

DESCRIPTION

IMAGE PROCESSING APPARATUS, IMAGE PROCESSING METHOD, AND COMPUTER-
READABLE RECORDING MEDIUM

5 TECHNICAL FIELD

[0001] The invention relates to an image processing apparatus and an image processing method for enabling a three-dimensional shape to be constructed from a plurality of images, and further relates to a computer-readable recording medium that has a program for realizing the apparatus and method recorded thereon.

10

BACKGROUND ART

[0002] In recent years, a technique for constructing the three-dimensional shape of a target captured in images have been garnering attention. SfM (Structure from Motion) is known as a representative example of such a technique. In SfM, shooting of a specific target is performed a plurality of times while changing the viewpoint of a camera, and the three-dimensional shape of the specific target is reconstructed from a plurality of obtained images.

15

[0003] Specifically, first, in SfM, feature amount (for example, SIFT feature amount or SURF feature amount) of each image are calculated, and feature points that have high robustness to enlargement/reduction, rotation, and a change in illuminance of the image are extracted. Next, matching of the extracted feature points is executed between the images, and pairs of matched feature points are extracted. Next, for example, using robust estimation, the geometric relations of pairs of feature points are calculated, and incorrect pairs of feature points are excluded.

20

[0004] Subsequently, a fundamental matrix is calculated based on the geometric relations of every several pairs of feature points, and the geometric relations of pairs of feature points are adjusted such that the differences between the calculated fundamental matrices are minimized. The three-dimensional shape (group of points) is then reconstructed based on the adjusted geometric relations. Note that methods for adjusting errors at this time include a processing technique called bundle adjustment.

25

[0005] Incidentally, in the above-described SfM, matching between feature points may be incorrect, and, in this case, a pair of feature points that do not match is extracted, decreasing the accuracy of the three-dimensional shape that is restored. Thus, Patent Document 1 and Patent Document 2 disclose a system for correcting the positions of corresponding feature points.

30

[0006] Specifically, in the system disclosed in Patent Document 1, first, two images of the same target object are disposed such that epipolar lines of the images are parallel with a scanning line,

and corresponding feature points are extracted from the images for each scanning line. Next, the system disclosed in Patent Document 1 accepts, from the user, input of a line segment for designating corresponding locations on each image.

[0007] Also, the system disclosed in Patent Document 1 determines whether or not each of the
5 corresponding feature points on a scanning line coincides with the intersection between this scanning line and the input line segment, and if the feature point does not coincide with the intersection, the position of the feature point is corrected to the position of the intersection. The system disclosed in Patent Document 1 then executes reconstruction of the three-dimensional shape using feature points whose positions have been corrected.

10 [0008] In addition, the system disclosed in Patent Document 2 first extracts combinations of corresponding feature points from a pair of images in which the same target object is captured. Next, when corresponding line segments or points are designated in each image of the pair, the system disclosed in Patent Document 2 calculates a numerical matrix based on the geometric relation between the line segments or the points.

15 [0009] Next, in one image in the pair of images, a feature point is indicated by p and an epipolar line on which the feature point p is present is indicated by l , and, in the other image, a corresponding feature point is indicated by p' and an epipolar line on which the feature point p' is present is indicated by l' . In this case, the system disclosed in Patent Document 2 calculates an epipolar line l' corresponding to the epipolar line l using the calculated numerical matrix, and
20 further calculates a distance d between the epipolar line l' and the feature point p' .

[0010] If the distance d is larger than or equal to a threshold value, the system disclosed in Patent Document 2 determines that the combination of the feature point p and the feature point p' is geometrically inconsistent. The system disclosed in Patent Document 2 then reconstructs the three-dimensional shape using calculated numerical matrices, while excluding combinations for
25 which the geometric relation has been determined as being inconsistent.

[0011] In this manner, with the system disclosed in Patent Document 1 or Patent Document 2, the problem of incorrect pairs of feature points being extracted is solved, and thus it is conceivable that a decrease in the accuracy of the three-dimensional shape that is restored is suppressed.

30 LIST OF PRIOR ART DOCUMENTS

PATENT DOCUMENT

[0012]

Patent Document 1: Japanese Patent No. 5311465

Patent Document 2: International Publication No. 2019/065784

DISCLOSURE OF THE INVENTION

PROBLEMS TO BE SOLVED BY THE INVENTION

[0013] However, in the system disclosed in Patent Document 1, the position of only a feature point on a scanning line that runs along epipolar lines of two images is corrected, and thus an incorrect pair of feature points that are not on the same scanning line are not excluded.

[0014] On the other hand, in the system disclosed in Patent Document 2, a geometric relation between designated line segments or points is obtained, and furthermore, incorrect pairs of feature points are excluded based on the distance between an epipolar line obtained based on this geometric relation and a feature point, and thus the problem with Patent Document 1 is solved. However, in the system disclosed in Patent Document 2, there is a problem in that, when the distance between an epipolar line obtained based on a geometric relation and an incorrect feature point is short, incorrect pairs of feature points cannot be excluded.

[0015] As described above, with the systems disclosed in Patent Document 1 and Patent Document 2, incorrect pairs of feature points are not insufficiently excluded, and the accuracy of the three-dimensional shape that is restored may decrease.

[0016] An example object of the invention is to provide an image processing apparatus, an image processing method, and a computer-readable recording medium that solve the aforementioned problem, and enable extraction of an incorrect combination of feature points to be suppressed when combinations of corresponding feature points are extracted from a plurality of images.

MEANS FOR SOLVING THE PROBLEMS

[0017] In order to achieve the above-described example object, an image processing apparatus according to an example aspect of the invention is an apparatus for constructing a three-dimensional shape of a target object from a plurality of images of the target object, the apparatus including:

- a feature point extraction unit configured to extract a combination of corresponding feature points from each of the plurality of images;

- a matrix calculation unit configured to, when two or more line segments or two or more points corresponding to each other are designated in at least two images out of the plurality of images, specify a geometric relation between the two or more line segments corresponding to each other or a geometric relation between the two or more points corresponding to each other, and calculate a numerical matrix that represents the specified geometric relation;

- an unnecessary-feature-point specifying unit configured to specify a combination of

feature points between which the geometric relation is inconsistent from among the extracted combinations of feature points, using the numerical matrix;

a three-dimensional coordinate calculation unit configured to calculate, for each combination of feature points excluding the specified combination of feature points, a camera matrix for each feature point constituting the combination, and further calculate three-dimensional coordinates of the target object that correspond to the feature points of the combination using the calculated camera matrices;

a suitability determination unit configured to calculate two-dimensional coordinates obtained by applying, for each combination of the feature points whose three-dimensional coordinates were calculated, a camera matrix corresponding to the image from which one feature point in the combination was extracted to the three-dimensional coordinates of the combination and thereby projecting the three-dimensional coordinates of the combination onto the image from which the one feature point was extracted, and determine whether or not the one feature point is appropriate based on a result of comparison between the calculated two-dimensional coordinates and two-dimensional coordinates of the one feature point; and

a three-dimensional shape constructing unit configured to construct a three-dimensional shape of the target object using the three-dimensional coordinates of the target object that correspond to feature points determined as being appropriate.

[0018] In addition, in order to achieve the above-described example object, an image processing method according to an example aspect of the invention is a method for constructing a three-dimensional shape of a target object from a plurality of images of the target object, the method including:

(a) a step of extracting a combination of corresponding feature points from each of the plurality of images;

(b) a step of, when two or more line segments or two or more points corresponding to each other are designated in at least two images out of the plurality of images, specifying a geometric relation between the two or more line segments corresponding to each other or a geometric relation between the two or more points corresponding to each other, and calculating a numerical matrix that represents the specified geometric relation;

(c) a step of specifying a combination of feature points between which the geometric relation is inconsistent from among the extracted combinations of feature points, using the numerical matrix;

(d) a step of calculating, for each combination of feature points excluding the specified combination of feature points, a camera matrix for each feature point constituting the combination,

and further calculating three-dimensional coordinates of the target object that correspond to the feature points of the combination using the calculated camera matrices;

(e) a step of calculating two-dimensional coordinates obtained by applying, for each combination of the feature points whose three-dimensional coordinates were calculated, a camera matrix corresponding to the image from which one feature point in the combination was extracted to the three-dimensional coordinates of the combination and thereby projecting the three-dimensional coordinates of the combination onto the image from which the one feature point was extracted, and determining whether or not the one feature point is appropriate based on a result of comparison between the calculated two-dimensional coordinates and two-dimensional coordinates of the one feature point; and

(f) a step of constructing a three-dimensional shape of the target object using the three-dimensional coordinates of the target object that correspond to feature points determined as being appropriate.

[0019] Furthermore, in order to achieve the above-described example object, a computer-readable recording medium according to an example aspect of the invention is a computer-readable recording medium that includes a program recorded thereon for constructing a three-dimensional shape of a target object from a plurality of images of the target object using a computer, the program including instructions that cause the computer to carry out:

(a) a step of extracting a combination of corresponding feature points from each of the plurality of images;

(b) a step of, when two or more line segments or two or more points corresponding to each other are designated in at least two images out of the plurality of images, specifying a geometric relation between the two or more line segments corresponding to each other or a geometric relation between the two or more points corresponding to each other, and calculating a numerical matrix that represents the specified geometric relation;

(c) a step of specifying a combination of feature points between which the geometric relation is inconsistent from among the extracted combinations of feature points, using the numerical matrix;

(d) a step of calculating, for each combination of feature points excluding the specified combination of feature points, a camera matrix for each feature point constituting the combination, and further calculating three-dimensional coordinates of the target object that correspond to the feature points of the combination using the calculated camera matrices;

(e) a step of calculating two-dimensional coordinates obtained by applying, for each combination of the feature points whose three-dimensional coordinates were calculated, a camera

matrix corresponding to the image from which one feature point in the combination was extracted to the three-dimensional coordinates of the combination and thereby projecting the three-dimensional coordinates of the combination onto the image from which the one feature point was extracted, and determining whether or not the one feature point is appropriate based on a result of comparison between the calculated two-dimensional coordinates and two-dimensional coordinates of the one feature point; and

(f) a step of constructing a three-dimensional shape of the target object using the three-dimensional coordinates of the target object that correspond to feature points determined as being appropriate.

ADVANTAGEOUS EFFECTS OF THE INVENTION

[0020] As described above, according to the invention, when combinations of corresponding feature points are extracted from a plurality of images, extraction of incorrect combinations of feature points can be inhibited.

BRIEF DESCRIPTION OF THE DRAWINGS

[0021] FIG. 1 is a block diagram schematically showing an image processing apparatus according to an example embodiment of the invention.

FIG. 2 is a block diagram showing a specific configuration of an image processing apparatus according to the example embodiment of the invention.

FIG. 3 is a diagram showing an example of a plurality of images to be processed in the example embodiment.

FIG. 4 is a diagram illustrating processing that is performed by a matrix calculation unit according to the example embodiment of the invention.

FIG. 5 is a diagram illustrating processing that is performed by a shape constructing unit according to the example embodiment of the invention.

FIG. 6 is a diagram showing an example of a pair of images from which combinations of feature points have been extracted.

FIG. 7 is a diagram showing an example of three-dimensional coordinates and rotation matrices of cameras of an initial pair of images, which are obtained from camera matrices.

FIG. 8 is a diagram showing an example of an image newly selected after the initial pair of images have been selected and combinations of feature points extracted from the newly selected image.

FIG. 9 is a diagram illustrating processing for re-projecting three-dimensional coordinates

of a feature point in a target object onto a two-dimensional image.

FIG. 10 is a flowchart showing operations of the image processing apparatus according to the example embodiment of the invention.

FIG. 11 is a block diagram showing an example of a computer that realizes the image processing apparatus according to the example embodiment of the invention.

MODE FOR CARRYING OUT THE INVENTION

Embodiment

[0022] An image processing apparatus according to an example embodiment of the invention will be described below with reference to FIGS. 1 to 11.

Apparatus Configuration

[0023] First, a schematic configuration of an image processing apparatus according to the example embodiment of the invention will be described with reference to FIG. 1. FIG. 1 is a block diagram schematically showing an image processing apparatus according to the example embodiment of the invention.

[0024] An image processing apparatus 10 according to the example embodiment shown in FIG. 1 is an apparatus for constructing a three-dimensional shape of a target object from a plurality of images of the target object. As shown in FIG. 1, the image processing apparatus 10 includes a feature point extraction unit 11, a matrix calculation unit 12, an unnecessary-feature-point specifying unit 13, a three-dimensional coordinate calculation unit 14, a suitability determination unit 15, and a three-dimensional shape constructing unit 16.

[0025] The feature point extraction unit 11 extracts combinations of corresponding feature points from a plurality of images. When two or more line segments or two or more points corresponding to each other are designated in at least two images out of the plurality of images, the matrix calculation unit 12 specifies the geometric relation between the two or more line segments corresponding to each other or the geometric relation between the two or more points corresponding to each other. The matrix calculation unit 12 also calculates a numerical matrix that represents the specified geometric relation.

[0026] The unnecessary-feature-point specifying unit 13 specifies a combination of feature points between which the geometric relation is inconsistent from among the extracted combinations of feature points, using the calculated numerical matrix. The three-dimensional coordinate calculation unit 14 calculates camera matrices of the feature points of each of the combinations excluding the specified combination of feature points. In addition, the three-

dimensional coordinate calculation unit 14 calculates three-dimensional coordinates of the target object that correspond to the feature points of the combinations using the calculated camera matrices.

[0027] The suitability determination unit 15 applies, for each combination of the feature points whose three-dimensional coordinates were calculated, a camera matrix corresponding to an image from which one feature point in the combination was extracted, to the three-dimensional coordinates of this combination. By applying this camera matrix, the suitability determination unit 15 calculates two-dimensional coordinates that are obtained by projecting the three-dimensional coordinates of this combination onto the image from which the above one feature point was extracted. Furthermore, the suitability determination unit 15 compares the calculated two-dimensional coordinates with the two-dimensional coordinates of the one feature point (two-dimensional coordinates in the image from which the one feature point was extracted), and determines whether or not the above one feature point is appropriate based on the comparison result.

[0028] The three-dimensional shape constructing unit 16 constructs a three-dimensional shape of the target object using the three-dimensional coordinates of the target object that correspond to feature points of combinations determined as being appropriate by the suitability determination unit 15.

[0029] As described above, in the present example embodiment, first, a combination of feature points between which the geometric relation obtained based on designated line segments or points is inconsistent is excluded, then, obtained three-dimensional coordinates of a target object are projected onto a two-dimensional image, and determination is performed as to whether or not each feature point is appropriate. Accordingly, in the present example embodiment, incorrect feature points are excluded twice, namely before and after calculation of three-dimensional coordinates.

Thus, according to the present example embodiment, when combinations of corresponding feature points are extracted from a plurality of images, it is possible to inhibit extraction of incorrect combinations of feature points. As a result, an accurate three-dimensional shape is constructed.

[0030] Next, the configuration of the image processing apparatus 10 according to the example embodiment will be described in more detail with reference to FIGS. 2 to 9. FIG. 2 is a block diagram showing a specific configuration of the image processing apparatus according to the example embodiment of the invention.

[0031] As shown in FIG. 2, in the present example embodiment, in addition to the feature point extraction unit 11, the matrix calculation unit 12, the unnecessary-feature-point specifying unit 13, the three-dimensional coordinate calculation unit 14, the suitability determination unit 15, and the

three-dimensional shape constructing unit 16, the image processing apparatus 10 also includes an image obtaining unit 17, a filtering unit 18, an input accepting unit 19, and a display unit 20. In FIG. 2, reference numeral 21 denotes a display device.

[0032] The image obtaining unit 17 obtains image data of a plurality of images that show a target for which a three-dimensional shape is to be constructed, from an external apparatus such as an image capturing apparatus, a terminal apparatus, or a storage device that holds the image data. FIG. 3 is a diagram showing an example of a plurality of images that are to be processed in the present example embodiment. In the example in FIG. 3, a pair of images are illustrated, but the number of images to be processed in the present example embodiment is not particularly limited.

[0033] In the present example embodiment, for example, the feature point extraction unit 11 calculates SIFT feature amounts or SURF feature amounts for each image, specifies feature points, and further extracts feature points that correspond to each other between the images as combinations of corresponding feature points. Note that, since two images are used in the example in FIG. 3, a combination of feature points is a pair of feature points. In FIG. 3, a circled portion indicates one feature point. In addition, hereinafter, two images from which combinations of corresponding feature points are extracted is referred to as a “pair of images”.

[0034] The filtering unit 18 calculates the geometric relation between feature points for each of the combinations of feature points, specifies an incorrect combination of feature points based on the calculation result, and excludes the specified combination of feature points. Similarly to conventional SfM, the filtering unit 18 performs filtering processing using Robust Estimation. In the present example embodiment, in addition to filtering that is performed by the unnecessary-feature-point specifying unit 13, which will be described later, the filtering unit 18 also performs filtering, and thus an incorrect combination of feature points is more reliably excluded.

[0035] When line segments or points corresponding to each other between the images that are to be processed are designated, the input accepting unit 19 accepts input of the designated line segments or points. In addition, when input of line segments corresponding to each other between the images is accepted, the input accepting unit 19 notifies the matrix calculation unit 12 of information regarding the line segments (the coordinates of the starting points and the ending points thereof). Furthermore, when input of points corresponding to each other between the images is accepted, the input accepting unit 19 notifies the matrix calculation unit 12 of information regarding the points (coordinates). Note that line segments or points may be designated by the user of the image processing apparatus 10 or another calculator.

[0036] When notified of the information regarding the line segments or the points between the images by the input accepting unit 19, the matrix calculation unit 12 specifies the geometric

relation between the line segments or the points based on the notified information, calculates a numerical matrix representing the specified geometric relation, and defines this numerical matrix as an absolute numerical matrix. When no numerical matrix can be calculated based on the information regarding the line segments or the points between the image notified by the input
 5 accepting unit 19, or information regarding line segments or points is not present depending on an image, the matrix calculation unit 12 can also calculate a numerical matrix based on combinations of feature points that were filtered after extraction was performed by the feature point extraction unit 11, and that have a smaller error. Note that the numerical matrix that is calculated in this case is a numerical matrix similar to those conventionally obtained, and is not an absolute
 10 numerical matrix.

[0037] Here, processing that is performed by the matrix calculation unit 12 will be described in more detail with reference to FIG. 4. FIG. 4 is a diagram illustrating processing that is performed by the matrix calculation unit according to the example embodiment of the invention. In FIG. 4, E indicates an epipolar plane, O_i indicates the central position of a camera that captured one image,
 15 and O'_i indicates the central position of a camera that captured the other image. In addition, in FIG. 4, parallelograms shown on the right and left indicate the frames of the images.

[0038] In addition, in the example in FIG. 4, a line segment L is designated in the one image, a line segment L' is designated in the other image, and the line segments correspond to each other. In this case, an epipolar line l_n that intersects with the starting point of the line segment L and an
 20 epipolar line l_{n+m} that intersects with the ending point are defined as absolute epipolar lines, and also, an epipolar line l'_n that intersects with the starting point of the line segment L' and an epipolar line l'_{n+m} that intersects with the ending point are defined as absolute epipolar lines.

[0039] Furthermore, an intersection x_i between the line segment L and the absolute epipolar line l_n and an intersection x'_i between the line segment L' and the absolute epipolar line l'_n are defined
 25 as a combination of absolute feature points. In addition, an intersection x_j between the line segment L and the absolute epipolar line l_{n+m} and an intersection x'_j between the line segment L' and the absolute epipolar line l'_{n+m} are also defined as a combination of absolute feature points.

[0040] Note that, as shown in FIG. 4, suitable epipolar lines l_{n+1} and l_{n+2} may also be set between the epipolar line l_n and the epipolar line l_{n+m} . In this case, similarly, suitable epipolar lines l'_{n+1}
 30 and l'_{n+2} are also set between the epipolar line l'_n and the epipolar line l'_{n+m} . In addition, in this case, intersections between the line segment L and the newly set epipolar lines l_{n+1} and l_{n+2} and intersections between the line segment L' and the newly set epipolar lines l'_{n+1} and l'_{n+2} are respectively regarded as combinations of absolute feature points. Note that the distance between epipolar lines is set to any value.

[0041] Therefore, in the example in FIG. 4, the matrix calculation unit 12 obtains combinations of absolute feature points as the geometric relation between the line segments, and calculates a fundamental matrix (see reference document) as an absolute numerical matrix using the obtained combinations of feature points, based on the relational equation in Formula 1 below. Note that, in Formula 1 below, x indicates a point on a two-dimensional plane obtained by projecting a point X in a three-dimensional space onto one image, and x' indicates a point on a two-dimensional plane obtained by projecting the point X in the three-dimensional space onto the other image. T indicates a transposed matrix. Also, F indicates a fundamental matrix. In addition, in the present example embodiment, a numerical matrix is not limited to a fundamental matrix, and any matrix that can represent a geometric relation may be used.

[0042]

$$x'^T F x = 0 \quad \dots (1)$$

[0043] Reference document: Richard Hartley and Andrew Zisserman, "Multiple View Geometry in Computer Vision Second Edition", Cambridge University Press, March 2004.

[0044] In the present example embodiment, the unnecessary-feature-point specifying unit 13 first specifies a combination of feature points between which the geometric relation is inconsistent, from combinations of feature points extracted by the feature point extraction unit 11 (excluding those excluded by the filtering unit 18), using the numerical matrix (fundamental matrix) calculated by the matrix calculation unit 12.

[0045] In addition, the unnecessary-feature-point specifying unit 13 can specify a combination that includes a feature point that overlaps the designated line segments or points, from the combinations of feature points extracted by the feature point extraction unit 11 (excluding those excluded by the filtering unit 18), and can also specify a combination of feature points that include a specified feature point, as feature points between which the geometric relation is inconsistent. This is because, when a feature point corresponding to a feature point that overlaps only one line segment or one point is not present on the other line segment or the other point, it is highly likely that a combination that includes this feature point is incorrect.

[0046] Here, processing that is performed by the unnecessary-feature-point specifying unit 13 will be described in more detail with reference to FIG. 5. FIG. 5 is a diagram illustrating processing that is performed by a shape constructing unit according to the example embodiment of the invention. In FIG. 5, constituent elements that are same as those in FIG. 4 are given the same reference signs as those in FIG. 4. In addition, in FIG. 5, determination is performed as to whether or not a combination of a feature point p_i and a feature point p'_i is incorrect.

[0047] In the example in FIG. 5, the unnecessary-feature-point specifying unit 13 calculates an epipolar line l'_1 corresponding to an epipolar line l_1 on which the feature point p_i is present, using the relationship shown in Formula 2 below related to the epipolar line.

[0048]

$$5 \quad l' = F \times \dots (2)$$

[0049] The unnecessary-feature-point specifying unit 13 then determines whether or not the distance d between the calculated epipolar line l'_1 and the feature point p'_i is larger than or equal to a threshold value, and if the distance d is larger than or equal to the threshold value, the unnecessary-feature-point specifying unit 13 determines that the combination of the feature point p_i and the feature point p'_i is incorrect. In this case, the unnecessary-feature-point specifying unit 13 specifies the combination of the feature point p_i and the feature point p'_i , as a combination of feature points between which the geometric relation is inconsistent.

[0050] The three-dimensional coordinate calculation unit 14 calculates three-dimensional coordinates for constructing a three-dimensional shape, using combinations of feature points that have not been specified by the unnecessary-feature-point specifying unit 13. In addition, at this time, the three-dimensional coordinate calculation unit 14 can also use corresponding points or points on corresponding line segments accepted by the input accepting unit 19, as a combination of feature points.

[0051] Here, processing for calculating three-dimensional coordinates that is performed by the three-dimensional coordinate calculation unit 14 will be described in more detail with reference to FIGS. 6 to 8. FIG. 6 is a diagram showing an example of a pair of images from which combinations of feature points have been extracted. FIG. 7 is a diagram showing an example of three-dimensional coordinates of cameras and rotation matrices of an initial pair of images, which are obtained based on camera matrices. FIG. 8 is a diagram showing an example of an image newly selected after the initial pair of images have been selected and combinations of feature points extracted from the newly selected image.

[0052] As shown in FIG. 6, the three-dimensional coordinate calculation unit 14 first selects an image 31 and an image 32 as a pair of images (initial pair of images). In this case, feature points (m_1 to m_5) extracted from the image 31 correspond to feature points (m'_1 to m'_5) extracted from the image 32. The feature points m_1 and m'_1 , m_2 and m'_2 , m_3 and m'_3 , m_4 and m'_4 , and m_5 and m'_5 respectively indicate combinations of feature points (hereinafter, also referred to as “pairs of feature points”). In addition, in the example in FIG. 6, the image 31 was captured by a camera 41, and the image 32 was captured by a camera 42. In FIG. 6, feature points M (M_1 to M_5)

indicate three-dimensional coordinates of a target object corresponding to respective feature points.

[0053] Subsequently, the three-dimensional coordinate calculation unit 14 calculates a camera matrix P of the camera 41 that captured the image 31 and a camera matrix P' of the camera 42 that captured the image 32 using the pairs of feature points (m_1 to m_5 and m'_1 to m'_5) extracted from the initial pair of images. In addition, when the position of the camera 41 is set as an origin, the camera matrix P and the camera matrix P' can be expressed by Formula 3 and Formula 4 below.

[0054]

$$P = [I \mid 0] \quad \dots (3)$$

[0055]

$$P' = [R \mid t] \quad \dots (4)$$

[0056] In Formula 3 above, I indicates a rotation matrix of the camera 41. Since the position of the camera 41 serves as the origin as shown in FIG. 7, then $I = (1, 1, 1)$. In addition, in Formula 4 above, R indicates a rotation matrix of the camera 42 ($R = (R_x, R_y, R_z)$). t indicates a translational matrix as described above, and corresponds to the three-dimensional coordinates of the position of the camera 42 ($t = (t_x, t_y, t_z)$).

[0057] Therefore, in this case, R and t can be calculated by an inverse operation based on the camera matrix P and the camera matrix P'. Specifically, the three-dimensional coordinate calculation unit 14 calculates R and t using the coordinates of feature points by solving equations indicated by Formulas 5 to 7 below. In Formulas 5 to 7, $\hat{m}$ indicates coordinates on an image A obtained by normalizing m (m_1 to m_5). Similarly, $\hat{m}'$ indicates coordinates on an image B obtained by normalizing m' (m'_1 to m'_5). E indicates an essential matrix, and K indicates a calibration matrix of the camera.

[0058]

$$E = [t]_M R = R [R^T t]_M \quad \dots (5)$$

[0059]

$$\hat{m}'^T E \hat{m} = 0 \quad \dots (6)$$

[0060]

$$\hat{m}_l = K^{-1} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad \dots (7)$$

[0061] In addition, the calibration matrix K can be obtained based on Formulas 8 and 9 below. Note that “ c_x, c_y ” indicates the central coordinates of the camera.

[0062]

$$K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad \dots (8)$$

5 [0063]

f: focal distance (pixel unit)

$$= \frac{\text{long side of screen (number of pixels)} \times \text{focal distance (fx or fy: screen longitudinal direction) (mm)}}{\text{sensor size (mm) of long side of screen}}$$

... (9)

10 [0064] Next, the three-dimensional coordinate calculation unit 14 calculates three-dimensional coordinates M (M_1 to M_5) of the extracted feature points using the three-dimensional coordinates and the rotation matrices of the positions of cameras. Specifically, the three-dimensional coordinate calculation unit 14 calculates the three-dimensional coordinates M by solving Formula 10 below. In addition, a matrix A in Formula 10 is expressed by Formula 11. In Formula 11,
15 p^{iT} indicates a row in the camera matrix P, and p'^{iT} indicates a row in the camera matrix P'.

[0065]

$$AM = 0 \quad \dots (10)$$

[0066]

$$A = \begin{bmatrix} xp^{3T} & - & p^{1T} \\ yp^{3T} & - & p^{2T} \\ x'p'^{3T} & - & p'^{1T} \\ y'p'^{3T} & - & p'^{2T} \end{bmatrix} \quad \dots (11)$$

20 [0067] Next, as shown in FIG. 8, the three-dimensional coordinate calculation unit 14 newly selects one image 33, from images from which feature points have been extracted other than the initial pair of images, and sets the newly selected image 33 and one image of the initial pair of images as a new pair of images. The image 33 is captured by a camera 43.

[0068] The three-dimensional coordinate calculation unit 14 then specifies feature points (m''_1 to m''_3) in the image 33 that correspond to the feature points in the image 32, and sets the feature points in the image 32 and the feature points in the image 33 as pairs of feature points. The three-dimensional coordinate calculation unit 14 then calculates a camera matrix P_n of the camera 43

25

that captured the image 33. The camera matrix P_n can be expressed by Formula 12 below.

[0069]

$$P_n = [R_n \mid t_n] \quad \dots (12)$$

[0070] Specifically, the three-dimensional coordinate calculation unit 14 calculates R_n and t_n in the camera matrix P_n of the camera 43 using the specified feature points in the image 33, by solving the equation indicated by Formula 13 below.

[0071]

$$M_i = t + d_i R \hat{m}_i \quad (i=1,2,3) \quad \dots (13)$$

[0072] In Formula 13, M_i indicates the three-dimensional coordinates of a feature point in the newly selected image 33 that corresponds to a feature point in the image 32. $\hat{m}_i$ indicates normalized coordinates of the feature point in the newly selected image 33. d_i indicates the distance between the camera 43 that captured the image 33 and the $\hat{m}_i$ as indicated by Formula 14 below.

[0073]

$$d_i = \|t - \hat{m}_i\| \quad \dots (14)$$

[0074] Next, the three-dimensional coordinate calculation unit 14 calculates the three-dimensional coordinates M_i of the specified feature points (m''_1 to m''_3) in the image 33 using calculated R_n and t_n of the camera matrix P_n of the camera 43. Also in this case, the three-dimensional coordinate calculation unit 14 calculates the three-dimensional coordinates M (M_1 to M_3) of the feature points by solving Formula 10 above. By performing the above-described processing, the three-dimensional coordinate calculation unit 14 can calculate the three-dimensional coordinates of the target object.

[0075] In the present example embodiment, for example, the suitability determination unit 15 re-projects a feature point in each pair of feature points onto the original two-dimensional image using the three-dimensional coordinates obtained from the feature point and the camera matrix corresponding to the image from which the feature point was extracted, and compares the two-dimensional coordinates of the position of the projected feature point with the two-dimensional coordinates of the position at the time of extraction. In the present example embodiment, the suitability determination unit 15 then calculates the difference between the former two-dimensional coordinates and the later two-dimensional coordinates, as a comparison result, and, if the calculated difference is smaller than or equal to a threshold value, the suitability determination

unit 15 determines that the target feature point is appropriate. On the other hand, if the calculated difference exceeds the threshold value, the suitability determination unit 15 determines that the target feature point is not appropriate.

[0076] In addition, in the present example embodiment, the suitability determination unit 15 can select all or some of feature points in the combinations of feature points whose three-dimensional coordinates have been calculated, and determine whether or not each of the selected feature points is appropriate. If a combination of feature points includes three feature points, for example, the suitability determination unit 15 may determine whether or not each of the three feature points is appropriate, or may select only two feature points from the three feature points, and determine whether or not each of the two selected feature points is appropriate.

[0077] Here, determination processing that is performed by the suitability determination unit 15 will be described in detail with reference to FIG. 9. FIG. 9 is a diagram illustrating processing for re-projecting three-dimensional coordinates of feature points of a target object onto a two-dimensional image. In addition, FIG. 9 shows an example of re-projecting one of the feature points in the image 32 shown in FIG. 6. Furthermore, in FIG. 9, three-dimensional coordinates corresponding to the feature point, that is, three-dimensional coordinates on a world coordinate system are indicated by (X_w, Y_w, Z_w) , and the two-dimensional coordinates of the position of the feature point at the time of extraction are indicated by (x_f, y_f) . In addition, the coordinates of the feature point on a camera coordinate system are indicated by (X_c, Y_c, Z_c) .

[0078] First, the suitability determination unit 15 calculates coordinates (X_c, Y_c, Z_c) of the feature point on the camera coordinate system based on the camera matrix $P' (= [R|t])$ and the three-dimensional coordinates (X_w, Y_w, Z_w) of the feature point using Formula 15 below.

[0079]

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = R \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + t \quad \dots (15)$$

[0080] Next, the suitability determination unit 15 normalizes the coordinates (X_c, Y_c, Z_c) of the feature point on the camera coordinate system using Formula 16 below.

[0081]

$$\begin{bmatrix} X'_c \\ Y'_c \end{bmatrix} = \begin{bmatrix} X_c/Z_c \\ Y_c/Z_c \end{bmatrix} \quad \dots (16)$$

[0082] Subsequently, the suitability determination unit 15 applies the coordinates normalized using Formula 16 above and the internal parameters of the camera (in the example in FIG. 9, a focal length f of the camera 42 and the coordinates (c_x, c_y) of the image central position) to Formula

17 below, and calculates two-dimensional coordinates (x_p, y_p) of the feature point when re-projected onto the image 32.

[0083]

$$\begin{bmatrix} x_p \\ y_p \end{bmatrix} = f \begin{bmatrix} X' \\ Y' \end{bmatrix} + \begin{bmatrix} c_x \\ c_y \end{bmatrix} \quad \dots (17)$$

5 [0084] Next, the suitability determination unit 15 calculates a difference d between the two-dimensional coordinates (x_p, y_p) of the re-projected feature point calculated based on Formula 17 above and the two-dimensional coordinates (x_f, y_f) of the position of the feature point at the time of extraction, using Formula 18 below.

[0085]

$$10 \quad d = \sqrt{(x_p - x_f)^2 + (y_p - y_f)^2} \quad \dots (18)$$

[0086] Thereafter, when the calculated difference d is smaller than or equal to a threshold value G , the suitability determination unit 15 determines that the target feature point is appropriate, and, if the calculated difference d exceeds the threshold value G , the suitability determination unit 15 determines that the target feature point is not appropriate.

15 [0087] In the present example embodiment, the suitability determination unit 15 excludes the three-dimensional coordinates corresponding to feature points determined as being not appropriate and combinations that include all of the feature points corresponding to these three-dimensional coordinates. In addition, in the present example embodiment, a configuration can also be adopted in which the suitability determination unit 15 excludes only feature points determined as being not appropriate, and does not exclude three-dimensional coordinates corresponding to the feature points determined as being not appropriate and feature points other than the feature points determined as being not appropriate, from among the combinations of feature points corresponding to these three-dimensional coordinates. In this case, when suitability determination for all of the feature points has been completed, the suitability determination unit 15 needs to exclude three-dimensional coordinates for which the number of feature points corresponding to each of all of the three-dimensional coordinates is fewer than two.

20

25

[0088] The three-dimensional shape constructing unit 16 recalculates three-dimensional coordinates using the remaining feature points, constructs point group data of the target object, and sets the point group data as a three-dimensional shape of the target object.

30 [0089] In addition, the display unit 20 displays the three-dimensional shape constructed by the three-dimensional shape constructing unit 16 on the screen of a display device 21. Specifically, the display unit 20 creates image data for displaying the constructed three-dimensional shape on

the two-dimensional screen, and outputs the created image data to the display device 21.

[0090] In addition, in the present example embodiment, the three-dimensional coordinate calculation unit 14 can also calculate provisional three-dimensional coordinates of a target object using combinations of feature points extracted by the feature point extraction unit 11, for example, using combinations of feature points subjected to only filtering performed by the filtering unit 18. In this case, the three-dimensional shape constructing unit 16 further constructs a provisional three-dimensional shape using the calculated provisional three-dimensional coordinates. The display unit 20 then displays the constructed provisional three-dimensional shape on the screen of the display device 21. In this case, the user can designate line segments or points corresponding to each other on the provisional three-dimensional shape displayed on the screen.

Apparatus Operation

[0091] Next, operations of the image processing apparatus 10 according to the example embodiment of the invention will be described with reference to FIG. 10. FIG. 10 is a flowchart showing operations of the image processing apparatus according to the example embodiment of the invention. In the following description, FIGS. 1 to 9 will be referred to as appropriate. In addition, in the present example embodiment, an image processing method is carried out by operating the image processing apparatus 10. Therefore, a description of the image processing method according to the present example embodiment is replaced with the following description of the operations of the image processing apparatus 10.

[0092] As shown in FIG. 10, first, the image obtaining unit 17 obtains, from an external apparatus, image data of a plurality of images of a target for which a three-dimensional shape is to be constructed (step A1).

[0093] Next, the feature point extraction unit 11 specifies feature points in an image corresponding to each piece of the image data obtained in step A1, and further extracts feature points corresponding to each other between images, as combinations of corresponding feature points (step A2).

[0094] Next, the filtering unit 18 calculates a geometric relation between feature points, for each of the combinations of feature points extracted in step A2, specifies an incorrect combination of feature points based on the calculation result, and excludes the specified combination of feature points (step A3).

[0095] Next, when filtering processing in step A3 is executed, the three-dimensional coordinate calculation unit 14 calculates provisional three-dimensional coordinates using combinations of filtered feature points, and furthermore, the three-dimensional shape constructing unit 16

constructs a provisional three-dimensional shape using the calculated provisional three-dimensional coordinates. Furthermore, the display unit 20 displays the constructed provisional three-dimensional shape on the screen of the display device 21 (step A4).

[0096] When step A4 is executed, and the provisional three-dimensional shape is displayed on the screen of the display device 21, the user designates line segments or points corresponding to each other between the images, via an input device (not illustrated in FIG. 2). Accordingly, the input accepting unit 16 accepts input of the designated line segments or points (step A5). The input accepting unit 16 also notifies the matrix calculation unit 12 of information regarding the designated line segments or points.

[0097] Next, when step A5 is executed, the matrix calculation unit 12 specifies a geometric relation between the line segments or points based on the information regarding the line segments or the points, and calculates a numerical matrix that represents the specified geometric relation (step A6).

[0098] Next, the unnecessary-feature-point specifying unit 13 specifies a combination of feature points between which the geometric relation is inconsistent from among the combinations of feature points that have been subjected to step A3, using the numerical matrix (fundamental matrix) calculated in step A6 (step A7).

[0099] Next, the three-dimensional coordinate calculation unit 14 calculates three-dimensional coordinates for constructing a three-dimensional shape, using combinations of feature points that were not specified in step A7 (step A8).

[0100] Next, the suitability determination unit 15 applies, for each of the combinations of feature points whose three-dimensional coordinates were calculated in step A8, a camera matrix corresponding to an image from which one feature point in the combination was extracted to the three-dimensional coordinates of this combination. The suitability determination unit 15 then calculates two-dimensional coordinates that are obtained by projecting the three-dimensional coordinates of this combination onto the image from which the above-described one feature point was extracted, by applying this camera matrix. Furthermore, the suitability determination unit 15 compares the calculated two-dimensional coordinates with the two-dimensional coordinates of the one feature point (two-dimensional coordinates on the image from which the feature point was extracted), and determines whether or not the above-described one feature point is appropriate based on the comparison result (step A9).

[0101] Next, in step A6, the three-dimensional shape constructing unit 16 constructs a three-dimensional shape of the target object using three-dimensional coordinates of the target object that correspond to feature points of combinations determined as being appropriate (step A10). The

display unit 20 then displays the three-dimensional shape constructed in step A10 on the screen of the display device 21 (step A11).

[0102] [Effects of example embodiment]

5 As described above, the image processing apparatus 10 according to the present example embodiment can exclude incorrect feature points, using the filtering unit 18, the unnecessary-feature-point specifying unit 13, and the suitability determination unit 15. Thus, when extracting combinations of corresponding feature points from a plurality of images, the image processing apparatus 10 can inhibit extraction of incorrect combinations of feature points. As a result,
10 according to the present example embodiment, it is possible to accurately construct a three-dimensional shape of a target object.

[0103] [Program]

The program according to the present example embodiment may be any program for
15 causing a computer to execute steps A1 to A11 shown in FIG. 10. By installing this program in a computer and executing the program, it is possible to realize the image processing apparatus 10 and the image processing method according to the present example embodiment. In this case, the processor of the computer functions as the feature point extraction unit 11, the matrix calculation unit 12, the unnecessary-feature-point specifying unit 13, the three-dimensional
20 coordinate calculation unit 14, the suitability determination unit 15, the three-dimensional shape constructing unit 16, the image obtaining unit 17, the filtering unit 18, and the input accepting unit 19, and performs processing.

[0104] In addition, the program according to the present example embodiment may be executed by a computer system constituted by a plurality of computers. In this case, for example, each
25 computer may function as one of the feature point extraction unit 11, the matrix calculation unit 12, the unnecessary-feature-point specifying unit 13, the three-dimensional coordinate calculation unit 14, the suitability determination unit 15, the three-dimensional shape constructing unit 16, the image obtaining unit 17, the filtering unit 18, and the input accepting unit 19.

[0105] Here, a computer that realizes the image processing apparatus by executing the program
30 according to the present example embodiment will be described with reference to FIG. 11. FIG. 11 is a block diagram showing an example of a computer that realizes the image processing apparatus according to the present example embodiment.

[0106] As shown in FIG. 11, a computer 110 includes a CPU 111, a main memory 112, a storage device 113, an input interface 114, a display controller 115, a data reader/writer 116, and a

communication interface 117. These units are connected via a bus 121 so as to be able to perform data communication with each other. Note that the computer 110 may include a GPU (Graphics Processing Unit) or a FPGA (Field-Programmable Gate Array) in addition to the CPU 111 or instead of the CPU 111.

5 [0107] The CPU 111 loads programs (codes) according to the present example embodiment stored in the storage device 113 to the main memory 112, and executes the programs in a predetermined order to perform various kinds of calculations. The main memory 112 is typically a volatile storage device such as a DRAM (Dynamic Random Access Memory). Also, the program according to the present example embodiment is provided in the state of being stored in
10 a computer-readable recording medium 120. Note that programs according to the present example embodiment may be distributed on the Internet that is connected via the communication interface 117.

[0108] Specific examples of the storage device 113 include a hard disk drive, and a semiconductor storage device such as a flash memory. The input interface 114 mediates data
15 transmission between the CPU 111 and an input device 118 such as a keyboard or a mouse. The display controller 115 is connected to a display device 119, and controls the display of the display device 119.

[0109] The data reader/writer 116 mediates data transmission between the CPU 111 and the recording medium 120, reads out programs from the recording medium 120, and writes the results
20 of processing performed by the computer 110 to the recording medium 120. The communication interface 117 mediates data transmission between the CPU 111 and another computer.

[0110] Specific examples of the recording medium 120 include general-purpose semiconductor storage devices such as a CF (Compact Flash (registered trademark)) and a SD (Secure Digital), a magnetic recording medium such as a flexible disk, and an optical recording medium such as a
25 CD-ROM (Compact Disk Read Only Memory).

[0111] Note that the image processing apparatus 10 according to the present example embodiment can also be realized by using hardware (for example, circuits) corresponding to the units, in place of a computer that has programs installed therein. Furthermore, a configuration may also be adopted in which a portion of the image processing apparatus 10 is realized by
30 programs, and the remaining portion is realized by hardware.

[0112] One or all of the above-described example embodiments can be expressed as, but are not limited to, Supplementary Note 1 to Supplementary Note 18 described below.

[0113] (Supplementary Note 1)

An image processing apparatus for constructing a three-dimensional shape of a target

object from a plurality of images of the target object, the apparatus including:

a feature point extraction unit configured to extract a combination of corresponding feature points from each of the plurality of images;

5 a matrix calculation unit configured to, when two or more line segments or two or more points corresponding to each other are designated in at least two images out of the plurality of images, specify a geometric relation between the two or more line segments corresponding to each other or a geometric relation between the two or more points corresponding to each other, and calculate a numerical matrix that represents the specified geometric relation;

10 an unnecessary-feature-point specifying unit configured to specify a combination of feature points between which the geometric relation is inconsistent from among the extracted combinations of feature points, using the numerical matrix;

a three-dimensional coordinate calculation unit configured to calculate, for each combination of feature points excluding the specified combination of feature points, a camera matrix for each feature point constituting the combination, and further calculate three-dimensional
15 coordinates of the target object that correspond to the feature points of the combination using the calculated camera matrices;

a suitability determination unit configured to calculate two-dimensional coordinates obtained by applying, for each combination of the feature points whose three-dimensional coordinates were calculated, a camera matrix corresponding to the image from which one feature
20 point in the combination was extracted to the three-dimensional coordinates of the combination and thereby projecting the three-dimensional coordinates of the combination onto the image from which the one feature point was extracted, and determine whether or not the one feature point is appropriate based on a result of comparison between the calculated two-dimensional coordinates and two-dimensional coordinates of the one feature point; and

25 a three-dimensional shape constructing unit configured to construct a three-dimensional shape of the target object using the three-dimensional coordinates of the target object that correspond to feature points determined as being appropriate.

[0114] (Supplementary Note 2)

The image processing apparatus according to Supplementary Note 1,

30 wherein the suitability determination unit calculates a difference between the calculated two-dimensional coordinates and the two-dimensional coordinates of the one feature point, as the comparison result, and determines that the one feature point is appropriate if the calculated difference is smaller than or equal to a threshold value, and that the one feature point is not appropriate if the calculated difference exceeds the threshold value.

[0115] (Supplementary Note 3)

The image processing apparatus according to Supplementary Note 1 or 2,

wherein the suitability determination unit selects, from each combination of feature points whose three-dimensional coordinates were calculated, all or one of the feature points constituting the combination, and determines, for each selected feature point, whether or not the selected feature point is appropriate.

[0116] (Supplementary Note 4)

The image processing apparatus according to any one of Supplementary Notes 1 to 3,

wherein the unnecessary-feature-point specifying unit further specifies a feature point that overlaps two or more designated line segments or points corresponding to each other, from among the feature points in the extracted combinations of feature points.

[0117] (Supplementary Note 5)

The image processing apparatus according to any one of Supplementary Notes 1 to 4, further including:

a display unit configured to display, on a screen, the three-dimensional shape constructed by the three-dimensional shape constructing unit.

[0118] (Supplementary Note 6)

The image processing apparatus according to Supplementary Note 5,

wherein the three-dimensional coordinate calculation unit calculates provisional three-dimensional coordinates of the target object using the combination of feature points extracted by the feature point extraction unit,

the three-dimensional shape constructing unit constructs a provisional three-dimensional shape of the target object using the calculated provisional three-dimensional coordinates, and

the display unit further displays the provisional three-dimensional shape on the screen.

[0119] (Supplementary Note 7)

An image processing method for constructing a three-dimensional shape of a target object from a plurality of images of the target object, the method including:

(a) a step of extracting a combination of corresponding feature points from each of the plurality of images;

(b) a step of, when two or more line segments or two or more points corresponding to each other are designated in at least two images out of the plurality of images, specifying a geometric relation between the two or more line segments corresponding to each other or a geometric relation between the two or more points corresponding to each other, and calculating a numerical matrix that represents the specified geometric relation;

(c) a step of specifying a combination of feature points between which the geometric relation is inconsistent from among the extracted combinations of feature points, using the numerical matrix;

(d) a step of calculating, for each combination of feature points excluding the specified
5 combination of feature points, a camera matrix for each feature point constituting the combination, and further calculating three-dimensional coordinates of the target object that correspond to the feature points of the combination using the calculated camera matrices;

(e) a step of calculating two-dimensional coordinates obtained by applying, for each
10 combination of the feature points whose three-dimensional coordinates were calculated, a camera matrix corresponding to the image from which one feature point in the combination was extracted to the three-dimensional coordinates of the combination and thereby projecting the three-dimensional coordinates of the combination onto the image from which the one feature point was extracted, and determining whether or not the one feature point is appropriate based on a result of comparison between the calculated two-dimensional coordinates and two-dimensional coordinates
15 of the one feature point; and

(f) a step of constructing a three-dimensional shape of the target object using the three-dimensional coordinates of the target object that correspond to feature points determined as being appropriate.

[0120] (Supplementary Note 8)

20 The image processing method according to Supplementary Note 7,
wherein, in the (e) step, a difference between the calculated two-dimensional coordinates and the two-dimensional coordinates of the one feature point is calculated as the comparison result, and it is determined that the one feature point is appropriate if the calculated difference is smaller than or equal to a threshold value, and that the one feature point is not appropriate if the calculated
25 difference exceeds the threshold value.

[0121] (Supplementary Note 9)

The image processing method according to Supplementary Note 7 or 8,
wherein, in the (e) step, from each combination of feature points whose three-dimensional coordinates were calculated, all or one of the feature points constituting the combination is selected,
30 and determination is performed as to, for each selected feature point, whether or not the selected feature point is appropriate.

[0122] (Supplementary Note 10)

The image processing method according to any one of Supplementary Notes 7 to 9,
wherein, in the (c) step, a feature point that overlaps two or more designated line segments

or points corresponding to each other is further specified from among the feature points in the extracted combinations of feature points.

[0123] (Supplementary Note 11)

The image processing method according to any one of Supplementary Notes 7 to 10,
5 further including:

(g) a step of displaying, on a screen, the three-dimensional shape constructed in the (f) step.

[0124] (Supplementary Note 12)

The image processing method according to Supplementary Note 11, further including:

10 (h) a step of calculating provisional three-dimensional coordinates of the target object using the combination of feature points extracted in the (a) step, constructing a provisional three-dimensional shape of the target object using the calculated provisional three-dimensional coordinates, and further displaying the provisional three-dimensional shape on the screen.

[0125] (Supplementary Note 13)

15 A computer-readable recording medium that includes a program recorded thereon for constructing a three-dimensional shape of a target object from a plurality of images of the target object using a computer, the program including instructions that cause the computer to carry out:

(a) a step of extracting a combination of corresponding feature points from each of the plurality of images;

20 (b) a step of, when two or more line segments or two or more points corresponding to each other are designated in at least two images out of the plurality of images, specifying a geometric relation between the two or more line segments corresponding to each other or a geometric relation between the two or more points corresponding to each other, and calculating a numerical matrix that represents the specified geometric relation;

25 (c) a step of specifying a combination of feature points between which the geometric relation is inconsistent from among the extracted combinations of feature points, using the numerical matrix;

(d) a step of calculating, for each combination of feature points excluding the specified combination of feature points, a camera matrix for each feature point constituting the combination,
30 and further calculating three-dimensional coordinates of the target object that correspond to the feature points of the combination using the calculated camera matrices;

(e) a step of calculating two-dimensional coordinates obtained by applying, for each combination of the feature points whose three-dimensional coordinates were calculated, a camera matrix corresponding to the image from which one feature point in the combination was extracted

to the three-dimensional coordinates of the combination and thereby projecting the three-dimensional coordinates of the combination onto the image from which the one feature point was extracted, and determining whether or not the one feature point is appropriate based on a result of comparison between the calculated two-dimensional coordinates and two-dimensional coordinates of the one feature point; and

(f) a step of constructing a three-dimensional shape of the target object using the three-dimensional coordinates of the target object that correspond to feature points determined as being appropriate.

[0126] (Supplementary Note 14)

The computer-readable recording medium according to Supplementary Note 13, wherein, in the (e) step, a difference between the calculated two-dimensional coordinates and the two-dimensional coordinates of the one feature point is calculated as the comparison result, and it is determined that the one feature point is appropriate if the calculated difference is smaller than or equal to a threshold value, and that the one feature point is not appropriate if the calculated difference exceeds the threshold value.

[0127] (Supplementary Note 15)

The computer-readable recording medium according to Supplementary Note 13 or 14, wherein, in the (e) step, from each combination of feature points whose three-dimensional coordinates were calculated, all or one of the feature points constituting the combination is selected, and determination is performed as to, for each selected feature point, whether or not the selected feature point is appropriate.

[0128] (Supplementary Note 16)

The computer-readable recording medium according to any one of Supplementary Notes 13 to 15,

wherein, in the (c) step, a feature point that overlaps two or more designated line segments or points corresponding to each other is further specified from among the feature points in the extracted combinations of feature points.

[0129] (Supplementary Note 17)

The computer-readable recording medium according to any one of Supplementary Note 13 to 16, further causing a computer to carry out:

(g) a step of displaying, on a screen, the three-dimensional shape constructed in the (f) step.

[0130] (Supplementary Note 18)

The computer-readable recording medium according to Supplementary Note 17, further

causing a computer to carry out:

(h) a step of calculating provisional three-dimensional coordinates of the target object using the combination of feature points extracted in the (a) step, constructing a provisional three-dimensional shape of the target object using the calculated provisional three-dimensional coordinates, and further displaying the provisional three-dimensional shape on the screen.

[0131] Although the invention has been described above with reference to an example embodiment, the invention is not limited to the above example embodiment. Various modifications understandable to a person skilled in the art can be made to configurations and details of the invention, within the scope of the invention.

[0132] This application is based upon and claims the benefit of priority from Japanese Patent Application No. 2019-189792, filed October 16, 2019, the disclosure of which is incorporated herein in its entirety by reference.

INDUSTRIAL APPLICABILITY

[0133] As described above, according to the invention, when combinations of corresponding feature points are extracted from a plurality of images, it is possible to inhibit extraction of incorrect combinations of feature points. The invention is useful in techniques for constructing a three-dimensional shape from a plurality of images, such as SfM.

LIST OF REFERENCE SIGNS

[0134]	10	Image processing apparatus
	11	Feature point extraction unit
	12	Matrix calculation unit
	13	Unnecessary-feature-point specifying unit
	14	Three-dimensional coordinate calculation unit
	15	Suitability determination unit
	16	Three-dimensional shape constructing unit
	17	Image obtaining unit
	18	Filtering unit
	19	Input accepting unit
	20	Display unit
	21	Display device
	110	Computer

	111	CPU
	112	Main memory
	113	Storage device
	114	Input interface
5	115	Display controller
	116	Data reader/writer
	117	Communication interface
	118	Input device
	119	Display device
10	120	Recording medium
	121	Bus

1. An image processing apparatus for constructing a three-dimensional shape of a target object from a plurality of images of the target object, the apparatus comprising:

5 a feature point extraction means for extracting a combination of corresponding feature points from each of the plurality of images;

a matrix calculation means for, when two or more line segments or two or more points corresponding to each other are designated in at least two images out of the plurality of images, specifying a geometric relation between the two or more line segments corresponding to each other or a geometric relation between the two or more points corresponding to each other, and calculating
10 a numerical matrix that represents the specified geometric relation;

an unnecessary-feature-point specifying means for specifying a combination of feature points between which the geometric relation is inconsistent from among the extracted combinations of feature points, using the numerical matrix;

15 a three-dimensional coordinate calculation means for calculating, for each combination of feature points excluding the specified combination of feature points, a camera matrix for each feature point constituting the combination, and further calculating three-dimensional coordinates of the target object that correspond to the feature points of the combination using the calculated camera matrices;

20 a suitability determination means for calculating two-dimensional coordinates obtained by applying, for each combination of the feature points whose three-dimensional coordinates were calculated, a camera matrix corresponding to the image from which one feature point in the combination was extracted to the three-dimensional coordinates of the combination and thereby projecting the three-dimensional coordinates of the combination onto the image from which the
25 one feature point was extracted, and determining whether or not the one feature point is appropriate based on a result of comparison between the calculated two-dimensional coordinates and two-dimensional coordinates of the one feature point; and

30 a three-dimensional shape constructing means for constructing a three-dimensional shape of the target object using the three-dimensional coordinates of the target object that correspond to feature points determined as being appropriate.

2. The image processing apparatus according to claim 1,

wherein the suitability determination means calculates a difference between the calculated two-dimensional coordinates and the two-dimensional coordinates of the one feature point, as the

comparison result, and determines that the one feature point is appropriate if the calculated difference is smaller than or equal to a threshold value, and that the one feature point is not appropriate if the calculated difference exceeds the threshold value.

- 5 3. The image processing apparatus according to claim 1 or 2,
 wherein the suitability determination means selects, from each combination of feature points whose three-dimensional coordinates were calculated, all or one of the feature points constituting the combination, and determines, for each selected feature point, whether or not the selected feature point is appropriate.

- 10 4. The image processing apparatus according to any one of claims 1 to 3,
 wherein the unnecessary-feature-point specifying means further specifies a feature point that overlaps two or more designated line segments or points corresponding to each other, from among the feature points in the extracted combinations of feature points.

- 15 5. The image processing apparatus according to any one of claims 1 to 4, further comprising:
 a display means for displaying, on a screen, the three-dimensional shape constructed by the three-dimensional shape constructing means.

- 20 6. The image processing apparatus according to claim 5,
 wherein the three-dimensional coordinate calculation means calculates provisional three-dimensional coordinates of the target object using the combination of feature points extracted by the feature point extraction means,
 the three-dimensional shape constructing means constructs a provisional three-
25 dimensional shape of the target object using the calculated provisional three-dimensional coordinates, and
 the display means further displays the provisional three-dimensional shape on the screen.

7. An image processing method for constructing a three-dimensional shape of a target object
30 from a plurality of images of the target object, the method comprising:
 extracting a combination of corresponding feature points from each of the plurality of images;
 when two or more line segments or two or more points corresponding to each other are designated in at least two images out of the plurality of images, specifying a geometric relation

between the two or more line segments corresponding to each other or a geometric relation between the two or more points corresponding to each other, and calculating a numerical matrix that represents the specified geometric relation;

specifying a combination of feature points between which the geometric relation is inconsistent from among the extracted combinations of feature points, using the numerical matrix;

calculating, for each combination of feature points excluding the specified combination of feature points, a camera matrix for each feature point constituting the combination, and further calculating three-dimensional coordinates of the target object that correspond to the feature points of the combination using the calculated camera matrices;

calculating two-dimensional coordinates obtained by applying, for each combination of the feature points whose three-dimensional coordinates were calculated, a camera matrix corresponding to the image from which one feature point in the combination was extracted to the three-dimensional coordinates of the combination and thereby projecting the three-dimensional coordinates of the combination onto the image from which the one feature point was extracted, and determining whether or not the one feature point is appropriate based on a result of comparison between the calculated two-dimensional coordinates and two-dimensional coordinates of the one feature point; and

constructing a three-dimensional shape of the target object using the three-dimensional coordinates of the target object that correspond to feature points determined as being appropriate.

8. The image processing method according to claim 7,

wherein, in the determining of whether or not the one feature point is appropriate, a difference between the calculated two-dimensional coordinates and the two-dimensional coordinates of the one feature point is calculated as the comparison result, and it is determined that the one feature point is appropriate if the calculated difference is smaller than or equal to a threshold value, and that the one feature point is not appropriate if the calculated difference exceeds the threshold value.

9. The image processing method according to claim 7 or 8,

wherein, in the determining of whether or not the one feature point is appropriate, from each combination of feature points whose three-dimensional coordinates were calculated, all or one of the feature points constituting the combination is selected, and determination is performed as to, for each selected feature point, whether or not the selected feature point is appropriate.

10. The image processing method according to any one of claims 7 to 9,
 wherein, in the specifying of a combination of feature points, a feature point that overlaps
 two or more designated line segments or points corresponding to each other is further specified
 from among the feature points in the extracted combinations of feature points.

5 11. The image processing method according to any one of claims 7 to 10, further comprising:
 displaying, on a screen, the three-dimensional shape constructed in the constructing the
 three-dimensional shape.

10 12. The image processing method according to claim 11, further comprising:
 calculating provisional three-dimensional coordinates of the target object using the
 combination of feature points extracted in the extracting the combinations of corresponding feature
 points, constructing a provisional three-dimensional shape of the target object using the calculated
 provisional three-dimensional coordinates, and further displaying the provisional three-
 15 dimensional shape on the screen.

13. A computer-readable recording medium that includes a program recorded thereon for
 constructing a three-dimensional shape of a target object from a plurality of images of the target
 object using a computer, the program including instructions that cause the computer to carry out:
 20 extracting a combination of corresponding feature points from each of the plurality of
 images;

when two or more line segments or two or more points corresponding to each other are
 designated in at least two images out of the plurality of images, specifying a geometric relation
 between the two or more line segments corresponding to each other or a geometric relation
 25 between the two or more points corresponding to each other, and calculating a numerical matrix
 that represents the specified geometric relation;

specifying a combination of feature points between which the geometric relation is
 inconsistent from among the extracted combinations of feature points, using the numerical matrix;
 calculating, for each combination of feature points excluding the specified combination
 30 of feature points, a camera matrix for each feature point constituting the combination, and further
 calculating three-dimensional coordinates of the target object that correspond to the feature points
 of the combination using the calculated camera matrices;

calculating two-dimensional coordinates obtained by applying, for each combination of
 the feature points whose three-dimensional coordinates were calculated, a camera matrix

corresponding to the image from which one feature point in the combination was extracted to the three-dimensional coordinates of the combination and thereby projecting the three-dimensional coordinates of the combination onto the image from which the one feature point was extracted, and determining whether or not the one feature point is appropriate based on a result of comparison
5 between the calculated two-dimensional coordinates and two-dimensional coordinates of the one feature point; and

constructing a three-dimensional shape of the target object using the three-dimensional coordinates of the target object that correspond to feature points determined as being appropriate.

10 14. The computer-readable recording medium according to claim 13,
wherein, in the determining of whether or not the one feature point is appropriate, a difference between the calculated two-dimensional coordinates and the two-dimensional coordinates of the one feature point is calculated as a comparison result, and it is determined that the one feature point is appropriate if the calculated difference is smaller than or equal to a
15 threshold value, and that the one feature point is not appropriate if the calculated difference exceeds the threshold value.

15. The computer-readable recording medium according to claim 13 or 14,
wherein, in the determining of whether or not the one feature point is appropriate, from
20 each combination of feature points whose three-dimensional coordinates were calculated, all or one of the feature points constituting the combination is selected, and determination is performed as to, for each selected feature point, whether or not the selected feature point is appropriate.

16. The computer-readable recording medium according to any one of claims 13 to 15,
25 wherein, in the specifying of a combination of feature points, a feature point that overlaps two or more designated line segments or points corresponding to each other is further specified from among the feature points in the extracted combinations of feature points.

17. The computer-readable recording medium according to any one of claims 13 to 16, the
30 program further including instructions that cause the computer to carry out:
displaying, on a screen, the three-dimensional shape constructed in the constructing the three-dimensional shape.

18. The computer-readable recording medium according to claim 17, the program further

including instructions that cause the computer to carry out:

- calculating provisional three-dimensional coordinates of the target object using the combination of feature points extracted in the extracting of the combination of corresponding feature points, constructing a provisional three-dimensional shape of the target object using the
- 5 calculated provisional three-dimensional coordinates, and further displaying the provisional three-dimensional shape on the screen.