



(11)

EP 4 035 969 B1

(12)

EUROPÄISCHE PATENTSCHRIFT

(45) Veröffentlichungstag und Bekanntmachung des Hinweises auf die Patenterteilung:
13.11.2024 Patentblatt 2024/46

(51) Internationale Patentklassifikation (IPC):
B61L 15/00 ^(2006.01) **B61L 27/20** ^(2022.01)
B61L 27/60 ^(2022.01)

(21) Anmeldenummer: **21154227.9**

(52) Gemeinsame Patentklassifikation (CPC):
B61L 15/0062; B61L 27/20; B61L 27/60

(22) Anmeldetag: **29.01.2021**

(54) **VERFAHREN ZUM TRAINIEREN EINER STEUERUNG FÜR EIN SCHIENENFAHRZEUG, STEUERUNG UND SCHIENENFAHRZEUG**

CONTROLLER, RAILWAY VEHICLE AND METHOD FOR TRAINING A CONTROLLER FOR A RAILWAY VEHICLE

PROCÉDÉ D'APPRENTISSAGE D'UN DISPOSITIF COMMANDE POUR VÉHICULE FERROVIAIRE, DISPOSITIF DE COMMANDE ET VÉHICULE FERROVIAIRE

(84) Benannte Vertragsstaaten:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR

(43) Veröffentlichungstag der Anmeldung:
03.08.2022 Patentblatt 2022/31

(73) Patentinhaber: **Siemens Mobility GmbH**
81739 München (DE)

(72) Erfinder:
• **Palmer, Andrew**
Spring Hill, 4004 (AU)
• **Weber, Marc Christian**
80339 München (DE)

(74) Vertreter: **Siemens Patent Attorneys**
Postfach 22 16 34
80506 München (DE)

(56) Entgegenhaltungen:
EP-A1- 3 552 921 EP-B1- 3 552 921
DE-A1- 102017 215 802

- **SHUANBAO YAO ETAL: "Optimization design for aerodynamic elements of high speed trains", COMPUTERS AND FLUIDS, PERGAMON PRESS, NEW YORK, NY, GB, vol. 95, 3 March 2014 (2014-03-03), pages 56 - 73, XP028844629, ISSN: 0045-7930, DOI: 10.1016/J.COMPFLUID.2014.02.018**

Anmerkung: Innerhalb von neun Monaten nach Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents im Europäischen Patentblatt kann jedermann nach Maßgabe der Ausführungsordnung beim Europäischen Patentamt gegen dieses Patent Einspruch einlegen. Der Einspruch gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist. (Art. 99(1) Europäisches Patentübereinkommen).

EP 4 035 969 B1

Beschreibung

[0001] Die Erfindung betrifft ein Verfahren zum Trainieren einer Steuerung für ein Schienenfahrzeug. Die Erfindung betrifft ferner eine Steuerung und ein Schienenfahrzeug mit einer Steuerung.

[0002] Für den automatischen Zugbetrieb ist eine Steuerung erforderlich, die eine gewünschte Trajektorie von einer übergeordneten Planungseinheit in die zum Fahren der Trajektorie erforderlichen Beschleunigungs- und Bremsbefehle übersetzen kann. Typischerweise erfordern die verwendeten Steuerungen einen erheblichen Abstimmungsaufwand, um sich an die spezifischen Eigenschaften der einzelnen Fahrzeuge anzupassen.

[0003] Typische Ansätze für Steuerungen basieren auf Proportional-Integral-Derivative (PID)-Reglern und Model Predictive Control (MPC)-Reglern. PID-Regler werden sehr häufig verwendet, da sie sehr einfach zu implementieren sind. Sie erfordern jedoch einen erheblichen Abstimmungsaufwand, wenn sie in einem neuen Fahrzeug eingesetzt werden. Außerdem berücksichtigen sie nicht mehr als den aktuellen Zeitschritt und können daher bei Systemen, bei denen es eine Verzögerung in der Reaktion des Systems auf die Steuereingaben gibt, eine schlechte Leistung aufweisen. In diesen Fällen werden typischerweise MPC-Regler eingesetzt. Die MPC-Steuerung setzt jedoch ein genaues Modell des zu steuernden Systems voraus. Bei komplexen Systemen kann dieses jedoch aufwändig sein beziehungsweise nicht mit ausreichender Präzision erreicht werden.

[0004] Aus der EP 3 552 921 A1 ist eine AUTONOME GESCHWINDIGKEITS-PLANUNG EINES AUF EINEN VORBESTIMMTEN PFAD BESCHRÄNKTEN BEWEGLICHEN AKTEURS bekannt.

[0005] Ferner ist in der DE 10 2017 215 802 A1 ein Fahrerassistenzsystem für Schienenfahrzeuge beschrieben.

[0006] Weiterhin wird über das Design von Hochgeschwindigkeitzügen in SHUANBAO YAO ET AL: "Optimization design for aerodynamic elements of high speed trains", COMPUTERS AND FLUIDS, PERGAMON PRESS, NEW YORK, NY, GB, Bd. 95, 3. März 2014 (2014-03-03), Seiten 56-73, berichtet.

[0007] Der Erfindung liegt die Aufgabe zugrunde, ein verbessertes Verfahren zum Trainieren einer Steuerung für ein Schienenfahrzeug, eine verbesserte Steuerung und ein Schienenfahrzeug mit einer Steuerung bereitzustellen.

[0008] Diese Aufgabe wird erfindungsgemäß durch ein Verfahren zum Trainieren einer Steuerung eines Schienenfahrzeugs, eine entsprechend trainierte Steuerung und ein Schienenfahrzeug mit einer trainierten Steuerung gemäß den unabhängigen Ansprüchen gelöst. Vorteilhafte Ausgestaltungen sind in den Unteransprüchen angegeben.

[0009] Nach einem Aspekt der Erfindung wird ein Verfahren zum Trainieren einer Steuerung eines Schienenfahrzeugs bereitgestellt, wobei das Verfahren umfasst:

- Bereitstellen von Trainingsdaten basierend auf Sensordaten eines Schienenfahrzeugs;
- Trainieren eines Surrogat-Modells des Schienenfahrzeugs basierend auf den Trainingsdaten bezüglich einer Relation zwischen einer Ansteuerung eines Antriebs des Schienenfahrzeugs und einer resultierenden Geschwindigkeit des Schienenfahrzeugs;
- Trainieren einer Aktionsauswahlregel basierend auf den Trainingsdaten und dem Surrogat-Modell unter Verwendung von maschinellem Lernen und unter Berücksichtigung wenigstens eines objektiven Steuerungsziels, wobei die Aktionsauswahlregel Steuerungsanweisungen zum Ansteuern des Antriebs des Schienenfahrzeugs umfasst, die eingerichtet sind, das Schienenfahrzeug aus einem ersten Geschwindigkeitszustand in einen zweiten Geschwindigkeitszustand zu beschleunigen; und
- Generieren einer trainierten Aktionsauswahlregel, wobei die trainierte Aktionsauswahlregel Steuerungsanweisungen umfasst, die eingerichtet sind, das Schienenfahrzeug zu beschleunigen und das Steuerungsziel zu erfüllen.

[0010] Hierdurch kann der technische Vorteil erreicht werden, dass ein verbessertes Verfahren zum Trainieren einer Steuerung eines Schienenfahrzeugs bereitgestellt werden kann. Die Steuerung des Schienenfahrzeugs wird hierbei unter Berücksichtigung von Techniken des Maschinenlernens trainiert. Hierzu werden auf Sensordaten eines Schienenfahrzeugs basierende Trainingsdaten bereitgestellt. Basierend auf den Trainingsdaten wird darauffolgend ein Surrogat-Modell des Schienenfahrzeugs in Bezug auf eine Relation zwischen einer Ansteuerung eines Antriebs des Schienenfahrzeugs und einer hieraus resultierenden Geschwindigkeit des Schienenfahrzeugs trainiert. Darauffolgend wird eine Aktionsauswahlregel basierend auf den Trainingsdaten und dem trainierten Surrogat-Modell mittels Techniken des verstärkenden Lernens und unter Berücksichtigung wenigstens eines Steuerungsziels trainiert und eine trainierte Aktionsauswahlregel generiert. Basierend auf der trainierten Aktionsauswahlregel kann das Schienenfahrzeug gesteuert und das jeweils vorbestimmte Steuerungsziel erreicht werden.

[0011] Ein Surrogat-Modell eines Schienenfahrzeugs ist im Sinne der Anmeldung ein virtuelles Modell eines realen Schienenfahrzeugs, das alle relevanten Merkmale des realen Schienenfahrzeugs darstellt. Ein Surrogat-Modell kann analog zu einem virtuellen Zwilling einer realen Maschine ausgebildet sein und den Betrieb einer realen Maschine bzw. eines Schienenfahrzeugs virtuell simulieren. Das Surrogat-Modell ist somit eingerichtet, ein Verfahren eines Schienenfahrzeugs durch eine entsprechende Ansteuerung zu simulieren. Das Surrogat-Modell kann beispielsweise als ein entsprechend trainiertes neuronales Netz ausgebildet sein, und kann darauf trainiert sein, das Fahrverhalten bzw. die steuerungs-

relevanten Eigenschaften des Schienenfahrzeugs zu simulieren bzw. darzustellen.

[0012] Eine Aktionsauswahlregel im Sinne der Anmeldung ist eine Mehrzahl von Steuerungsanweisungen, die eingerichtet sind, das Schienenfahrzeug aus einem ersten Geschwindigkeitszustand in einen zweiten Geschwindigkeitszustand zu beschleunigen. Die ersten und zweiten Geschwindigkeitszustände können hierbei ein Ist-Zustand, in dem sich das Schienenfahrzeug zu einem bestimmten Zeitpunkt befindet, und ein Soll-Zustand sein, in den das Schienenfahrzeug durch Ansteuern des Antriebs zu überführen ist. Die ersten und zweiten Geschwindigkeitszustände können alternativ hierzu zwei zeitlich nacheinander folgende Zustände des Schienenfahrzeugs sein, in die das Schienenfahrzeug beim Ansteuern durch die Steuerung überführt wird. Durch Ausführen der Steuerungsanweisungen der Aktionsauswahlregel durch die Steuerung kann das jeweilige Schienenfahrzeug somit unter Berücksichtigung des zu erreichenden Steuerungsziels gesteuert werden.

[0013] Steuerungsanweisung sind im Sinne der Anmeldung Anweisung bzw. Befehle zum Steuern des Schienenfahrzeugs. Steuerungsanweisungen können beispielsweise das Beschleunigen oder Abbremsen des Schienenfahrzeugs inklusive detaillierter Ansteuerung des Antriebs umfassen. Darüber hinaus können Steuerungsanweisungen das Schalten in verschiedene Gänge oder in einen Energiesparmodus des Antriebs umfassen. Darüber hinaus können in den Steuerungsanweisungen verschiedene steuerungsrelevante Aspekte in Form entsprechender Anweisungen berücksichtigt sein.

[0014] Steuerungsziele sind im Sinne der Anmeldung Ziele, die durch das Ansteuern der Steuerung zu erreichen sind. Steuerungsziele können beispielsweise eine Geschwindigkeit sein, auf die durch Ansteuern der Steuerung das Schienenfahrzeug beschleunigt werden soll. Steuerungsziele können alternativ oder zusätzlich für verschiedene zukünftige Zeitpunkte definiert sein, beispielsweise über Geschwindigkeitstrajektorien definiert sein. Alternativ können Steuerungsziele einen Energieverbrauch definieren, der bei der Ansteuerung des Schienenfahrzeugs zu erreichen oder nicht zu überschreiten ist. Die Steuerungsziele sind im Sinne der Anmeldung basierend auf den Trainingsdaten definiert, die wiederum basierend auf Sensordaten eines Schienenfahrzeugs generiert sind. Steuerungsziele können hierbei aus den Sensordaten des Schienenfahrzeugs generiert sein. Die Steuerungsziele können sich aus den Sensordaten ergeben und beispielsweise zu erzielende Geschwindigkeiten oder Energieverbräuche beschreiben, gemäß denen das Schienenfahrzeug während der Aufnahme der Sensordaten gesteuert würde. Steuerungsziele können auch während der Ausführung der trainierten Aktionsauswahlregel zur Steuerung des Schienenfahrzeugs durch übergeordnete Systeme, beispielsweise entsprechende Planungsmodule zur Steuerung des Schienenfahrzeugs, definiert oder geändert werden.

[0015] Nach einer Ausführungsform ist das maschinelle

Lernen als bestärkendes Lernen ausgebildet.

[0016] Hierdurch kann der technische Effekt erreicht werden, dass ein effizientes Trainieren der Aktionsauswahlregel ermöglicht ist.

[0017] Bestärkendes Lernen (Reinforcement Learning) ist im Sinne der Anmeldung ein Bereich des maschinellen Lernens, der sich mit dem Trainieren von operationellen Teilnehmern beschäftigt, gewünschte Aktionen auszuführen, um somit den Teilnehmer aus einem Ist-Zustand in einen gewünschten Soll-Zustand zu überführen. Das Trainieren des Teilnehmers wird hierbei unter Berücksichtigung eines Steuerungsziels getätigt, das durch die Ausführung der Handlung des Teilnehmers zu erreichen ist. Das Training kann ein Maximieren einer Belohnungsfunktion berücksichtigen, durch die dem Teilnehmer die ausführende Handlung und das zu erreichende Ziel dargestellt ist.

[0018] Nach einer Ausführungsform umfasst das Trainieren: Randomisiertes Abändern des wenigstens einen auf den Trainingsdaten basierenden objektiven Steuerungsziels und Definieren von abgeänderten Steuerungszielen; und Trainieren der Aktionsauswahlregel in Bezug auf Erfüllung der abgeänderten Steuerungsziele.

[0019] Hierdurch kann der technische Vorteil erreicht werden, dass ein präziseres Training der Steuerung des Schienenfahrzeugs bereitgestellt werden kann. Durch das randomisierte Abändern des wenigstens einen Steuerungsziels und das damit verbundene Definieren von abgeänderten Steuerungszielen können die Trainingsdaten zum Trainieren der Steuerung effektiver ausgenutzt werden. Durch das Abändern der Steuerungsziele und das Generieren bzw. Definieren von abgeänderten Steuerungszielen können Steuerungsziele definiert werden, die durch die Trainingsdaten gestützt, jedoch nicht auf diese beschränkt sind.

[0020] Die geänderten Steuerungsziele ergeben somit Steuerungsziele, die nicht auf Trainingsdaten basieren, sondern Steuerungsziele darstellen, die über die Steuerungsziele hinausgehen, gemäß denen das Schienenfahrzeug während der Aufnahme der Sensordaten angesteuert wurde. Durch das Trainieren der Aktionsauswahlregel basierend auf den abgeänderten Steuerungszielen kann eine verbesserte trainierte Aktionsauswahlregel generiert werden, die Steuerungsanweisungen für eine erhöhte Anzahl verschiedener Steuerungsziele definiert. Hierdurch kann eine verbesserte Steuerung bereitgestellt werden.

[0021] Nach einer Ausführungsform umfasst das Trainieren der Aktionsauswahlregel ein Maximieren einer Belohnungsfunktion, wobei die Belohnungsfunktion für eine Aktionsauswahlregel maximal ist, die das objektive Steuerungsziel und/oder die abgeänderten Steuerungsziele erfüllt.

[0022] Hierdurch kann der technische Vorteil erreicht werden, dass ein möglichst präzises Training der Aktionsauswahlregel gemäß den Techniken des bestärkenden Lernens erreicht werden kann.

[0023] Nach einer Ausführungsform berücksichtigt die

Belohnungsfunktion eine Differenz zwischen einem durch Ausführen einer Steuerungsaktion der Aktionsauswahlregel erzielten Geschwindigkeitszustand und dem objektiven Steuerungsziel und/oder den abgeänderten Steuerungszielen.

[0024] Hierdurch kann der technische Vorteil erreicht werden, dass ein effizientes Training der Aktionsauswahlregel und eine präzise trainierte Aktionsauswahlregel bereitgestellt werden kann, die präzise das zu erreichende Steuerungsziel erfüllt.

[0025] Nach einer Ausführungsform wird das Maximieren der Belohnungsfunktion durch ein künstliches neuronales Netz ausgeführt.

[0026] Hierdurch kann der technische Vorteil erreicht werden, dass ein effizientes Training der Aktionsauswahlregel bzw. der Steuerung des Schienenfahrzeugs bereitgestellt werden kann.

[0027] Nach einer Ausführungsform umfasst das Steuerungsziel eine Sollgeschwindigkeit des Schienenfahrzeugs und/oder einen Sollenergieverbrauch und/oder eine Sollbeschleunigung und/oder ein verschleißarmes Beschleunigungs- und/oder Bremsverhalten.

[0028] Hierdurch kann der technische Vorteil erreicht werden, dass eine effizient trainierte Aktionsauswahlregel und damit verbunden eine effizient trainierte Steuerung eines Schienenfahrzeugs bereitgestellt werden kann. Die derart trainierte Aktionsauswahlregel umfasst hierbei Steuerungsanweisungen, die geeignet sind, das Schienenfahrzeug unter Berücksichtigung der genannten Steuerungsziele zu steuern.

[0029] Nach einer Ausführungsform werden die Trainingsdaten während eines Verfahrens des Schienenfahrzeugs aufgenommen und umfassen Sensordaten von Zustandsvariablen, Steuerungsaktionen und Geschwindigkeitstrajektorien, wobei die Zustandsvariablen Geschwindigkeitsdaten, Beschleunigungsdaten, Ortsdaten, Spezifikationsdaten des Antriebs und/oder des Schienenfahrzeugs umfassen, wobei die Steuerungsaktionen Antriebs- und/oder Bremsbetätigungen umfassen, und wobei die Geschwindigkeitstrajektorien entsprechende zeitliche Geschwindigkeitsentwicklungen des Schienenfahrzeugs beschreiben.

[0030] Hierdurch kann der technische Vorteil erreicht werden, dass durch umfassende Trainingsdaten ein präzises Training der Steuerung ermöglicht ist.

[0031] Nach einem zweiten Aspekt der Erfindung wird eine Steuerung für ein Schienenfahrzeug bereitgestellt, wobei die Steuerung nach einem Verfahren zum Trainieren einer Steuerung eines Schienenfahrzeugs nach einer der voranstehenden Ausführungsformen trainiert ist, und wobei die Steuerung eingerichtet ist, das Schienenfahrzeug unter Ausführung der trainierten Aktionsauswahlregel zu steuern.

[0032] Hierdurch kann eine verbesserte Steuerung für ein Schienenfahrzeug bereitgestellt werden, die unter Verwendung von Methoden des maschinellen Lernens, insbesondere des bestärkenden Lernens, trainiert ist. Die Steuerung weist hierzu eine trainierte Aktionsaus-

wahlregel auf, die gemäß dem erfindungsgemäßen Verfahren zum Trainieren einer Steuerung eines Schienenfahrzeugs gemäß den oben genannten Ausführungsformen trainiert ist. Basierend auf der trainierten Aktionsauswahlregel, die eine Mehrzahl von Steuerungsanweisungen zum Steuern des Schienenfahrzeugs unter Berücksichtigung verschiedener Steuerungsziele aufweist, kann das Schienenfahrzeug zum Erreichen der jeweiligen Steuerungsziele angesteuert werden. Durch das Trainieren der Aktionsauswahlregel basierend auf dem Surrogat-Modell des Schienenfahrzeugs kann die Steuerung für beliebige Schienenfahrzeuge, die dem Surrogat-Modell entsprechen, trainiert werden, sodass bei einem neuen Schienenfahrzeug eine aufwändige Anpassung des zur Steuerung verwendeten Modells einer Steuerung, wie dies bei einem Model Predictive Control (MPC)-Regler notwendig ist, entfallen kann. Hierdurch kann eine präzise, zuverlässige und variable Steuerung für Schienenfahrzeuge bereitgestellt werden.

[0033] Nach einem dritten Aspekt wird ein Schienenfahrzeug mit einer Steuerung nach einer der voranstehenden Ausführungsformen bereitgestellt.

[0034] Hierdurch kann ein Schienenfahrzeug mit einer verbesserten Steuerung mit den obengenannten Vorteilen bereitgestellt werden.

[0035] Nach einem vierten Aspekt wird ein Computerprogrammprodukt umfassend Befehle bereitgestellt, die bei der Ausführung des Programms durch eine Datenverarbeitungseinheit diese veranlassen, das Verfahren zum Trainieren einer Steuerung für ein Schienenfahrzeug nach einer der voranstehenden Ausführungsformen auszuführen.

[0036] Die oben beschriebenen Merkmale und Vorteile dieser Erfindung sowie die Art und Weise, wie diese erreicht werden, werden klarer und deutlicher verständlich durch die Erläuterungen der folgenden, stark vereinfachten, schematischen Darstellungen bevorzugter Ausführungsbeispiele. Hierbei zeigen:

FIG 1 ein Flussdiagramm eines Verfahrens zum Trainieren einer Steuerung eines Schienenfahrzeugs gemäß einer Ausführungsform;

FIG 2 eine schematische Darstellung eines Schienenfahrzeugs mit einer Steuerung nach einer Ausführungsform; und

FIG 3 eine schematische Darstellung eines Computerprogrammprodukts.

[0037] FIG 1 zeigt ein Flussdiagramm eines Verfahrens 100 zum Trainieren einer Steuerung 200 eines Schienenfahrzeugs 201 gemäß einer Ausführungsform.

[0038] Zum Trainieren der Steuerung 200 des Schienenfahrzeugs 201 werden zunächst in einem ersten Verfahrensschritt 101 Trainingsdaten bereitgestellt. Die Trainingsdaten basieren hierbei auf Sensordaten eines Schienenfahrzeugs 201 und umfassen in der gezeigten

Ausführungsform Zustandsvariablen 217, Steuerungsaktionen 219 und Geschwindigkeitstrajektorien 221. Die Trainingsdaten 203, insbesondere die Sensordaten des Schienenfahrzeugs 201, können beispielsweise während eines Fahrens des Schienenfahrzeugs 201 oder eines vergleichbaren Schienenfahrzeugs durch eine entsprechende Sensorik aufgenommen sein.

[0039] Zustandsvariablen 217 beschreiben im Sinne der Anmeldung Punkte innerhalb eines Zustandsraums, der verschiedene Zustände des Schienenfahrzeugs 201 beschreibt. Insbesondere können die Zustandsvariablen 217 Geschwindigkeitsdaten, Beschleunigungsdaten und/oder Ortsdaten des Schienenfahrzeugs 201 umfassen, die während des Verfahrens des Schienenfahrzeugs 201 oder des vergleichbaren Schienenfahrzeugs aufgenommen wurden. Das vergleichbare Schienenfahrzeug kann beispielsweise ein Schienenfahrzeug identischen Typs sein. Darüber hinaus können die Zustandsvariablen 217 Spezifikationsdaten des Antriebs und/oder des Schienenfahrzeugs 201 umfassen, die beispielsweise die Art des Antriebs, maximale Leistung bzw. Drehzahl und andere Parameter des Antriebs oder eine Größe bzw. ein Gewicht des Schienenfahrzeugs 201 umfassen.

[0040] Steuerungsaktionen 219 sind im Sinne der Anmeldung Aktionen, die während des Verfahrens des Schienenfahrzeugs 201 zum Steuern des Schienenfahrzeugs 201 durch die Steuerung 200 ausgeführt werden. Steuerungsaktionen 219 können beispielsweise die Betätigung des Gaspedals oder des Steuerungshebels eines Schienenfahrzeugs bzw. die Betätigung der Bremse des Schienenfahrzeugs 201 umfassen.

[0041] Geschwindigkeitstrajektorien 221 sind im Sinne der Anmeldung durch den Zustandsraum verlaufende Spuren von zeitlich aufeinander folgenden Geschwindigkeitswerten, die während des Verfahrens des Schienenfahrzeugs 201 erreicht wurden.

[0042] Nach Bereitstellen der Trainingsdaten 203 wird in einem weiteren Verfahrensschritt 103 ein Surrogat-Modell 205 des Schienenfahrzeugs 201 generiert bzw. trainiert und dem Surrogat-Modell 205 eine Relation zwischen Ansteuerungen eines Antriebs 207 des Schienenfahrzeugs 201 und einer hieraus resultierenden Geschwindigkeit des Schienenfahrzeugs 201 gelernt.

[0043] Basierend auf den Trainingsdaten 203, die während des Verfahrens des Schienenfahrzeugs 201 aufgenommen wurden, wird dem Surrogat-Modell 205, das eine virtuelle Kopie des Schienenfahrzeugs 201 darstellt, mittels Methoden des Maschinenslernens trainiert, welche Steuerungsaktionen 219, Betätigung des Gaspedals oder Betätigung der Bremse, zu welchen Endgeschwindigkeiten des Schienenfahrzeugs 201, die basierend auf den entsprechenden Zustandsvariablen 217 ermittelt werden, führen. Das derart trainierte Surrogat-Modell 205 ermöglicht somit eine Simulation der Steuerung 200 des Schienenfahrzeugs 201, bei der durch Ausführung entsprechender Steuerungsfunktionen 219 entsprechende Geschwindigkeitstrajektorien 221 erzielt

werden können.

[0044] In einem weiteren Verfahrensschritt 105 wird eine Aktionsauswahlregel 209 basierend auf den Trainingsdaten 203 und dem trainierten Surrogat-Modell 205 unter Verwendung von Methoden des bestärkenden Lernens und unter Berücksichtigung wenigstens eines objektiven Steuerungsziels 211 trainiert. Die Aktionsauswahlregel 209 umfasst hierbei Steuerungsanweisungen zum Steuern des Antriebs 207 des Schienenfahrzeugs 201. Die Steuerungsanweisungen sind hierbei eingerichtet, das Schienenfahrzeug 201 aus einem ersten Geschwindigkeitszustand in einen zweiten Geschwindigkeitszustand zu beschleunigen.

[0045] Die Steuerungsanweisungen können beispielsweise den Steuerungsaktionen 219 entsprechen, die als Trainingsdaten 203 während des Verfahrens des Schienenfahrzeugs 201 aufgenommen wurden, und das Beschleunigen bzw. Abbremsen des Schienenfahrzeugs umfassen. Die Aktionsauswahlregel 209 kann hierzu eine Mehrzahl von Steuerungsanweisungen aufweisen, mittels denen der Antrieb 207 des Schienenfahrzeugs 201 angesteuert werden kann. Die Steuerungsanweisungen sind hierbei derart ausgebildet, dass während des Ansteuerns des Antriebs 207 des Schienenfahrzeugs 201 das objektive Steuerungsziel 211 erreicht wird. Das objektive Steuerungsziel 211 kann hierbei beispielsweise eine zu erreichende Endgeschwindigkeit, ein gewünschter Energieverbrauch oder eine maximale Beschleunigung des Schienenfahrzeugs 201 sein, die jeweils während des Ansteuerns des Schienenfahrzeugs 201 zu erreichen bzw. einzuhalten sind. Das objektive Steuerungsziel 211 kann ebenfalls durch die Trainingsdaten 203 bereitgestellt sein. Beispielsweise kann das Steuerungsziel 211 durch die aufgezeichneten Geschwindigkeitstrajektorien 221 der Trainingsdaten 203 dargestellt sein.

[0046] Die ersten und zweiten Geschwindigkeitszustände können hierbei jeweils ein Ist-Zustand bzw. ein Soll-Zustand des Schienenfahrzeugs 201 sein, wobei das Schienenfahrzeug 201 durch das Ansteuern gemäß den Steuerungsanweisungen der Aktionsauswahlregel 209 aus dem Ist-Zustand in den Soll-Zustand zu überführen ist. Alternativ hierzu können die ersten und zweiten Geschwindigkeitszustände zwei zeitlich nacheinander eintretende Zustände des Zustandsraums sein, in die das Schienenfahrzeug 201 durch Ansteuern gemäß den Steuerungsanweisungen der Aktionsauswahlregel zu überführen ist.

[0047] Zum Trainieren der Aktionsauswahlregel 209 durch bestärkendes Lernen wird eine beliebig ausgewählte Aktionsauswahlregel mit beliebigen Steuerungsanweisungen basierend auf den Trainingsdaten 203 inklusive der Zustandsvariablen 217 des Zustandsraums des Schienenfahrzeugs 201 und unter Berücksichtigung des jeweils ausgewählten Steuerungsziels 211 trainiert bzw. optimiert, sodass die Aktionsauswahlregel 209 eingerichtet ist, das ausgewählte Steuerungsziel 211 zu erreichen. In der gezeigten Ausführungsform wird das Trai-

ning der beliebig gewählten Aktionsauswahlregel 209 in einem weiteren Verfahrensschritt 111 durch ein Maximieren einer entsprechend eingerichteten Belohnungsfunktion ausgeführt. Die Belohnungsfunktion kann hierbei beispielsweise eine Differenz zwischen einem durch Ausführen einer Steuerungsaktion 219 der Aktionsauswahlregel 209 erzielten Geschwindigkeitszustand und dem objektiven Steuerungsziel 211 definiert sein. Das Training der Aktionsauswahlregel 209 erfolgt somit dadurch, dass die Steuerungsanweisungen bzw. die Aktionsauswahlregel 209 derart modifiziert werden, dass die entsprechend definierte Belohnungsfunktion einen maximalen Wert erreicht. Eine Aktionsauswahlregel 209 mit einer maximalen Belohnungsfunktion ist demzufolge in der Lage, das ausgewählte Steuerungsziel 211 zu erreichen. Durch das derartige Trainieren der Aktionsauswahlregel 209 kann der durch die Trainingsdaten 203 definierte Zustandsraum des Schienenfahrzeugs 201, in dem verschiedene Zustände des Schienenfahrzeugs 201 angeführt sind, durchquert werden, um die optimale Aktionsauswahlregel 209 zu bestimmen, die eingerichtet ist, durch Ausführung der entsprechenden Steuerungsanweisungen das Schienenfahrzeug 201 in einer optimierten Trajektorie in Zustände zu überführen, die eine optimale Steuerung 200 und das Erreichen des vorbestimmten objektiven Steuerungsziels 211 gewährleisten.

[0048] Gemäß der gezeigten Ausführungsform wird zum Trainieren der Aktionsauswahlregel 209 in einem Verfahrensschritt 109 das objektive Steuerungsziel 211 randomisiert geändert und geänderte Steuerungsziele 215 generiert. Durch das randomisierte Ändern des Steuerungsziels 211 können geänderte Steuerungsziele 215 generiert werden, die von den Trainingsdaten 203 abweichen können.

[0049] Beispielsweise kann das objektive Steuerungsziel 211 durch eine Geschwindigkeitstrajektorie 221 der Trainingsdaten 203 gebildet sein. Die jeweilige Geschwindigkeitstrajektorie 221 kann hierbei durch die Sensordaten des Schienenfahrzeugs 201 gestützt sein, die während des Verfahrens des Schienenfahrzeugs 201 aufgenommen wurden. Durch Verändern einzelner Werte der Geschwindigkeitstrajektorie 221 können somit abgeänderte Geschwindigkeitstrajektorien 221 als abgeänderte Steuerungsziele 215 generiert werden, wobei die abgeänderten Geschwindigkeitstrajektorien 221 nicht vollständig durch die Trainingsdaten 203 gestützt sind und Geschwindigkeitswerte aufweisen, die während des Verfahrens des Schienenfahrzeugs 201 und der Aufnahme der jeweiligen Geschwindigkeitswerte des Schienenfahrzeugs 201 abweichen. Durch das Ändern der Geschwindigkeitstrajektorie 221 können somit Punkte im Zustandsraum des Schienenfahrzeugs 201 erreicht werden, für die keine expliziten Trainingsdaten 203 generiert wurden.

[0050] Alternativ hierzu können auch mehrere Steuerungsziele 211, 215 beim Training der Aktionsauswahlregel 209 berücksichtigt werden, sodass die Aktionsauswahlregel 209 eingerichtet ist, eine Mehrzahl von Steu-

erungszielen 211, 215 zu erfüllen. Die Steuerungsziele können hierbei neben der zu erreichenden Endgeschwindigkeit des Schienenfahrzeugs 201 beispielsweise ein Energieverbrauch des Schienenfahrzeugs 201 oder eine maximal zulässige Beschleunigung des Schienenfahrzeugs 201 umfassen, die während der Steuerung 200 des Schienenfahrzeugs 201 zu erfüllen bzw. zu berücksichtigen sind.

[0051] Nach dem Trainieren der Aktionsauswahlregel 209 im Verfahrensschritt 105 wird basierend auf der Ausgangs-Aktionsauswahlregel 209 eine entsprechend trainierte Aktionsauswahlregel 213 generiert, die Steuerungsanweisungen umfasst, die eingerichtet sind, das Schienenfahrzeug 201 zu beschleunigen und das Steuerungsziel 211 bzw. die abgeänderten Steuerungsziele 215 zu erfüllen.

[0052] Das Trainieren der Aktionsauswahlregel 209 bzw. das Maximieren der Belohnungsfunktion kann durch eine trainierte künstliche Intelligenz, beispielsweise durch ein entsprechend trainiertes neuronales Netz, durchgeführt werden. Durch Maximieren der entsprechend eingerichteten Belohnungsfunktion kann somit eine Aktionsauswahlregel 213 generiert werden, die die entsprechenden Steuerungsziele 211, 215 erfüllt und somit eine optimierte Steuerung 200 des Schienenfahrzeugs 201 gewährleistet.

[0053] FIG 2 zeigt eine schematische Darstellung eines Schienenfahrzeugs 201 mit einer Steuerung 200 nach einer Ausführungsform.

[0054] Fig. 2 zeigt ein Schienenfahrzeug 201 mit einer Steuerung 200, wobei die Steuerung 200 eine gemäß dem erfindungsgemäßen Verfahren 100 zum Trainieren einer Steuerung 200 eines Schienenfahrzeugs 201 trainierte Aktionsauswahlregel 213 umfasst. Das Schienenfahrzeug 201 umfasst ferner einen Antrieb 207 und einen Sensor 223. Über den Sensor 223 können Sensorwerte des Antriebs 207 aufgenommen werden, um eine Ansteuerung des Antriebs 207 mittels der Steuerung 200 zu verfolgen. Das Schienenfahrzeug 201 umfasst ferner eine Planungseinheit 225, die mit der Steuerung 200 verbunden ist. Zur Ansteuerung des Schienenfahrzeugs 201 kann somit die Steuerung 200 von der Planungseinheit 225 eine entsprechende Geschwindigkeitstrajektorie 221 empfangen, die einen geplanten Geschwindigkeitsverlauf des Verfahrens des Schienenfahrzeugs 201 beschreibt. Durch Ausführen der durch die trainierte Aktionsauswahlregel 213 definierten Steuerungsanweisungen kann somit die Steuerung 200 die entsprechenden Steuerungsanweisungen ausführen, die geeignet sind, das Schienenfahrzeug 201 gemäß der geplanten Geschwindigkeitstrajektorie 221 der Planungseinheit 225 zu beschleunigen. Die Steuerung 200 des Schienenfahrzeugs 201 kann hierbei verschiedene Steuerungsziele 211, 215 berücksichtigen. Beispielsweise kann das Schienenfahrzeug 201 unter einem vorbestimmten maximalen Energieverbrauch angesteuert werden. Alternativ oder zusätzlich hierzu kann das Schienenfahrzeug 201 unter Berücksichtigung einer maximal zulässigen

Beschleunigung angesteuert werden. Indem die entsprechend trainierte Aktionsauswahlregel 213 für eine Vielzahl verschiedener Steuerungsziele 211, 215 entsprechende Steuerungsanweisungen umfasst, die geeignet sind, das Schienenfahrzeug 201 unter Berücksichtigung der jeweiligen Steuerungsziele 211, 215 anzusteuern, können während des Betriebs des Schienenfahrzeugs 201 und insbesondere bei bereits installierter Steuerung 200 entsprechende Steuerungsziele 211, 215 abgeändert werden, die während des Steuerns des Schienenfahrzeugs 201 erreicht werden sollen. Eine Änderung der Steuerung 200 und insbesondere eine Anpassung der jeweiligen definierten Steuerungsanweisungen ist aufgrund des Trainings der trainierten Aktionsauswahlregel 213 nicht erforderlich.

[0055] FIG 3 zeigt eine schematische Darstellung eines Computerprogrammprodukts 300.

[0056] Figur 3 zeigt ein Computerprogrammprodukt 300, umfassend Befehle, die bei der Ausführung des Programms durch eine Recheneinheit dieses veranlassen, das Verfahren 100 nach einer der oben genannten Ausführungsformen auszuführen. Das Computerprogrammprodukt 300 ist in der gezeigten Ausführungsform auf einem Speichermedium 301 gespeichert. Das Speichermedium 301 kann hierbei ein beliebiges aus dem Stand der Technik bekanntes Speichermedium sein.

[0057] Obwohl die Erfindung im Detail durch das bevorzugte Ausführungsbeispiel näher illustriert und beschrieben wurde, so ist die Erfindung nicht durch die offenbarten Beispiele eingeschränkt und andere Variationen können vom Fachmann hieraus abgeleitet werden, ohne den Schutzbereich der Erfindung zu verlassen, der durch die Ansprüche definiert ist.

Patentansprüche

1. Verfahren (100) zum Trainieren einer Steuerung (200) eines Schienenfahrzeugs (201), umfassend:

- Bereitstellen (101) von Trainingsdaten (203) basierend auf Sensordaten eines Schienenfahrzeugs (201);
- Trainieren (103) eines Surrogat-Modells (205) des Schienenfahrzeugs (201) basierend auf den Trainingsdaten (203) bezüglich einer Relation zwischen einer Ansteuerung eines Antriebs (207) des Schienenfahrzeugs (201) und einer resultierenden Geschwindigkeit des Schienenfahrzeugs (201);
- Trainieren (105) einer Aktionsauswahlregel (209) basierend auf den Trainingsdaten (203) und dem Surrogat-Modell (205) unter Verwendung von maschinellem Lernen und unter Berücksichtigung wenigstens eines objektiven Steuerungsziels (211), wobei die Aktionsauswahlregel (209) Steuerungsanweisungen zum Ansteuern des Antriebs (207) des Schienen-

fahrzeugs (201) umfasst, die eingerichtet sind, das Schienenfahrzeug (201) aus einem ersten Geschwindigkeitszustand in einen zweiten Geschwindigkeitszustand zu beschleunigen; und
- Generieren (107) einer trainierten Aktionsauswahlregel (213), wobei die trainierte Aktionsauswahlregel (213) Steuerungsanweisungen umfasst, die eingerichtet sind, das Schienenfahrzeug (201) zu beschleunigen und das Steuerungsziel (211) zu erfüllen.

2. Verfahren (100) nach Anspruch 1, wobei das maschinelle Lernen als bestärkendes Lernen ausgebildet ist.

3. Verfahren (100) nach Anspruch 1 oder 2, wobei das Trainieren (105) der Aktionsauswahlregel (209) umfasst:

Randomisiertes Abändern (109) des wenigstens einen auf den Trainingsdaten (203) basierenden objektiven Steuerungsziels (211) und Definieren von abgeänderten Steuerungszielen (215); und
Trainieren der Aktionsauswahlregel (209) in Bezug auf Erfüllung der abgeänderten Steuerungsziele (215).

4. Verfahren (100) nach Anspruch 3, wobei das Trainieren (105) der Aktionsauswahlregel (209) umfasst: Maximieren (111) einer Belohnungsfunktion, wobei die Belohnungsfunktion für eine Aktionsauswahlregel (209) maximal ist, die das objektive Steuerungsziel (211) und/oder die abgeänderten Steuerungsziele (215) erfüllt.

5. Verfahren (100) nach Anspruch 4, wobei die Belohnungsfunktion eine Differenz zwischen einem durch Ausführen einer Steuerungsanweisung der Aktionsauswahlregel (209) erzielten Geschwindigkeitszustand und dem objektiven Steuerungsziel (211) und/oder den abgeänderten Steuerungszielen (215) berücksichtigt.

6. Verfahren (100) nach Anspruch 1 oder 2, wobei das Trainieren (105) der Aktionsauswahlregel (209) umfasst:
Maximieren (111) einer Belohnungsfunktion, wobei die Belohnungsfunktion für eine Aktionsauswahlregel (209) maximal ist, die das objektive Steuerungsziel (211) erfüllt.

7. Verfahren (100) nach Anspruch 6, wobei die Belohnungsfunktion eine Differenz zwischen einem durch Ausführen einer Steuerungsanweisung der Aktionsauswahlregel (209) erzielten Geschwindigkeitszustand und dem objektiven Steuerungsziel (211) berücksichtigt.

8. Verfahren (100) nach einem der Ansprüche 4 bis 7, wobei das Maximieren (111) der Belohnungsfunktion durch ein künstliches neuronales Netz ausgeführt wird.
9. Verfahren (100) nach einem der voranstehenden Ansprüche 3, 4 oder 5, wobei das Steuerungsziel (211) und/oder die abgeänderten Steuerungsziele (215) eine Sollgeschwindigkeit des Schienenfahrzeugs (201) und/oder einen Sollenergieverbrauch und/oder eine Sollbeschleunigung und/oder ein verschleißarmes Beschleunigungs- und/oder Bremsverhalten umfasst.
10. Verfahren (100) nach einem der voranstehenden Ansprüche 6, 7 oder 8, wobei das Steuerungsziel (211) eine Sollgeschwindigkeit des Schienenfahrzeugs (201) und/oder einen Sollenergieverbrauch und/oder eine Sollbeschleunigung und/oder ein verschleißarmes Beschleunigungs- und/oder Bremsverhalten umfasst.
11. Verfahren (100) nach einem der voranstehenden Ansprüche, wobei die Trainingsdaten (203) während eines Verfahrens des Schienenfahrzeugs (201) aufgenommen werden und Sensordaten von Zustandsvariablen (217), Steuerungsaktionen (219) und Geschwindigkeitstrajektorien (221) umfassen, wobei die Zustandsvariablen (217) Geschwindigkeitsdaten, Beschleunigungsdaten, Ortsdaten, Spezifikationsdaten des Antriebs und/oder des Schienenfahrzeugs (201) umfassen, wobei die Steuerungsaktionen (219) Antriebs- und/oder Bremsbetätigungen umfassen, und wobei die Geschwindigkeitstrajektorien (221) entsprechende zeitliche Geschwindigkeitsentwicklungen des Schienenfahrzeugs (201) beschreiben.
12. Steuerung (200) für ein Schienenfahrzeug (201), wobei die Steuerung (200) nach einem Verfahren (100) zum Trainieren einer Steuerung (200) eines Schienenfahrzeugs (201) nach einem der voranstehenden Ansprüche 1 bis 11 trainiert ist, und wobei die Steuerung (200) eingerichtet ist, das Schienenfahrzeug (201) unter Ausführung der trainierten Aktionsauswahlregel (213) zu steuern.
13. Schienenfahrzeug (201) mit einer Steuerung (200) nach Anspruch 12.
14. Computerprogrammprodukt (300) umfassend Befehle, die bei der Ausführung des Programms durch eine Datenverarbeitungseinheit diese veranlassen, das Verfahren (100) nach einem der voranstehenden Ansprüche 1 bis 11 auszuführen.

Claims

1. Method (100) for training a controller (200) of a rail vehicle (201) including:
 - providing (101) training data (203) based on sensor data for a rail vehicle (201);
 - training (103) a surrogate model (205) of the rail vehicle (201) based on the training data (203) with respect to a relationship between an actuation of a drive (207) of the rail vehicle (201) and a resulting speed of the rail vehicle (201);
 - training (105) an action selection rule (209) based on the training data (203) and the surrogate model (205) using machine learning and taking into account at least one objective control target (211), wherein the action selection rule (209) includes control instructions for actuating the drive (207) of the rail vehicle (201) configured to accelerate the rail vehicle (201) from a first speed state into a second speed state; and
 - generating (107) a trained action selection rule (213), wherein the trained action selection rule (213) includes control instructions configured to accelerate the rail vehicle (201) and to meet the control target (211).
2. Method (100) according to claim 1, wherein the machine learning is embodied as reinforcement learning.
3. Method (100) according to claim 1 or 2, wherein the training (105) of the action selection rule (209) includes:
 - randomised modification (109) of the at least one objective control target (211) based on the training data (203) and the definition of modified control targets (215); and
 - training the action selection rule (209) in terms of meeting the modified control targets (215).
4. Method (100) according to claim 3, wherein the training (105) of the action selection rule (209) includes: maximising (111) a reward function, wherein the reward function is maximum for an action selection rule (209) that meets the objective control target (211) and/or the modified control targets (215).
5. Method (100) according to claim 4, wherein the reward function takes account of a difference between a speed state achieved by executing a control instruction of the action selection rule (209) and the objective control target (211) and/or the modified control targets (215).
6. Method (100) according to claim 1 or 2, wherein the training (105) of the action selection rules (209) in-

cludes: maximising (111) a reward function, wherein the reward function is maximum for an action selection rule (209) that meets the objective control target (211).

7. Method (100) according to claim 6, wherein the reward function takes into account a difference between a speed state achieved by executing a control instruction of the action selection rule (209) and the objective control target (211).
8. Method (100) according to one of claims 4 to 7, wherein the maximisation (111) of the reward function is executed by an artificial neural network.
9. Method (100) according to one of the preceding claims 3, 4 or 5, wherein the control target (211) and/or the modified control targets (215) include a target speed of the rail vehicle (201) and/or a target energy consumption and/or a target acceleration and/or low-wear acceleration behaviour and/or braking behaviour.
10. Method (100) according to one of the preceding claims 6, 7 or 8, wherein the control target (211) includes a target speed of the rail vehicle (201) and/or a target energy consumption and/or a target acceleration and/or low-wear acceleration behaviour and/or braking behaviour.
11. Method (100) according to one of the preceding claims, wherein the training data (203) is recorded while the rail vehicle (201) is moving and includes sensor data for state variables (217), control actions (219) and speed trajectories (221), wherein the state variables (217) include speed data, acceleration data, location data, specification data for the drive and/or the rail vehicle (201), wherein the control actions (219) include drive and/or brake actuations, and wherein the speed trajectories (221) describe corresponding temporal speed developments of the rail vehicle (201).
12. Controller (200) for a rail vehicle (201), wherein the controller (200) is trained in accordance with a method (100) for training a controller (200) of a rail vehicle (201) according to one of the preceding claims 1 to 11, and wherein the controller (200) is configured to control the rail vehicle (201) by executing the trained action selection rule (213).
13. Rail vehicle (201) with a controller (200) according to claim 12.
14. Computer program product (300) including commands which, when the program is executed by a data processing unit, cause it to execute the method (100) according to one of the preceding claims 1 to

11.

Revendications

1. Procédé (100) d'apprentissage d'un dispositif (200) de commande d'un véhicule (201) ferroviaire, comprenant :
 - se procurer (101) des données (203) d'apprentissage reposant sur des données de capteur d'un véhicule (201) ferroviaire ;
 - faire subir un apprentissage (103) à un modèle (205) de substitution du véhicule (201) ferroviaire sur la base des données (203) d'apprentissage en ce qui concerne une relation entre une commande d'un entraînement (207) du véhicule (201) ferroviaire et une vitesse qui s'ensuit du véhicule (201) ferroviaire ;
 - faire subir un apprentissage (105) à une règle (209) de sélection d'action sur la base des données (203) d'apprentissage et du modèle (205) de substitution en utilisant de l'enseignement automatique et en tenant compte d'au moins une cible (211) de commande objective, dans lequel la règle (209) de sélection d'action comprend des instructions de commande pour la commande de l'entraînement (207) du véhicule (201) ferroviaire, qui sont agencées pour accélérer le véhicule (201) ferroviaire d'un premier état de vitesse à un deuxième état de vitesse ; et
 - création (107) d'une règle (213) de sélection d'action ayant subi un apprentissage, dans lequel la règle (213) de sélection d'action ayant subi un apprentissage comprend des instructions de commande, qui sont agencées pour accélérer le véhicule (201) ferroviaire et satisfaire la cible (211) de commande.
2. Procédé (100) suivant la revendication 1, dans lequel l'apprentissage automatique est constitué sous la forme d'un apprentissage par renforcement.
3. Procédé (100) suivant la revendication 1 ou 2, dans lequel l'apprentissage (105) de la règle (209) de sélection d'action comprend :
 - modification (109) aléatoire de la cible (211) de commande objective reposant sur les données (203) d'apprentissage et définition de cibles (215) de commande modifiées ;
 - et
 - faire subir un apprentissage à la règle (209) de sélection d'action par rapport à la satisfaction des cibles (215) de commande modifiées.
4. Procédé (100) suivant la revendication 3, dans lequel l'apprentissage (105) de la règle (209) de sé-

lection d'action comprend :

rendre maximum (111) une fonction de rétribution, dans lequel la fonction de rétribution est maximum pour une règle (209) de sélection d'action, qui satisfait la cible (211) de commande objective et/ou les cibles (215) de commande modifiées.

5. Procédé (100) suivant la revendication 4, dans lequel la fonction de rétribution tient compte d'une différence entre un état de vitesse atteint en exécutant une instruction de commande de la règle (209) de sélection d'action et la cible (211) de commande objective et/ou les cibles (215) de commande modifiées. 5
6. Procédé (100) suivant la revendication 1 ou 2, dans lequel l'apprentissage (105) subi par la règle (209) de sélection d'action comprend :
rendre maximum (111) une fonction de rétribution, dans lequel la fonction de rétribution est maximum pour une règle (209) de sélection d'action, qui satisfait la cible (211) de commande objective. 10
7. Procédé (100) suivant la revendication 6, dans lequel la fonction de rétribution tient compte d'une différence entre un état de vitesse obtenu en exécutant une instruction de commande de la règle (209) de sélection d'action et la cible (211) de commande objective. 15
8. Procédé (100) suivant l'une des revendications 4 à 7, dans lequel on rend maximum (111) la fonction de rétribution par un réseau neuronal artificiel. 20
9. Procédé (100) suivant l'une des revendications 3, 4 ou 5 précédentes, dans lequel la cible (211) de commande et/ou les cibles (215) de commande modifiées comprennent une vitesse de consigne du véhicule (201) ferroviaire et/ou une consommation d'énergie de consigne et/ou une accélération de consigne et/ou un comportement d'accélération et/ou de freinage avec peu d'usure. 25
10. Procédé (100) suivant l'une des revendications 6, 7 ou 8 précédentes, dans lequel la cible (211) de commande comprend une vitesse de consigne du véhicule (201) ferroviaire et/ou une consommation d'énergie de consigne et/ou une accélération de consigne et/ou un comportement en accélération et/ou en freinage avec peu d'usure. 30
11. Procédé (100) suivant l'une des revendications précédentes, dans lequel on enregistre les données (203) d'apprentissage pendant un déplacement du véhicule (201) ferroviaire et des données de capteur comprennent des variables (217) d'état, des actions (219) de commande et des trajectoires (221) de vitesse, dans lequel les variables (217) d'état com- 35

prennent des données de vitesse, des données d'accélération, des données de localisation, des données de spécification de l'entraînement et/ou du véhicule (201) ferroviaire, dans lequel les actions (219) de commande comprennent des actionnements d'entraînement et/ou de frein, et dans lequel les trajectoires (221) de vitesse décrivent des déroulements de vitesse correspondants dans le temps du véhicule (201) ferroviaire.

12. Dispositif (200) de commande d'un véhicule (201) ferroviaire, dans lequel le dispositif (200) de commande subit un apprentissage par un procédé (100) d'apprentissage d'un dispositif (200) de commande d'un véhicule (201) ferroviaire suivant l'une des revendications 1 à 11 précédentes, et dans lequel le dispositif (200) de commande est agencé pour commander le véhicule (201) ferroviaire en exécutant la règle (213) de sélection d'action ayant subi un apprentissage. 40
13. Véhicule (201) ferroviaire comprenant un dispositif (200) de commande suivant la revendication 12. 45
14. Produit (300) de programme d'ordinateur, comprenant des instructions qui, lors de l'exécution du programme par une unité de traitement de données, fait que celle-ci exécute le procédé (100) suivant l'une des revendications 1 à 11 précédentes. 50

FIG 1

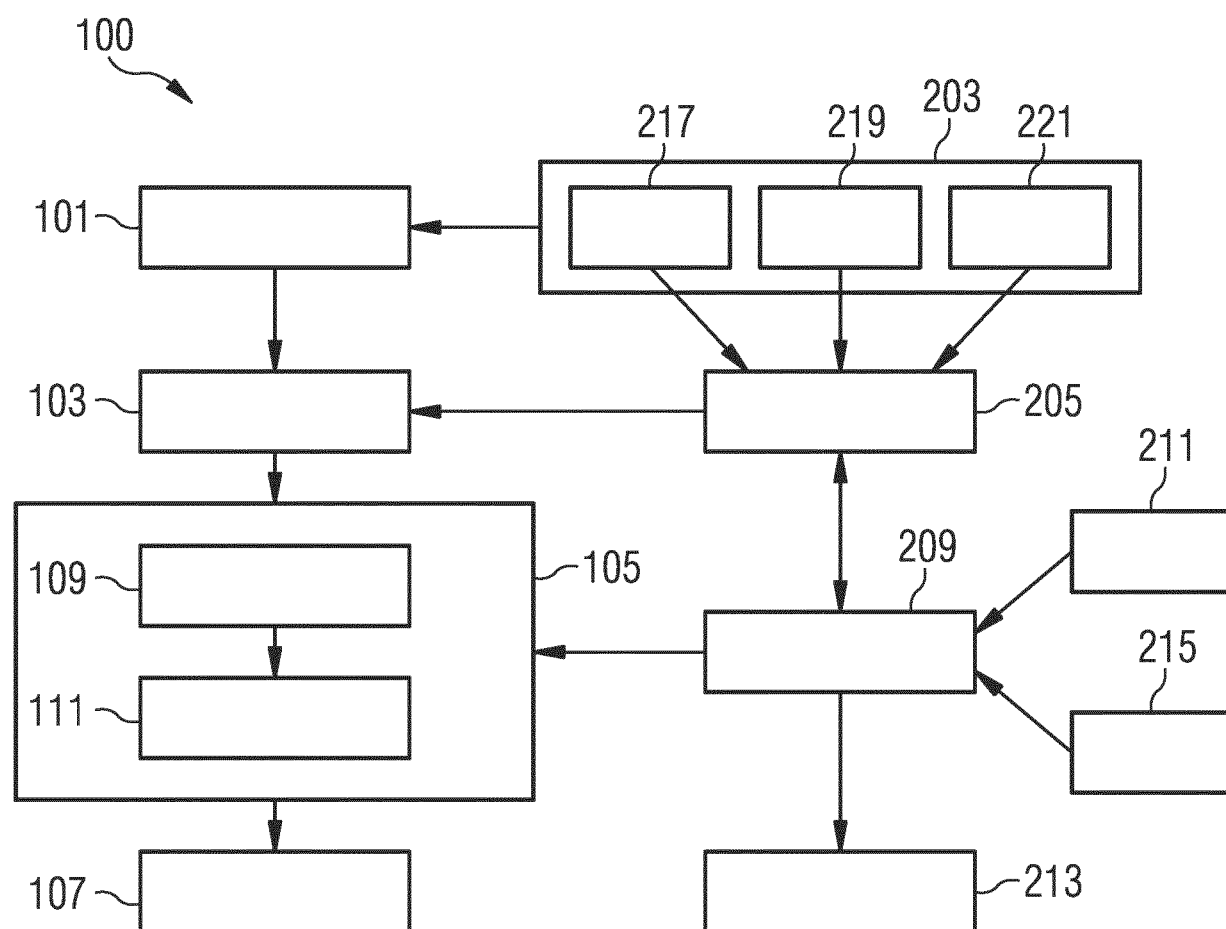


FIG 2

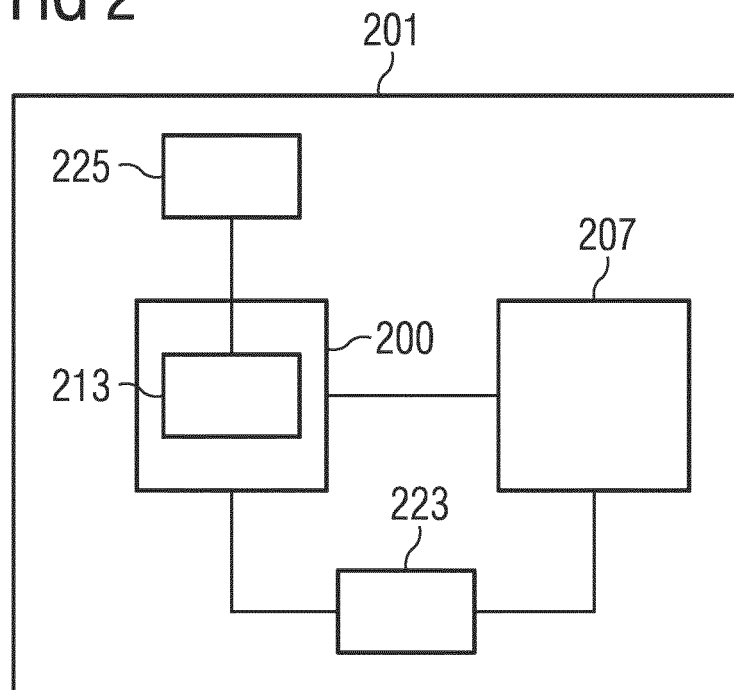
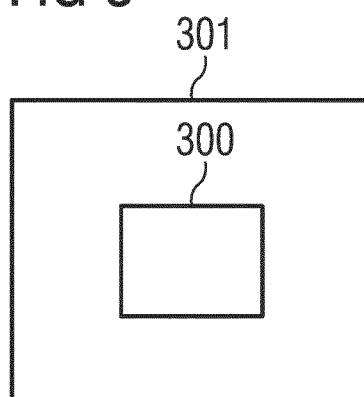


FIG 3



IN DER BESCHREIBUNG AUFGEFÜHRTE DOKUMENTE

Diese Liste der vom Anmelder aufgeführten Dokumente wurde ausschließlich zur Information des Lesers aufgenommen und ist nicht Bestandteil des europäischen Patentdokumentes. Sie wurde mit größter Sorgfalt zusammengestellt; das EPA übernimmt jedoch keinerlei Haftung für etwaige Fehler oder Auslassungen.

In der Beschreibung aufgeführte Patentedokumente

- EP 3552921 A1 [0004]
- DE 102017215802 A1 [0005]

In der Beschreibung aufgeführte Nicht-Patentliteratur

- Optimization design for aerodynamic elements of high speed trains. **SHUANBAO YAO et al.** COMPUTERS AND FLUIDS. PERGAMON PRESS, 03. März 2014, vol. 95, 56-73 [0006]