



US 20210406781A1

(19) **United States**

(12) **Patent Application Publication**
KURAUCHI et al.

(10) **Pub. No.: US 2021/0406781 A1**

(43) **Pub. Date: Dec. 30, 2021**

(54) **TRAINING DEVICE, METHOD AND PROGRAM**

Publication Classification

(71) Applicant: **NIPPON TELEGRAPH AND TELEPHONE CORPORATION**,
Tokyo (JP)

(51) **Int. Cl.**
G06N 20/20 (2006.01)
G06K 9/62 (2006.01)
(52) **U.S. Cl.**
CPC **G06N 20/20** (2019.01); **G06K 9/6256**
(2013.01)

(72) Inventors: **Yuki KURAUCHI**, Tokyo (JP); **Naoto ABE**, Tokyo (JP); **Hitoshi SESHIMO**, Tokyo (JP)

(57) **ABSTRACT**

(73) Assignee: **NIPPON TELEGRAPH AND TELEPHONE CORPORATION**,
Tokyo (JP)

It is possible to learn a model for accurately estimating a label indicating the condition of data. A batch size, which is a unit of learning data used in machine learning, is set to a predetermined size, learning data having the batch size in a learning data set is used to learn a first model for determining, for each label, the likelihood of the label, and time-series likelihood data that is the likelihood of each label at each time is output for each piece of the learning data. The time-series likelihood data for each piece of the learning data given a correct label is input, the batch size is set to a size larger than the predetermined size in a first learning unit 24, and machine learning is performed to learn a second model for outputting one of the labels based on a change in the likelihood of each label.

(21) Appl. No.: **17/293,458**

(22) PCT Filed: **Nov. 12, 2019**

(86) PCT No.: **PCT/JP2019/044332**

§ 371 (c)(1),

(2) Date: **May 12, 2021**

(30) **Foreign Application Priority Data**

Nov. 16, 2018 (JP) 2018-215917

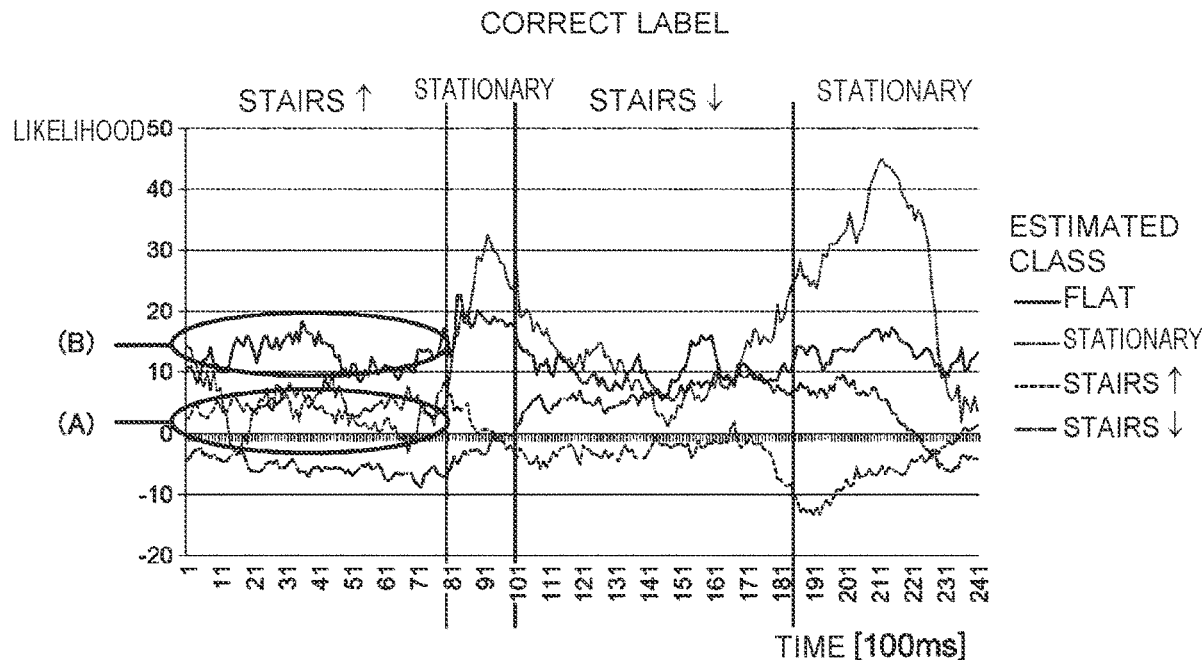


Fig. 1

CORRECT LABEL

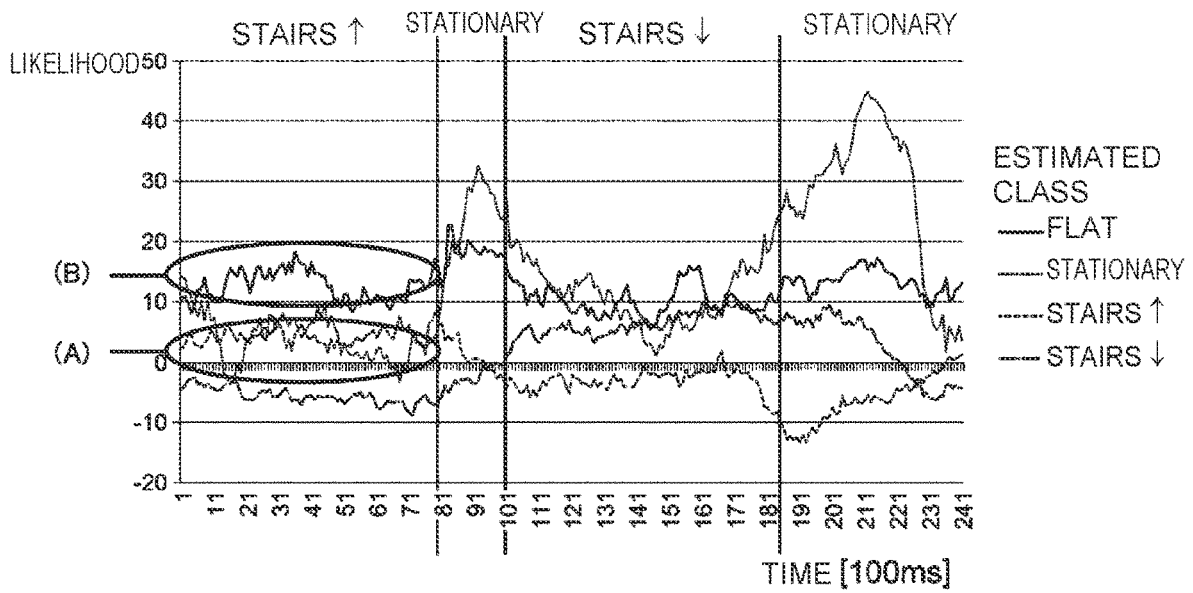


Fig. 2

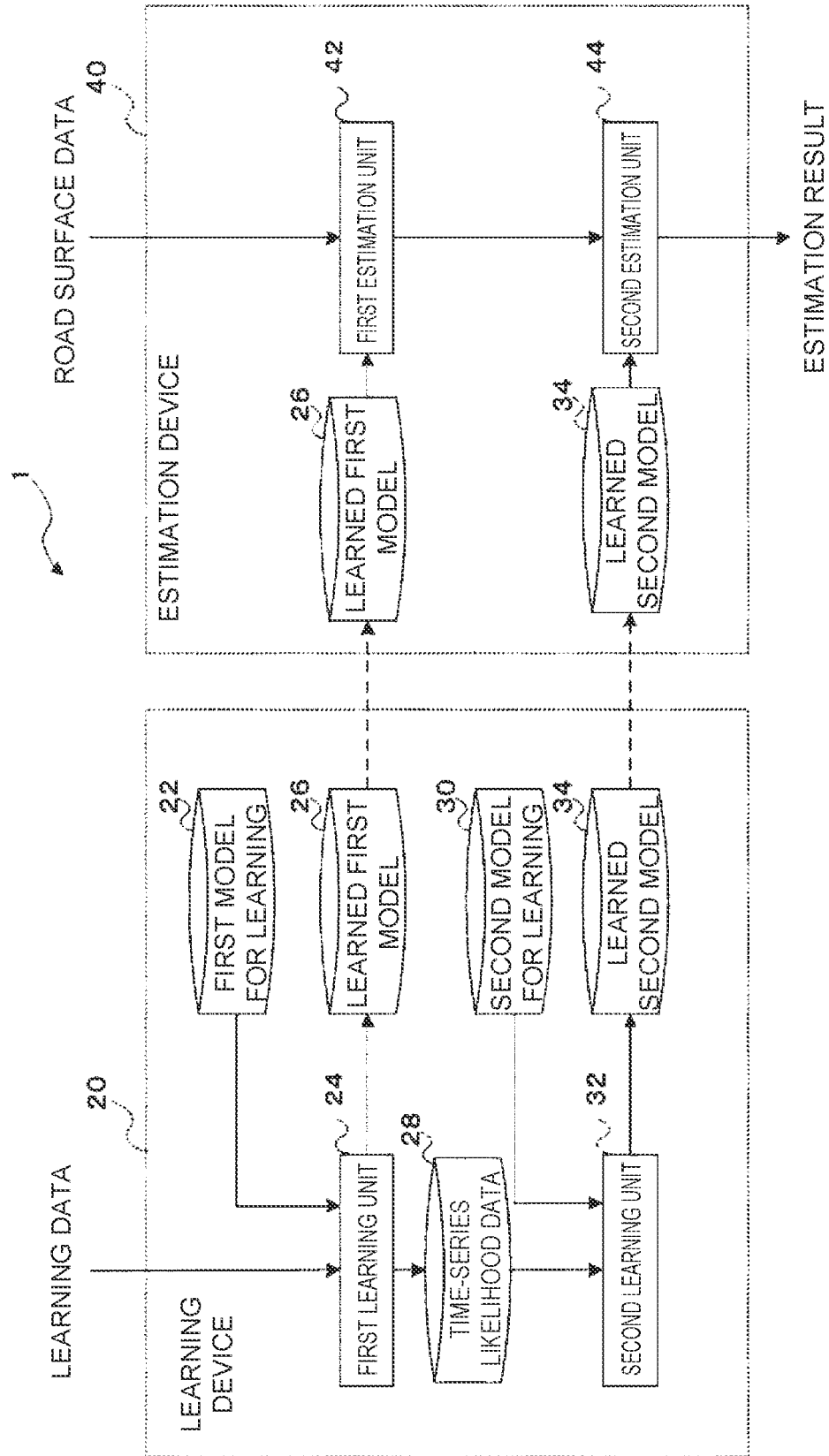


Fig. 3

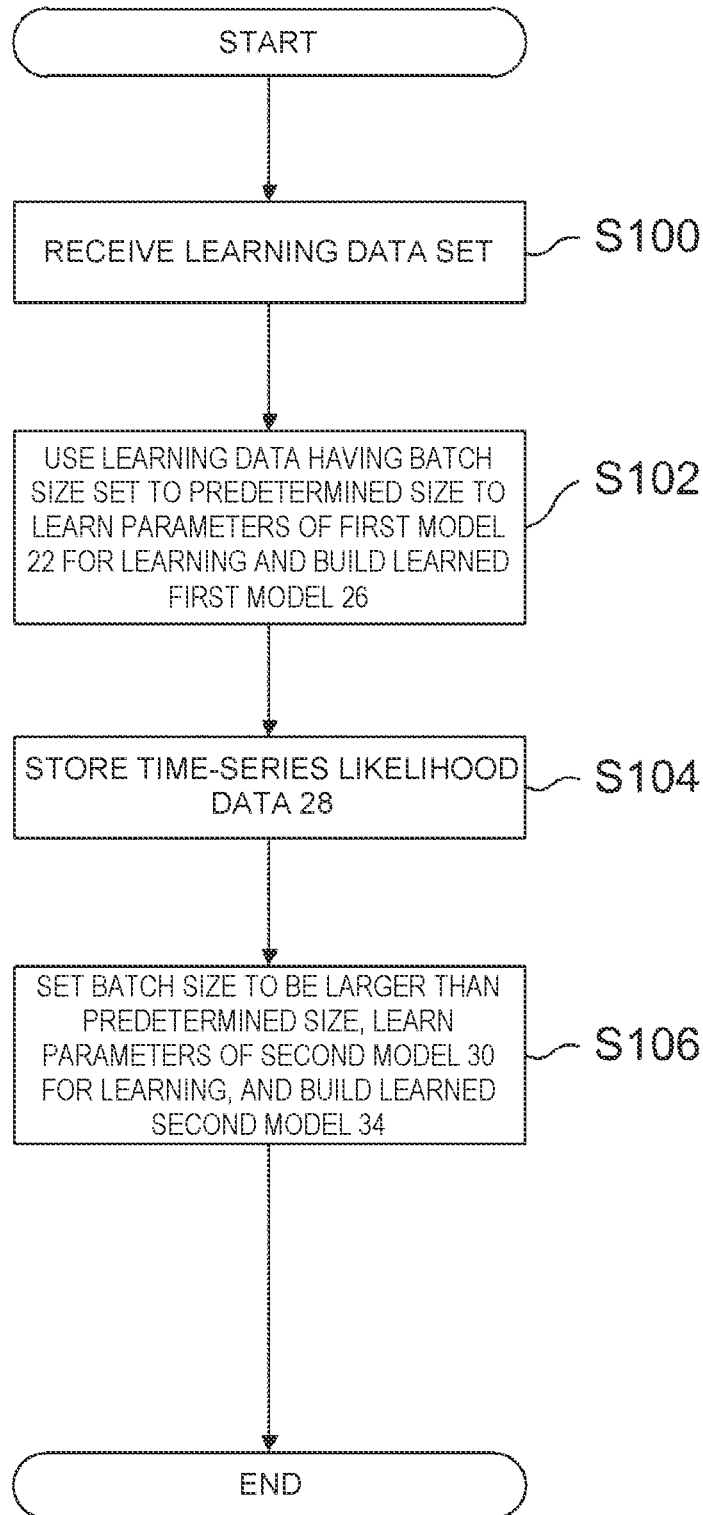
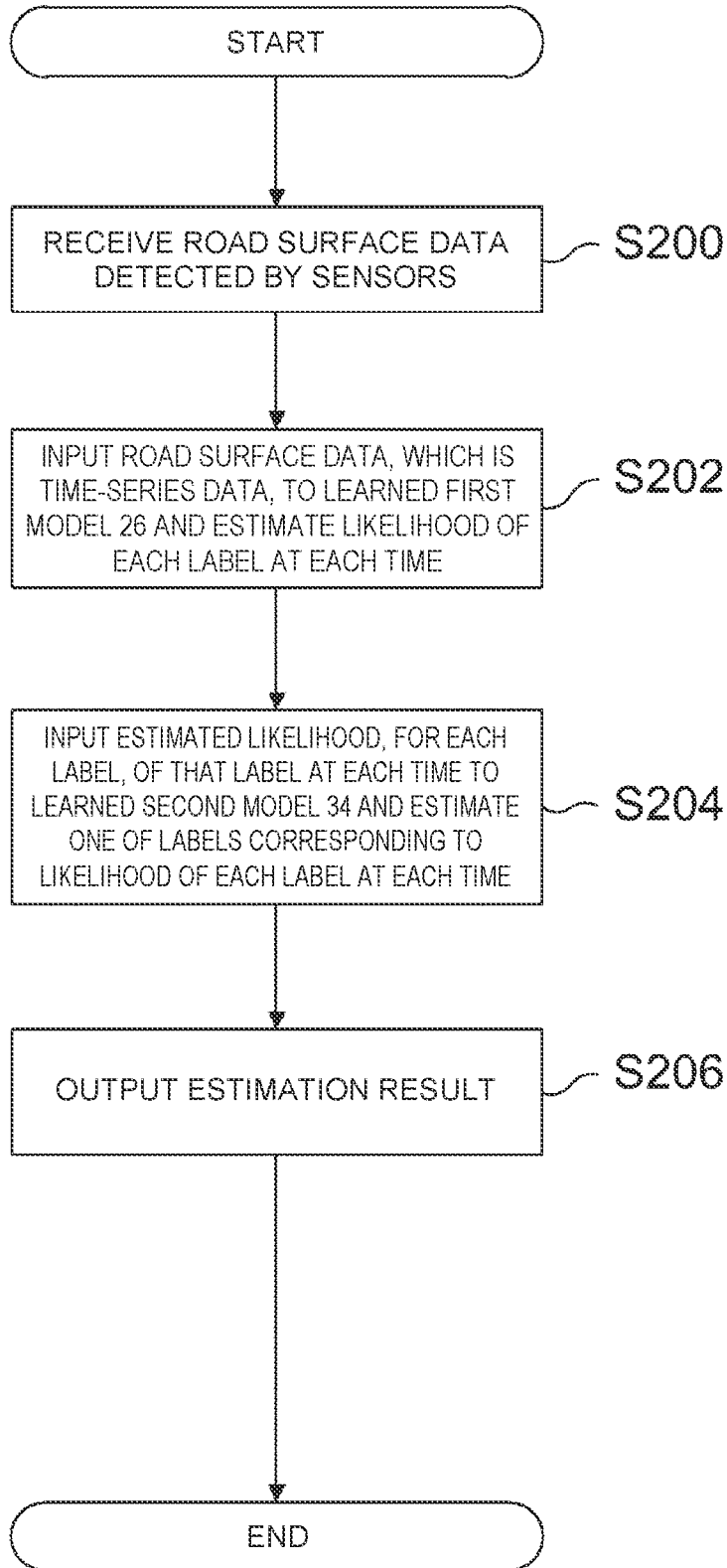


Fig. 4



TRAINING DEVICE, METHOD AND PROGRAM

TECHNICAL FIELD

[0001] The present invention relates to a learning device, a method, and a program, and particularly to a learning device, a method, and a program for estimating the state of a target.

BACKGROUND ART

[0002] There have been studied techniques for using a sensor mounted on a moving body such as an automobile, a pedestrian, or a wheelchair moving on a road surface such as a sidewalk or a roadway to estimate the condition (step, slope, etc.) of the road surface on which the moving body moves (see, e.g., Non-Patent Literatures 1 and 2).

CITATION LIST

Non-Patent Literature

[0003] Non-Patent Literature 1: Akihiro Miyata, Iori Araki, Tongshun Wang, and Tenshi Suzuki, "A Study on Barrier Detection Using Sensor Data of Unimpaired Walkers", IPSJ Journal (2018)

[0004] Non-Patent Literature 2: "Unevenness on Road Surface is Detected Using Acceleration Sensor of Smartphone Mounted on Expressway Bus, and Verification Test is Conducted", (online), (searched on Nov. 6, 2018), Internet <URL: <https://sgforum.impress.co.jp/news/3595>>

SUMMARY OF THE INVENTION

Technical Problem

[0005] The estimation of the condition of a road surface as described above is often performed using a model built by machine learning using learning data. However, depending on the condition of the road surface, there is a problem that a desired estimation result cannot be obtained and the estimation accuracy is not sufficient.

[0006] The present invention has been made in view of the above circumstances, and aims at providing a learning device, a method, and a program capable of learning a model for accurately estimating a label indicating the condition of data.

Means for Solving the Problem

[0007] In order to achieve the above object, a learning device according to a first invention is configured to include: a first learning unit that receives, as input, a learning data set including learning data that is time-series data and is given one of a plurality of types of labels as a correct label at each time, sets a batch size which is a unit of learning data used in machine learning to a predetermined size, uses learning data having the batch size in the learning data set to perform predetermined machine learning to learn a first model for estimating a label, and outputs an estimation result of the label at each time for each piece of learning data; and a second learning unit that receives, as input, the estimation result of the label at each time for each piece of the learning data given a correct label, sets the batch size to a size larger than the predetermined size, and performs predetermined

machine learning to learn a second model for outputting one of the labels based on the estimation result of the label at each time.

[0008] A learning method according to a second invention is characterized by including and executing: a step of a first learning unit receiving, as input, a learning data set including learning data that is time-series data and is given one of a plurality of types of labels as a correct label at each time, setting a batch size, which is a unit of learning data used in machine learning, to a predetermined size, using learning data having the batch size in the learning data set to perform predetermined machine learning to learn a first model for estimating a label, and outputting an estimation result of the label at each time for each piece of learning data; and a step of a second learning unit receiving, as input, the estimation result of the label at each time for each piece of the learning data given a correct label, setting the batch size to a size larger than the predetermined size, and performing predetermined machine learning to learn a second model for outputting one of the labels based on the estimation result of the label at each time.

[0009] A program according to a third invention is a program for causing a computer to function as each unit of the learning device according to the first invention.

Effects of the Invention

[0010] According to the learning device, the method, and the program of the present invention, it is possible to obtain an effect that a model for accurately estimating a label indicating the condition of data can be learned.

BRIEF DESCRIPTION OF DRAWINGS

[0011] FIG. 1 is a diagram showing an example of a result of estimation of the condition of a road surface on which a moving body moves by conventional machine learning.

[0012] FIG. 2 is a block diagram showing a configuration of an estimation system including a learning device and an estimation device according to an embodiment of the present invention.

[0013] FIG. 3 is a flowchart showing a processing routine in the learning device according to the embodiment of the present invention.

[0014] FIG. 4 is a flowchart showing a processing routine in the estimation device according to the embodiment of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

[0015] An embodiment of the present invention will be described below in detail with reference to the drawings.

Overview of Embodiment of the Present Invention

[0016] First, an overview of the embodiment of the present invention will be described. In the embodiment of the present invention, the estimation accuracy is improved by performing learning by stepwise setting a batch size in machine learning.

[0017] FIG. 1 shows an example of a result of estimation of the condition of a road surface on which a moving body moves by conventional machine learning. FIG. 1 shows a correct label and an estimation result of likelihood of each estimated class, and is a graph in which the vertical axis represents likelihood and the horizontal axis represents the

time of time series. The time of time series is in units of 100 ms. Here, the unit of 100 ms is the unit time for shifting a batch at the time of learning, and the graph in FIG. 1 shows that there is a time transition of 100 ms per unit. Hereinafter, a “time” shown in the embodiment of the present invention represents a unit, time of 100 ms. The estimated classes for the condition of the road surface on which the moving body moves are “flat” indicating a flat road, “stationary” indicating that the moving body is in a stationary state, “stairs ↑” indicating ascending stairs, and “stairs ↓” indicating descending stairs. The correct label is assigned to each time of the time series, but in the likelihood in the estimation result, the likelihood of the desired label may not always be the highest. For example, the correct label for times of 1-71 of the time series is “stairs ↑”, but the likelihood of the label “flat” shown in (B) is higher than the likelihood of the label “stairs ↑” shown in (A), resulting in a problem that the correct result cannot be obtained. One of the causes of such an estimation result is that the number of correct labels in learning data is biased. For example, if the labels are subdivided and a label “2 cm step ↑” indicating a 2 cm ascending step is used, such a label is expected to be low in appearance frequency in the learning data.

[0018] Further, in machine learning, learning data is divided into batches per batch size for learning. In general, there is a problem that the estimation accuracy decreases when a batch defined as having the batch size does not include all the labels, while the learning accuracy decreases when the batch size is too large. Therefore, in the embodiment of the present invention, learning is performed so as to correct the estimation accuracy of a label that is low in appearance frequency using two-stage learning that makes use of the high-accuracy learning in the case of a small batch size and also performs learning in an enlarged batch size. For example, if there is a label that appears once every 1000 times, with a unit of 100 ms as one batch, the batch size is set to 1024, which is larger than 1000, or to 2048, 4096, etc. to have some margin. This makes it possible to estimate a desired label even when the learning data is biased.

[0019] The embodiment of the present invention will be described based on the above premises.

Configuration according to Embodiment of the Present Invention

[0020] Next, a configuration according to the embodiment of the present invention will be described. As shown in FIG. 2, an estimation system 1 according to the embodiment of the present invention is configured to include a learning device 20 and an estimation device 40. Each of the learning device 20 and the estimation device 40 can be configured with a computer including a CPU, a RAM, and a ROM that stores a program for executing the processing of the operation described later and various data.

[0021] First, the learning device 20 will be described. A learning data set composed of learning data that is time-series data is input to the learning device 20. Road surface data detected in time series by a plurality of types of sensors that detect the state of the moving body that is the estimation target is used as the learning data. Further, the learning data is given one of a plurality of types of labels as the correct label at each time. The labels are types of the condition of the road surface on which the moving body moves, such as “flat”, “stationary”, “stairs ↑”, “stairs ↓”, etc. By subdividing the condition of the road surface, a label such as “2 cm

step ↑” may be used. In this embodiment, it is assumed to learn a model for determining the likelihood of each label from input data obtained from road surface data by shifting a window having a width of 1500 ms in increments of 100 ms. The correct class is the condition of the road surface corresponding to the point of 750 ms, which is the center of the window. When learning and estimation are performed in real time, the window width may be shorter than 1500 ms. As the sensors, various sensors such as an acceleration sensor, a gyro sensor, a geomagnetic sensor, a gravity sensor, an atmospheric pressure sensor, and a tilt sensor can be used as appropriate according to the estimation target.

[0022] The learning device 20 includes a first learning unit 24 and a second learning unit 32. The first learning unit 24 uses a first model 22 for learning to build a learned first model 26. The first model 22 for learning is a model for determining, for each label, the likelihood of that label. The second learning unit 32 uses a second model 30 for learning to build a learned second model 34. Various machine learning models can be used as the first model 22 for learning and the second model 30 for learning, such as CNN (convolutional neural network), RNN (recurrent neural network), LSTM (long short-term memory), and SVM (support vector machine).

[0023] The first learning unit 24 receives the learning data set as input, sets the batch size, which is a unit of learning data used in machine learning, to a predetermined size, uses learning data having the predetermined batch size in the learning data set to perform machine learning to learn parameters of the first model 22 for learning, and builds the learned first model 26. Specifically, in the machine learning by the first learning unit 24, it is sufficient to perform learning by setting, for example, the batch size to 64 to 512, the number of learning times to 500, the number of epochs (the number of times of iterating a learning unit in which the number of learning times is set to 1) to 50, etc. Further, the parameters of the first model 22 for learning are learned from input data, which is obtained from the learning data by shifting the window having a width of 1500 ms in increments of 100 ms, and the correct label so that the likelihood of the correct label is highest. Further, the first learning unit 24 gives the correct label to the likelihood of each label at each time for each piece of the learning data that is obtained in the learning process, and stores it as time-series likelihood data 28. The correct label is given, for example, by giving the correct label corresponding to the time to each of the likelihoods of the labels of “flat”, “stationary”, “stairs ↑”, and “stairs ↓”. For example, in the case of the likelihood of each label for times of 1-71 shown in FIG. 1, “stairs ↑” is given as the correct label.

[0024] The second learning unit 32 receives the time-series likelihood data 28 as input, sets the batch size to a size larger than the predetermined size in the first learning unit 24, and performs machine learning to learn parameters of the second model 30 for learning for outputting one of the labels based on an estimation result of the label at each time. Then, the second learning unit 32 builds the learned second model 34. Specifically, 1024, 2048, 4096, etc., which is a size larger than the batch size of 64 to 512 used in the first learning unit 24, is used as the batch size. Note that the number of learning times and the number of epochs may be the same as or different from those in the machine learning by the first learning unit 24.

[0025] Next, the estimation device 40 will be described. Road surface data detected in time series by the sensors mounted on a moving body moving on a road surface is input to the estimation device 40. It is assumed that the state of the moving body at each time is detected in time series using the road surface data. The estimation device 40 estimates a label using the learned first model 26 for determining, for each of the plurality of types of labels, the likelihood of the label, and the learned second model 34 for outputting one of the labels based on the estimation result of the label at each time. The estimated label is “flat”, “stationary”, “stairs ↑”, “stairs ↓”, etc. used as the correct labels for the learning data in the learning device 20 described above.

[0026] The first estimation unit 42 inputs the road surface data, which is time-series data, to the learned first model 26, estimates the likelihood, for each label, of the label at each time, and outputs it to the second estimation unit 44. Specifically, the learned first model 26 is used to estimate the likelihood of each label for each piece of the input data obtained from the road surface data by shifting a window having a width of 1500 ms in increments of 100 ms, thereby determining the likelihood, for each label, of the label at each time.

[0027] The second estimation unit 44 inputs the likelihood, for each label, of the label at each time, which is estimated in the first estimation unit 42, to the learned second model 34, and estimates one of the labels corresponding to the likelihood of each label at each time.

Operation according to Embodiment of the Present Invention

[0028] Next, the operation of the estimation system 1 according to the embodiment of the present invention will be described.

[0029] First, the operation of the learning device 20 will be described with reference to the flowchart in FIG. 3.

[0030] In step S100, the learning device 20 receives input of the learning data set composed of the learning data, which is time-series data. The learning data is the road surface data detected by the sensors in time series.

[0031] In step S102, the first learning unit 24 receives the learning data set as input, sets the batch size, which is a unit of learning data used in machine learning, to a predetermined size, uses learning data having the predetermined batch size in the learning data set to perform machine learning to learn parameters of the first model 22 for learning, and builds the learned first model 26.

[0032] In step S104, the first learning unit 24 gives the correct label to the likelihood of each label at each time for each piece of the learning data that is obtained in the learning process, and stores it as the time-series likelihood data 28.

[0033] In step S106, the second learning unit 32 receives the time-series likelihood data 28 as input, sets the batch size to a size larger than the predetermined size in the first learning unit 24, and performs machine learning to learn parameters of the second model 30 for learning for outputting one of the labels based on an estimation result of the label at each time. Then, the second learning unit 32 builds the learned second model 34.

[0034] Next, the operation of the estimation device 40 will be described with reference to the flowchart in FIG. 4.

[0035] In step S200, the estimation device 40 receives input of the road surface data detected by the sensors in time series.

[0036] In step S202, the first estimation unit 42 inputs the road surface data, which is time-series data, to the learned first model 26, estimates the likelihood, for each label, of the label at each time, and outputs it to the second estimation unit 44.

[0037] In step S204, the second estimation unit 44 inputs the likelihood, for each label, of the label at each time, which is estimated in the first estimation unit 42, to the learned second model 34, and estimates one of the labels corresponding to the likelihood of each label at each time.

[0038] In step S206, the estimation device 40 outputs the label estimation result obtained in step S204.

[0039] As described above, the estimation system 1 according to the embodiment of the present invention uses the learning device 20 to set the batch size, which is a unit of learning data used in machine learning, to the predetermined size, and learn the first model for determining, for each label, the likelihood of the label. Further, the estimation system 1 according to the embodiment of the present invention outputs the time-series likelihood data, which is the likelihood of each label at each time, for each piece of the learning data. Further, the time-series likelihood data for each piece of the learning data given the correct label is input, and machine learning is performed to learn the second model for outputting one of the labels based on a change in the likelihood of each label. This makes it possible to learn a model for accurately estimating a label indicating the condition of data.

[0040] Further, the estimation device 40 is used to input the time-series data to the learned first model for determining, for each of the plurality of types of labels, the likelihood of the label, and estimate the likelihood, for each label, of the label at each time. In addition, the estimated likelihood, for each label, of the label at each time is input to the learned second model for outputting one of the labels based on the estimation result of the label at each time, the batch size is set to a size larger than the predetermined size, and one of the labels corresponding to the likelihood of each label at each time is estimated. This makes it possible to accurately estimate a label indicating the condition of data.

[0041] Further, the learning device 20 and the estimation device 40 can be implemented using a computer. Such a computer can be implemented by storing a program, in which the processing content for implementing each function of the learning device 20 and the estimation device 40 is written, in a storage unit of the computer in advance, and causing a CPU of the computer to read and execute this program.

[0042] Note that the present invention is not limited to the above-described embodiment, but various modifications and applications can be made without departing from the spirit of the present invention.

[0043] For example, although the case where the road surface data is used as input has been described as an example, it is not limited thereto, and the embodiment of the present invention can be applied to any data detected at each time of a time series.

REFERENCE SIGNS LIST

[0044] 1 Estimation system
 [0045] 20 Learning device

- [0046] 22 First model for learning
- [0047] 24 First learning unit
- [0048] 26 Learned first model
- [0049] 28 Time-series likelihood data
- [0050] 30 Second model for learning
- [0051] 32 Second learning unit
- [0052] 34 Learned second model
- [0053] 40 Estimation device
- [0054] 42 First estimation unit
- [0055] 44 Second estimation unit

1. A learning device comprising:
 a first learner configured to:
 receive, as input, a learning data set including learning data that is time-series data and is given one of a plurality of types of labels as a correct label at each time,
 set a batch size, which is a unit of learning data used in machine learning, to a predetermined size,
 use learning data having the batch size in the learning data set to perform predetermined machine learning to learn a first model for estimating a label, and output an estimation result of the label at each time for each piece of learning data; and
 a second learner configured to:
 receive, as input, the estimation result of the label at each time for each piece of the learning data given a correct label,
 set the batch size to a size larger than the predetermined size, and
 perform predetermined machine learning to learn a second model for outputting one of the labels based on the estimation result of the label at each time.

2. The learning device according to claim 1, wherein the learning data is detected data that is detected in time series by a sensor that detects a state of a target, and the label is a type of condition of a road surface on which a target moves.

3. A learning method comprising:
 receiving, by a first learner, as input, a learning data set including learning data that is time-series data and is given one of a plurality of types of labels as a correct label at each time, setting a batch size, which is a unit of learning data used in machine learning, to a predetermined size, using learning data having the batch size in the learning data set to perform predetermined machine learning to learn a first model for estimating a label, and outputting an estimation result of the label at each time for each piece of learning data; and
 receiving, by a second learner, as input, the estimation result of the label at each time for each piece of the learning data given a correct label, setting the batch size to a size larger than the predetermined size, and performing predetermined machine learning to learn a second model for outputting one of the labels based on the estimation result of the label at each time.

4. A computer-readable non-transitory recording medium storing a computer-executable program instructions that when executed by a processor cause a computer system to:
 receive, by a first learner, as input, a learning data set including learning data that is time-series data and is given one of a plurality of types of labels as a correct label at each time, setting a batch size, which is a unit of learning data used in machine learning, to a predetermined size, using learning data having the batch size in the learning data set to perform predetermined

machine learning to learn a first model for estimating a label, and outputting an estimation result of the label at each time for each piece of learning data; and
 receive, by a second learner, as input, the estimation result of the label at each time for each piece of the learning data given a correct label, setting the batch size to a size larger than the predetermined size, and performing predetermined machine learning to learn a second model for outputting one of the labels based on the estimation result of the label at each time.

5. The learning device according to claim 1, wherein the label includes at least one of: flat, stationary, ascending stairs, or descending stairs.

6. The learning device according to claim 1, wherein the batch size used for learning of the second model is larger than the batch size used for learning the first model.

7. The learning device according to claim 1, wherein the first model is distinct from the second model.

8. The learning device according to claim 1, wherein the learning of the first model and the second model improves accuracy of estimating a label that is low appearance frequency.

9. The learning device according to claim 2, wherein the sensor includes at least one of:
 a gyro sensor,
 a geomagnetic sensor,
 a gravity sensor,
 an atmospheric pressure sensor, or
 a tilt sensor.

10. The learning method according to claim 3, wherein the learning data is detected data that is detected in time series by a sensor that detects a state of a target, and the label is a type of condition of a road surface on which a target moves.

11. The learning method according to claim 3, wherein the label includes at least one of: flat, stationary, ascending stairs, or descending stairs.

12. The learning method according to claim 3, wherein the batch size used for learning of the second model is larger than the batch size used for learning the first model.

13. The learning method according to claim 3, wherein the first model is distinct from the second model.

14. The learning method according to claim 3, wherein the learning of the first model and the second model improves accuracy of estimating a label that is low appearance frequency.

15. The learning method according to claim 10, wherein the sensor includes at least one of:
 a gyro sensor,
 a geomagnetic sensor,
 a gravity sensor,
 an atmospheric pressure sensor, or
 a tilt sensor.

16. The computer-readable non-transitory recording medium according to claim 4, wherein the learning data is detected data that is detected in time series by a sensor that detects a state of a target, and the label is a type of condition of a road surface on which a target moves.

17. The computer-readable non-transitory recording medium according to claim 4, wherein the label includes at least one of: flat, stationary, ascending stairs, or descending stairs.

18. The computer-readable non-transitory recording medium according to claim 4, wherein the batch size used

for learning of the second model is larger than the batch size used for learning the first model.

19. The computer-readable non-transitory recording medium according to claim **4**, wherein the learning of the first model and the second model improves accuracy of estimating a label that is low appearance frequency.

20. The computer-readable non-transitory recording medium according to claim **16**, wherein the sensor includes at least one of:

- a gyro sensor,
- a geomagnetic sensor,
- a gravity sensor,
- an atmospheric pressure sensor, or
- a tilt sensor.

* * * * *