

(19)



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS  
ESPAÑA



(11) Número de publicación: **3 010 352**

(51) Int. Cl.:

**B61L 15/00** (2006.01)

**B61L 27/20** (2012.01)

**B61L 27/60** (2012.01)

(12)

## TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Fecha de presentación y número de la solicitud europea: **29.01.2021 E 21154227 (9)**

(97) Fecha y número de publicación de la concesión europea: **13.11.2024 EP 4035969**

---

(54) Título: **Procedimiento para entrenar un controlador para un vehículo ferroviario, controlador y vehículo ferroviario**

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:  
**02.04.2025**

(73) Titular/es:

**SIEMENS MOBILITY GMBH (100.00%)**  
Otto-Hahn-Ring 6  
81739 München, DE

(72) Inventor/es:

**PALMER, ANDREW y  
WEBER, MARC CHRISTIAN**

(74) Agente/Representante:

**CARVAJAL Y URQUIJO, Isabel**

**ES 3 010 352 T3**

---

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

## DESCRIPCIÓN

Procedimiento para entrenar un controlador para un vehículo ferroviario, controlador y vehículo ferroviario

La presente invención hace referencia a un procedimiento para entrenar un controlador para un vehículo ferroviario. La invención además hace referencia a un controlador y a un vehículo ferroviario con un controlador.

5 Para el funcionamiento automático de los trenes se requiere un controlador que pueda trasladar una trayectoria deseada, desde una unidad de planificación superordinada, a las órdenes de aceleración y de frenado requeridas para recorrer la trayectoria. Habitualmente, los controladores utilizados requieren un esfuerzo de adaptación considerable para adaptarse a las propiedades específicas de los vehículos individuales.

10 Los principios básicos para los controles se basan en controladores proporcionales, integrales y derivativos (PID) y en controladores de control predictivo por modelo (MPC). Los controladores PID se utilizan con mucha frecuencia, ya que pueden implementarse con mucha facilidad. Sin embargo, requieren un esfuerzo considerable para la adaptación cuando se utilizan en un nuevo vehículo. Además, éstos no consideran más que el intervalo de tiempo real y, por tanto, no presentan un buen rendimiento en sistemas en los que existe un retardo en la reacción del sistema en cuanto a las entradas de control. En esos casos, habitualmente se utilizan 15 controladores MPC. No obstante, el controlador MPC presupone un modelo preciso del sistema que debe controlarse. En los sistemas complejos esto puede ser complicado o no puede lograrse con una precisión suficiente.

Por la solicitud EP 3 552 921 A1 se conoce una planificación de velocidad autónoma de un actuador móvil limitado a una ruta predeterminada.

20 Además, en la solicitud DE 10 2017 215 802 A1 se describe un sistema de asistencia para el conductor para vehículos ferroviarios.

Asimismo, en SHUANBAO YAO ET AL: "Optimization design for aerodynamic elements of high speed trains", COMPUTERS AND FLUIDS, PERGAMON PRESS, NEW YORK, NY, GB, tomo 95, del 3 de marzo de 2014 (2014-03-03), páginas 56-73, se informa sobre el diseño de trenes de alta velocidad.

25 El objeto de la presente invención consiste en proporcionar un procedimiento mejorado para entrenar un controlador para un vehículo ferroviario, un controlador mejorado y un vehículo ferroviario con un controlador.

Dicho objeto, según la invención, se soluciona mediante un procedimiento para entrenar un controlador de un vehículo ferroviario, mediante un controlador entrenado de modo correspondiente y un vehículo ferroviario con un controlador entrenado, según las reivindicaciones independientes. En las reivindicaciones dependientes se indican configuraciones ventajosas.

Según un aspecto de la invención se proporciona un procedimiento para entrenar un controlador de un vehículo ferroviario, donde el procedimiento comprende:

- la puesta a disposición de datos de entrenamiento en base a datos de sensor de un vehículo ferroviario;

35 - el entrenamiento de un modelo de sustitución del vehículo ferroviario en base a los datos de entrenamiento con respecto a una relación entre un control de un accionamiento del vehículo ferroviario y una velocidad resultante del vehículo ferroviario;

- el entrenamiento de una regla de selección de acción en base a los datos de entrenamiento y al modelo de sustitución, utilizando aprendizaje automático y considerando al menos una finalidad de control objetiva, donde la regla de selección de acción comprende instrucciones de control para controlar el accionamiento del vehículo ferroviario, que están configuradas para acelerar el vehículo ferroviario, desde un primer estado de velocidad, 40 a un segundo estado de velocidad; y

- la generación de una regla de selección de acción entrenada, donde la regla de selección de acción entrenada comprende instrucciones de control que están configuradas para acelerar el vehículo ferroviario y cumplir con la finalidad de control.

45 Gracias a esto puede alcanzarse la ventaja técnica de que puede proporcionarse un procedimiento mejorado para entrenar un controlador de un vehículo ferroviario. El controlador del vehículo ferroviario se entrena en este caso considerando técnicas del aprendizaje automático. Para ello se proporcionan datos de entrenamiento que se basan en datos de sensor de un vehículo ferroviario. A continuación, en base a los datos de

entrenamiento, un modelo de sustitución del vehículo ferroviario se entrena con respecto a una relación entre un control de un accionamiento del vehículo ferroviario y una velocidad resultante de ello, del vehículo ferroviario. A continuación, se entrena una regla de selección de acción en base a los datos de entrenamiento y al modelo de sustitución entrenado, mediante técnicas del aprendizaje de refuerzo, y considerando al menos una finalidad de control, y se genera una regla de selección de acción entrenada. En base a la regla de selección de acción entrenada, el vehículo ferroviario puede controlarse y puede alcanzarse la finalidad de control respectivamente predeterminada.

En el sentido de la solicitud, un modelo de sustitución de un vehículo ferroviario es un modelo virtual de un vehículo ferroviario real, que representa todas las características relevantes del vehículo ferroviario real. Un modelo de sustitución puede estar diseñado de forma análoga a un gemelo virtual de una máquina real y puede simular virtualmente el funcionamiento de una máquina real o de un vehículo ferroviario. De este modo, el modelo de sustitución está configurado para simular un procedimiento de un vehículo ferroviario mediante un control correspondiente. El modelo de sustitución, por ejemplo, puede estar diseñado como una red neuronal entrenada de modo correspondiente, y puede estar entrenado para simular o representar el comportamiento de marcha o las propiedades del vehículo ferroviario relevantes para el control.

Una regla de selección de acción, en el sentido de la solicitud, es una pluralidad de instrucciones de control que están configuradas para acelerar el vehículo ferroviario, desde un primer estado de velocidad, a un segundo estado de velocidad. El primer y el segundo estado de velocidad pueden ser en este caso un estado real, en el que se encuentra el vehículo ferroviario en un instante determinado, y un estado objetivo, al que debe pasarse el vehículo ferroviario mediante el control del accionamiento. El primer y el segundo estado de velocidad, de manera alternativa, pueden ser dos estados consecutivos en el tiempo del vehículo ferroviario, a los que se pasa el vehículo ferroviario en el caso de un control mediante el controlador. Mediante la ejecución de las instrucciones de control de la regla de selección de acción, por medio del controlador, de este modo, el respectivo vehículo ferroviario puede controlarse considerando la finalidad de control que debe alcanzarse.

Las instrucciones de control, en el sentido de la solicitud, son instrucciones u órdenes para controlar el vehículo ferroviario. Las instrucciones de control, por ejemplo, pueden comprender la aceleración o el frenado del vehículo ferroviario, incluyendo un control detallado del accionamiento. Además, las instrucciones de control pueden comprender la comutación a distintas marchas o a un modo de ahorro de energía del accionamiento. Además, en las instrucciones de control pueden estar considerados distintos aspectos, relevantes para el control, en forma de instrucciones correspondientes.

Las finalidades de control, en el sentido de la solicitud, son objetivos que deben ser alcanzados mediante el control del controlador. Las finalidades de control, por ejemplo, pueden ser una velocidad a la que debe acelerarse el vehículo ferroviario mediante el control del controlador. Las finalidades de control, de manera alternativa o adicional, pueden estar definidas para distintos instantes futuros, por ejemplo mediante trayectorias de velocidad. De manera alternativa, las finalidades de control pueden definir un consumo de energía que debe alcanzarse o que no debe superarse al controlar el vehículo ferroviario. Las finalidades de control, en el sentido de la solicitud, están definidas en base a los datos de entrenamiento, que a su vez se han generado en base a datos de sensor de un vehículo ferroviario. Las finalidades de control, en este caso, pueden estar generadas a partir de los datos de sensor del vehículo ferroviario. Las finalidades de control pueden resultar de los datos de sensor y, por ejemplo, describir velocidades o consumos de energía que deben alcanzarse, según los cuales se controlaría el vehículo ferroviario durante el registro de los datos de sensor. Las finalidades de control también pueden definirse o modificarse durante la ejecución de la regla de selección de acción entrenada para el control del vehículo ferroviario mediante sistemas superordinados, por ejemplo módulos de planificación correspondientes para controlar el vehículo ferroviario.

Según una forma de ejecución, el aprendizaje automático está diseñado como aprendizaje de refuerzo.

Gracias a esto puede alcanzarse el efecto técnico de que se posibilita un entrenamiento eficiente de la regla de selección de acción.

El aprendizaje de refuerzo (Reinforcement Learning), en el sentido de la solicitud, es un área del aprendizaje automático que se ocupa del entrenamiento de participantes operacionales para ejecutar acciones deseadas, para de ser modo pasar el participante desde un estado real a un estado objetivo deseado. El entrenamiento del participante se realiza aquí considerando una finalidad de control que puede alcanzarse mediante la ejecución de la acción del participante. El entrenamiento puede considerar una maximización de una función de recompensa, mediante la cual se presenta al participante la acción que debe realizarse y el objetivo que debe alcanzarse.

Según una forma de ejecución, el entrenamiento comprende: la modificación randomizada de al menos una finalidad de control objetiva que se basa en los datos de entrenamiento y la definición de finalidades de control

modificadas; y el entrenamiento de la regla de selección de acción con respecto al cumplimiento de las finalidades de control modificadas.

Gracias a esto puede alcanzarse la ventaja técnica de que puede proporcionarse un entrenamiento más preciso del controlador del vehículo ferroviario. Mediante la modificación randomizada de al menos una finalidad de control y la definición, asociada a ello, de finalidades de control modificadas, los datos de entrenamiento pueden utilizarse de manera efectiva para el entrenamiento del controlador. Mediante la modificación de las finalidades de control y la generación o definición de finalidades de control modificadas, pueden definirse finalidades de control respaldadas por datos de entrenamiento, pero que no están limitadas a los mismos.

De este modo, de las finalidades de control modificadas resultan finalidades de control que no se basan en datos de entrenamiento, sino que representan finalidades de control que exceden las finalidades de control, según las que fue controlado el vehículo ferroviario durante el registro de los datos de sensor. Mediante el entrenamiento de la regla de selección de acción, en base a las finalidades de control modificadas, puede generarse una regla de selección de acción mejor entrenada, que define las instrucciones de control para un mayor número de finalidades de control diferentes. Gracias a esto puede proporcionarse un control mejorado.

Según una forma de ejecución, el entrenamiento de la regla de selección de acción comprende una maximización de una función de recompensa, donde la función de recompensa es máxima para una regla de selección de acción que cumple con la finalidad de control objetiva y/o con las finalidades de control modificadas.

Gracias a esto puede alcanzarse la ventaja técnica de que puede alcanzarse un entrenamiento lo más preciso posible de la regla de selección de acción según las técnicas del aprendizaje de refuerzo.

Según una forma de ejecución, la función de recompensa considera una diferencia entre un estado de velocidad logrado mediante la ejecución de una acción de control de la regla de selección de acción y la finalidad de control objetiva y/o las finalidades de control modificadas.

Gracias a esto puede alcanzarse la ventaja técnica de que puede proporcionarse un entrenamiento eficiente de la regla de selección de acción y una regla de selección de acción entrenada precisa, que cumple con precisión la finalidad de control que debe alcanzarse.

Según una forma de ejecución, la maximización de la función de recompensa se realiza mediante una red neuronal artificial.

Gracias a esto puede alcanzarse la ventaja técnica de que puede proporcionarse un entrenamiento eficiente de la regla de selección de acción, así como del controlador del vehículo ferroviario.

Según una forma de ejecución, la finalidad de control comprende una velocidad objetivo del vehículo ferroviario y/o un consumo de energía objetivo y/o una aceleración objetivo y/o un comportamiento de aceleración y/o de frenado con poco desgaste.

Gracias a esto puede alcanzarse la ventaja técnica de que puede proporcionarse una regla de selección de acción entrenada de manera eficiente y, asociado a ello, un controlador entrenado de manera eficiente, de un vehículo ferroviario. La regla de selección de acción entrenada de ese modo comprende aquí instrucciones de control que son adecuadas para controlar el vehículo ferroviario considerando las finalidades de control mencionadas.

Según una forma de ejecución, los datos de entrenamiento se registran durante un desplazamiento del vehículo ferroviario y comprenden datos de sensor de variables de estado, acciones de control y trayectorias de velocidad, donde las variables de estado comprenden datos de velocidad, datos de aceleración, datos de ubicación, datos de especificación del accionamiento y/o del vehículo ferroviario, donde las acciones de control comprenden operaciones de accionamiento y/o de frenado, y donde las trayectorias de velocidad describen desarrollos de velocidad correspondientes en el tiempo del vehículo ferroviario.

Gracias a esto puede alcanzarse la ventaja técnica de que mediante datos de entrenamiento completos es posible un entrenamiento preciso del controlador.

De acuerdo con un segundo aspecto de la invención se proporciona un controlador para un vehículo ferroviario, donde el controlador está entrenado según un procedimiento para entrenar un controlador de un vehículo ferroviario según una de las formas de ejecución precedentes, y donde el controlador está configurado para controlar el vehículo ferroviario ejecutando la regla de selección de acción entrenada.

De este modo puede proporcionarse un controlador mejorado para un vehículo ferroviario, que está entrenado utilizando métodos del aprendizaje automático, en particular del aprendizaje de refuerzo. Para ello, el controlador presenta una regla de selección de acción entrenada, que está entrenada según el procedimiento según la invención para entrenar un controlador de un vehículo ferroviario, según las formas de ejecución antes mencionadas. En base a la regla de selección de acción entrenada, que presenta una pluralidad de instrucciones de control para controlar el vehículo ferroviario considerando distintas finalidades de control, el vehículo ferroviario puede controlarse para alcanzar las respectivas finalidades de control. Mediante el entrenamiento de la regla de selección de acción en base al modelo de sustitución del vehículo ferroviario, el controlador puede entrenarse para cualquier vehículo ferroviario que corresponda al modelo de sustitución, de manera que, en el caso de un nuevo vehículo ferroviario, puede omitirse una adaptación compleja del modelo de un controlador, utilizado para el controlador, como se requiere en el caso de un controlador de control predictivo por modelo (MPC). Gracias a esto puede proporcionarse un controlador preciso, fiable y variable para vehículos ferroviarios.

Según un tercer aspecto se proporciona un vehículo ferroviario con un controlador según una de las formas de ejecución precedentes.

Gracias a esto puede proporcionarse un vehículo ferroviario con un controlador mejorado, con las ventajas antes mencionadas.

Según un cuarto aspecto se proporciona un producto de programa informático que comprende comandos que, al ser ejecutado el programa mediante una unidad de procesamiento de datos, disponen al mismo a realizar el procedimiento para entrenar un controlador para un vehículo ferroviario según una de las formas de ejecución precedentes.

Las características y ventajas de esta invención, descritas anteriormente, así como el modo de alcanzar las mismas, se aclaran y se vuelven más compresibles mediante las explicaciones de las siguientes representaciones, muy simplificadas, de ejemplos de ejecución preferentes. Muestran:

Figura 1 un diagrama de flujo de un procedimiento para entrenar un controlador de un vehículo ferroviario, según una forma de ejecución;

Figura 2 una representación esquemática de un vehículo ferroviario con un controlador según una forma de ejecución; y

Figura 3 una representación esquemática de un producto de programa informático.

La Figura 1 muestra un diagrama de flujo de un procedimiento 100 para entrenar un controlador 200 de un vehículo ferroviario 201, según una forma de ejecución.

Para el entrenamiento del controlador 200 del vehículo ferroviario 201 en primer lugar, en una primera etapa del procedimiento 101, se proporcionan datos de entrenamiento. Los datos de entrenamiento en este caso se basan en datos de sensor de un vehículo ferroviario 201 y, en la forma de ejecución mostrada, comprenden variables de estado 217, acciones de control 219 y trayectorias de velocidad 221. Los datos de entrenamiento 203, en particular los datos de sensor del vehículo ferroviario 201, por ejemplo, pueden registrarse durante una marcha del vehículo ferroviario 201 o de un vehículo ferroviario comparable, mediante un sistema de sensores correspondientes.

Las variables de estado 217, en el sentido de la solicitud, describen puntos dentro de un espacio de estado que describe distintos estados del vehículo ferroviario 201. En particular, las variables de estado 217 pueden comprender datos de velocidad, datos de aceleración y/o datos de ubicación del vehículo ferroviario 201, que se registraron durante el desplazamiento del vehículo ferroviario 201 o del vehículo ferroviario comparable. El vehículo ferroviario comparable, por ejemplo, puede ser un vehículo ferroviario de un tipo idéntico. Además, las variables de estado 217 pueden comprender datos de especificación del accionamiento y/o del vehículo ferroviario 201, que por ejemplo comprenden el tipo de accionamiento, la potencia o velocidad máxima, u otros parámetros del accionamiento, o un tamaño o un peso del vehículo ferroviario 201.

Las acciones de control 219, en el sentido de la solicitud, son acciones que realiza el controlador 200 durante el desplazamiento del vehículo ferroviario 201, para controlar el vehículo ferroviario 201. Las acciones de control 219, por ejemplo, pueden comprender el accionamiento del pedal del acelerador o de la palanca de control de un vehículo ferroviario, o el accionamiento del freno del vehículo ferroviario 201.

Las trayectorias de velocidad 221, en el sentido de la solicitud, son líneas que se extienden por el espacio de estado, de valores de velocidad consecutivos en el tiempo, que se alcanzaron durante el desplazamiento del vehículo ferroviario 201.

5      Despues de la puesta a disposición de los datos de entrenamiento 203, en otra etapa del procedimiento 103, se genera, así como se entrena, un modelo de sustitución 205 del vehículo ferroviario 201, y el modelo de sustitución 205 aprende una relación entre controles de un accionamiento 207 del vehículo ferroviario 201 y una velocidad del vehículo ferroviario 201, resultante de ello.

10     En base a los datos de entrenamiento 203 que se registraron durante el desplazamiento del vehículo ferroviario 201, el modelo de sustitución 205, que representa una copia virtual del vehículo ferroviario 201, se entrena mediante métodos del aprendizaje automático, que conducen a acciones de control 219, al accionamiento del pedal del acelerador o al accionamiento del freno, con respecto a velocidades finales del vehículo ferroviario 201 que se determinan en base a las variables de estado 217 correspondientes. El modelo de sustitución 205 entrenado de ese modo, con ello, permite una simulación del controlador 200 del vehículo ferroviario 201, en donde mediante la ejecución de funciones de control 219 correspondientes pueden alcanzarse trayectorias de velocidad 221 correspondientes.

20     En otra etapa del procedimiento 105, se entrena una regla de selección de acción 209 en base a los datos de entrenamiento 203 y al modelo de sustitución 205 entrenado, utilizando métodos del aprendizaje de refuerzo y considerando al menos una finalidad de control objetiva 211. La regla de selección de acción 209 comprende aquí instrucciones de control para controlar el accionamiento 207 del vehículo ferroviario 201. En este caso, las instrucciones de control están configuradas para acelerar el vehículo ferroviario 201, desde un primer estado de velocidad, a un segundo estado de velocidad.

25     Las instrucciones de control, por ejemplo, pueden corresponder a las acciones de control 219 que fueron registradas como datos de entrenamiento 203 durante el desplazamiento del vehículo ferroviario 201, y que comprenden la aceleración o el frenado del vehículo ferroviario. La regla de selección de acción 209, para ello, puede presentar una pluralidad de instrucciones de control, mediante las que puede controlarse el accionamiento 207 del vehículo ferroviario 201. Las instrucciones de control, en este caso, están diseñadas de manera que durante el control del accionamiento 207 del vehículo ferroviario 201 se alcanza la finalidad de control objetiva 211. En este caso, por ejemplo, la finalidad de control objetiva 211 puede ser una velocidad final que debe alcanzarse, un consumo de energía deseado o una aceleración máxima del vehículo ferroviario 201, que respectivamente deben alcanzarse u observarse durante el control del vehículo ferroviario 201. La finalidad de control objetiva 211 igualmente puede estar proporcionada mediante los datos de entrenamiento 203. Por ejemplo, la finalidad de control 211 puede estar representada mediante las trayectorias de velocidad 221 registradas de los datos de entrenamiento 203.

35     Los primeros y segundos estados de velocidad, en este caso, pueden ser respectivamente un estado real o un estado objetivo del vehículo ferroviario 201, donde el vehículo ferroviario 201, mediante el control según las instrucciones de control de la regla de selección de acción 209, puede pasarse desde el estado real al estado objetivo. Alternativamente con respecto a ello, el primer y el segundo estado de velocidad pueden ser dos estados del espacio de estado que se presentan de forma consecutiva en el tiempo, a los que puede pasarse el vehículo ferroviario 201, mediante el control según las instrucciones de control de la regla de selección de acción.

40     Para el entrenamiento de la regla de selección de acción 209 mediante aprendizaje automático, se entrena y optimiza una regla de selección de acción seleccionada del modo deseado, con cualquier instrucción de control en base a los datos de entrenamiento 203, incluyendo las variables de estado 217 del espacio de estado del vehículo ferroviario 201 y considerando la finalidad de control 211 respectivamente seleccionada, de modo que la regla de selección de acción 209 está configurada para alcanzar la finalidad de control 211 seleccionada. En la forma de ejecución mostrada, el entrenamiento de cualquier regla de selección de acción 209 seleccionada se realiza en otra etapa del procedimiento 111, mediante una maximización de una función de recompensa configurada de modo correspondiente. La función de recompensa, en este caso, por ejemplo, puede estar definida como una diferencia entre un estado de velocidad alcanzado mediante la realización de una acción de control 219 de la regla de selección de acción 209 y la finalidad de control objetiva 211. El entrenamiento de la regla de selección de acción 209, con ello, tiene lugar de manera que las instrucciones de control, así como la regla de selección de acción 209, se modifican de manera que la función de recompensa, definida de modo correspondiente, alcanza un valor máximo. Por consiguiente, una regla de selección de acción 209 con una función de recompensa máxima puede alcanzar la finalidad de control 211 seleccionada. Mediante el entrenamiento de esa clase de la regla de selección de acción 209 puede atravesarse el espacio de estado del vehículo ferroviario 201, definida por los datos de entrenamiento 203, en donde se indican diferentes estados del vehículo ferroviario 201, para determinar la regla de selección de acción 209 óptima, que está configurada para, mediante la ejecución de las instrucciones de control correspondientes, pasar el vehículo ferroviario 201,

en una trayectoria optimizada, a estados que garanticen un controlador 200 óptimo y el alcance de la finalidad de control objetiva 211 predeterminada.

Según la forma de ejecución mostrada, para el entrenamiento de la regla de selección de acción 209, en una etapa del procedimiento 109, la finalidad de control objetiva 211 se modifica de forma randomizada y se generan finalidades de control 215 modificadas. Mediante la modificación randomizada de la finalidad de control 211 pueden generarse finalidades de control 215 modificadas que pueden diferir de los datos de entrenamiento 203.

Por ejemplo, la finalidad de control objetiva 211 puede estar formada por una trayectoria de velocidad 221 de los datos de entrenamiento 203. La respectiva trayectoria de velocidad 221 puede estar respaldada por los datos de sensor del vehículo ferroviario 201, que fueron registrados durante el desplazamiento del vehículo ferroviario 201. Mediante una modificación de valores individuales de la trayectoria de velocidad 221, de este modo, pueden generarse trayectorias de velocidad 221 modificadas, como finalidades de control 215 modificadas, donde las trayectorias de velocidad 221 modificadas no están respaldadas por completo por los datos de entrenamiento 203, y presentan valores de velocidad que difieren durante el desplazamiento del vehículo ferroviario 201 y el registro de los respectivos valores de velocidad del vehículo ferroviario 201. Modificando la trayectoria de velocidad 221, de este modo, pueden alcanzarse puntos en el espacio de estado del vehículo ferroviario 201, para los que no se generaron datos de entrenamiento 203 explícitos.

Alternativamente, en el entrenamiento de la regla de selección de acción 209 también pueden considerarse varias finalidades de control 211, 215, de modo que la regla de selección de acción 209 esté configurada para cumplir con una pluralidad de finalidades de control 211, 215. Las finalidades de control, junto con la velocidad final del vehículo ferroviario 201 que debe alcanzarse, por ejemplo, pueden comprender un consumo de energía del vehículo ferroviario 201 o una aceleración máxima admisible del vehículo ferroviario 201, que deben cumplirse u observarse con el controlador 200 del vehículo ferroviario 201.

Después del entrenamiento de la regla de selección de acción 209, en la etapa del procedimiento 105, en base a la regla de selección de acción inicial 209, se genera una regla de selección de acción 213 entrenada de modo correspondiente, que comprende instrucciones de control que están configuradas para acelerar el vehículo ferroviario 201 y para cumplir con la finalidad de control 211 o con las finalidades de control 215 modificadas.

El entrenamiento de la regla de selección de acción 209, así como la maximización de la función de recompensa, puede realizarse mediante una inteligencia artificial entrenada, por ejemplo mediante una red neuronal entrenada de forma correspondiente. Con ello, mediante la maximización de la función de recompensa configurada de modo correspondiente puede generarse una regla de selección de acción 213 que cumpla con las finalidades de control 211, 215 correspondientes, garantizando un controlador 200 optimizado del vehículo ferroviario 201.

La figura 2 muestra una representación de un vehículo ferroviario 201 con un controlador 200 según una forma de ejecución.

La figura 2 muestra un vehículo ferroviario 201 con un controlador 200, donde el controlador 200 comprende una regla de selección de acción 213 entrenada según el procedimiento 100 según la invención, para entrenar un controlador 200 de un vehículo ferroviario 201. El vehículo ferroviario 201 comprende además un accionamiento 207 y un sensor 223. Mediante el sensor 223 pueden registrarse valores de sensor del accionamiento 207, para hacer el seguimiento de un control del accionamiento 207, mediante el controlador 200. El vehículo ferroviario 201 comprende además una unidad de planificación 225 que está conectada al controlador 200. Para controlar el vehículo ferroviario 201, de este modo, el controlador 200 puede recibir desde la unidad de planificación 225 una trayectoria de velocidad 221 correspondiente, que describe una curva de velocidad planificada del desplazamiento del vehículo ferroviario 201. Mediante la ejecución de las instrucciones de control definidas por la regla de selección de acción 213 entrenada, de este modo, el controlador 200 puede realizar las instrucciones de control correspondientes que sean adecuadas para acelerar el vehículo ferroviario 201 según la trayectoria de velocidad 221 planificada de la unidad de planificación 225. El controlador 200 del vehículo ferroviario 201, de este modo, puede considerar diferentes finalidades de control 211, 215. Por ejemplo, el vehículo ferroviario 201 puede controlarse mediante un consumo de energía máximo predeterminado. De manera alternativa o adicional con respecto a ello, el vehículo ferroviario 201 puede controlarse considerando una aceleración máxima admisible. Si la regla de selección de acción 213 entrenada de modo correspondiente, para una pluralidad de finalidades de control 211, 215 diferentes, comprende instrucciones de control correspondientes que son adecuadas para controlar el vehículo ferroviario 201 considerando las respectivas finalidades de control 211, 215, durante el funcionamiento del vehículo ferroviario 201 y en particular, al ya estar instalado el controlador 200, pueden modificarse finalidades de control 211, 215 correspondientes, que deben alcanzarse durante el control del vehículo ferroviario 201. Una modificación del

controlador 200, y en particular una adaptación de las instrucciones de control respectivamente definidas, no son necesarias debido al entrenamiento de la regla de acción 213 entrenada.

La FIG 3 muestra una representación esquemática de un producto de programa informático 300.

La Figura 3 muestra un producto de programa informático 300 que comprende comandos que, al ser ejecutado el programa mediante una unidad informática, disponen al mismo a realizar el procedimiento 100 según una de las formas de ejecución antes mencionadas. El producto de programa informático 300, en la forma de ejecución mostrada, está almacenado en un medio de almacenamiento 301. En este caso, el medio de almacenamiento 301 puede ser cualquier medio de almacenamiento conocido por el estado de la técnica.

Si bien la invención fue ilustrada y descrita en detalle mediante el ejemplo de ejecución preferente, la invención no está limitada por los ejemplos descritos, y el experto puede deducir de éstos otras variaciones, sin abandonar el alcance de protección de la invención, que está definido mediante las reivindicaciones.

## REIVINDICACIONES

1. Procedimiento (100) para entrenar un controlador (200) de un vehículo ferroviario (201), que comprende:
    - la puesta a disposición (101) de datos de entrenamiento (203) en base a datos de sensor de un vehículo ferroviario (201);
- 5 - el entrenamiento (103) de un modelo de sustitución (205) del vehículo ferroviario (201) en base a los datos de entrenamiento (203) con respecto a una relación entre un control de un accionamiento (207) del vehículo ferroviario (201) y una velocidad resultante del vehículo ferroviario (201);
- 10 - el entrenamiento (105) de una regla de selección de acción (209) en base a los datos de entrenamiento (203) y al modelo de sustitución (205), utilizando aprendizaje automático y considerando al menos una finalidad de control objetiva (211), donde la regla de selección de acción (209) comprende instrucciones de control para controlar el accionamiento (207) del vehículo ferroviario (201), que están configuradas para acelerar el vehículo ferroviario (201), desde un primer estado de velocidad, a un segundo estado de velocidad; y
- 15 - la generación (107) de una regla de selección de acción (213) entrenada, donde la regla de selección de acción (213) entrenada comprende instrucciones de control que están configuradas para acelerar el vehículo ferroviario (201) y cumplir con la finalidad de control (211).
2. Procedimiento (100) según la reivindicación 1, donde el aprendizaje automático está diseñado como aprendizaje de refuerzo.
  3. Procedimiento (100) según la reivindicación 1 ó 2, donde el entrenamiento (105) de la regla de selección de acción (209) comprende:
    - la modificación randomizada (109) de al menos una finalidad de control objetiva (211) que se basa en los datos de entrenamiento (203) y la definición de finalidades de control (215) modificadas; y
    - el entrenamiento de la regla de selección de acción (209) con respecto al cumplimiento de las finalidades de control (215) modificadas.
- 25 4. Procedimiento (100) según la reivindicación 3, donde el entrenamiento (105) de la regla de selección de acción (209) comprende: la maximización (111) de una función de recompensa, donde la función de recompensa es máxima para una regla de selección de acción (209) que cumple con la finalidad de control objetiva (211) y/o con las finalidades de control (215) modificadas.
- 30 5. Procedimiento (100) según la reivindicación 4, donde la función de recompensa considera una diferencia entre un estado de velocidad logrado mediante la ejecución de una instrucción de control de la regla de selección de acción (209) y la finalidad de control objetiva (211) y/o las finalidades de control (215) modificadas.
- 35 6. Procedimiento (100) según la reivindicación 1 ó 2, donde el entrenamiento (105) de la regla de selección de acción (209) comprende:
  - la maximización (111) de una función de recompensa, donde la función de recompensa es máxima para una regla de selección de acción (209) que cumple con la finalidad de control objetiva (211).
- 35 7. Procedimiento (100) según la reivindicación 6, donde la función de recompensa considera una diferencia entre un estado de velocidad logrado mediante la ejecución de una instrucción de control de la regla de selección de acción (209) y la finalidad de control objetiva (211).
- 40 8. Procedimiento (100) según una de las reivindicaciones 4 a 7, donde la maximización (111) de la función de recompensa se realiza mediante una red neuronal artificial.
- 40 9. Procedimiento (100) según una de las reivindicaciones 3, 4 ó 5 precedentes, donde la finalidad de control (211) y/o las finalidades de control (215) modificadas comprenden una velocidad objetivo del vehículo ferroviario (201) y/o un consumo de energía objetivo y/o una aceleración objetivo y/o un comportamiento de aceleración y/o de frenado con poco desgaste.

10. Procedimiento (100) según una de las reivindicaciones 6, 7 u 8 precedentes, donde la finalidad de control (211) comprende una velocidad objetivo del vehículo ferroviario (201) y/o un consumo de energía objetivo y/o una aceleración objetivo y/o un comportamiento de aceleración y/o de frenado con poco desgaste.
- 5      11. Procedimiento (100) según una de las reivindicaciones precedentes, donde los datos de entrenamiento (203) se registran durante un desplazamiento del vehículo ferroviario (201) y comprenden datos de sensor de variables de estado (217), acciones de control (219) y trayectorias de velocidad (221), donde las variables de estado (217) comprenden datos de velocidad, datos de aceleración, datos de ubicación, datos de especificación del accionamiento y/o del vehículo ferroviario (201), donde las acciones de control (219) comprenden operaciones de accionamiento y/o de frenado, y donde las trayectorias de velocidad (221) describen desarrollos de velocidad correspondientes en el tiempo del vehículo ferroviario (201).
- 10      12. Controlador (200) para un vehículo ferroviario (201), donde el controlador (200) está entrenado según un procedimiento (100) para entrenar un controlador (200) de un vehículo ferroviario (201) según una de las reivindicaciones 1 a 11 precedentes, y donde el controlador (200) está configurado para controlar el vehículo ferroviario (201) ejecutando la regla de selección de acción (213) entrenada.
- 15      13. Vehículo ferroviario (201) con un controlador (200) según la reivindicación 12.
14. Producto de programa informático (300) que comprende órdenes que, al ser ejecutado el programa por una unidad de procesamiento de datos, disponen al mismo a realizar el procedimiento (100) según una de las reivindicaciones 1 a 11 precedentes.

## DIBUJOS

FIG 1

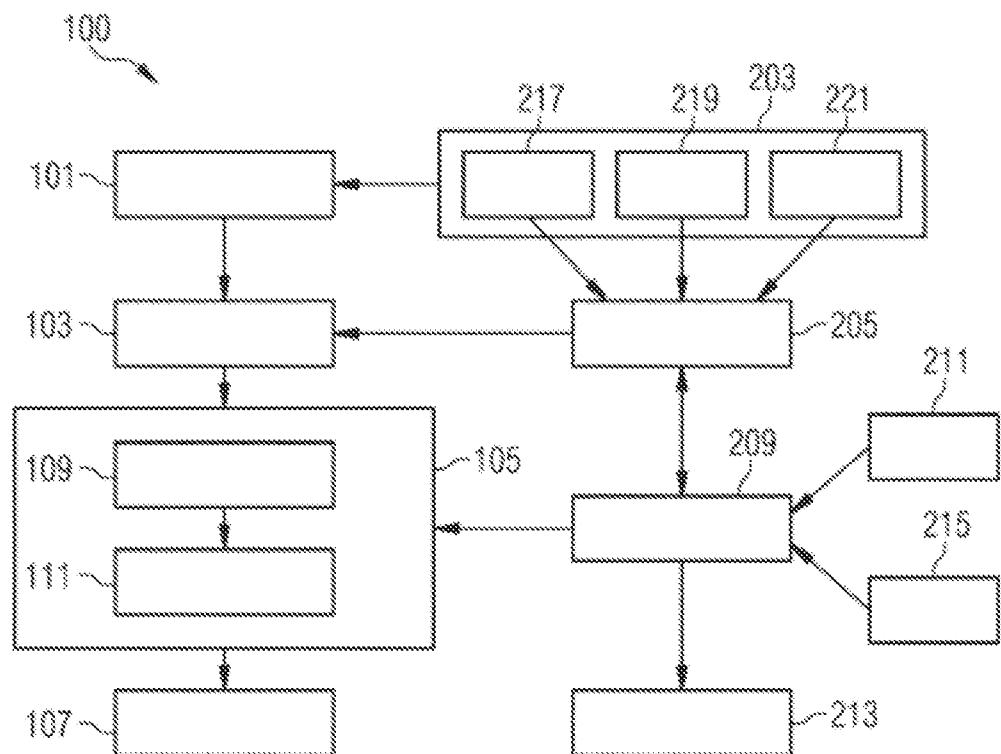


FIG 2

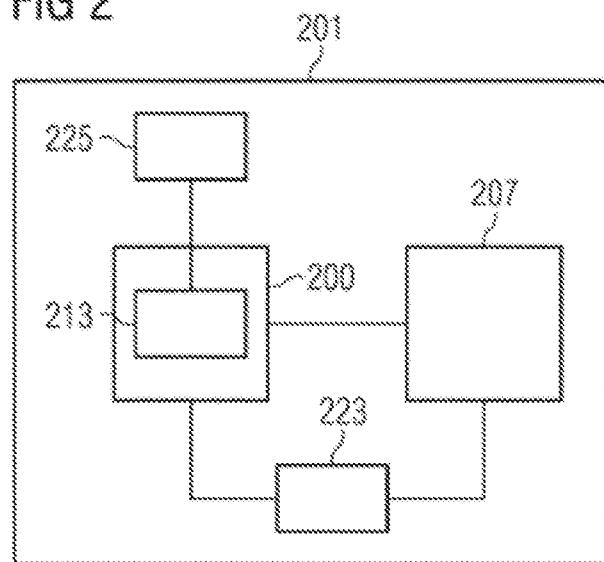


FIG 3

