(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
29 November 2018 (29.11.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/216490 A1

(51) International Patent Classification:
B25J 9/16 (2006.01) *G05D 1/02* (2006.01)
G05B 13/02 (2006.01) *G06N 3/08* (2006.01)
G05D 1/00 (2006.01)

(21) International Application Number:
PCT/JP2018/018133

(22) International Filing Date:
10 May 2018 (10.05.2018)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2017-104528 26 May 2017 (26.05.2017) JP

(71) Applicant: **OMRON CORPORATION** [JP/JP]; 801, Minamifudodo-cho, Horikawahigashiiru, Shiokoji-dori, Shimogyo-ku, Kyoto-shi, Kyoto, 6008530 (JP).

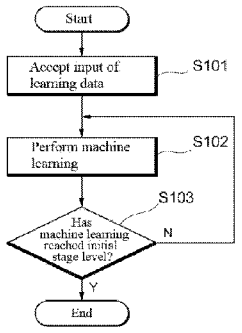
(72) Inventor: **ANDO, Tanichi**; c/o OMRON Automotive Electronics Co. Ltd., 6368, Okusa-Nenjozaka, Komaki-shi, Aichi, 4850802 (JP).

(74) Agent: **INABA, Yoshiyuki** et al.; TMI ASSOCIATES, 23rd Floor, Roppongi Hills Mori Tower, 6-10-1, Roppongi, Minato-ku, Tokyo, 1066123 (JP).

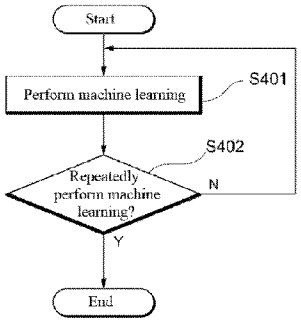
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG,

(54) Title: LEARNING APPARATUS, LEARNING CONTROL METHOD, PROGRAM THEREFOR

[Fig. 5]



[Fig. 8]



(57) Abstract: In order to provide a technique for shortening the time required for a learning apparatus to achieve a learning purpose, without performing manual manipulation, a learning apparatus configured to learn control of an operation involved in a predetermined task includes: a learning data accepting unit configured to accept learning data containing a learning purpose and an allowable requirement for a learning operation performed when learning the control; a neural network configured to perform learning based on the learning data; and an output unit configured to output a learning result obtained by the neural network, wherein the neural network performs a first learning process for achieving an initial stage of the learning purpose, performs a second learning process for learning a control range in which a learning operation matches an allowable requirement, based on a result of the first learning process, and performs a third learning process for achieving the learning purpose within the control range based on a result of the second learning process.

WO 2018/216490 A1

[Continued on next page]

MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

Description

Title of Invention: LEARNING APPARATUS, LEARNING CONTROL METHOD, PROGRAM THEREFOR

Technical Field

[0001] The present invention relates to a learning apparatus, a learning control method, and a program therefor.

(CROSS-REFERENCES TO RELATED APPLICATIONS)

This application claims priority to Japanese Patent Application No. 2017-104528 filed May 26, 2017, the entire contents of which are incorporated herein by reference.

Background Art

[0002] Conventionally, artificial intelligence technology (hereinafter, referred to as "AI technology") such as neural networks has been widely researched (see JP H6-289918A, for example). In particular, with the rise of AI technology called deep learning, for example, the recognition rates of techniques for recognizing targets based on images have improved rapidly in these years, and the recognition rate of classifying images is almost higher than that of humans. Deep learning technology is expected to be applied not only to image recognition but also to various other fields such as speech recognition, personal authentication, behavior prediction, summary writing, machine translation, monitoring, autonomous driving, failure prediction, sensor data analysis, music genre determination, content generation, and security systems.

[0003] In machine learning such as deep learning, machines can be trained to attain a predetermined ability. At this time, learning apparatuses that perform machine learning repeatedly perform learning operations until they attain a predetermined ability.

[0004] For example, JP H6-289918A discloses a learning control method for robots. In the learning control method described in JP H6-289918A, an input value that is to be supplied to a driving unit of a robot is modified based on a difference between a target motion targeted in a robot operation set in advance by a person and an actual motion when the robot actually operates.

Citation List

Patent Literature

[0005] PTL 1: JP H6-289918A

[0006] In learning apparatuses that control actuators based on information from many sensors, such as control of an engine and driving of an automobile, chemical plants, or the like, control and sensor information output affect each other, and thus more complex learning has to be performed in order to attain a control method. Accordingly, in learning apparatuses that perform such complex learning, it is not easy for a person

to set a target value of the control amount in advance as in JP H6-289918A. On the other hand, in the case of training learning apparatuses without setting a target value, it is necessary to repeat trial and error processing for an extremely large number of times, which is inefficient.

Summary of Invention

Technical Problem

[0007] It is an object of the present invention to shorten the time required for a learning apparatus to achieve a learning purpose, without performing manual manipulation.

Solution to Problem

[0008] An aspect of the present invention is directed to a learning apparatus configured to learn control of an operation involved in a predetermined task, the apparatus including: a learning data accepting unit configured to accept learning data containing a learning purpose and an allowable requirement for a learning operation performed when learning the control; a neural network configured to perform learning based on the learning data; and an output unit configured to output a learning result obtained by the neural network, wherein the neural network performs a first learning process for achieving an initial stage of the learning purpose, performs a second learning process for learning a control range in which the learning operation matches the allowable requirement, based on a result of the first learning process, and performs a third learning process for achieving the learning purpose within the control range based on a result of the second learning process.

[0009] With this configuration, learning of a control range in which a learning operation matches an allowable requirement is performed before a third learning process for achieving a learning purpose is performed. Accordingly, the apparatus itself can perform learning in a control range in which a learning operation matches an allowable requirement, without a person setting conditions that restrict control operations. Thus, it is possible to achieve the learning purpose in a shorter period of time.

[0010] It is possible that the output unit outputs the result of the second learning process. With this aspect, a learning result of the control range in which a learning operation matches an allowable requirement can be used by other learning apparatuses as well.

[0011] It is possible that the learning apparatus is a learning apparatus configured to learn control of a series of operations involved in a predetermined task, the learning apparatus further includes a classifying unit configured to divide the task into a plurality of scenes, and, for each of the divided scenes, specify a partial operation that is to be performed in the scene, from among the series of operations, and the neural network performs the second learning process and the third learning process for each partial operation.

[0012] With this aspect, the learning apparatus can classify an operation involved in the learning into partial operations, which is classifying an operation into smaller units, according to the scenes, and perform learning for each classified partial operation. Accordingly, it is possible to achieve the learning purpose in a shorter period of time.

[0013] An aspect of the present invention is directed to an autonomous driving control learning apparatus configured to learn control of a series of operations involved in autonomous driving of a vehicle that does laps of a predetermined course, the apparatus including: a learning data accepting unit configured to accept learning data containing a learning purpose in which a purpose is to do a predetermined number of laps of the course within a predetermined period of time, and an allowable requirement for a learning operation performed when learning the control; a neural network configured to perform learning based on the learning data; and an output unit configured to output a learning result obtained by the neural network, wherein the neural network performs a first learning process for achieving running of one lap of the course by the vehicle, performs a second learning process for learning a control range in which the learning operation matches the allowable requirement, based on a result of the first learning process, and performs a third learning process for achieving the learning purpose within the control range based on a result of the second learning process.

[0014] An aspect of the present invention is directed to a robot control learning apparatus configured to learn control of a series of operations involved in a task of carrying a predetermined target to a predetermined placement position, and stacking and wrapping a predetermined number of the targets, the apparatus including: a learning data accepting unit configured to accept learning data containing a learning purpose in which a purpose is to complete a series of operations as soon as possible, and an allowable requirement for a learning operation performed when learning the control; a neural network configured to perform learning based on the learning data; and an output unit configured to output a learning result obtained by the neural network, wherein the neural network performs a first learning process for achieving control of the series of operations involved in the task, performs a second learning process for learning a control range in which the learning operation matches the allowable requirement, based on a result of the first learning process, and performs a third learning process for achieving the learning purpose within the control range based on a result of the second learning process.

[0015] An aspect of the present invention is directed to a learning method for learning control of an operation involved in a predetermined task, the method including: accepting, by a computer, learning data containing a learning purpose and an allowable requirement for a learning operation performed when learning the control; performing, by the computer, learning based on the learning data; and outputting, by the computer,

a learning result obtained in the learning, wherein performing learning includes performing, by the computer, a first learning process for achieving an initial stage of the learning purpose, performing, by the computer, a second learning process for learning a control range in which the learning operation matches the allowable requirement, based on a result of the first learning process, and performing, by the computer, a third learning process for achieving the learning purpose within the control range based on a result of the second learning process.

[0016] An aspect of the present invention is directed to a program for causing a computer configured to learn control of an operation involved in a predetermined task to execute: accepting learning data containing a learning purpose and an allowable requirement for a learning operation performed when learning the control; performing learning based on the learning data; and outputting a learning result obtained in the learning, wherein performing learning includes performing a first learning process for achieving an initial stage of the learning purpose, performing a second learning process for learning a control range in which the learning operation matches the allowable requirement, based on a result of the first learning process, and performing a third learning process for achieving the learning purpose within the control range based on a result of the second learning process.

[0017] An aspect of the present invention is directed to an apparatus configured to perform a predetermined task, the apparatus including: a first sensor configured to sense information necessary for an operation of the apparatus to perform the task; an actuator; a second sensor configured to sense a change in a status of the apparatus caused by the actuator; a control unit configured to control the actuator based on sensor values output from the first sensor and the second sensor; and a storage unit configured to store a learning result obtained by the above-described learning apparatus, wherein the control unit determines a control amount that corresponds to the sensor values output from the first sensor and the second sensor, within the control range, based on the learning result stored in the storage unit.

Advantageous Effects of Invention

[0018] According to the present invention, it is possible to provide a technique for shortening the time required for a learning apparatus to achieve a learning purpose, without performing manual manipulation.

Brief Description of Drawings

[0019] [fig.1]FIG. 1 is a block diagram showing a schematic configuration of a learning apparatus in a first embodiment.

[fig.2]FIG. 2 is a schematic diagram showing a course for autonomous driving of a vehicle that is controlled by the learning apparatus in the first embodiment.

[fig.3]FIG. 3 is a flowchart showing the outline of the processing performed by the learning apparatus in the first embodiment.

[fig.4]FIG. 4 is a block diagram showing a detailed configuration of the learning apparatus in the first embodiment.

[fig.5]FIG. 5 is a flowchart showing details of the processing performed by the learning apparatus in the first embodiment.

[fig.6]FIG. 6 is a flowchart showing details of the processing performed by the learning apparatus in the first embodiment.

[fig.7]FIG. 7 is a flowchart showing details of the processing performed by the learning apparatus in the first embodiment.

[fig.8]FIG. 8 is a flowchart showing details of the processing performed by the learning apparatus in the first embodiment.

[fig.9]FIG. 9 is a diagram showing an example of the hardware configuration of the learning apparatus in the first embodiment.

[fig.10]FIG. 10 is a block diagram showing a schematic configuration of a learning apparatus in a second embodiment.

Description of Embodiments

[0020] First Embodiment

Hereinafter, embodiments of the present invention will be described in detail with reference to the drawings. Note that the same constituent elements are denoted by the same reference numerals, and a description thereof may not be repeated. The following embodiments are merely illustrative for the sake of explanation, and are not intended to limit the present invention thereto. Various modifications can be made without departing from the gist thereof.

[0021] 1. System Outline

Hereinafter, the outline of the system in this embodiment will be described with reference to FIGS. 1 to 3. FIG. 1 is a block diagram showing a schematic configuration of a learning apparatus 1 according to this embodiment. the learning apparatus 1 learns a predetermined task. For example, the learning apparatus 1 according to this embodiment is mounted in a vehicle controlled to drive autonomously (hereinafter, alternatively referred to simply as a "vehicle") 90, and learns control of the vehicle 90 for autonomous driving along a predetermined course (see FIG. 2). At this time, the learning apparatus 1 is provided with learning data, for example, from an operator or the like. The learning data is data containing, for example, a learning purpose, learning requirements, and an allowable requirement as follows.

Learning Purpose

- Do ten laps of a course within predetermined period of time and reach goal

Learning Requirements

- Do not slide off track
- Running direction is clockwise
- Reach goal
- "Do one lap of course and reach goal" in initial stage level

Allowable Requirement

- Do not come in contact with another object

[0022] Note that the task is a matter that is required to be achieved through an operation involved in the learning ("operation involved in the learning" in this embodiment refers to various types of control necessary for autonomous driving of the vehicle 90, and may be considered as being an operation that is performed by the vehicle 90 through the various types of control), and refers to doing laps of a course in this embodiment. Furthermore, the learning purpose is a standard that is to be achieved by the task, and refers to "do ten laps of a course within predetermined period of time and reach goal" as described above in this embodiment. Thus, in this embodiment, it is considered that making the task performable is provided as a learning requirement, in learning in the initial stage level.

[0023] Furthermore, the allowable requirement is a requirement that is required to be satisfied through an operation (learning operation) that is performed during learning. That is to say, it is meant that, during learning, for example, the vehicle is allowed to slide off the track, but is not allowed to come in contact with another object. Note that a learning purpose, learning requirements, and an allowable requirement are not limited thereto.

[0024] In this embodiment, the learning data is input by an operator or the like from a later-described learning data input/output unit 21, but the allowable requirement may be automatically generated by the learning apparatus 1 based on specification values or the like of components.

[0025] Furthermore, in the description below, the learning apparatus 1 is described as an apparatus including a computer such as a PC (personal computer) or a server apparatus, but there is no limitation to this, and it may be realized by, for example, any built-in device that has a processor, a RAM, a ROM, and the like. Furthermore, the configuration of elements installed in the apparatuses is not limited to those realized by software. The apparatuses may have a configuration realized by hardware. For example, a later-described neural network 22 may be configured by an electronic circuit such as a custom LSI (large-scale integration) or an FPGA (field-programmable gate array).

[0026] As shown in FIG. 1, the learning apparatus 1 includes a control unit 10, a machine learning unit 20, an operation classifying unit 30, and a storage unit 40.

[0027] In the vehicle 90, the control unit 10 is connected to a control sensor 91, an actuator 92, and a status detection sensor 93 provided outside the learning apparatus 1. The control unit 10 controls the actuator 92 in response to output from the control sensor 91 and the status detection sensor 93, thereby performing autonomous driving of the vehicle 90.

[0028] The control sensor 91 is a sensor group for controlling the autonomous driving of the vehicle 90. For example, the control sensor 91 is configured by a sensor for sensing obstacles outside the vehicle or a sensor for detecting the state of the road surface such as a vehicle-mounted camera or a laser, for example. Meanwhile, the status detection sensor 93 is a sensor group for detecting the control status of the vehicle 90 in an autonomous driving state. For example, the status detection sensor 93 is configured by a vibration sensor, a noise sensor, a fuel consumption sensor, a vehicle speed sensor, an acceleration sensor, a yaw rate sensor, or the like.

[0029] The actuator 92 is controlled by the control unit 10 for autonomous driving of the vehicle 90. The actuator 92 is configured by, for example, an accelerator actuator, a brake actuator, a steering actuator, or the like. The accelerator actuator controls the vehicle driving force by controlling the throttle opening degree in response to a control signal from the control unit 10. The brake actuator controls the braking force on vehicle wheels by controlling the operation amount of the brake pedal in response to a control signal from the control unit 10. The steering actuator performs control of vehicle steering action by controlling driving of a steering assistance motor of an electronic power steering system in response to a control signal from the control unit 10.

[0030] Next, the procedure in which the learning apparatus 1 performs learning will be roughly described with reference to FIG. 3. The processing in each step will be described later in detail. FIG. 3 is a flowchart showing the outline of the processing flow when the learning apparatus 1 performs learning. First, as a learning initial stage (S1), learning is performed for the purpose of making the task performable (i.e., making an operation that satisfies the learning requirements in the initial stage performable). In order to learn control of autonomous driving along a predetermined course, the learning apparatus 1 in this embodiment is provided with "do one lap of course and reach goal" as a learning requirement in the initial stage.

[0031] After the purpose in the initial stage has been cleared, operation classification (S2) is performed. In this stage, the content of the learning performed in the learning initial stage S1 is analyzed, so that the task is divided into a plurality of portions based on predetermined parameter (hereinafter, a portion obtained by dividing a task is alternatively referred to as a "scene"), and, in each divided scene, an operation that is to be performed in that scene (hereinafter, alternatively referred to as a "partial operation") is

specified from among the series of operations involved in the task. The predetermined parameter for dividing the task is, for example, a displacement amount in the operation involved in the learning of the task, or the environment in which the operation involved in the learning of the task is performed (the time elapsed from when the task is started, a relative position from where the task is started, etc.). In this embodiment, a relative position from where the task is started is used as the predetermined parameter. That is to say, in this embodiment, the learning apparatus 1 divides the task into scenes based on the position in a course, and classifies the series of operations involved in the learning into scenes based on an operation performed in units of a course corresponding to each divided scene. The efficiency of learning can be increased by performing the learning in units of partial operations classified according to the scenes. Increasing the efficiency of learning may refer to, for example, shortening the time required from when the learning is started to when the learning purpose is achieved.

[0032] In the next step after the operations have been classified, learning of an operation allowable range (S3) is performed for each classified partial operation. The operation allowable range refers to an appropriate range of a control amount that is given, during learning, to a control target whose control is learned by the control unit 10 through learning. More specifically, the operation allowable range refers to, in a case where the control unit 10 causes a control target to operate (perform a learning operation) during learning with a predetermined control amount, the range of a control amount in which the learning operation matches the allowable requirement. In this embodiment, the operation allowable range is, in a case where the control unit 10 controls the actuator 92 with a predetermined control amount, the range of a control amount in which the vehicle 90 can operate without coming in contact with another object as a result of the actuator 92 operating the vehicle 90 based on the given control amount. Specifically, as the operation allowable range, it is possible to learn values ranging from the minimum value to the maximum value of a control amount of an actuator (control of a brake, an accelerator, steering, and the like). If the learning of an operation allowable range in which the allowable requirement is satisfied is performed in advance, learning can be performed without performing an operation departing from the allowable requirements when performing learning of optimal control in subsequent steps. Accordingly, the efficiency of the learning can be further increased.

[0033] In a learning final stage (S4), the learning is optimized. In this stage, in a state where partial operations classified into scenes and learned are combined, learning for optimally performing the operation from start to end is performed. In this embodiment, learning through doing ten laps of a course within a predetermined period of time and reaching a goal is performed as the learning in the final stage.

[0034] 2. Detailed Processing

Next, the processing of the learning apparatus 1 in each step will be described in detail with reference to FIGS. 4 to 8. FIG. 4 is a block diagram showing a detailed configuration of the learning apparatus 1 according to this embodiment. As shown in FIG. 4, the machine learning unit 20 includes a learning data input/output unit 21, a neural network 22, and a learning result output unit 23. The operation classifying unit 30 includes a control data extracting unit 31 and an operation classifying result extracting unit 32.

Hereinafter, the processing of each unit will be described in detail in each step of FIG. 3.

[0035] 2-1. Learning Initial Stage

FIG. 5 is a flowchart showing details of the processing flow in the learning initial stage in step S1 shown in FIG. 3. First, in the learning initial stage (first learning process), the learning data input/output unit 21 accepts learning data (S101). The learning data is data containing, for example, a learning purpose, learning requirements, and an allowable requirement as described above.

[0036] In the next step (S102), machine learning is performed. In this embodiment, conditions that restrict individual control operations are not specified in advance, and thus learning of the control operation is performed by the learning apparatus 1 itself. Specifically, the control unit 10 operates the actuator 92 by setting a random control amount thereto. At this time, it is natural that the vehicle 90 cannot drive along the course, and thus the vehicle 90 drives in a haphazard way while sliding off the track, for example. The control unit 10 reads output (hereinafter, alternatively referred to as "sensor value") from the control sensor 91 and the status detection sensor 93 for the control amount given at random, and stores the data (the control amount and the sensor values) in the storage unit 40. The neural network 22 refers to the storage unit 40 and reads the stored control amount and sensor values, and performs learning of a control operation that matches the learning requirements through deep learning (S102).

[0037] In the learning requirements, "do one lap of course and reach goal" is set as the purpose in the initial stage level. Accordingly, in the learning apparatus 1, for example, when it is determined based on output from the control sensor 91 that the vehicle has done one lap of the course and reached the goal, the machine learning is judged to have reached the initial stage level (S103: Y), and the learning in the initial stage is ended.

[0038] 2-2. Operation Classification

FIG. 6 is a flowchart showing details of the processing flow in the operation classification in step S2 shown in FIG. 3. First, when performing the operation classification processing, the control data extracting unit 31 extracts the sensor value of the control sensor 91 when the learning initial stage ends and the control amount of the actuator 92 and the sensor value of the status detection sensor 93 corresponding thereto, from the

storage unit 40 (S201). the control data extracting unit 31 inputs the extracted values to the neural network 22 as learning data.

[0039] Next, the neural network 22 performs machine learning based on the learning data input by the control data extracting unit 31 (S202). At this time, the neural network 22 divides the course-running operation into a predetermined number of divided scenes.

[0040] Hereinafter, the processing in which the neural network 22 classifies the course-running operation into scenes will be described in more detail. The neural network 22 classifies the course-running operation into scenes based on scene vectors and operation vectors. In this embodiment, the scene vectors express a scene of the task that is performed by the vehicle 90. The scene vectors are acquired, for example, from sensor values (e.g., a position (or distance) from the start point, and a direction from the start point) that are output by the control sensor 91. For example, assuming the xy coordinates taking the start point as the origin, the scene vector at a point l can be expressed as $(l_x, _y)$.

[0041] Meanwhile, the operation vectors express the control status of the driving vehicle 90. The operation vectors are acquired, for example, from sensor values (e.g., velocity, acceleration, angular velocity, angular acceleration, etc.) that are output by the status detection sensor 93. For example, the operation vector at the point l can be expressed as (v_l, a_l) using a velocity v and an acceleration a at the point l.

[0042] The neural network 22 divides the task into scenes based on the scene vector $(l_x, _y)$, and learns, for each divided scene, operation classification that is to be learned in that scene based on the operation vector (v_l, a_l). Accordingly, the learning apparatus 1 can learn a partial operation that corresponds to a scene, by judging in which scene the learning apparatus 1 is currently present. For example, the neural network 22 focuses on the position that is expressed by the scene vector as well as the point at which the operation vector changes, thereby finding acceleration, deceleration, change of direction, and the like of the operation of the vehicle 90, so that the series of operations can be classified into operations corresponding to the scenes based on the change point. Also, for example, the neural network 22 can learn the operation classification based on the similarity levels of the operation vectors.

[0043] In the example of the course shown in FIG. 2, the task is divided into scenes corresponding to five courses (a) to (e). The partial operations classified into scenes are, for example, as follows.

Scene (a): first straight partial operation (e.g., control of deceleration timing, driving position, and the like when reaching the subsequent first corner)

Scene (b): first corner partial operation (e.g., control of steering at a corner, acceleration timing when entering the second straight, and the like)

Scene (c): second straight partial operation (e.g., control of deceleration timing,

driving position, and the like when reaching the subsequent second corner)
Scene (d): second corner partial operation (e.g., control of steering at a corner, acceleration timing, when entering the third straight and the like)
Scene (e): third straight partial operation (e.g., control of acceleration and the like when entering the first straight)

[0044] Note that the neural network 22 can preferably sort the divided scenes in the time-series order.

[0045] The operation classifying result extracting unit 32 extracts the partial operation classification learned by the neural network 22, and stores it in the storage unit 40 (S203).

[0046] 2-3. Learning of Operation Allowable Range

FIG. 7 is a flowchart showing details of the processing flow in the learning of an operation allowable range (second learning process) in step S3 shown in FIG. 3. First, the learning data input/output unit 21 selects a partial operation from among the partial operations classified in the processing in step S2, referring to the storage unit 40, and extracts the control amount for the actuator 92 necessary for this selected partial operation. The learning data input/output unit 21 performs control in the extracted control amount, and judges whether or not the allowable requirement has been satisfied as a result of the control, for example, based on output from the status detection sensor 93 and the like. At this time, for example, under the allowable requirement of not coming in contact with another object, it is judged whether or not the vehicle has come in contact with another object, whether or not the vehicle has stopped due to that contact, and the like. The learning data input/output unit 21 reads the extracted control amount and information (judgment result) as to whether or not the learning operation becomes to match the allowable requirement as a result of the control, as learning data, and gives them to the neural network 22 as learning data. The neural network 22 performs learning through deep learning based on the given learning data (S301). The neural network 22 extracts the minimum value and the maximum value of the control amount of the actuator 92 in which the allowable requirement is satisfied, as a result of the learning. The learning result output unit 23 stores the extracted minimum value and maximum value, as the operation allowable range, in the storage unit 40.

[0047] The learning apparatus 1 performs the processing in step S301 (and S302) on all classified partial operations (S303). Note that the learning result output unit 23 can output the stored operation allowable range, as a learning result of the operation allowable range. Accordingly, for example, it is possible to accept a learning result of the operation allowable range, from another learning apparatus 1' that has a similar configuration, thereby omitting the second learning process.

[0048] Although not essential, after learning the operation allowable range, for all classified partial operations, the learning apparatus 1 may perform learning again through the

series of operations (S304). Accordingly, faster course-running control can be performed.

[0049] In this manner, the learning apparatus 1 according to this embodiment performs learning of control with which the learning operation is within the operation allowable range, for classified partial operations, and thus it is possible to perform learning while avoiding such control in the subsequent learning. Accordingly, it is possible to perform a more efficient learning process.

[0050] 2-4. Optimization Learning

FIG. 8 is a flowchart showing details of the processing flow in the optimization learning (third learning process) in step S4 shown in FIG. 3. In the optimization learning, the learning that has been performed in the steps up to S3 is optimized, and thus learning for achieving the learning purpose ("do ten laps of a course within predetermined period of time and reach goal" in this embodiment) given as learning data when the learning was started is performed. In the optimization learning, it is possible to perform the learning within the operation allowable range learned in step S3. At this time, the learning data input/output unit 21 refers to the storage unit 40 and extracts learning data (which is set by an operator) input in the learning initial stage (S1 in FIG. 3). Furthermore, the learning data input/output unit 21 extracts a status of the neural network 22 after learning an operation allowable range referring to the storage unit 40. The learning data input/output unit 21 sets the extracted data to the control unit 10.

[0051] The control unit 10 outputs the control amount for the actuator 92, based on the set data described above, and acquires sensor values of the control sensor 91 and the status detection sensor 93 corresponding thereto. The control unit 10 stores the control amount and the sensor values output therefor, in the storage unit 40. At this time, it is judged whether or not the control amount that is output to the actuator 92 is within the stored operation allowable range. The control amount is output if it is within the operation allowable range, and the control amount is not output if it is out of the operation allowable range.

[0052] The neural network 22 reads the control amount and the sensor values stored by the control unit 10 in the above-described processing, and performs learning through deep learning (S401). Accordingly, the neural network 22 can more efficiently learn a control operation that matches the learning requirements from the start to the end of the operation (i.e., from the start to the goal of the course), in a state where an operation allowable range has been learned. The processing in step S401 is repeatedly performed until the entire learning is optimized (S402). A result of the optimization learning is extracted by the learning result output unit 23, and stored in the storage unit 40. Accordingly, in the optimization learning, it is possible to perform learning within the operation allowable range.

[0053] In this manner, with the learning apparatus 1 according to this embodiment, the learning apparatus 1 itself can classify an operation involved in the learning into partial operations, and perform learning. Accordingly, individual optimization can be performed for each classified operation, and the learning can be performed more efficiently (i.e., in a shorter period of time). Moreover, with the learning apparatus 1 according to this embodiment, when learning a partial operation, first, an operation allowable range is learned. Accordingly, the learning can be efficiently performed without a person setting detailed conditions in advance for each operation.

[0054] Hardware Configuration

Hereinafter, an example of the hardware configuration in a case where the learning apparatus 1 described above is realized by a computer 800 will be described with reference to FIG. 9. Note that the configuration of each apparatus may be realized by a plurality of separate devices.

[0055] As shown in FIG. 9, the computer 800 includes a processor 801, a memory 803, a storage apparatus 805, an input interface unit (input I/F unit) 807, a data interface unit (data I/F unit) 809, a communication interface unit (communication I/F unit) 811, and a display apparatus 813.

[0056] The processor 801 controls various types of processing in the computer 800 by executing programs stored in the memory 803. For example, as a result of the processor 801 executing programs stored in the memory 803, the control unit 10, the machine learning unit 20, the operation classifying unit 30, and the like of the learning apparatus 1 can be realized.

[0057] The memory 803 is a storage medium, for example, such as a RAM (random access memory). The memory 803 temporarily stores program code of programs that are executed by the processor 801, and data that is required at the time when the programs are executed.

[0058] The storage apparatus 805 is, for example, an auxiliary memory such as a hard disk drive (HDD) or a solid state drive, or a non-volatile storage medium such as a flash memory. The storage apparatus 805 stores various programs for realizing the operating system or the above-described configurations. These programs and data are loaded onto the memory 803 as necessary, and referred to by the processor 801. For example, the storage unit 40 described above is realized by the storage apparatus 805.

[0059] The input I/F unit 807 is a device for accepting input from an administrator. Specific examples of the input I/F unit 807 include a keyboard, a mouse, a touch panel, various sensors, a wearable device, and the like. The input I/F unit 807 may be connected to the computer 800, for example, via an interface such as a USB (universal serial bus).

[0060] The data I/F unit 809 is a device for inputting data from the outside of the computer 800. Specific examples of the data I/F unit 809 include drive devices and the like for

reading data stored in various storage media. The data I/F unit 809 may be provided external to the computer 800. In that case, the data I/F unit 809 is connected to the computer 800, for example, via an interface such as a USB.

[0061] The communication I/F unit 811 is a device for performing wired or wireless data communication with an apparatus outside the computer 800 via the Internet N. the communication I/F unit 811 may be provided external to the computer 800. In that case, the communication I/F unit 811 is connected to the computer 800, for example, via an interface such as a USB.

[0062] The display apparatus 813 is a device for displaying various types of information. Specific examples of the display apparatus 813 include a liquid crystal display, an organic EL (electro-luminescence) display, a wearable display, and the like. the display apparatus 813 may be provided external to the computer 800. In that case, the display apparatus 813 is connected to the computer 800, for example, via a display cable or the like.

[0063] Second Embodiment

In the first embodiment, an example was described in which the learning apparatus 1 is used for the vehicle 90 controlled to drive autonomously. However, the apparatus to which the learning apparatus 1 is applied is not limited to the example shown in the first embodiment, and the learning apparatus 1 can be applied to various apparatuses. In this embodiment, an example will be described in which the learning apparatus 1 is applied to the control of a robot whose task is to wrap paper cups. In the second embodiment, mainly differences from the first embodiment will be described.

[0064] First, differences from the first embodiment in terms of the system configuration according to this embodiment will be described with reference to FIG. 10. The configuration of the learning apparatus 1 is as in the first embodiment. On the other hand, the configuration outside the learning apparatus 1 in this embodiment is such that a control sensor 91' is configured by a sensor group for performing a pick-and-place operation. Specifically, it is configured by a workpiece detection sensor (image sensor), a robot holding force sensor, or the like. Furthermore, the control sensor 91' includes an image recognition algorithm, and can recognize the shape of a workpiece that it is holding. The other portions of the configuration outside the learning apparatus 1 are as described in the first embodiment.

[0065] Next, differences between learning according to this embodiment and learning according to the first embodiment will be described. The pick-and-place operation that is a task according to this embodiment refers to an operation that is performed following the procedure below.

1) Hold only one paper cup, and place it at predetermined position

2) Repeat the above-stated operation (hold + place) for predetermined number of

times, thereby stacking predetermined number of paper cups

3) Place one group of predetermined number of stacked paper cups in bag

4) Twist end of bag, and clip opening of bag with band

5) Place one bag whose opening has been clipped with band, in cardboard box

6) Repeat the above-stated operations 1) to 5) for predetermined number of times, thereby placing another bag whose opening has been clipped with band, in cardboard box

7) Repeat the above-stated operation 6) for predetermined number of times, and close lid of cardboard box in which bags are contained

[0066] In learning of robot control according to this embodiment, a given learning purpose, learning requirements, and allowable requirement are as follows.

[0067] Learning Purpose

- Complete the above-stated series of operations within predetermined period of time

Learning Requirements

- Do not allow cups to stick out of bag or cardboard box

- "Complete series of operations" (i.e., "making series of operations involved in the task controllable") in initial stage level

Allowable Requirement

- Do not damage paper cup, bag, cardboard box

[0068] With the learning apparatus 1 to which the above-described learning data is given according to this embodiment, the minimum value and the maximum value of a force applied when holding a cardboard box or a paper cup are learned as the operation allowable range.

[0069] Above, embodiments of the present invention have been described. The foregoing embodiments are for the purpose of facilitating understanding of the present invention, and are not to be interpreted as limiting the present invention. The invention can be altered and improved without departing from the gist thereof. For example, the steps in the above-described processing flows can be partially omitted, be rearranged in any desired order, or be executed in parallel, as long as doing so does not cause conflict in the processing content.

[0070] In the foregoing embodiments, examples were described in which the system according to the present invention is used to manage abilities acquired by machines according to AI technology such as deep learning, but the present invention is not limited thereto, and can be applied to a wide variety of fields. Examples thereof include distinguishing non-defective products from defective products, foods, machine parts, chemical products, drugs, and the like in various fields including industrial fields, fishery fields, agricultural fields, forestry fields, service industries, and medical and health fields. Furthermore, the present invention can be applied to cases where AI

technology is applied to products in embedding fields, systems such as social systems using IT techniques, analysis of big data, classifying processing in a wide variety of control apparatuses, and the like.

[0071] Note that, in this specification, a "portion", "unit" or "procedure" does not merely mean a physical configuration, and there are also cases where processing that is performed by a "portion" or the like is realized by software. Furthermore, processing that is performed by one "portion" or the like or apparatus may be performed by two or more physical configurations or apparatuses, and processing that is performed by two or more "portions" or the like or apparatuses may be performed by one physical unit or apparatus.

[0072] Note that part of or the entirety of the foregoing embodiments may be described as in Additional Remarks below, but there is no limitation to this.

Additional Remark 1

A learning apparatus including at least one hardware processor,

wherein the hardware processor

accepts learning data containing a learning purpose and an allowable requirement for a learning operation performed when learning control,

performs learning based on the learning data, and

outputs a learning result obtained by a neural network,

wherein the performing learning includes performing a first learning process for achieving an initial stage of the learning purpose, performing a second learning process for learning a control range in which the learning operation matches the allowable requirement, based on a result of the first learning process, and performing a third learning process for achieving the learning purpose within the control range based on a result of the second learning process.

Additional Remark 2

A learning method that is performed by at least one hardware processor for executing a step of performing learning, the method including:

a step of accepting learning data containing a learning purpose and an allowable requirement for a learning operation performed when learning control;

a step of performing learning based on the learning data; and

a step of outputting a learning result obtained in the step of performing learning,

wherein the step of performing learning includes a step of performing a first learning process for achieving an initial stage of the learning purpose, performing a second learning process for learning a control range in which the learning operation matches the allowable requirement, based on a result of the first learning process, and performing a third learning process for achieving the learning purpose within the control range based on a result of the second learning process.

Claims

[Claim 1] A learning apparatus configured to learn control of an operation involved in a predetermined task, the apparatus comprising:
a learning data accepting unit configured to accept learning data containing a learning purpose and an allowable requirement for a learning operation performed when learning the control;
a neural network configured to perform learning based on the learning data; and
an output unit configured to output a learning result obtained by the neural network,
wherein the neural network performs a first learning process for achieving an initial stage of the learning purpose, performs a second learning process for learning a control range in which the learning operation matches the allowable requirement, based on a result of the first learning process, and performs a third learning process for achieving the learning purpose within the control range based on a result of the second learning process.

[Claim 2] The learning apparatus according to claim 1, wherein the output unit outputs the result of the second learning process.

[Claim 3] The learning apparatus according to claim 1,
wherein the learning apparatus is a learning apparatus configured to learn control of a series of operations involved in a predetermined task,
the learning apparatus further comprises a classifying unit configured to divide the task into a plurality of scenes, and, for each of the divided scenes, specify a partial operation that is to be performed in the scene, from among the series of operations, and
the neural network performs the second learning process and the third learning process for each partial operation.

[Claim 4] An autonomous driving control learning apparatus configured to learn control of a series of operations involved in autonomous driving of a vehicle that does laps of a predetermined course, the apparatus comprising:
a learning data accepting unit configured to accept learning data containing a learning purpose in which a purpose is to do a predetermined number of laps of the course within a predetermined period of time, and an allowable requirement for a learning operation performed when learning the control;

a neural network configured to perform learning based on the learning data; and

an output unit configured to output a learning result obtained by the neural network,

wherein the neural network performs a first learning process for achieving running of one lap of the course by the vehicle, performs a second learning process for learning a control range in which the learning operation matches the allowable requirement, based on a result of the first learning process, and performs a third learning process for achieving the learning purpose within the control range based on a result of the second learning process.

[Claim 5] A robot control learning apparatus configured to learn control of a series of operations involved in a task of carrying a predetermined target to a predetermined placement position, and stacking and wrapping a predetermined number of the targets, the apparatus comprising:

a learning data accepting unit configured to accept learning data containing a learning purpose in which a purpose is to complete the task as soon as possible, and an allowable requirement for a learning operation performed when learning the control;

a neural network configured to perform learning based on the learning data; and

an output unit configured to output a learning result obtained by the neural network,

wherein the neural network performs a first learning process for achieving control of the series of operations involved in the task, performs a second learning process for learning a control range in which the learning operation matches the allowable requirement, based on a result of the first learning process, and performs a third learning process for achieving the learning purpose within the control range based on a result of the second learning process.

[Claim 6] A learning method for learning control of an operation involved in a predetermined task, the method comprising:

accepting, by a computer, learning data containing a learning purpose and an allowable requirement for a learning operation performed when learning the control;

performing, by the computer, learning based on the learning data; and

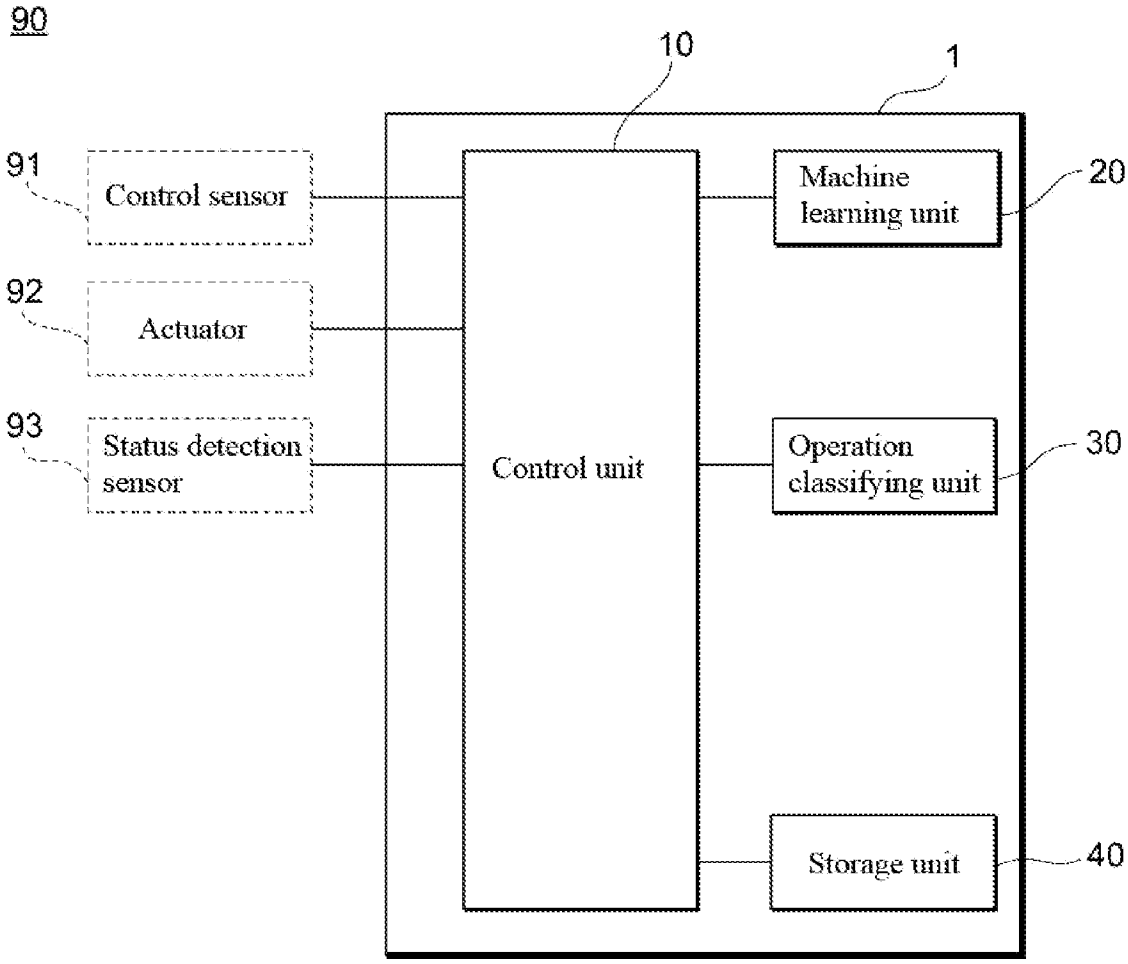
outputting, by the computer, a learning result obtained in the learning,

wherein performing learning includes performing, by the computer, a first learning process for achieving an initial stage of the learning purpose, performing, by the computer, a second learning process for learning a control range in which the learning operation matches the allowable requirement, based on a result of the first learning process, and performing, by the computer, a third learning process for achieving the learning purpose within the control range based on a result of the second learning process.

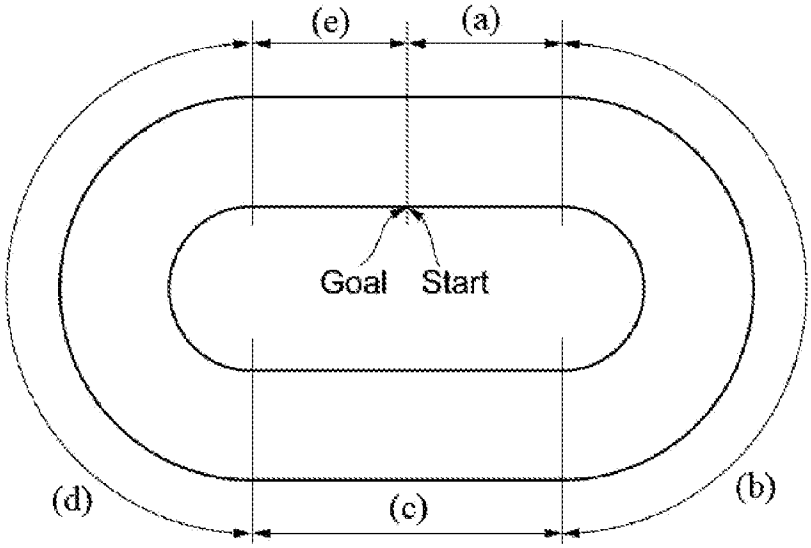
[Claim 7] A program for causing a computer configured to learn control of an operation involved in a predetermined task to execute:
accepting learning data containing a learning purpose and an allowable requirement for a learning operation performed when learning the control;
performing learning based on the learning data; and
outputting a learning result obtained in the learning,
wherein performing learning includes performing a first learning process for achieving an initial stage of the learning purpose, performing a second learning process for learning a control range in which the learning operation matches the allowable requirement, based on a result of the first learning process, and performing a third learning process for achieving the learning purpose within the control range based on a result of the second learning process.

[Claim 8] An apparatus configured to perform a predetermined task, the apparatus comprising:
a first sensor configured to sense information necessary for an operation of the apparatus to perform the task;
an actuator;
a second sensor configured to sense a change in a status of the apparatus caused by the actuator;
a control unit configured to control the actuator based on sensor values output from the first sensor and the second sensor; and
a storage unit configured to store a learning result obtained by the learning apparatus according to any one of claims 1 to 3,
wherein the control unit determines a control amount that corresponds to the sensor values output from the first sensor and the second sensor, within the control range, based on the learning result stored in the storage unit.

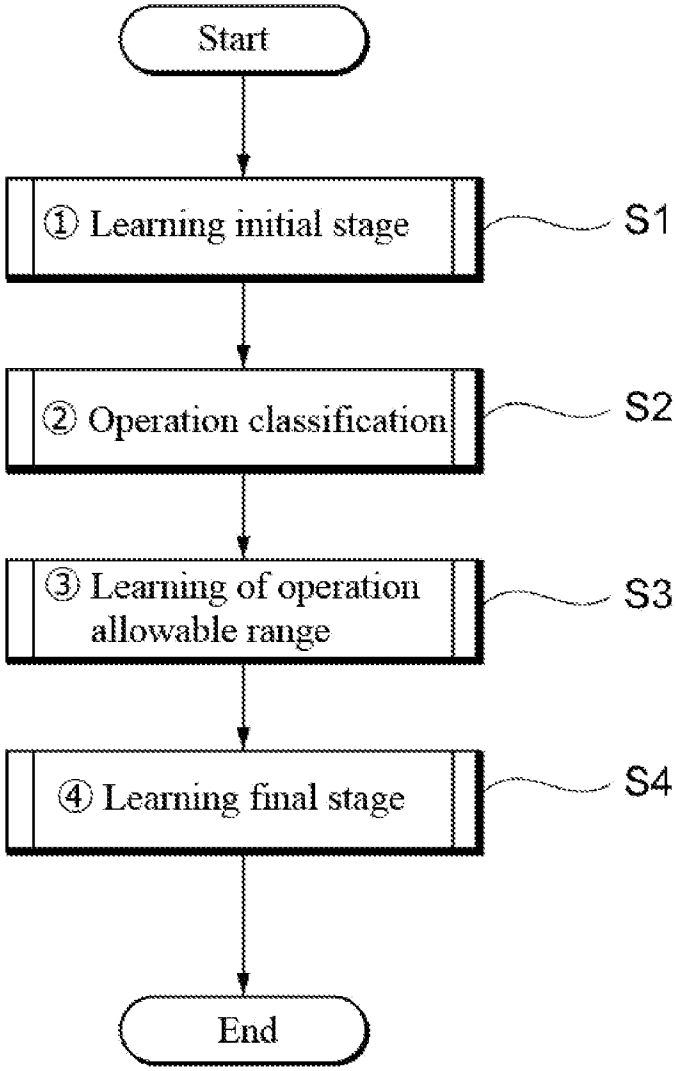
[Fig. 1]


90
10
1
91
Control sensor
92
Actuator
93
Status detection sensor
Control unit
Machine learning unit
20
Operation classifying unit
30
Storage unit
40

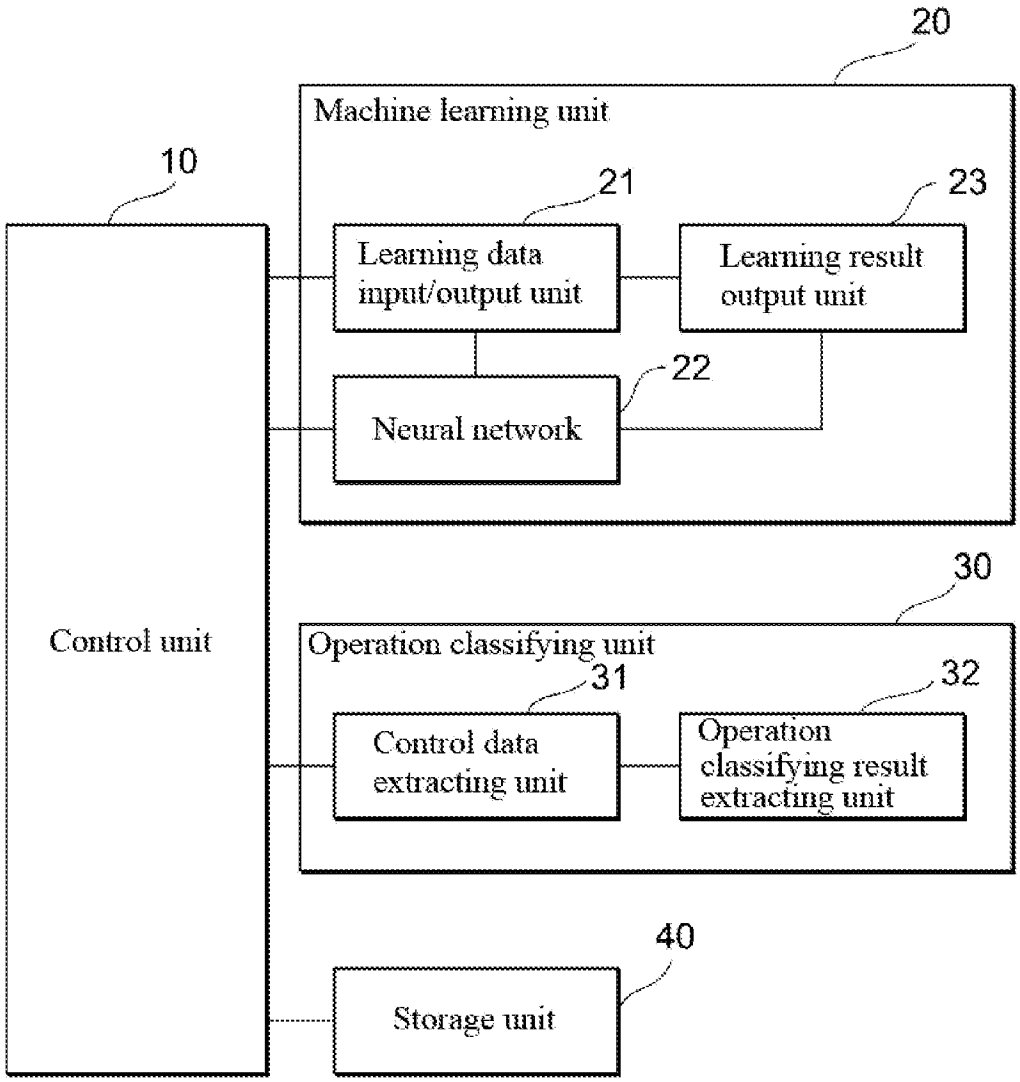
[Fig. 2]


(e)
(a)
Goal
Start
(d)
(c)
(b)

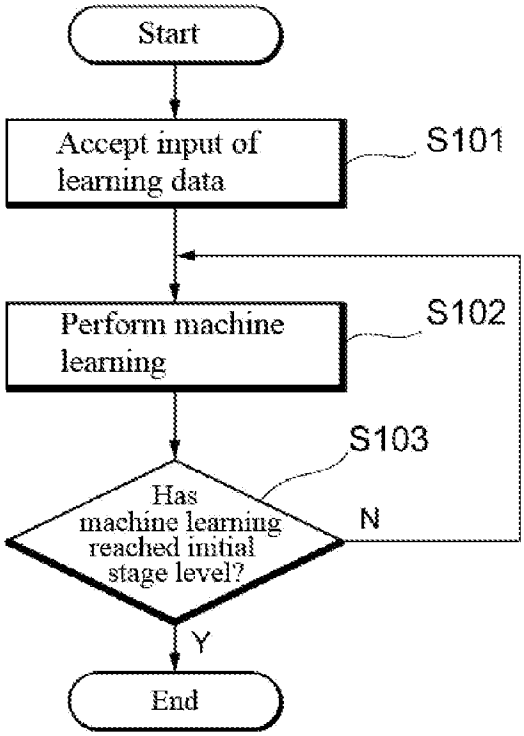
[Fig. 3]


Start
① Learning initial stage
S1
② Operation classification
S2
③ Learning of operation allowable range
S3
④ Learning final stage
S4
End

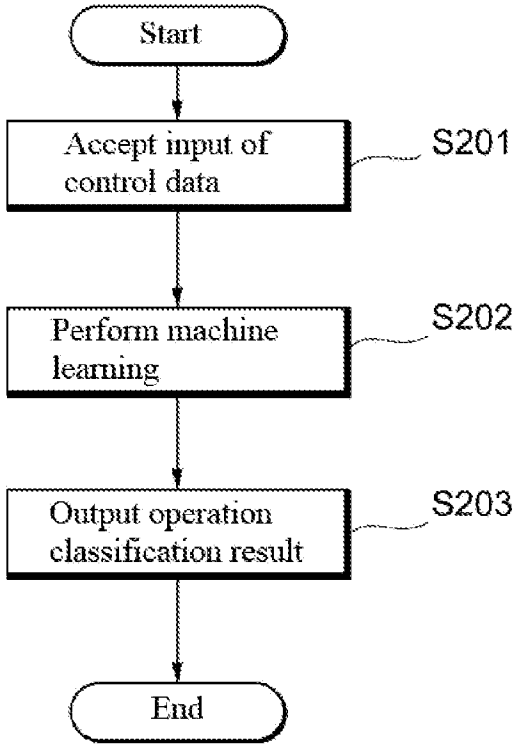
[Fig. 4]


1
20
Machine learning unit
10
21
23
Learning data input/output unit
Learning result output unit
22
Neural network
30
Control unit
Operation classifying unit
31
32
Control data extracting unit
Operation classifying result extracting unit
40
Storage unit

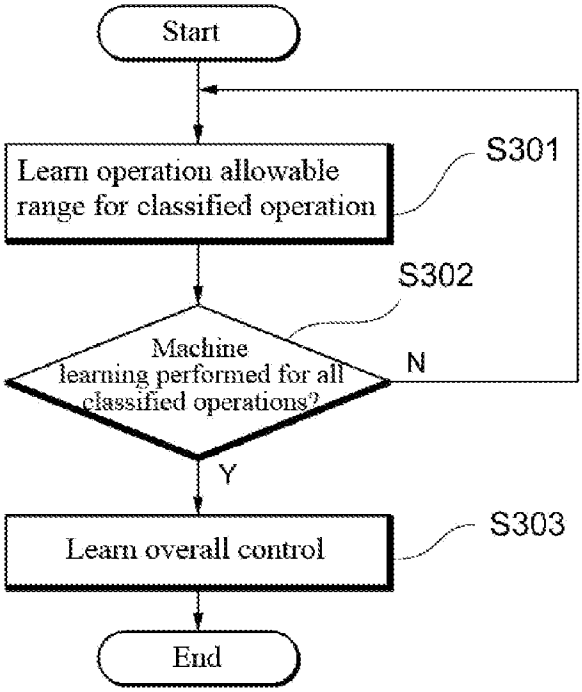
[Fig. 5]


Start
Accept input of learning data
S101
Perform machine learning
S102
S103
Has machine learning reached initial stage level?
N
Y
End

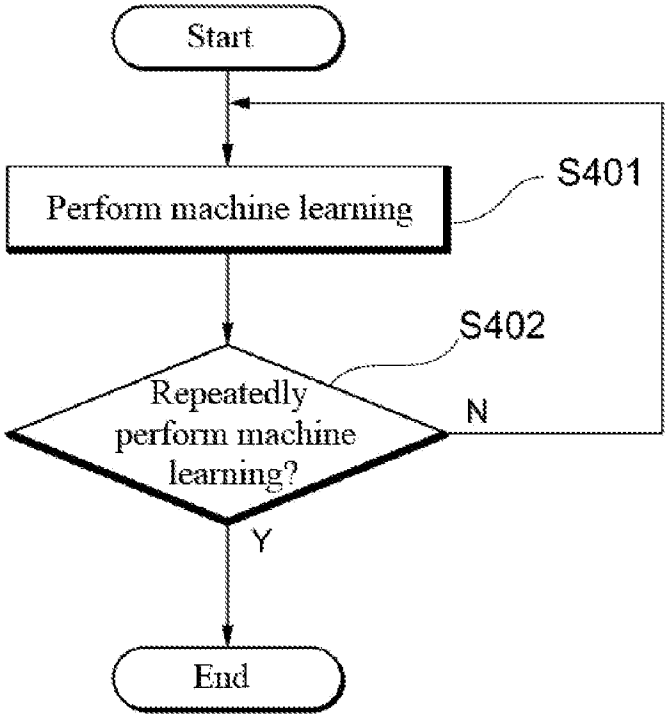
[Fig. 6]


Start
Accept input of control data
S201
Perform machine learning
S202
Output operation classification result
S203
End

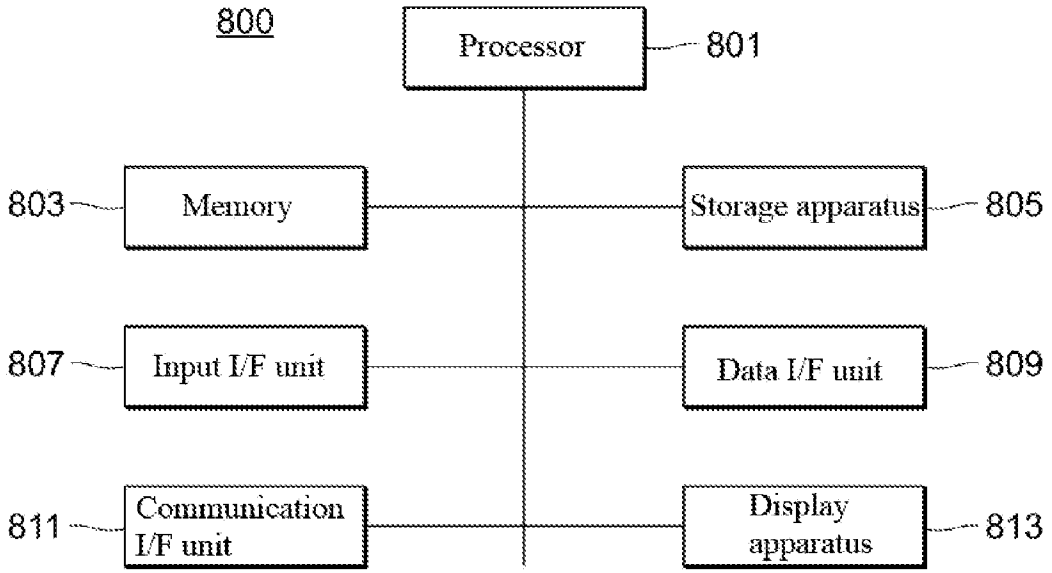
[Fig. 7]


Start
Learn operation allowable range for classified operation
S301
S302
Machine learning performed for all classified operations?
N
Y
Learn overall control
S303
End

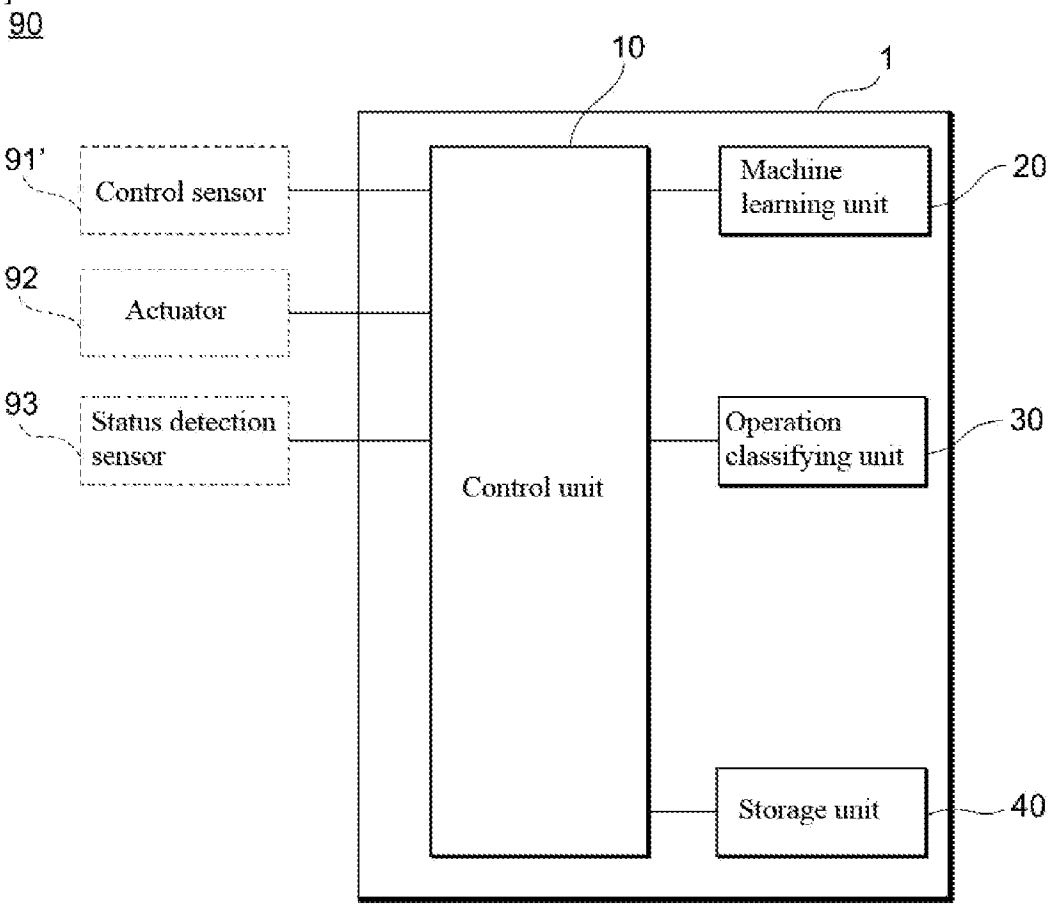
[Fig. 8]


Start
Perform machine learning
S401
S402
Repeatedly perform machine learning?
N
Y
End

[Fig. 9]


800
Processor
801
803
Memory
Storage apparatus
805
807
Input I/F unit
Data I/F unit
809
811
Communication I/F unit
Display apparatus
813

[Fig. 10]


90
10
1
91'
Control sensor
Machine learning unit
20
92
Actuator
93
Status detection sensor
Control unit
Operation classifying unit
30
Storage unit
40

INTERNATIONAL SEARCH REPORT

International application No
PCT/JP2018/018133

A. CLASSIFICATION OF SUBJECT MATTER
INV. B25J9/16 G05B13/02 G05D1/00 G05D1/02 G06N3/08
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
B25J G05B G05D G06N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 9 015 093 B1 (COMMONS MICHAEL LAMPORT [US]) 21 April 2015 (2015-04-21)	1,2,4-8
Y	column 1 column 14 - column 15 column 19 - column 20 column 46 - column 48 figures 1,11-12	3

Y	US 2015/127155 A1 (PASSOT JEAN-BAPTISTE [US] ET AL) 7 May 2015 (2015-05-07)	3
A	paragraph [0005] paragraph [0057] - paragraph [0058] paragraph [0108] - paragraph [0109]	4,5

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
9 August 2018	27/08/2018

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Bassi, Luca

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2018/018133

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☒ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-8

Learning apparatus

1.1. claim: 8

Automatically controlled apparatus

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/JP2018/018133

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 9015093 B1	21-04-2015	US 9015093 B1	21-04-2015
		US 9875440 B1	23-01-2018
US 2015127155 A1	07-05-2015	US 2015127154 A1	07-05-2015
		US 2015127155 A1	07-05-2015
		US 2017203437 A1	20-07-2017