



US 20200096317A1

(19) **United States**(12) **Patent Application Publication**
Nakazato et al.(10) **Pub. No.: US 2020/0096317 A1**(43) **Pub. Date: Mar. 26, 2020**(54) **THREE-DIMENSIONAL MEASUREMENT
APPARATUS, PROCESSING METHOD, AND
NON-TRANSITORY COMPUTER-READABLE
STORAGE MEDIUM**(71) Applicant: **CANON KABUSHIKI KAISHA,**
Tokyo (JP)(72) Inventors: **Yusuke Nakazato,** Tokyo (JP);
Kazuhiko Kobayashi, Yokohama-shi
(JP)(21) Appl. No.: **16/681,115**(22) Filed: **Nov. 12, 2019****Related U.S. Application Data**(60) Division of application No. 15/800,605, filed on Nov.
1, 2017, which is a continuation of application No.
13/515,106, filed on Jun. 11, 2012, now Pat. No.
9,841,271, filed as application No. PCT/JP2011/
054206 on Feb. 18, 2011.**Foreign Application Priority Data**

Feb. 24, 2010 (JP) 2010-039201

Publication Classification(51) **Int. Cl.**
G01B 11/03 (2006.01)
G06T 7/73 (2006.01)
G01B 11/26 (2006.01)
(52) **U.S. Cl.**
CPC **G01B 11/03** (2013.01); **G06T 2207/10028**
(2013.01); **G01B 11/26** (2013.01); **G06T 7/75**
(2017.01)(57) **ABSTRACT**

A three-dimensional measurement apparatus selects points corresponding to geometric features of a three-dimensional shape model of a target object, projects a plurality of selected points corresponding to the geometric features onto a range image based on approximate values indicating the position and orientation of the target object and imaging parameters at the time of imaging of the range image, searches regions of predetermined ranges respectively from the plurality of projected points for geometric features on the range image which correspond to the geometric features of the three-dimensional shape model, and associates these geometric features with each other. The apparatus then calculates the position and orientation of the target object using differences of distances on a three-dimensional space between the geometric features of the three-dimensional shape model and those on the range image, which are associated with each other.

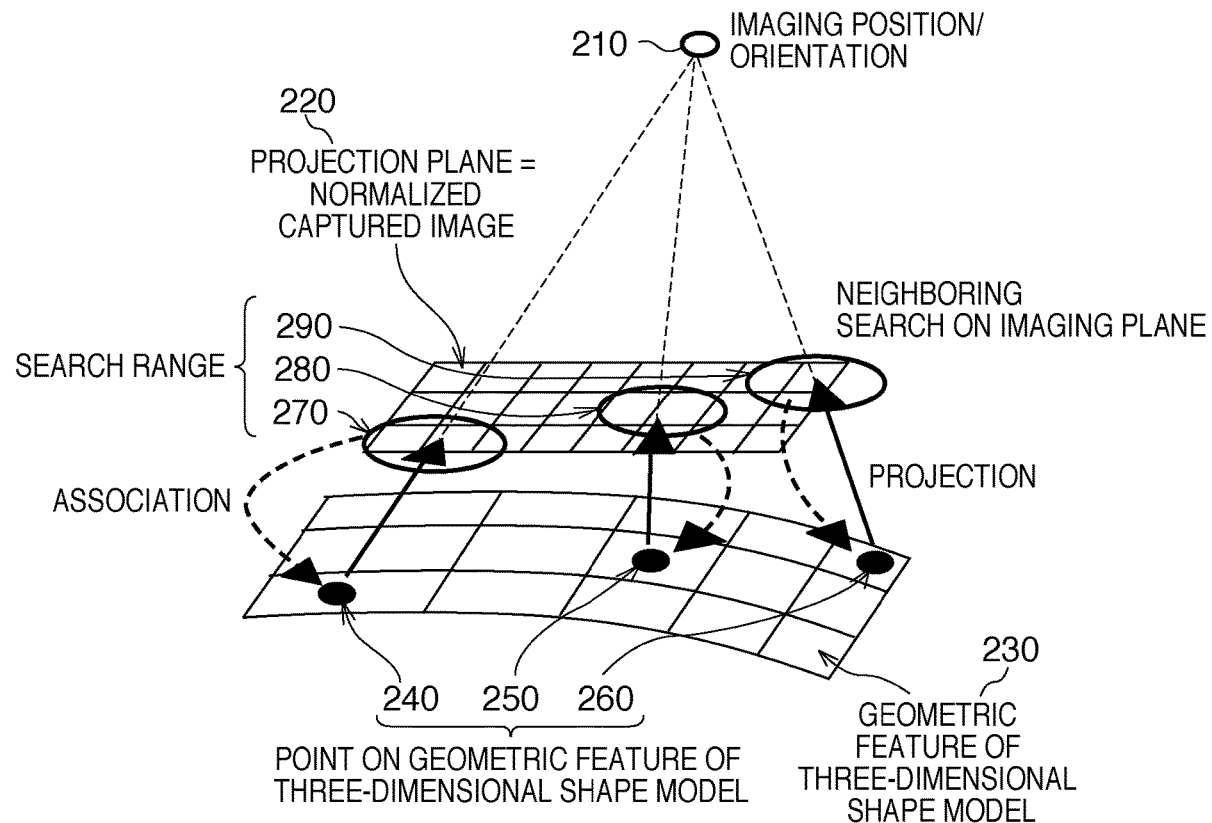


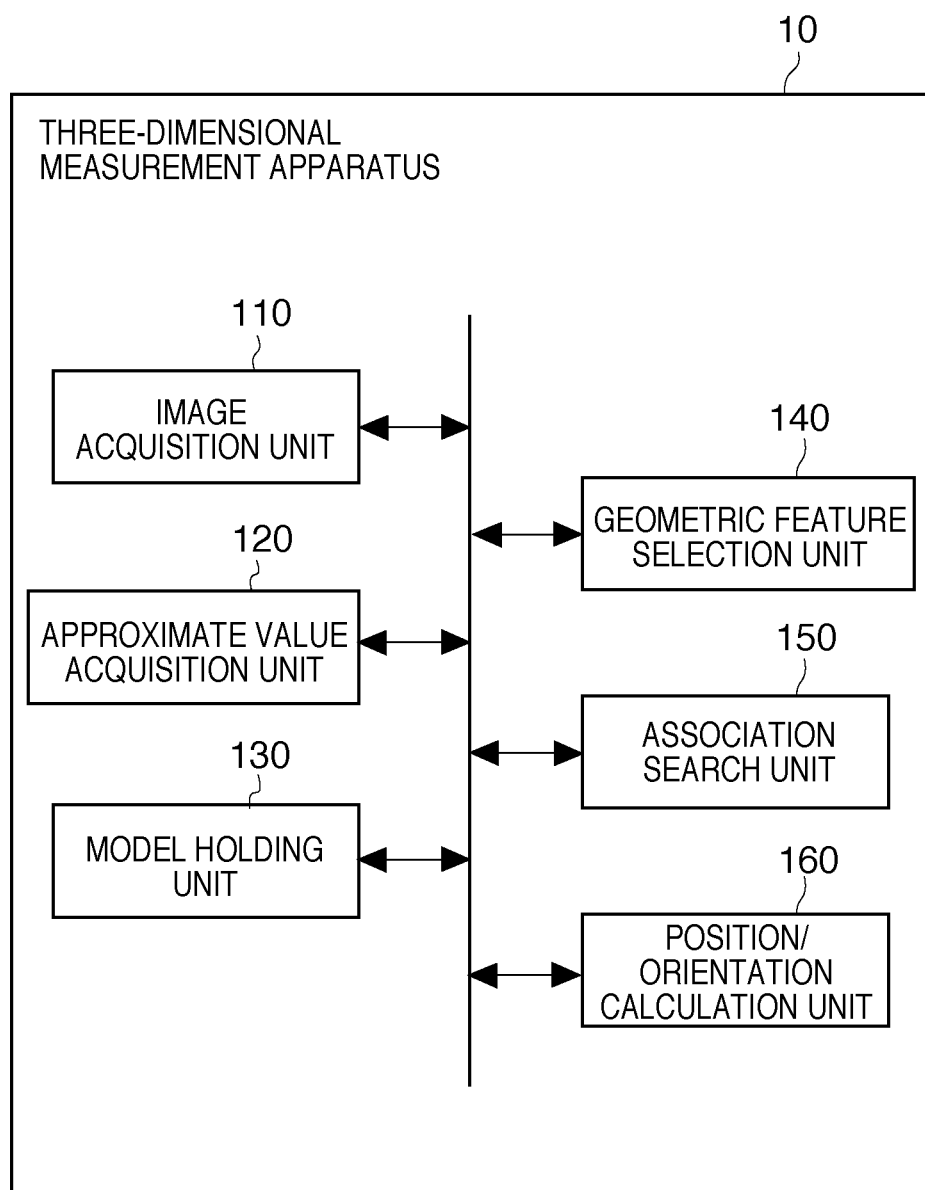
FIG. 1

FIG. 2

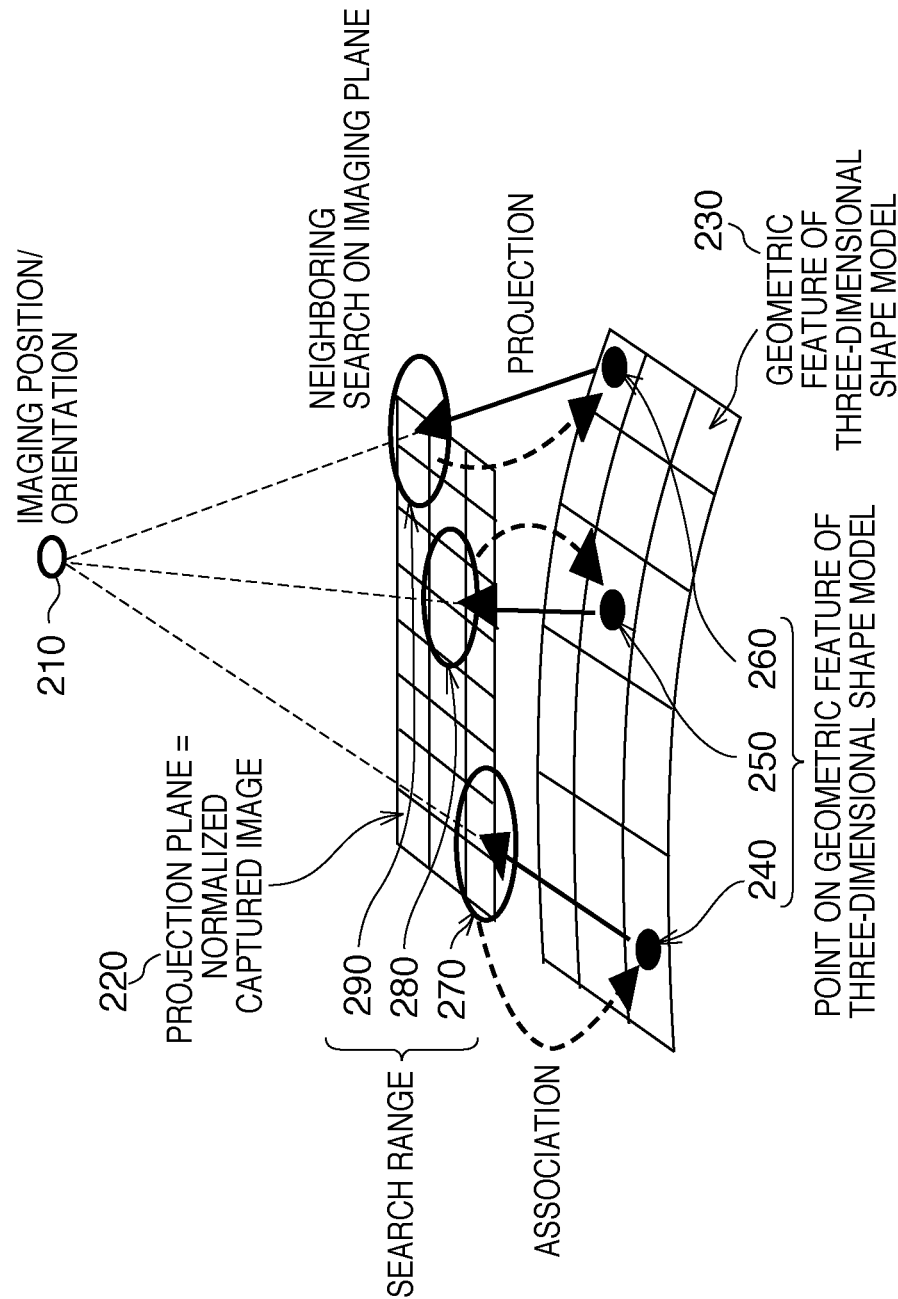


FIG. 3

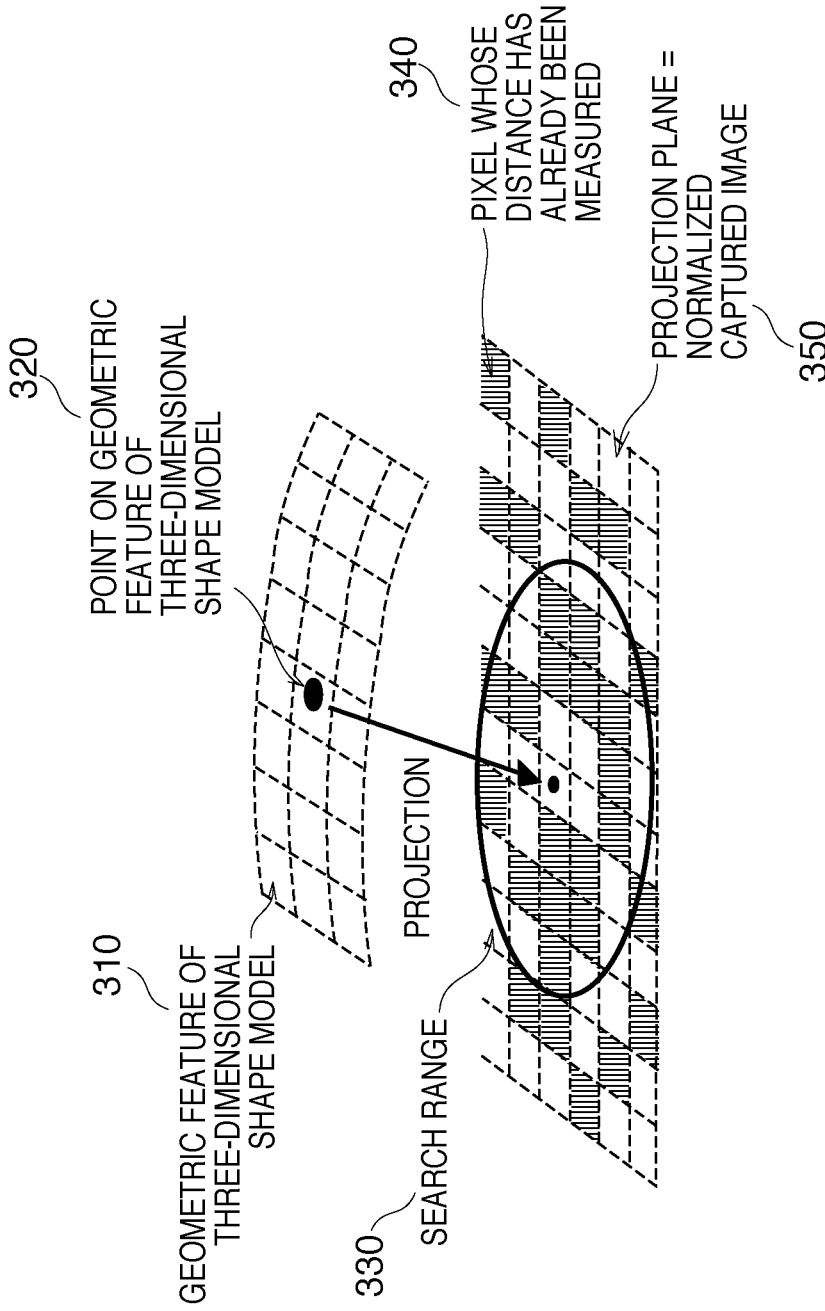


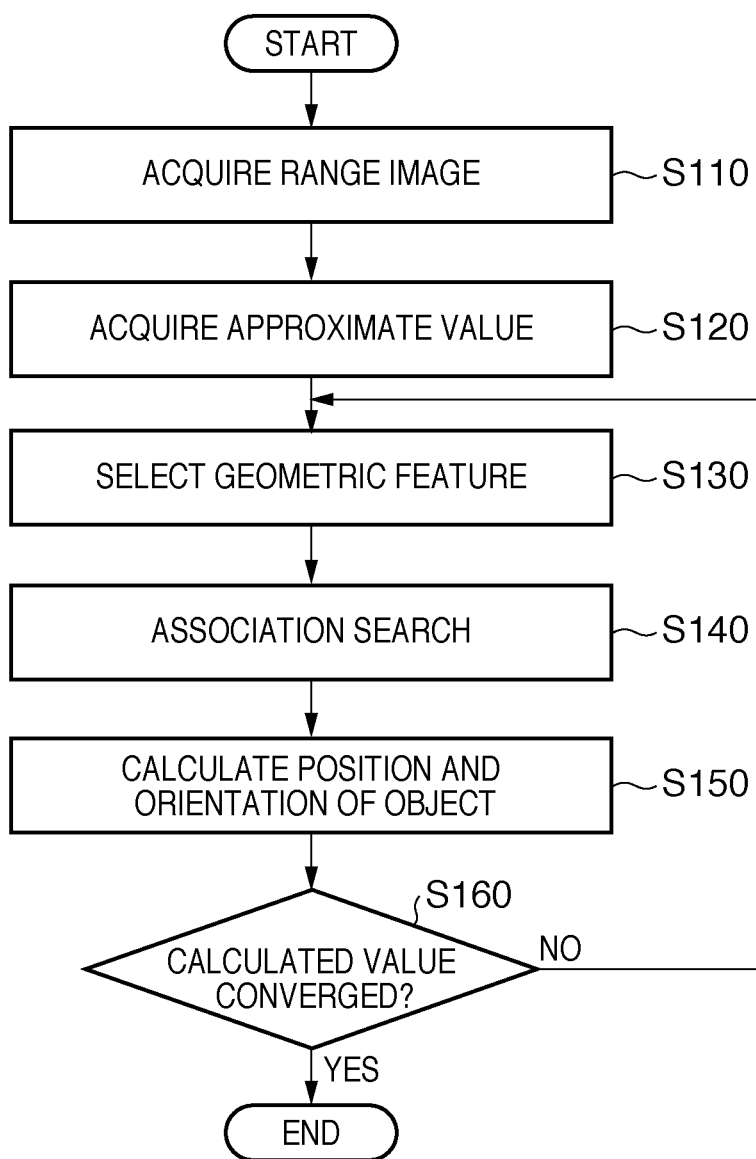
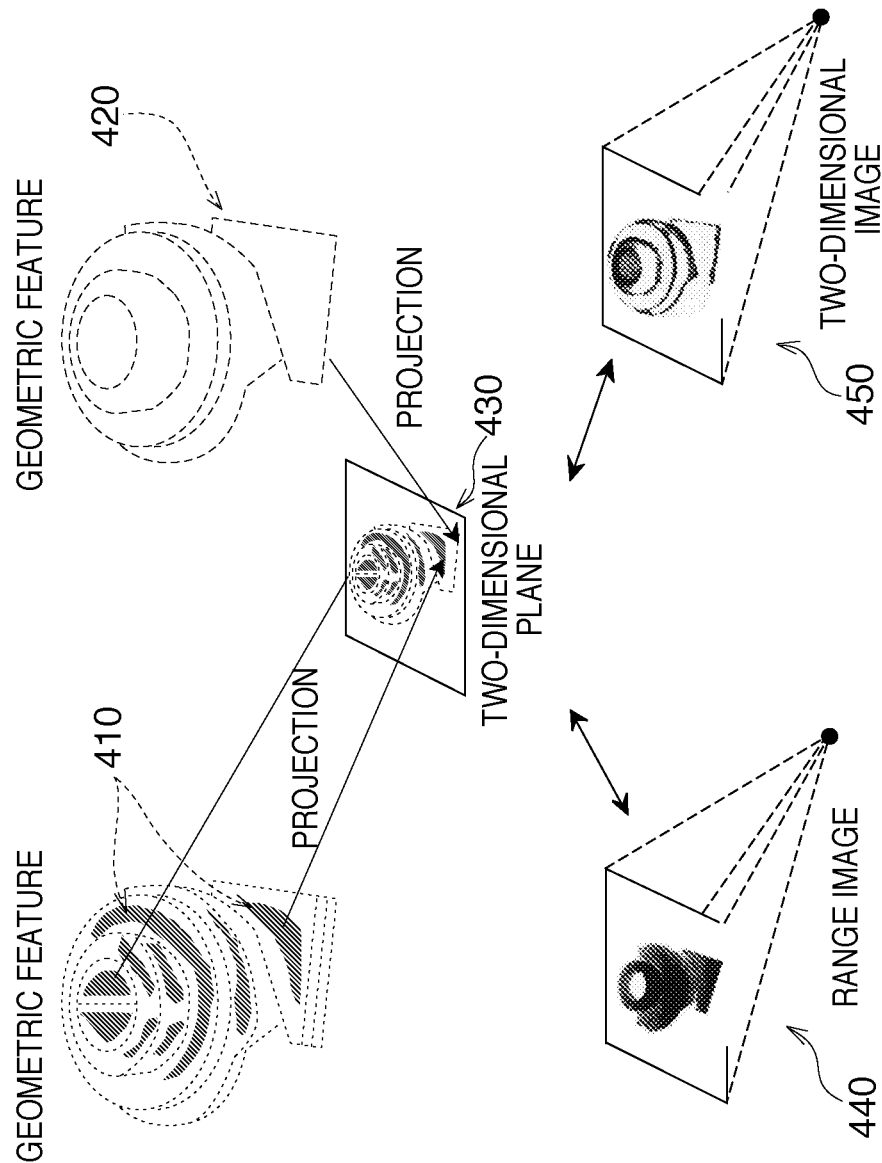
FIG. 4

FIG. 5



**THREE-DIMENSIONAL MEASUREMENT
APPARATUS, PROCESSING METHOD, AND
NON-TRANSITORY COMPUTER-READABLE
STORAGE MEDIUM**

TECHNICAL FIELD

[0001] The present invention relates to a three-dimensional measurement apparatus, processing method, and non-transitory computer-readable storage medium.

BACKGROUND ART

[0002] In recent years, along with the growth of robot technologies, robots perform complicated tasks (for example, assembling processes of industrial products), which were, up until now, performed manually. Such robots assemble parts by gripping them using an end effector such as a hand.

[0003] In order to control the robot to grip a part, a relative position and orientation between the part to be gripped and the robot (hand) have to be measured (estimated). Such position and orientation measurements are used not only when the robot grips a part but also in various purposes such as self position estimation required for the robot to autonomously move and registration between a real space and virtual object in augmented reality.

[0004] In the position and orientation measurements, methods using a two-dimensional image captured by a camera and a range image obtained from a range sensor are known. For example, measurements by means of model fitting are known. In the measurements, a three-dimensional shape model of an object is fitted to features detected from the two-dimensional image or to the range image.

[0005] In the model fitting to a two-dimensional image, the position and orientation are measured by fitting a projected image obtained by projecting a three-dimensional shape model onto an image to features detected from a two-dimensional image. In the model fitting to a range image, respective points of the range image are converted into a three-dimensional point group having three-dimensional coordinates, and a three-dimensional shape model is fitted to the three-dimensional point group on a three-dimensional space, thereby measuring the position and orientation.

[0006] As a method of measuring the position and orientation using a two-dimensional image, a method of measuring the position and orientation of a camera using edges is known (T. Drummond and R. Cipolla, "Real-time visual tracking of complex structures," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 932-946, 2002. (to be referred to as reference 1 hereinafter)). With this method, a three-dimensional shape of an object is expressed by a set of line segments (wire frame model), and projected images of three-dimensional line segments are fitted to edges detected on an image, thereby measuring the position and orientation of the object. More specifically, three-dimensional line segments are projected onto the image based on approximate values of the position and orientation, so as to detect edges in the vicinity of the projected line segments. Next, the position and orientation of a target object are measured by nonlinear optimization, so as to minimize a sum total of distances on the image between

the projected images of the line segments based on the approximate values of the position and orientation, and the corresponding edges.

[0007] On the other hand, as a method of measuring the position and orientation using a range image, a method using an ICP (Iterative Closest Point) algorithm is known (P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 14, no. 2, pp. 239-256, 1992. (to be referred to as reference 2 hereinafter)). With this method, the position and orientation of an object are measured by fitting a three-dimensional shape model of the object to three-dimensional point group data converted from a range image. Processing for searching geometric features of the three-dimensional shape model closest to three-dimensional points based on approximate values of the position and orientation, and updating the position and orientation so as to minimize a sum total of distances between points and the geometric features of the three-dimensional model is iteratively executed.

[0008] The method of measuring the position and orientation of an object using a range image requires heavy-load arithmetic processing upon searching for planes of the three-dimensional shape model corresponding to respective points of the point group data. In order to cope with this, Japanese Patent Laid-Open No. 2006-202152 (to be referred to as reference 3 hereinafter) discloses a technique that speeds up association search processing at the time of registration of a plurality of range images. In this method, small planes (meshes) assigned with index values are fitted to a range image, the index values are converted into non-overlapping colors to obtain colors of the meshes, and an index image is generated by rendering the meshes based on the imaging position and orientation of the range image. Then, a three-dimensional point group converted from another range image is projected onto the index image based on the position and orientation at the time of imaging, and corresponding colors on the index image are acquired from the coordinate values of the projected three-dimensional points. After that, the index values corresponding to the acquired colors are inversely converted to associate the meshes with the three-dimensional points converted from the range image. In this manner, the association search processing is speeded up.

[0009] In the aforementioned registration method (reference 3) using the index image, in order to register the positions of a plurality of range images, all three-dimensional points extracted from the range images are projected onto a two-dimensional image to associate them with each other. For this reason, for example, when the position and orientation of a target object are measured from an image including an object other than the target object, since all measurement points including those of the object other than the target object have to be densely projected, resulting in wasteful processing.

[0010] In case of the method of reference 3, depending on three-dimensional shape models, the model has to be remeshed so as to prevent meshes from being crushed at the time of projection. For this reason, it is difficult to use a CAD model of a target object intact as a three-dimensional shape model. Furthermore, in order to render meshes on a two-dimensional plane at high speed, dedicated hardware such as a GPU (Graphic Processing Unit) is required.

[0011] Furthermore, in case of the method of reference 1, all line segments as geometric features of a three-dimensional shape model as well as those on a back surface (portion which is not measured) of a target object have to be projected onto an image. Hence, the method of reference 1 also requires wasteful processing.

[0012] In this case, the method using a two-dimensional image is suited to, for example, an environment including many artificial objects on the premise of lines, and the method using a range image is suited to, for example, an object having a plurality of smooth planes.

[0013] Since the method using a two-dimensional image and that using a range image require different properties of information to be measured, the position and orientation measurement precision is expected to be improved by combining the model fitting to a two-dimensional image and that to a range image. In the model fitting to a two-dimensional image, processing for projecting geometric features of a three-dimensional shape model such as edges onto a two-dimensional image, and searching for corresponding geometric features on the two-dimensional image is executed, as described above. In the model fitting to a range image, measurement points are projected in place of geometric features of a three-dimensional shape model, as described above. That is, these methods use different association methods.

[0014] For this reason, the association processing cannot be executed using an identical framework, and has to be executed independently upon measuring the position and orientation of a target object using both a two-dimensional image and range image.

SUMMARY OF INVENTION

[0015] The present invention provides a technique which allows to speed up position and orientation calculations of a target object by efficiently making arithmetic operations executed upon associating a three-dimensional shape and measurement data with each other.

[0016] According to a first aspect of the present invention, there is provided a three-dimensional measurement apparatus comprising; image acquisition means for acquiring a range image in which pixels represent distances from a target object; approximate value acquisition means for acquiring approximate values indicating a position and an orientation of the target object; selection means for selecting points corresponding to geometric features of a three-dimensional shape model of the target object; projection means for projecting a plurality of points corresponding to the geometric features selected by the selection means onto the range image based on the approximate values indicating the position and the orientation of the target object and imaging parameters at the time of imaging of the range image, search means for searching regions of predetermined ranges respectively from the plurality of projected points for geometric features on the range image which correspond to the geometric features of the three-dimensional shape model, and associating the geometric features with each other; and position/orientation calculation means for calculating a position and an orientation of the target object using differences of distances on a three-dimensional space between the geometric features of the three-dimensional shape model and the geometric features on the range image, which are associated with each other by the search means.

[0017] According to a second aspect of the present invention, there is provided a processing method for a three-dimensional measurement apparatus for calculating a position and an orientation of a target object, characterized by comprising: a step of acquiring a range image in which pixels represent distances from a target object; a step of acquiring approximate values indicating a position and an orientation of the target object; a step of selecting points corresponding to geometric features of a three-dimensional shape model of the target object; a step of projecting a plurality of selected points corresponding to the geometric features onto the range image based on the approximate values indicating the position and the orientation of the target object and imaging parameters at the time of imaging of the range image; a step of searching regions of predetermined ranges respectively from the plurality of projected points for geometric features on the range image which correspond to the geometric features of the three-dimensional shape model, and associating the geometric features with each other; and a step of calculating a position and an orientation of the target object using differences of distances on a three-dimensional space between the geometric features of the three-dimensional shape model and the geometric features on the range image, which are associated with each other.

[0018] According to a third aspect of the present invention, there is provided a non-transitory computer-readable storage medium storing a computer program for controlling a computer incorporated in a three-dimensional measurement apparatus for calculating a position and an orientation of a target object, to function as: image acquisition means for acquiring a range image in which pixels represent distances from a target object; approximate value acquisition means for acquiring approximate values indicating a position and an orientation of the target object; selection means for selecting points corresponding to geometric features of a three-dimensional shape model of the target object; projection means for projecting a plurality of points corresponding to the geometric features selected by the selection means onto the range image based on the approximate values indicating the position and the orientation of the target object and imaging parameters at the time of imaging of the range image, search means for searching regions of predetermined ranges respectively from the plurality of projected points for geometric features on the range image which correspond to the geometric features of the three-dimensional shape model, and associating the geometric features with each other; and position/orientation calculation means for calculating a position and an orientation of the target object using differences of distances on a three-dimensional space between the geometric features of the three-dimensional shape model and the geometric features on the range image, which are associated with each other by the search means.

[0019] Further features of the present invention will be apparent from the following description of exemplary embodiments (with reference to the attached drawings).

BRIEF DESCRIPTION OF DRAWINGS

[0020] The accompanying drawings, which are incorporated in and constitute a part of the specification, illustrate embodiments of the invention, and together with the description, serve to explain the principles of the invention.

[0021] FIG. 1 is a block diagram showing an example of the arrangement of a three-dimensional measurement apparatus according to an embodiment of the present invention;

[0022] FIG. 2 is a view showing an example of association processing by an association search unit 150 shown in FIG. 1;

[0023] FIG. 3 is a view showing an example of association processing by the association search unit 150 shown in FIG. 1;

[0024] FIG. 4 is a flowchart showing an example of the operation in a three-dimensional measurement apparatus 10 shown in FIG. 1; and

[0025] FIG. 5 is a view showing an example of association processing according to the second embodiment.

DESCRIPTION OF EMBODIMENTS

[0026] An exemplary embodiment(s) of the present invention will now be described in detail with reference to the drawings. It should be noted that the relative arrangement of the components, the numerical expressions and numerical values set forth in these embodiments do not limit the scope of the present invention unless it is specifically stated otherwise.

First Embodiment

[0027] FIG. 1 is a block diagram showing an example of the arrangement of a three-dimensional measurement apparatus according to an embodiment of the present invention. Note that a three-dimensional measurement apparatus 10 incorporates a computer. The computer includes a main control unit such as a CPU, and storage units such as a ROM (Read Only Memory), RAM (Random Access Memory), and HDD (Hard Disk Drive). The computer may also include input/output units such as buttons and a display or touch panel, and a communication unit such as a network card. Note that these components are connected via, for example, a bus, and are controlled when the main control unit executes programs stored in the storage units.

[0028] The three-dimensional measurement apparatus 10 projects (arbitrary) points on geometric features of a three-dimensional shape model onto a two-dimensional plane, and searches for geometric features on the two-dimensional plane (range image), which exist within regions of predetermined ranges from the projected points on the two-dimensional plane. Then, the geometric features of the three-dimensional shape model are associated with those on the range image, and the position and orientation of a target object are calculated based on the associated geometric features.

[0029] The three-dimensional measurement apparatus 10 is configured to include an image acquisition unit 110, approximate value acquisition unit 120, model holding unit 130, geometric feature selection unit 140, association search unit 150, and position/orientation calculation unit 160.

[0030] The image acquisition unit 110 acquires a range image. In this embodiment, the image acquisition unit 110 is implemented by, for example, a range sensor which captures a range image. However, the present invention is not limited to this. For example, the image acquisition unit 110 may adopt an arrangement for simply acquiring a range image captured by an external range sensor. The range image is an image in which each pixel has depth information (distance information). The range sensor can adopt an active system

that captures reflected light of laser light or slit light with which a target is irradiated, and measures a distance by triangulation. Note that the range sensor need only capture a range image, and its system is not particularly limited. For example, the range sensor may adopt a time-of-flight system using a time of flight of light, or a passive system which calculates, by triangulation, depths of respective pixels from images captured by a stereo camera. Assume that imaging parameters (a field angle, resolution, focal length, etc. at the time of imaging) of the range sensor are given. The imaging parameters are used in the association search unit 150 (to be described later).

[0031] The approximate value acquisition unit 120 acquires approximate values of the position and orientation of an object with respect to the three-dimensional measurement apparatus 10. In this embodiment, the position and orientation of the object with respect to the three-dimensional measurement apparatus 10 express those of the object with reference to the image acquisition unit 110. However, the image acquisition unit 110 need not always be used as a reference. For example, when the relative position and orientation of the object to a coordinate system of the image acquisition unit 110 are given, and they remain unchanged, another portion of the three-dimensional measurement apparatus 10 may be used as a reference.

[0032] In this embodiment, as the approximate values of the position and orientation of the object, measurement values which were measured by the three-dimensional measurement apparatus 10 previously (for example, at an immediately preceding timing) are used. However, the approximate values of the position and orientation of the object need not always be such values. For example, time-series filtering processing (for example, a linear filter or Kalman filter) is applied to the previously measured measurement values of the position and orientation of the object to estimate the motion velocity and angular velocity of the object. Then, values obtained by predicting the position and orientation of the object based on the estimation results may be used as the approximate values. Alternatively, the position and orientation of the object obtained from a sensor may be used as the approximate values. In this case, the sensor need only measure the position and orientation of the object in six degrees of freedom, and its system (for example, a magnetic system, optical system, or ultrasonic system) is not particularly limited. Note that when an approximate position and orientation where the object is placed are determined in advance, these values may be used as the approximate values. Furthermore, by applying object recognition processing to a captured image or range image obtained by capturing an image of a visual scene including the object, the approximate values of the position and orientation of the object may be measured.

[0033] The model holding unit 130 holds three-dimensional geometric model data (to be also referred to as a three-dimensional shape model hereinafter) of a target object (to be also simply referred to as an object hereinafter). The three-dimensional shape model can use the same model as in the conventional method, and a detailed description thereof will not be given. For example, the three-dimensional shape model is defined by pieces of information of planes configured by connecting point sets or points, and pieces of information of line segments which configure each plane.

[0034] The geometric feature selection unit **140** selects points on arbitrary geometric features from the three-dimensional shape model of the target object. In this embodiment, assume that the three-dimensional shape model is configured by a plurality of NURBS (Non-Uniform Rational B-Spline) curved surfaces, and geometric features of this model are sets of points and normals uniformly sampled from the respective NURBS curved surfaces. Of course, the present invention is not limited such specific geometric features. For example, the geometric features of the three-dimensional shape model may be the NURBS curved surfaces themselves or meshes when the three-dimensional shape model is expressed by meshes.

[0035] The geometric feature selection unit **140** may perform the above selection based on information that prescribes geometric features which are easily measured from the three-dimensional shape model or measurement data and those which are effective to calculate the position and orientation. That is, the geometric features can be those which can be used when the position/orientation calculation unit **160** calculates the position and orientation of the target object. Since the geometric feature selection unit **140** selects arbitrary geometric features, even when the position and orientation are to be measured for an image which includes an object other than the target object, association processing can be executed without increasing a calculation amount.

[0036] The geometric feature selection unit **140** selects points on geometric features of the three-dimensional shape model based on the imaging parameters of the range sensor and the approximate values of the position and orientation of the object acquired by the approximate value acquisition unit **120**. Based on them, only geometric features which can be measured from a range image are selected.

[0037] More specifically, the three-dimensional shape model is rendered from all directions, and geometric features which are viewable from the respective directions are registered in association with the directions. Then, the geometric feature selection unit **140** selects geometric features registered in directions closest to a line of sight vector, which is calculated based on the approximate values of the position and orientation of the object, and the imaging parameters. Note that inner products between vectors of the respective directions and normals to geometric features may be compared, and only geometric features whose normals oppose the direction vectors may be registered.

[0038] Note that the geometric feature selection unit **140** may select points on geometric features based on a user instruction. That is, the user may manually select points on geometric features with reference to a GUI that displays the three-dimensional shape model of the target object. Alternatively, the center of a surface of each geometric feature may be selected as a point of that geometric feature of the three-dimensional shape model. In addition, points detected by a feature detector disclosed in A. E. Johnson and M. Hebert, "Efficient multiple model recognition in cluttered 3-d scenes," Proc. Computer Vision and Pattern Recognition, pp. 671-677, 1998. may be selected as points on geometric features of the three-dimensional shape model. Furthermore, points which are uniform on a range image may be selected based on the approximate values of the position and orientation of the object acquired by the approximate value acquisition unit **120**. More specifically, the three-dimensional shape model is projected onto a two-dimensional image based on the approximate values of

the position and orientation of the object and the imaging parameters of the range sensor, and points on the three-dimensional shape model, which are uniformly sampled on the two-dimensional image, are inversely projected onto a three-dimensional space. Then, points on geometric features, which become uniform on range image, are calculated and selected. In addition, the selection method is not particularly limited as long as points corresponding to geometric features of a three-dimensional shape model are selected from the surface of the model.

[0039] The association search unit **150** associates geometric features of the three-dimensional shape model with those on a range image. Then, the association search unit **150** outputs the associated sets of the geometric features of the three-dimensional shape model and those on the range image.

[0040] As shown in FIG. 2, the association search unit **150** projects a point **240** selected by the geometric feature selection unit **140** onto a two-dimensional plane (projection plane) **220**. Note that this projection is performed based on the given imaging parameters (field angle, resolution, focal length, etc.) of a range image, and the approximate values of the position and orientation of the object acquired by the approximate value acquisition unit **120**. Formula (1) below is used to convert three-dimensional coordinates (x, y, z) of a point selected by the geometric feature selection unit **140** into coordinates (u, v) on the two-dimensional plane (projection plane) by perspective projection.

$$(u, v)^T = (f \cdot x/z, f \cdot y/z)^T \quad (1)$$

where f is a focal length.

[0041] Pixels on the projection plane and range image uniquely correspond to each other, and a range **270** in the vicinity of the two-dimensional coordinates on the range image corresponding to the point (projected point) projected onto the projection plane is searched for a geometric feature to be associated on the range image. Since the projection calculation of a point can be attained by formula (1), an association search can be a two-dimensional search around the projected point. For this reason, since the calculation amount can be reduced, association processing between the geometric features of the three-dimensional shape model and those on a range image can be speeded up.

[0042] A geometric feature on a range image is searched for within a region Ω **330** formed by an elliptic region having the projected point as the center, as shown in FIG. 3. Note that the search range may be a pixel of the projected point or may have a rectangular shape having an arbitrary size. A size and shape of the search region Ω **330** may be decided based on at least any of, for example, the approximate values of the position and orientation of the object, the imaging parameters of the image acquisition unit **110**, and a motion of the object or image acquisition unit **110**. That is, the search range Ω **330** may be a rectangular or elliptic region decided based on the line of sight direction of the image acquisition unit **110** and the normal direction of a geometric feature, or an elliptic region having a motion vector as a major axis.

[0043] In this embodiment, assume that a geometric feature on a range image is defined by coordinates of a three-dimensional point obtained from depth information (distance information) included in a pixel corresponding to the projected point. As a search method of a geometric feature on a range image, a measurement point within the

search range Ω 330 having a shortest distance from a projected point of a measurement point 340 can be selected. Formula (2) is used to calculate the shortest distance of coordinates g_i on a range image of a measurement point measured at a projected point $p=(u, v)^T$.

$$\operatorname{argmin}\{\|g_i - p\|\}, g_i \in \Omega \quad (2)$$

[0044] Note that this search method need not always be adopted. For example, feature amounts of respective pixels within the search region Ω 330 may be calculated by a feature detector, and a feature amount most similar to that of a geometric feature of the three-dimensional shape model may be selected. Alternatively, for example, a median value or average value of coordinates on a three-dimensional space of measurement points within the search region Ω 330, that is, coordinates on the space, which are calculated by statistical processing may be selected. In addition, a geometric feature on a range image is not limited to a three-dimensional point, but it may be parameters of a curved surface or plane, which are fitted to measurement points within the search region Ω , as long as the geometric feature can be used when the position/orientation calculation unit 160 calculates the position and orientation of the target object.

[0045] Also, association error excluding processing may be executed. That is, a normal to a geometric feature selected on a range image is calculated, and the calculated normal is compared with that to a geometric feature of the three-dimensional shape model. If these normal directions are different beyond a predetermined value, these geometric features are not associated with each other. Note that the normal to the geometric feature on the range image can be calculated by applying principal component analysis to a three-dimensional point group adjacent to the geometric feature, and deciding a third principal component as a normal. When the association error excluding processing is executed, errors which associate measurement data of different geometric features with those of the three-dimensional shape model can be eliminated.

[0046] The position/orientation calculation unit 160 calculates the position and orientation of the target object using the sets of the geometric features of the three-dimensional shape model and those on the range image, which are associated with each other by the association search unit 150. In the calculations of the position and orientation of the target object in this embodiment, arithmetic processing is repeated to optimize values obtained by the arithmetic processing. In this manner, differences of distances on the three-dimensional space between the sets of the geometric features of the three-dimensional shape model and those on the range image, which are associated with each other, are minimized.

[0047] In this embodiment, the geometric features of the three-dimensional shape model are sets of coordinates on the three-dimensional space and pieces of normal information, and those on the range image are coordinates of three-dimensional points. A geometric feature of the three-dimensional shape model is considered as a small plane, and a distance between the plane on the three-dimensional shape model and a three-dimensional point calculated from a range image is used as a difference. Then, by minimizing this difference, the position and orientation of the target object are measured.

[0048] In this case, the difference depends on an expression method of a geometric feature. For example, if a geometric feature of the three-dimensional shape model is a point, and that on a range image is a point, the difference may be a distance between these points. Alternatively, for example, when a geometric feature of the three-dimensional shape model is a plane and that on a range image is a plane, the difference may be a distance between these planes. Alternatively, for example, when a geometric feature of the three-dimensional shape model is a plane and that on a range image is a point, the difference may be a distance between the plane and point. Furthermore, the difference may be a distance which is approximately calculated using, for example, an implicit polynomial. That is, a calculation method is not particularly limited as long as the arrangement for measuring the position and orientation of the target object using an evaluation function based on differences of the distances between the geometric features of the three-dimensional shape model and those on the range image is adopted.

[0049] An example of the operation in the three-dimensional measurement apparatus 10 shown in FIG. 1 will be described below using the flowchart shown in FIG. 4. In this case, the sequence of processing upon measuring the position and orientation of an object will be explained.

[0050] The three-dimensional measurement apparatus 10 controls the image acquisition unit 110 to capture a range image including depth information (distance information) from the range sensor. Thus, the range image is acquired (S110).

[0051] After completion of acquisition of the range image, the three-dimensional measurement apparatus 10 controls the approximate value acquisition unit 120 to acquire the approximate values of the position and orientation of the object with respect to the three-dimensional measurement apparatus 10 (S120). As described above, the approximate values of the position and orientation of the object can use, for example, the previously measured position and orientation, predicted values of the position and orientation by adding tinges of estimation results of motions such as a velocity and angular velocity, measurement values by another sensor, and the object recognition results from a two-dimensional image and range image. Note that this embodiment uses the previously measured approximate values of the position and orientation as the approximate values of the position and orientation of the object. Note that the order of the processes in steps S110 and S120 may be replaced.

[0052] Next, the three-dimensional measurement apparatus 10 starts position/orientation measurement processing of the target object using, for example, the range image acquired by the process in step S110 and the approximate values acquired by the process in step S120. The position/orientation measurement processing according to this embodiment uses a Levenberg-Marquardt method. More specifically, the approximate values of the position and orientation of the object are repetitively corrected by iterative arithmetic operations. In this way, the differences of the distances on the three-dimensional space between the geometric features of the three-dimensional shape model and those on the range image are minimized, and values obtained as a result of this processing are determined as the position and orientation of the target object. Of course, the present invention is not limited to the Levenberg-Marquardt

method. For example, other optimization methods such as a Gauss-Newton method and steepest descent method may be used. Alternatively, other nonlinear optimization calculation methods such as a conjugate gradient method may be used.

[0053] After the position/orientation measurement processing is started, the three-dimensional measurement apparatus **10** acquires the approximate values of the target object obtained by the process in step **S120** as initial values of the position/orientation measurement processing. Then, the three-dimensional measurement apparatus **10** controls the geometric feature selection unit **140** to select geometric features of the three-dimensional shape model based on the approximate values (**S130**).

[0054] Upon completion of the selection, the three-dimensional measurement apparatus **10** controls the association search unit **150** to project the geometric features of the three-dimensional shape model onto the range image (two-dimensional plane) based on the approximate values of the position and orientation of the target object and the imaging parameters (field angle, resolution, focal length, etc.) of the range sensor. Then, the association search unit **150** associates the geometric features of the three-dimensional shape model with those on the range image, and outputs the associated sets of the geometric features of the three-dimensional shape model and those on the range image (**S140**). The three-dimensional measurement apparatus **10** controls the position/orientation calculation unit **160** to update the approximate values of the position and orientation of the target object by the nonlinear optimization method (**S150**).

[0055] Finally, the three-dimensional measurement apparatus **10** controls the position/orientation calculation unit **160** to make convergence determination. If calculated values are converged (YES in step **S160**), the processing ends; otherwise (NO in step **S160**), the process returns to step **S130**, and the aforementioned processes in steps **S130** to **S150** are repeated. Note that in the convergence determination, for example, if differences between square sums of error vectors of the position and orientation before and after the update process are nearly "0", it can be determined that the calculated values are converged.

[0056] As described above, according to the first embodiment, upon associating the geometric features of the three-dimensional shape model and range image with each other, points on the geometric features of the three-dimensional shape model are selected, and the selected points on the geometric features are projected onto the range image. Then, by searching for geometric features which exist within regions of predetermined ranges from the projected points on the range image, the geometric features of the three-dimensional shape model are associated with those on the range image. In this way, arithmetic operations can be efficiently executed upon associating the three-dimensional shape model with measurement data.

Second Embodiment

[0057] The second embodiment will be described below. The second embodiment will explain a case in which a two-dimensional image is input as an input image in addition to a range image. FIG. **5** is a view showing an overview of processing for projecting geometric features of a three-dimensional shape model onto a two-dimensional plane, and searching a two-dimensional image and range image for corresponding geometric features.

[0058] For both of a range image **440** and two-dimensional image **450**, geometric features **410** and **420** of a three-dimensional shape model are projected onto a two-dimensional plane (projection plane) **430** to search for corresponding geometric features on the images from the two-dimensional plane, thereby associating those geometric features with each other.

[0059] For this reason, the efficiency of arithmetic processing required to associate the three-dimensional shape model and measurement data with each other can be improved. Assume that in this embodiment, the range image **440** and two-dimensional image **450** are captured at the same viewpoint. In this case, the geometric features on the two-dimensional image can be relatively easily compared with those on the range image, thus allowing to determine a jump edge (isolated edge) in addition to elimination of association errors due to the influence of, for example, a shadow.

[0060] In this case, the overall arrangement of a three-dimensional measurement apparatus **10** according to the second embodiment is the same as that of the first embodiment, but processes in respective units are slightly different. Such differences will be described below. Note that the following description will be given while focusing on the different processes.

[0061] An image acquisition unit **110** acquires a two-dimensional image and range image. The range image is captured by a range sensor in the same manner as in the first embodiment. The two-dimensional image is captured by a camera (imaging apparatus) which is used to capture a normal two-dimensional image. The two-dimensional image to be captured may be a density image or a color image. A camera model such as a focal length and principal point position of a camera and lens distortion parameters is calibrated in advance by a method disclosed in, for example, R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," IEEE Journal of Robotics and Automation, vol. RA-3, no. 4, 1987.

[0062] In the second embodiment, for example, a range sensor of an active system, which is exemplified in the first embodiment, is used. This is because the range sensor of the active system, that using a stereo camera, and the like can also simultaneously capture a normal two-dimensional image. For this reason, the two-dimensional image and range image can be captured from the same viewpoint. However, the two-dimensional image and range image need not always be captured from the same viewpoint. Note that when the imaging positions and orientations of imaging apparatuses in the two-dimensional image and range image are closer to each other, and the geometrical relationship between the two apparatuses is given, one of the two-dimensional image and range image may undergo projection transformation based on the geometrical relationship between the two apparatuses, and an image obtained by that transformation may be used as that captured from the same viewpoint.

[0063] An association search unit **150** associates the geometric features of the three-dimensional shape model, which are selected by the geometric feature selection unit **140**, with those on the two-dimensional image and range image, and outputs the associated sets of the geometric features of the three-dimensional shape model and those on the range image and two-dimensional image. The association search

unit **150** attains the association processing by projecting the geometric features of the three-dimensional shape model, which are selected by the geometric feature selection unit **140**, onto a two-dimensional plane (projection plane), as in the first embodiment. Note that the association processing between the three-dimensional shape model and range image is the same as that in the first embodiment. The association processing between the three-dimensional shape model and two-dimensional image can use edges as geometric features. Since the association processing based on the edges can use the conventional technique, a detailed description thereof will not be given. Note that the geometric features may be feature points detected by a feature detector disclosed in C. Harris and M. Stephens, "A combined corner and edge detector," Proc. 4th Alvey Vision Conf., pp. 147-151, Manchester, U.K., 1988. In addition, the geometric features can be those which can be used when a position/orientation calculation unit **160** measures the position and orientation of the target object.

[0064] Since the two-dimensional image and range image according to this embodiment are those captured from the same viewpoint, pixels on the projection plane, two-dimensional image, and range image uniquely correspond to each other. For this reason, as in the first embodiment, using a predetermined range (a region) around a projected point as a search region Ω , a geometric feature on an image corresponding to that of the three-dimensional shape model is two-dimensionally searched for. The search region Ω may be a rectangular or elliptic region as in the first embodiment. Note that upon searching for edges on the two-dimensional image, a line having an arbitrary length may undergo a linear search in the normal direction of an edge.

[0065] Furthermore, association error excluding processing described in the first embodiment may be applied to the two-dimensional image. The association error excluding processing is executed based on, for example, comparison between feature amounts and that between edge directions. If features which correspond to the edges or feature amounts detected on the two-dimensional image cannot be detected from the range image, the association error excluding processing may be executed to skip the association processing since it is determined that they are the edges or feature amounts caused by the influence of, for example, a shadow.

[0066] Also, processing for determining whether or not an edge detected on the two-dimensional image is a jump edge may be executed using the range image. Since the range image and two-dimensional image are captured from the same viewpoint, the association error excluding processing can be efficiently executed using both the two-dimensional image and range image.

[0067] The position/orientation calculation unit **160** calculates the position and orientation of the target object using the sets of the geometric features of the three-dimensional shape model and those on the two-dimensional image and range image, which are associated with each other by the association search unit **150**. Upon calculations of the position and orientation of the target object in this embodiment, arithmetic processing is repeated to optimize the values obtained by that processing. In this way, evaluation functions based on both of differences of distances on a three-dimensional space between the geometric features of the three-dimensional shape model and those on the two-dimensional space, and differences of distances on the three-

dimensional space between the geometric features of the three-dimensional shape model and those on the range image are minimized.

[0068] Note that since the two-dimensional image does not include any depth information, the geometric features of the two-dimensional image are projected onto the three-dimensional space in, for example, a state without any depth information. Of course, depth information may be assigned to each of the geometric features of the two-dimensional image using a predetermined algorithm or predetermined values.

[0069] In addition, the measurement method is not particularly limited as long as it measures the position and orientation of the target object based on evaluation functions based on the differences between the geometric features of the three-dimensional shape model and those on the two-dimensional image and range image. Note that the sequence of position/orientation measurement processing according to the second embodiment is the same as that in the first embodiment, and a description thereof will not be repeated.

[0070] Note that in the second embodiment, the geometric features to be associated of the three-dimensional shape model may be curved surfaces, planes, or curves. However, in order to project complicated shapes such as NURBS curved surfaces popularly used in a CAD model onto a two-dimensional plane, tessellation or processing for clipping curved surfaces using contour lines is required, and a huge time is required for arithmetic processing unless dedicated hardware for geometric calculations such as a GPU is available. Hence, as in the first embodiment, the geometric features of the three-dimensional shape model are expressed as points or line segments, and the points or line segments are projected onto the two-dimensional plane. With this projection, the association processing can be speeded up without requiring any dedicated hardware.

[0071] As described above, according to the second embodiment, arithmetic operations executed upon associating the three-dimensional shape model and measurement data with each other can be efficiently made as in the first embodiment.

[0072] The representative embodiments of the present invention have been exemplified. However, the present invention is not limited to the aforementioned and illustrated embodiments, and allows modifications as needed without departing from the scope of its gist.

[0073] Note that the present invention can adopt embodiments in the form of, for example, a system, apparatus, method, program, or storage medium. More specifically, the present invention may be applied to either a system configured by a plurality of devices or an apparatus including a single device.

Other Embodiments

[0074] Aspects of the present invention can also be realized by a computer of a system or apparatus (or devices such as a CPU or MPU) that reads out and executes a program recorded on a memory device to perform the functions of the above-described embodiment(s), and by a method, the steps of which are performed by a computer of a system or apparatus by, for example, reading out and executing a program recorded on a memory device to perform the functions of the above-described embodiment(s). For this purpose, the program is provided to the computer for example via a network or from a recording medium of

various types serving as the memory device (for example, computer-readable storage medium).

[0075] While the present invention has been described with reference to exemplary embodiments, it is to be understood that the invention is not limited to the disclosed exemplary embodiments. The scope of the following claims is to be accorded the broadest interpretation so as to encompass all such modifications and equivalent structures and functions.

[0076] This application claims the benefit of Japanese Patent Application No. 2010-039201 filed on Feb. 24, 2010, which is hereby incorporated by reference herein in its entirety.

[0077] 1-7. (canceled)

8. A three-dimensional measurement apparatus comprising:

- a processor; and
- a memory having stored therein a program which, when executed by the processor, causes the apparatus to:
 - (1) acquire (a) a range image in which pixels represent distances from a target object and (b) a two-dimensional image representing a shape of the target object;
 - (2) acquire approximate values of a position and an orientation of the target object;
 - (3) (a) associate a plurality of geometric features on the range image with a plurality of first geometric features of a three-dimensional shape model and (b) associate a plurality of geometric features on the two-dimensional image with a plurality of second geometric features of the three-dimensional shape model, based on the approximate values of the position and the orientation of the target object; and
 - (4) derive a position and an orientation of the target object based on results of (a) the association between the plurality of geometric features on the range image and the plurality of first geometric features of the three-dimensional shape model and (b) the association between the plurality of geometric features on the two-dimensional image and the plurality of second geometric features of the three-dimensional shape model.

9. The apparatus according to claim **8**, wherein a normal direction to a geometric feature of the three-dimensional shape model is compared with a normal direction to a geometric feature detected from a predetermined range of a projected point on the range image, and when the two normal directions are different beyond a predetermined value, association processing of the plurality of geometric features is skipped.

10. The apparatus according to claim **8**, wherein the two-dimensional image is captured from the same viewpoint as the range image, and

wherein edges of the target object are detected from the two-dimensional image as the geometric features of the two-dimensional image, and association processing between an edge which is not included in the geometric features on the range image and a corresponding geometric feature of the three-dimensional shape model upon execution of the association processing between

the detected edges and the geometric features of the three-dimensional shape model is skipped.

11. The apparatus according to claim **8**, wherein the geometric features of the three-dimensional shape model are expressed by points or line segments.

12. A processing method for a three-dimensional measurement apparatus for calculating a position and an orientation of a target object, the method comprising:

- a step of acquiring (a) a range image in which pixels represent distances from a target object and (b) a two-dimensional image representing a shape of the target object;
- a step of acquiring approximate values of a position and an orientation of the target object;
- (a) associating a plurality of geometric features on the range image with a plurality of first geometric features of the three-dimensional shape model (b) associating a plurality of geometric features on the two-dimensional image with a plurality of second geometric features of the three-dimensional shape model, based on the approximate values of the position and the orientation of the target object; and
- a step of deriving a position and an orientation of the target object based on results of (a) the association between the plurality of geometric features on the range image and the plurality of first geometric features of the three-dimensional shape model and (b) the association between the plurality of geometric features on the two-dimensional image and the plurality of second geometric features of the three-dimensional shape model.

13. A non-transitory computer-readable storage medium storing a computer program for controlling a computer incorporated in a three-dimensional measurement apparatus for calculating a position and an orientation of a target object, to cause the computer to:

- acquire (a) a range image in which pixels represent distances from a target object and (b) a two-dimensional image representing a shape of the target object;
- acquire approximate values of a position and an orientation of the target object;
- (a) associate a plurality of geometric features on the range image with a plurality of first geometric features of the three-dimensional shape model and (b) associate a plurality of geometric features on the two-dimensional image with a plurality of second geometric features of the three-dimensional shape model, based on the approximate values of the position and the orientation of the target object; and
- derive a position and an orientation of the target object based on results of (a) the association between the plurality of geometric features on the range image and the plurality of first geometric features of the three-dimensional shape model and (b) the association between the plurality of geometric features on the two-dimensional image and the plurality of second geometric features of the three-dimensional shape model.

* * * * *