

DESCRIPTION

[TITLE]

IMAGE ANALYSIS APPARATUS, IMAGE ANALYSIS METHOD, AND COMPUTER PROGRAM

[TECHNICAL FIELD]

[0001]

The present disclosure relates to an image analysis apparatus, an image analysis method, and a computer program.

[BACKGROUND ART]

[0002]

The diagnostic system described in Patent Literature 1 has a function to automatically determine and detect deteriorated areas by performing diagnostic processing by a computer based on continuous aerial images taken by a flying body of past and present subjects. The subjects are buildings or infrastructure facilities, etc. A predetermined area on the surface of the subject is a diagnostic target area. Based on the user's input operation, the diagnostic target area of the subject is set as a basic setting.

[0003]

The diagnostic system described in Patent Literature 1 has an SFM processing unit. The SFM processing unit applies SFM processing to a plurality of input images to restore the 3D structure and output a three-dimensional model.

[Citation List]

Patent Literature

[0004]

Patent Literature 1: Japanese Patent Application Laid-Open Publication No. 2019-070631

[SUMMARY OF INVENTION]

[Technical Problem]

[0005]

Generally, there may be cases where a diagnostic target area is searched from a three-dimensional model of a subject, and the diagnostic target area is diagnosed.

[0006]

However, it is often difficult to search for a target region such as a diagnostic target region from a three-dimensional model and to identify the position of the target region on the

three-dimensional model.

[0007]

Therefore, the present disclosure has been made in view of the foregoing and has its object of providing an image analysis apparatus, an image analysis method, and a computer program that make it possible to identify a position of a target region on a three-dimensional model by a simple process.

[Solution to Problem]

[0008]

According to the present disclosure, there is provided an image analysis apparatus comprising: a model generation unit that generates a three-dimensional model of a subject based on a plurality of two-dimensional images generated by imaging the subject from a plurality of different imaging positions; an image placement unit that places, among the plurality of two-dimensional images, a target two-dimensional image in a virtual space based on imaging three-dimensional coordinates and imaging posture information, the imaging three-dimensional coordinates being three-dimensional coordinates in the virtual space where the three-dimensional model is placed and indicating an imaging position when the target two-dimensional image is generated, and the imaging posture information indicating a posture of a camera that performs imaging; and a target coordinate calculation unit that calculates second target three-dimensional coordinates indicating a position where a straight line passing through the imaging three-dimensional coordinates and first target three-dimensional coordinates contacts the three-dimensional model, wherein the first target three-dimensional coordinates indicate a position in the virtual space of a target image included in the target two-dimensional image.

[0009]

Also, according to the present disclosure, there is provided an image analysis method comprising: generating a three-dimensional model of a subject based on a plurality of two-dimensional images generated by imaging the subject from a plurality of different imaging positions; acquiring imaging three-dimensional coordinates and imaging posture information, the imaging three-dimensional coordinates being three-dimensional coordinates in a virtual space where the three-dimensional model is placed and indicating an imaging position when a target two-dimensional image among the plurality of two-dimensional images is generated, and the imaging posture information indicating a posture of a camera that performs imaging; placing the target two-dimensional image in the virtual space based on the imaging three-dimensional coordinates and the imaging posture information; calculating first target three-dimensional coordinates, which are three-dimensional coordinates in the virtual space of a target image included in the target two-dimensional image; and calculating second target three-dimensional

coordinates, which are three-dimensional coordinates of a position where a straight line passing through the imaging three-dimensional coordinates and the first target three-dimensional coordinates contacts the three-dimensional model.

[0010]

Furthermore, according to the present disclosure, there is provided a computer program causing a computer to perform: generating a three-dimensional model of a subject based on a plurality of two-dimensional images generated by imaging the subject from a plurality of different imaging positions; acquiring imaging three-dimensional coordinates and imaging posture information, the imaging three-dimensional coordinates being three-dimensional coordinates in a virtual space where the three-dimensional model is placed and indicating an imaging position when a target two-dimensional image among the plurality of two-dimensional images is generated, and the imaging posture information indicating a posture of a camera that performs imaging; placing the target two-dimensional image in the virtual space based on the imaging three-dimensional coordinates and the imaging posture information; calculating first target three-dimensional coordinates, which are three-dimensional coordinates in the virtual space of a target image included in the target two-dimensional image; and calculating second target three-dimensional coordinates, which are three-dimensional coordinates of a position where a straight line passing through the imaging three-dimensional coordinates and the first target three-dimensional coordinates contacts the three-dimensional model.

[Advantageous Effects of Invention]

[0011]

According to the present disclosure, it is possible to provide an image analysis apparatus, an image analysis method, and a computer program that make it possible to identify a position of a target region on a three-dimensional model by a simple process.

[BRIEF DESCRIPTION OF DRAWINGS]

[0012]

[FIG. 1]

FIG. 1 is a block diagram showing an example configuration of an image analysis system according to an embodiment of the present invention.

[FIG. 2]

FIG. 2 is a block diagram showing an example configuration of a server according to the embodiment.

[FIG. 3]

FIG. 3 is a perspective view schematically showing a camera, a two-dimensional

image, and a subject in real space according to the embodiment.

[FIG. 4]

FIG. 4 is a perspective view schematically showing an imaging position and a three-dimensional model of a subject in a virtual space according to the embodiment.

[FIG. 5]

FIG. 5 is a perspective view schematically showing an imaging position, a target two-dimensional image, and a three-dimensional model in a virtual space according to the embodiment.

[FIG. 6]

FIG. 6A is a side view schematically showing an imaging position represented in a virtual space according to the embodiment. FIG. 6B is a side view schematically showing a target two-dimensional image placed in the virtual space. FIG. 6C is a side view schematically showing a target image and a bounding box included in the target two-dimensional image placed in the virtual space, and a target region included in the three-dimensional model. FIG. 6D is a side view schematically showing a state in which a straight line passing through the imaging position and the target image contacts the three-dimensional model in the virtual space.

[FIG. 7]

FIG. 7 is a front view showing a target two-dimensional image placed in a virtual space according to the embodiment.

[FIG. 8]

FIG. 8 is a flowchart showing an image analysis method according to the embodiment.

[FIG. 9]

FIG. 9 is a flowchart showing a three-dimensional model generation process according to the embodiment.

[FIG. 10]

FIG. 10 is a flowchart showing an image analysis method according to a modification of the embodiment.

[DESCRIPTION OF EMBODIMENTS]

[0013]

The following is a detailed description of a preferred embodiment of the present disclosure with reference to the accompanying drawings. In this specification and the drawings, duplicated explanations will be omitted by applying the same reference numerals to configuration elements that have substantially the same functional configuration.

[0014]

FIG. 1 shows an example configuration of an image analysis system SYS according to

an embodiment of the present invention. As shown in FIG. 1, the image analysis system SYS includes a server 1. The server 1 corresponds to an example of an "image analysis apparatus" of the present disclosure.

[0015]

The server 1 generates a three-dimensional model of a subject to be placed in a virtual space based on a plurality of two-dimensional images generated by imaging the subject from a plurality of different imaging positions. In the virtual space, the server 1 calculates a position where a straight line passing through the imaging position and a target image included in a single two-dimensional image contacts the three-dimensional model (that is, a position of a target region on the three-dimensional model). As a result, the position of the target region on the three-dimensional model is identified. In the embodiment, instead of executing a process to search for the target region across the entire area of the three-dimensional model, the position of the target region on the three-dimensional model is calculated from a camera position and a single two-dimensional image. Therefore, it is possible to identify the position of the target region on the three-dimensional model by a simple process. Details on this will be described below.

[0016]

The image analysis system SYS also includes at least one terminal 2, at least one moving apparatus 3, or at least one imaging device 5.

[0017]

The server 1, the terminal 2, the moving apparatus 3, and the imaging device 5 are connected to a network NW. The network NW includes, for example, the Internet, a closed network, a public telephone network, a LAN (Local Area Network), and a short-range wireless network.

[0018]

The terminal 2 is, for example, a personal computer (e.g., a notebook PC, a desktop PC, or a tablet).

[0019]

The moving apparatus 3 is, for example, an unmanned moving apparatus or a manned moving apparatus. The unmanned moving apparatus is, for example, an unmanned aerial vehicle such as a drone, an unmanned ground machine, an unmanned underwater vehicle, or an unmanned surface vehicle. The unmanned ground machine is, for example, an unmanned ground vehicle or an unmanned ground machine modeled after a living organism (e.g., a snake-type unmanned ground machine). The manned moving apparatus is, for example, an aircraft, an automobile, a ship, or a submarine. The moving apparatus 3 includes a camera 4.

[0020]

The imaging device 5 includes a camera 6. The imaging device 5 is, for example, a portable terminal such as a smartphone. The imaging device 5 may be, for example, the camera 6 itself.

[0021]

The cameras 4 and 6 image a subject and generate video data that represents a moving image including an image of the subject. The moving image is a sequence of consecutive two-dimensional images. The cameras 4 and 6 may also generate a plurality of still image data items each representing a still image including an image of the subject. A still image is a two-dimensional image.

[0022]

The video data is hereinafter referred to as "video data 511" (FIG. 2). The still image data items are referred to as "still image data set 512" (FIG. 2).

[0023]

Hereinafter, when it is not necessary to distinguish between the cameras 4 and 6, they are collectively referred to as "camera CM."

[0024]

The subject is not particularly limited as long as it can be imaged by the camera CM, and the size, shape, pattern, and color of the subject are not particularly limited. For example, the subject is one or more movable or immovable properties. The subject is, for example, one or more objects. Typically, the subject is one or more stationary objects. The stationary object is, for example, an artificial object or a natural object. The artificial object is, for example, a structure, a machine, an electronic device, or a work of authorship. The structure is, for example, a building or an infrastructure facility. The building is, for example, a high-rise building or a residential house. The infrastructure facility is a facility for building social infrastructure. For example, the infrastructure facility is a road, a bridge, a road traffic facility, a power generation facility, a power distribution facility, a water treatment facility, or a gas distribution facility. The machine is, for example, an automobile, a work vehicle, a train, an aircraft, a ship, a submarine, or a robot. The natural object is, for example, a tree, a forest, a ground surface, a cliff, a coast, or a river.

[0025]

Furthermore, the subject includes one or more targets to be identified for their position on the three-dimensional model. The targets are not particularly limited and can be set arbitrarily, such as the entire region included in the subject, a part of the region, the entire object included in the subject, and a part of the object, etc.

[0026]

The terminal 2 acquires the video data 511 or the still image data set 512 generated by

the camera CM. For example, the terminal 2 receives the video data 511 or the still image data set 512 transmitted from the camera CM via the network NW. For example, the terminal 2 acquires the video data 511 or the still image data set 512 from a removable media. The terminal 2 transmits the video data 511 or the still image data set 512 to the server 1 via the network NW. The moving apparatus 3 and the imaging device 5 may also transmit the video data 511 or the still image data set 512 to the server 1 via the network NW.

[0027]

FIG. 2 is a block diagram showing an example configuration of the server 1 in FIG. 1. As shown in FIG. 2, the server 1 includes an arithmetic unit 10, a communication unit 40, and a storage unit 50. The server 1 may also include an input unit 20 and a display unit 30.

[0028]

The input unit 20 is an input device for inputting various types of information to the arithmetic unit 10. For example, the input unit 20 is a keyboard and a pointing device, or a touch panel.

[0029]

The display unit 30 displays various types of information. The display unit 30 is, for example, a liquid crystal display or an organic electroluminescent display.

[0030]

The communication unit 40 is connected to the network NW. The communication unit 40 communicates with external devices connected to the network NW. The external devices are, for example, the terminal 2, the moving apparatus 3, and the imaging device 5. The communication unit 40 is a communication device that communicates according to a predetermined communication protocol, including, for example, a network interface controller. The predetermined communication protocol is, for example, a protocol compliant with Ethernet (registered trademark) and the Internet Protocol Suite.

[0031]

The communication unit 40 receives, for each subject, the video data 511 or the still image data set 512 generated by imaging the subject from the terminal 2, the moving apparatus 3, and the imaging device 5 via the network NW.

[0032]

The storage unit 50 includes a storage device that stores data and computer programs. The storage unit 50 includes a main storage device such as a semiconductor memory or the like, and an auxiliary storage device such as a semiconductor memory, a hard disk drive, or the like. The storage unit 50 may also include removable media, such as an optical disk. The storage unit 50 may be, for example, a non-transitory computer readable storage medium.

[0033]

The storage unit 50 stores a first database 51, a second database 52, and a trained model 53.

[0034]

The first database 51 includes the video data 511 and the still image data set 512 received by the communication unit 40. In the first database 51, the video data 511 and the still image data set 512 are associated with attribute information (hereinafter referred to as "attribute information AT"). The attribute information AT includes, for example, user information, imaging conditions, and subject information. The user information includes, for example, identification information of a user of the terminal 2, the moving apparatus 3, or the imaging device 5. The imaging conditions include, for example, date and time of imaging and calibration information of the camera CM. The calibration information includes, for example, information about lens distortion of the camera CM, a focal length, and information about a position of a principal point. The lens distortion information includes, for example, information about distortion. The imaging conditions may include information about an imaging location. The imaging location is indicated, for example, by position coordinates of the camera CM obtained by GPS (Global Positioning System) or GNSS (Global Navigation Satellite System). The subject information includes, for example, subject identification information. The subject information may include coordinates of a ground control point (GCP: Ground Control Point).

[0035]

The second database 52 includes a plurality of data sets 520. Each of the plurality of data sets 520 includes a plurality of two-dimensional images 150, a three-dimensional model 301 (point cloud data), a plurality of imaging position information items 521, and a plurality of imaging posture information items 522. In the second database 52, the two-dimensional images 150, the three-dimensional model 301, the imaging position information items 521, and the imaging posture information items 522 are associated with each other and with the attribute information AT. These details are described later.

[0036]

The trained model 53 is a computer program. The trained model 53 includes, for example, a neural network. Details of the trained model 53 are described later.

[0037]

The arithmetic unit 10 performs various calculations. The arithmetic unit 10 includes a processor such as a CPU (Central Processing Unit), a GPU (Graphics Processing Unit), or the like.

[0038]

Specifically, the arithmetic unit 10 includes a model generation unit 11, a distortion correction unit 12, an image placement unit 13, a target detection unit 14, a target coordinate

calculation unit 15, and a display control unit 16. For example, the arithmetic unit 10 functions as the model generation unit 11, the distortion correction unit 12, the image placement unit 13, the target detection unit 14, the target coordinate calculation unit 15, and the display control unit 16 by executing a computer program stored in the storage unit 50.

[0039]

Next, with reference to FIGS. 2 to 6, the process performed by the arithmetic unit 10 will be described. FIG. 3 is a perspective view schematically showing a camera CM, a two-dimensional image 150, and a subject 300 in real space RS. As shown in FIG. 3, the subject 300 includes a target 310. The two-dimensional image 150 includes a subject image 200 representing the subject 300. The subject image 200 includes a target image 210 representing the target 310 depending on an imaging position 100. In the embodiment, the single camera CM is moved while imaging the subject 300 to generate the two-dimensional images 150. FIG. 4 is a perspective view schematically showing an imaging position 101 and a three-dimensional model 301 of the subject 300 in a virtual space VS. In FIG. 4, the imaging position 101 is shown by a black circle. For ease of understanding, a region A of the three-dimensional model 301 is shown enlarged.

[0040]

As shown in FIGS. 2 to 4, the model generation unit 11 of the arithmetic unit 10 generates the three-dimensional model 301 of the subject 300 based on the plurality of two-dimensional images 150 generated by imaging the subject 300 from the plurality of different imaging positions 100. The three-dimensional model 301 includes a target region 311 representing the target 310 of the subject 300. The three-dimensional model 301 is placed in the virtual space VS. The three-dimensional model 301 represents a three-dimensional shape of the subject 300. The three-dimensional model 301 is formed by point cloud data. The point cloud data is data representing a point cloud 321. The point cloud 321 is an aggregation of multiple points 322. The point cloud data includes three-dimensional coordinates of each point 322. The point cloud data may further include one or more of color information (e.g., RGB values) of each point 322, normal vector information of each point, and reflection intensity information.

[0041]

As an example, the model generation unit 11 generates the three-dimensional model 301 by performing SfM (Structure from Motion) processing. SfM processing refers to a process that generates the three-dimensional model 301 by utilizing the principle of triangulation based on the plurality of two-dimensional images 150 generated by imaging the subject 300 having a plurality of feature points from the plurality of imaging positions 100. SfM processing preferably includes bundle adjustment. The bundle adjustment refers to a process of minimizing reprojection error.

[0042]

Specifically, the model generation unit 11 first acquires the plurality of two-dimensional images 150 including the subject image 200 from the video data 511 or the still image data set 512 in the first database 51. Next, the model generation unit 11 extracts feature points of the subject 300 from each of the two-dimensional images 150 and associates the feature points across the two-dimensional images 150 (feature point matching). The associated feature points are called tie points. In the feature point matching, the model generation unit 11 employs, for example, SIFT (Scale Invariant Feature Transformation), SURF (Speeded-Up Robust Features), or FAST (Features from Accelerated Segment Test).

[0043]

Next, the model generation unit 11 estimates the imaging position 100 (the position of the camera CM) and a posture (orientation) of the camera CM when each of the two-dimensional images 150 is generated based on the associated feature points, and calculates three-dimensional coordinates of a three-dimensional point corresponding to a feature point based on the estimation results. In this case, the model generation unit 11 optimizes the estimation results of the imaging position 100 and the posture of the camera CM, as well as the three-dimensional coordinates of the three-dimensional point, by performing the bundle adjustment. The model generation unit 11 calculates three-dimensional coordinates of a plurality of three-dimensional points corresponding to a plurality of feature points. The three-dimensional coordinates in this case indicate coordinates in a three-dimensional coordinate system CS set in the virtual space VS. The three-dimensional coordinate system CS is defined by mutually orthogonal X, Y, and Z axes. The three-dimensional coordinate system CS may be a coordinate system with an origin at a predetermined position in the virtual space VS, or the three-dimensional coordinate system CS may be a coordinate system to which geospatial coordinates (coordinates on ground) or actual dimension information is assigned.

[0044]

The plurality of three-dimensional points constitute the point cloud 321. In other words, each three-dimensional point is each point 322 constituting the point cloud 321. In the above manner, the model generation unit 11 generates the point cloud data. In this case, for example, the point cloud data represents a sparse point cloud 321. Since the three-dimensional model 301 is constituted by the point cloud data, the generation of the point cloud data is synonymous with the generation of the three-dimensional model 301.

[0045]

The model generation unit 11 may add one or more of color information, luminance information, normal vector information, and reflection intensity information to each point 322 constituting the point cloud 321. In addition to SfM processing, the model generation unit 11

may also perform MVS (Multi View Stereo) processing. MVS processing refers to a process that calculates depth and normal for each pixel in each of the two-dimensional images 150 by multi-view stereo measurement, integrates them, and generates a dense point cloud 321 of the subject 300. Therefore, the model generation unit 11 can generate point cloud data representing the dense point cloud 321 by performing MVS processing. In this case, the three-dimensional model 301 is formed by the dense point cloud 321.

[0046]

As described above, in the process of generating the three-dimensional model 301, the model generation unit 11 estimates the imaging position 100 and the posture of the camera CM for each of the two-dimensional images 150.

[0047]

The imaging position 100 indicates the position of the camera CM when the subject 300 is imaged and the two-dimensional image 150 is generated. The estimated result of the imaging position 100 (hereinafter referred to as "imaging position information item 521") is represented by three-dimensional coordinates in the three-dimensional coordinate system CS (hereinafter referred to as "imaging three-dimensional coordinates"). As shown in FIG. 4, in the embodiment, the imaging position 100 in the real space RS is shown as "imaging position 101" in the virtual space VS. The imaging position information item 521 is estimated for each of the two-dimensional images 150.

[0048]

The posture of the camera CM indicates the posture (orientation) of the camera CM when the subject 300 is imaged and the two-dimensional image 150 is generated. The estimated result of the posture of the camera CM (hereinafter referred to as "imaging posture information item 522") is represented by a rotation angle of the camera CM around the X axis, a rotation angle of the camera CM around the Y axis, and a rotation angle of the camera CM around the Z axis in the three-dimensional coordinate system CS. In FIG. 4, for simplification of the drawing, the imaging posture information items 522 are shown by arrows. The imaging posture information item 522 is estimated for each of the two-dimensional images 150. The model generation unit 11 may also perform a process to reduce the effect of gimbal lock when estimating the posture of the camera CM. The gimbal lock is a phenomenon in which three degrees of freedom of rotation become two degrees of freedom when a three-dimensional posture is expressed in terms of Euler angles.

[0049]

It is also preferable for the model generation unit 11 to assign a surface to the three-dimensional model 301 based on the point cloud data. The process of assigning the surface is, as an example, a process of converting the point cloud data constituting the three-dimensional

model 301 into mesh data. In this case, for example, the model generation unit 11 generates a plurality of polygonal (e.g., triangular) facets (e.g., polygons) formed by connecting points 322 of the point cloud data constituting the three-dimensional model 301, and represents the three-dimensional model 301 by the plurality of polygonal facets. In this case, for example, the model generation unit 11 generates TIN (Triangulated Irregular Network) data based on the point cloud data, and represents the three-dimensional model 301 by the TIN data. As a result, the three-dimensional model 301 is represented by an aggregation of triangular facets.

[0050]

The process of assigning a surface to the three-dimensional model 301 is not particularly limited and may be the following process. For example, the model generation unit 11 assigns a facet to each point 322 constituting the point cloud data. In this case, the facet is, for example, a circular facet centered on each point 322 constituting the point cloud data. Alternatively, for example, the model generation unit 11 assigns a surface to the three-dimensional model 301 formed by the point cloud data by performing an octree algorithm. In this case, specifically, the model generation unit 11 represents the three-dimensional model 301 by multiple cubes by recursively dividing the three-dimensional model 301 into eight cubes (octants). Alternatively, for example, the model generation unit 11 assigns a surface to the three-dimensional model 301 by converting the point cloud data to surface data.

[0051]

The model generation unit 11 may also add material information based on the two-dimensional image 150 to the surface assigned to the three-dimensional model 301. The material information is information that includes a color and a pattern of the subject 300. For example, the model generation unit 11 may perform a process of mapping a texture to each facet (e.g., each polygon) constituting the mesh data of the three-dimensional model 301. In this case, the mesh data is, for example, TIN data.

[0052]

The storage unit 50 stores the three-dimensional model 301 (the point cloud data forming the three-dimensional model 301) in the second database 52. Furthermore, the storage unit 50 stores the plurality of two-dimensional images 150, the plurality of imaging position information items 521, and the plurality of imaging posture information items 522 in the second database 52 in association with the three-dimensional model 301. In the second database 52, the two-dimensional images 150 are stored in association with the imaging position information items 521 and the imaging posture information items 522.

[0053]

As long as the imaging position information item 521 and the imaging posture information item 522 can be obtained, the method of generating the three-dimensional model

301 is not particularly limited. For example, the model generation unit 11 may generate the three-dimensional model 301 by 3D Gaussian Splatting.

[0054]

Continuing to refer to FIG. 2, the distortion correction unit 12 of the arithmetic unit 10 will be described. The distortion correction unit 12 acquires from the second database 52 a two-dimensional image 150 to be processed (hereinafter referred to as "target two-dimensional image 150A") from among the plurality of two-dimensional images 150 used in generating the three-dimensional model 301.

[0055]

The distortion correction unit 12 also acquires the calibration information of the camera CM from the second database 52. The distortion correction unit 12 corrects distortion of the target two-dimensional image 150A based on the calibration information. The distortion in this case is, for example, lens distortion. The lens distortion is, for example, a type of distortion.

[0056]

Next, with reference to FIGS. 2, 5, and 6A, the image placement unit 13 of the arithmetic unit 10 will be described. FIG. 5 and FIGS. 6A to 6D render the camera CM for ease of understanding.

[0057]

FIG. 5 is a perspective view schematically showing the imaging position 101, a target two-dimensional image 151, and the three-dimensional model 301 in the virtual space VS. As shown in FIG. 5, in the embodiment, the target two-dimensional image 150A placed in the virtual space VS is described as "target two-dimensional image 151". Also, the subject image 200 and the target image 210 of the target two-dimensional image 150A placed in the virtual space VS are described as "subject image 201" and "target image 211," respectively.

[0058]

FIG. 6A is a side view schematically showing the imaging position 101 represented in the virtual space VS. FIG. 6B is a side view schematically showing the target two-dimensional image 151 placed in the virtual space VS.

[0059]

First, as shown in FIG. 2, the image placement unit 13 acquires the imaging position information item 521 and the imaging posture information item 522 associated with the target two-dimensional image 150A from the second database 52. The imaging position information item 521 indicates imaging three-dimensional coordinates (aX , aY , aZ). As shown in FIG. 5 and FIG. 6A, the imaging position 101 in the virtual space VS is indicated by the imaging three-dimensional coordinates (aX , aY , aZ).

[0060]

Next, as shown in FIGS. 5 and 6B, the image placement unit 13 places the target two-dimensional image 151, in which distortion has been corrected, in the virtual space VS based on the imaging position information item 521 and the imaging posture information item 522.

[0061]

Specifically, the image placement unit 13 places the target two-dimensional image 151 in the virtual space VS at a position separated by a predetermined distance L along the optical axis 523 from the imaging three-dimensional coordinates (aX, aY, aZ) . In this case, the image placement unit 13 places the target two-dimensional image 151 in the virtual space VS so that the target two-dimensional image 151 is orthogonal to an optical axis 523 of the camera CM and the size of the target two-dimensional image 151 corresponds to an angle of view θ of the camera CM.

[0062]

In this case, the image placement unit 13 calculates three-dimensional coordinates (bX, bY, bZ) , (cX, cY, cZ) , (dX, dY, dZ) in the virtual space VS where at least three corners 161, 162, 163 of four corners 161, 162, 163, 164 of the target two-dimensional image 151 are placed. Then, the image placement unit 13 places the target two-dimensional image 151 in the virtual space VS by placing at least three corners 161-163 of the four corners 161-164 of the target two-dimensional image 151 at the three-dimensional coordinates (bX, bY, bZ) to (dX, dY, dZ) .

[0063]

Next, the target detection unit 14 of the arithmetic unit 10 will be described with reference to FIGS. 2, 5, 6C, and 7. FIG. 6C is a side view schematically showing the target image 211 and a bounding box 221 included in the target two-dimensional image 151 placed in the virtual space VS, and the target region 311 included in the three-dimensional model 301. FIG. 7 is a front view showing the target two-dimensional image 151 placed in the virtual space VS. In FIG. 7, the target two-dimensional image 151 is viewed from the imaging position 101.

[0064]

As shown in FIG. 2, preferably, the target detection unit 14 utilizes the trained model 53 that is constructed by being trained with training data items. The trained model 53 is constructed by being trained with the training data items. In this case, the training is supervised learning. As an example, the trained model 53 employs YOLO (You Only Look Once), which is an object detection algorithm based on a convolutional neural network. The object detection algorithm is not particularly limited and may be, for example, R-CNN (Regions with Convolutional Neural Network) or SSD (Single Shot Multi-Box Detector). The training data items each include a two-dimensional image including a target image and an information tag. The information tag is attached to the two-dimensional image. The information tag includes a

bounding box surrounding the target image included in a subject image in the two-dimensional image. The information tag may further include a name of the target image. The bounding box corresponds to an example of "information for identifying the target image" of the present disclosure.

[0065]

The object detection algorithm may be an algorithm that performs segmentation. In this case, the object detection algorithm is, for example, "U-NET" that performs semantic segmentation, "Mask R-CNN" that performs instance segmentation, or "Panoptic Feature Pyramid Network" that performs panoptic segmentation.

[0066]

As shown in FIGS. 2, 5, 6C, and 7, the target detection unit 14, as an example, causes the trained model 53 to detect the target image 211 included in the target two-dimensional image 151 and sets the bounding box 221 surrounding the target image 211. In this way, the target detection unit 14 causes the trained model 53 to detect the target image 211 and identify the target image 211. In other words, when the target detection unit 14 inputs the target two-dimensional image 151 to the trained model 53, the trained model 53 outputs the target two-dimensional image 151 that includes the target image 211 surrounded by the bounding box 221. In FIG. 6C, the bounding box 221 is exaggerated to make the drawing easier to view.

[0067]

Although the target detection unit 14 employs the trained model 53, for example, a rule-based algorithm may be employed to detect the target image 211.

[0068]

Next, with reference to FIGS. 2, 5, 6C, 6D, and 7, the target coordinate calculation unit 15 will be described. FIG. 6D is a side view schematically showing a state in which a straight line 400 passing through the imaging position 101 and the target image 211 in the virtual space VS contacts the three-dimensional model 301.

[0069]

As shown in FIGS. 2, 5, 6C, and 7, the target coordinate calculation unit 15 calculates three-dimensional coordinates (eX, eY, eZ) in the virtual space VS of the target image 211 included in the target two-dimensional image 151. Hereinafter, the three-dimensional coordinates (eX, eY, eZ) are referred to as "first target three-dimensional coordinates (eX, eY, eZ)".

[0070]

Specifically, the target coordinate calculation unit 15 calculates the first target three-dimensional coordinates (eX, eY, eZ) based on the three-dimensional coordinates (bX, bY, bZ) to (dX, dY, dZ) of at least three corners 161 to 163 among the four corners 161-164 of the target

two-dimensional image 151.

[0071]

Preferably, as shown in FIG. 7, the first target three-dimensional coordinates (eX , eY , eZ) are indicated by three-dimensional coordinates in the virtual space VS of a specific point 224 in the bounding box 221. The specific point 224 is, as an example, the intersection of a diagonal line 222 and a diagonal line 223 of the bounding box 221. Thus, the specific point 224 is determined based on the bounding box 221. The specific point 224 is a point inside the target image 211. The bounding box 221 is a rectangle surrounding the target image 211.

[0072]

In this case, specifically, the target coordinate calculation unit 15 calculates two-dimensional coordinates (u , v) of the specific point 224 in the bounding box 221 in the target two-dimensional image 151. The two-dimensional coordinates (u , v) are coordinates of a screen coordinate system SC. The screen coordinate system SC is defined by mutually orthogonal U-axis and V-axis. An origin of the screen coordinate system SC is not particularly limited, but is set, for example, at the corner 161 of the target two-dimensional image 151. The origin may be set, for example, at a center of the target two-dimensional image 151.

[0073]

Then, based on the three-dimensional coordinates (bX , bY , bZ) to (dX , dY , dZ) of at least three corners 161 to 163 among the four corners 161 to 164 of the target two-dimensional image 151, the target coordinate calculation unit 15 transforms the two-dimensional coordinates (u , v) of the specific point 224 in the screen coordinate system SC into the first target three-dimensional coordinates (eX , eY , eZ), which are three-dimensional coordinates.

[0074]

The specific point 224 is determined based on the bounding box 221, but is not particularly limited as long as it is a point inside the target image 211.

[0075]

Next, as shown in FIGS. 5 and 6D, the target coordinate calculation unit 15 calculates three-dimensional coordinates (fX , fY , fZ) of a position where the straight line 400 passing through the imaging three-dimensional coordinates (aX , aY , aZ) and the first target three-dimensional coordinates (eX , eY , eZ) contacts the three-dimensional model 301. The three-dimensional coordinates (fX , fY , fZ) in this case are referred to as "second target three-dimensional coordinates (fX , fY , fZ)". In other words, the second target three-dimensional coordinates (fX , fY , fZ) are three-dimensional coordinates of an intersection point when the straight line 400 intersects the three-dimensional model 301. The straight line 400 is, for example, a virtual straight line.

[0076]

The second target three-dimensional coordinates (fX, fY, fZ) indicate three-dimensional coordinates of the target region 311 on the three-dimensional model 301. The target region 311 corresponds to the target image 211 and is a region that reproduces the target 310 in the subject 300 on the three-dimensional model 301.

[0077]

Specifically, the target coordinate calculation unit 15 calculates the second target three-dimensional coordinates (fX, fY, fZ), which are the position where the straight line 400 contacts the surface of the three-dimensional model 301. In other words, the second target three-dimensional coordinates (fX, fY, fZ) are three-dimensional coordinates of an intersection point when the straight line 400 intersects the surface of the three-dimensional model 301.

[0078]

For example, in a case where a plurality of polygonal (e.g., triangular) facets are assigned to the three-dimensional model 301 with the points 322 of the point cloud 321 serving as vertices, the target coordinate calculation unit 15 sets, as the second target three-dimensional coordinates, an average value of three-dimensional coordinates of a plurality of vertices defining the facet that the straight line 400 contacts among the plurality of polygonal facets. Alternatively, for example, the target coordinate calculation unit 15 sets, as the second target three-dimensional coordinates, three-dimensional coordinates of a vertex closest to the position where the straight line 400 contacts among the plurality of vertices defining the facet that the straight line 400 contacts.

[0079]

For example, in a case where a circular facet centered at each point 322 constituting the point cloud 321 is assigned to the three-dimensional model 301, the target coordinate calculation unit 15 sets, as the second target three-dimensional coordinates, three-dimensional coordinates of the point 322 at the center of the circular facet that the straight line 400 contacts.

[0080]

For example, in a case where a plurality of cubes are assigned to the three-dimensional model 301 by an octree algorithm, the target coordinate calculation unit 15 sets, as the second target three-dimensional coordinates, three-dimensional coordinates of the point 322 encompassed by the cube having a facet that the straight line 400 contacts. In a case where a plurality of points 322 are encompassed in the cube, an average value of three-dimensional coordinates of the plurality of points 322 is set as the second target three-dimensional coordinates.

[0081]

As explained above with reference to FIGS. 5 and 6, according to the embodiment, the target coordinate calculation unit 15 calculates the second target three-dimensional coordinates

(fX, fY, fZ), which are the three-dimensional coordinates of the target region 311 on the three-dimensional model 301, by determining the position where the straight line 400, which passes through the imaging three-dimensional coordinates (aX, aY, aZ) and the first target three-dimensional coordinates (eX, eY, eZ) of the target two-dimensional image 151, contacts the three-dimensional model 301 in the virtual space VS. Thus, in the embodiment, instead of performing a process to search for a target region across the entire area of a three-dimensional model, the position of the target region 311 on the three-dimensional model 301 is identified based on the imaging three-dimensional coordinates (aX, aY, aZ) and the single target two-dimensional image 151. In other words, the position of the target region 311 on the three-dimensional model 301 is identified by processing a two-dimensional image (the target two-dimensional image 151). Therefore, it is possible to identify the position of the target region 311 on the three-dimensional model 301 by a simple process.

[0082]

In addition, according to the embodiment, the position of the target region 311 on the three-dimensional model 301 is automatically identified by the arithmetic unit 10. Therefore, human work to visually search for the target region 311 on the three-dimensional model 301 can be omitted. Thus, the human workload can be reduced.

[0083]

In particular, the imaging position information item 521 (the imaging three-dimensional coordinates (aX, aY, aZ)) and the imaging posture information item 522 are necessarily calculated as byproducts when generating the three-dimensional model 301. Accordingly, it is possible to suppress the occurrence of additional processing for identifying the target region 311 on the three-dimensional model 301.

[0084]

Also, according to the embodiment, the model generation unit 11 assigns the surface to the three-dimensional model 301. Therefore, the straight line 400 always contacts the three-dimensional model 301. Thus, it is possible to reliably calculate the second target three-dimensional coordinates (fX, fY, fZ), which indicate the position of the target region 311, without depending on the density of the point cloud 321.

[0085]

Furthermore, according to the embodiment, the distortion correction unit 12 corrects distortion of the target two-dimensional image 151. Therefore, even when the target image 211 is present in a region where distortion occurs in the target two-dimensional image 151, it is possible to calculate with high accuracy the two-dimensional coordinates (u, v) indicating the position of the target image 211 and the first target three-dimensional coordinates (eX, eY, eZ).

[0086]

Furthermore, according to the embodiment, the target detection unit 14 inputs the target two-dimensional image 151 to the trained model 53 to have the target image 211 detected. In other words, the trained model 53 is constructed by being trained with the two-dimensional images 150 as training data items. Therefore, compared to a case where a trained model for detecting a target image is constructed by being trained with a three-dimensional model, it is possible to reduce the workload on an operator in creating training data, the volume of training data, and the amount of calculations for training. As a result, it is possible to reduce the cost and training time for constructing the trained model 53. In particular, the target image 211 is detected on a two-dimensional image (the target two-dimensional image 151) instead of analyzing a three-dimensional model to detect a target region. Therefore, according to the embodiment, it is possible to identify the target region 311 on the three-dimensional model 301 by a faster and simpler process compared to the case where a target region is detected by analyzing a three-dimensional model.

[0087]

Furthermore, according to the embodiment, the first target three-dimensional coordinates (eX, eY, eZ) are three-dimensional coordinates of the specific point 224 determined based on the bounding box 221. In other words, the specific point 224 is determined as the intersection of diagonal line 222 and diagonal line 223 of the bounding box 221. Therefore, it is possible to easily determine the specific point 224 by a simple process compared to a case where the specific point 224 is determined at any point in the target image 211.

[0088]

Next, an image analysis method according to the embodiment will be described with reference to FIGS. 2 and 8. FIG. 8 is a flowchart showing the image analysis method. The image analysis method is performed by the server 1. As shown in FIG. 8, the image analysis method includes Steps S1 to S9. A computer program stored in the storage unit 50 causes the arithmetic unit 10 to execute Steps S1 to S9. In other words, a computer program product realizes Steps S1 to S9 when the computer program is executed by the arithmetic unit 10. The arithmetic unit 10 corresponds to an example of a "computer" of the present disclosure.

[0089]

As shown in FIGS. 2 and 8, first, in Step S1, the model generation unit 11 generates the three-dimensional model 301 of the subject 300 based on the plurality of two-dimensional images 150 generated by imaging the subject 300 from the plurality of different imaging positions 100.

[0090]

Next, in Step S2, the distortion correction unit 12 acquires the target two-dimensional image 150A from the second database 52 from among the plurality of two-dimensional images

150.

[0091]

Next, in Step S3, the distortion correction unit 12 corrects distortion of the target two-dimensional image 150A based on the calibration information. The target two-dimensional image 150A whose distortion has been corrected is stored in the storage unit 50.

[0092]

Next, in Step S4, the image placement unit 13 acquires the imaging position information item 521 and the imaging posture information item 522 from the second database 52. The imaging position information item 521 indicates the position of the camera CM at the time of imaging by means of the imaging three-dimensional coordinates (aX , aY , aZ) in the three-dimensional coordinate system CS. The imaging posture information item 522 indicates the posture of the camera CM at the time of imaging by the rotation angles in the three-dimensional coordinate system CS.

[0093]

Next, in Step S5, the image placement unit 13 places the target two-dimensional image 150A, whose distortion has been corrected, in the virtual space VS based on the imaging position information item 521 and the imaging posture information item 522. The target two-dimensional image 150A placed in the virtual space VS is referred to as "target two-dimensional image 151".

[0094]

Next, in Step S6, the target detection unit 14 causes the trained model 53 to detect the target image 211 included in the target two-dimensional image 151 and sets the bounding box 221 surrounding the target image 211.

[0095]

Next, in Step S7, the target coordinate calculation unit 15 calculates the first target three-dimensional coordinates (eX , eY , eZ), which indicate three-dimensional coordinates of the specific point 224 in the bounding box 221.

[0096]

Next, in Step S8, the target coordinate calculation unit 15 calculates the second target three-dimensional coordinates (fX , fY , fZ) that indicate a position where the straight line 400 passing through the imaging three-dimensional coordinates (aX , aY , aZ) indicated by the imaging position information item 521 and the first target three-dimensional coordinates (eX , eY , eZ) contacts the three-dimensional model 301. The second target three-dimensional coordinates (fX , fY , fZ) indicate three-dimensional coordinates of the target region 311 on the three-dimensional model 301. In this way, a position of the target 310 included in the subject 300 is identified as the target region 311 on the three-dimensional model 301.

[0097]

Next, in Step S9, the display control unit 16 causes the three-dimensional model 301 to be displayed on the terminal 2 via the communication unit 40 in response to a request from a user of the terminal 2. For example, the display control unit 16 causes the three-dimensional model 301 to be displayed on the terminal 2 in a manner that allows operation. For example, the display control unit 16 causes the terminal 2 to display the target two-dimensional image 151 superimposed on the three-dimensional model 301. For example, the display control unit 16 causes the three-dimensional model 301 to be displayed on the terminal 2 in a posture such that the target region 311 on the three-dimensional model 301 is displayed on the terminal 2. In this case, for example, the display control unit 16 causes the target region 311 on the three-dimensional model 301 to be displayed on the terminal 2 in an enlarged state. Therefore, according to the embodiment, the user of the terminal 2 can omit the task of visually searching for the target region 311 on the three-dimensional model 301. In other words, it is possible to reduce the workload of the user. When Step S9 is completed, the image analysis method ends.

[0098]

As described above with reference to FIG. 8, according to the image analysis method of the embodiment, the position of the target region 311 on the three-dimensional model 301 is identified by calculating the position where the straight line 400 contacts the three-dimensional model 301 in the virtual space VS. Therefore, it is possible to identify the position of the target region 311 by a simple process.

[0099]

FIG. 9 is a flowchart showing the three-dimensional model generation process in Step S1 of FIG. 8. As shown in FIG. 9, the three-dimensional model generation process includes Steps S11 to S15.

[0100]

First, in Step S11, the model generation unit 11 acquires the plurality of two-dimensional images 150 including the subject image 200 from the first database 51.

[0101]

Next, in Step S12, the model generation unit 11 extracts feature points of the subject 300 from each of the two-dimensional images 150 and associates feature points across the two-dimensional images 150.

[0102]

Next, in Step S13, the model generation unit 11 estimates the imaging position 100 and the posture of the camera CM when each two-dimensional image 150 is generated based on the associated feature points. Then, the model generation unit 11 calculates three-dimensional coordinates of three-dimensional points corresponding to the feature points based on the

estimation results of the imaging position 100 (the imaging position information items 521) and the posture of the camera CM (the imaging posture information items 522). As a result, the point cloud data representing the point cloud 321 consisting of the plurality of three-dimensional points (the plurality of points 322) is generated.

[0103]

Next, in Step S14, the model generation unit 11 assigns the surface to the three-dimensional model 301 based on the point cloud data.

[0104]

Next, in Step S15, the model generation unit 11 adds material information based on the two-dimensional images 150 to the surface assigned to the three-dimensional model 301. The process then returns to Step S2 in FIG. 8.

[0105]

(Modification)

Referring to FIGS. 2 and 10, a modification of the embodiment is described. In the modification, the target 310 included in the subject 300 is a portion of the subject 300 that represents a specific state. The following is a description of the main points in which the modification differs from the above embodiment.

[0106]

In the modification, the target image 211 shown in FIG. 7 is an image of a portion of the subject 300 that represents a specific state (hereinafter referred to as "specific state portion"). Therefore, according to the modification, the specific state portion of the subject 300 is identified as the target region 311 on the three-dimensional model 301. In other words, the position of the target region 311 representing the specific state portion of the subject 300 is calculated as the second target three-dimensional coordinates on the three-dimensional model 301. As a result, human work to visually search for the target region 311 representing the specific state portion on the three-dimensional model 301 can be omitted. In addition, The risk of missing the detection of the target region 311 representing the specific state portion can be reduced.

[0107]

The specific state portion of the subject 300 is, for example, a portion of the entirety or a portion of an interest region of the subject 300 that represents an abnormal state that is not a normal state (hereinafter referred to as "abnormal state portion"). Therefore, in this case, the target image 211 is an image of the abnormal state portion. As a result, the need for human work to visually search for the target region 311 that represents the abnormal state portion on the three-dimensional model 301 can be omitted. In addition, the risk of missing the detection of the target region 311 that represents the abnormal state portion can be reduced.

[0108]

As an example, the target image 211 represents a region of the target two-dimensional image 151 that has changed relative to a reference two-dimensional image. The region that has changed relative to the reference two-dimensional image is, for example, a region that has undergone deformation, mutation, displacement, increase, decrease, or disappearance relative to the reference two-dimensional image. The reference two-dimensional image is, for example, a two-dimensional image of the subject 300 imaged earlier than the target two-dimensional image 151. Alternatively, the reference two-dimensional image is, for example, a two-dimensional image generated by imaging the subject 300 in a reference state. The reference state is, for example, a state where the entirety or the interest region of the subject 300 is in a normal state. In this example, the target image 211 is an image of a portion of the entirety or a portion of the interest region of the subject 300 that represents an abnormal state.

[0109]

Next, an image analysis method related to the modification is described with reference to FIG. 2 and FIG. 10. FIG. 10 is a flowchart showing the image analysis method according to the modification. The image analysis method is performed by the server 1. As shown in FIG. 10, the image analysis method includes Steps S21 to S23. A computer program stored in the storage unit 50 causes the arithmetic unit 10 to execute Steps S21 to S23. In other words, a computer program product realizes Steps S21 to S23 when the computer program is executed by the arithmetic unit 10.

[0110]

First, in Step S21, the target detection unit 14 executes a target image detection process. The target image detection process indicates a process of detecting a target image 210 from video data 511 of a subject 300, a two-dimensional image 150 extracted from the video data 511, a still image data set 512, or a two-dimensional image 150 extracted from the still image data set 512. Hereinafter, the video data 511, the two-dimensional image 150 extracted from the video data 511, the still image data set 512, and the still image data set 512 are collectively referred to as "detection target image data."

[0111]

For example, the target detection unit 14 inputs the detection target image data to a trained model. In a case where the trained model detects the target image 210 in the detection target image data, the trained model outputs information indicating that the detection target image data includes the target image 210. In a case where the trained model does not detect the target image 210 from the detection target image data, the trained model outputs information indicating that the detection target image data does not include the target image 210.

[0112]

The trained model is a computer program that is constructed by being trained with training data items. The trained model includes, for example, a neural network. The training is, for example, supervised learning. The training data item is the detection target image data and "information that indicates inclusion of a target image."

[0113]

Next, in Step S22, the target detection unit 14 determines whether or not the target image 210 is detected in the target image detection process.

[0114]

When a negative determination is made in Step S22 (NO), the image analysis method ends.

[0115]

On the other hand, when an affirmative determination is made in Step S22 (YES), the process proceeds to Step S23.

[0116]

Next, in Step S23, the arithmetic unit 10 identifies the target region 311 on the three-dimensional model 301 corresponding to the target image 210. The process of Step S23 is similar to the processes of Steps S1 to S9 shown in FIG. 8. When Step S23 is completed, the image analysis method ends.

[0117]

As described above with reference to FIG. 10, according to the modification, the processing of Step S23 is executed only when the target image 210 is detected. Therefore, the processing load of the arithmetic unit 10 can be reduced. In particular, when the target image 210 is not detected, the three-dimensional model 301 is not generated, thus reducing the processing load of the arithmetic unit 10. Also, in this case, an increase in the amount of data stored in the storage unit 50 can be suppressed.

[0118]

The image analysis method shown in FIG. 10 can also be applied to the above embodiment described with reference to FIGS. 1 to 9.

[0119]

While preferred embodiments and modifications of the present disclosure have been described in detail with reference to the accompanying drawings, the technical scope of the present disclosure is not limited to these examples. It is apparent that a person having ordinary skill in the technical field of the present disclosure can conceive of various modifications or alterations within the scope of the technical concept described in the claims, and it is naturally understood that these also fall within the technical scope of the present disclosure.

[0120]

The apparatus or system described in this specification may be realized as a single apparatus or may be realized by a plurality of apparatuses (e.g., cloud servers), some or all of which are connected via a network. For example, some or all of the model generation unit 11, the distortion correction unit 12, the image placement unit 13, the target detection unit 14, the target coordinate calculation unit 15, and the display control unit 16 may be realized by the same computer or server. Also, for example, the model generation unit 11, the distortion correction unit 12, the image placement unit 13, the target detection unit 14, the target coordinate calculation unit 15, and the display control unit 16 may be realized by the terminal 2. For example, the model generation unit 11, the distortion correction unit 12, the image placement unit 13, the target detection unit 14, the target coordinate calculation unit 15, and the display control unit 16 may be realized by separate computers or servers. For example, the first database 51, the second database 52, and the trained model 53 may be stored in separate storage devices or servers.

[0121]

A series of processes by the apparatus described in this specification may be realized using software, hardware, or a combination of software and hardware. It is possible to create a computer program to realize each function of the arithmetic unit 10 according to the present embodiment and to implement it in a PC or the like. Also, a computer-readable recording medium storing such a computer program can be provided. The recording medium is, for example, a magnetic disk, an optical disk, a magneto-optical disk, a flash memory, or the like. Also, the computer program described above may be distributed via a network, for example, without using a recording medium.

[0122]

The processes described in this specification using flowchart diagrams do not necessarily have to be executed in the illustrated order. Some processing steps may be executed in parallel. Also, additional processing steps may be adopted, and some processing steps may be omitted.

[0123]

The effects described in this specification are merely explanatory or illustrative and are not limiting. In other words, the technology according to the present disclosure may achieve other effects that are apparent to those skilled in the art from the description in this specification, either together with or in place of the above effects.

[0124]

The following configurations also belong to the technical scope of the present disclosure.

[0125]

(Item 1)

An image analysis apparatus comprising:

a model generation unit that generates a three-dimensional model of a subject based on a plurality of two-dimensional images generated by imaging the subject from a plurality of different imaging positions;

an image placement unit that places, among the plurality of two-dimensional images, a target two-dimensional image in a virtual space based on imaging three-dimensional coordinates and imaging posture information, the imaging three-dimensional coordinates being three-dimensional coordinates in the virtual space where the three-dimensional model is placed and indicating an imaging position when the target two-dimensional image is generated, and the imaging posture information indicating a posture of a camera that performs imaging; and

a target coordinate calculation unit that calculates second target three-dimensional coordinates indicating a position where a straight line passing through the imaging three-dimensional coordinates and first target three-dimensional coordinates contacts the three-dimensional model, wherein

the first target three-dimensional coordinates indicate a position in the virtual space of a target image included in the target two-dimensional image.

[0126]

(Item 2)

The image analysis apparatus according to item 1, wherein

the three-dimensional model is composed of point cloud data,

the model generation unit assigns a surface to the three-dimensional model based on the point cloud data, and

the target coordinate calculation unit calculates the second target three-dimensional coordinates, which are the position where the straight line contacts the surface.

[0127]

(Item 3)

The image analysis apparatus according to item 1 or 2, further comprising

a distortion correction unit that corrects distortion of the target two-dimensional image, wherein

the target coordinate calculation unit calculates the first target three-dimensional coordinates based on the target two-dimensional image that has been corrected for distortion.

[0128]

(Item 4)

The image analysis apparatus according to any one of items 1 to 3, further comprising a target detection unit that identifies the target image by causing a trained model to

detect the target image included in the target two-dimensional image, the trained model being constructed by being trained with training data items, wherein

the training data items each include a two-dimensional image including a target image and an information tag, and

the information tag includes information for identifying the target image.

[0129]

(Item 5)

The image analysis apparatus according to item 4, wherein

the target detection unit causes the trained model to detect the target image and sets a bounding box surrounding the target image, and

the first target three-dimensional coordinates indicate the three-dimensional coordinates of a specific point in the virtual space determined based on the bounding box.

[0130]

(Item 6)

The image analysis apparatus according to any one of items 1 to 5, wherein

the target image is an image of a portion of the subject that represents a specific state.

[0131]

(Item 7)

An image analysis method comprising:

generating a three-dimensional model of a subject based on a plurality of two-dimensional images generated by imaging the subject from a plurality of different imaging positions;

acquiring imaging three-dimensional coordinates and imaging posture information, the imaging three-dimensional coordinates being three-dimensional coordinates in a virtual space where the three-dimensional model is placed and indicating an imaging position when a target two-dimensional image among the plurality of two-dimensional images is generated, and the imaging posture information indicating a posture of a camera that performs imaging; placing the target two-dimensional image in the virtual space based on the imaging three-dimensional coordinates and the imaging posture information;

calculating first target three-dimensional coordinates, which are three-dimensional coordinates in the virtual space of a target image included in the target two-dimensional image; and

calculating second target three-dimensional coordinates, which are three-dimensional coordinates of a position where a straight line passing through the imaging three-dimensional coordinates and the first target three-dimensional coordinates contacts the three-dimensional model.

[0132]

(Item 8)

A computer program causing a computer to perform:

generating a three-dimensional model of a subject based on a plurality of two-dimensional images generated by imaging the subject from a plurality of different imaging positions;

acquiring imaging three-dimensional coordinates and imaging posture information, the imaging three-dimensional coordinates being three-dimensional coordinates in a virtual space where the three-dimensional model is placed and indicating an imaging position when a target two-dimensional image among the plurality of two-dimensional images is generated, and the imaging posture information indicating a posture of a camera that performs imaging; placing the target two-dimensional image in the virtual space based on the imaging three-dimensional coordinates and the imaging posture information;

calculating first target three-dimensional coordinates, which are three-dimensional coordinates in the virtual space of a target image included in the target two-dimensional image; and

calculating second target three-dimensional coordinates, which are three-dimensional coordinates of a position where a straight line passing through the imaging three-dimensional coordinates and the first target three-dimensional coordinates contacts the three-dimensional model.

[Industrial Applicability]

[0133]

The present disclosure provides an image analysis apparatus, an image analysis method, and a computer program, and has industrial applicability.

[Reference Signs List]

[0134]

1 Server (Image analysis apparatus), 11 Model generation unit, 12 Distortion correction unit, 13 Image placement unit, 14 Target detection unit, 15 Target coordinate calculation unit, 16 Display control unit, 53 Trained model

CLAIMS

1. An image analysis apparatus comprising:

a model generation unit that generates a three-dimensional model of a subject based on a plurality of two-dimensional images generated by imaging the subject from a plurality of different imaging positions;

an image placement unit that places, among the plurality of two-dimensional images, a target two-dimensional image in a virtual space based on imaging three-dimensional coordinates and imaging posture information, the imaging three-dimensional coordinates being three-dimensional coordinates in the virtual space where the three-dimensional model is placed and indicating an imaging position when the target two-dimensional image is generated, and the imaging posture information indicating a posture of a camera that performs imaging; and

a target coordinate calculation unit that calculates second target three-dimensional coordinates indicating a position where a straight line passing through the imaging three-dimensional coordinates and first target three-dimensional coordinates contacts the three-dimensional model, wherein

the first target three-dimensional coordinates indicate a position in the virtual space of a target image included in the target two-dimensional image.

2. The image analysis apparatus according to claim 1, wherein

the three-dimensional model is composed of point cloud data,

the model generation unit assigns a surface to the three-dimensional model based on the point cloud data, and

the target coordinate calculation unit calculates the second target three-dimensional coordinates, which are the position where the straight line contacts the surface.

3. The image analysis apparatus according to claim 1 or 2, further comprising

a distortion correction unit that corrects distortion of the target two-dimensional image, wherein

the target coordinate calculation unit calculates the first target three-dimensional coordinates based on the target two-dimensional image that has been corrected for distortion.

4. The image analysis apparatus according to claim 1 or 2, further comprising

a target detection unit that identifies the target image by causing a trained model to detect the target image included in the target two-dimensional image, the trained model being constructed by being trained with training data items, wherein

the training data items each include a two-dimensional image including a target image and an information tag, and

the information tag includes information for identifying the target image.

5. The image analysis apparatus according to claim 4, wherein
the target detection unit causes the trained model to detect the target image and sets a bounding box surrounding the target image, and

the first target three-dimensional coordinates indicate the three-dimensional coordinates of a specific point in the virtual space determined based on the bounding box.

6. The image analysis apparatus according to claim 1 or 2, wherein
the target image is an image of a portion of the subject that represents a specific state.

7. An image analysis method comprising:
generating a three-dimensional model of a subject based on a plurality of two-dimensional images generated by imaging the subject from a plurality of different imaging positions;
acquiring imaging three-dimensional coordinates and imaging posture information, the imaging three-dimensional coordinates being three-dimensional coordinates in a virtual space where the three-dimensional model is placed and indicating an imaging position when a target two-dimensional image among the plurality of two-dimensional images is generated, and the imaging posture information indicating a posture of a camera that performs imaging;
placing the target two-dimensional image in the virtual space based on the imaging three-dimensional coordinates and the imaging posture information;
calculating first target three-dimensional coordinates, which are three-dimensional coordinates in the virtual space of a target image included in the target two-dimensional image;
and

calculating second target three-dimensional coordinates, which are three-dimensional coordinates of a position where a straight line passing through the imaging three-dimensional coordinates and the first target three-dimensional coordinates contacts the three-dimensional model.

8. A computer program causing a computer to perform:
generating a three-dimensional model of a subject based on a plurality of two-dimensional images generated by imaging the subject from a plurality of different imaging positions;

acquiring imaging three-dimensional coordinates and imaging posture information, the imaging three-dimensional coordinates being three-dimensional coordinates in a virtual space where the three-dimensional model is placed and indicating an imaging position when a target two-dimensional image among the plurality of two-dimensional images is generated, and the imaging posture information indicating a posture of a camera that performs imaging;

placing the target two-dimensional image in the virtual space based on the imaging three-dimensional coordinates and the imaging posture information;

calculating first target three-dimensional coordinates, which are three-dimensional coordinates in the virtual space of a target image included in the target two-dimensional image; and

calculating second target three-dimensional coordinates, which are three-dimensional coordinates of a position where a straight line passing through the imaging three-dimensional coordinates and the first target three-dimensional coordinates contacts the three-dimensional model.