SPECIFICATION

TITLE OF INVENTION

PROGRAM AND DATA PROCESSING METHOD

TECHNICAL FIELD

[0001] This invention relates to techniques for assisting analysis of trends shown by measured values.

BACKGROUND ART

[0002] As an example of a technique for assisting analysis of trends shown by measured values, Patent Document 1 discloses a technique of calculating an estimated value indicative of a propulsive performance of a ship in a given weather environment based on data recorded while the ship is sailing at sea, and correcting the estimated value based on measured values indicative of a sea weather environment.

PRIOR ART DOCUMENT

PATENT DOCUMENT

[0003] Patent Document 1: Japanese Unexamined Patent Application No. 2018-34585

SUMMARY OF THE INVENTION

PROBLEM TO BE SOLVED BY THE INVENTION

[0004] To evaluate a performance of a vehicle, machine, etc., it is common practice to analyze a trend shown by measured values based on a relationship between different items. For example, fuel efficiency is analyzed based on a relationship between a sailing speed and fuel consumption. The measured values used in such a case may include multiple groups of measured values that show different relationships

between different items due to different conditions existing at the time of measurement, such as a group of values measured during strong wind and a group of values measured during light wind, etc. In such a case, it is desirable to filter the measured values such that filtered measured values show a relationship between different items with high reproducibility. Accordingly, the purpose of the present invention is to assist in filtering measured values such that filtered measured values show a relationship between different items with high reproducibility.

MEANS FOR SOLVING THE PROBLEM

[0005] To solve the problem described above, the present invention includes, as a first aspect, a program causing a computer to execute: a step for acquiring pre-extracted datasets, each of the pre-extracted datasets containing values of one or more items of a first group and values of one or more items of a second group, the values contained in the pre-extracted dataset being measured in a same time period, and a step for extracting post-extracted datasets from among the pre-extracted datasets, each of the post-extracted datasets containing values of one or more items of the second group satisfying a predetermined extraction condition.

[0006] The present invention includes, as a second aspect, a program according to the first aspect, causing the computer to execute: a step for judging whether a number of the post-extracted datasets and a statistical variation in values of one or more items of the first group contained in the post-extracted datasets satisfy a predetermined judgment condition.

[0007] The present invention includes, as a third aspect, a program according to the second aspect, wherein the judgment condition is satisfied when a reduction ratio is greater than a predetermined threshold, the reduction ratio being a ratio of a first reduction to a second reduction, the first reduction being a reduction in statistical variations in values of the one or

more items of the first group in the extraction executed at the step for extracting the post-extracted datasets, the second reduction being a reduction in numbers of datasets in the extraction executed at the step for extracting the post-extracted datasets.

[0008] The present invention includes, as a fourth aspect, a program according to the second or third aspect, wherein when the post-extracted datasets are first post-extracted datasets, the program causes the computer to execute: a step for extracting second post-extracted datasets from among the first post-extracted datasets, each of the second post-extracted datasets containing values of one or more items of the second group, each of the values being within a range predetermined for a corresponding item, and a step for judging whether a number of the second post-extracted datasets and statistical variation in values of the one or more items of the first group contained in the second post-extracted datasets satisfy a predetermined judgment condition.

[0009] The present invention includes, as a fifth aspect, a program according to the fourth aspect, wherein the one or more items of the second group used at the step for extracting the first post-extracted datasets and the one or more items of the second group used at the step for extracting the second post-extracted datasets are different.

[0010] The present invention includes, as a sixth aspect, a program according to the fourth or fifth aspect, causing the computer to execute: a step for repeatedly executing the step for extracting and the step for judging, using the second post-extracted datasets extracted at a preceding step for extracting as the first post-extracted datasets in a subsequent step for extracting, until at least one of the following, satisfies a predetermined end condition: a statistical variation in values of one or more items of the first group contained in the post-extracted datasets extracted at the last step for extracting, a number of the post-extracted datasets extracted at

the last step for extracting, a reduction in statistical variations in values of the one or more items of the first group at the last step for extracting, and a reduction in numbers of datasets at the last step for extracting.

[0011] The present invention includes, as a seventh aspect, a program according to any one of the second to sixth aspect, causing the computer to execute: a step for calculating a relational expression among the items of the first group based on values contained in the post-extracted datasets, and a step for determining the statistical variation used at the step for judging based on the values contained in the post-extracted datasets and the relational expression.

[0012] The present invention includes, as an eighth aspect, a program according to any one of the first to seventh aspect, causing the computer to execute: a step for calculating a relational expression among the items of the second group based on values contained in the pre-extracted datasets, and a step for determining a statistical variation in values of the items of the second group based on the values contained in the pre-extracted datasets and the relational expression, wherein at the step for extracting, datasets containing values of the items of the second group whose difference from the relational expression satisfies the extraction condition are extracted as the post-extracted datasets.

[0013] The present invention includes, as an ninth aspect, a program according to any one of the first to eighth aspect, wherein at the step for extracting, the extraction condition is adjusted such that a reduction in the numbers of datasets in the extraction satisfies a predetermined condition.

[0014] The present invention includes, as an tenth aspect, a program according to any one of the first to ninth aspect, wherein at the step for extracting, the extraction condition is adjusted such that a reduction in statistical variations in values of the one or more items of the first group in the extraction satisfies a predetermined condition.

[0015] The present invention includes, as an eleventh aspect, a program according to the ninth or tenth aspect, causing the computer to execute: a step for repeatedly executing the step for extracting and the step for judging, using the post-extracted datasets extracted at a preceding step for extracting as the pre-extracted datasets at a subsequent step for extracting, wherein at the step for extracting, the extraction condition is changed based on a number of steps for extracting already executed.

[0016] The present invention includes, as an twelfth aspect, a program according to any one of the first to eleventh aspect, wherein at the step for acquiring, plural groups of the pre-extracted datasets are acquired, at the step for executing, the extraction is executed for each of the plural groups of the pre-extracted datasets, and the program causes the computer to execute: a step for calculating an indicator, for each of plural groups of the post-extracted datasets extracted from any of the plural groups of the pre-extracted datasets, based on a number of the post-extracted datasets and a statistical variation in values of the one or more items of the first group of the post-extracted datasets, and a step for judging whether the indicators calculated for the plural groups of the post-extracted datasets satisfy a predetermined condition.

[0017] The present invention includes, as a thirteenth aspect, a method executed by a data processing device comprising: a step for acquiring pre-extracted datasets, each of the pre-extracted datasets containing values of one or more items of a first group and values of one or more items of a second group, values contained in the pre-extracted dataset being measured in a same time period, and a step for extracting post-extracted datasets from among the pre-extracted datasets, each of the post-extracted datasets containing values of one or more items of the second group satisfying a predetermined extraction condition.

EFFECTS OF THE INVENTION

[0018] According to the present invention, a user is assisted in filtering measured values such that filtered measured values show a relationship between different items with high reproducibility.

BRIEF EXPLANATION OF THE DRAWINGS

[0019] [FIG 1] Figure 1 shows an example of a hardware configuration of a performance analyzing device according to an exemplary embodiment of the present invention.

[FIG 2] Figure 2 shows a functional configuration realized by the performance analyzing device.

[FIG 3] Figure 3 shows examples of stored measured values.

[FIG 4] Figure 4 shows an example of an item selection screen.

[FIG 5] Figure 5 shows examples of stored datasets.

[FIG 6] Figure 6 is a graph showing an example of a calculated relational expression.

[FIG 7] Figure 7 shows an example of an operating flow in an extraction process executed by the performance analyzing device.

[FIG 8] Figure 8 is an example of graph showing a transition of a statistical variation and a number of datasets.

[FIG 9] Figure 9 is a graph showing an example of datasets used in a performance analysis.

[FIG 10] Figure 10 is a graph showing an example of datasets used in a performance analysis.

[FIG 11] Figure 11 is a graph showing an example of datasets used in a performance analysis.

[FIG 12] Figure 12 is an example of graph showing a transition of a statistical variation and a number of datasets due to filtering of datasets.

[FIG 13] Figure 13 is another example of graph showing a transition of a statistical variation and a number of datasets due to filtering datasets.

MODES FOR CARRYING OUT THE INVENTION

[0020] [1] EXEMPLARY EMBODIMENT

Figure 1 shows an example of a hardware configuration of Performance Analyzing Device 10 according to an exemplary embodiment of the present invention. Performance Analyzing Device 10 is a device that executes various types of data processing to analyze a performance of a vehicle, machinery, etc. In this exemplary embodiment, Performance Analyzing Device 10 executes various types of data processing to analyze a performance of a ship. Performance Analyzing Device 10 is a computer comprising Processor 11, Memory 12, Storage 13, Communication Device 14, Data Input Device 15, Data Output Device 16 and Data Bus 17.

[0021] Processor 11 controls the entire computer by executing data processing in accordance with programs such as an operating system. Processor 11 comprises a central processing unit (CPU) equipped with an interface for peripheral devices, a control unit, an arithmetic unit, a register, etc. Processor 11 reads out programs from at least one of Storage 13 and Communication Device 14 to Memory 12, and executes various types of data processing in accordance with the programs.

[0022] Memory 12 is a computer-readable recording medium, such as a read only memory (ROM) and a random access memory (RAM). Storage 13 is a computer-readable storage medium such as a hard disk drive. Communication Device 14 is a device that performs data communications with other computers via at least one of wired and wireless networks.

[0023] Data Input Device 15 is a device that receives data input from external devices such as a switch, a button, a sensor, etc. The external devices include an imaging device such as a digital camera that takes pictures. Data Output Device 16 is a device that outputs data to external devices

such as a display, a speaker, an LED lamp, etc.

[0024] Data Input Device 15 and Data Output Device 16 may be configured as a single device such as a touch screen equipped with a stacked display and touch panel. Each device, such as Processor 11 and Memory 12, is connected with Data Bus 17 and communicates with other units of Performance Analyzing Device 10 via Data Bus 17. Data Bus 17 may be constituted of a single bus or of plural busses with one for each unit.

[0025] Functions of Performance Analyzing Device 10 are realized when Processor 11 executes data processing according to a program read from Storage 13 to Memory 12, and controls operations of other units of Performance Analyzing Device 10 such as data communications of Communication Device 14.

[0026] Figure 2 illustrates a functional configuration of Performance Analyzing Device 10. Performance Analyzing Device 10 comprises, as functional components, Measured Value Storing Unit 101, Item Selection Receiving Unit 102, Dataset Extracting Unit 103, Condition Sufficiency Judging Unit 104, Relational Expression Calculating Unit 105 and Processing End Judging Unit 106. Measured Value Storing Unit 101 stores measured values of plural items related to an object of performance analysis. In this exemplary embodiment, Measured Value Storing Unit 101 stores values measured during one voyage of the ship, a performance of which is analyzed. In this specification, one voyage means a voyage performed by the ship from departure from a port to arrival at a next port.

[0027] In detail, Measured Value Storing Unit 101 stores measured values of plural items including, for example, a sailing speed (log speed and/or over-the-ground speed), fuel consumption, engine speed (RPM), ship displacement, trim, change in sailing speed, wind speed, wind direction, wave height, current speed, current direction, pitch, roll, rudder angle, etc. Values are measured by a sensor on the ship, by a crew of the ship or by

an external system such as a weather forecasting system, and the measured values are inputted to Performance Analyzing Device 10. Input of the measured values to Performance Analyzing Device 10 may be performed automatically via Communication Device 14 or manually via Data Input Device 15.

[0028] Measured Value Storing Unit 101 stores measured values of each item for plural measurement periods. Figure 3 illustrates examples of measured values stored in Measured Value Storing Unit 101. As shown in Figure 3, Measured Value Storing Unit 101 stores, for example, measured values of each item for measurement periods t1, t2, t3, t4, t5, etc., in correspondence with a name of ship where the values were measured. In the examples of measured values, the sign at the end of each measured value such as "-t1" indicates a measurement period in which the value was measured.

[0029] A time interval of measured values stored in Measured Value Storing Unit 101 is, for example, a shortest time interval (hereinafter, "minimum measurement interval") among measurement time intervals for each item. For example, if the minimum measurement interval is once every minute, Measured Value Storing Unit 101 stores measured values of each item every minute. Measured Value Storing Unit 101 stores values complemented by interpolation or by extrapolation from available measured values before and after a measurement period for which no measured value is available.

[0030] Measured Value Storing Unit 101 may, for example, interpolate an average of available measured values before and after measurement periods as a value for the measurement period for which no measured value is available. Measured Value Storing Unit 101 may also calculate an approximate formula indicating a time-series change of measured values, and interpolate or extrapolate a value that the approximate

formula indicates for each measurement period as a value for the measurement period for which no measured value is available. With regard to some items, such as ship displacement, a value is measured only once a day because a value of the item rarely fluctuates during navigation. For those items, Measured Value Storing Unit 101 may store the value measured on the day for all of measurement periods in the day.

[0031] For example, if a minimum measurement interval is every 10 minutes, and measurement intervals of two items are the same every 10 minutes but measurement timings of the items are not the same, Measured Value Storing Unit 101 stores the measured values of the items measured at different timings as values measured in the same measurement period. Namely, values of different items of the same measurement period may be measured at different timings in the measurement period. Alternatively, Measured Value Storing Unit 101 may calculate interpolated values or extrapolated values of different items for the same timings in each measurement period, and store the interpolated values or the extrapolated values of different items for the measurement period.

[0032] Item Selection Receiving Unit 102 receives selections of items for analysis and selections of items for filtering from among plural items of measured values related to performance analysis. Items for analysis are two or more items relations of which are determined and used for the performance analysis. Items for filtering are two or more items that are used for filtering datasets containing measured values of items for analysis that adequately represent the aforementioned relations. In this invention, items for analysis are examples of "items of a first group," and items for filtering are examples of "items of a second group."

[0033] Item Selection Receiving Unit 102 displays an "item selection screen" that receives operations made by a user for selecting items for analysis and items for filtering. Figure 4 shows an example of item selection

screen. Item selection screen A1 shown in Figure 4 contains Input Field C1 for receiving a name of ship to be analyzed, Input Field C2 for receiving one or more names of items to be included in the items for analysis, Input Field C3 for receiving one or more names of items to be included in the items for filtering, Input Field C4 for receiving a period in which values to be used for performance analysis are measured, and Button B1 for receiving an instruction to start the performance analysis.

[0034] Input of data to each input field in the item selection screen is made by a user of Performance Analyzing Device 10, such as a user who is in charge of performance analysis of ships. The input of data to each input field may be made by typing text, checking selected items from a pull-down list showing candidate items, dragging selected items from a list of candidate items, etc. Item Selection Receiving Unit 102 receives data in each input field in the item selection screen as data indicating a ship selected by the user, items selected by the user as items for analysis, items selected by the user as items for filtering, or a period determined by the user as a period for selecting datasets to be used for the performance analysis, when Button B1 is pressed by the user to start performance analysis.

[0035] When Item Selection Receiving Unit 102 receives the data input by the user to the item selection screen, it sends the received data to Dataset Extracting Unit 103. Dataset Extracting Unit 103 extracts, from among pre-extracted datasets such as datasets stored in Measured Value Storing Unit 101, post-extracted datasets containing values of the items for filtering within a specified range for each item. Each dataset stored in Measured Value Storing Unit 101 contains values of plural items measured in the same measurement period.

[0036] Dataset Extracting Unit 103 reads from Measured Value Storing Unit 101 datasets that are associated with the name of ship indicated by the

data sent from Item Selection Receiving Unit 102, contain values of the items for analysis and the items for filtering indicated by the data sent from Item Selection Receiving Unit 102, and contain values measured in a measurement period within the period indicated by the data sent from Item Selection Receiving Unit 102. Dataset Extracting Unit 103 causes Memory 12 to temporarily store the datasets read from Measured Value Storing Unit 101 as pre-extracted datasets.

[0037] Figure 5 shows examples of pre-extracted datasets stored in Memory 12. Figure 5 shows, for example, values of the items for analysis, i.e. log speed and fuel consumption, and values of the items for filtering, i.e. wind speed and wind direction, measured in measurement period t1, i.e. LOG-t1, FOC-t1, Wind-t1 and WD-t1, are stored as a dataset named DS-t1 in Memory 12.

[0038] Similarly, datasets named DS-t2, DS-t3, DS-t4, DS-t5, etc., corresponding to measurement periods t2, t3, t4, t5, etc., are stored in Memory 12. The area in Memory 12 where pre-extracted datasets as shown in Figure 5 are temporality stored is hereafter referred to as a "pre-extracted dataset-storing area." The pre-extracted datasets that are read out from Measured Value Storing Unit 101 by Dataset Extracting Unit 103 and temporarily stored in the pre-extracted dataset-storing area are hereafter referred to as "initial pre-extracted datasets."

[0039] Dataset Extracting Unit 103 extracts post-extracted datasets containing values within the specified range for each of the items for filtering from the initial pre-extracted datasets stored in the pre-extracted dataset-storing area, and causes Memory 12 to store the post-extracted datasets in an area of Memory 12 other than the pre-extracted dataset-storing area named as a "post-extracted dataset-storing area." In this embodiment, Dataset Extracting Unit 103 executes the extraction of datasets by adjusting the specified ranges for each of the items for filtering such that

a number of datasets excluded in the extraction is approximately a predetermined amount.

[0040] In the following description, the predetermined amount of datasets excluded in the extraction is defined as 90% of the number of pre-extracted datasets, and wind speed is used as one of the items for filtering. Dataset Extracting Unit 103 determines a range less than a predetermined percentage, for example 90% of the maximum wind speed of the pre-extracted datasets as a provisional range of wind speed for filtering datasets. When a range that is less than 90% of the maximum wind speed of the pre-extracted datasets is determined as the provisional range for filtering, Dataset Extracting Unit 103 excludes datasets that contain measured values of wind speeds greater than 90% of the maximum wind speed of the pre-extracted datasets, and tentatively extracts datasets that contain measured values of wind speeds less than 90% of the maximum wind speed.

[0041] When a number of extracted datasets does not reach a predetermined amount, i.e. 10% of the number of pre-extracted datasets, Dataset Extracting Unit 103 tentatively extracts datasets that contain measured values of wind speeds that are less than 80% of the maximum wind speed of the pre-extracted datasets. Dataset Extracting Unit 103 repeatedly limits the range of wind speed for extracting datasets and tentatively extracts datasets containing values of wind speeds within the limited range until a number of datasets excluded in the tentative extraction reaches the predetermined amount, i.e. 10% of the number of pre-extracted datasets. When a number of datasets excluded in the tentative extraction reaches the predetermined amount, Dataset Extracting Unit 103 determines the datasets extracted in the final tentative extraction as post-extracted datasets for use in the subsequent processing, and stops the repeated tentative extractions.

[0042] A method of limiting a range of a value of an item for filtering is not limited to the above-mentioned method. For example, Dataset Extracting Unit 103 may repeatedly limit the range of wind speed by any percentage of increase, such as 1%, rather than 10%. Further, Dataset Extracting Unit 103 may re-extract datasets using a range (for example, a range less than 85% of the maximum value) that is greater than a range used in a previous extraction (for example, a range less than 80% of the maximum value) when a number of datasets excluded in the previous extraction is much larger than the predetermined amount.

[0043] In connection with the predetermined amount of excluded datasets in the extraction, Dataset Extracting Unit 103 may use a fixed number instead of a fixed percentage of a number of pre-extracted datasets. After Dataset Extracting Unit 103 completes the extraction of datasets for wind speed as described above, it performs extractions of datasets for each of items for filtering other than wind speed in the same way as for wind speed. Dataset Extracting Unit 103 causes Memory 12 to temporarily store each group of post-extracted datasets extracted for each of the items for filtering in different post-extracted dataset-storing areas prepared for each of the items for filtering.

[0044] Condition Sufficiency Judging Unit 104 judges whether a number of post-extracted datasets and a statistical variation in measured values of the items for analysis contained in the post-extracted datasets satisfy a predetermined extraction condition. The extraction condition is a condition used for filtering datasets. A statistical variation is an indicator of variation in measured values in the present embodiment. The smaller the statistical variation in the measured values is, the clearer the trend shown by the measured values is. Accordingly, the smaller the statistical variation in the measured values is, the better the measured values are for performance analysis. Details of statistical variation in measured values

will be explained later.

[0045] In this embodiment, a reduction ratio is defined as a ratio of a reduction in statistical variations in measured values to a reduction in numbers of datasets in the extraction. Condition Sufficiency Judging Unit 104 judges that the extraction condition is satisfied when a reduction ratio is larger than a predetermined threshold. Condition Sufficiency Judging Unit 104 calculates a reduction in numbers of datasets by subtracting the number of post-extracted datasets stored in the post-extracted dataset-storing area from the number of pre-extracted datasets stored in the pre-extracted dataset-storing area, for each of the items for filtering.

[0046] Condition Sufficiency Judging Unit 104 instructs Relational Expression Calculating Unit 105 to calculate a relational expression among the items for analysis based on measured values of the pre-extracted datasets or the post-extracted datasets. The relational expression is used for calculating a reduction in statistical variations in values of each of the items for analysis. When Relational Expression Calculating Unit 105 receives the instruction, it calculates a relational expression among the items for analysis by use of the pre-extracted datasets stored in the pre-extracted datasets. Relational Expression Calculating Unit 105 also calculates a relational expression among the items for analysis by use of the post-extracted datasets stored in the post-extracted dataset-storing area for each of the items for filtering.

[0047] Relational Expression Calculating Unit 105 calculates the relational expression according to, for example, a curve fitting method. The curve fitting method is a method for determining a curve expressing a relation among parameters using a regression analysis such as a nonlinear least squares method, a nonlinear least absolute value method, etc. The type of regression analysis may be any one of a normal regression analysis, a principal component regression analysis, a geometric mean regression

analysis, etc. The type of curve expressing a relation among the items for analysis may be any one of a curve following the square-cubic low expressed as y=ax^3, a more generic curve expressed as y=ax^b, etc.

[0048] The curve calculated by Relational Expression Calculating Unit 105 according to the curve fitting method may be a curve expressed by a polynomial, a curve with an intercept, etc. A type of curve calculated by Relational Expression Calculating Unit 105 may be selected by the user from among plural types of curve. In this case, the item selection screen shown in Figure 4 may contain an input field for receiving a type of curve selected by the user. Relational Expression Calculating Unit 105 may calculate a relational expression that represents a relationship between, for example, values of log speed and values of fuel consumption using the datasets illustrated in Figure 5.

[0049] Figure 6 shows an example of a curve representing a relationship between values of log speed and values of fuel consumption. In Figure 6, Coordinate Space G1 is a coordinate space with a horizontal axis indicating log speed and a vertical axis indicating fuel consumption. Each of the plots in Coordinate Space G1 indicates any of the combinations of values of log speed and fuel consumption contained in the pre-extracted datasets shown in Figure 5. Curve D1 in Coordinate Space G1 indicates a relational expression calculated by Relational Expression Calculating Unit 105 based on the values contained in the pre-extracted datasets. Relational Expression Calculating Unit 105 calculates a relational expression for the pre-extracted datasets. Relational Expression Calculating Unit 105 also calculates a relational expression for the post-extracted datasets for each of the items for filtering. Relational Expression Calculating Unit 105 sends the calculated relational expressions to Condition Sufficiency Judging Unit 104.

[0050] Condition Sufficiency Judging Unit 104 calculates a reduction in

statistical variations in values of the items for analysis in the extraction using the notified relational expressions for each of the items for filtering. More specifically, Condition Sufficiency Judging Unit 104 calculates, for each of the relational expressions handed over from Relational Expression Calculating Unit 105, a statistical variation, such as a mean, a standard deviation, etc., of differences in the direction of a Y-axis (or differences in the direction of an X-axis, shortest distances, etc.) between the values of items for analysis and the curve indicating the relationship among the items for analysis. Namely, Condition Sufficiency Judging Unit 104 calculates a statistical variation in values of the items for analysis contained in the pre-extracted datasets, which is referred to hereafter as a pre-extraction statistical variation. Condition Sufficiency Judging Unit 104 also calculates, for each of the items for filtering, a statistical variation in values of the items for analysis contained in the post-extracted datasets that are extracted based on values of the item for filtering, which is referred to hereafter as a post-extraction statistical variation. Then, Condition Sufficiency Judging Unit 104 calculates, for each of the items for filtering, a reduction in statistical variations by subtracting the post-extraction statistical variation from the pre-extraction statistical variation.

[0051] Then, Condition Sufficiency Judging Unit 104 calculates, for each of the items for filtering, a reduction ratio by dividing the reduction in statistical variations by the reduction in numbers of datasets. Condition Sufficiency Judging Unit 104 judges, for each of the items for filtering, whether the extraction condition is satisfied based on the reduction ratio.

[0052] In this embodiment, Condition Sufficiency Judging Unit 104 selects the item for filtering whose reduction ratio is the largest among the reduction ratios of the items for filtering, and judges that only the extraction condition for the selected item for filtering is satisfied.

[0053] As described above, in this embodiment, Condition Sufficiency Judging Unit 104 judges whether the extraction condition is satisfied based on the statistical variation represented by differences between the values of the items for analysis and the relational expression calculated by Relational Expression Calculating Unit 105. Condition Sufficiency Judging Unit 104 sends the post-extracted datasets corresponding to the selected item for filtering to Dataset Extracting Unit 103.

[0054] Dataset Extracting Unit 103 repeats the extraction process described above using the post-extracted datasets sent from Condition Sufficiency Judging Unit 104 as new pre-extracted datasets.

[0055] For example, Dataset Extracting Unit 103 instructs Memory 12 to delete the initial pre-extracted datasets from the pre-extracted dataset-storing area and store the new pre-extracted datasets in the pre-extracted dataset-storing area. Dataset Extracting Unit 103 also instructs Memory 12 to delete the post-extracted datasets from each of the post-extracted dataset-storing areas corresponding to any of the items for filtering. Alternatively, Dataset Extracting Unit 103 may instruct Memory 12 to prepare a new pre-extracted dataset-storing area and new post-extracted dataset-storing areas, and store the new pre-extracted datasets in the new pre-extracted dataset-storing area. When Dataset Extracting Unit 103 completes the extraction process using the new pre-extracted datasets, Condition Sufficiency Judging Unit 104 judges which extraction condition is satisfied based on the reduction ratio calculated for new post-extracted datasets extracted from the new pre-extracted datasets by use of the extraction condition for each of the items for filtering.

[0056] As the extraction process and the judgment process are repeated as described above, a number of the last selected post-extracted datasets, i.e. a number of the latest pre-extracted datasets, is gradually reduced. Each time a set of the extraction process and the judgment process is executed,

Condition Sufficiency Judging Unit 104 notifies a number of the last selected post-extracted datasets to Processing End Judging Unit 106.

[0057] Processing End Judging Unit 106 judges whether the extraction process and the judgment process, which are executed repeatedly, should be terminated. In this embodiment, Processing End Judging Unit 106 judges that a predetermined end condition is satisfied when the number of the last selected post-extracted datasets notified from Condition Sufficiency Judging Unit 104 is less than a predetermined threshold. When Processing End Judging Unit 106 judges that the end condition is satisfied, it instructs Condition Sufficiency Judging Unit 104 to terminate the judgment process. Following the instruction from Processing End Judging Unit 106, Condition Sufficiency Judging Unit 104 does not send the last selected post-extracted datasets to Dataset Extracting Unit 103. As described above, Dataset Extracting Unit 103 and Condition Sufficiency Judging Unit 104 repeatedly execute the extraction process and the judgment processes until the end condition is satisfied.

[0058] As described above, Performance Analyzing Device 10 executes the extraction process to extract post-extracted datasets from pre-extracted datasets. Figure 7 illustrates an example of an operation flow in the extraction process executed by Performance Analyzing Device 10. The extraction process is initiated when a user in charge of performance analysis instructs Performance Analyzing Device 10 to display the item selection screen shown in Figure 4. First, at Step S11, Performance Analyzing Device 10 (Item Selection Receiving Unit 102) receives data input to the input fields in the item selection screen, such as items for analysis input to Input Field C2 and items for filtering input to Input Field C3.

[0059] Then, at Step S12, Performance Analyzing Device 10 (Dataset Extracting Unit 103) selects one of the items for filtering as an item of

focus for filtering, and extracts post-extracted datasets from the initial pre-extracted datasets by use of the extraction condition corresponding to the item of focus for filtering. Then, at Step S13, Performance Analyzing Device 10 judges whether there are one or more items for filtering that have not yet been selected in already executed Step S12.

[0060] When Performance Analyzing Device 10 judges that there are one or more items for filtering that have not yet been selected in already executed Step S12, it repeats the processes at Step S12 and subsequent steps. When Performance Analyzing Device 10 judges that there are no items for filtering that have not yet been selected in already executed Step S12, the process at Step S14 is executed. At Step S14, Performance Analyzing Device 10 (Condition Sufficiency Judging Unit 104) judges which extraction condition is satisfied among the extraction conditions each for any of the items for filtering, based on a number of post-extracted datasets and a statistical variation of values of the items for analysis in the post-extracted datasets.

[0061] Then, at Step S15, The Performance Analyzing Device 10 (Processing End Judging Unit 106) judges whether the end condition is satisfied. When Performance Analyzing Device 10 judges that the end condition is not satisfied, it repeats the processes at Step S12 and subsequent steps. When Performance Analyzing Device 10 judges that the end condition is satisfied, it terminates the series of processes following the flow shown in Figure 7. The series of processes executed in the flow shown in Figure 7 is referred to hereafter as a set of repeated extractions.

[0062] Figure 8 is an example of a graph showing a transition of a statistical variation and a number of datasets while the set of repeated extractions is executed. The coordinate space shown in Figure 8 has a vertical axis indicating a statistical variation and a horizontal axis indicating a number of datasets. Black Circle E0 in the graph indicates a combination of the

statistical variation of values of the items for analysis and the number of datasets in connection with the initial pre-extracted datasets. White Inverted Triangle E1 in the graph indicates a combination of the statistical variation and the number of datasets in connection with the first post-extracted datasets of the first extraction process in the set of repeated extractions corresponding to the first item for filtering. Similarly, White Square E2, White Triangle E3 and Black Diamond E4 respectively indicate combinations of the statistical variation and the number of datasets regarding the first post-extracted datasets corresponding to the second item for filtering, the third item for filtering and the fourth item for filtering.

[0063] Hereinafter, the first item for filtering shall be trim, the second item for filtering shall be wave height, the third item for filtering shall be current speed, and the fourth item for filtering shall be wind speed. In this embodiment, the extraction condition is determined such that a reduction in numbers of datasets in each extraction process is approximately a predetermined number N. The graph shows that a size relation of reductions in statistical variations corresponding to the items of filtering in the first extraction process is trim, wave height, current speed and wind speed in ascending order. Namely, when the extraction condition regarding wind speed is used, the maximum reduction in statistical variations is achieved. Therefore, the first post-extracted datasets extracted by use of the extraction condition regarding wind speed is selected as a new pre-extracted datasets, i.e. the second pre-extracted datasets used for the second extraction process.

[0064] As shown in the graph, a size relation of reductions in statistical variations corresponding to the items of filtering in the second extraction process is wave height, trim, wind speed and current speed in ascending order. Therefore, the second post-extracted datasets extracted by use of

the extraction condition regarding current speed is selected as the third pre-extracted datasets used for the third extraction process. The graph also shows that a size relation of reductions in statistical variations corresponding to the items of filtering in the third extraction process is current speed, wind speed, trim and wave height in ascending order. Therefore, the third post-extracted datasets extracted by use of the extraction condition regarding wave height is selected.

[0065] Threshold Th1 shown on the horizontal axis in Figure 8 indicates the threshold used in the end condition. In the example shown in the graph of Figure 8, after the third extraction process, a number of each of the third post-extracted datasets corresponding to the items for filtering is less than Threshold Th1. Therefore, the set of repeated extractions is terminated. Dataset Extracting Unit 103 outputs to a predetermined destination the last selected datasets, i.e. the third post-extracted datasets extracted by use of the extraction condition regarding wave height.

[0066] The predetermined destination may be, for example, Storage 13, an external device connected to Data Output Device 16 such as a display, an external storage device, etc. In this embodiment, the datasets output to the predetermined destination are used for performance analysis of the ship. Figures 9 and 10 are each graphs showing examples of datasets used in the performance analysis of the ship.

[0067] The graphs of Figures 9 and 10 show log speeds of the ship measured each month. A faster log speed indicates a higher performance in terms of fuel efficiency.

[0068] Datasets used to draw Graph F1 of Figure 9 are the initial pre-extracted datasets fuel consumptions of which are within a predetermined range. Datasets used to draw Graph F2 of Figure 10 are the selected last post-extracted datasets in the set of repeated extractions.

[0069] In this example, maintenance was performed on the ship in the fourth

month, and the log speed of the ship increased after the maintenance. This indicates that maintenance reduced propulsion resistance during sailing of the ship, and fuel efficiency improved. It is not readily apparent Graph F1 shown in Figure 9 that fuel consumption had decreased at the timing when the maintenance was performed because the statistical variation in log speeds is large.

[0070] On the other hand, it is readily apparent from Graph F2 in Figure 10 that fuel consumption had decreased at the timing when the maintenance was performed because statistical variation in log speeds is small. Thus, as the statistical variation of measured values decreases, trends in measured values become more apparent. According to this embodiment, datasets are filtered based on measured values of the items for filtering contained in the datasets such that statistical variation of measured values of the items for analysis contained in the filtered datasets decrease. As a result, datasets are obtained that show a trend in values of the items for analysis more clearly than the original datasets.

[0071] In general, when extraction of datasets is properly performed a statistical variation in measured values decreases as a number of values is reduced. However, a problem arises when a number of datasets after extraction becomes too small. Graph F4 shown in Figure 11 comparative to Graph F2 shown in Figure 10 is illustrative of the problem. Datasets used to draw Graph F4 of Figure 11 are datasets extracted from the initial pre-extracted datasets used to draw Graph F1 using an extraction condition that filters out a large number of datasets in one extraction process. The number of plots in Graph F4 is smaller than the number of plots in Graph F2.

[0072] Graph F4 shows an unreliable trend such that the longer the elapsed time after the maintenances is, the higher the fuel efficiency is. This problem may be caused by the existence of abnormal values in the initial

pre-extraction data, and by the existence of correlations between the item for filtering used in the extraction, and the log speed.

[0073] When a number of datasets after the extraction process is excessively small, reliability of values contained in remaining datasets cannot be ensured, since the values may be subject to the influence of abnormal values contained in the initial pre-extracted datasets. According to Performance Analyzing Device 10, datasets are extracted in such a way that a number of datasets filtered out in the set of repeated extractions is minimized to reduce a statistical variation in values of the items for analysis, by selecting the item for filtering used for each of the extraction processes in the set of repeated extractions from among the plural items for filtering based on reduction ratios corresponding to the items for filtering. As a result, a sufficient number of datasets remain after the set of repeated extractions performed to obtain the target statistical variation.

[0074] For example, a relationship between log speed and fuel consumption should be maintained constant under constant conditions. However, values of log speed and fuel consumption that are measured during sailing of the ship include values that do not accurately reflect a relationship between log speed and fuel consumption due to various factors each of which change in a variety of ways during sailing. Datasets for performance analysis of a ship should include as few datasets as possible that show a relationship among values of items for analysis under circumstances that rarely occur, and as many datasets as possible that show a relationship among values of items for analysis under circumstances that occur frequently. Circumstances that occur frequently are referred to hereinafter as "usual circumstances," and the circumstances that rarely occur are referred to hereinafter as "unusual circumstances."

[0075] A relationship between values of log speed and values of fuel

consumption measured under usual circumstances is more reproducible than other relationships. When datasets showing a relationship between values of log speed and values of fuel consumption measured under unusual circumstances are excluded, a reduction in statistical variations is larger than when datasets showing a relationship between values of log speed and values of fuel consumption measured under usual circumstances are excluded, and a statistical variation in values of log speed and values of fuel consumption is reduced.

[0076] Performance Analyzing Device 10 uses a reduction ratio to specify the extraction condition that achieves a most efficient reduction of datasets, and uses the specified extraction condition for the extraction of datasets. As a result, datasets containing values measured under unusual circumstances are preferentially excluded, compared to a case where the extraction condition used for the extraction is not determined based on the reduction ratio. Thus, datasets extracted by Performance Analyzing Device 10 contain values of the items for analysis that show a relationship that is more reproducible than that for other relationships.

[0077] Items for filtering, which are used for filtering out datasets containing values measured under unusual circumstances, are not limited to items that have an obvious influence on a relationship between log speed and fuel consumption, such as wind speed, current speed, etc. Even if changes in circumstances affecting the relationship between log speed and fuel consumption and changes in measured values are coincidentally linked to each other, if the reduction ratio satisfies the extraction conditions, datasets extracted by Performance Analyzing Device 10 include measured values showing a reproducible relationship between log speed and fuel consumption.

[0078] In this embodiment, a number of datasets is not reduced by one reduction process, but rather a number of datasets is gradually reduced

by plural reduction processes using different extraction conditions as shown in the flow in Figure 8. By reducing a small number of datasets in each extraction and repeating reductions using different extraction conditions, datasets containing values measured under different unusual circumstances can be removed with high probability.

[0079] In this embodiment, Performance Analyzing Device 10 judges whether an extraction condition is satisfied and whether an end condition is satisfied. Therefore, reliable datasets can be obtained consistently without being affected by an ability or experience of a user in charge of performance analysis of the ship. In this embodiment, a number of datasets to be reduced in a single extraction is limited, as a result of which a number of datasets is prevented from becoming too small.

[0080] In this embodiment, datasets are extracted such that statistical variation, which is represented by differences between a relational expression among the items for analysis and measured values of the items for analysis, is reduced. Namely, the relational expression among the items for analysis is used as the reference for calculating the statistical variation. As a result, values of the items for analysis contained in the datasets after extraction more accurately show a relationship among the items for analysis.

[0081] [2] Modifications

The above-described embodiment is an exemplary embodiment of the present invention, and may be modified in various ways. Following are examples of modifications of the above-described embodiment. Two or more of the above-described embodiment and the following modifications may be combined.

[0082] [2-1] End condition

In the above-described embodiment, Processing End Judging Unit 106 judges that the set of repeated extractions should be terminated when a

number of datasets is less than a predetermined threshold. Processing End Judging Unit 106 may judge that the set of repeated extractions should be terminated when a different condition is satisfied. For example, Processing End Judging Unit 106 may judge that the set of repeated extractions should be terminated when a statistical variation in measured values of the items for analysis contained in the last post-extracted datasets satisfies a predetermined condition. For example, when a statistical variation in measured values of the items for analysis contained in the last post-extracted datasets is less than a predetermined threshold, Processing End Judging Unit 106 may judge that the set of repeated extractions should be terminated.

[0083] Processing End Judging Unit 106 may judge that the set of repeated extractions should be terminated when a reduction in statistical variations in values of the items for analysis in the last extraction satisfies a predetermined condition. For example, when a reduction in statistical variations in values of the items for analysis in the last extraction is more than a predetermined threshold, Processing End Judging Unit 106 may judge that the set of repeated extractions should be terminated. Processing End Judging Unit 106 may judge that the set of repeated extractions should be terminated when a reduction in numbers of datasets in the last extraction satisfies a predetermined condition. For example, when a reduction in numbers of datasets in the last extraction is more than a predetermined threshold, Processing End Judging Unit 106 may judge that the set of repeated extractions should be terminated.

[0084] When a number of datasets or a reduction in numbers of datasets is used in the end condition, the set of repeated extractions can be terminated before a number of remaining datasets becomes too small. In such cases, termination of the set of repeated extractions may occur when a statistical variation is not small enough. On the other hand, when a

statistical variation or a reduction in statistical variations is used in the end condition, a statistical variation can be terminated after a statistical variation becomes sufficiently small.

[0085] Processing End Judging Unit 106 may judge whether the set of repeated extractions should be terminated based on two or more of a statistical variation in values of the items for analysis contained in the post-extracted datasets extracted in the last extraction, a number of the post-extracted datasets extracted in the last extraction, a reduction in statistical variations in values of the items for analysis in the last extraction, and a reduction in numbers of datasets in the last extraction. For example, when the statistical variation is more than a threshold and the number of datasets is more than another threshold, Processing End Judging Unit 106 may judge that the set of repeated extractions should be terminated.

[0086] Processing End Judging Unit 106 may calculate an indicator using two or more of the above four parameters, and judge whether the set of repeated extractions should be terminated based on the indicator. For example, when an indicator obtained by adding a reduction in statistical variations multiplied by a predetermined factor and a number of datasets multiplied by another predetermined factor is more than a predetermined threshold, Processing End Judging Unit 106 may judge that the set of repeated extractions should be terminated.

[0087] Depending on how the thresholds are set, it is possible, for example, that a number of datasets may be reduced to an extent that extraction of datasets may result in a situation where the reduction ratio or statistical variation increases. To avoid such a situation, Processing End Judging Unit 106 may judge that the set of repeated extractions should be terminated when a reduction ratio or a statistical variation increases in the last extraction process for more than a predetermined number of items for

filtering, in addition to the end condition being satisfied. In this case, Dataset Extracting Unit 103 may select the post-extracted datasets of the last extraction process with the highest reduction ratio as the final datasets to be output, or Dataset Extracting Unit 103 may select the pre-extracted datasets of the last extraction process as the final datasets to be output.

[0088] According to this modification, reliable datasets can be stably extracted without dependence on knowledge or experience of a user in charge of performance analysis. Further, according to this modification, it is possible to extract datasets with a good balance between prevention of an excessive reduction of numbers of datasets and improvement of reduction in statistical variations, by combining two or more parameters.

[0089] [2-2] Reduction of numbers of datasets in each extraction process

In the above-described embodiment, extraction conditions are adjusted such that a reduction in numbers of datasets in each extraction process matches or approximates a predetermined number. Alternatively, extraction conditions may be adjusted such that a reduction in statistical variations in each extraction process matches or approximates a predetermined value. In this modification, Dataset Extracting Unit 103 adjusts a range of values of the item for filtering in the extraction condition such that a reduction in statistical variations in each extraction process matches or approximates a predetermined value.

[0090] In this modification, Dataset Extracting Unit 103 tentatively extracts, for each of the items for filtering, post-extracted datasets from pre-extracted datasets. Relational Expression Calculating Unit 105 calculates, for each of the group of post-extracted datasets, a relational expression among the items for analysis based on values of the items for analysis contained in the post-extracted datasets. Dataset Extracting Unit 103 calculates, for each of the group of post-extracted datasets, a reduction in statistical variations in the tentative reduction process by use of the

relational expression in the same way as in the above-described embodiment. When the reduction in statistical variations in the tentative reduction process does not approximate the predetermined value, Dataset Extracting Unit 103 changes, for each of the items for filtering, the range of values of the item for filtering in the extraction condition such that a reduction in statistical variations approaches the predetermined value. More specifically, when the reduction in statistical variations in the tentative reduction process is substantially smaller than the predetermined value, Dataset Extracting Unit 103 widens the range in the extraction condition, and when the reduction in statistical variations in the tentative reduction process is substantially larger than the predetermined value, Dataset Extracting Unit 103 narrows down the range in the extraction condition. Dataset Extracting Unit 103 repeats the tentative extraction and the adjustment of extraction condition until the reduction in statistical variations substantially matches the predetermined value for all of the items for filtering.

[0091] Figure 12 is an example of a graph showing a transition of a statistical variation and a number of datasets while the set of repeated extractions is executed according to this modification. In Figure 12, as in the example in Figure 8, Black Circle E0 indicating a combination of the statistical variation and the number of the initial pre-extracted datasets, White Inverted Triangle E1 indicating a combination of the statistical variation and the number of the post-extracted datasets of each extraction process using the extraction condition in connection with trim, White Square E2 indicating a combination of the statistical variation and the number of the post-extracted datasets of each extraction process using the extraction condition in connection with wave height, White Triangle E3 indicating a combination of the statistical variation and the number of the post-extracted datasets of each extraction process using the extraction

condition in connection with current speed, and Black Diamond E4 indicating a combination of the statistical variation and the number of the post-extracted datasets of each extraction process using the extraction condition in connection with wind speed, are shown. In this modification, the extraction condition is determined such that a reduction in statistical variations in values of the items for analysis contained in datasets in each extraction process becomes approximately a predetermined value σ.

[0092] The graph shows that a size relation of reductions in numbers of datasets corresponding to the items of filtering in the first extraction process is trim, wave height, current speed and wind speed in ascending order. In this case, a reduction ratio of datasets extracted using the extraction condition in connection with wind speed becomes the largest. Therefore, the post-extracted datasets of the first extraction process extracted using the extraction condition in connection with wind speed is selected as the second pre-extracted datasets used in the second extraction process. For the second and third extraction processes, similarly, datasets with the smallest reduction in numbers of datasets is selected.

[0093] In this modification, Processing End Judging Unit 106 judges that the set of repeated extractions should be terminated when a statistical variation is less than a predetermined threshold. In the graph of Figure 12, Threshold Th2 used in the end condition is shown on the vertical axis, and a statistical variation becomes less than Threshold Th2 after the third extraction process. Therefore, Performance Analyzing Device 10 terminates the set of repeated extractions after the third extraction process, and the last post-extracted datasets a number of which is the largest is selected as the final datasets to be output.

[0094] In this modification, when the threshold used in the end condition is fixed, a number of extraction processes executed in the set of repeated extractions is also fixed. In the above-described embodiment, when a

fixed threshold such as Threshold Th1 shown in Figure 8 is used in the end condition, a number of extraction processes executed in the set of repeated extractions is also fixed.

[0095] In this modification, a threshold of a number of datasets such as Threshold Th1 shown in Figure 8 may be used in the end condition instead of Threshold Th2. Similarly, in the above-described embodiment, a threshold of statistical variation such as Threshold Th2 shown in Figure 12 may be used in the end condition instead of Threshold Th1. In either case, a number of datasets, a statistical variation in values of the items for analysis, or their reduction can be reduced to the threshold used in the end condition.

[0096] [2-3] Change in reduction in each extraction process

In the above-described embodiment, a reduction in numbers of datasets in each extraction process is substantially fixed. The reduction in numbers of datasets may be changed in each extraction process. In the above modification, a reduction in statistical variations in each extraction process is substantially fixed. The reduction in statistical variations may be changed in each extraction process. For example, Dataset Extracting Unit 103 may change the extraction condition based on a number of extraction processes that have already been executed in the set of repeated extractions.

[0097] Figure 13 is an example of a graph showing a transition of a statistical variation and a number of datasets while the set of repeated extractions is executed in this modification. The information represented by Black Circle E0, White Inverted Triangle E1, White Square E2, White Triangle E3 and Black Diamond E4 in the graph in Figure 13 is the same as those in Figures 8 and 12. In the example of Figure 13, when at least one of a statistical variation for any of the items for filtering reaches Threshold Th3 the set of repeated extractions is terminated.

[0098] In this modification, Dataset Extracting Unit 103 adjusts the extraction condition used for the first extraction process such that a reduction in numbers of datasets becomes approximately N1, adjusts the extraction condition used for the second extraction process such that a reduction in numbers of datasets becomes approximately N2, which is smaller than N1, and adjusts the extraction condition used for the third extraction process such that a reduction in numbers of datasets becomes approximately N3, which is smaller than N2. Namely, until the statistical variation is reduced to a certain degree, a large number of datasets are excluded in one extraction process, and when the statistical variation approaches the threshold, a small number of datasets are excluded in one extraction process.

[0099] Similarly, Dataset Extracting Unit 103 may adjust the extraction condition such that a reduction in statistical variations reaches a threshold that becomes smaller with the number of extraction processes already executed. In this case, the set of repeated extractions may be terminated when a number of the last post-extracted datasets reaches a threshold. According to this modification, without increasing a number of extraction processes in a set of repeated extractions, datasets showing a reliable relationship among the items for analysis are obtained

[0100] [2-4] Change of items for filtering

In the above-described embodiment, the items for filtering are not changed when the set of repeated extractions is executed. However, the items for filtering may be changed when the set of repeated extractions is executed. In this modification, Dataset Extracting Unit 103 may change the items for filtering for each of extraction processes executed in the set of repeated extractions. For example, Dataset Extracting Unit 103 may remove from the current items for filtering an item corresponding to the post-extracted datasets selected as the pre-extracted datasets for a next

extraction process, and add a new item to the items for filtering, each time an extraction process is completed.

[0101] For example, in the example in Figure 8, Dataset Extracting Unit 103 removes wind speed from the items for filtering after the first extraction process is completed, since the post-extracted datasets extracted using the extraction condition in connection with wind speed are selected as the second pre-extracted datasets. Then, Dataset Extracting Unit 103 adds a new item, such as engine revolution speed, to the items for filtering for the second extraction process. It may occur that, once a range of values of an item has been limited, further limiting the range of values of the same item may result in a significant reduction in numbers of datasets.

[0102] According to this modification, an item already used for filtering out unreliable datasets is not repeatedly used in a set of repeated extractions and thus the above problem is avoided. Instead of replacing an item for filtering already used for the extraction process with another item, the item may simply be removed from the items for filtering, or one or more new items may be added to the items for filtering without removing the item already used for the extraction process. In either case, as long as the items for filtering are appropriately changed, reliable datasets can be efficiently extracted.

[0103] [2-5] How to limit measured values of the items for filtering

A method for limiting measured values of the items for filtering is not limited to that described in the above-described embodiment. For example, Dataset Extracting Unit 103 may extract datasets in which measured values of two or more items for filtering satisfy a predetermined condition.

[0104] For example, Dataset Extracting Unit 103 may extract, from the pre-extracted datasets, datasets containing measured values of wind direction indicating a head wind and measured values of wind speed indicating

values that are less than the first threshold, and containing measured values of wind direction indicating a tail wind and measured values of a wind speed indicating values that are less than the second threshold, or containing measured values of a wind direction indicating a side wind and measured values of a wind speed indicating values that are less than the third threshold. For example, Dataset Extracting Unit 103 may calculate, for each of the pre-extracted datasets, an indicator by adding a measured value of pitch and a measured value of roll, and extract, from the pre-extracted datasets, datasets in which indicators are less than a threshold.

[0105] A relational expression among the items for filtering may be used for extracting datasets. For example, Relational Expression Calculating Unit 105 calculates a relational expression among two or more items for filtering based on measured values of the items for filtering contained in the pre-extracted datasets. Then, Dataset Extracting Unit 103 extracts datasets by excluding, from the pre-extracted datasets, datasets containing values of the items for filtering with a large difference from the relational expression such that a statistical variation in differences between values of the items for filtering and the relational expression becomes less than a threshold.

[0106] For example, Dataset Extracting Unit 103 calculates a relational expression expressing a relationship between engine revolution speed and change in sailing speed based on values of these items for filtering contained in the pre-extracted datasets. Then, Dataset Extracting Unit 103 calculates a statistical variation, such as a mean, a standard deviation, etc., of differences between values of the items for filtering and a curve indicating the relational expression (differences in the direction of a Y-axis, differences in the direction of an X-axis, or shortest distances, etc.). When the statistical variation is more than a threshold, Dataset Extracting

Unit 103 removes datasets from the pre-extracted datasets in the order of their difference from the curve indicating the relational expression until a statistical variation in values of the items for filtering contained in the remaining datasets reaches the threshold. The datasets that remain when the statistical variation reaches the threshold are used as post-extracted datasets.

[0107] When there is a correlation among values of two or more items for filtering, the larger the statistical variation in differences between measured values and the relational expression, the more likely it is that a large number of abnormal values due to failure of a measuring device, misjudgment of a measurer, etc. are contained in the pre-extracted datasets. In many cases, the abnormal values are not detected by monitoring measured values of only one of the items for filtering. According to this modification, such abnormal values can be eliminated by using the relational expression as a reference.

[0108] [2-6] Extraction condition

Extraction conditions different from those used in the above-described embodiment may be used. In the above-described embodiment, Condition Sufficiency Judging Unit 104 selects post-extracted datasets with the largest reduction ratio as the datasets satisfying the extraction condition. Alternatively, for example, Condition Sufficiency Judging Unit 104 may select a predetermined number of groups of post-extracted datasets in decreasing order of reduction ratios.

[0109] In such a case, Dataset Extracting Unit 103 executes the next extraction process for each of the selected groups of post-extracted datasets. For example, if two groups of post-extracted datasets are selected by Condition Sufficiency Judging Unit 104 in each extraction process and extraction processes are executed three times for the set of repeated extractions, eight groups of post-extracted datasets are obtained as a

result of the set of repeated extractions. In such a case, all of the post-extracted datasets may be used for performance analysis, or one only may be selected for use for performance analysis.

[0110] For example, Condition Sufficiency Judging Unit 104 may select post-extracted datasets with a smallest statistical variation or a largest number of datasets as the final datasets to be output. In the above-described embodiment, Condition Sufficiency Judging Unit 104 compares reduction ratios calculated for each of the items for filtering to judge whether the extraction condition is satisfied. Instead of using a relative condition, an absolute condition may be used as the extraction condition.

[0111] For example, Condition Sufficiency Judging Unit 104 judges that the extraction condition is satisfied when the reduction ratio is larger than a predetermined threshold. In such a case, when plural groups of datasets are extracted, Condition Sufficiency Judging Unit 104 may select one of them as the final datasets to be output. Condition Sufficiency Judging Unit 104 may use an extraction condition including a combination of an absolute condition and a relative condition.

[0112] In such a case, when there are plural groups of datasets that satisfy the absolute condition, Condition Sufficiency Judging Unit 104 may select, for example, a group of datasets that satisfies the relative condition as the datasets to be used in the next extraction process. Whichever extraction condition is used, reliable datasets showing a correct relationship among the items for analysis are obtained without need to depend on a knowledge and experience of a user.

[0113] [2-7] Number of extraction processes

In the above-described embodiment, the extraction process is executed more than once for the set of repeated extractions. The extraction process may be executed only once for the set of repeated extractions. In such a case, when one group of post-extracted datasets of the first execution

process is selected to be output, the set of repeated extractions is terminated. Even in such a case, since one group of datasets is selected from among plural groups of datasets based on efficiency of extraction expressed by a reduction ratio, a large number of datasets with only a small statistical variation can be output. Thus, according to this modification, datasets that show a reliable relationship among the items for analysis are obtained.

[0114] [2-8] Datasets to be output

In the above-described embodiment, Dataset Extracting Unit 103 outputs datasets that are judged to satisfy the extraction condition. However, Dataset Extracting Unit 103 may output datasets that are judged not to satisfy the extraction condition, in addition to datasets that are judged to satisfy the extraction condition. In such a case, the user may select one group of datasets from among the plural groups of datasets output from Performance Analyzing Device 10 for performance analysis. In such a case, as long as the user appropriately selects a group of datasets, the user can obtain datasets containing measured values that show a reproducible relationship among the items for analysis. Namely, according to this modification, the user can readily obtain datasets that are useful for performance analysis with the assistance of Performance Analyzing Device 10.

[0115] When Condition Sufficiency Judging Unit 104 determines datasets to be output, datasets that are useful for performance analysis are obtained regardless of a knowledge and experience of the user. However, a skilled user who has extensive knowledge and experience may select a better group of datasets for performance analysis than a group of datasets automatically selected by Performance Analyzing Device 10. According to this modification, the knowledge and experience of a skilled user can be utilized to select datasets.

[0116] [2-9] Datasets containing values measured during several different voyages

In the above-described embodiment, datasets are extracted from ones that contain values measured during the same voyage of a ship whose performance is analyzed. Datasets may be extracted from ones that contain values measured during any one of plural different voyages of the ship.

[0117] In this modification, Measured Value Storing Unit 101 stores datasets classified into plural groups corresponding to different voyages of the same ship whose performance is analyzed. These groups are referred to hereinafter as voyage groups. Each of the voyage groups contains plural datasets that contain values measured during a voyage corresponding to the group. Dataset Extracting Unit 103 executes the first extraction process for each of the voyage group. As a result, plural groups of post-extracted datasets each corresponding to any of the items for filtering are extracted for each of plural voyage groups. Condition Sufficiency Judging Unit 104 calculates, for each of the groups of post-extracted datasets, a statistical variation in values of the items for analysis contained in the datasets. Then, Condition Sufficiency Judging Unit 104 calculates an indicator for each of the items for filtering using the statistical variations calculated for the item for filtering.

[0118] For example, Condition Sufficiency Judging Unit 104 calculates the reduction ratio, i.e. a ratio of reduction in statistical variations in values of the items for analysis contained in datasets to reduce a number of datasets in the extraction process for each of the groups of post-extracted datasets. Then, Condition Sufficiency Judging Unit 104 calculates an average of the reduction ratios for each of the items for filtering as an indicator of reduction efficiency.

[0119] Condition Sufficiency Judging Unit 104 selects, for each of the voyage

group, the post-extracted datasets corresponding to the item for filtering an indicator of which is the largest as the pre-extracted datasets for the second extraction process.

[0120] Similar to the above-described embodiment, in this modification, Dataset Extracting Unit 103 and Condition Sufficiency Judging Unit 104 repeatedly execute the extraction process and the selection process until the end condition is satisfied. A method for calculating the indicator of reduction efficiency is not limited to the above method. For example, Condition Sufficiency Judging Unit 104 may calculate a sum of the reduction ratios, a value obtained by multiplying the reduction ratios, etc., as the indicator of reduction efficiency.

[0121] The voyage groups may correspond to voyages of different ships, and each voyage group may contain datasets containing values measured during a voyage of a corresponding ship. In such a case, datasets containing values measured under unusual circumstances are excluded, and datasets containing values showing a reproducible relationship among the items for analysis are obtained.

[0122] The voyage groups may correspond to different time periods, and each voyage group may contain datasets containing values measured during the voyage of the ship in the corresponding time period. In this modification, Condition Sufficiency Judging Unit 104 calculates the indicator of reduction efficiency for each of the different time periods. The indicators may be used to analyze changes over time in parameters that affect a performance of the ship.

[0123] [2-10] Type of statistical variation

In the above-described embodiment, a statistical variation in differences between values of the items for analysis and a relational expression showing a relationship among the items for analysis is used for executing the set of repeated extractions. Types of statistical variation

used for executing the set of repeated extractions are not limited to that in the above-described embodiment. For example, a statistical variation in values of one item for analysis may be used for executing the set of repeated extractions.

[0124] [2-11] Number of items for analysis

In the above-described embodiment, log speed and fuel consumption are used as examples of the items for analysis. A number of items for analysis is not limited to two, and one or more than two items for items for analysis may be used.

[0125] [2-12] How to determine and how to obtain items for analysis and items for filtering

In the above-described embodiment, the items for analysis and the items for filtering are selected by the user from candidate items measured values of which are stored in Measured Value Storing Unit 101. A method to determine the items for analysis and the items for filtering is not limited to that in the above-described embodiment. The items for analysis and the items for filtering may be predetermined.

[0126] In the above-described embodiment, measured values of the items for analysis and measured values of the items for filtering are stored in Measured Value Storing Unit 101. A place where these values are stored is not limited to Measured Value Storing Unit 101. These values may be stored in an external system with which Performance Analyzing Device 10 can communicate, and Dataset Extracting Unit 103 may obtain the values from the external system.

[0127] [2-13] Functional configuration

A functional configuration of a device according to the present invention is not limited to that shown in Figure 2. For example, in the above-described embodiment, Dataset Extracting Unit 103 outputs the final datasets. A unit shown in Figure 2 other than Dataset Extracting Unit

103 or a unit that is not shown in Figure 2 may output the final datasets. For example, the extraction process executed by Dataset Extracting Unit 103 and the selection process executed by Condition Sufficiency Judging Unit 104 in the above-described embodiment may be executed by a single unit.

[0128] [2-14] Hardware configuration

A hardware configuration of a device according to the present invention is not limited to that shown in Figure 1. For example, Performance Analyzing Device 10 may consist of two or more separate devices. In such a case, the devices that constitute Performance Analyzing Device 10 may contain a resource comprising a cloud service.

[0129] [2-15] Categories of Invention

The present invention can be understood as a data processing device exemplified by Performance Analyzing Device 10 in the above-described embodiment, a data processing method exemplified by the method consisting of steps executed by the data processing device, and a program for causing a computer to execute the steps of the data processing method. The program may be provided in a form of a recording medium such as an optical disk on which the program is stored, or may be provided in a downloadable form to a computer via a network such as the Internet and installed in the computer.

DESCRIPTION OF REFERENCE NUMERALS

[0130] 10: Performance Analyzing Device

101: Measured Value Storing Unit

102: Item Selection Receiving Unit 102

103: Dataset Extracting Unit 103

104: Condition Sufficiency Judging Unit 104

105: Relational Expression Calculating Unit 105

106: Processing End Judging Unit 106

CLAIMS

[Claim 1] A program causing a computer to execute:

a step for acquiring pre-extracted datasets, each of the pre-extracted datasets containing values of one or more items of a first group and values of one or more items of a second group, the values contained in the pre-extracted dataset being measured in a same time period,

a step for extracting first post-extracted datasets from among the pre-extracted datasets, each of the first post-extracted datasets containing values of one or more items of the second group satisfying a predetermined extraction condition,

a step for judging whether a number of the first post-extracted datasets and a statistical variation in values of one or more items of the first group contained in the first post-extracted datasets satisfy a predetermined judgment condition,

a step for extracting second post-extracted datasets from among the first post-extracted datasets, each of the second post-extracted datasets containing values of one or more items of the second group, each of the values being within a range predetermined for a corresponding item, and

a step for judging whether a number of the second post-extracted datasets and statistical variation in values of the one or more items of the first group contained in the second post-extracted datasets satisfy a predetermined judgment condition,

wherein

the one or more items of the second group used at the step for extracting the first post-extracted datasets and the one or more items of the second group used at the step for extracting the second post-

extracted datasets are different.

[Claim 2] A program according to Claim 1 causing the computer to execute:

a step for repeatedly executing the step for extracting and the step for judging, using the second post-extracted datasets extracted at a preceding step for extracting as the first post-extracted datasets in a subsequent step for extracting, until at least one of the following, satisfies a predetermined end condition:

a statistical variation in values of one or more items of the first group contained in the post-extracted datasets extracted at the last step for extracting,

a number of the post-extracted datasets extracted at the last step for extracting,

a reduction in statistical variations in values of the one or more items of the first group at the last step for extracting, and

a reduction in numbers of datasets at the last step for extracting.