

DESCRIPTION

PICKING DEVICE AND IMAGE GENERATION PROGRAM

Technical Field

[0001]

The present invention relates to a picking device and an image generation program capable of accurately identifying a position of a picking target object.

Background Art

[0002]

In the related art, in a shipment stage from a factory, a distribution base, or the like, work such as sorting, transshipment, and unloading of goods has been automated by a robot.

[0003]

For example, Patent Literature 1 has previously been proposed as a technique for automatically confirming and measuring a box-shaped picking target object to be gripped. Patent Literature 1 discloses a configuration in which a box-shaped picking target object to be gripped is accurately gripped by combining two pieces of information of a two-dimensional image obtained by performing photographing by a two-dimensional camera and three-dimensional point group information acquired by a three-dimensional camera or the like.

Citation List

Patent Literature

[0004]

Patent Literature 1: JP 2019-136828 A

Summary of Invention

Technical Problem

[0005]

As in Patent Literature 1, it is possible to improve a detection accuracy of a position of the picking target object by combining two pieces of information of the two-dimensional image obtained by performing photographing by the two-dimensional camera and the three-dimensional point group information acquired by the three-dimensional camera or the like, but it is still likely that the position of the picking target object is erroneously detected even if the two pieces of information are combined. For example, when the picking target objects are stacked precisely without a gap, there is a concern that a boundary of the picking target object on a top surface cannot be identified only with the three-dimensional point group information. In addition, for example, in a case of a configuration in which the position of the picking target object is detected by performing image recognition processing by AI on the

two-dimensional image obtained by performing photographing by the two-dimensional camera, if the two-dimensional image includes a blown highlight portion due to reflection of light, a dark portion caused by being shaded by another object, or the like, there is a concern that the position of the picking target object is erroneously detected.

[0006]

The present invention has been made in view of the above problems, and an object of the present invention is to provide a picking device and an image generation program capable of accurately identifying a position of a picking target object.

Solution to Problem

[0007]

A picking device according to the present invention includes: a picking robot including a hand portion that picks a picking target object; a two-dimensional camera that two-dimensionally photographs the picking target object; a three-dimensional camera that acquires three-dimensional point group information indicating distance information to the picking target object; and a control unit that controls a picking operation of the picking robot, in which the control unit includes a two-dimensional photographed image acquisition function of controlling the

two-dimensional camera to acquire at least one or more pieces of two-dimensional photographed image data obtained by photographing the entire plurality of picking target objects that are stacked, a three-dimensional point group information acquisition function of controlling the three-dimensional camera to acquire three-dimensional point group information on the entire plurality of the picking target objects that are stacked, a calibration function of commonizing coordinate information between the two-dimensional photographed image data and the three-dimensional point group information based on the picking target object as a target, a top surface candidate region extraction function of extracting top surface candidate regions of the plurality of picking target objects from the three-dimensional point group information, a region-of-interest setting function of setting a region corresponding to the top surface candidate region in the two-dimensional photographed image data as a region of interest, an AI determination image acquisition function of selecting or synthesizing the two-dimensional photographed image data in which luminance information in the region of interest satisfies a predetermined condition based on at least one or more pieces of the two-dimensional photographed image data to obtain AI determination image data, a center coordinate information acquisition function of acquiring

center coordinate information of each picking target object in the AI determination image data by inputting the AI determination image data to a learned model learned in advance about identifying the picking target object from input image data and outputting the center coordinate information of each picking target object, and a picking control function of controlling a picking operation and a transfer operation of the picking target object by the picking robot based on the center coordinate information and height information acquired from the three-dimensional point group information.

[0008]

Also, the picking device according to the present invention sets, in the AI determination image acquisition function, a condition where an average luminance value of pixels in the region of interest is equal to or more than a predetermined lower limit threshold and/or equal to or less than a predetermined upper limit threshold, as a predetermined condition to be satisfied by the AI determination image data.

[0009]

Also, the picking device according to the present invention acquires, in the AI determination image acquisition function, the AI determination image data by determining whether the predetermined condition is

satisfied for a plurality of pieces of the two-dimensional photographed image data acquired by performing photographing by the two-dimensional camera a plurality of times while changing an exposure condition.

[0010]

Also, the picking device according to the present invention generates, in the AI determination image acquisition function, the AI determination image data in which the luminance information in the region of interest satisfies the predetermined condition by executing an HDR processing based on a plurality of pieces of the two-dimensional photographed image data acquired by performing photographing by the two-dimensional camera a plurality of times while changing an exposure condition.

[0011]

Also, in the picking device according to the present invention, the two-dimensional camera is installed near the hand portion of the picking robot to be movable integrally with the hand portion.

[0012]

Also, in the picking device according to the present invention, the three-dimensional camera is installed near the hand portion of the picking robot to be movable integrally with the hand portion.

[0013]

Also, in the picking device according to the present invention, a stereo camera is adopted instead of the two-dimensional camera and the three-dimensional camera, and acquisition of the two-dimensional photographed image data and acquisition of the three-dimensional point group information are realized by the stereo camera.

[0014]

An image generation program according to the present invention causes a computer to realize processing of generating AI determination image data to be used in a configuration in which processing of determining two-dimensional photographed image data obtained by photographing a picking target object to identify the picking target object and calculate center coordinate information of the picking target object is executed based on a learned model (AI) learned in advance by machine learning, and causes the computer to realize: a two-dimensional photographed image acquisition function of controlling a two-dimensional camera to acquire at least one or more pieces of two-dimensional photographed image data obtained by photographing the entire plurality of picking target objects that are stacked; a three-dimensional point group information acquisition function of controlling a three-dimensional camera to acquire three-dimensional point group information on the entire plurality

of the picking target objects that are stacked; a calibration function of commonizing coordinate information between the two-dimensional photographed image data and the three-dimensional point group information based on the picking target object as a target; a top surface candidate region extraction function of extracting top surface candidate regions of the plurality of the picking target objects from the three-dimensional point group information; a region-of-interest setting function of setting a region corresponding to the top surface candidate region in the two-dimensional photographed image data as a region of interest; and an AI determination image acquisition function of selecting or synthesizing the two-dimensional photographed image data in which luminance information in the region of interest satisfies a predetermined condition based on at least one or more pieces of the two-dimensional photographed image data to obtain AI determination image data.

Advantageous Effects of Invention

[0015]

According to the present invention, AI determination image data to be input to a learned model learned in advance about identifying a picking target object from input image data and outputting center coordinate

information of each picking target object is obtained by selecting or synthesizing two-dimensional photographed image data in which luminance information in a region of interest corresponding to a top surface candidate region extracted from three-dimensional point group information satisfies a predetermined condition, and thus it is possible to acquire the center coordinate information (position information) of the picking target object with high accuracy by using the AI determination image data not including blown highlights or dark portions.

Brief Description of Drawings

[0016]

Fig. 1 is an explanatory diagram for describing an overall configuration of a picking device corresponding to at least one of embodiments of the present invention.

Fig. 2 is a block diagram illustrating an example of a configuration of the picking device corresponding to at least one of the embodiments of the present invention.

Fig. 3 is a flowchart illustrating an example of a flow of picking processing of a pickup target object corresponding to at least one of the embodiments of the present invention.

Fig. 4 is an image diagram illustrating an example of two-dimensional photographed image data acquired by the

picking device corresponding to at least one of the embodiments of the present invention.

Fig. 5 is an image diagram illustrating an example of three-dimensional point group information acquired by the picking device corresponding to at least one of the embodiments of the present invention.

Fig. 6 is an image diagram illustrating an example of a state in which a region of interest is set in the two-dimensional photographed image data in the picking device corresponding to at least one of the embodiments of the present invention.

Fig. 7 is an image diagram illustrating an example of AI determination image data generated by the picking device corresponding to at least one of the embodiments of the present invention.

Fig. 8 is an image diagram illustrating an example of a state in which center coordinate information output by AI is superimposed and displayed on the two-dimensional photographed image data in the picking device corresponding to at least one of the embodiments of the present invention.

Fig. 9 is an image diagram illustrating an example of a photographed image obtained by performing photographing a plurality of times while changing an exposure time in the picking device corresponding to at least one of the

embodiments of the present invention.

Fig. 10 is an explanatory diagram for describing a state of determining a predetermined condition focusing only on the region of interest of the two-dimensional photographed image data in the picking device corresponding to at least one of the embodiments of the present invention.

Fig. 11 is an explanatory diagram for describing an example of generating the AI determination image data by HDR processing in the picking device corresponding to at least one of the embodiments of the present invention.

Description of Embodiments

[0017]

[First Embodiment]

Hereinafter, an example of a picking device according to a first embodiment of the present invention is described with reference to the drawings.

[0018]

Fig. 1 is an explanatory diagram for describing an overall configuration of the picking device corresponding to at least one of embodiments of the present invention. As illustrated in Fig. 1, a picking device 100 includes at least a picking robot 10, a two-dimensional camera 20, a three-dimensional camera 30, and a control unit 40. The

picking device 100 of the present example is a device that identifies an accurate position of a top surface of a pickup target object by the control unit 40 based on two-dimensional photographed image data obtained by performing photographing by the two-dimensional camera 20 and three-dimensional point group information acquired by the three-dimensional camera 30, and controls the picking robot 10 based on the identified information to automatically execute work such as transshipment of the picking target object.

[0019]

Fig. 2 is a block diagram illustrating an example of a configuration of the picking device corresponding to at least one of the embodiments of the present invention. As illustrated in Fig. 2, the picking device 100 at least includes the picking robot 10, the two-dimensional camera 20, the three-dimensional camera 30, and the control unit 40.

[0020]

The picking robot 10 is a robot including a hand portion that picks a picking target object. In the present example, the picking robot 10 including a suction hand that sucks and lifts a central portion of the picking target object is described, but the present invention is not limited thereto. For example, the picking robot may

include a hand that grips the picking target object by holding a side surface of the picking target object. The operation of the picking robot 10 is controlled by the control unit 40 described below.

[0021]

The two-dimensional camera 20 has a configuration for two-dimensionally photographing the picking target object. An installation position and a photographing angle can be set in various ways as long as the picking target object can be photographed. However, in the case of the configuration in which a top surface of the picking target object is sucked and gripped as in the present example, it is preferable to install the two-dimensional camera 20 to photograph the picking target object from the above so that the top surface of the picking target object appears. Specifically, it is considered to install the two-dimensional camera 20 in the vicinity of the hand portion at a distal end of the picking robot 10. The vicinity of the hand portion refers to a position where the two-dimensional camera 20 can move integrally with the hand portion and a position where the two-dimensional camera 20 does not affect a picking work even when being moved in conjunction with the hand portion. Further, the two-dimensional camera 20 is not limited to being installed integrally with the hand portion, and the two-dimensional

camera 20 may be installed in a scaffold by setting up a scaffold near the picking target object separately from the picking robot 10 not to interfere with an operation range of the picking robot 10. The execution of photographing of the two-dimensional camera 20 is controlled by the control unit 40 described below.

[0022]

The three-dimensional camera 30 is configured to acquire three-dimensional point group information indicating distance information to the picking target object. Any means may be used as long as the three-dimensional point group information can be acquired, and for example, a time-of-flight camera (TOF camera) can be considered. In addition, a configuration in which the distance information is extracted by using a stereo camera may be possible, or a configuration in which the three-dimensional point group information is acquired by using light detection and ranging, laser imaging detection and ranging (LiDAR) may be possible. The execution of photographing of the three-dimensional camera 30 is controlled by the control unit 40 described below. An installation position and a photographing angle of the three-dimensional camera 30 can be set in various ways as long as the picking target object can be photographed. However, in the case of the configuration in which the top

surface of the picking target object is sucked and gripped as in the present example, it is preferable to install the camera to photograph the picking target object from the above so that the top surface of the picking target object appears. Specifically, it is considered to install the three-dimensional camera 30 in the vicinity of the hand portion at the distal end of the picking robot 10. The vicinity of the hand portion refers to a position where the three-dimensional camera 30 can move integrally with the hand portion and a position where the three-dimensional camera 30 does not affect the picking work even when being moved in conjunction with the hand portion. Further, the three-dimensional camera 30 is not limited to being installed integrally with the hand portion, and the three-dimensional camera 30 may be installed in a scaffold by setting up a scaffold near the picking target object separately from the picking robot 10 not to interfere with the operation range of the picking robot 10. Note that, in the example of Fig. 1, the two-dimensional camera 20 and the three-dimensional camera 30 are installed adjacent to each other to have substantially the same photographing angle.

[0023]

The control unit 40 is configured to control a photographing operation by the two-dimensional camera 20, a

photographing operation by the three-dimensional camera 30, and a picking operation of the picking robot. The control unit 40 may be configured by any means as long as control described below can be executed and is realized by, for example, a computer, a microcomputer, a programmable logic controller (PLC), or a combination thereof. Among the means, the computer may have a configuration of including a central processing unit (CPU), a memory, and a storage such as a hard disk drive and an SSD, that are normally included in a general computer, including an input device such as a mouse and a keyboard, an output device such as a display and a printer, and a communication device for connecting to a communication network, and connecting the units via a bus. Processing in each unit of an information processing device 10 is realized by reading a program for executing processing in each of the units from a memory and executing the program in a CPU, a GPU, or a field programmable gate array (FPGA) that functions as a control circuit (processing circuit, processing circuitry). In other words, a processor (processing circuit) is configured to be able to execute each processing of each device by executing the program. In addition, a configuration may be considered in which processing up to acquisition of center coordinate information of the picking target object is executed by a computer and operation control of the picking

robot 10 by using the center coordinate information is realized using a PLC. In addition, it is obvious that various pieces of data used for the control unit 40 and various pieces of data obtained as a result of processing by the control unit 40 are appropriately stored in a storage unit (not illustrated).

[0024]

As illustrated in Fig. 2, the control unit 40 includes a two-dimensional photographed image acquisition function 41, a three-dimensional point group information acquisition function 42, a calibration function 43, a top surface candidate region extraction function 44, a region-of-interest setting function 45, an AI determination image acquisition function 46, a center coordinate information acquisition function 47, and a picking control function 48.

[0025]

The two-dimensional photographed image acquisition function 41 has a function of controlling the two-dimensional camera 20 to acquire at least one or more pieces of the two-dimensional photographed image data obtained by photographing the entire plurality of picking target objects that are stacked. Furthermore, the two-dimensional photographed image acquisition function 41 can acquire a plurality of pieces of the two-dimensional photographed image data having different exposure times by

changing the exposure time in photographing when performing photographing a plurality of times.

[0026]

The three-dimensional point group information acquisition function 42 has a function of controlling the three-dimensional camera 30 to acquire the three-dimensional point group information on the entire plurality of picking target objects that are stacked. Here, the three-dimensional point group information is information obtained by calculating a distance to an object existing in a photographing range for each measurement point, and it is possible to identify a surface and a shape of the object by analyzing information of a point group.

[0027]

The calibration function 43 has a function of commonizing coordinate information between the two-dimensional photographed image data and the three-dimensional point group information based on the picking target object as a target. It is possible to mutually convert the coordinate information between the two-dimensional photographed image data and the three-dimensional point group information by executing commonization of the coordinate information, so-called calibration, based on the pickup target object commonly appearing between the two-dimensional photographed image

data and the three-dimensional point group information. As a result, it is also possible to perform processing of redrawing the three-dimensional point group information at the same field angle as the two-dimensional photographed image data. By the processing of commonizing the coordinate information in the calibration function 43, it is possible to acquire XY coordinates of each point on the top surface of the pickup target object appearing in the two-dimensional photographed image data and a Z coordinate (distance information in a vertical direction) that is a distance from a photographing point by the two-dimensional camera 20 to the top surface of the pickup target object. [0028]

The top surface candidate region extraction function 44 has a function of extracting top surface candidate regions of the plurality of picking target objects from the three-dimensional point group information. By the processing of commonizing the coordinate information in the calibration function 43, the distance information to each point of the pickup target object becomes clear. Here, when a search range is gradually expanded in a Z coordinate direction from the photographing point, it is possible to detect a portion where the plurality of point groups exist planarly near a certain Z coordinate. As in the present example, when a picking target object is photographed from above and

the picking is performed by the picking robot 10 from above, it is preferable to extract a planar portion closest in distance from a photographing position of the three-dimensional camera 30 in the Z direction as the top surface candidate region. Note that, when detecting a portion where the point group exists planarly, a portion having a margin in a distance in the Z direction may be searched for. A distance range of the point group in the Z direction in which the point group is estimated to exist in the same plane can be appropriately set, according to a work environment.

[0029]

The region-of-interest setting function 45 has a function of setting a region corresponding to the top surface candidate region in the two-dimensional photographed image data as a region of interest. The top surface candidate region is extracted from the three-dimensional point group information in the top surface candidate region extraction function 44 and the coordinate information is commonized between the two-dimensional photographed image data and the three-dimensional point group information, and thus a portion on the two-dimensional photographed image data corresponding to the top surface candidate region can be set as the region of interest. In the setting of the region of interest, the

range of the region of interest needs to be set so that only pixel information existing inside the region of interest can be extracted.

[0030]

The AI determination image acquisition function 46 has a function of obtaining AI determination image data by selecting or synthesizing two-dimensional photographed image data in which luminance information in the region of interest satisfies a predetermined condition based on at least one or more pieces of the two-dimensional photographed image data. The AI determination image acquisition function 46 focuses not on the entire two-dimensional photographed image data but on the luminance information of the pixel included in the region of interest and selects or synthesizes the two-dimensional photographed image data in which the luminance information satisfies a predetermined condition to obtain the AI determination image data. Here, it is conceivable that the predetermined condition regarding the luminance information is, for example, a predetermined condition where an average luminance value for all pixels in the region of interest is equal to or more than a predetermined lower limit threshold and/or equal to or less than a predetermined upper limit threshold (when both the upper limit and the lower limit are set, the average luminance value is within a

predetermined range) to exclude images in which a region-of-interest portion is excessively dark or excessively bright. Furthermore, for example, it is considered to set a predetermined condition that the number of pixels having luminance values equal to or less than a predetermined value (for example, the luminance value in the case of 8 bits is 10 or less) is not equal to or does not exceed a predetermined number to exclude images in which a dark portion exists in the region-of-interest portion.

Furthermore, for example, it is considered to set a predetermined condition that the number of pixels having luminance values equal to or more than a predetermined value (for example, the luminance value in the case of 8 bits is 245 or more) is not equal to or does not exceed a predetermined number to exclude images in which blown highlights exist in the region-of-interest portion.

[0031]

Furthermore, in the AI determination image acquisition function 46, it is considered to acquire the AI determination image data by determining whether the predetermined condition is satisfied for a plurality of pieces of the two-dimensional photographed image data acquired by performing photographing by the two-dimensional camera 20 a plurality of times while changing an exposure condition. That is, it is considered that a plurality of

times of photographing is performed by the two-dimensional camera 20 in advance to acquire a plurality of pieces of the two-dimensional photographed image data in which the exposure condition is changed, and it is determined whether a predetermined condition is satisfied for a plurality of pieces of the two-dimensional photographed image data, whereby the most suitable two-dimensional photographed image data is selected and set as the AI determination image data.

[0032]

Furthermore, in the AI determination image acquisition function 46, it is considered to generate the AI determination image data in which the luminance information in the region of interest satisfies a predetermined condition by executing an HDR processing based on a plurality of pieces of the two-dimensional photographed image data acquired by performing photographing by the two-dimensional camera 20 a plurality of times while changing the exposure condition. Here, the high-dynamic-range (HDR) processing refers to processing of selecting portions with optimum brightness (without blown highlights or blocked-up shadows) from a plurality of pieces of the two-dimensional photographed image data obtained by performing photographing with various periods of exposure times and synthesizing the selected portions into one image. This is

also referred to as HDR synthesis. Since the HDR processing is a known technique, details thereof are omitted, but by performing the HDR processing, it is possible to acquire an image suitable for AI determination image data without blown highlights or blocked-up shadows. Whether the luminance information in the region of interest satisfies a predetermined condition is determined for the image data obtained by the HDR processing, and when the luminance information satisfies the predetermined condition, the image data obtained by the HDR processing is set as the AI determination image data.

[0033]

The center coordinate information acquisition function 47 has a function of acquiring the center coordinate information of each picking object in the AI determination image data by inputting the AI determination image data to a learned model that has learned in advance about identifying the picking target object from input image data and outputting the center coordinate information of each picking target object. Here, the learned model refers to a model that learned about outputting center coordinate information when image data in which a picking target object appears is input by performing machine learning by using a plurality of teacher data sets that are sets of image data in which the picking target object appears and

correct answer data of center coordinate information of the picking target object in the image data. The learned model may be any model, and for example, a learned model configured with a neural network is considered. Various images can be adopted as the teacher data set used for learning, and for example, learning may be performed by using image data including blown highlights or blocked-up shadows in the learning stage. Furthermore, the present invention is not limited to the case of outputting the center coordinate information, and a learned model that has learned to output the coordinate information of positions of four corners of the top surface of the box-shaped picking target object may be adopted, or a learned model that has learned to output both the coordinate information of the positions of the four corners of the top surface of the picking target object and the center coordinate information may be adopted. That is, it is preferable to learn the learned model so that information required for controlling the control target picking robot 10 to execute the picking work can be acquired.

[0034]

The picking control function 48 has a function of controlling the picking operation and a transfer operation of the picking target object by the picking robot based on the center coordinate information and height information

acquired from the three-dimensional point group information. The center coordinate information and the height information (Z coordinate information) are acquired as information required for the picking operation and the transfer operation of the picking target object. Here, when a plurality of pickup target objects having top surfaces at the same height exist, it is possible to appropriately set which picking target object is to be selected to execute the picking operation and the transfer operation. As an example, setting a rule to select the pickup target object closest to the hand portion is considered. The control after the pickup target object is gripped by the hand portion can be appropriately executed based on a known technique.

[0035]

Fig. 3 is a flowchart illustrating an example of a flow of picking processing of the pickup target object corresponding to at least one of the embodiments of the present invention. As illustrated in Fig. 3, the picking processing is started by acquiring one or more pieces of the two-dimensional photographed image data obtained by photographing the picking target object in the control unit 40 of the picking device 100 (step S101). Next, the control unit 40 acquires the three-dimensional point group information obtained by photographing the picking target

object (step S102). Next, the control unit 40 executes calibration for commonizing coordinate information between the two-dimensional photographed image data and the three-dimensional point group information (step S103). Next, the control unit 40 extracts top surface candidate regions from the three-dimensional point group information (step S104). Next, the control unit 40 sets a position corresponding to the top surface candidate region in the two-dimensional photographed image data as the region of interest (step S105). Next, the control unit 40 acquires the AI determination image data by selecting or generating image data in which the luminance information in the region of interest satisfies the predetermined condition based on one or more pieces of the two-dimensional photographed image data (step S106). Next, the control unit 40 acquires the center coordinate information of the picking target object by inputting the AI determination image data to the learned model (AI) (step S107). Then, the control unit 40 controls the picking operation and the transfer operation of the picking target object by the picking robot based on the acquired center coordinate information and the height information (step S108) and ends the picking processing.

[0036]

Next, the operation of the picking device 100 according to the present example is described by using a specific

image as an example.

[0037]

Fig. 4 is an image diagram illustrating an example of the two-dimensional photographed image data acquired by the picking device corresponding to at least one of the embodiments of the present invention. The image diagram illustrated in Fig. 4 represents two-dimensional photographed image data obtained by photographing a state in which three box-shaped pickup target objects are placed at the lower stage and two box-shaped pickup target objects are placed thereon by the two-dimensional camera 20.

[0038]

Fig. 5 is an image diagram illustrating an example of the three-dimensional point group information acquired by the picking device corresponding to at least one of the embodiments of the present invention. The image diagram illustrated in Fig. 5 represents three-dimensional point group information obtained by photographing the same pickup target object as in Fig. 4 with the three-dimensional camera 30, one image represents three-dimensional point group information obtained by performing photographing from a diagonal direction, one image represents three-dimensional point group information drawn at a field angle of the two-dimensional camera, and the remaining one image represents extracted top surface candidate regions.

[0039]

Fig. 6 is an image diagram illustrating an example of a state in which the region of interest is set in the two-dimensional photographed image data in the picking device corresponding to at least one of the embodiments of the present invention. The image diagram illustrated in Fig. 6 illustrates a state in which the region of interest is set in the two-dimensional photographed image data and is a masking image in which portions other than the region of interest are masked.

[0040]

Fig. 7 is an image diagram illustrating an example of the AI determination image data generated by the picking device corresponding to at least one of the embodiments of the present invention. Since what can be adopted as the AI determination image data is the two-dimensional photographed image data in which the luminance information in the region of interest satisfies the predetermined condition, it is determined whether the predetermined condition is satisfied for one or more pieces of the two-dimensional photographed image data based on the range of the region of interest set as illustrated in Fig. 6, and the two-dimensional photographed image data satisfying the predetermined condition is adopted as the AI determination image data.

[0041]

Fig. 8 is an image diagram illustrating an example of a state in which the center coordinate information output by the AI is superimposed and displayed on the two-dimensional photographed image data in the picking device corresponding to at least one of the embodiments of the present invention. The image diagram illustrated in Fig. 8 illustrates a state in which the center coordinate information output by the AI is superimposed and displayed on the two-dimensional photographed image data. As illustrated in Fig. 8, if the center coordinate information (the coordinate information of the four corners may be included as necessary) of the pickup target object can be acquired with high accuracy, the picking operation can be accurately controlled. To acquire the center coordinate information (the coordinate information of the four corners may be included as necessary) with high accuracy, the image data adopted as the AI determination image data is required to be an image with high determination accuracy without blown highlights or blocked-up shadows.

[0042]

Fig. 9 is an image diagram illustrating an example of a photographed image obtained by performing photographing a plurality of times while changing an exposure time in the picking device corresponding to at least one of the

embodiments of the present invention. As illustrated in Fig. 9, when photographing is performed a plurality of times while changing the exposure time, it can be understood that the state of image data obtained is different according to the exposure time. In the example illustrated in Fig. 9, three patterns are displayed, and it can be said that the image data of an intermediate exposure time is optimal. However, it can be said that which exposure time is optimal when change patterns of the exposure time is further increased differs depending on a photographing environment.

[0043]

Fig. 10 is an explanatory diagram for describing a state of determining the predetermined condition focusing only on the region of interest of the two-dimensional photographed image data in the picking device corresponding to at least one of the embodiments of the present invention. When the region of interest is set for the two-dimensional photographed image data illustrated in Fig. 9, the region of interest is set as in the image diagram illustrated in Fig. 10. For a plurality of pieces of the two-dimensional photographed image data in which the exposure time is changed as illustrated in Fig. 9, it may be determined whether the luminance information in the region of interest set as illustrated in Fig. 10 satisfies

the predetermined condition, to adopt the two-dimensional photographed image data satisfying the predetermined condition as the AI determination image data.

[0044]

Fig. 11 is an explanatory diagram for describing an example of generating the AI determination image data by the HDR processing in the picking device corresponding to at least one of the embodiments of the present invention. In unprocessed two-dimensional photographed image data illustrated in Fig. 11, a dark portion that has shadows not dark as blocked-up shadows and a blown highlight occurrence portion exists. In contrast, image data in which the dark portion and a blown highlight portion are eliminated can be generated by performing the HDR processing (HDR synthesis) by using a plurality of pieces of the two-dimensional photographed image data in which the exposure time is changed. Whether the luminance information in the region of interest satisfies the predetermined condition is determined for the image data obtained by the HDR processing, and when the luminance information satisfies the predetermined condition, the image data obtained by the HDR processing may be adopted as the AI determination image data.

[0045]

As described above, the picking device 100 according to

the present invention includes: a picking robot including a hand portion that picks a picking target object; a two-dimensional camera that two-dimensionally photographs the picking target object; a three-dimensional camera that acquires three-dimensional point group information indicating distance information to the picking target object; and a control unit that controls a picking operation of the picking robot, in which the control unit includes a two-dimensional photographed image acquisition function of controlling the two-dimensional camera to acquire at least one or more pieces of two-dimensional photographed image data obtained by photographing the entire plurality of picking target objects that are stacked, a three-dimensional point group information acquisition function of controlling the three-dimensional camera to acquire three-dimensional point group information on the entire plurality of picking target objects that are stacked, a calibration function of commonizing coordinate information between the two-dimensional photographed image data and the three-dimensional point group information based on the picking target object as a target, a top surface candidate region extraction function of extracting top surface candidate regions of the plurality of picking target objects from the three-dimensional point group information, a region-of-interest setting function of

setting a region corresponding to the top surface candidate region in the two-dimensional photographed image data as a region of interest, an AI determination image acquisition function of selecting or synthesizing the two-dimensional photographed image data in which luminance information in the region of interest satisfies a predetermined condition based on at least one or more pieces of the two-dimensional photographed image data to obtain AI determination image data, a center coordinate information acquisition function of acquiring center coordinate information of each picking target object in the AI determination image data by inputting the AI determination image data to a learned model learned in advance about identifying the picking target object from input image data and outputting the center coordinate information of each picking target object, and a picking control function of controlling a picking operation and a transfer operation of the picking target object by the picking robot based on the center coordinate information and height information acquired from the three-dimensional point group information, so that AI determination image data to be input to the learned model learned in advance about identifying the picking target object from the input image data and outputting the center coordinate information of each picking target object is obtained by selecting or synthesizing the two-dimensional

photographed image data in which the luminance information in the region of interest corresponding to the top surface candidate region extracted from the three-dimensional point group information satisfies the predetermined condition. Therefore, it is possible to acquire the center coordinate information (position information) of the picking target object with high accuracy by using the AI determination image data not including blown highlights or dark portions.

[0046]

Also, in the AI determination image acquisition function, a condition where an average luminance value of pixels in the region of interest is equal to or more than a predetermined lower limit threshold and/or equal to or less than a predetermined upper limit threshold is set as a predetermined condition to be satisfied by the AI determination image data. Therefore, it is possible to adopt the two-dimensional photographed image data excluding the two-dimensional photographed image data in which the average luminance value is excessively low or excessively high as the AI determination image data.

[0047]

Also, in the AI determination image acquisition function, the AI determination image data is acquired by determining whether the predetermined condition is satisfied for a plurality of pieces of the two-dimensional

photographed image data acquired by performing photographing by the two-dimensional camera a plurality of times while changing an exposure condition. Therefore, the two-dimensional photographed image data with an optimum exposure time in a photographing environment can be adopted as the AI determination image data.

[0048]

Also, in the AI determination image acquisition function, the AI determination image data in which the luminance information in the region of interest satisfies the predetermined condition is generated by executing the HDR processing based on a plurality of pieces of the two-dimensional photographed image data acquired by performing photographing by the two-dimensional camera a plurality of times while changing an exposure condition. Therefore, it is possible to generate the AI determination image data with high determination accuracy without blown highlights or blocked-up shadows by the HDR processing.

[0049]

[Second Embodiment]

In the first embodiment, the description has been given assuming that the two-dimensional photographed image data is acquired by using the two-dimensional camera 20 and the three-dimensional point group information is acquired by using the three-dimensional camera 30, but the present

invention is not limited thereto. For example, a stereo camera may be adopted, and both acquisition of the two-dimensional photographed image data and acquisition of the three-dimensional point group information may be realized by the stereo camera.

[0050]

According to the stereo camera, two pieces of the two-dimensional photographed image data are acquired by performing photographing at two different photographing angles and depth information (distance information) is acquired by stereo matching by using the two pieces of the two-dimensional photographed image data to generate the three-dimensional point group information. By using one of the two pieces of the two-dimensional photographed image data used here instead of the two-dimensional photographed image data acquired by the two-dimensional camera 20 in the first embodiment, both the acquisition of the two-dimensional photographed image data and the acquisition of the three-dimensional point group information can be realized by the stereo camera. The stereo camera can also perform the calibration function between the two-dimensional photographed image data and the three-dimensional point group information.

Reference Signs List

[0051]

100 picking device

10 picking robot

20 two-dimensional camera

30 three-dimensional camera

40 control unit

41 two-dimensional photographed image acquisition function

42 three-dimensional point group information acquisition
function

43 calibration function

44 top surface candidate region extraction function

45 region-of-interest setting function

46 AI determination image acquisition function

47 center coordinate information acquisition function

48 picking control function

CLAIMS

1. A picking device comprising:

a picking robot including a hand portion that picks a picking target object;

a two-dimensional camera that two-dimensionally photographs the picking target object;

a three-dimensional camera that acquires three-dimensional point group information indicating distance information to the picking target object; and

a control unit that controls a picking operation of the picking robot, wherein

the control unit includes

a two-dimensional photographed image acquisition function of controlling the two-dimensional camera to acquire at least one or more pieces of two-dimensional photographed image data obtained by photographing an entire plurality of picking target objects that are stacked,

a three-dimensional point group information acquisition function of controlling the three-dimensional camera to acquire three-dimensional point group information on the entire plurality of picking target objects that are stacked,

a calibration function of commonizing coordinate information between the two-dimensional photographed image data and the three-dimensional point group information

based on the picking target object as a target,

a top surface candidate region extraction function of extracting top surface candidate regions of the plurality of picking target objects from the three-dimensional point group information,

a region-of-interest setting function of setting a region corresponding to the top surface candidate region in the two-dimensional photographed image data as a region of interest,

an AI determination image acquisition function of selecting or synthesizing the two-dimensional photographed image data in which luminance information in the region of interest satisfies a predetermined condition based on at least one or more of the two-dimensional photographed image data to obtain AI determination image data,

a center coordinate information acquisition function of acquiring center coordinate information of each picking target object in the AI determination image data by inputting the AI determination image data to a learned model learned in advance about identifying the picking target object from input image data and outputting the center coordinate information of each picking target object, and

a picking control function of controlling a picking operation and a transfer operation of the picking target

object by the picking robot based on the center coordinate information and height information acquired from the three-dimensional point group information.

2. The picking device according to claim 1, wherein

in the AI determination image acquisition function, a condition where an average luminance value of pixels in the region of interest is equal to or more than a predetermined lower limit threshold and/or equal to or less than a predetermined upper limit threshold is set as a predetermined condition to be satisfied by the AI determination image data.

3. The picking device according to claim 1 or 2, wherein

in the AI determination image acquisition function, the AI determination image data is acquired by determining whether the predetermined condition is satisfied for a plurality of pieces of the two-dimensional photographed image data acquired by performing photographing by the two-dimensional camera a plurality of times while changing an exposure condition.

4. The picking device according to claim 1 or 2, wherein

in the AI determination image acquisition function, the AI determination image data in which the luminance information in the region of interest satisfies a predetermined condition is generated by executing an HDR processing based on a plurality of pieces of the two-

dimensional photographed image data acquired by performing photographing by the two-dimensional camera a plurality of times while changing an exposure condition.

5. The picking device according to any one of claims 1 to 4, wherein

the two-dimensional camera is installed near the hand portion of the picking robot to be movable integrally with the hand portion.

6. The picking device according to any one of claims 1 to 4, wherein

the three-dimensional camera is installed near the hand portion of the picking robot to be movable integrally with the hand portion.

7. The picking device according to any one of claims 1 to 6, wherein

a stereo camera is adopted instead of the two-dimensional camera and the three-dimensional camera, and acquisition of the two-dimensional photographed image data and acquisition of the three-dimensional point group information are realized by the stereo camera.

8. An image generation program for causing a computer to realize processing of generating AI determination image data to be used in a configuration in which processing of determining two-dimensional photographed image data obtained by photographing a picking target object to

identify the picking target object and calculate center coordinate information of the picking target object is executed based on a learned model (AI) learned in advance by machine learning, the image generation program causing the computer to realize:

a two-dimensional photographed image acquisition function of controlling a two-dimensional camera to acquire at least one or more pieces of two-dimensional photographed image data obtained by photographing an entire plurality of picking target objects that are stacked;

a three-dimensional point group information acquisition function of controlling a three-dimensional camera to acquire three-dimensional point group information on the entire plurality of picking target objects that are stacked;

a calibration function of commonizing coordinate information between the two-dimensional photographed image data and the three-dimensional point group information based on the picking target object as a target;

a top surface candidate region extraction function of extracting top surface candidate regions of the plurality of the picking target objects from the three-dimensional point group information;

a region-of-interest setting function of setting a region corresponding to the top surface candidate region in

the two-dimensional photographed image data as a region of interest; and

an AI determination image acquisition function of selecting or synthesizing the two-dimensional photographed image data in which luminance information in the region of interest satisfies a predetermined condition based on at least one or more of the two-dimensional photographed image data to obtain AI determination image data.