

(19)日本国特許庁(JP)

(12)公開特許公報(A)

(11)公開番号

特開2023-54776

(P2023-54776A)

(43)公開日 令和5年4月14日(2023.4.14)

(51)国際特許分類

G 0 6 Q 50/08 (2012.01)

F I

G 0 6 Q 50/08

テーマコード(参考)

5 L 0 4 9

審査請求 未請求 請求項の数 10 O L 外国語出願 (全18頁)

(21)出願番号 特願2022-159232(P2022-159232)
 (22)出願日 令和4年10月3日(2022.10.3)
 (31)優先権主張番号 10 2021 211 152.1
 (32)優先日 令和3年10月4日(2021.10.4)
 (33)優先権主張国・地域又は機関
 ドイツ(DE)

(71)出願人 390023711
 ローベルト ボツシュ ゲゼルシャフト
 ミット ベシユレンクテル ハフツング
 ROBERT BOSCH GMBH
 ドイツ連邦共和国 シュツツガルト (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

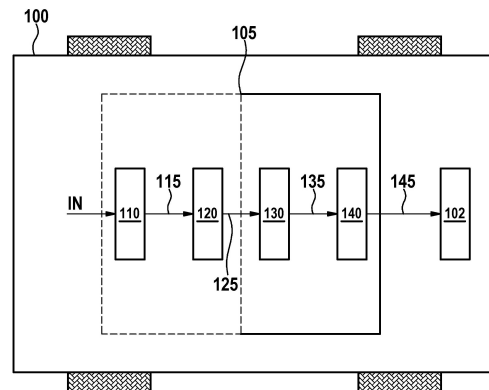
(54)【発明の名称】 車両を制御するように構成されたストラテジを生成するための強化学習用エージェントに対してトレーニングデータを提供するための装置及びコンピュータ実装された方法、車両

(57)【要約】 (修正有)

【課題】車両を制御するストラテジを生成する強化学習用エージェントに対してトレーニングデータを提供するための装置、方法及びプログラムを提供する。

【解決手段】車両100において、装置105の初期化ユニット110は、粒状材料と作業環境との選択された物理的特徴値INを使用して、環境モデル115を作成する。選択された物理的特徴値INは、初期化ユニットによって装置105の外部のデータソースから入力される。対応付けユニット120は、確率密度関数に基づく相対的な高さマップを含む環境モデルを、初期化ユニットから受信又はロードする。強化学習用エージェント130は、トレーニングデータ125を、対応付けユニット120から受信する。任意選択肢の制御ユニット140は、車両を制御するためのストラテジ135を強化学習用エージェントから受信し、ストラテジに基づいた制御信号145を少なくとも1つのアクチュエータ102に出力する。

【選択図】図1



10

【特許請求の範囲】

【請求項 1】

車両（100）、特にオフロード車両（100）を制御するために適したストラテジ（135）を生成するための強化学習用エージェント（130）に対してトレーニングデータ（125）を提供するためのコンピュータ実装された方法（200）であって、前記車両（100）は、作業環境内の粒状材料と相互作用するように構成されており、

当該方法（200）は、

前記粒状材料を含む前記作業環境の環境モデル（115）を初期化するステップ（210）であって、前記粒状材料と前記作業環境との選択された物理的特徴値（IN）が、少なくとも1つの確率密度関数における変数として使用され、前記粒状材料を含む前記作業環境の実際高さ状態と目標高さ状態との間の相対的な高さマップが、前記確率密度関数を使用して生成される、ステップ（210）と、

目標状態を達成するために、前記環境モデル（115）の前記確率密度関数の変化と報酬値とを、前記車両（100）の複数のアクションの集合のうちのそれぞれのアクションに対応付けるステップ（220）であって、前記トレーニングデータ（125）を提供するために、前記環境モデル（115）の前記相対的な高さマップと車両位置とが、それぞれのアクションごとに更新される、ステップ（220）と、
を含む、方法（200）。

【請求項 2】

前記初期化するステップ（210）において、

前記確率密度関数は、連続多変量ガウス確率密度関数であり、及び/又は、

前記選択された物理的特徴値は、前記粒状材料の比質量と、前記粒状材料の拡布パラメータと、前記粒状材料のそれぞれの山の体積と、環境に対するそれぞれの山の中心点の位置とを含む、

請求項 1 に記載の方法（200）。

【請求項 3】

前記対応付けるステップ（220）において、前記作業環境が複数の境界枠に分割され、

前記確率密度関数の前記変化は、それぞれの境界枠ごとに個別に計算される、

請求項 1 又は 2 に記載の方法（200）。

【請求項 4】

前記対応付けるステップ（220）において、アクションが前進並進であるか、後退並進であるか又は回転であるかに応じて、及び/又は、前記車両位置が前記作業環境の許容領域内に位置しているか又は許容領域外に位置しているかに応じて、前記アクションにそれぞれ異なる変化及び/又はそれぞれ異なる報酬値が対応付けられる、

請求項 1 乃至 3 のいずれか一項に記載の方法（200）。

【請求項 5】

車両（100）、特にオフロード車両（100）を制御するように構成されたストラテジ（135）を生成するための方法（300）であって、前記車両（100）は、作業環境内の粒状材料と相互作用するように構成されており、

当該方法（300）は、

請求項 1 乃至 4 のいずれか一項に記載の方法（200）によるトレーニングデータ（125）を提供するステップ（325）と、

前記トレーニングデータ（125）を使用して、強化学習用エージェント（130）によって前記ストラテジ（135）を生成するステップ（330）と、

を含む、方法（300）。

【請求項 6】

当該方法（300）は、前記ストラテジ（135）によって決定された制御信号（145）を用いて、前記車両（100）の少なくとも1つのアクチュエータ（102）を制御するステップ（340）を含む、

10

20

30

40

50

請求項 5 に記載の方法 (3 0 0) 。

【請求項 7】

請求項 1 乃至 6 のいずれか一項に記載の方法 (2 0 0 , 3 0 0) のステップを、対応するユニット (1 1 0 , 1 2 0 , 1 3 0 , 1 4 0) において実施及び / 又は制御するように構成されている装置 (1 0 5) 。

【請求項 8】

作業環境内の粒状材料と相互作用するように構成されている車両 (1 0 0) 、特にオフロード車両 (1 0 0) であって、

当該車両 (1 0 0) は、請求項 7 に記載の装置 (1 0 5) を含む、
車両 (1 0 0) 。

10

【請求項 9】

請求項 1 乃至 6 のいずれか一項に記載の方法 (2 0 0 , 3 0 0) のステップを実施及び / 又は制御するために構成されているコンピュータプログラム。

【請求項 1 0】

請求項 9 に記載のコンピュータプログラムが保存されている機械可読記憶媒体。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

本発明は、独立請求項に記載の装置又は方法に関する。コンピュータプログラムも、本発明の対象である。

20

【背景技術】

【0 0 0 2】

自律車両との関連において、粒状シミュレーション又は粒状材料のシミュレーションは、一般的に 3 種類のシミュレーションに分類することができ、例えば、固体力学及び経験式と、離散要素法と、高さマップアプローチとに分類することができる。特に、上記の方法は、粒状材料と車両との間の相互作用を特定するために反復アルゴリズムを使用する。さらに、これらのシミュレーションをオフラインで計算し、続いて、数値流体力学 (C F D) のような他の数値的な方法の場合と同様に、実行して結果を観察することができるが、このような方法は、一般的に、強化学習 (B L 、英語では reinforcement learning) にとって不十分である。なぜなら、シミュレーションは、B L エージェントをトレーニングするために複数回実行されなければならないからである。米国特許第 1 0 4 8 1 6 0 3 号明細書 (US10481603B2) は、オフロード車両のための移動軌道計画アルゴリズムを開示している。

30

【0 0 0 3】

独国特許出願公開第 1 9 8 5 9 1 6 9 号明細書 (DE19859169A1) からは、ジョイントで結合された複数の機械的なリンクを有する機械からなるシステムが公知である。処理システムは、少なくとも 1 つのスクリプトを実行する。それぞれのスクリプトは、機械の動きを規定する少なくとも 1 つの変パラメータを有する。学習アルゴリズムは、所望の結果と、結果に影響を与える測定された条件とに基づいて少なくとも 1 つの変パラメータの値を修正する。

40

【0 0 0 4】

独国特許出願公開第 1 0 2 0 1 4 2 0 3 3 1 2 号明細書 (DE102014203312A1) からは、道路の勾配を領域にわたってモデル化することができることが公知である。車両の運転中に所定のレートで勾配値のシーケンスが生成され、それぞれの勾配値は、その後に到達されるそれぞれの勾配領域を識別する。勾配値のシーケンスに対する応答として、マルコフ連鎖道路勾配モデルが更新され、このモデルは、それぞれの確率要素を、それぞれの所定の勾配領域からそれぞれの後続する勾配領域への遷移事象の行列の形態で表す。

【0 0 0 5】

独国特許出願公開第 1 1 2 0 0 9 0 0 1 5 5 4 号明細書翻訳文 (DE112009001554 T5) からは、工事現場での材料の山の上で車両を動作させるための方法が公知であり、

50

材料は、工事現場において開口部を介して排出され、同方法は、山の表面を検出することと、材料の排出に起因する山の表面上の障害ゾーンを、検出された表面と、開口部の既知の位置とに基づいて識別することと、障害ゾーンを知らせる信号を車両に伝送することとを含む。

【先行技術文献】

【特許文献】

【0006】

【特許文献1】米国特許第10481603号明細書(US10481603B2)

【特許文献2】独国特許出願公開第19859169号明細書(DE19859169A1)

【特許文献3】独国特許出願公開第102014203312号明細書(DE102014203312A1) 10

【特許文献4】独国特許出願公開第112009001554号明細書翻訳文(DE112009001554T5)

【発明の概要】

【発明が解決しようとする課題】

【0007】

発明の開示

このような背景を踏まえて、本明細書において提示されるアプローチは、それぞれ主請求項に記載の方法と、当該方法を使用する装置と、対応するコンピュータプログラムとを提供する。独立請求項において規定された対象の有利な実施形態及び改善形態は、従属請求項に記載された特徴によって可能となる。 20

【課題を解決するための手段】

【0008】

実施形態によれば、例えば、強化学習のトレーニングのための簡略化された粒状シミュレーションを提供することができる。シミュレーションは、例えばガウス分布に基づくことができ、この場合、粒状材料が、連続多変量ガウス確率密度関数(PDF)等のように挙動することを前提とすることができる。シミュレーションのために、粒状材料の物理学と数学的ガウス係数との間に、利用可能な関係が見出された。さらに、粒状材料と車両との間の相互作用を、ガウスPDFの変化として形式化することができる。特に、環境を高さマップによって表すことができ、ここで、 $H(x, y) = h$ は、位置 (x, y) における地面高さが h と同じであることを意味する。さらに、車両の位置は、車両の現在地と地軸に対する方位角との組合せとして仮定することができる。このシミュレーションは、例えば、2つの主要なアルゴリズムを含み得るものであり、すなわち、一方は、領域又は環境を初期化するためのアルゴリズムであって、この場合、粒状材料の山の位置のリストに基づいて、初期時の環境高さマップを生成することができ、そして他方は、更新された領域を計算するためのさらなるアルゴリズムであって、この場合、車両の移動又はアクションに基づいて新たな環境を特定することができる。 30

【0009】

実施形態によれば、例えば、特にオフロードでの移動軌道計画アルゴリズムのために使用することができる簡略化されたシミュレーションを提供することができる。特に、本明細書において提示されるアプローチは、複雑な物理学を検出すると同時に、短時間しかかからない簡単な自動シミュレータを維持するような、このような形式のシミュレーションを作成することが可能である。シミュレーション又はモデルは、強化学習用アルゴリズム(BLアルゴリズム)をトレーニングするために使用可能である。したがって、シミュレーションが、以下のもの、すなわち `reset`, `step`, `is_done`, `make_reward_calculate`を含むことが求められている `Gym-API` を改良することが提案される。特に、BLエージェントをトレーニングするためにシミュレーションを複数回実施しなければならないBLの場合には、従前の方法は、不十分である可能性がある。このような理由から、トレーニングにおいて使用することができ、かつ、例えば工事現場での1作業日の作業を最短時間でシミュレートすることができるような、簡略化 50

されたシミュレータを提供することができる。

【0010】

本明細書において提示されるアプローチによるシミュレーションは、それぞれの車両移動ごとの多重計算を省略することができるので、これまで挙げたアプローチよりも有利であろう。さらに、簡単な分布によってシミュレーションの複雑さが増加する可能性がある。したがって、例えば、車両の移動がPの確率でこぼれをもたらすという確率関数を追加することにより、車両移動にこぼれを追加することが可能である。本明細書において提示されるアプローチは、車両と環境との間の最も重要な相互作用特徴を、その際に簡略性を失うことなくシミュレートすることができる。さらに、このような形式のシミュレーションは、有利には、BLEージェントをトレーニングするために使用可能である。というのも、車両と地面との間のすべての相互作用が、BLEージェントの成功にとって重要であるとは限らないが、このような形式のシミュレーションは、最適なアクションを選択するためにBLトレーニングにとって必要である最も重要な相互作用を検出するからである。粒状材料の特徴を、有利には、ガウスPDFの作成のために容易に使用することができる数学的な変数に変換することも可能である。

10

【0011】

車両、特にオフロード車両を制御するように構成されたストラテジを生成するための強化学習用エージェントに対してトレーニングデータを提供するためのコンピュータ実装された方法であって、車両は、作業環境内の粒状材料と相互作用するように構成されており、

20

当該方法は、

粒状材料を含む作業環境の環境モデルを初期化するステップであって、粒状材料と作業環境との選択された物理的特徴値が、少なくとも1つの確率密度関数の変数として使用され、粒状材料を含む作業環境の実際高さ状態と目標高さ状態との間の相対的な高さマップが、確率密度関数を使用して生成される、ステップと、

目標状況を達成するために、環境モデルの確率密度関数の変化と報酬値とを、車両の複数のアクションの集合のうちのそれぞれのアクションに対応付けるステップであって、トレーニングデータを提供するために、環境モデルの相対的な高さマップと車両位置とが、それぞれのアクションごとに更新される、ステップと、
を含む、方法が提案される。

30

【0012】

車両又はオフロード車両は、ブルドーザ、平土機、転圧機、ダンプカー、掘削機、建設車両、重機、重量物用車、又は、環境との相互作用、特に粒状材料との相互作用が含まれる複数のタスクを有する他の種類の車両であってよい。このような種類のタスク又はアクションの例は、敷均し、粒状材料の荷降ろし、地面の転圧、粒状材料の除去等である。粒状材料は、土、砂、雪、瓦礫、石、及び/又は、類似の粒子状の材料を含み得る。ストラテジは、車両のための少なくとも1つの計画された移動軌道を含み得る。環境モデルは、確率密度関数に基づいた相対的な高さのマップを含み得る。報酬値は、正又は負又はゼロであってよい。確率密度関数の変化は、ゼロとは異なる絶対値を有し得るものであり、又は、ゼロに等しくてもよい。

40

【0013】

1つの実施形態によれば、初期化するステップにおいて、確率密度関数は、連続多変量ガウス確率密度関数であってよい。追加的又は代替的に、初期化するステップにおいて、選択された物理的特徴値は、粒状材料の比質量と、粒状材料の拡布パラメータと、粒状材料のそれぞれの山の体積と、環境に対するそれぞれの山の中心点の位置とを含み得る。このような実施形態は、粒状材料の複雑な物理学を検出できると同時に、短時間しかかからない簡単な自動シミュレータを提供することができるという利点を提供する。

【0014】

1つの実施形態によれば、対応付けるステップにおいて、作業環境を複数の境界枠に分割することができ、確率密度関数の変化をそれぞれの境界枠ごとに個別に計算することが

50

できる。このような実施形態は、車両のアクションに起因する環境の変化を、比較的わずかな計算コストで迅速に特定することができるという利点を提供する。

【0015】

1つの実施形態によれば、対応付けるステップにおいて、アクションが前進並進であるか、後退並進であるか又は回転であるかに応じて、追加的又は代替的に、車両位置が作業環境の許容領域内に位置しているか又は許容領域外に位置しているかに応じて、アクションにそれぞれ異なる変化、及び、追加的又は代替的にそれぞれ異なる報酬値を対応付けることができる。このような実施形態は、トレーニングエージェントにとって車両のアクションが容易に評価可能及び/又は実施可能となるという利点を提供する。

【0016】

本明細書においては、車両、特にオフロード車両を制御するように構成されたストラテジを生成するための方法であって、車両は、作業環境内の粒状材料と相互作用するように構成されており、当該方法は、

本明細書において提示される提供するための方法によるトレーニングデータを提供するステップと、

トレーニングデータを使用して、強化学習用エージェントによってストラテジを生成するステップと、
を含む、方法も提案される。

【0017】

生成するステップにおいては、トレーニングデータは、強化学習用エージェントによって、車両の少なくとも1つの最適なアクションを特定するために使用可能である。

【0018】

1つの実施形態においては、当該方法は、ストラテジによって決定された制御信号を用いて、車両の少なくとも1つのアクチュエータを制御するためのステップも含む。このような実施形態は、車両がアクション及びタスクを効率的に実施するように、車両を制御することができるという利点を提供する。

【0019】

本明細書において提示されるそれぞれの方法は、例えば、ソフトウェア若しくはハードウェアの形態で、又は、ソフトウェアとハードウェアとの混合形態で、例えば、装置又はコントローラにおいて実装可能である。

【0020】

本明細書において提示されるアプローチは、本明細書において提示される方法の変形例のステップを実施するように、制御するように、又は、対応する手段を用いて実行するように構成されている装置も企図している。本発明の基礎となる目的は、装置の形態である本発明のこの実施形態によっても迅速かつ効率的に達成可能である。

【0021】

この目的のために、当該装置は、信号又はデータを処理するための少なくとも1つの計算ユニット、信号又はデータを保存するための少なくとも1つのメモリユニット、センサからセンサ信号を読み出し、又は、データ又は制御信号をアクチュエータに出力する、センサ又はアクチュエータとの少なくとも1つのインタフェース、及び/又は、通信プロトコルに埋め込まれたデータを読み出す又は出力する少なくとも1つの通信インタフェースを含み得る。計算ユニットは、例えば信号プロセッサ、マイクロコントローラ等であってよく、メモリユニットは、フラッシュメモリ、EEPROM、又は、磁気的なメモリユニットであってよい。通信インタフェースは、データを無線及び/又は有線で読み出す又は出力するように構成可能であり、この場合、有線のデータを読み出す又は出力することができる通信インタフェースは、例えばデータを電氣的又は光学的に、対応するデータ伝送線路から読み出すことができ、又は、対応するデータ伝送線路に出力することができる。

【0022】

装置は、センサ信号を処理し、このセンサ信号に応じて制御信号及び/又はデータ信号を出力する電氣的な装置であると理解可能である。装置は、ハードウェア及び/又はソフ

10

20

30

40

50

トウェアの形態で実現されたインタフェースを含み得る。ハードウェアの形態で実現されている場合には、インタフェースは、例えば、装置の複数の異なる機能を含むいわゆるシステム A S I C の一部であってよい。しかしながら、インタフェースは、個々の集積回路であってもよいし、又は、少なくとも部分的にディスクリット部品からなっている場合もある。ソフトウェアの形態で実現されている場合には、インタフェースは、例えばマイクロコントローラ上において他のソフトウェアモジュールと一緒に実現されているソフトウェアモジュールであってよい。

【 0 0 2 3 】

本明細書において提示されるアプローチは、作業環境内の粒状材料と相互作用するように構成されている車両、特にオフロード車両であって、当該車両は、本明細書において提示される装置を含む、車両も提供する。

10

【 0 0 2 4 】

半導体メモリ、ハードディスク、又は、光学メモリのような機械可読担体上若しくは記憶媒体上に保存することができるプログラムコードを備えたコンピュータプログラム製品又はコンピュータプログラムであって、特にコンピュータ上又は装置上において実施された場合に、前述した実施形態のうちの1つによる方法のステップを実施するために、実装するために及び/又は制御するために使用されるコンピュータプログラム製品又はコンピュータプログラムも、有利である。

【 0 0 2 5 】

1つの実施形態は、機械学習のためのツールチェーンにおける前置部分として使用可能である。この実施形態は、以下に挙げる用途のために使用することができる機械学習システムをトレーニングするためのトレーニングデータを生成するための方法として使用可能である。このようにして機械学習システムがトレーニングされた後、この機械学習システムは、以下に説明するように後続の目的のために使用可能となる。機械学習システムは、オフロード車両のデータを分析するために、かつ、地面と相互作用するタスクのために使用可能である。機械学習システムは、1つ又は複数の連続値を決定するために、すなわち、例えば車両のアクションに起因する地面の変化に対する回帰を実施するために使用可能である。この場合、これらの実施形態は、特定の物理的なシステム、すなわちオフロード車両と接続される。1つの実施形態によって得られたシミュレーションを用いて、オフロード車両のための最適なストラテジを作成するために強化学習用エージェントをトレーニングすることができる。したがって、オフロード車両、例えばロボットなどのコンピュータ制御機械、車両、又は、地面と相互作用する他の車両のような物理的なシステムを制御するための制御信号を計算又は生成することができる。

20

30

【 0 0 2 6 】

本明細書において提示されるアプローチの実施形態を、図面に示し、以下の記載においてより詳細に説明する。

【 図面の簡単な説明 】

【 0 0 2 7 】

【 図 1 】 車両内の装置の1つの実施形態の概略図である。

【 図 2 】 車両を制御するように構成されたストラテジを生成するための強化学習用エージェントに対してトレーニングデータを提供するための方法の1つの実施形態のフロー図である。

40

【 図 3 】 車両を制御するように構成されたストラテジを生成するための方法の1つの実施形態のフロー図である。

【 図 4 】 車両の一連のアクションのシミュレーションの概略図である。

【 図 5 】 1つの実施形態によるシミュレーションプロセスのフロー図である。

【 図 6 】 1つの実施形態によるシミュレーションプロセスのフロー図である。

【 発明を実施するための形態 】

【 0 0 2 8 】

本発明の有利な実施形態の以下の説明においては、複数の異なる図面に示されている同

50

様に作用する要素には同一又は類似の参照符号を付しており、これらの要素についての説明の繰り返しは省略するものとする。

【0029】

図1は、車両100内の装置105の1つの実施形態の概略図を示す。車両100は、作業環境内の粒状材料と相互作用するように構成されている。特に、車両100は、オフロード車両であり、例えば、ブルドーザ、平土機、転圧機、ダンプカー、掘削機、建設車両、重機、重量物用車、又は、環境との相互作用、特に粒状材料との相互作用が含まれる複数のタスクを有する他の種類の車両である。車両100は、装置105を含む。車両100は、車両100のアクションを引き起し又は実施する少なくとも1つのアクチュエータ102をさらに含む。装置105とアクチュエータ102とは、信号通信又はデータ通信のために互いに接続されている。

10

【0030】

装置105は、車両100のような車両を制御するように構成されたストラテジ135を生成するための強化学習用エージェント130に対してトレーニングデータ125を提供するように構成されており、及び/又は、装置105は、車両100を制御するためのストラテジ135を生成するように構成されている。換言すれば、装置105は、図2による方法及び/又は図3による方法のステップを、対応するユニットにおいて実施及び/又は制御するように構成されている。装置105は、初期化ユニット110及び対応付けユニット120と、場合によっては強化学習用エージェント130又は生成ユニットも含む。1つの実施形態によれば、装置105は、制御ユニット140又はコントローラをさらに含む。

20

【0031】

装置105の提供部分は、初期化ユニット110及び対応付けユニット120を含む。装置105の生成部分は、強化学習用エージェント130及び任意選択肢の制御ユニット140を含む。図1の図示によれば、初期化ユニット110、対応付けユニット120、強化学習用エージェント130、及び、任意選択肢の制御ユニット140はすべて、装置105のうちの、車両100の内部に配置されている部分として実現されている。他の実施形態によれば、初期化ユニット110及び対応付けユニット120、すなわち提供部分を、装置105のうちの、物理的に車両100の外部に配置されている部分として実現してもよい。

30

【0032】

初期化ユニット110は、車両100の、粒状材料を含む作業環境の環境モデル115を初期化するように構成されている。初期化ユニット110は、粒状材料と作業環境との選択された物理的特徴値INを使用して、環境モデル115を作成するように構成されている。選択された物理的特徴値INは、初期化ユニット110によって装置105の外部のデータソースから入力又はロードされる。選択された物理的特徴値INは、初期化ユニット110によって少なくとも1つの確率密度関数に関する変数として使用される。粒状材料を含む作業環境の実際高さ状態と目標高さ状態との間の相対的な高さマップが、初期化ユニット110によって確率密度関数を使用して生成される。初期化ユニット110は、環境モデル115を対応付けユニット120に出力するように構成されている。

40

【0033】

1つの実施形態によれば、初期化ユニット110は、確率密度関数として連続多変量ガウス確率密度関数を使用する。選択された物理的特徴値INは、特に、粒状材料の比質量と、粒状材料の拡布パラメータと、粒状材料のそれぞれの山の体積と、環境に対するそれぞれの山の中心点の位置とを含み得る。

【0034】

対応付けユニット120は、確率密度関数に基づく相対的な高さマップを含む環境モデル115を、初期化ユニット110から受信又はロードするように構成されている。対応付けユニット120は、車両の複数のアクションの集合のうちのそれぞれのアクションの目標状態を達成するために、環境モデル115の確率密度関数の変化と報酬値とを対応付

50

けるように構成されている。環境モデル 1 1 5 の相対的な高さマップと、車両位置とは、トレーニングデータ 1 2 5 を提供するために、対応付けユニット 1 2 0 によってそれぞれのアクションごとに更新される。換言すれば、対応付けユニット 1 2 0 は、環境モデル 1 1 5 を使用してトレーニングデータ 1 2 5 を提供するように構成されている。

【 0 0 3 5 】

1 つの実施形態によれば、対応付けユニット 1 2 0 は、作業環境を複数の境界枠に分割し、確率密度関数の変化をそれぞれの境界枠ごとに個別に計算するように構成されている。追加的に又は代替的に、対応付けユニット 1 2 0 は、1 つの実施形態によれば、アクションが前進並進であるか、後退並進であるか又は回転であるかに応じて、及び / 又は、車両位置が作業環境の許容領域内に位置しているか又は許容領域外に位置しているかに応じて、アクションにそれぞれ異なる変化及び / 又はそれぞれ異なる報酬値を対応付けるように構成されている。

10

【 0 0 3 6 】

強化学習用エージェント 1 3 0 は、トレーニングデータ 1 2 5 を、対応付けユニット 1 2 0 から、すなわち装置 1 0 5 の提供部分から受信又はロードするように構成されている。強化学習用エージェント 1 3 0 は、トレーニングデータ 1 2 5 を使用して、強化学習によってストラテジ 1 3 5 を生成するように構成されている。エージェント 1 3 0 は、強化学習のためにストラテジ 1 3 5 を出力するようにさらに構成されている。

【 0 0 3 7 】

任意選択肢の制御ユニット 1 4 0 は、ストラテジ 1 3 5 を、強化学習用エージェント 1 3 0 から受信又はロードするように構成されている。制御ユニット 1 4 0 は、ストラテジ 1 3 5 に基づいて制御信号 1 4 5 を決定するようにさらに構成されている。制御ユニット 1 4 0 は、制御信号 1 4 5 を用いて、車両 1 0 0 の少なくとも 1 つのアクチュエータ 1 0 2 を制御するようにさらに構成されている。この目的のために、制御ユニット 1 4 0 は、制御信号 1 4 5 を少なくとも 1 つのアクチュエータ 1 0 2 に出力するように構成されている。他の実施形態によれば、制御ユニット 1 4 0 は、装置 1 0 5 から分離されたユニットであってよい。

20

【 0 0 3 8 】

図 2 は、車両、特にオフロード車両を制御するように構成されたストラテジを生成するための強化学習用エージェントに対してトレーニングデータを提供するための方法 2 0 0 の 1 つの実施形態のフロー図を示す。車両は、作業環境内の粒状材料と相互作用するように構成されている。提供するための方法 2 0 0 は、図 1 による装置又は類似の装置を用いて又はこれに関連して実施可能である。提供するための方法 2 0 0 は、コンピュータ実装された方法である。提供するための方法 2 0 0 は、初期化するステップ 2 1 0 と、対応付けるステップ 2 2 0 とを含む。

30

【 0 0 3 9 】

初期化するステップ 2 1 0 においては、粒状材料を含む作業環境の環境モデルが初期化される。初期化のために、粒状材料と作業環境との選択された物理的特徴値が、少なくとも 1 つの確率密度関数の変数として使用される。さらに、粒状材料を含む作業環境の実際高さ状態と目標高さ状態との間の相対的な高さマップが、確率密度関数を使用して作成される。続いて、対応付けるステップ 2 2 0 においては、目標状況を達成するために、環境モデルの確率密度関数の変化と報酬値とが、車両の複数のアクションの集合のうちのそれぞれのアクションに対応付けられる。トレーニングデータを提供するために、環境モデルの相対的な高さマップと車両位置とが、それぞれのアクションごとに更新される。

40

【 0 0 4 0 】

図 3 は、車両、特にオフロード車両を制御するように構成されたストラテジを生成するための方法 3 0 0 の 1 つの実施形態のフロー図を示す。車両は、作業環境内の粒状材料と相互作用するように構成されている。生成するための方法 3 0 0 は、図 1 による装置又は類似の装置を用いて又はこれに関連して実施可能である。生成するための方法 3 0 0 は、提供するステップ 3 2 5 と、生成するステップ 3 3 0 とを含む。

50

【 0 0 4 1 】

提供するステップ 3 2 5 においては、図 2 に示されている提供するための方法又は類似の方法のステップを実施することによってトレーニングデータが提供される。続いて、生成するステップ 3 3 0 においては、強化学習用エージェントを用いて、トレーニングデータを使用してストラテジが生成される。1つの実施形態によれば、生成するための方法 3 0 0 は、制御するステップ 3 4 0 も含む。制御するステップ 3 4 0 においては、生成するステップ 3 3 0 で生成されたストラテジを使用して決定された制御信号を用いて、車両の少なくとも1つのアクチュエータが制御される。

【 0 0 4 2 】

図 4 は、車両の一連のアクションのシミュレーションの概略図を示す。このシミュレーションは、図 1 に示されている装置又は類似の装置によって、及び/又は、図 2 及び/又は図 3 に示されている方法又は類似の方法を実施することによって作成される。具体的には、図 4 は、車両 1 0 0 に関するシミュレーション例、この場合には例えば7つのアクションの集合によって砂山を敷均しするブルドーザに関するシミュレーション例の8つの概略図を示す。図 4 は、それぞれ環境モデル 1 1 5 の相対的な高さマップを表す8つの線図又は部分図 $T = 0$ 、 $T = 1$ 、 $T = 2$ 、 $T = 3$ 、 $T = 4$ 、 $T = 5$ 、 $T = 6$ 及び $T = 7$ を示し、ここで、 $T = 0$ は、初期化された状況としての最初の状況を表し、後続の線図 $T = 1$ 、 $T = 2$ 、 $T = 3$ 、 $T = 4$ 、 $T = 5$ 、 $T = 6$ 及び $T = 7$ は、車両 1 0 0 のそれぞれ1つのアクションを順次に示す。

【 0 0 4 3 】

このシミュレーションは、車両 1 0 0 としてのブルドーザと、地面との間の相互作用を示し、この場合、ブルドーザは、砂山を敷均しすることが求められている。ここに示されている砂山は、互いに隣り合って並んでいる6つの部分砂山から形成され、最終的な砂山は、これらの砂山のすべてのガウス PDF の合計である。それぞれの線図 $T = 1$ 、 $T = 2$ 、 $T = 3$ 、 $T = 4$ 、 $T = 5$ 、 $T = 6$ 及び $T = 7$ は、車両 1 0 0 が砂山を移動させて敷均しするそれぞれ1つの異なる時点を表す。車両 1 0 0 の毎回の前進移動が、ブルドーザのブレードが地面上に置かれた状態で実施されて、地面を敷均しするということを前提とする。見て取れるように、毎回の前進移動のたびに地面のさらなる部分が敷均しされていき、車両 1 0 0 が目下の移動に対する最終位置に到着するとすぐに、ブレード内の残りの地面がガウス PDF として新たに分布される（例えば、 $T = 2$ を参照のこと）。他の実施形態においては、例えばガウス混合モデルのようなより複雑な表現を使用してもよいが、一方では、迅速なアプローチを維持するために、他方では、地面との相互作用の良好な表現を達成するために、ガウス PDF を使用することによって数学的な計算を簡略化することができるということに留意すべきである。

【 0 0 4 4 】

以下においては、前述した図面を参照して、実施形態と、実施形態の背景に対する利点とについて部分的に換言しながらまとめ、及び/又は、より詳細に説明することとする。

【 0 0 4 5 】

トレーニングデータ 1 2 5 及び/又はストラテジ 1 3 5 は、例えば地面と相互作用するあらゆる任意の車両 1 0 0 に対して使用可能であるが、本明細書の重点は、砂山を所定の高さにすることが求められている特定の車両種類にある。本説明中においては、シミュレーションにおける環境モデル 1 1 5 の高さマップは、実際の砂の状態と所望の最終的な砂の高さとの間の相対的なマップであるということが前提とされる。さらに、グリッドを作成して、グリッド単位とメートルとの間の換算を、以下の換算： $(x, y)_{grid} = (x_m * [pixel / meter], y_m * [pixel / meter])$ のように定義することによって、タスクが簡略化される。

【 0 0 4 6 】

実施形態をより良好に理解するためには、多変量ガウス PDF を理解すべきである。この場合にはデカルト座標である2つの変数 (x, y) が、両方とも正規分布から取り出されていると仮定すると、この場合における確率密度関数は、これら2つの変数が独立同一

10

20

30

40

50

分布に従っているという仮定のもとで、

【数 1】

$$f(x, y) = \frac{V}{2\pi\sigma_x\sigma_y} * e^{-\frac{1}{2}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]}$$

のようになる。

【0047】

この関数は、それぞれの (x , y) 点における地面の高さを特定するために使用される。この関数 V について積分が行われ、その積分が粒状材料の山の体積を表すことに留意されたい。これらの変数に対する単位は、V [m³] ; x [m] ; y [m] である。砂は均質であるので、山がすべての方向に均一に分布していることがさらに仮定され、したがって、以下の関係： x = y = が仮定される。この場合には、高さ分布は、

【数 2】

$$h(x, y) = \frac{V}{2\pi\sigma^2} * e^{-\frac{1}{2}\left[\left(\frac{x-\mu_x}{\sigma}\right)^2 + \left(\frac{y-\mu_y}{\sigma}\right)^2\right]}$$

まで低減する。

【0048】

すなわち、山の最大高さは、H_{max} = v / 2² である。さらに、(r ,) における (x , y) を、(x - μ_x) / = r * cos 及び (y - μ_y) / = r * sin を使用して変化させると、以下の関係：

【数 3】

$$h(r, \theta) = H_{max} * e^{-\frac{1}{2}r^2}$$

が得られる。

【0049】

車両 100 として、ブルドーザの簡単な機構が前提とされ、この場合、3つの可能な速度段階が存在し、ブルドーザの速度とブレード上の負荷との間には線形の関係がある。車両速度と負荷との間の簡単な関係は、アクションにかかる合計時間に影響を与え、その選択されたアクションの間に移動させられる砂の体積に基づいて計算される。さらに、ブレードは、このブレードが所定の時間内に敷均しすることができる最大の体積である既知の体積を有していることが前提とされる。

【0050】

シミュレーションを初期化するために、提供するための方法 200 を初期化するステップ 210 が実施される。構成データを用いてシミュレーションを初期化することができ、この構成データにおいて、以下の情報、すなわち、砂の比質量、

【数 4】

$$\left[\frac{kg}{m^3} \right]$$

と、砂の拡布パラメータ (SSP) と、それぞれの砂山についての砂山の体積 [m³] 及び砂山の中心点の位置 (x , y) [m , m] とが定義されている。これらのパラメータに基づいて、ガウス PDF 係数と砂パラメータとの間の以下の関係：

【数 5】

$$SSP = \frac{r_{max}^3}{V}$$

10

20

30

40

50

に基づいて、作業環境を定義することができ、ここで、 V は、砂山の体積 [m^3] であり、 r_{max} は、砂山の半径 [m] であり、ただし、 $r_{max} = 3$ である。なぜなら、6 が、領域全体の 99.7% を占めるからである。シミュレーションのための入力は、それぞれの砂山の SSP 及び体積であるので、これらの 2 つのパラメータが指定されると対応する式 $g(V, SSP) = (H_{max}, r_{max})$ を返す、オフライン機能を作成することが可能である。

【0051】

これらのパラメータが判明するとすぐに、初期時の環境又は環境モデル 115 を、以下の部分ステップにおいて作成することができ、すなわち、

- ・実際の面積の大きさと分解能とを用いて空白の高さマップを作成する： $h [meter / pixel]$ ステップと、
- ・山のリスト内のそれぞれの山ごとに、
 - ・上に挙げた方程式を使用して、砂山の高さ

【数6】

$$h_{s_i}$$

のマップを作成して、

- ・砂山の高さの目下のマップを、全体高さのマップに加算する：

【数7】

$$h = h + h_{s_i}$$

ステップと、

- ・相対的な高さマップを作成するために、目下の高さマップから所望の最終的な敷均し高さを減算するステップと

において作成することができる。このプロセスが終了すると、それぞれの点 (x, y) において除去されるべき粒状材料の量が判明している、領域の高さマップがもたらされる。

【0052】

次いで、トレーニングデータ 125 を提供するために、次のステップを計算することが必要である。シミュレーションにおいてステップを計算するために、提供するための方法 200 の対応付けるステップ 220 において、車両 100 の移動に基づいて地面の変化が特定される。車両 100 がブルドーザである場合であって、かつ、敷均しタスクである場合には、ブルドーザの移動によって地面が変化する。というのも、ブレードが砂山を前方へ押し出して、表面を可能な限り平らにするからである。この変化は、車両 100 が移動している領域を、複数の比較的小さな境界枠に分割し、続いて、それぞれの領域に対する相互作用の影響を別個に計算することによってモデル化される。以下に説明する図 5 及び図 6 においては、車両 100 のアクションの部分ステップによる、車両 100 の移動のシミュレーションロジックを見て取ることができる。

【0053】

シミュレーションにおいては、車両 100 の、特にブルドーザのアクションに対するアクション空間は、2 つの可能なアクション、すなわち (a) 前進/後退と、(b) 回転とを含み、すなわち、ニューラルネットワーク (NN) は、(2 x 1) ベクトルを出力する。第 1 のエントリにおける正の値は、前進を意味し、その一方で、負の値は、後退を意味する。アクションベクトルの第 2 のエントリにおける正の値は、時計回りの回転を表し、負の値は、反時計回りの回転を表す。状況と、対応する好ましいアクションとが存在する場合、強化学習用エージェント 130 は、このアクションを実施する。砂の移動がこれらのアクションの結果であると理解するためには、並進移動と回転移動とを区別すべきである。

【0054】

並進の場合、新たな車両位置が計算される。新たな位置が有効でない場合、すなわち許

10

20

30

40

50

容領域外に位置している場合には、大きな負の報酬が得られ、新たなシナリオが開始される。新たな位置が有効である場合には、拡布が行われる。ここで、慣性座標系におけるこのシミュレーション反復視野の内部の領域を、このステップ境界についての示唆として計算する。続いて、この領域内の高さマップを計算する。この領域内の砂が、拡布される。次いで、シミュレータは、この増分する単位面積の合計体積を計算し、ブルドーザは、合計体積を、この単位面積の平均値に分布させることができる。並進運動が終了し、ブレードが砂で覆われている場合には、シミュレータ、すなわち装置 105 の対応するユニットは、ブレード上に残っている体積を有するさらなるガウス砂山を生成する。続いて、シミュレータは、このガウス砂山を刈り揃えて、最適でない拡布パターンをシミュレートする。最後に、シミュレータは、このアクションのために必要とされる時間を計算して、報酬を得る。後退走行又は転回は、砂、すなわち高さマップを変化させず、負の報酬をもたらすだけである。

10

【0055】

回転の場合には、回転角度が計算される。車両の DCM (方向余弦行列) は、走行方向を表しているので更新される。回転運動は、砂の分布を変化させないので、小さな負の報酬が追加される。車両 100 と砂との間に相互作用が存在しない場合には、ブルドーザが回転し、したがって、回転中には正の報酬を得ることができない。

【0056】

図 5 は、1つの実施形態によるシミュレーションプロセス 500 のフロー図を示す。シミュレーションプロセス 500 は、図 1 による装置によって、及び / 又は、図 2 による提供するための方法及び / 又は図 3 による生成するための方法に関連して実施可能である。換言すれば、図 5 は、シミュレーションのアクションステップを示す。

20

【0057】

ブロック 502 は、車両姿勢、ここではブルドーザ姿勢と、高さマップと、ブレード姿勢との入力である。プロセス 500 は、ブロック 502 から、ストラテジネットワークを表すブロック 504 にさらに進み、アクションを表すブロック 506 にさらに進み、アクションが並進であるかどうか判定されるブロック 508 にさらに進む。アクションが並進である場合には、プロセス 500 は、ブロック 508 から、車両が前進移動しているかどうか判定されるブロック 510 にさらに進む。車両が前進移動している場合には、プロセス 500 は、ブロック 512 において続行される。ブロック 512 は、以下の命令：

30

- (現在地 - 最終的な地点 > ϵ) の間に、
- 1 . 積分の面積要素に対する高さマップを作成する
 - 2 . この面積要素の新たな高さを計算する
 - 3 . 更新された位置 = 位置 + Δ pos
 - 4 . さらに 1 へ

を含む。

【0058】

ブロック 512 から、又は、アクションが並進でない場合 (ブロック 508 を参照のこと)、若しくは、車両が前進移動していない場合 (ブロック 510 を参照のこと) には、プロセス 500 は、ブロック 514 にさらに進む。ブロック 514 は、以下の命令：

40

- 1 . ステップの合計時間を計算するステップ
- 2 . ブルドーザ姿勢及びブレード姿勢を更新するステップ
- 3 . 高さマップを更新するステップ
- 4 . 報酬 = - 時間を計算するステップ

を終了することを含む。

【0059】

プロセス 500 は、ブロック 514 から、シミュレートされたアクションが実施されるかどうかチェックされるブロック 516 にさらに進む。このチェックは、以下の基準：

- 1 . Max (局所的な熱画像マップ) < ϵ
- 2 . $N_{\text{ステップ}} > N_{\text{ステップ}_{max}}$

50

に依存している。

【 0 0 6 0 】

アクションが実施される場合には、プロセス 5 0 0 は、新たなシナリオを生成するためのリセットを表すブロック 5 1 8 にさらに進む。アクションが実施されない場合には、プロセス 5 0 0 は、ブロック 5 0 2 に戻る。

【 0 0 6 1 】

図 6 は、1つの実施形態によるシミュレーションプロセス 6 0 0 のフロー図を示す。シミュレーションプロセス 6 0 0 は、図 1 による装置によって、及び/又は、図 2 による提供するための方法及び/又は図 3 による生成するための方法に関連して実施可能である。換言すれば、図 6 は、上述したシミュレーションステップのためのアルゴリズム、又は、それぞれの比較的小さな領域若しくは境界枠のためのロジックを示す。

10

【 0 0 6 2 】

ブロック 6 0 2 においては、車両経路、ここではブルドーザ経路が、車体軸線に沿った小さな長方形に分割される。プロセス 6 0 0 は、ブロック 6 0 2 から、長方形リスト内のそれぞれの長方形ごとに、除去されるべき長方形の体積が計算されるブロック 6 0 4 にさらに進む。体積がゼロより大きい場合には、プロセス 6 0 0 は、ブロック 6 0 6 にさらに進み、ブロック 6 0 6 においては、ブレード内の新たな体積と、新たな地面高さとは、

$$V_{\text{ブレード}} = V_{\text{ブレード}} + \text{体積}$$

$$\text{高さマップ}(x, y) = \text{目標高さ}$$

のように計算される。

20

【 0 0 6 3 】

体積がゼロより小さい場合には、プロセス 6 0 0 は、ブロック 6 0 8 にさらに進む。ブロック 6 0 8 は、

$V_{\text{ブレード}} > \text{体積}$ の場合には：

$$\text{高さマップ}(x, y) = \text{体積}$$

$$V_{\text{ブレード}} = V_{\text{ブレード}} - \text{目標高さ}$$

そうでない場合には：

$$\text{高さマップ}(x, y) = \text{ブレード内の平均体積}$$

$$V_{\text{ブレード}} = 0$$

を含む。

30

【 0 0 6 4 】

次いで、プロセス 6 0 0 は、ブロック 6 0 6 及びブロック 6 0 8 から、次の長方形への移行を表すブロック 6 1 0 にさらに進む。

【 0 0 6 5 】

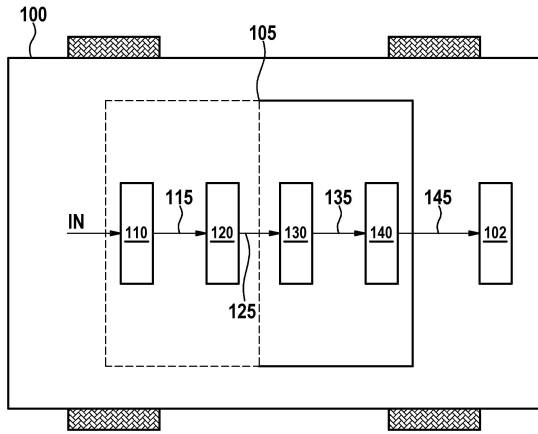
ある実施形態が、第 1 の特徴と第 2 の特徴との間に「及び/又は」の接続詞を含む場合には常に、このことは、この実施形態が、第 1 の変形例においては、第 1 の特徴及び第 2 の特徴の両方を含み、さらなる変形例においては、第 1 の特徴又は第 2 の特徴のいずれかを含むということの意味する。

40

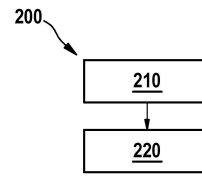
50

【 図面 】

【 図 1 】

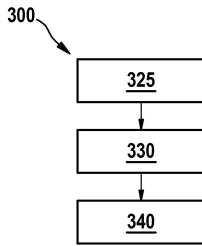


【 図 2 】

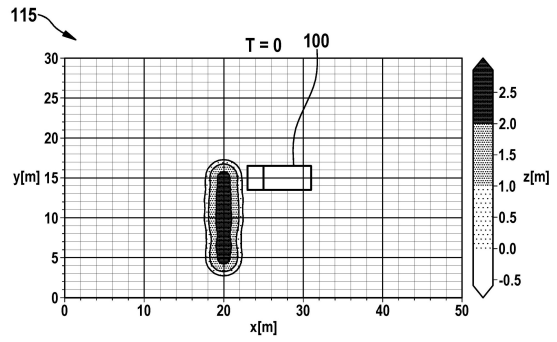


10

【 図 3 】

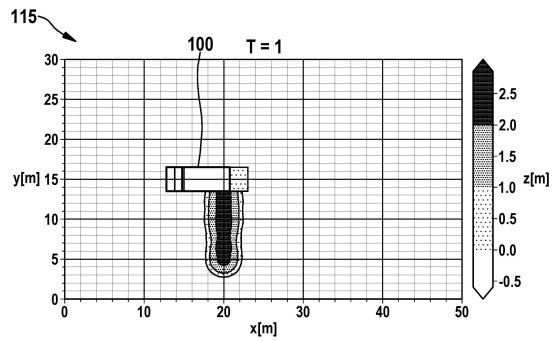


【 図 4 - 1 】



20

a



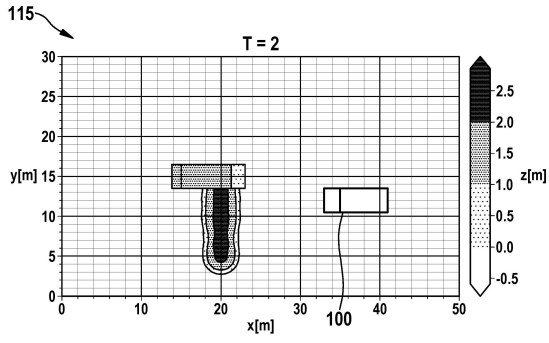
30

b

40

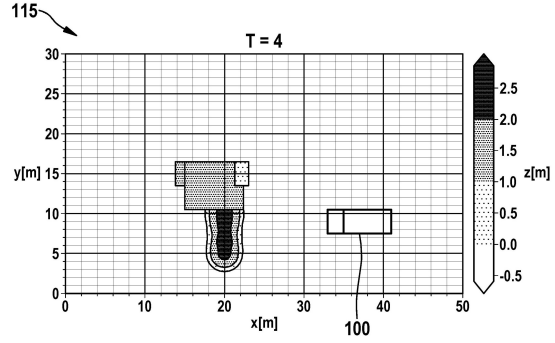
50

【 図 4 - 2 】



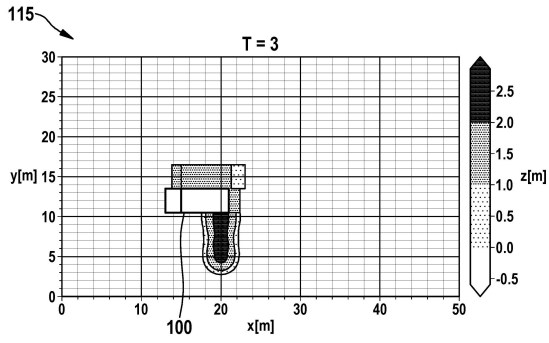
c

【 図 4 - 3 】

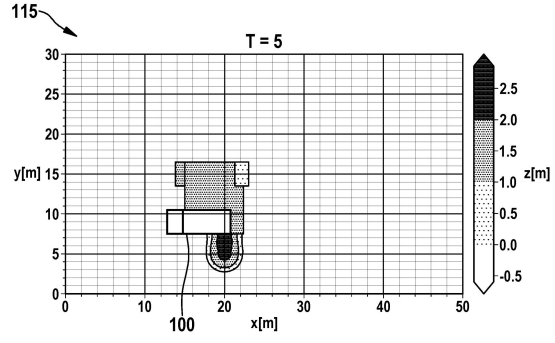


e

10



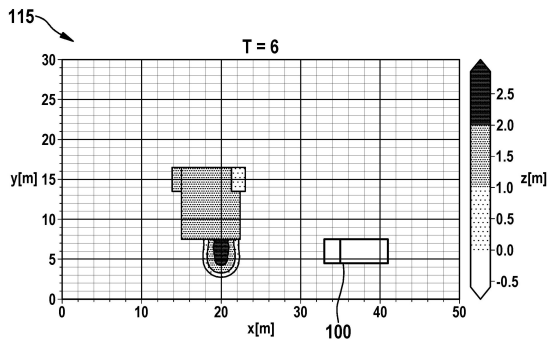
d



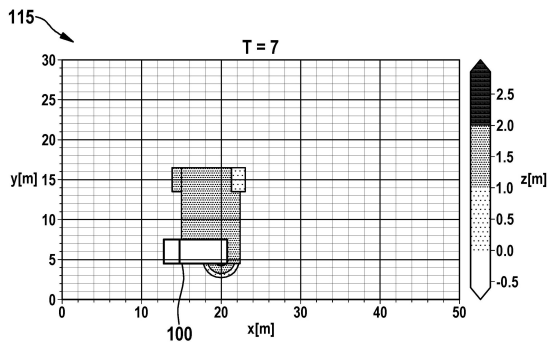
f

20

【 図 4 - 4 】

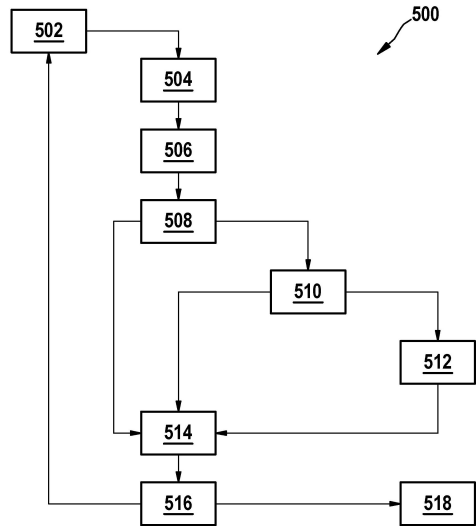


g



h

【 図 5 】

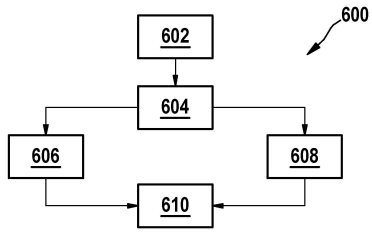


30

40

50

【 図 6 】



【 外国語明細書 】

10

2023054776000018.pdf

20

30

40

50

フロントページの続き

(54)【発明の名称】 を制御するように構成されたストラテジを生成するための装置及び方法、並びに、車両
弁理士 永島 秀郎

(74)代理人 100162880
弁理士 上島 類

(72)発明者 チャナ ロス
イスラエル国 ジフロン・ヤアコヴ アフ・アル・ピチェン ストリート 45 / 7

(72)発明者 ヤコフ ミロン
イスラエル国 ハイファ ヤコフ ショハム ストリート 21 / 9

(72)発明者 ドータン ディ カストロ
イスラエル国 ハイファ シュロムジオン ハマルカ ナンバー 17

Fターム(参考) 5L049 CC07