



(12)

Patentschrift

(21) Aktenzeichen: **10 2021 211 152.1**

(22) Anmeldetag: **04.10.2021**

(43) Offenlegungstag: –

(45) Veröffentlichungstag
der Patenterteilung: **15.12.2022**

(51) Int Cl.: **B60W 50/00 (2006.01)**

Innerhalb von neun Monaten nach Veröffentlichung der Patenterteilung kann nach § 59 Patentgesetz gegen das Patent Einspruch erhoben werden. Der Einspruch ist schriftlich zu erklären und zu begründen. Innerhalb der Einspruchsfrist ist eine Einspruchsgebühr in Höhe von 200 Euro zu entrichten (§ 6 Patentkostengesetz in Verbindung mit der Anlage zu § 2 Abs. 1 Patentkostengesetz).

(73) Patentinhaber:
Robert Bosch Gesellschaft mit beschränkter Haftung, 70469 Stuttgart, DE

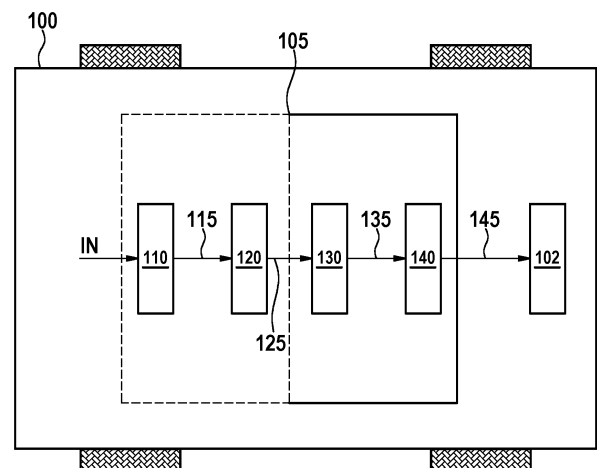
(72) Erfinder:
Ross, Chana, Zichron Yaakov, IL; Di Castro, Dotan, Haifa, IL; Miron, Yakov, Haifa, IL

(56) Ermittelte Stand der Technik:

DE	198 59 169	A1
DE	10 2014 203 312	A1
DE	11 2009 001 554	T5
US	10 481 603	B2

(54) Bezeichnung: **Vorrichtung und computer-implementiertes Verfahren zum Bereitstellen von Trainingsdaten für einen Agenten für bestärkendes Lernen zum Erzeugen einer zur Steuerung eines Fahrzeugs ausgelegten Strategie, Vorrichtung und Verfahren zum Erzeugen einer zur Steuerung eines Fahrzeugs ausgelegten Strategie und Fahrzeug**

(57) Zusammenfassung: Die Erfindung betrifft ein computer-implementiertes Verfahren zum Bereitstellen von Trainingsdaten (125) für einen Agenten (130) für bestärkendes Lernen zum Erzeugen einer Strategie (135), die zum Steuern eines Fahrzeugs (100) ausgelegt ist, das zur Interaktion mit granulearem Material in einer Arbeitsumgebung ausgelegt ist. Ein Umgebungsmodell (115) der das granulare Material enthaltenden Arbeitsumgebung wird initialisiert. Ausgewählte physikalische Merkmalswerte (IN) des granularen Materials und der Arbeitsumgebung werden als Variablen mindestens einer Wahrscheinlichkeitsdichtefunktion verwendet, die zum Erzeugen einer relativen Höhenkarte zwischen einem Ist-Höhenstatus und einem Soll-Höhenstatus der das granulare Material enthaltenden Arbeitsumgebung verwendet wird. Zum Erreichen des Sollstatus werden jeder Aktion eines Satzes von Aktionen des Fahrzeugs (100) eine Änderung der Wahrscheinlichkeitsdichtefunktion des Umgebungsmodells (115) und ein Belohnungswert zugeordnet. Für jede Aktion werden die relative Höhenkarte des Umgebungsmodells (115) und eine Fahrzeugposition aktualisiert, um die Trainingsdaten (125) bereitzustellen.



Beschreibung

Stand der Technik

[0001] Die Erfindung betrifft eine Vorrichtung oder ein Verfahren gemäß den unabhängigen Ansprüchen. Ein Computerprogramm ist ebenfalls Gegenstand der vorliegenden Erfindung.

[0002] Im Zusammenhang mit autonomen Fahrzeugen lassen sich granulare Simulationen oder Simulationen von granularem Material im Allgemeinen in drei Arten von Simulationen unterteilen, zum Beispiel: Festkörpermechanik und empirische Formeln, Diskrete-Elemente-Verfahren und Höhenkartenansatz. Insbesondere verwenden die genannten Verfahren iterative Algorithmen, um die Interaktion zwischen dem granularen Material und dem Fahrzeug zu ermitteln. Darüber hinaus können diese Simulationen offline berechnet und anschließend ausgeführt werden, um die Ergebnisse zu betrachten, ähnlich wie bei anderen numerischen Verfahren wie Computational Fluid Dynamics (CFD), einem Verfahren, das im Allgemeinen für das bestärkende Lernen (BL, engl. reinforcement learning) unzureichend ist, da die Simulation viele Male ausgeführt werden muss, um den BL-Agenten zu trainieren. US 10 481 603 B2 offenbart einen Bewegungsbahnplanungsalgorithmus für Geländefahrzeuge.

[0003] Aus der DE 198 59 169 A1 ist ein System bekannt, welches aus einer Maschine mit mehreren mechanischen Verbindungen, die an Verbindungspunkten zusammengefügt sind, besteht. Ein Verarbeitungssystem führt mindestens ein Skript aus. Jedes Skript hat mindestens einen variablen Parameter, der eine Bewegung der Maschine definiert. Ein Lernalgorithmus modifiziert den Wert mindestens eines variablen Parameters auf der Grundlage eines gewünschten Ergebnisses und gemessener Bedingungen, die das Ergebnis beeinflussen.

[0004] Aus der DE 10 2014 203 312 A1 ist bekannt, dass Straßenneigung über eine Region modelliert werden können. Eine Abfolge von Neigungswerten wird während des Betriebs des Fahrzeugs an einer vorbestimmten Rate erzeugt, wobei jeder Neigungswert einen jeweiligen Neigungsbereich, der dann angetroffen wird, identifiziert. Ein Markov-Ketten-Straßenneigungsmodell wird als Reaktion auf die Abfolge von Neigungswerten aktualisiert, wobei das Modell jeweilige Wahrscheinlichkeitselemente in einer Matrix von Übergangsereignissen von jedem vorbestimmten Neigungsbereich zu einem jeweiligen darauf folgenden Neigungsbereich darstellt.

[0005] Aus der DE 11 2009 001 554 T5 ist ein Verfahren zum Betreiben eines Fahrzeugs auf einem Haufen eines Materials auf einer Baustelle bekannt, wobei das Material durch eine Öffnung auf der Bau-

stelle abgegeben wird und das Verfahren umfasst: Erfassen einer Oberfläche des Haufens, Identifizieren einer Störzone auf der Oberfläche des Haufens, die durch die Abgabe des Materials verursacht wird, basierend auf der erfassten Oberfläche und einer bekannten Position der Öffnung und Übermitteln eines die Störzone angegebenden Signals an das Fahrzeug.

Offenbarung der Erfindung

[0006] Vor diesem Hintergrund stellt der hier vorgestellte Ansatz ein Verfahren, eine Vorrichtung, die dieses Verfahren anwendet, und ein entsprechendes Computerprogramm gemäß den Hauptansprüchen bereit. Vorteilhafte Ausführungsformen und Verbesserungen des im unabhängigen Anspruch definierten Gegenstandes sind durch die in den abhängigen Ansprüchen angegebenen Merkmale möglich.

[0007] Gemäß Ausführungsformen kann beispielsweise eine vereinfachte granulare Simulation für das Training von bestärkendem Lernen bereitgestellt werden. Die Simulation kann beispielsweise auf gaußschen Verteilungen beruhen, wobei davon ausgegangen werden kann, dass sich das granulare Material wie eine kontinuierliche multivariate gaußsche Wahrscheinlichkeitsdichtefunktion (PDF) oder dergleichen verhält. Für die Simulation wurde eine Beziehung zwischen der Physik des granularen Materials und den mathematischen Gauß-Koeffizienten gefunden, die genutzt werden kann. Darüber hinaus kann eine Interaktion zwischen dem granularen Material und dem Fahrzeug als Änderung der gaußschen PDF formalisiert werden. Insbesondere kann die Umgebung durch eine Höhenkarte dargestellt werden, wobei $H(x, y) = h$ bedeutet, dass die Bodenhöhe am Ort (x, y) gleich h ist. Außerdem kann die Position des Fahrzeugs als Kombination aus dem Standort des Fahrzeugs und dem Azimutwinkel relativ zur Weltachse angenommen werden. Die Simulation kann beispielsweise zwei Hauptalgorithmen umfassen: einen zur Initialisierung des Bereichs oder der Umgebung, wobei die anfängliche Umgebungshöhenkarte basierend auf einer Liste von Positionen von Haufen granularen Materials erzeugt werden kann, und einen weiteren zur Berechnung eines aktualisierten Bereichs, wobei die neue Umgebung basierend auf der Fahrzeugbewegung oder -aktion ermittelt werden kann.

[0008] Gemäß Ausführungsformen kann beispielsweise eine vereinfachte Simulation bereitgestellt werden, die für Bewegungsbahnplanungsalgorithmen insbesondere im Gelände verwendet werden kann. Insbesondere gelingt es dem hier vorgestellten Ansatz, diese Art von Simulationen zu erstellen, die komplexe Physik zu erfassen und gleichzeitig einen einfachen automatischen Simulator beizubehalten, der nur kurze Zeit läuft. Die Simulation oder das

Modell kann zum Trainieren eines Algorithmus für bestärkendes Lernen (BL-Algorithmus) verwendet werden. Daher wird vorgeschlagen, eine Gym-API zu pflegen, in der die Simulation Folgendes enthalten sollte: Rücksetzen, Schritt, wird_ausgeführt, erstellen und Belohnung_berechnen. Insbesondere für BL, bei der die Simulation viele Male durchlaufen werden muss, um den BL-Agenten zu trainieren, ist die bisherige Verfahrensweise möglicherweise nicht ausreichend. Aus diesem Grund kann ein vereinfachter Simulator bereitgestellt werden, der beim Trainieren eingesetzt werden kann und zum Beispiel die Arbeit eines Arbeitstages auf der Baustelle in kürzester Zeit simulieren kann.

[0009] Die Simulation gemäß dem hier vorgestellten Ansatz kann vorteilhafter sein als die bisher genannten Ansätze, da sie auf Mehrfachberechnungen für jede Fahrzeugbewegung verzichten kann. Darüber hinaus kann die Komplexität der Simulation mit einfachen Verteilungen erhöht werden. So ist es beispielsweise möglich, der Fahrzeugbewegung Verschüttungen hinzuzufügen, indem eine Wahrscheinlichkeitsfunktion hinzugefügt wird, bei der die Bewegung des Fahrzeugs mit der Wahrscheinlichkeit P zu Verschüttungen führt. Der hier vorgestellte Ansatz kann die wichtigsten Interaktionsmerkmale zwischen dem Fahrzeug und der Umgebung simulieren, ohne dabei an Einfachheit einzubüßen. Darüber hinaus kann diese Art der Simulation vorteilhaft für das Trainieren eines BL-Agenten eingesetzt werden, da sie die wichtigsten Interaktionen erfasst, die für das BL-Training erforderlich sind, um eine optimale Aktion zu wählen, während nicht alle Interaktionen zwischen dem Fahrzeug und dem Boden für den Erfolg von BL-Agenten wichtig sind. Es ist auch möglich, Merkmale des granularen Materials vorteilhaft in mathematische Variablen umzuwandeln, die leicht zur Erstellung der gaußschen PDFs verwendet werden können.

[0010] Vorgestellt wird ein computer-implementiertes Verfahren zum Bereitstellen von Trainingsdaten für einen Agenten für bestärkendes Lernen zum Erzeugen einer Strategie, die zum Steuern eines Fahrzeugs, insbesondere eines Geländefahrzeugs, ausgelegt ist, wobei das Fahrzeug dazu ausgelegt ist, mit granularem Material in einer Arbeitsumgebung zu interagieren, wobei das Verfahren die folgenden Schritte umfasst:

Initialisieren eines Umgebungsmodells der das granulare Material enthaltenden Arbeitsumgebung, wobei ausgewählte physikalische Merkmalswerte des granularen Materials und der Arbeitsumgebung als Variablen mindestens einer Wahrscheinlichkeitsdichtefunktion verwendet werden, wobei eine relative Höhenkarte zwischen einem Ist-Höhenstatus und einem Soll-Höhenstatus der das granulare Material enthaltenden Arbeitsumgebung unter Verwen-

dung der Wahrscheinlichkeitsdichtefunktion erzeugt wird; und

Zuordnen einer Änderung der Wahrscheinlichkeitsdichtefunktion des Umgebungsmodells und eines Belohnungswerts zu jeder Aktion eines Satzes von Aktionen des Fahrzeugs zum Erreichen des Sollzustands, wobei die relative Höhenkarte des Umgebungsmodells und eine Fahrzeugposition für jede Aktion aktualisiert werden, um die Trainingsdaten bereitzustellen.

[0011] Bei dem Fahrzeug oder Geländefahrzeug kann es sich um eine Planierraupe, einen Bulldozer, einen Verdichter, einen Kipper, einen Bagger, ein Baufahrzeug, eine Schwermaschine, ein Schwerlastfahrzeug oder eine andere Art von Fahrzeug handeln, das eine Anzahl von Aufgaben hat, die eine Interaktion mit der Umgebung, insbesondere mit dem granularen Material, beinhalten. Beispiele für diese Art von Aufgaben oder Aktionen sind Planieren, Abladen von granulearem Material, Verdichten einer Fläche, Entfernen von granulearem Material usw. Das granulare Material kann Erde, Sand, Schnee, Schutt, Steine und/oder ähnliches partikelförmiges Material umfassen. Die Strategie kann mindestens eine geplante Bewegungsbahn für das Fahrzeug beinhalten. Das Umgebungsmodell kann die Karte der relativen Höhe basierend auf der Wahrscheinlichkeitsdichtefunktion enthalten. Ein Belohnungswert kann positiv, negativ oder null sein. Die Änderung der Wahrscheinlichkeitsdichtefunktion kann einen von null verschiedenen Absolutwert aufweisen oder gleich null sein.

[0012] Gemäß einer Ausführungsform kann im Schritt des Initialisierens die Wahrscheinlichkeitsdichtefunktion eine kontinuierliche multivariate gaußsche Wahrscheinlichkeitsdichtefunktion sein. Zusätzlich oder alternativ können im Schritt des Initialisierens die ausgewählten physikalischen Merkmalswerte eine spezifische Masse des granularen Materials, einen Ausbreitungsparameter des granularen Materials, ein Volumen jedes Haufens von granulearem Material und eine Position jedes Haufenmittelpunkts relativ zur Umgebung umfassen. Eine solche Ausführungsform bietet den Vorteil, dass die komplexe Physik des granularen Materials erfasst werden kann und gleichzeitig ein einfacher automatischer Simulator zur Verfügung steht, der nur kurze Zeit läuft.

[0013] Gemäß einer Ausführungsform kann die Arbeitsumgebung im Schritt des Zuordnens in eine Mehrzahl von Begrenzungsrahmen unterteilt werden, wobei die Änderung der Wahrscheinlichkeitsdichtefunktion für jeden Begrenzungsrahmen einzeln berechnet werden kann. Eine solche Ausführungsform bietet den Vorteil, dass die Veränderung der Umgebung aufgrund einer Aktion des Fahrzeugs

schnell und mit geringerem Rechenaufwand ermittelt werden kann.

[0014] Gemäß einer Ausführungsform können im Schritt des Zuordnens verschiedene Änderungen und zusätzlich oder alternativ verschiedene Belohnungswerte einer Aktion zugeordnet werden, je nachdem, ob es sich bei der Aktion um eine Vorwärtstranslation, eine Rückwärtstranslation oder eine Drehung handelt, und zusätzlich oder alternativ je nachdem, ob sich die Fahrzeugposition innerhalb oder außerhalb eines zulässigen Bereichs der Arbeitsumgebung befindet. Eine solche Ausführungsform bietet den Vorteil, dass Aktionen des Fahrzeugs für den Trainingsagenten leicht auswertbar und/oder umsetzbar sind.

[0015] Ebenfalls wird hier ein Verfahren zum Erzeugen einer Strategie vorgestellt, die dazu ausgelegt ist, ein Fahrzeug, insbesondere ein Geländefahrzeug, zu steuern, wobei das Fahrzeug dazu ausgelegt ist, mit granularem Material in einer Arbeitsumgebung zu interagieren, wobei das Verfahren die folgenden Schritte umfasst:

Bereitstellen von Trainingsdaten gemäß einem hier vorgestellten Verfahren zum Bereitstellen; und

Erzeugen der Strategie mittels eines Agenten für bestärkendes Lernen unter Verwendung der Trainingsdaten.

[0016] Im Schritt des Erzeugens können die Trainingsdaten durch den Agenten für bestärkendes Lernen dazu verwendet werden, mindestens eine optimale Aktion des Fahrzeugs zu ermitteln.

[0017] Gemäß einer Ausführungsform umfasst das Verfahren auch einen Schritt zum Steuern mindestens eines Aktuators des Fahrzeugs mittels eines Steuersignals, das anhand der Strategie bestimmt wird. Eine solche Ausführungsform bietet den Vorteil, dass das Fahrzeug so gesteuert werden kann, dass es Aktionen und Aufgaben auf effiziente Weise ausführt.

[0018] Jedes hier vorgestellte Verfahren kann zum Beispiel in Software oder Hardware oder in einer Mischung aus Software und Hardware, zum Beispiel in einer Vorrichtung oder einem Controller, implementiert werden.

[0019] Der hier vorgestellte Ansatz sieht auch eine Vorrichtung vor, die dazu ausgelegt ist, die Schritte einer Variante eines hier vorgestellten Verfahrens auszuführen, zu steuern oder mit entsprechenden Mitteln umzusetzen. Der der Erfindung zugrunde liegende Zweck kann auch mit dieser Ausführungsform der Erfindung in Form einer Vorrichtung schnell und effizient erreicht werden.

[0020] Zu diesem Zweck könnte die Vorrichtung mindestens eine Recheneinheit zur Verarbeitung von Signalen oder Daten, mindestens eine Speichereinheit zur Speicherung von Signalen oder Daten, mindestens eine Schnittstelle zu einem Sensor oder einem Aktuator, die Sensorsignale aus dem Sensor liest oder Daten oder Steuersignale an den Aktuator ausgibt, und/oder mindestens eine Kommunikationsschnittstelle zum Lesen oder Ausgeben von Daten, die in ein Kommunikationsprotokoll eingebettet sind, umfassen. Die Recheneinheit könnte beispielsweise ein Signalprozessor, ein Mikrocontroller oder dergleichen sein, wobei die Speichereinheit ein Flash-Speicher, ein EEPROM oder eine magnetische Speichereinheit sein könnte. Die Kommunikationsschnittstelle könnte dazu ausgelegt sein, Daten drahtlos und/oder drahtgebunden zu lesen oder auszugeben, wobei eine Kommunikationsschnittstelle, die in der Lage ist, drahtgebundene Daten zu lesen oder auszugeben, beispielsweise Daten elektrisch oder optisch aus einer entsprechenden Datenübertragungsleitung lesen oder an eine entsprechende Datenübertragungsleitung ausgeben könnte.

[0021] Eine Vorrichtung kann als elektrische Einrichtung verstanden werden, die Sensorsignale verarbeitet und abhängig davon Steuer- und/oder Datensignale ausgibt. Die Vorrichtung könnte eine in Hardware und/oder Software realisierte Schnittstelle umfassen. Bei einer Realisierung in Hardware könnten die Schnittstellen beispielsweise Teil eines so genannten System-ASICs sein, das verschiedene Funktionen der Vorrichtung enthält. Die Schnittstellen könnten aber auch einzelne integrierte Schaltungen sein oder zumindest teilweise aus diskreten Bauteilen bestehen. Bei einer Realisierung in Software könnten die Schnittstellen Softwaremodule sein, die zum Beispiel auf einem Mikrocontroller zusammen mit anderen Softwaremodulen realisiert sind.

[0022] Der hier vorgestellte Ansatz stellt auch ein Fahrzeug, insbesondere ein Geländefahrzeug, bereit, das dazu ausgelegt ist, mit granularem Material in einer Arbeitsumgebung zu interagieren, wobei das Fahrzeug die hier vorgestellte Vorrichtung umfasst.

[0023] Vorteilhaft ist auch ein Computerprogrammprodukt oder Computerprogramm mit Programmcode, das auf einem maschinenlesbaren Träger oder Speichermedium, wie einem Halbleiterspeicher, Festplattenspeicher oder optischem Speicher, gespeichert sein kann und zur Ausführung, Implementierung und/oder Steuerung der Schritte des Verfahrens gemäß einer der zuvor beschriebenen Ausführungsformen dient, insbesondere wenn das Programmprodukt oder Programm auf einem Computer oder einer Vorrichtung ausgeführt wird.

[0024] Eine Ausführungsform könnte als vorgelagerter Teil in einer Werkzeugkette für maschinelles Lernen verwendet werden. Sie könnte als Verfahren zum Erzeugen von Trainingsdaten für das Trainieren eines maschinellen Lernsystems dienen, das für die nachfolgend genannten Anwendungen verwendet werden kann. Nachdem ein maschinelles Lernsystem auf diese Weise trainiert wurde, könnte es, wie im Folgenden beschrieben, für nachgelagerte Zwecke eingesetzt werden. Es könnte für die Analyse von Daten eines Geländefahrzeugs und für Bodeninteraktionsaufgaben verwendet werden. Es könnte zum Bestimmen eines kontinuierlichen Wertes oder mehrerer kontinuierlicher Werte verwendet werden, d. h. zur Durchführung einer Regression, z. B. in Bezug auf die Veränderung des Bodens aufgrund von Fahrzeugaktionen. Dabei werden die Ausführungsformen mit einem bestimmten physischen System, nämlich einem Geländefahrzeug, in Verbindung gebracht. Mithilfe der Simulation, die mittels einer Ausführungsform erhalten wurde, kann ein Agent für bestärkendes Lernen trainiert werden, um eine optimale Strategie für ein Geländefahrzeug zu erstellen. So kann ein Steuersignal zum Steuern eines physischen Systems, wie eines Geländefahrzeugs, einer computergesteuerten Maschine, wie eines Roboters, eines Fahrzeugs oder eines anderen Fahrzeugs, das mit dem Boden interagiert, berechnet oder erzeugt werden.

[0025] Ausführungsformen des hier vorgestellten Ansatzes sind in den Zeichnungen dargestellt und werden in der nachfolgenden Beschreibung näher erläutert.

Fig. 1 zeigt eine schematische Ansicht einer Ausführungsform einer Vorrichtung in einem Fahrzeug;

Fig. 2 zeigt ein Flussdiagramm einer Ausführungsform eines Verfahrens zum Bereitstellen von Trainingsdaten für einen Agenten für bestärkendes Lernen zum Erzeugen einer Strategie, die zum Steuern eines Fahrzeugs ausgelegt ist;

Fig. 3 zeigt ein Flussdiagramm einer Ausführungsform eines Verfahrens zum Erzeugen einer Strategie, die zum Steuern eines Fahrzeugs ausgelegt ist;

Fig. 4 zeigt eine schematische Darstellung der Simulation einer Reihe von Aktionen eines Fahrzeugs;

Fig. 5 zeigt ein Flussdiagramm eines Simulationsprozesses gemäß einer Ausführungsform; und

Fig. 6 zeigt ein Flussdiagramm eines Simulationsprozesses gemäß einer Ausführungsform.

[0026] In der nachfolgenden Beschreibung vorteilhafter Ausführungsformen der vorliegenden Erfindung werden ähnlich wirkende Elemente, die in den verschiedenen Figuren dargestellt sind, mit gleichen oder ähnlichen Bezugszahlen bezeichnet, wobei eine wiederholte Beschreibung dieser Elemente unterbleiben soll.

[0027] **Fig. 1** zeigt eine schematische Ansicht einer Ausführungsform einer Vorrichtung 105 in einem Fahrzeug 100. Das Fahrzeug 100 ist dazu ausgelegt, mit granularem Material in einer Arbeitsumgebung zu interagieren. Insbesondere ist das Fahrzeug 100 ein Geländefahrzeug, zum Beispiel eine Planierraupe, ein Bulldozer, ein Verdichter, ein Kipper, ein Bagger, ein Baufahrzeug, eine Schwermaschine, ein Schwerlastfahrzeug oder eine andere Art von Fahrzeug, das eine Anzahl von Aufgaben hat, die eine Interaktion mit der Umgebung, insbesondere mit dem granularen Material, beinhalten. Das Fahrzeug 100 umfasst die Vorrichtung 105. Darüber hinaus umfasst das Fahrzeug 100 mindestens einen Aktuator 102, der eine Aktion des Fahrzeugs 100 bewirkt oder durchführt. Die Vorrichtung 105 und der Aktuator 102 sind zur Signal- oder Datenkommunikation miteinander verbunden.

[0028] Die Vorrichtung 105 ist dazu ausgelegt, Trainingsdaten 125 für einen Agenten 130 für bestärkendes Lernen zum Erzeugen einer Strategie 135 bereitzustellen, die dazu ausgelegt ist, ein Fahrzeug wie das Fahrzeug 100 zu steuern, und/oder dazu ausgelegt ist, eine Strategie 135 zum Steuern des Fahrzeugs 100 zu erzeugen. Mit anderen Worten, die Vorrichtung 105 ist dazu ausgelegt, die Schritte eines Verfahrens gemäß **Fig. 2** und/oder eines Verfahrens gemäß **Fig. 3** in entsprechenden Einheiten auszuführen und/oder zu steuern. Die Vorrichtung 105 umfasst eine Initialisierungseinheit 110 und eine Zuordnungseinheit 120 und gegebenenfalls auch einen Agenten 130 für bestärkendes Lernen oder eine Erzeugungseinheit. Gemäß einer Ausführungsform umfasst die Vorrichtung 105 ferner eine Steuereinheit 140 oder einen Controller.

[0029] Ein Bereitstellungsteil der Vorrichtung 105 umfasst die Initialisierungseinheit 110 und die Zuordnungseinheit 120. Ein Erzeugungsteil der Vorrichtung 105 umfasst den Agenten für bestärkendes Lernen 130 und die optionale Steuereinheit 140. Gemäß der Darstellung in **Fig. 1** sind die Initialisierungseinheit 110, die Zuordnungseinheit 120 und der Agent für bestärkendes Lernen 130 sowie die optionale Steuereinheit 140 als Teile der Vorrichtung 105 realisiert, die sämtlich im Fahrzeug 100 angeordnet sind. Gemäß einer anderen Ausführungsform könnten die Initialisierungseinheit 110 und die Zuordnungseinheit 120, d. h. der Bereitstellungsteil, als ein Teil der Vorrichtung 105 realisiert sein, die physisch außerhalb des Fahrzeugs 100 angeordnet ist.

[0030] Die Initialisierungseinheit 110 ist dazu ausgelegt, ein Umgebungsmodell 115 der das granulare Material enthaltenden Arbeitsumgebung des Fahrzeugs 100 zu initialisieren. Die Initialisierungseinheit 110 ist dazu ausgelegt, das Umgebungsmodell 115 unter Verwendung ausgewählter physikalischer Merkmalswerte IN des granularen Materials und der Arbeitsumgebung zu erstellen. Die ausgewählten physikalischen Merkmalswerte IN werden durch die Initialisierungseinheit 110 aus einer Datenquelle außerhalb der Vorrichtung 105 eingegeben oder geladen. Die ausgewählten physikalischen Merkmalswerte IN werden durch die Initialisierungseinheit 110 als Variablen für mindestens eine Wahrscheinlichkeitsdichtefunktion verwendet. Eine relative Höhenkarte zwischen einem Ist-Höhenstatus und einem Soll-Höhenstatus der das granulare Material enthaltenden Arbeitsumgebung wird durch die Initialisierungseinheit 110 unter Verwendung der Wahrscheinlichkeitsdichtefunktion erzeugt. Die Initialisierungseinheit 110 ist dazu ausgelegt, das Umgebungsmodell 115 an die Zuordnungseinheit 120 auszugeben.

[0031] Gemäß einer Ausführungsform verwendet die Initialisierungseinheit 110 eine kontinuierliche multivariate gaußsche Wahrscheinlichkeitsdichtefunktion als Wahrscheinlichkeitsdichtefunktion. Die ausgewählten physikalischen Merkmalswerte IN könnten insbesondere eine spezifische Masse des granularen Materials, einen Ausbreitungsparameter des granularen Materials, ein Volumen jedes Haufens granularen Materials und eine Position jedes Haufenmittelpunkts in Bezug auf die Umgebung umfassen.

[0032] Die Zuordnungseinheit 120 ist dazu ausgelegt, das Umgebungsmodell 115, das die auf der Wahrscheinlichkeitsdichtefunktion basierende relative Höhenkarte enthält, aus der Initialisierungseinheit 110 zu empfangen oder zu laden. Die Zuordnungseinheit 120 ist dazu ausgelegt, zum Erreichen des Sollstatus jeder Aktion eines Satzes von Aktionen des Fahrzeugs eine Änderung in der Wahrscheinlichkeitsdichtefunktion des Umgebungsmodells 115 und einen Belohnungswert zuzuordnen. Die relative Höhenkarte des Umgebungsmodells 115 und eine Fahrzeugposition werden durch die Zuordnungseinheit 120 für jede Aktion aktualisiert, um die Trainingsdaten 125 bereitzustellen. Mit anderen Worten, die Zuordnungseinheit 120 ist dazu ausgelegt, die Trainingsdaten 125 unter Verwendung des Umgebungsmodells 115 bereitzustellen.

[0033] Gemäß einer Ausführungsform ist die Zuordnungseinheit 120 dazu ausgelegt, die Arbeitsumgebung in eine Vielzahl von Begrenzungsrahmen zu unterteilen und die Änderung der Wahrscheinlichkeitsdichtefunktion individuell für jeden Begrenzungsrahmen zu berechnen. Zusätzlich oder alterna-

tiv ist die Zuordnungseinheit 120 gemäß einer Ausführungsform dazu ausgelegt, einer Aktion unterschiedliche Änderungen und/oder unterschiedliche Belohnungswerte zuzuordnen, je nachdem, ob es sich bei der Aktion um eine Vorwärtstranslation, eine Rückwärtstranslation oder eine Drehung handelt, und/oder je nachdem, ob sich die Fahrzeugposition innerhalb oder außerhalb eines zulässigen Bereichs der Arbeitsumgebung befindet.

[0034] Der Agent für bestärkendes Lernen 130 ist dazu ausgelegt, die Trainingsdaten 125 aus der Zuordnungseinheit 120, d. h. dem Bereitstellungsteil der Vorrichtung 105, zu empfangen oder zu laden. Der Agent 130 für bestärkendes Lernen ist dazu ausgelegt, die Strategie 135 durch bestärkendes Lernen unter Verwendung der Trainingsdaten 125 zu erzeugen. Außerdem ist der Agent 130 für bestärkendes Lernen dazu ausgelegt, die Strategie 135 auszugeben.

[0035] Die optionale Steuereinheit 140 ist dazu ausgelegt, die Strategie 135 aus dem Agenten 130 für bestärkendes Lernen zu empfangen oder zu laden. Die Steuereinheit 140 ist außerdem dazu ausgelegt, anhand der Strategie 135 ein Steuersignal 145 zu bestimmen. Darüber hinaus ist die Steuereinheit 140 dazu ausgelegt, den mindestens einen Aktuator 102 des Fahrzeugs 100 mittels des Steuersignals 145 zu steuern. Zu diesem Zweck ist die Steuereinheit 140 dazu ausgelegt, das Steuersignal 145 an den mindestens einen Aktuator 102 auszugeben. Gemäß einer anderen Ausführungsform könnte die Steuereinheit 140 eine von der Vorrichtung 105 getrennte Einheit sein.

[0036] Fig. 2 zeigt ein Flussdiagramm einer Ausführungsform eines Verfahrens 200 zum Bereitstellen von Trainingsdaten für einen Agenten für bestärkendes Lernen zum Erzeugen einer Strategie, die zum Steuern eines Fahrzeugs, insbesondere eines Geländefahrzeugs, ausgelegt ist. Das Fahrzeug ist dazu ausgelegt, mit granularem Material in einer Arbeitsumgebung zu interagieren. Das Verfahren 200 zum Bereitstellen ist mittels oder in Verbindung mit der Vorrichtung aus Fig. 1 oder einer ähnlichen Vorrichtung ausführbar. Das Verfahren 200 zum Bereitstellen ist ein computer-implementiertes Verfahren. Das Verfahren 200 zum Bereitstellen umfasst einen Schritt 210 des Initialisierens und einen Schritt 220 des Zuordnens.

[0037] Im Schritt 210 des Initialisierens wird ein Umgebungsmodell der das granulare Material enthaltenden Arbeitsumgebung initialisiert. Zum Initialisieren werden ausgewählte physikalische Merkmalswerte des granularen Materials und der Arbeitsumgebung als Variablen mindestens einer Wahrscheinlichkeitsdichtefunktion verwendet. Darüber hinaus wird eine relative Höhenkarte zwischen

einem Ist-Höhenstatus und einem Soll-Höhenstatus der das granulare Material enthaltenden Arbeitsumgebung unter Verwendung der Wahrscheinlichkeitsdichtefunktion erstellt. Anschließend werden im Schritt 220 des Zuordnens eine Änderung der Wahrscheinlichkeitsdichtefunktion des Umgebungsmodells und ein Belohnungswert jeder Aktion eines Satzes von Aktionen des Fahrzeugs zum Erreichen des Sollzustands zugeordnet. Die relative Höhenkarte des Umgebungsmodells und eine Fahrzeugposition werden für jede Aktion aktualisiert, um die Trainingsdaten zu bereitzustellen.

[0038] Fig. 3 zeigt ein Flussdiagramm einer Ausführungsform eines Verfahrens 300 zum Erzeugen einer Strategie, die zum Steuern eines Fahrzeugs, insbesondere eines Geländefahrzeugs, ausgelegt ist. Das Fahrzeug ist dazu ausgelegt, mit granularem Material in einer Arbeitsumgebung zu interagieren. Das Verfahren 300 zum Erzeugen ist mittels oder in Verbindung mit der Vorrichtung aus Fig. 1 oder einer ähnlichen Vorrichtung ausführbar. Das Verfahren 300 zum Erzeugen umfasst einen Schritt 325 des Bereitstellens und einen Schritt 330 des Erzeugens.

[0039] Im Schritt 325 des Bereitstellens werden Trainingsdaten bereitgestellt, indem die Schritte des in Fig. 2 gezeigten Verfahrens zum Bereitstellen oder eines ähnlichen Verfahrens ausgeführt werden. Anschließend wird im Schritt 330 des Erzeugens die Strategie mithilfe eines Agenten für bestärkendes Lernen unter Verwendung der Trainingsdaten erzeugt. Gemäß einer Ausführungsform umfasst das Verfahren 300 zum Erzeugen auch einen Schritt 340 des Steuerns. Im Schritt 340 des Steuerns wird mindestens ein Aktuator des Fahrzeugs mittels eines Steuersignals gesteuert, das unter Verwendung der im Schritt 330 des Erzeugens erzeugten Strategie bestimmt wird.

[0040] Fig. 4 zeigt eine schematische Darstellung der Simulation einer Reihe von Aktionen eines Fahrzeugs. Die Simulation wird durch die in Fig. 1 gezeigte Vorrichtung oder eine ähnliche Vorrichtung und/oder durch Ausführen des in Fig. 2 und/oder Fig. 3 gezeigten Verfahrens oder eines ähnlichen Verfahrens erstellt. Konkret zeigt Fig. 4 acht schematische Darstellungen eines Simulationsbeispiels für ein Fahrzeug 100, in diesem Fall eine Planieraupe, die einen Sandhaufen mit einem Satz von zum Beispiel sieben Aktionen planiert. Fig. 4 zeigt acht Diagramme bzw. Teildarstellungen T=0, T=1, T=2, T=3, T=4, T=5, T=6 und T=7, die jeweils eine relative Höhenkarte eines Umgebungsmodells 115 darstellen, wobei T=0 einen Anfangszustand als initialisiert bezeichnet und die nachfolgenden Diagramme T=1, T=2, T=3, T=4, T=5, T=6 und T=7 nacheinander jeweils eine Aktion des Fahrzeugs 100 zeigen.

[0041] Diese Simulation zeigt die Interaktion zwischen einer Planieraupe als Fahrzeug 100 und dem Boden, wobei die Planieraupe einen Sandhaufen planieren soll. Der hier gezeigte Sandhaufen wird aus sechs nahe beieinander liegenden Teilsandhaufen gebildet, und der endgültige Sandhaufen ist die Summe aller gaußschen PDFs der Sandhaufen. Jedes Diagramm T=1, T=2, T=3, T=4, T=5, T=6 und T=7 stellt einen anderen Zeitpunkt dar, zu dem das Fahrzeug 100 den Sandhaufen bewegt und planiert. Wir gehen davon aus, dass jede Vorwärtsbewegung des Fahrzeugs 100 mit der Schaufel der Planieraupe auf dem Boden erfolgt und den Boden planiert. Wie man sieht, wird bei jeder Vorwärtsbewegung ein weiterer Teil des Bodens planiert, und sobald das Fahrzeug 100 seine Endposition für die aktuelle Bewegung erreicht hat, wird der verbleibende Boden in der Schaufel als gaußsche PDF neu verteilt, siehe zum Beispiel T=2. Es ist anzumerken, dass gemäß anderen Ausführungsformen komplexere Darstellungen verwendet werden können, wie z. B. gaußsche Mischungsmodelle, aber die mathematischen Berechnungen können durch Verwendung einer gaußschen PDF vereinfacht werden, um einerseits einen schnellen Ansatz beizubehalten und andererseits eine gute Darstellung der Bodeninteraktion zu erreichen.

[0042] Unter Bezugnahme auf die zuvor beschriebenen Figuren sollen im Folgenden Ausführungsformen und Vorteile vor dem Hintergrund von Ausführungsformen zusammengefasst und/oder näher erläutert werden, teilweise mit anderen Worten.

[0043] Die Trainingsdaten 125 und/oder die Strategie 135 können für jedes beliebige Fahrzeug 100 verwendet werden, das zum Beispiel mit dem Boden interagiert; der Schwerpunkt der Beschreibung liegt jedoch auf einem bestimmten Fahrzeugtyp, der Sandhaufen auf eine bestimmte Höhe bringen soll. In der Erklärung wird davon ausgegangen, dass die Höhenkarte des Umgebungsmodells 115 in der Simulation eine relative Karte zwischen dem Ist-Sandstatus und der gewünschten endgültigen Sandhöhe ist. Darüber hinaus wird die Aufgabe durch Erstellen eines Rasters und Definieren einer Umrechnung zwischen der Rastereinheit und Metern wie der folgenden Umrechnung vereinfacht:

$$(x, y)_{grd} = \left(x_m * \left[\frac{pixel}{meter} \right], y_m * \left[\frac{pixel}{meter} \right] \right).$$

[0044] Um die Ausführungsformen besser zu verstehen, sollte man die multivariate Gauß-PDF verstehen. Nehmen wir an, dass zwei Variablen (x, y), die in diesem Fall kartesische Koordinaten sind, beide einer Normalverteilung entnommen sind, so lautet die Wahrscheinlichkeitsdichtefunktion für diesen Fall unter der Annahme, dass sie identisch und unabhängig voneinander verteilt sind:

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

[0045] Diese Funktion wird verwendet, um die Höhe des Bodens an jedem (x, y) Punkt zu bestimmen. Man beachte, dass das Integral über diese Funktion V ist und das Volumen des Haufens des granularen Materials darstellt. Die Einheiten für diese Größen sind $V[\text{m}^3]$; $\sigma_x[\text{m}]$; $\sigma_y[\text{m}]$. Da der Sand homogen ist, wird außerdem angenommen, dass die Haufen in allen Richtungen gleichmäßig verteilt sind, weshalb die folgende Beziehung angenommen wird: $\sigma_x = \sigma_y = \sigma$. In diesem Fall reduziert sich die Höhenverteilung

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

[0046] Die maximale Höhe des Haufens beträgt also

$$H_{max} = \frac{V}{2\pi\sigma^2}. \text{ Wenn man außerdem } (x, y) \text{ in } (r, \theta)$$

unter Verwendung von $\frac{x-\mu_x}{\sigma} = r * \cos \theta$ und

$\frac{y-\mu_y}{\sigma} = r * \sin \theta$ ändert, erhält man die Beziehung

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

[0047] Es wird von einer einfachen Mechanik einer Planierraupe als Fahrzeug 100 ausgegangen, wobei es drei mögliche Geschwindigkeitsstufen gibt und eine lineare Beziehung zwischen der Geschwindigkeit der Planierraupe und der Last auf der Schaufel besteht. Die einfache Beziehung zwischen der Fahrzeuggeschwindigkeit und der Last wirkt sich auf die Gesamtzeit aus, die eine Aktion dauert, und wird basierend auf dem Sandvolumen berechnet, das während der gewählten Aktion bewegt wird. Außerdem wird davon ausgegangen, dass die Schaufel ein bekanntes Volumen hat, welches das maximale Volumen ist, das sie in einer bestimmten Zeit planieren kann.

[0048] Zur Initialisierung der Simulation wird der Schritt 210 des Initialisierens des Verfahrens 200 zum Bereitstellen durchgeführt. Die Simulation könnte mit einer Konfigurationsdatei initialisiert werden, in der die folgenden Angaben definiert sind:

spezifische Sandmasse $\left[\frac{\text{kg}}{\text{m}^3}\right]$, Sandausbreitungsparameter (SSP) und für jeden Sandhaufen: Volumen des Sandhaufens $[\text{m}^3]$ und Position des Sandhaufenmittelpunkts $(x, y)[\text{m}, \text{m}]$. Basierend auf diesen Parametern kann die Arbeitsumgebung anhand der folgenden Beziehungen zwischen den gaußschen PDF-Koeffizienten und den Sandpara-

metern definiert werden: $SSP = \frac{r_{max}^2}{V}$, wobei V das Volumen des Sandhaufens $[\text{m}^3]$ ist, r_{max} der Radius des Sandhaufens $[\text{m}]$ ist, mit $r_{max} = 3\sigma$, da 6σ 99,7 % des gesamten Bereichs ausmacht. Da die Eingabe für die Simulation das SSP und das Volumen jedes Sandhaufens ist, ist es möglich, eine Offline-Funktion zu erstellen, die bei Vorgabe dieser beiden Parameter den entsprechenden Ausdruck $g(V, SSP) = (\sigma, H_{max}, r_{max})$ zurückgibt.

[0049] Sobald diese Parameter bekannt sind, kann die anfängliche Umgebung oder das Umgebungsmodell 115 in den folgenden Teilschritten erstellt werden:

- Erstellen einer leeren Höhenkarte, mit der tatsächlichen Größe der Fläche und der Auflösung: $\delta h \left[\frac{\text{meter}}{\text{pixer}} \right]$.
- Für jeden Haufen in der Liste der Haufen:
 - Erstellen eine Karte der Sandhaufenhöhe h_s , unter Verwendung der oben genannten Gleichungen.
 - Addieren der aktuellen Karte der Sandhaufenhöhe zur Karte der vollen Höhe: $h = h + h_s$.
 - Subtrahieren der gewünschten endgültigen Planierhöhe von der aktuellen Höhenkarte, um die relative Höhenkarte zu erstellen.

[0050] Am Ende dieses Prozesses liegt eine Höhenkarte des Bereichs vor, bei der an jedem Punkt (x, y) die Menge des zu entfernenden granularen Materials bekannt ist.

[0051] Um die Trainingsdaten 125 bereitzustellen, ist es dann notwendig, den nächsten Schritt zu berechnen. Um den Schritt in der Simulation zu berechnen, wird im Schritt 220 des Zuordnens des Verfahrens 200 zum Bereitstellen die Veränderung des Bodens aufgrund der Bewegung des Fahrzeugs 100 bestimmt. Im Falle der Planierraupe als Fahrzeug 100 und des Planierauftrags verändert die Bewegung der Planierraupe den Boden, da die Schaufel den Sandhaufen vor sich herschiebt und die Oberfläche nach Möglichkeit abflacht. Diese Änderungen werden modelliert, indem der Bereich, in dem sich das Fahrzeug 100 bewegt, in kleinere Begrenzungsrahmen unterteilt und anschließend die Wirkung der Interaktion auf jeden Bereich separat berechnet wird. In **Fig. 5** und **Fig. 6**, die nachfolgend beschrieben werden, ist die Simulationslogik der Bewegung des Fahrzeugs 100 durch einen Teilschritt seiner Aktion zu sehen.

[0052] In der Simulation umfasst der Aktionsraum für Aktionen des Fahrzeugs 100, insbesondere einer Planierraupe, zwei mögliche Aktionen: (a) vor-

wärts/rückwärts und (b) Drehung, d. h. das neuronale Netz (NN) gibt einen (2x1)-Vektor aus. Ein positiver Wert im ersten Eintrag bedeutet vorwärts, während ein negativer Wert rückwärts bedeutet. Ein positiver Wert im zweiten Eintrag der Aktionsvektoren repräsentiert eine Drehung im Uhrzeigersinn und ein negativer Wert eine Drehung gegen den Uhrzeigersinn. Bei Vorliegen eines Zustands und der entsprechenden bevorzugten Aktion führt der Agent 130 für bestärkendes Lernen diese Aktion aus. Um die Sandbewegung als Ergebnis dieser Aktionen zu verstehen, sollte zwischen Translations- und Rotationsbewegung unterschieden werden.

[0053] Im Falle einer Translation wird eine neue Fahrzeugposition berechnet. Falls die neue Position nicht gültig ist, d. h. außerhalb einer erlaubten Region liegt, erhält man eine große negative Belohnung und beginnt mit einem neuen Szenario. Ist die neue Position gültig, wird eine Ausbreitung vorgenommen. Nun berechnet man die Region innerhalb dieses Simulationsiterationensichtfeldes in Inertialkoordinaten als Hinweis auf diese Schrittgrenzen. Anschließend berechnet man die Höhenkarte innerhalb dieser Region. Der Sand in dieser Region wird ausgebreitet. Der Simulator berechnet dann das Gesamtvolumen dieser inkrementellen Flächeneinheit und lässt die Planierraupe es auf den Mittelwert dieser Flächeneinheit verteilen. Wenn die Translationsbewegung beendet ist und falls die Schaufel mit Sand belegt ist, erzeugt der Simulator, d. h. die entsprechende Einheit der Vorrichtung 105, einen weiteren gaußschen Sandhaufen mit dem verbleibenden Volumen auf der Schaufel. Anschließend beschneidet er ihn, um ein nicht optimales Ausbreitungsmuster zu simulieren. Abschließend berechnet der Simulator die Zeit, die für diese Aktion benötigt wird, um die Belohnung zu erhalten. Rückwärtsfahren oder Umkehren verändert den Sand, d. h. die Höhenkarte, nicht und bringt nur eine negative Belohnung.

[0054] Im Falle einer Drehung wird der Drehwinkel berechnet. Die DCM (Richtungskosinusmatrix) des Fahrzeugs wird aktualisiert, da sie die Fahrtrichtung darstellt. Eine kleine negative Belohnung wird hinzugefügt, da die Rotationsbewegungen die Sandverteilung nicht verändern. Die Planierraupe dreht sich, wenn keine Interaktion zwischen dem Fahrzeug 100 und dem Sand vorhanden ist, sodass während der Drehung keine positive Belohnung erzielt werden kann.

[0055] Fig. 5 zeigt ein Flussdiagramm eines Simulationsprozesses 500 gemäß einer Ausführungsform. Der Simulationsprozess 500 ist durch die Vorrichtung aus Fig. 1 und/oder in Verbindung mit dem Verfahren zum Bereitstellen aus Fig. 2 und/oder dem Verfahren zum Erzeugen aus Fig. 3 ausführbar. Mit anderen

Worten, Fig. 5 zeigt einen Aktionsschritt der Simulation.

[0056] Ein Block 502 stellt eine Eingabe der Fahrzeugstellung, hier eine Planierraupenstellung, der Höhenkarte und der Schaufelstellung dar. Von Block 502 schreitet der Prozess 500 weiter zu einem Block 504, der ein Strategienetzwerk darstellt, und weiter zu einem Block 506, der eine Aktion darstellt, und weiter zu einem Block 508, in dem entschieden wird, ob die Aktion eine Translation ist. Handelt es sich bei der Aktion um eine Translation, schreitet der Prozess 500 von Block 508 weiter zu Block 510, in dem entschieden wird, ob sich das Fahrzeug vorwärts bewegt. Bewegt sich das Fahrzeug vorwärts, wird der Prozess 500 bei Block 512 fortgesetzt. Block 512 enthält die folgenden Anweisungen:

Während (Standort - endgültiger Standort > eps)

1. Erzeugen einer Höhenkarte für das Integrationsflächenelement
2. Berechnen der neuen Höhe dieses Flächenelements
3. Aktualisierte Position = Position + delta_pos
4. Weiter zu 1

[0057] Von Block 512 oder falls die Aktion keine Translation ist, siehe Block 508, oder falls das Fahrzeug nicht vorwärts fährt, siehe Block 510, schreitet der Prozess 500 weiter zu Block 514. Block 514 enthält die folgenden Anweisungen: Abschließen des Schritts:

1. Berechnen der Gesamtzeit des Schritts
2. Aktualisieren der Planierraupen- und Schaufelstellung
3. Aktualisieren der Höhenkarte
4. Berechnen der Belohnung = - Zeit

[0058] Von Block 514 schreitet der Prozess 500 weiter zu Block 516, in dem geprüft wird, ob die simulierte Aktion durchgeführt wird. Diese Prüfung hängt von den folgenden Kriterien ab:

1. Max(Lokale Wärmebildkarte) < eps
2. N_Schritte > N_Schritte_max

[0059] Wird die Aktion durchgeführt, schreitet der Prozess 500 weiter zu einem Block 518, der eine Rücksetzung zum Erzeugen eines neuen Szenarios darstellt. Wird die Aktion nicht durchgeführt, kehrt der Prozess 500 zu Block 502 zurück.

[0060] Fig. 6 zeigt ein Flussdiagramm eines Simulationsprozesses 600 gemäß einer Ausführungsform. Der Simulationsprozess 600 ist durch die Vorrichtung

aus **Fig. 1** und/oder in Verbindung mit dem Verfahren zum Bereitstellen aus **Fig. 2** und/oder dem Verfahren zum Erzeugen aus **Fig. 3** ausführbar. Mit anderen Worten, **Fig. 6** zeigt einen Algorithmus für den oben erwähnten Simulationsschritt oder eine Logik für jeden kleineren Bereich oder Begrenzungsrahmen.

[0061] In Block 602 wird der Fahrzeugweg, hier der Weg der Planierraupe, in kleine Rechtecke in der Karosserieachse unterteilt. Von Block 602 schreitet der Prozess 600 für jedes Rechteck in der Liste der Rechtecke weiter zu Block 604, in dem das Volumen des zu entfernenden Rechtecks berechnet wird. Ist das Volumen größer als null, schreitet der Prozess 600 weiter zu einem Block 606, in dem das neue Volumen in der Schaufel und die neue Bodenhöhe wie folgt berechnet werden:

$$V_{\text{Schaufel}} = V_{\text{Schaufel}} + \text{Volumen}$$

$$\text{Höhe}_{\text{Karte}}(x, y) = \text{Sollhöhe}$$

Ist das Volumen kleiner als null, schreitet der Prozess 600 weiter zu Block 608. Block 608 enthält Folgendes:

Wenn $V_{\text{Schaufel}} > \text{Volumen}$:

$$\text{Höhe}_{\text{Karte}}(x, y) = \text{Volumen}$$

$$V_{\text{Schaufel}} = V_{\text{Schaufel}} + \text{Sollhöhe}$$

Andernfalls:

$$\text{Höhe}_{\text{Karte}}(x, y) = \text{mittleres Volumen in Schaufel}$$

$$V_{\text{Schaufel}} = 0$$

[0062] Von Block 606 und von Block 608 schreitet der Prozess 600 dann weiter zu Block 610, der den Übergang zum nächsten Rechteck darstellt.

[0063] Wann immer eine Ausführungsform eine „und/oder“-Verknüpfung zwischen einem ersten Merkmal und einem zweiten Merkmal enthält, bedeutet dies, dass die Ausführungsform in einer ersten Variante sowohl das erste Merkmal als auch das zweite Merkmal und in einer weiteren Variante entweder das erste Merkmal oder das zweite Merkmal umfasst.

Patentansprüche

1. Computer-implementiertes Verfahren (200) zum Bereitstellen von Trainingsdaten (125) für einen Agenten (130) für bestärkendes Lernen zum Erzeugen einer Strategie (135), die geeignet ist, ein Fahrzeug (100), insbesondere ein Geländefahrzeug (100), zu steuern, wobei das Fahrzeug (100) dazu ausgelegt ist, mit granularem Material in einer

Arbeitsumgebung zu interagieren, wobei das Verfahren (200) die folgenden Schritte umfasst:

Initialisieren (210) eines Umgebungsmodells (115) der Arbeitsumgebung enthaltend das granulare Material, wobei ausgewählte physikalische Merkmalswerte (IN) des granularen Materials und der Arbeitsumgebung als Variablen durch mindestens eine Wahrscheinlichkeitsdichtefunktion verwendet werden, wobei eine relative Höhenkarte zwischen einem Ist-Höhenstatus und einem Soll-Höhenstatus der das granulare Material enthaltenden Arbeitsumgebung unter Verwendung der Wahrscheinlichkeitsdichtefunktion erzeugt wird; und Zuordnen (220) einer Änderung der Wahrscheinlichkeitsdichtefunktion des Umgebungsmodells (115) und eines Belohnungswerts zu jeder Aktion eines Satzes von Aktionen des Fahrzeugs (100) zum Erreichen des Soll-Status, wobei die relative Höhenkarte des Umgebungsmodells (115) und eine Fahrzeugposition für jede Aktion aktualisiert werden, um die Trainingsdaten (125) bereitzustellen.

2. Verfahren (200) gemäß einem der vorhergehenden Ansprüche, wobei in dem Schritt (210) des Initialisierens die Wahrscheinlichkeitsdichtefunktion eine kontinuierliche multivariate gaußsche Wahrscheinlichkeitsdichtefunktion ist und/oder die ausgewählten physikalischen Merkmalswerte eine spezifische Masse des granularen Materials, einen Ausbreitungsparameter des granularen Materials, ein Volumen jedes Haufens von granularem Material und eine Position jedes Haufenmittelpunkts relativ zur Umgebung umfassen.

3. Verfahren (200) gemäß einem der vorhergehenden Ansprüche, wobei im Schritt (220) des Zuordnens die Arbeitsumgebung in eine Mehrzahl von Begrenzungsrahmen unterteilt wird, wobei die Änderung der Wahrscheinlichkeitsdichtefunktion für jeden Begrenzungsrahmen individuell berechnet wird.

4. Verfahren (200) gemäß einem der vorhergehenden Ansprüche, wobei im Schritt (220) des Zuordnens unterschiedliche Änderungen und/oder unterschiedliche Belohnungswerte abhängig davon, ob die Aktion eine Vorwärtstranslation, eine Rückwärtstranslation oder eine Drehung ist und/oder ob die Fahrzeugposition innerhalb oder außerhalb eines erlaubten Bereichs der Arbeitsumgebung liegt, einer Aktion zugeordnet werden.

5. Verfahren (300) zum Erzeugen einer Strategie (135), die dazu ausgelegt ist, ein Fahrzeug (100), insbesondere ein Geländefahrzeug (100), zu steuern, wobei das Fahrzeug (100) dazu ausgelegt ist, mit granularem Material in einer Arbeitsumgebung zu interagieren, wobei das Verfahren (300) die folgenden Schritte umfasst:
Bereitstellen (325) von Trainingsdaten (125) gemäß einem Verfahren (200) gemäß einem der vorhergeh-

enden Ansprüche; und Erzeugen (330) der Strategie (135) mittels eines Agenten (130) für bestärkendes Lernen unter Verwendung der Trainingsdaten (125).

6. Verfahren (300) gemäß Anspruch 5, umfassend einen Schritt (340) des Steuerns mindestens eines Aktuators (102) des Fahrzeugs (100) mittels eines Steuersignals (145), das mittels der Strategie (135) bestimmt wird.

7. Vorrichtung (105), die dazu ausgelegt ist, die Schritte eines Verfahrens (200, 300) gemäß einem der vorhergehenden Ansprüche in entsprechenden Einheiten (110, 120, 130, 140) auszuführen und/oder zu steuern.

8. Fahrzeug (100), insbesondere ein Geländefahrzeug (100), das für die Interaktion mit granularem Material in einer Arbeitsumgebung ausgelegt ist, wobei das Fahrzeug (100) eine Vorrichtung (105) gemäß Anspruch 7 umfasst.

9. Computerprogramm, das dazu ausgelegt ist, die Schritte eines Verfahrens (200, 300) gemäß einem der vorhergehenden Ansprüche auszuführen und/oder zu steuern.

10. Maschinenlesbares Speichermedium, auf dem das Computerprogramm gemäß Anspruch 9 gespeichert ist.

Es folgen 6 Seiten Zeichnungen

Anhängende Zeichnungen

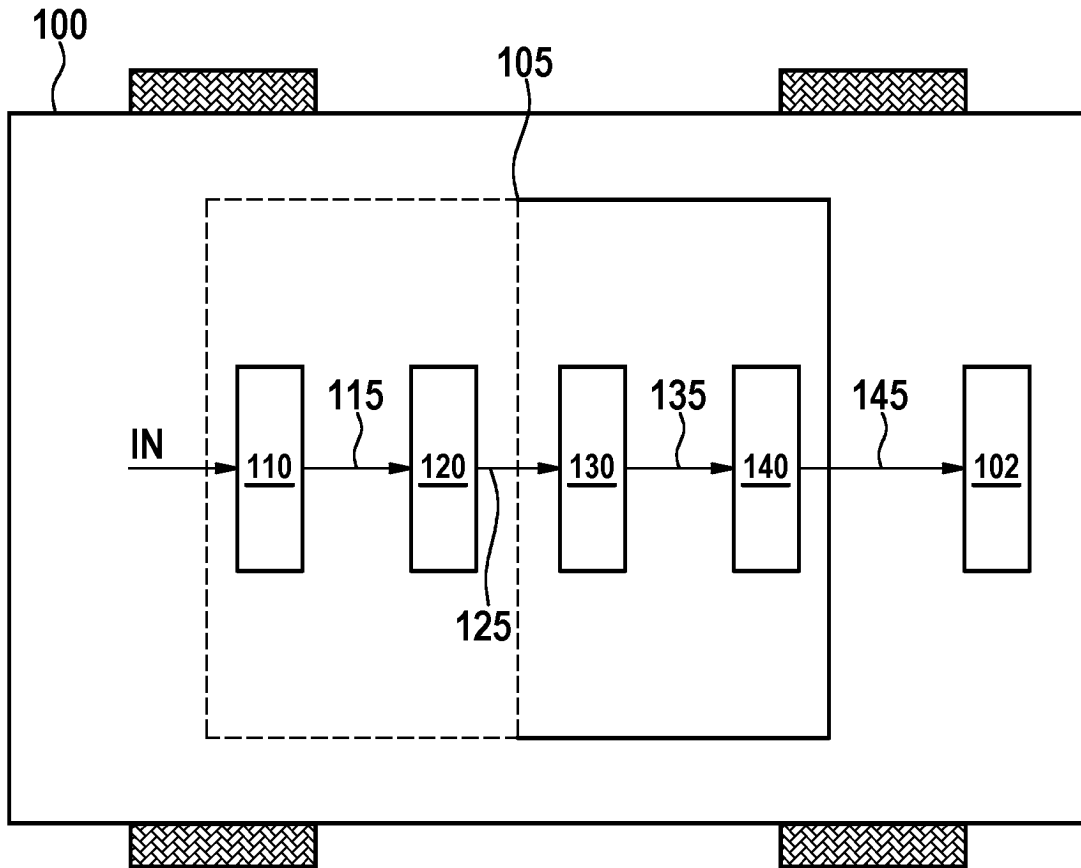


Fig. 1

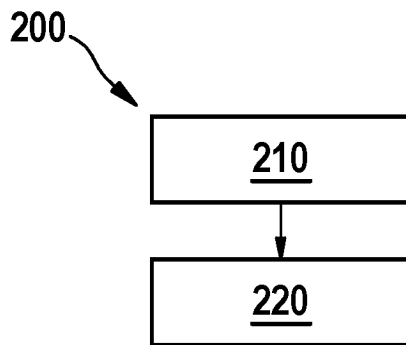


Fig. 2

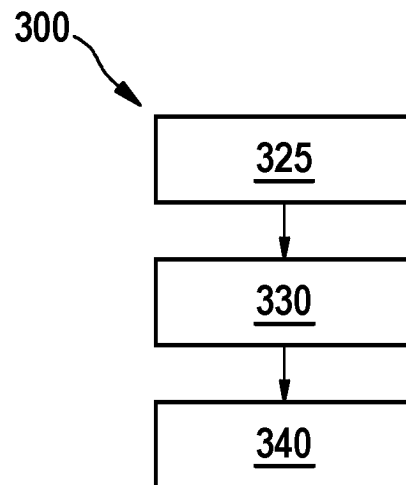


Fig. 3

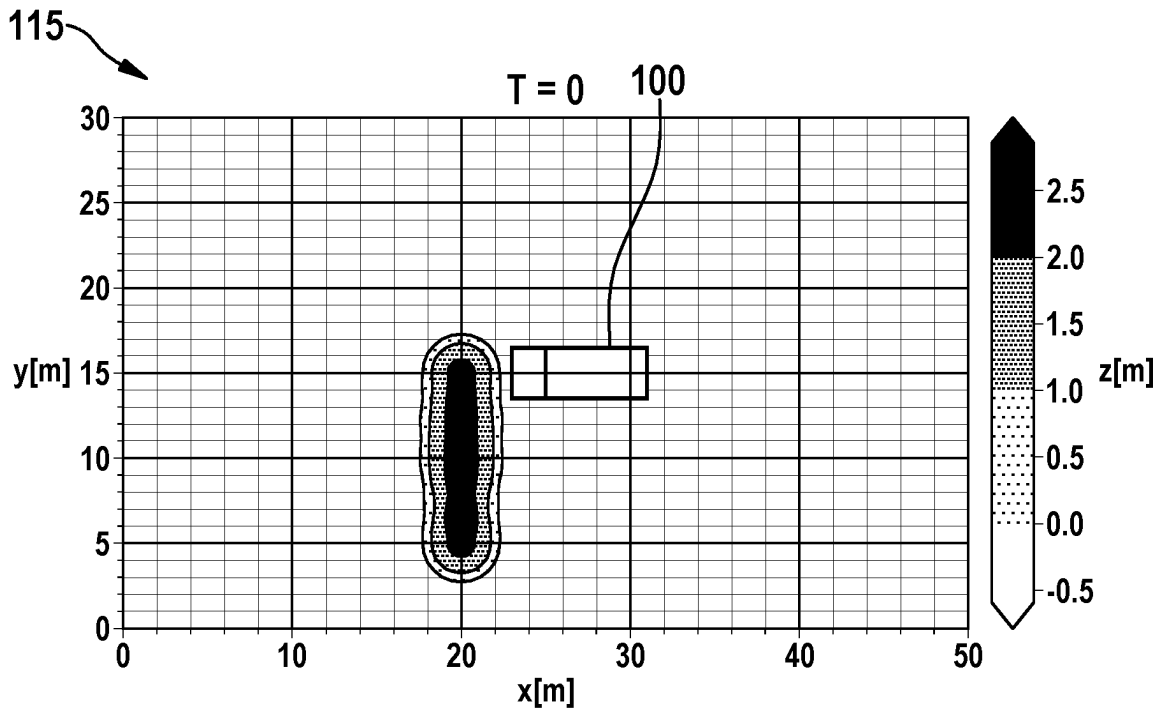


Fig. 4a

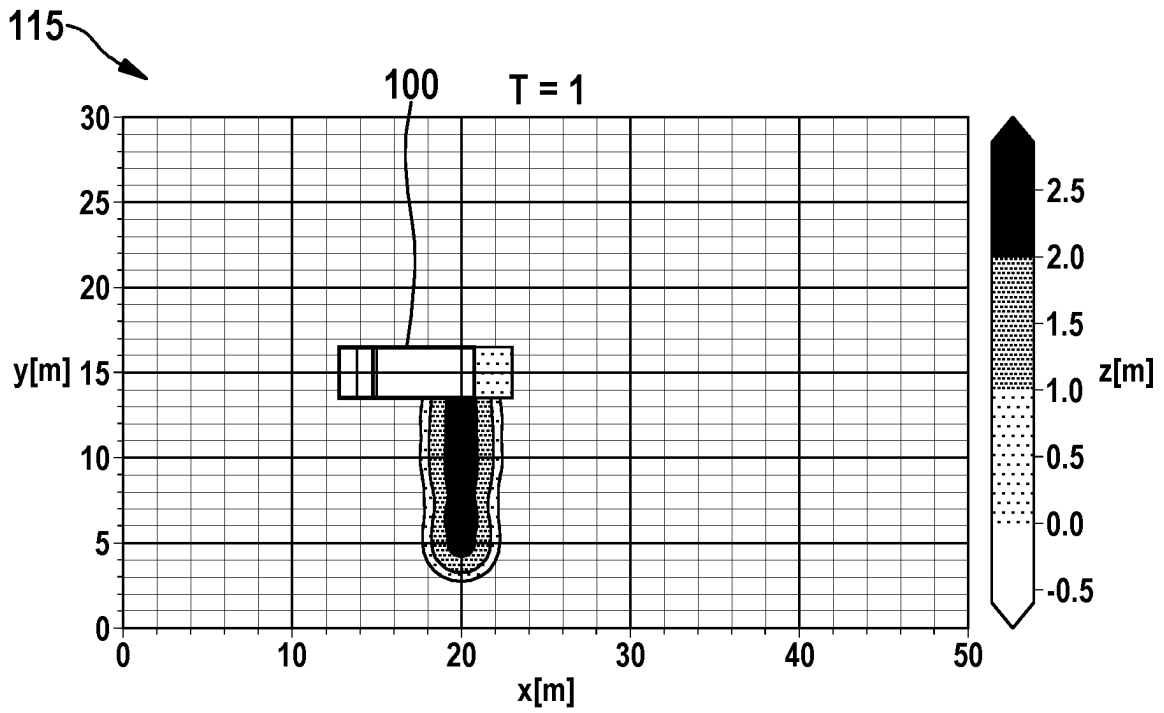


Fig. 4b

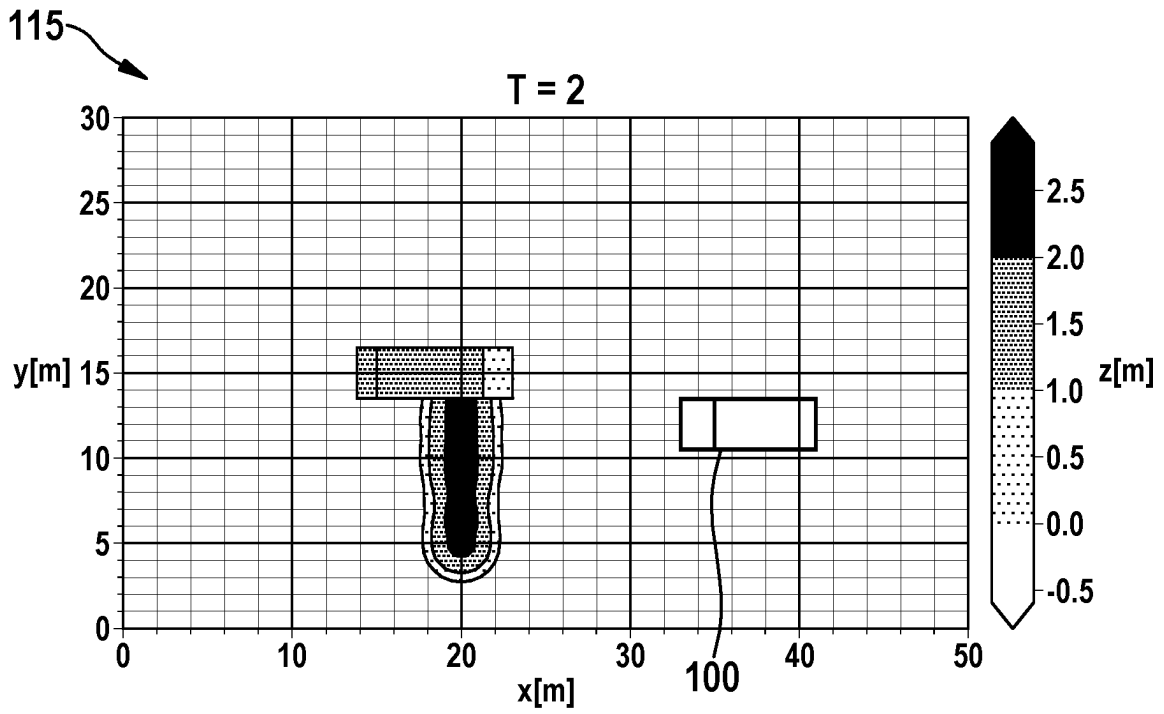


Fig. 4c

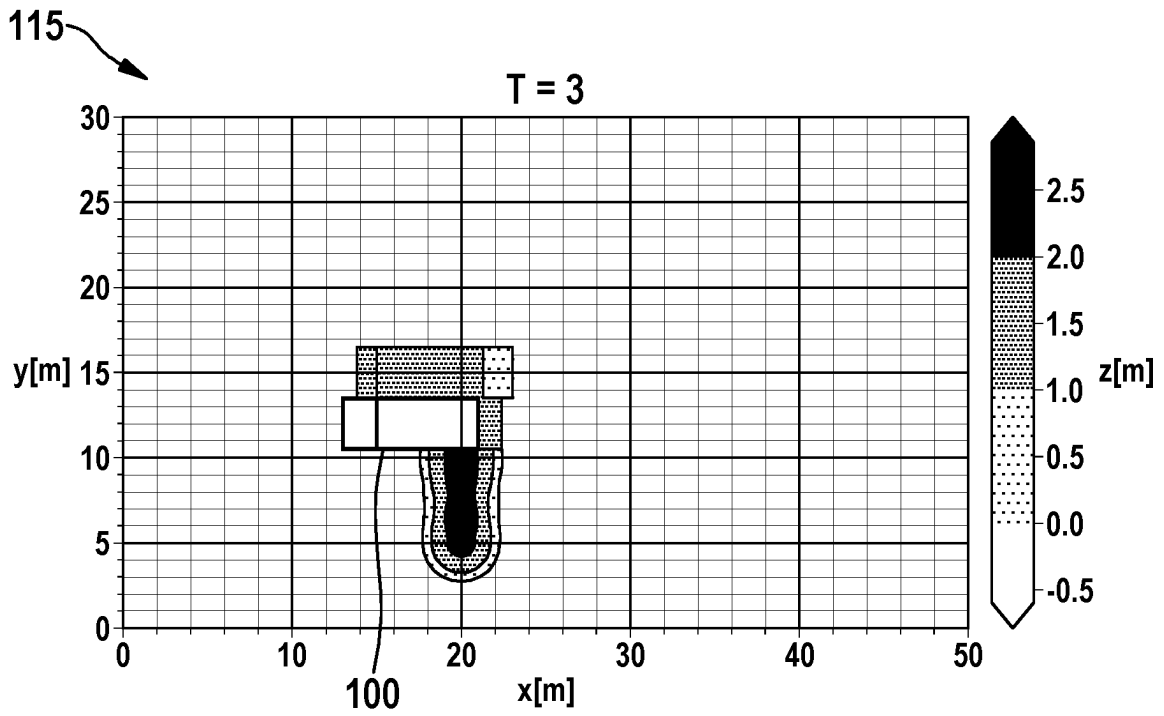


Fig. 4d

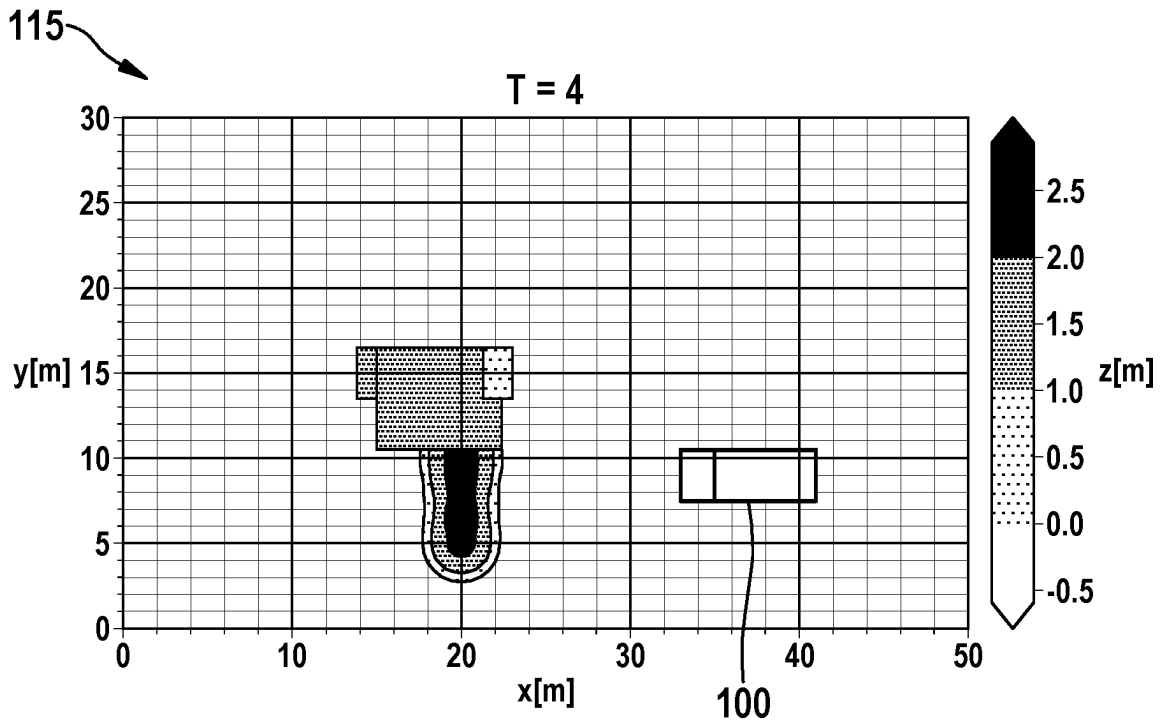


Fig. 4e

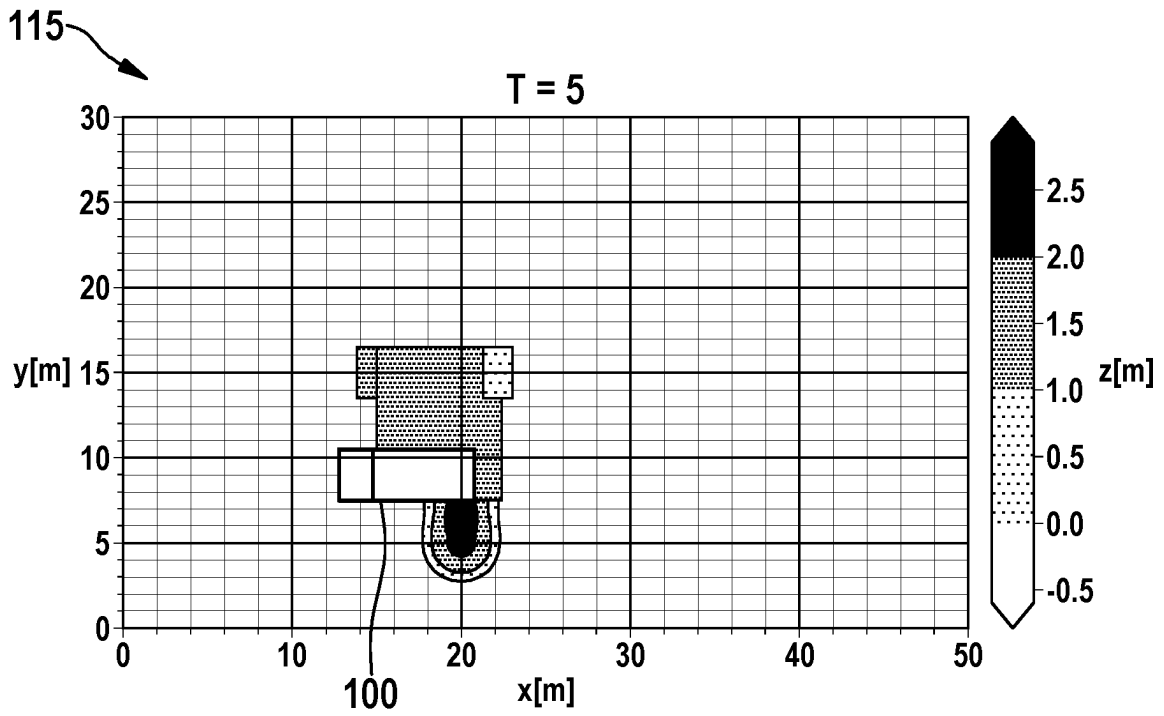


Fig. 4f

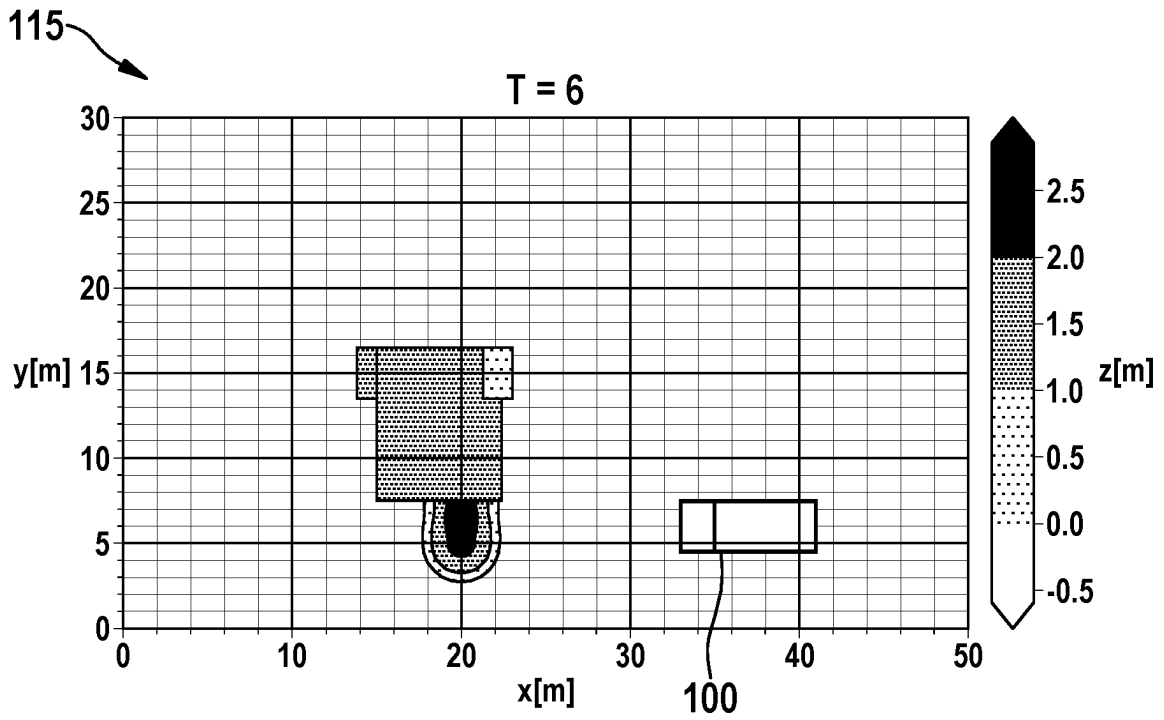


Fig. 4g

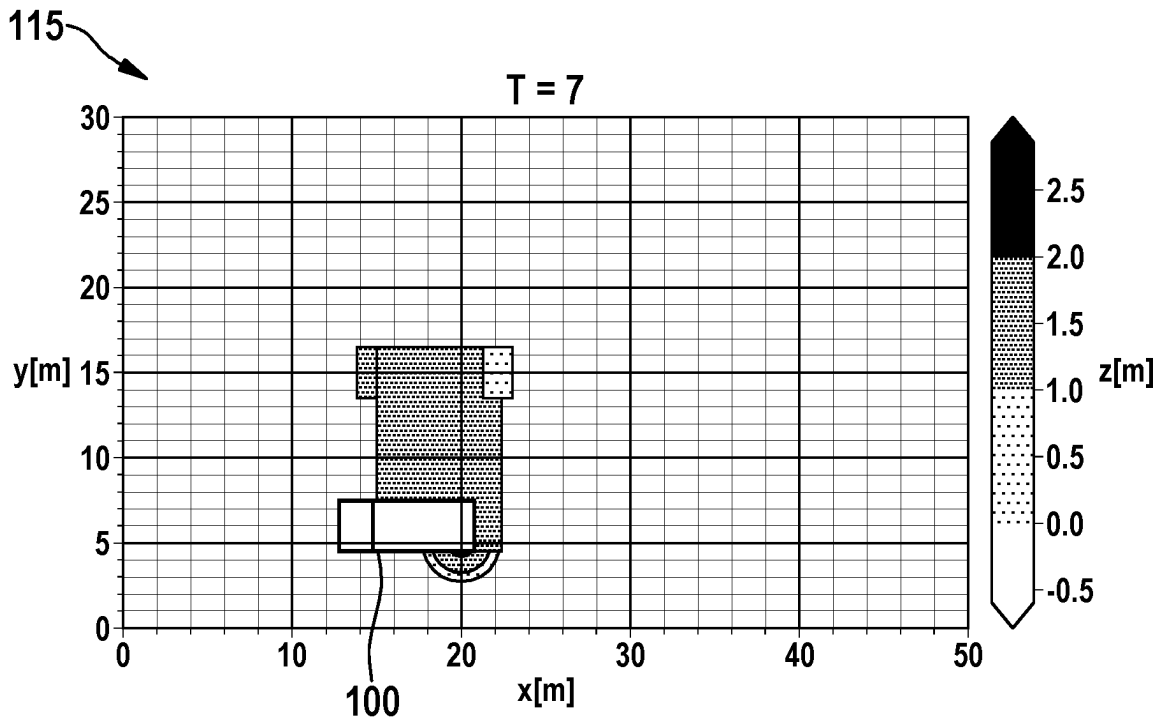


Fig. 4h

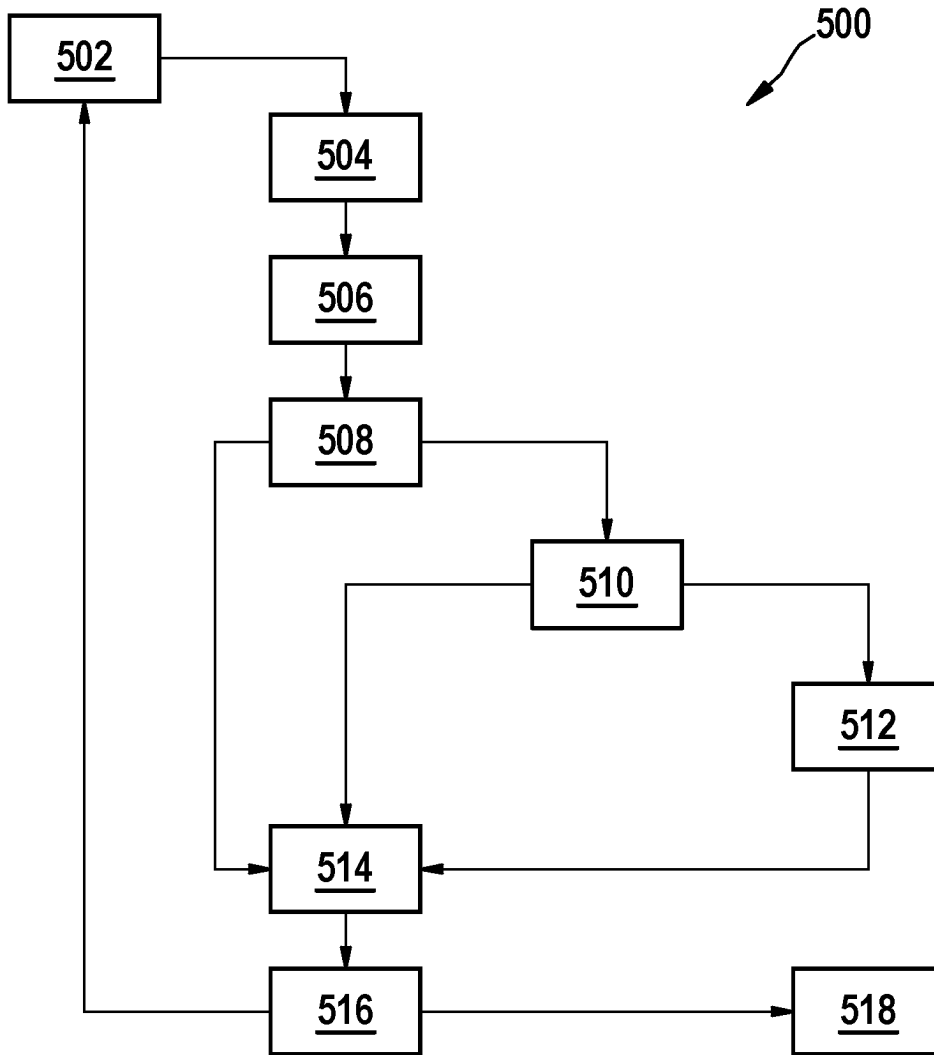


Fig. 5

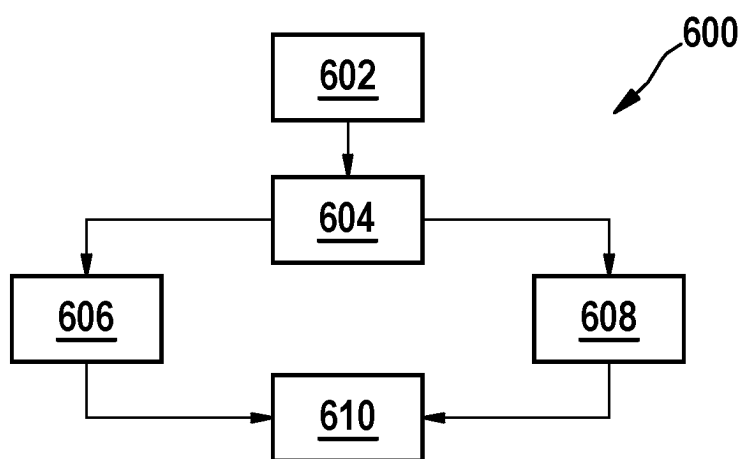


Fig. 6