール408、通信モジュール422、位置モジュール418、移動モジュール426、メモリ424、及びコンピュータ可読媒体414を含む。送受信機416はアンテナ434に接続されている。送受信機416は様々な装置と伝送媒体を介して通信する。例えば、送受信機416は通信を介してロボット装置428のオペレーターからの命令を受信してもよい。本明細書中に議論するように、オペレーターはロボット装置428から離れた位置にいてもよい。別の例として、送受信機416はキーフレームマッチングに関する情報を視覚化モジュール408からオペレーターへ送信してもよい。

30

【0111】

ロボットシステム400はコンピュータ可読媒体414に接続されたプロセッサ420を含む。プロセッサ420は、コンピュータ可読媒体414に記憶され、本開示による機能を提供するソフトウェアの実行を含む処理を行う。ソフトウェアは、プロセッサ420により実行された時、ロボットシステム400により、ロボット装置428又はモジュール402、408、414、416、418、420、422、424、426のような特定の装置に対して説明された様々な機能を実行させる。コンピュータ可読媒体414は、ソフトウェアを実行した時にプロセッサ420により操作されるデータを記憶するためにも使用されてもよい。

40

【0112】

センサモジュール402は、第1のセンサ406及び第2のセンサ404のような異なるセンサを介して測定値を得るために使用されてもよい。第1のセンサ406は、2D画

50

像を撮影するためのステレオカメラ又はRGBカメラのような視覚センサであってもよい。第2のセンサ404はLiDARセンサ、RADARセンサ、又はRGB-Dセンサのような測距センサであってもよい。当然、本開示の態様は上記のセンサに限られず、例えば、温度、音波、及び/又はレーザー等のその他の型のセンサもまたセンサ404、406のどちらかとして考えられる。第1のセンサ406及び第2のセンサ404による測定値は、本明細書中に説明した機能を実装するために、コンピュータ可読媒体414と併せて、プロセッサ420、センサモジュール402、通信モジュール422、位置モジュール418、移動モジュール426、メモリ424、のうち1つ以上により処理されてもよい。ある構成では、第1のセンサ406及び第2のセンサ404により捕捉されたデータは送受信機416を介して供給ビデオとしてオペレーターへ送信されてもよい。第1のセンサ406及び第2のセンサ404はロボット装置428へ接続されていてもよく、又はロボット装置428と通信状態にあってもよい。

【0113】

位置モジュール418はロボット装置428の位置を判断するために使用されてもよい。例えば、位置モジュール418は、ロボット装置428の位置を判断するために、全地球測位システム(GPS)を使用してもよい。通信モジュール422は送受信機416を介した通信を促進するために使用されてもよい。例えば、通信モジュール422は、Wi-Fi、long term evolution(LTE)、3G等のような異なる無線プロトコルを介した通信能力を提供してもよい。通信モジュール422はまた、ロボットシステム400のモジュールではない、ロボット装置428のその他の構成要素と通信するために使用されてもよい。

【0114】

移動モジュール426は、ロボット装置428の、及び/又はロボット装置428の構成要素(例えば手足、手、等)の移動を促進するために使用されてもよい。例えば、移動モジュール426は手足438及び/又は車輪432の動きを制御してもよい。別の例として、移動モジュール426はエンジン又はバッテリーのようなロボット装置428の電力源と通信中であってもよい。＼

【0115】

ロボットシステム400はまた、ロボット装置428及び視覚化モジュール408の操作に関連するデータを記憶するためのメモリ424を含む。モジュールはプロセッサ420内で実行されるソフトウェアモジュール、コンピュータ可読媒体414及び/又はメモリ424に常駐/記憶されるもの、プロセッサ420に接続された1つ以上のハードウェアモジュール、又はそれらの組み合わせであってもよい。

【0116】

視覚化モジュール408はセンサモジュール402、送受信機416、プロセッサ420、通信モジュール422、位置モジュール418、移動モジュール426、メモリ424、及びコンピュータ可読媒体414と通信可能であってもよい。ある構成では、視覚化モジュール408はロボット装置が意図する経路を環境中に生成する。意図する経路はロボット操作センターのような遠隔地へ送信されてもよい。遠隔地にて意図する経路に対応する視覚化を表示してもよい。

【0117】

視覚化モジュール408はまた環境中の運転可能領域を識別する。運転可能領域は、過去に走行した経路に基づいてニューラルネットワークにより判断されてもよい。訓練、奥行情報、及びその他の情報に基づいて、環境中の領域は運転不可能として識別されてもよい。意図する経路はロボット操作センターのような遠隔地へ送信されてもよい。遠隔地は意図する経路に対応する視覚化を表示してもよい。

【0118】

ある構成では、視覚化モジュール408は環境中の第2の領域を運転可能として識別する情報を受信する。情報は遠隔地からの通信を介して受信されてもよい。遠隔地にいるユーザーが環境の画像中のピクセルを運転可能としてマークする。同様に、視覚化モジュー

10

20

30

40

50

ル 4 0 8 はある領域が運転不可能であることを識別する情報を受信してもよい。運転可能及び／又は運転不可能な領域に関する情報は、以前に視覚化モジュール 4 0 8 によって識別された領域を修正するために受信されてもよい。例えば、視覚化モジュール 4 0 8 はある領域が運転不可能であると間違ってマークし得る。受信情報により、間違ったマーキングが修正され得る。

【 0 1 1 9 】

視覚化モジュール 4 0 8 は、1つ以上の過去に走行した経路に基づいて意図する経路及び運転可能領域を判断するように訓練されてもよい。運転可能領域は、意図する経路に重ね合わせる形で表示されることが意図される。遠隔地からの更新を受信した後に、視覚化モジュール 4 0 8 は受信した更新に基づいて第 3 の領域を運転可能として識別する。

10

【 0 1 2 0 】

ある構成では、視覚化モジュール 4 0 8 は遠隔地にて現在の画像とキーフレーム画像の比較を表示するための情報を提供する。比較は、現在の画像とキーフレーム画像で一致したピクセルを識別する視覚化されたマークを含む。加えて、視覚化モジュール 4 0 8 は遠隔地にて現在の画像とキーフレームセット中の各キーフレームの間のマッチングの確信度を表示するための情報を提供してもよい。更に、視覚化モジュール 4 0 8 は遠隔地にて1つ以上の姿勢デルタ条件を1つ以上の閾値と比較するグラフを表示するための情報を提供してもよい。情報は奥行情報、クリーク情報、姿勢情報、R M S エラー情報、及び／又はその他の情報を含んでもよい。

【 0 1 2 1 】

図 5 は本開示の態様による教示された動作のグラフシーケンス 5 0 0 の例を示す。図 5 に示すように、グラフシーケンス 5 0 0 は開始ノード 5 0 2 と終了ノード 5 0 4 を含む。グラフシーケンス 5 0 0 はセンシングした視覚入力、音声入力、又はその他の条件に基づいて枝分かれ又はループしてもよい。

20

【 0 1 2 2 】

例えば、図 5 に示すように、開始ノード 5 0 2 の後、ロボットは「listen_for_object」の動作を実行してもよい。この例では、ロボットは、カップ又は瓶に対応する視覚又は音声入力をセンシングしたかどうかを判断する。この例では、センシングした入力がカップに対応するか又は瓶に対応するかどうかに基づいて異なる動作シーケンスが実行される。本開示の態様は図 5 に示した動作に限定されない。

30

【 0 1 2 3 】

図 6 は本開示の態様によるロボットシステムのためのソフトウェアモジュールの例を示す。図 6 のソフトウェアモジュールは、プロセッサ 4 2 0 、通信モジュール 4 2 2 、位置モジュール 4 1 8 、センサモジュール 4 0 2 、移動モジュール 4 2 6 、メモリ 4 2 4 、キーフレームモジュール 4 0 8 、及びコンピュータ可読媒体 4 1 4 のような図 4 のハードウェアシステムの1つ以上の構成要素を使用してもよい。本開示の態様は図 6 に示したモジュールに限定されない。

【 0 1 2 4 】

図 6 に示すように、ロボットは音声データ 6 0 4 及び／又は画像データ 6 0 2 を受信してもよい。画像データ 6 0 2 は R G B - D 画像であってもよい。音声ネットワーク 6 0 6 はあるインターバルでウェイクワードを聞いていてもよい。音声ネットワーク 6 0 6 はウェイクアードを検出し生のオーディオデータ 6 0 4 からキーワードを抽出するために生のオーディオデータ 6 0 4 を受信する。

40

【 0 1 2 5 】

密なエンベディングネットワーク 6 0 8 のようなニューラルネットワークは、画像データ 6 0 2 を受信する。画像データ 6 0 2 はあるインターバルにて受信されてもよい。密なエンベディングネットワーク 6 0 8 は画像データ 6 0 2 を処理し、画像データ 6 0 2 のエンベディング 6 1 0 を出力する。エンベディング 6 1 0 及び画像データ 6 0 2 はボクセルマップ 6 2 6 を生成するために組み合わされてもよい。エンベディング 6 1 0 もまたキーフレームマッチャ 6 1 2 に入力されてもよい。

50

【 0 1 2 6 】

キーフレームマッチャー 6 1 2 はエンベディング 6 1 0 を複数のキーフレームと比較する。エンベディング 6 1 0 がキーフレームのエンベディングと対応する場合に、一致するキーフレームが識別される。エンベディング 6 1 0 はピクセルディスクリプタ、奥行情報、及びその他の情報を含んでもよい。

【 0 1 2 7 】

タスクモジュール 6 1 4 は 1 つ以上のタスクグラフ 6 1 6 を受信してもよい。タスクモジュール 6 1 4 はキーフレームマッチャー 6 1 2 からの要求に対する応答を提供する。キーフレームマッチャー 6 1 2 は、一致したキーフレームへタスクを一致させる。タスクはタスクグラフ 6 1 6 から判断されてもよい。

10

【 0 1 2 8 】

タスクモジュール 6 1 4 はまた、動作モジュール(behavior module) 6 1 8 へ動作要求を送信してもよい。動作モジュール 6 1 8 はタスクモジュール 6 1 4 へ動作ステータスを提供する。加えて、動作モジュール 6 1 8 は一致したキーフレームに関する情報、及びキーフレームマッチャー 6 1 2 からの対応するタスクを要求してもよい。キーフレームマッチャー 6 1 2 は一致したキーフレームに関する情報、及び対応するタスクを動作モジュール 6 1 8 に提供する。動作モジュール 6 1 8 はボクセルマップ 6 2 6 からボクセルを受信してもよい。

【 0 1 2 9 】

ある構成では、動作モジュール 6 1 8 は、動作計画要求に応答して動作計画器 6 2 0 からの動作計画を受信する。動作モジュール 6 1 8 はまた、部品制御モジュール 6 2 2 より部品状況を受信する。動作モジュール 6 1 8 は、部品状況を受信したことに応答して部品制御モジュール 6 2 2 へ部品命令を送信する。最後に、部品制御モジュール 6 2 2 は関節制御モジュール 6 2 4 から関節状況を受信する。部品制御モジュール 6 2 2 は、関節状況を受信したことに応答して関節制御モジュール 6 2 4 へ関節命令を送信する。

20

【 0 1 3 0 】

図 7 は本開示の態様によるロボット装置により生成されたデータを可視化する方法 7 0 0 を示す。随意の構成では、ブロック 7 0 2 にて、ロボット装置は 1 つ以上の過去に走行した経路に基づいて意図する経路を判断するよう訓練される。過去に走行した経路は自律的に走行するか、及び / 又は人間の入力に基づいて走行する。

30

【 0 1 3 1 】

ブロック 7 0 4 にて、制御センターはロボット装置の意図する経路を環境中に表示する。環境は制御センターから遠隔であってもよい。加えて、環境はロボット装置を訓練するために使用した環境から異なっても、又は同じであってもよい。意図する経路はロボット装置により生成され、制御センターへ送信される。

【 0 1 3 2 】

ブロック 7 0 6 にて、制御センターはロボット装置にとって運転可能と識別された環境中の第 1 の領域を表示する。第 1 の領域は意図する経路に重ね合わせて表示されてもよい。加えて、意図する経路は第 1 の領域内に存在する。第 1 の領域はロボット装置により生成され、制御センターへ送信される。

40

【 0 1 3 3 】

ブロック 7 0 8 にて、制御センターは環境中の第 2 の領域を運転可能として識別するための入力を受信する。加えて、又は代わりに、入力により領域が運転不可能としてマークされてもよい。すなわち、ロボット装置によって判断された第 1 の領域が、入力により調整されてもよい。入力はタッチスクリーンディスプレイ、マウス、又はキーボードなどの入力装置により、人間のオペレーターにより提供されてもよい。入力は環境の画像中のピクセルを運転可能又は運転不可能としてマークすることを含んでもよい。

【 0 1 3 4 】

ブロック 7 1 0 にて、制御センターは第 2 の領域をロボット装置へ送信する。随意の構成では、第 2 の領域を受信したことに応答し、ブロック 7 1 2 にて、ロボット装置は第 3

50

の領域を運転可能と判断する。第3の領域は異なる環境にあってもよい。SVMの重みは第2の領域（例えばユーザーにより編集された情報）に基づいて更新される。SVMは教師あり学習モデルと称されてもよい。

【0135】

随意の構成では、ブロック714にて、制御センターは現在の画像とキーフレーム画像との比較（図3A及び図3B参照）を表示する。比較は、現在の画像とキーフレーム画像で一致したピクセルを識別する視覚化されたマークを含んでもよい。更に別の随意の構成では、ブロック716にて、制御センターは現在の画像と、キーフレームのセット中の各キーフレーム画像とのマッチングの確信度（図3A及び図3B参照）を表示する。

【0136】

上に説明した方法の様々な操作は、対応する機能を実行可能な任意の適切な手段により実行されてもよい。手段は回路、特定用途向け集積回路（ASIC）、又はプロセッサを限定せず含む、様々なハードウェア及び／又はソフトウェアコンポーネント及び／又はモジュールを含んでもよい。図に示す操作があるときは、これらの操作は、概ね類似した番号を割り振られた対応する機能的構成要素を有してもよい。

10

【0137】

本明細書中に使用されるように、「判断」は多岐にわたるアクションを含む。例えば、「判断」は算出、計算、処理、導出、調査、検索（例えば表、データベース又はその他の構造の中を検索）、究明等を含み得る。加えて、「判断」は受信（例えば情報を受信すること）、アクセス（例えばメモリ中のデータにアクセスすること）等を含み得る。更に、「判断」は、解決、選出、選択、確立等を含み得る。

20

【0138】

本明細書中に使用するように、「のうち少なくとも1つ」のフレーズは、項目のリストから、単一の項目を含む、項目の任意の組み合わせを指す。例えば、「a、b、又はcのうち少なくとも1つ」はa、b、c、a-b、a-c、b-c、a-b-cを含むことが意図される。

【0139】

本開示に関連して説明される様々な例示的な論理ブロック、モジュール、および回路は、本開示に従って構成されたプロセッサ、デジタル信号プロセッサ（DSP）、特定用途向け集積回路（ASIC）、Field Programmable Gate Array（FPGA）信号又は他のプログラマブル論理デバイス（PLD）、個別ゲート又はトランジスタ論理、個別ハードウェアコンポーネント、又は本明細書中に説明した機能を実行するよう設計された上記の任意の組み合わせにより実装又は実行されてもよい。プロセッサは、本明細書中の説明のように構成されるマイクロプロセッサ、コントローラ、マイクロコントローラ、又は状態マシンであってもよい。プロセッサはまた、例えばDSPとマイクロプロセッサとの組み合わせ、複数のマイクロプロセッサ、DSPコアと組み合わせた1つ以上のマイクロプロセッサ、又は本明細書で説明するその他の特別な構成といった、計算装置の組み合わせとして実装されてもよい。

30

【0140】

本開示と関連して説明される方法のステップ又はアルゴリズムはハードウェア、プロセッサにより実行されるソフトウェアモジュール、又はこの2つの組み合わせ内にて直接具現化されてもよい。ソフトウェアモジュールは、ランダムアクセスメモリ（RAM）、リードオンリーメモリ（ROM）、フラッシュメモリ、Erasable Programmable Read Only Memory（EPROM）、Electrically Erasable Programmable Read-Only Memory（EEPROM）、レジスタ、ハードディスク、取り外し可能ディスク、CD-ROM又は他の光ディスク記憶装置、磁気ディスク記憶装置又はその他の磁気記憶装置、命令又はデータ構造の形態で所望のプログラムコードを搬送又は記憶することに使用可能でコンピュータによってアクセス可能なその他の任意の媒体を含む、記憶装置、又は機可読械媒体に存在してもよい。ソフトウェアモジュールは、单一の命令、又は多数の命令を備えることもあ

40

50

り、複数の異なるコードセグメント、異なるプログラム間、及び複数の記憶媒体に分散されていてもよい。プロセッサが記憶媒体へ情報を書き込んだり記憶媒体から情報を読み出したりできるようにプロセッサに記憶媒体が接続されていてもよい。代わりに、記憶媒体はプロセッサと一体となっていてもよい。

【0141】

本明細書にて開示する方法は、開示された方法を実現するための1つ以上のステップ又はアクションを含む。方法のステップ及び／又はアクションは特許請求の範囲から逸脱することなく互いに入れ替えてよい。言い換えれば、ステップ又はアクションの特定の順序が特定されていない限り、特定のステップ及び／又はアクションの順序及び／又は用途は特許請求の範囲から逸脱することなく変更されてもよい。

10

【0142】

説明された機能はハードウェア、ソフトウェア、ファームウェア、又はこれらの任意の組み合わせにより実装されてもよい。ハードウェアで実装する場合、ハードウェア構成の例は装置中に処理システムを備えてよい。処理システムはバスアーキテクチャを用いて実装してもよい。バスは処理システムの特定の用途及び全体的な設計制約に応じて任意の数の相互接続バス及びブリッジを含んでもよい。バスはプロセッサ、機械可読媒体、及びバスインターフェースを含む様々な回路を接続してもよい。バスインターフェースは、その他の物の中でも、ネットワークアダプタを処理システムにバスを介して接続することに使用されてもよい。ネットワークアダプタは信号処理機能を実装するために使用されてもよい。特定の態様では、ユーザーインターフェース（例えばキーパッド、ディスプレイ、マウス、ジョイスティック等）もまたバスに接続されてもよい。バスはタイミングソース、周辺機器、電圧制御、電源管理回路等のような、当業者に既知であり、故にこれ以上の説明は行わないその他の様々な回路も接続してもよい。

20

【0143】

プロセッサはバスの管理、及び機械可読媒体に記憶されたソフトウェアの実行を含む処理を担当してもよい。ソフトウェアは、ソフトウェア、ファームウェア、ミドルウェア、マイクロコード、ハードウェア記述言語、又はその他の呼び方に関わらず、命令、データ、又はそれらの任意の組み合わせを意味すると解釈するものとする。

【0144】

ハードウェア実装にて、機械可読媒体はプロセッサとは別個の処理システムの一部であつてもよい。しかしながら、当業者が容易に理解するように、機械可読媒体、又はその任意の一部は、処理システムの外部にあってもよい。例えば、機械可読媒体は通信線、データによって変調された搬送波、及び／又は装置から切り離されたコンピュータ製品を含んでも良く、これらは全てバスインターフェースを介してプロセッサによりアクセスされてもよい。代わりに、又は加えて、機械可読媒体、又はその一部は、キャッシュ及び／又は特別なレジスタファイルが存在し得る場合のようにプロセッサに統合されていてもよい。議論した様々な構成要素は、ローカルな構成要素のように特別な位置を有するように説明されたが、それらは分散コンピューティングシステムの一部として構成される特定の構成要素のように様々な方法により構成されてもよい。

30

【0145】

処理システムはプロセッサ機能を提供する1つ以上のマイクロプロセッサ、及び機械可読媒体の少なくとも一部を提供する及び外部メモリにより構成されてもよく、それらは全て外部のバスアーキテクチャによりサポート回路を通して連結されてもよい。代わりに、処理システムは、本明細書にて説明したニューロンモデル及びニューラルシステムのモデルを実装するために、1つ以上のニューロモーフィックプロセッサを備えていてもよい。別の代替として、処理システムは、プロセッサを有する特定用途向け集積回路（ASIC）、バスインターフェース、ユーザーインターフェース、サポート回路、及び単一のチップに統合された機械可読媒体の少なくとも一部、又は1つ以上のField Programmable Gate Array (FPGA)、プログラマブル論理デバイス（PLD）、コントローラ、状態マシン、ゲート論理、個別ハードウェア構成要素、又はその他

40

50

の適切な回路、又は本開示内で説明された様々な機能を実行可能な回路の任意の組み合わせにより実装されてもよい。当業者は、特定の用途、及びシステム全体に課される全体的な設計制約により、説明された処理システムの機能をどのように実装することが最適なのかを認識するであろう。

【 0 1 4 6 】

機械可読媒体は数々のソフトウェアモジュールを備えていてもよい。ソフトウェアモジュールは送信モジュール及び受信モジュールを含んでいてもよい。各ソフトウェアモジュールは単一の記憶装置内に存在してもよく、又は複数の記憶装置に渡って分散されていてもよい。例えば、トリガとなるイベントが起こったときにソフトウェアモジュールがハードドライブから R A M にロードされてもよい。ソフトウェアモジュールの実行中、プロセッサは、アクセス速度を上げるために、幾つかの命令をキャッシングヘロードしてもよい。プロセッサにより実行するため、1つ以上のキャッシングラインがその後特殊用途レジスタファイルにロードされてもよい。ソフトウェアモジュールの以下の機能を参照すれば、ソフトウェアモジュールによる命令の実行時にプロセッサにより機能が実施されることが理解されよう。更に、本開示の態様によりプロセッサ、コンピュータ、マシン、又はこのような態様を実装するその他のシステムの機能が改善することが理解されるべきである。

【 0 1 4 7 】

ソフトウェアに実装されれば、機能は1つ以上の命令又はコードとしてコンピュータ可読媒体上に記憶又は転送されてもよい。コンピュータ可読媒体には、コンピュータの記憶装置と、コンピュータプログラムがある場所から別の場所へ転送することを促進する任意の記憶装置を含む通信メディアとの両方が含まれる。加えて、任意の接続をコンピュータ可読媒体と呼ぶのが適切である。例えば、ソフトウェアが同軸ケーブル、光ファイバケーブル、ツイストペア、デジタル加入者線 (D S L)、又は赤外線 (I R)、ラジオ、及びマイクロ波などの無線技術を使用してウェブサイト、サーバー、又は他の遠隔ソースから送信される場合は、同軸ケーブル、光ファイバケーブル、ツイストペア、 D S L 、又は赤外線、ラジオ、及びマイクロ波などの無線技術はメディアの定義に含まれる。本明細書に使用される d i s k 及び d i s c は、コンパクトディスク (C D)、レーザーディスク (登録商標)、光ディスク、デジタル多用途ディスク (D V D)、フロッピー (登録商標) ディスク、ブルーレイ (登録商標) ディスクを含み、 d i s k は通常磁気的にデータを再生し、 d i s c はレーザーを使用して光学的にデータを再生する。従って、幾つかの態様では、コンピュータ可読媒体は非一時的なコンピュータ可読媒体 (例えは有形媒体) を備えていてもよい。加えて、その他の態様では、コンピュータ可読媒体は一時的なコンピュータ可読媒体 (例えは信号) を備えていてもよい。上記の組み合わせはコンピュータ可読媒体の範囲に含まれるべきものとする。

【 0 1 4 8 】

従って、特定の態様は本明細書中に提示された操作を実行するコンピュータプログラム製品を備えていてもよい。例えば、そのようなコンピュータプログラム製品は命令を記憶 (及び / 又は暗号化) したコンピュータ可読媒体を備えていてもよく、命令は本明細書中に説明した操作を実行するために1つ以上のプロセッサにより実行可能である。特定の態様では、コンピュータプログラム製品は包装資材を含んでもよい。

【 0 1 4 9 】

更に、モジュール並びに / 又は本明細書中に説明した方法及び技術を実行するその他の適切な手段は、必要に応じてダウンロード並びに / 又はユーザー端末及び / 又はベースステーションにより取得可能であることが理解されるべきである。例えば、本明細書中に説明された方法を実行するための手段の転送を促進するために、そのような装置をサーバーに接続することができる。代わりに、本明細書中に説明した様々な方法は、記憶手段を装置に接続するか、又は記憶手段を装置に提供することによりユーザー端末及び / 又はベースステーションが様々な方法を取得することができる形で、記憶手段を介して提供することができる。更に、本明細書中に説明した方法及び技術を装置に提供するその他の任意の技術を使用することができる。

10

20

30

40

50

【 0 1 5 0 】

特許請求の範囲は上記に示された正確な構成及び構成要素に限定されないことが理解されるべきである。上に説明した方法及び装置の配置、操作、並びに詳細に対して、特許請求の範囲から逸脱せずに、様々な修正、変更及び変形がなされ得る。

本明細書に開示される発明は以下の態様を含む。

〔態様 1〕

ロボット装置により生成されたデータを可視化する方法であって、

前記ロボット装置の意図する経路を環境中に表示することと、

前記ロボット装置により運転可能として識別された前記環境中の第 1 の領域を表示することと、

前記環境中の第 2 の領域を運転可能として識別するための入力を受信することと、

前記第 2 の領域を前記ロボット装置へ送信することと、

を含む、方法。

10

〔態様 2〕

過去に走行した経路に基づいて前記意図する経路を判断するように前記ロボット装置を訓練することを更に含む、態様 1 に記載の方法。

〔態様 3〕

前記意図する経路に重ね合わせて前記第 1 の領域を表示することを更に含む、態様 1 に記載の方法。

〔態様 4〕

前記入力は前記第 2 の領域に対応するピクセルにマークを付することを含む、態様 1 に記載の方法。

20

〔態様 5〕

前記ロボット装置が前記受信した第 2 の領域に基づいて第 3 の領域を運転可能として判断する、態様 1 に記載の方法。

〔態様 6〕

現在の画像とキーフレーム画像との比較を表示することを更に含み、前記比較は前記現在の画像と前記キーフレーム画像との一致ピクセルを識別する可視化されたマークを有する、態様 1 に記載の方法。

〔態様 7〕

現在の画像とキーフレームのセットの中の各キーフレームとの間のマッチングの確信度を表示することを更に含む、態様 1 に記載の方法。

30

〔態様 8〕

前記意図する経路及び前記第 1 の領域を前記ロボット装置から遠方の位置にて表示する、態様 1 に記載の方法。

〔態様 9〕

メモリと、

前記メモリに接続された少なくとも 1 つのプロセッサを備え、前記少なくとも 1 つのプロセッサは、

ロボット装置の意図する経路を環境中に表示し、

前記ロボット装置により運転可能と識別された前記環境中の第 1 の領域を表示し、

前記環境中の第 2 の領域を運転可能として識別するための入力を受信し、

前記第 2 の領域を前記ロボット装置へ送信する、

よう構成されている、ロボット装置により生成されたデータを可視化する装置。

40

〔態様 10〕

前記少なくとも 1 つのプロセッサは更に、過去に走行した経路に基づいて前記意図する経路を判断するよう前記ロボット装置を訓練するよう構成されている、態様 9 に記載の装置。

〔態様 11〕

前記少なくとも 1 つのプロセッサは更に、前記意図する経路に重ね合わせて前記第 1 の領

50

域を表示するように構成されている、態様 9 に記載の装置。

[態様 1 2]

前記入力は前記第 2 の領域に対応するピクセルにマークを付することを含む、態様 9 に記載の装置。

[態様 1 3]

前記ロボット装置が前記受信した第 2 の領域に基づいて第 3 の領域を運転可能として判断する、態様 9 に記載の装置。

[態様 1 4]

前記少なくとも 1 つのプロセッサは更に、現在の画像とキーフレーム画像との比較を表示するように構成され、前記比較は前記現在の画像と前記キーフレーム画像との一致ピクセルを識別する可視化されたマークを含む、態様 9 に記載の装置。

10

[態様 1 5]

前記少なくとも 1 つのプロセッサは更に、現在の画像とキーフレームのセットの中の各キーフレームとの間のマッチング確信度を表示するように構成されている、態様 9 に記載の装置。

[態様 1 6]

前記意図する経路及び前記第 1 の領域を前記ロボット装置から遠方の位置にて表示する、態様 9 に記載の装置。

[態様 1 7]

ロボット装置により生成されたデータを可視化するためのプログラムコードを記録した非一時的なコンピュータ可読媒体であって、

20

前記プログラムコードはプロセッサにより実行され、

前記ロボット装置の意図する経路を環境中に表示するプログラムコードと、

前記ロボット装置により運転可能として識別された前記環境中の第 1 の領域を表示するプログラムコードと、

前記環境中の第 2 の領域を運転可能として識別するための入力を受信するプログラムコードと、

前記第 2 の領域を前記ロボット装置へ送信するプログラムコードと、
を含む、非一時的なコンピュータ可読媒体。

30

[態様 1 8]

前記プログラムコードは更に、過去に走行した経路に基づいて前記意図する経路を判断するよう前記ロボット装置を訓練するプログラムコードを有する、態様 1 7 に記載の非一時的なコンピュータ可読媒体。

[態様 1 9]

前記プログラムコードは更に、前記意図する経路に重ね合わせて前記第 1 の領域を表示するプログラムコードを有する、態様 1 7 に記載の非一時的なコンピュータ可読媒体。

[態様 2 0]

前記入力は前記第 2 の領域に対応するピクセルにマークを付することを含む、態様 1 7 に記載の非一時的なコンピュータ可読媒体。

40

50

【図面】

【図1】

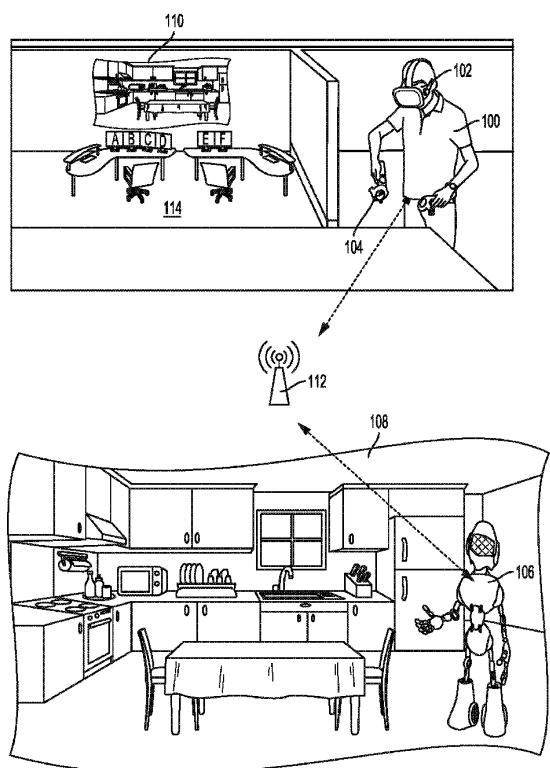


FIG. 1

【図2A】

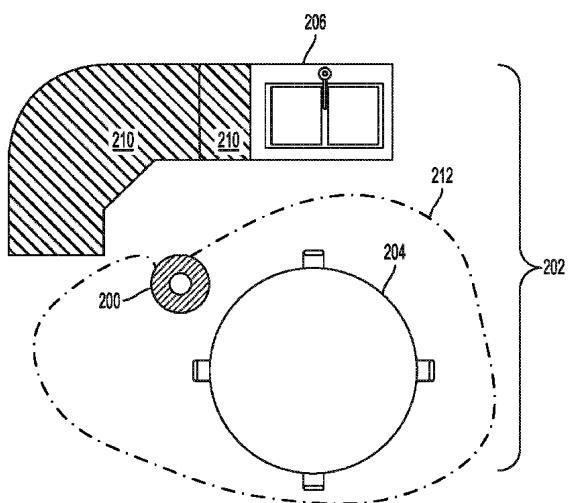


FIG. 2A

10

20

【図2B】

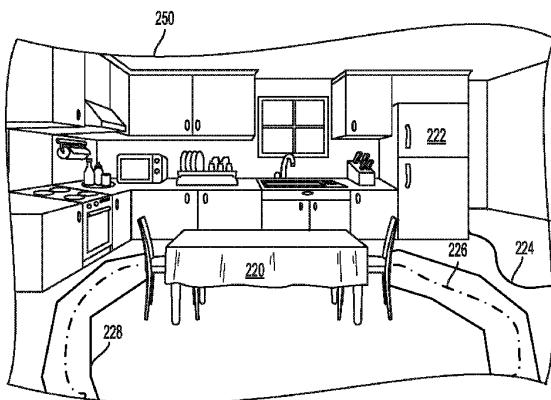
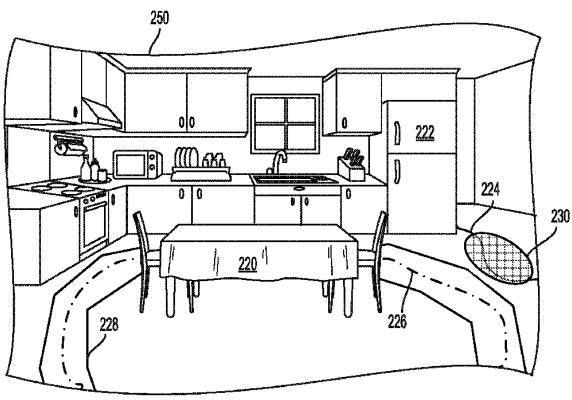


FIG. 2B

【図2C】



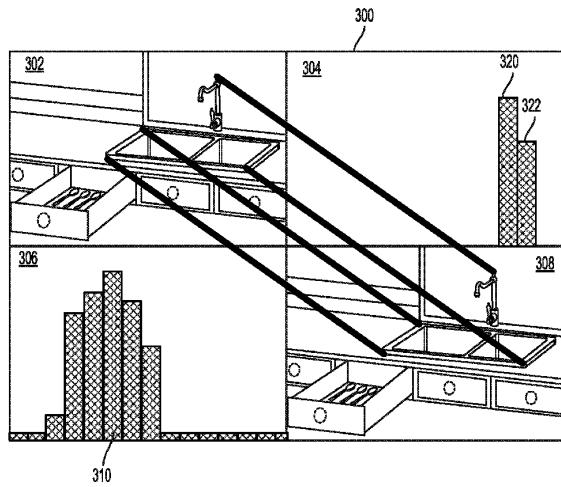
30

40

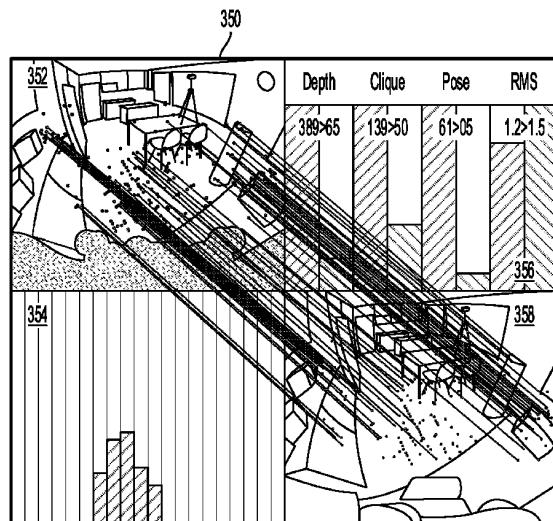
FIG. 2C

50

【図 3 A】



【図 3 B】



10

FIG. 3B

FIG. 3A

20

【図 4】

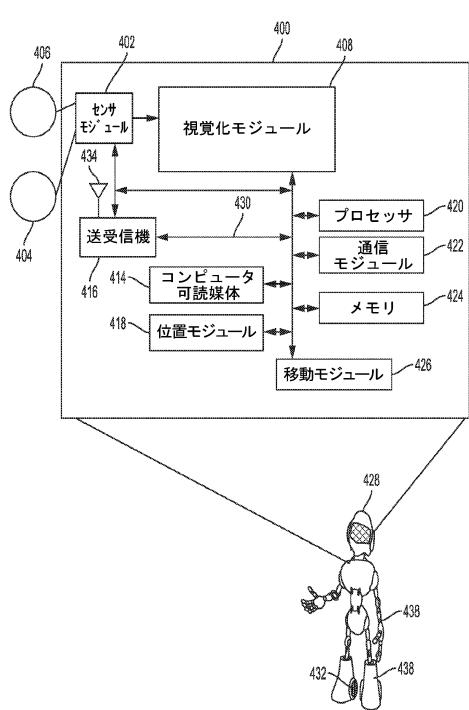
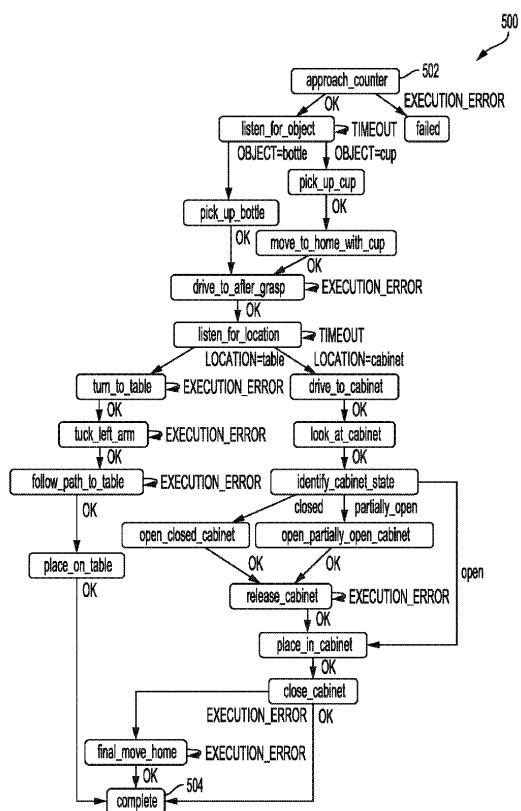


FIG. 4

【図 5】



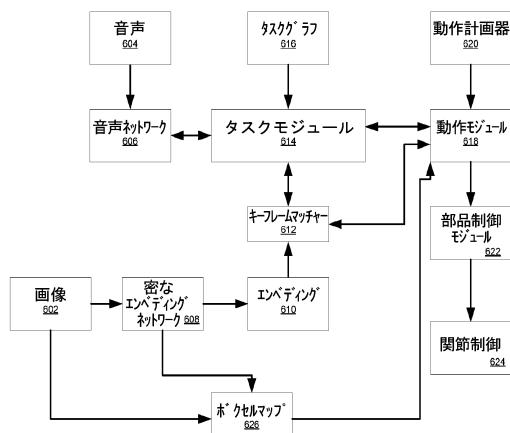
30

40

FIG. 5

50

【図6】



【図7】

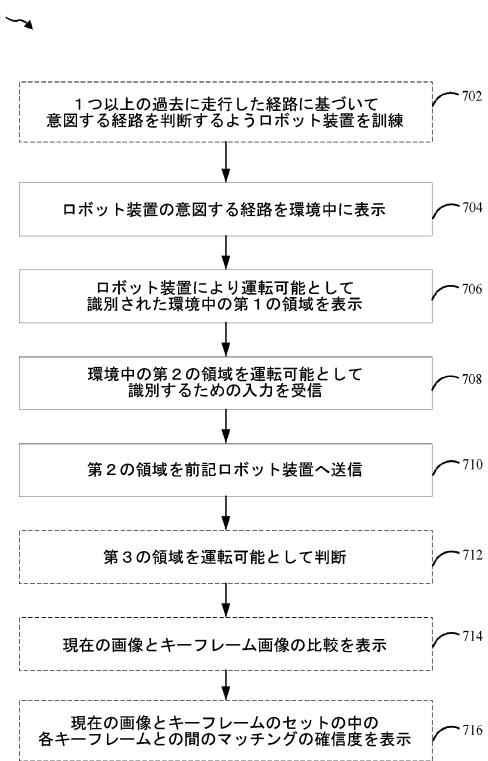


FIG. 6

FIG. 7

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 62/877,793

(32) 優先日 令和1年7月23日(2019.7.23)

(33) 優先権主張国・地域又は機関

米国(US)

(31) 優先権主張番号 16/570,540

(32) 優先日 令和1年9月13日(2019.9.13)

(33) 優先権主張国・地域又は機関

米国(US)

弁理士 関根 宣夫

(74) 代理人 100133835

弁理士 河野 努

(72) 発明者 ジェレミー マー

アメリカ合衆国, カリフォルニア 94022, ロスアルトス, エル カミノ リアル 4440

審査官 渡邊 捷太郎

(56) 参考文献 特開2019-106060 (JP, A)

特開2014-085829 (JP, A)

国際公開第2019/138640 (WO, A1)

特開2017-033399 (JP, A)

国際公開第2018/158248 (WO, A1)

特開2010-152835 (JP, A)

米国特許出願公開第2019/0015985 (US, A1)

特開2019-121364 (JP, A)

特表2019-515407 (JP, A)

米国特許出願公開第2016/0165795 (US, A1)

(58) 調査した分野 (Int.Cl., DB名)

G05D 1/00

G06F 3/01

G06F 3/0481

G06T 19/00