

MODEL RECONSTRUCTION METHOD, APPARATUS AND DEVICE, AND STORAGE MEDIUM

TECHNICAL FIELD

[0001] Embodiments of the disclosure relate to the field of image processing, and particularly to a model reconstruction method, apparatus and device, and a storage medium.

BACKGROUND

[0002] A model reconstruction technology or motion capture technology for an object to be reconstructed usually needs expensive optical equipment, and a special reflecting ball needs to be arranged at a specific joint point (key point) of the object to be reconstructed to facilitate the acquisition by the optical equipment. As a result, the device cost is relatively high, and it is unlikely to obtain a complete Three-Dimensional (3D) model.

SUMMARY

[0003] The embodiments of the disclosure provide a model reconstruction method, apparatus and device, and a storage medium.

[0004] A first aspect provides a model reconstruction method, which may include the following operations.

[0005] Scanned images of multiple viewing angles corresponding to an object to be reconstructed are acquired.

[0006] First Three-Dimensional (3D) key point information of the object to be reconstructed is determined based on the scanned images of the multiple viewing angles.

[0007] An initial parametric model is adjusted based on the first 3D key point

information of the object to be reconstructed to obtain a target parametric model characterizing a real posture of the object to be reconstructed.

[0008] In some embodiments, the scanned image may include a texture map. The operation that first 3D key point information of the object to be reconstructed is determined based on the scanned images of the multiple viewing angles may include the following operations.

[0009] A camera parameter and the texture map that correspond to each of the multiple viewing angles are acquired.

[0010] Two-Dimensional (2D) key point detection is performed on the texture map of each viewing angle to obtain 2D key point information corresponding to the texture map of each viewing angle.

[0011] The first 3D key point information of the object to be reconstructed is determined based on the camera parameter corresponding to each viewing angle and the 2D key point information corresponding to the texture map of each viewing angle.

[0012] In some embodiments, the operation that the first 3D key point information of the object to be reconstructed is determined based on the camera parameter corresponding to each viewing angle and the 2D key point information corresponding to the texture map of each viewing angle may include the following operations.

[0013] Position information of each 2D key point in the texture map of each viewing angle is determined based on the 2D key point information corresponding to the texture map of each viewing angle.

[0014] For each 2D key point, position information of a 3D key point corresponding to the 2D key point in a world coordinate system is determined based on the position information of the 2D key point in the texture map of each viewing angle and the camera parameter corresponding to each viewing angle, the first 3D key point information including the position information of the 3D key point corresponding to each 2D key point in the world coordinate system.

- [0015] In some embodiments, the scanned image may include a depth map. The method may further include the following operations.
- [0016] 3D point cloud data of the object to be reconstructed is constructed in a world coordinate system based on the scanned images of the multiple viewing angles.
- [0017] A first mesh model of the object to be reconstructed is constructed based on the 3D point cloud data.
- [0018] A surface texture of the first mesh model of the object to be reconstructed is determined based on the first mesh model and the scanned images of the multiple viewing angles.
- [0019] In some embodiments, the operation that 3D point cloud data of the object to be reconstructed is constructed in a world coordinate system based on the scanned images of the multiple viewing angles may include the following operations.
- [0020] A camera parameter corresponding to each of the multiple viewing angles is acquired.
- [0021] For each viewing angle, texture information and depth information of each of multiple scanned pixels are determined based on the texture map and depth map in the scanned image corresponding to the viewing angle, and a cloud point set corresponding to the viewing angle is generated based on the camera parameter corresponding to the viewing angle and the texture information and depth information of each scanned pixel, the point cloud set including position information of each scanned pixel in the world coordinate system.
- [0022] The 3D point cloud data of the object to be reconstructed is constructed based on the point cloud set corresponding to each viewing angle.
- [0023] In some embodiments, the scanned image may include a texture map. The operation that a surface texture of the first mesh model of the object to be reconstructed is determined based on the first mesh model and the scanned images of the multiple viewing angles may include the following operations.

- [0024] For any model triangular facet in the first mesh model, the model triangular facet is projected into the texture map of at least one viewing angle, a projected triangular facet corresponding to the model triangular facet in each texture map and texture information of the projected triangular facet are determined, and texture information of the model triangular facet is determined based on the texture information of each projected triangular facet.
- [0025] The surface texture of the first mesh model is generated based on the texture information of each model triangular facet in the first mesh model.
- [0026] In some embodiments, the operation that the model triangular facet is projected into the texture map of at least one viewing angle and a projected triangular facet corresponding to the model triangular facet in each texture map and texture information of the projected triangular facet are determined may include the following operations.
- [0027] For each viewing angle, projected vertex coordinates of the projected triangular facet corresponding to model vertex coordinates of the model triangular facet in the viewing angle are determined in the world coordinate system based on the model vertex coordinates of the model triangular facet and a camera parameter corresponding to the viewing angle.
- [0028] The projected triangular facet and the texture information of the projected triangular facet are determined in the texture map corresponding to the viewing angle based on the projected vertex coordinates of the projected triangular facet in the viewing angle.
- [0029] In some embodiments, the scanned image may further include a depth map. The operation that the projected triangular facet and the texture information of the projected triangular facet are determined in the texture map corresponding to the viewing angle based on the projected vertex coordinates of the projected triangular facet in the viewing angle may include the following operations.
- [0030] Reconstructed depth information corresponding to each projected pixel in the projected triangular facet is generated based on the model vertex coordinates of the model triangular facet and the camera parameter corresponding to the viewing angle.

- [0031] Original depth information corresponding to each projected pixel in the projected triangular facet is determined based on depth information of each scanned pixel in the depth map.
- [0032] A valid pixel in the projected triangular facet is determined based on the reconstructed depth information and the original depth information corresponding to each projected pixel in the projected triangular facet, a distance between the reconstructed depth information and the original depth information of the valid pixel being less than a preset depth distance threshold.
- [0033] The texture information of the projected triangular facet is determined based on the texture map and the valid pixel in the projected triangular facet.
- [0034] In some embodiments, the operation that an initial parametric model is adjusted based on the first 3D key point information of the object to be reconstructed to obtain a target parametric model characterizing a real posture of the object to be reconstructed may include the following operations.
- [0035] A target energy function is constructed based on the first 3D key point information, the target energy function being configured to determine a difference between the initial parametric model and the real posture of the object to be reconstructed, the target energy function at least including a first energy function, and the first energy function being configured to characterize a difference between second 3D key point information corresponding to the initial parametric model and the first 3D key point information.
- [0036] The initial parametric model is adjusted by minimizing the target energy function so as to obtain the target parametric model.
- [0037] In some embodiments, the first energy function may be constructed by the following operations.
- [0038] A second mesh model corresponding to the initial parametric model is determined based on an initial model parameter of the initial parametric model.
- [0039] The second 3D key point information corresponding to the initial parametric

model is determined based on the second mesh model.

[0040] The first energy function is constructed based on a 3D key point distance between the first 3D key point information and the second 3D key point information.

[0041] In some embodiments, the target energy function may further include at least one of a second energy function or a third energy function, the second energy function being configured to characterize a difference between the second mesh model corresponding to the initial parametric model and a third mesh model corresponding to the target object, and the third energy function being configured to constrain a rotation angle of a 3D key point in the initial parameter model.

[0042] In some embodiments, the target energy function may include the second energy function. The method may further include that: the third mesh model corresponding to the object to be reconstructed is acquired, the third mesh model being configured to characterize a body shape feature of the object to be reconstructed.

[0043] The second energy function may be constructed by the following operations.

[0044] A second model vertex set corresponding to the second mesh model and a third model vertex set corresponding to the third mesh model are acquired, the second model vertex set including 3D coordinates of each second model vertex in the second mesh model, and the third model vertex set including 3D coordinates of each third model vertex in the third mesh model.

[0045] A first distance corresponding to each second model vertex in the third model vertex set and a second distance corresponding to each third model vertex in the second model vertex set are determined based on the 3D coordinates of each third model vertex and the 3D coordinates of each second model vertex.

[0046] The second energy function is constructed based on the first distance corresponding to each second model vertex and the second distance corresponding to each third model vertex.

[0047] In some embodiments, the target energy function may include the third energy

function. The third energy function may be constructed by the following operations.

- [0048] A rotation threshold of at least one 3D key point is acquired, the rotation threshold including a rotation range of the 3D key point corresponding to each rotation axis.
- [0049] The third energy function is constructed based on the rotation threshold of the at least one 3D key point.
- [0050] In some embodiments, the scanned image may further include a depth map. The operation that the first 3D key point information of the object to be reconstructed is determined based on the camera parameter corresponding to each viewing angle and the 2D key point information corresponding to the texture map of each viewing angle may include the following operations.
- [0051] Position information of each 2D key point in the texture map of each viewing angle is determined based on the 2D key point information corresponding to the texture map of each viewing angle.
- [0052] For each viewing angle, intermediate 3D key point information of each 2D key point in the viewing angle in a world coordinate system is determined based on the position information of each 2D key point in the texture map of the viewing angle and based on the depth map and the camera parameter that correspond to the viewing angle.
- [0053] For each 2D key point, position information of a 3D key point corresponding to the 2D key point in the world coordinate system is determined based on the intermediate 3D key point information corresponding to the 2D key point at each viewing angle, the first 3D key point information including the position information of the 3D key point corresponding to each 2D key point in the world coordinate system.
- [0054] In some embodiments, the method further includes the following operations.
- [0055] A preset action set is acquired, the action set including multiple preset actions and second 3D key point information corresponding to each preset action.
- [0056] A preset action corresponding to target 3D key point information is determined

as a real action of the object to be reconstructed, the target 3D key point information matching the first 3D key point information of the object to be reconstructed.

[0057] A second aspect provides a model reconstruction apparatus, which may include an acquisition module, a determination module, and an adjustment module.

[0058] The acquisition module may be configured to acquire scanned images of multiple viewing angles corresponding to an object to be reconstructed.

[0059] The determination module may be configured to determine first 3D key point information of the object to be reconstructed based on the scanned images of the multiple viewing angles.

[0060] The adjustment module may be configured to adjust an initial parametric model based on the first 3D key point information of the object to be reconstructed to obtain a target parametric model characterizing a real posture of the object to be reconstructed.

[0061] A third aspect provides a model reconstruction device, which may include a memory and a processor. The memory may store a computer program capable of running in the processor. The processor may execute the computer program to implement the steps in the method.

[0062] A fourth aspect provides a computer storage medium, which may store one or more programs capable of being executed by one or more processors to implement the steps in the method.

[0063] In the embodiments of the disclosure, the scanned images of the object to be reconstructed at the multiple viewing angles may be acquired to determine the first 3D key point information of the object to be reconstructed. Compared with a method in a conventional art that needs the arrangement of an apparatus such as a reflecting ball at a key point position of the object to be reconstructed, the disclosure has the advantages that the 3D key point information of the object to be reconstructed may be determined based on the scanned images of the multiple viewing angles without any reflecting ball, so that not only is implementation facilitated, but also the used scanning device cost is reduced. In addition, the initial parametric model is adjusted based on the obtained first 3D key point

information of the object to be reconstructed to obtain the target parametric model, so that the conversion from a 3D key point to a parametric model may be implemented, and the obtained target parametric model may reflect the real posture of the object to be reconstructed.

BRIEF DESCRIPTION OF THE DRAWINGS

[0064] FIG. 1 is a flowchart of a model reconstruction method according to an embodiment of the disclosure.

[0065] FIG. 2 is a flowchart of a model reconstruction method according to an embodiment of the disclosure.

[0066] FIG. 3 is a flowchart of a model reconstruction method according to an embodiment of the disclosure.

[0067] FIG. 4 is a flowchart of a model reconstruction method according to an embodiment of the disclosure.

[0068] FIG. 5 is a flowchart of a model reconstruction method according to an embodiment of the disclosure.

[0069] FIG. 6 is a flowchart of a model reconstruction method according to an embodiment of the disclosure.

[0070] FIG. 7 is a composition structure diagram of a model reconstruction apparatus according to an embodiment of the disclosure.

[0071] FIG. 8 is a schematic diagram of a hardware entity of a model reconstruction device according to an embodiment of the disclosure.

DETAILED DESCRIPTION

[0072] The technical solutions of the disclosure will be specifically described below through the embodiments and in combination with the drawings in detail. The following specific embodiments may be combined. The same or similar concepts or processes will not be elaborated in some embodiments.

[0073] It is to be noted that, in the embodiments of the disclosure, "first", "second" and the like are adopted to distinguish similar objects and not intended to describe a target sequence or order. In addition, the technical solutions recorded in the embodiments of the disclosure may be combined freely without conflicts.

[0074] Before the embodiments of the disclosure are further described in detail, nouns and terms involved in the embodiments of the disclosure will be described. The nouns and terms involved in the embodiments of the disclosure are suitable to be explained as follows.

[0075] 1) Image coordinate system: an image coordinate system is a coordinate system established taking a top left vertex of an image acquired by a camera as an origin and millimeter or pixel as a unit. An x axis and y axis of the image coordinate system are length and width directions of the acquired image.

[0076] 2) Camera coordinate system: a camera coordinate system is a 3D rectangular coordinate system established taking a focus center of the camera as an origin and an optical axis as a z axis. An x axis of the camera coordinate system is parallel to that of the image coordinate system of the acquired image, while a y axis is parallel to that of the image coordinate system of the acquired image.

[0077] 3) World coordinate system: a world coordinate system may describe a position of the camera in a real world and a position of an object in the image acquired by the camera in the real world. The camera coordinate system may be converted into the world coordinate system based on a position and pose of the camera in the world coordinate system. Under normal conditions, an x axis of the world coordinate system horizontally points to due east, a y axis horizontally points to due north, and a z axis is vertically upward.

[0078] 4) An RGB image (R represents Red, i.e., red, G represents Green, i.e., green, and B represents Blue, i.e., blue) refers to an image obtained according to an RGB color mode. As a color standard in industry, the RGB color mode obtains various colors by

changing and superimposing three RGB color channels. RGB are colors representing the three RGB channels. This standard almost includes all colors visually perceivable to humans, and is one of the most extensively used color systems now.

[0079] 5) Depth map: in 3D computer graphics, a depth map is a map or map channel including information about distances between surfaces of a scene object at viewpoints. The depth map is similar to a gray-scale map, but each pixel value thereof is a practical distance between a sensor and the object.

[0080] 6) RGBD image (R represents Red, i.e., red, G represents Green, i.e., green, B represents Blue, i.e., blue, and D represents Depth, i.e., depth): an RGBD image includes an RGB image and a depth map. The RGB image and the depth map are registered, so pixels are in one-to-one correspondence. Therefore, the pixels in the RGB image and the depth map may be combined to generate the RGBD image.

[0081] 7) A Skinned Multi-Person Linear (SMPL) model is a parametric human body model. The model is driven by an SMPL parameter. The SMPL parameter includes a shape parameter β and a pose parameter θ . The shape parameter includes 10 parameters characterizing a height, weight, head-body ratio, etc., of a human body. The pose parameter includes 72 parameters corresponding to 24 joint points.

[0082] An embodiment of the disclosure provides a model reconstruction method. Model reconstruction of an object to be reconstructed may be completed based on scanned images with relatively low cost. The model reconstruction method provided in the embodiment of the disclosure is applied to an electronic device.

[0083] An exemplary application of an electrode device provided in an embodiment of the disclosure will be described below. The electronic device provided in the embodiment of the disclosure may be implemented as various types of user terminals (referred to as a terminal hereinafter) such as a notebook computer, a tablet computer, a desktop computer, a set-top box, and a mobile device (e.g., a mobile phone, a portable music player, a personal digital assistant, a dedicated messaging device, and a portable game device), or may be implemented as a server.

[0084] Referring to FIG. 1, FIG. 1 is an optional flowchart of a model reconstruction

method according to an embodiment of the disclosure. Descriptions will be made in combination with the operations shown in FIG. 1.

[0085] At S101, scanned images of multiple viewing angles corresponding to an object to be reconstructed are acquired.

[0086] In some embodiments, multiple scanning devices may be arranged in a real scene. Each scanning device may obtain a scanned image of a viewing angle. It is to be noted that, during the model reconstruction of the object to be reconstructed, the multiple scanning devices need to simultaneously scan the object to be reconstructed so as to obtain the scanned images of the multiple viewing angles.

[0087] In some embodiments, the multiple scanning devices may be fixed to a scanning bracket in advance. When the object to be reconstructed is scanned, the multiple scanning devices in the scanning bracket may surround the object to be reconstructed so as to obtain the scanned images of the multiple viewing angles from different viewing angles. For example, when there are four scanning devices, the scanning devices may be arranged in front of, behind and on the left and right of the object to be reconstructed respectively. When there are eight scanning devices, the scanning devices may be arranged in front of, behind and on the left, right, front-left, front-right, back-left and back-right of the object to be reconstructed respectively. Of course, multiple layers of scanning devices may be arranged to further obtain scanned images of the object to be reconstructed from more comprehensive viewing angles. The multiple layers of scanning devices may be arranged according to different heights. For example, when the object to be reconstructed is a person, a first-layer scanning device may be arranged at a leg height, and the first-layer scanning device may scan leg images of the object to be reconstructed from multiple viewing angles. A second-layer scanning device may be arranged at a body height, and the second-layer scanning device may scan body images of the object to be reconstructed from multiple viewing angles. A third-layer scanning device may be arranged at a head height, and the third scanning device may scan head images of the object to be reconstructed from multiple viewing angles.

[0088] In some embodiments, the object to be reconstructed may be a person, an animal, a plant, a vehicle, etc.

- [0089] In some embodiments, the scanned image corresponding to each viewing angle may include multiple scanned pixels. Each scanned pixel includes a color channel and a depth channel. For example, the scanning device is an RGBD camera. In such case, the scanned image includes three RGB color channels of each scanned pixel and a depth channel of each scanned pixel.
- [0090] At S102, first 3D key point information of the object to be reconstructed is determined based on the scanned images of the multiple viewing angles.
- [0091] The first 3D key point information is multiple 3D key points corresponding to the object to be reconstructed and 3D coordinates of each 3D key point in a world coordinate system.
- [0092] In some embodiments, 2D coordinates and depth information of a 2D key point corresponding to any viewing angle may be determined based on the scanned image corresponding to the viewing angle. The 2D coordinates of the 2D key point in an image coordinate system of the scanned image may be converted into 3D coordinates of a 3D key point corresponding to the 2D key point in the world coordinate system in combination with a camera parameter corresponding to the viewing angle.
- [0093] In some embodiments, 2D coordinates of each 2D key point in each viewing angle may be determined based on the scanned images corresponding to at least two viewing angles. 3D coordinates of a 3D key point corresponding to each 2D key point in the world coordinate system may be obtained in combination with a camera parameter corresponding to each viewing angle.
- [0094] At S103, an initial parametric model is adjusted based on the first 3D key point information of the object to be reconstructed to obtain a target parametric model characterizing a real posture of the object to be reconstructed.
- [0095] In some embodiments, the initial parametric model is a preset parametric model. The parametric model is a parameter-driven 3D model. For example, the object to be reconstructed is a person. In such case, the corresponding parametric model may be SMPL, DMPL, SMPL+H, and SMPL-x. For the ease of describing the disclosure, descriptions will be made in the following embodiment taking the condition that the parametric model

is SMPL as an example.

[0096] In some embodiments, the initial parametric model is a human body model determined based on an initial model parameter. The model parameter includes a shape parameter and a pose parameter. The shape parameter includes 10 parameters characterizing a height, weight, head-body ratio, etc., of a human body. The pose parameter includes 72 parameters corresponding to 24 joint points (the parameter corresponding to each joint point is represented by a 3D rotational vector, so there are included totally 24×3 parameters). The initial model parameter is adjusted such that a difference between second 3D key point information corresponding to a parametric model corresponding to an adjusted model parameter and the first 3D key point information satisfies a preset condition. The adjusted parametric model is determined as the target parametric model characterizing the real posture of the object to be reconstructed.

[0097] The second 3D key point information is multiple 3D key points corresponding to the parametric model and 3D coordinates of each 3D key point in the world coordinate system. A distance between the 3D coordinates of each 3D key point in the first 3D key point information and the 3D coordinates in the second 3D key point information is calculated to determine a difference between the second 3D key point information and the first 3D key point information.

[0098] In the embodiments of the disclosure, the scanned images of the object to be reconstructed at the multiple viewing angles may be acquired to determine the first 3D key point information of the object to be reconstructed. Compared with a method in the conventional art that needs the arrangement of an apparatus such as a reflecting ball at a key point position of the object to be reconstructed, the disclosure has the advantages that the 3D key point information of the object to be reconstructed may be determined based on the scanned images of the multiple viewing angles without any reflecting ball, so that not only is implementation facilitated, but also the used scanning device cost is reduced. In addition, the initial parametric model is adjusted based on the obtained first 3D key point information of the object to be reconstructed to obtain the target parametric model, so that the conversion from a 3D key point to a parametric model may be implemented, and the obtained target parametric model may reflect the real posture of the object to be

reconstructed.

[0099] Referring to FIG. 2, FIG. 2 is an optional flowchart of a model reconstruction method according to an embodiment of the disclosure. Based on FIG. 1, S102 in FIG. 1 may include S201 to S203. Descriptions will be made in combination with the operations shown in FIG. 2.

[00100] At S201, a camera parameter and the texture map that correspond to each of the multiple viewing angles are acquired.

[00101] The scanned image corresponding to each viewing angle may include a texture map configured to characterize a visual feature of the object to be reconstructed. For example, the scanned image is an RGBD image. In such case, a size of the scanned image is 800*600. That is, the scanned image includes 480,000 scanned pixels. Each scanned pixel includes three RGB color channels and a depth channel. A texture map in the scanned image includes the three RGB color channels of each scanned pixel. A depth map in the scanned image includes the depth channel of each scanned pixel.

[00102] The camera parameter includes a camera intrinsic parameter and a camera extrinsic parameter. The camera intrinsic parameter is configured to determine 3D coordinates of each scanned pixel in the scanned image in a camera coordinate system. The camera extrinsic parameter is configured to convert the 3D coordinates of the scanned pixel in the camera coordinate system into 3D coordinates in the world coordinate system.

[00103] At S202, 2D key point detection is performed on the texture map of each viewing angle to obtain 2D key point information corresponding to the texture map of each viewing angle.

[00104] In some embodiments, the texture map of each viewing angle may be detected by a preset 2D key point detection model so as to obtain the 2D key point information corresponding to the texture map of each viewing angle. The 2D key point information corresponding to the texture map of each viewing angle includes multiple 2D key points of the object to be reconstructed in the texture map and position information of each 2D key point in the texture map. The position information may be 2D coordinates of the 2D key point in the texture map.

- [00105] At S203, the first 3D key point information of the object to be reconstructed is determined based on the camera parameter corresponding to each viewing angle and the 2D key point information corresponding to the texture map of each viewing angle.
- [00106] In some embodiments, the operation that the first 3D key point information of the object to be reconstructed is determined based on the camera parameter corresponding to each viewing angle and 2D key point information corresponding to the texture map of each viewing angle may be implemented by S2031 to S2032.
- [00107] At S2031, position information of each 2D key point in the texture map of each viewing angle is determined based on the 2D key point information corresponding to the texture map of each viewing angle.
- [00108] In some embodiments, for the texture map of each viewing angle, the position information of each 2D key point of the object to be reconstructed in the texture map may be determined based on the 2D key point information in the texture map. For example, if the object to be reconstructed corresponds to N 2D key points, 2D key point information of each texture map may include 2D coordinates of each 2D coordinates, i.e., 2D coordinates of each 2D key point among P1, P2 to PN.
- [00109] At S2032, for each 2D key point, position information of a 3D key point corresponding to the 2D key point in a world coordinate system is determined based on the position information of the 2D key point in the texture map of each viewing angle and the camera parameter corresponding to each viewing angle, the first 3D key point information including the position information of the 3D key point corresponding to each 2D key point in the world coordinate system.
- [00110] In some embodiments, for each 2D key point, a triangularization process is performed by use of the position information (i.e., 2D coordinates) of the 2D key point at each viewing angle and the camera parameter (intrinsic parameter and extrinsic parameter) corresponding to the viewing angle to obtain position information (i.e., 3D coordinates) of the 3D key point corresponding to the 2D key point in the world coordinate system.
- [00111] For example, for any 2D key point, such as P1, position information (X1, Y1) corresponding to P1 at each viewing angle and a camera parameter corresponding to each

viewing angle are acquired. Position information and camera parameters corresponding to at least two viewing angles need to be acquired so as to determine position information of a 3D key point corresponding to P1 in the world coordinate system. Moreover, more accurate position information of the 3D key point in the world coordinate system may be obtained in case of more viewing angles. Based on the method, the position information of the 3D key point corresponding to each 2D key point in the world coordinate system may be obtained to further obtain the first 3D key point information.

[00112] In some other embodiments, the operation that the first 3D key point information of the object to be reconstructed is determined based on the camera parameter corresponding to each viewing angle and 2D key point information corresponding to the texture map of each viewing angle may be implemented by S2033 to S2035.

[00113] At S2033, position information of each 2D key point in the texture map of each viewing angle is determined based on the 2D key point information corresponding to the texture map of each viewing angle.

[00114] In some embodiments, for the texture map of each viewing angle, the position information of each 2D key point of the object to be reconstructed in the texture map may be determined based on the 2D key point information in the texture map. For example, if the object to be reconstructed corresponds to N 2D key points, 2D key point information of each texture map may include 2D coordinates of each 2D coordinates, i.e., 2D coordinates of each 2D key point among P1, P2 to PN.

[00115] At S2034, for each viewing angle, intermediate 3D key point information of each 2D key point in the viewing angle in a world coordinate system is determined based on the position information of each 2D key point in the texture map of the viewing angle and based on the depth map and the camera parameter that correspond to the viewing angle.

[00116] In some embodiments, for each 2D key point, depth data corresponding to the 2D key point at each viewing angle may be determined at first based on the position information (i.e., 2D coordinates) of the 2D key point corresponding to each viewing angle and the depth map corresponding to each viewing angle. For each viewing angle, the intermediate 3D key point information (i.e., 3D coordinates) of the 2D key point in the

viewing angle in the world coordinate system may be determined based on the depth data of the 2D key point, the position information (i.e., 2D coordinates corresponding to the viewing angle) and the camera parameter (intrinsic parameter and extrinsic parameter) corresponding to the viewing angle.

[00117] At S2035, for each 2D key point, position information of a 3D key point corresponding to the 2D key point in the world coordinate system is determined based on the intermediate 3D key point information corresponding to the 2D key point at each viewing angle, the first 3D key point information including the position information of the 3D key point corresponding to each 2D key point in the world coordinate system.

[00118] In some embodiments, after the intermediate 3D key point information corresponding to each 2D key point at each viewing angle is obtained, for each 2D key point, target 3D coordinates are determined in the world coordinate system such that a distance sum between the target 3D coordinates and the intermediate 3D key point information corresponding to the 2D key point at each viewing angle is minimum. The target 3D coordinates are determined as the position information of the 3D key point corresponding to the 2D key point in the world coordinate system. The target 3D coordinates may be set as O . The intermediate 3D key point information corresponding to the multiple viewing angles may be set as N_1 to N_n . The distance sum may be represented as $\sum_1^N \|O - N_n\|$. The distance sum may be minimized to obtain the target 3D coordinates O .

[00119] For example, for any 2D key point, such as P2, position information (X2, Y2) corresponding to P2 at each viewing angle and a depth map corresponding to each viewing angle may be acquired to obtain depth data corresponding to P2 at each viewing angle. For each viewing angle, intermediate 3D key point information (X', Y', Z') corresponding to P2 at each viewing angle may be determined based on the depth data, position information (X2, Y2) and camera parameter corresponding to the viewing angle. Position information of a 3D key point corresponding to P2 in the world coordinate system may be determined in combination with the intermediate 3D key point information (X', Y', Z') corresponding to each viewing angle. Based on the method, the position information of the 3D key point corresponding to each 2D key point in the world coordinate system may be obtained to

further obtain the first 3D key point information.

[00120] In some embodiments, a real action of the object to be reconstructed may further be determined after the first 3D key point information of the object to be reconstructed is obtained. The method further includes the following operations.

[00121] A preset action set is acquired, the action set including multiple preset actions and second 3D key point information corresponding to each preset action. A preset action corresponding to target 3D key point information is determined as a real action of the object to be reconstructed, the target 3D key point information matching the first 3D key point information of the object to be reconstructed.

[00122] In some embodiments, the multiple preset actions are preset. The second 3D key point information corresponding to each preset action is stored, the second 2D key point information including multiple second 3D key points and 3D coordinates of each second 3D key point. After the first 3D key point information is obtained, the 3D coordinates of each 3D key point in the first 3D key point information may be matched in the action set to obtain the target 3D key point information. The preset action corresponding to the target 3D key point information is determined as the real action of the object to be reconstructed so as to further complete the action recognition of the object to be reconstructed.

[00123] Referring to FIG. 3, FIG. 3 is an optional flowchart of a model reconstruction method according to an embodiment of the disclosure. Based on any above embodiment, the method may further include S301 to S303. Descriptions will be made in combination with the operations shown in FIG. 3.

[00124] At S301, 3D point cloud data of the object to be reconstructed is constructed in a world coordinate system based on the scanned images of the multiple viewing angles.

[00125] The scanned image includes three RGB color channels and depth channel of each scanned pixel.

[00126] In some embodiments, the operation that 3D point cloud data of the object to be reconstructed is constructed in a world coordinate system based on the scanned images of the multiple viewing angles may be implemented by S3011 to S3013.

- [00127] At S3011, a camera parameter corresponding to each of the multiple viewing angles is acquired.
- [00128] The camera parameter includes a camera intrinsic parameter and a camera extrinsic parameter. The camera intrinsic parameter is configured to determine 3D coordinates of each scanned pixel in the scanned image in a camera coordinate system. The camera extrinsic parameter is configured to convert the 3D coordinates of the scanned pixel in the camera coordinate system into 3D coordinates in the world coordinate system.
- [00129] At S3012, for each viewing angle, texture information and depth information of each of multiple scanned pixels are determined based on the texture map and depth map in the scanned image corresponding to the viewing angle, and a cloud point set corresponding to the viewing angle is generated based on the camera parameter corresponding to the viewing angle and the texture information and depth information of each scanned pixel, the point cloud set including position information of each scanned pixel in the world coordinate system.
- [00130] Depth information of each scanned pixel may be determined according to the depth map. The depth information is a value of the depth channel.
- [00131] In some embodiments, for each viewing angle, position information of a scanning point corresponding to each scanned pixel in the world coordinate system may be determined based on the position information and depth information of each scanned pixel in the scanned image and the camera parameter corresponding to the viewing angle.
- [00132] At S3013, the 3D point cloud data of the object to be reconstructed is constructed based on the point cloud set corresponding to each viewing angle.
- [00133] In some embodiments, before the scanned images of the multiple viewing angles corresponding to the object to be reconstructed are acquired, namely when any object to be reconstructed does not appear at each viewing angle, scanned images of a present real scene corresponding to multiple views are acquired at first, and 3D point cloud data of the scene is constructed.
- [00134] After the point cloud set corresponding to each viewing angle is obtained,

namely after the point cloud set corresponding to a scene including the object to be reconstructed at each viewing angle is obtained, the point cloud set corresponding to each viewing angle may be directly fused, and 3D point data of the scene in a fused point cloud set is removed, so as to further obtain the 3D point cloud data of the object to be reconstructed.

[00135] In some other embodiments, after the 3D point cloud data of the scene including the fused point cloud data is removed, a denoising process is further needed to eliminate noise points therein, so as to obtain the 3D point cloud data of the object to be reconstructed.

[00136] At S302, a first mesh model of the object to be reconstructed is constructed based on the 3D point cloud data.

[00137] In some embodiments, surface reconstruction may be performed on the 3D point cloud data by use of a Screened Poisson Surface Reconstruction (SPSR) method to obtain the first mesh model of the object to be reconstructed.

[00138] At S303, a surface texture of the first mesh model of the object to be reconstructed is determined based on the first mesh model and the scanned images of the multiple viewing angles.

[00139] In some embodiments, the operation that 3D point cloud data of the object to be reconstructed is constructed in a world coordinate system based on the scanned images of the multiple viewing angles may be implemented by S3031 to S3033.

[00140] At S3031, for any model triangular facet in the first mesh model, the model triangular facet is projected into the texture map of at least one viewing angle, a projected triangular facet corresponding to the model triangular facet in each texture map and texture information of the projected triangular facet are determined, and texture information of the model triangular facet is determined based on the texture information of each projected triangular facet.

[00141] Taking any model triangular facet $\triangle ABC$ in the first mesh model as an example, a camera parameter corresponding to at least one viewing angle is acquired. Descriptions are made taking a first viewing angle as an example. Three model vertexes (A, B, C) of the

model triangular facet $\triangle ABC$ are converted to the camera coordinate system based on a camera parameter corresponding to the first viewing angle, and then are converted into three projected vertexes (D, E, F) of a projected triangular facet of a scanned image (texture map) of the first viewing angle. Projected vertex coordinates of each projected vertex are obtained. An obtained triangular facet $\triangle DEF$ is a projected triangular facet obtained by projecting the model triangular facet $\triangle ABC$ to the first viewing angle. Therefore, texture information of the projected triangular facet $\triangle DEF$ may be determined based on a position of the projected triangular facet $\triangle DEF$ in the scanned image (texture map). Based on the method, texture information of a projected triangular facet $\triangle DEF$ corresponding to the model triangular facet $\triangle ABC$ in each viewing angle may be obtained.

[00142] In some embodiments, for any model triangular facet, the surface texture of the model triangular facet may not be observed at some viewing angles. Therefore, before the model triangular facet is projected into the texture map corresponding to the at least one viewing angle, the multiple viewing angles need to be screened the model triangular facet to remove an invalid viewing angle in the multiple viewing angles so as to obtain the at least one viewing angle. A method for removing the invalid viewing angle includes that: a viewing angle vector of each viewing angle is determined based on the camera parameter corresponding to each viewing angle; an orientation vector of the model triangular facet is determined based on vertex information of the model triangular facet, the orientation vector being perpendicular to the model triangular facet and pointing to a texture presentation direction of the model triangular facet; and an angle between the viewing angle vector and the orientation vector is determined, and the viewing angle is determined as an invalid viewing angle when the angle is smaller than a preset threshold. The preset threshold may be set to be 90 degrees. When the angle between the viewing angle vector and the orientation vector is 180 degrees, it indicates that the model triangular facet is vertically observed at the viewing angle. When the angle decreases gradually until the angle between the viewing angle vector and the orientation vector is 90 degrees, it indicates that the model triangular facet just may not be observed at the viewing angle. When the angle between the viewing angle vector and the orientation vector is 0 to 90 degrees, it indicates that the model triangular facet may not be observed at the viewing angle.

[00143] In some embodiments, after the texture information of the projected triangular facet corresponding to the model triangular facet in each viewing angle is obtained, a texture confidence of the projected triangular facet corresponding to each viewing angle may be determined based on an angle corresponding to each viewing angle. If the angle corresponding to the viewing angle is larger, the texture confidence is higher. Weighted summation is performed on the texture information of the projected triangular facet corresponding to each viewing angle based on the texture confidence corresponding to each viewing angle to obtain the texture information of the model triangular facet.

[00144] At S3032, the surface texture of the first mesh model is generated based on the texture information of each model triangular facet in the first mesh model.

[00145] Texture information of each model triangular facet in the first mesh model is sequentially determined based on the above-mentioned method to further form the surface texture of the first mesh model.

[00146] Based on the embodiment of the disclosure, the surface texture of the first mesh model may be restored based on the scanned images acquired at the multiple viewing angles. In addition, the texture information of each model triangular facet may be determined based on the scanned image of at least one viewing angle, so that the authenticity of the texture information of the model triangular facet is improved, and the overall authenticity of the surface texture of the first mesh model is further improved.

[00147] Referring to FIG. 4, FIG. 4 is an optional flowchart of a model reconstruction method according to an embodiment of the disclosure. Based on FIG. 3, S3031 in FIG. 3 may include S401 to S402. Descriptions will be made in combination with the operations shown in FIG. 4.

[00148] At S401, for each viewing angle, projected vertex coordinates of the projected triangular facet corresponding to model vertex coordinates of the model triangular facet in the viewing angle are determined in the world coordinate system based on the model vertex coordinates of the model triangular facet and a camera parameter corresponding to the viewing angle.

[00149] In some embodiments, taking any model triangular facet $\triangle ABC$ in the first mesh model as an example, a camera parameter corresponding to at least one viewing angle is acquired. Descriptions are made taking a first viewing angle as an example. Three model vertexes (A, B, C) of the model triangular facet $\triangle ABC$ are converted to the camera coordinate system based on a camera parameter corresponding to the first viewing angle, and then are converted into three projected vertexes (D, E, F) of a projected triangular facet of a scanned image (texture map) of the first viewing angle. Projected vertex coordinates of each projected vertex are obtained. An obtained triangular facet $\triangle DEF$ is a projected triangular facet obtained by projecting the model triangular facet $\triangle ABC$ to the first viewing angle.

[00150] At S402, the projected triangular facet and the texture information of the projected triangular facet are determined in the texture map corresponding to the viewing angle based on the projected vertex coordinates of the projected triangular facet in the viewing angle.

[00151] In some embodiments, texture information of the projected triangular facet $\triangle DEF$ may be determined based on the projected vertex coordinates of the three projected vertexes (D, E, F) of the projected triangular facet $\triangle DEF$ in the scanned image (texture map). Based on the method, texture information of a projected triangular facet $\triangle DEF$ corresponding to the model triangular facet $\triangle ABC$ in each viewing angle may be obtained.

[00152] In order to avoid a texture reconstruction error caused by an excessive difference between the reconstructed model triangular facet in the first mesh model and the real object to be reconstructed, i.e., a local recess or bump, checking is needed during the texture reconstruction of each model triangular facet, to avoid the acquisition of wrong texture information. In some embodiments, original depth information of each projected vertex in the projected triangular facet may be determined based on a depth map corresponding to the viewing angle. Reconstructed depth information of the projected vertex corresponding to each model vertex is determined based on the original depth information of each projected vertex, the camera parameter corresponding to the viewing angle and the model

vertex coordinates of each model vertex in the model triangular facet. Whether the texture information of the projected triangular facet corresponding to the viewing angle is available for the determination of the texture information of the model triangular facet is determined based on the reconstructed depth information and the original depth information corresponding to each model vertex.

[00153] In some embodiments, for each projected vertex, a depth distance of the projected vertex is determined based on the original depth information and reconstructed depth information of the projected vertex. When the depth distance of each projected vertex satisfies a preset accuracy condition, the texture information of the projected triangular facet corresponding to the viewing angle is available for the determination of the texture information of the model triangular facet. When the depth distance of each projected vertex does not satisfy the preset accuracy condition, the texture information of the projected triangular facet corresponding to the viewing angle is unavailable for the determination of the texture information of the model triangular facet, and the texture information of the projected triangular facet corresponding to the viewing angle is not considered during the calculation of the texture information of the model triangular facet. The preset accuracy condition may be that the depth distance of each projected vertex is smaller than a preset depth distance threshold. Alternatively, the preset accuracy condition may be that an average value of the depth distances of the projected vertexes is smaller than the preset depth distance threshold. No limits are made thereto in the disclosure.

[00154] In some embodiments, the operation that the projected triangular facet and the texture information of the projected triangular facet are determined in the texture map corresponding to the viewing angle based on the projected vertex coordinates of the projected triangular facet in the viewing angle may be implemented by S4021 to S4024.

[00155] At S4021, reconstructed depth information corresponding to each projected pixel in the projected triangular facet is generated based on the model vertex coordinates of the model triangular facet and the camera parameter corresponding to the viewing angle.

[00156] The projected vertex coordinates and reconstructed depth information of each projected vertex in the projected triangular facet corresponding to the viewing angle may be determined based on the model vertex coordinates of the model triangular facet and

the camera parameter corresponding to the viewing angle. All projected pixels (including the projected vertexes) in the projected triangular facets and position information (i.e., 2D coordinates) of each projected pixel in the scanned image (texture map) may be determined based on the projected vertex coordinates of each projected vertex. Reconstructed depth information of each projected pixel may be obtained by linear calculation in combination with the reconstructed depth information of each projected vertex. For example, if depth information of two projected vertexes is 8 and 10 respectively, depth information of the only projected pixel in the two projected vertexes is 9.

[00157] At S4022, original depth information corresponding to each projected pixel in the projected triangular facet is determined based on depth information of each scanned pixel in the depth map.

[00158] The original depth information corresponding to each projected pixel may be acquired based on the position information of each projected pixel in the projected triangular facet in the depth map.

[00159] At S4023, a valid pixel in the projected triangular facet is determined based on the reconstructed depth information and the original depth information corresponding to each projected pixel in the projected triangular facet, a distance between the reconstructed depth information and the original depth information of the valid pixel being less than a preset depth distance threshold.

[00160] For any projected pixel, a distance between the reconstructed depth information and the original depth information of the projected pixel may be determined. The projected pixel is determined as a valid pixel when the distance is smaller than a preset depth distance threshold. The projected pixel is determined as an invalid pixel when the distance is more than or equal to the preset depth distance threshold. It is to be noted that, for a projected triangular facet at a viewing angle, all projected pixels may be valid pixels, or all the projected pixels are invalid pixels, or part of projected pixels are valid pixels while the other part of projected pixels are invalid pixels.

[00161] At S4024, the texture information of the projected triangular facet is determined based on the texture map and the valid pixel in the projected triangular facet.

[00162] Pixel information (e.g., values of three RGB channels in the above-mentioned example) of each valid pixel may be acquired from the texture map based on position information of each valid pixel in the projected triangular facet to further obtain the texture information of the projected triangular facet.

[00163] Referring to FIG. 5, FIG. 5 is an optional flowchart of a model reconstruction method according to an embodiment of the disclosure. Based on any above-mentioned embodiment, such as FIG. 1, S103 in FIG. 1 may be updated to S501 to S502. Descriptions will be made in combination with the operations shown in FIG. 5.

[00164] At S501, a target energy function is constructed based on the first 3D key point information, the target energy function being configured to determine a difference between the initial parametric model and the real posture of the object to be reconstructed, the target energy function at least including a first energy function, and the first energy function being configured to characterize a difference between second 3D key point information corresponding to the initial parametric model and the first 3D key point information.

[00165] In some embodiments, after the initial parametric model is acquired, the second 3D key point information corresponding to the initial parametric model may be determined based on a pose parameter corresponding to the initial parametric model. The second 3D key point information is the same as the first 3D key point information in 3D key point count and joint information corresponding to key points. It is to be noted that, when the pose parameter is different from a 3D key point count of the first 3D key point information and joint information corresponding to key points, the pose parameter may be converted into second 3D key point information of the same format by a preset conversion method so as to facilitate the comparison between the initial parametric model and the first mesh model.

[00166] In some embodiments, the construction of the first energy function may be implemented by S5011 to S5013.

[00167] At S5011, a second mesh model corresponding to the initial parametric model is determined based on an initial model parameter of the initial parametric model.

[00168] At S5012, the second 3D key point information corresponding to the initial

parametric model is determined based on the second mesh model.

[00169] At S5013, the first energy function is constructed based on a 3D key point distance between the first 3D key point information and the second 3D key point information.

[00170] The first energy function may be constructed through formula (1).

[00171] $E_{\mathcal{P}}(\theta, \beta, t) = \frac{1}{P} \sum_{\mathcal{P}} \|\mathcal{T}(\mathcal{M}(\theta, \beta, t)) - \mathcal{P}_{\mathcal{P}}\|$ Formula (1).

[00172] θ represents a pose parameter. β represents a shape parameter. t represents a translation parameter. $\mathcal{M}(\cdot)$ represents a parametric model. The second mesh model corresponding to the initial parametric model may be determined based on the pose parameter, the shape parameter and the translation parameter. $\mathcal{T}(\cdot)$ may determine the second 3D key point information based on the second mesh model corresponding to the initial parametric model. $\mathcal{P}_{\mathcal{P}}$ represents a $\mathcal{P}$ -th 3D key point in the first 3D key point information. Position information of the $\mathcal{P}$ -th key point in the first 3D key point information and position information of a second 3D key point may be determined through formula (1) to further obtain a 3D key point distance corresponding to the $\mathcal{P}$ -th 3D key point. An averaging process may be performed on P 3D key points to obtain the first energy function $E_{\mathcal{P}}(\theta, \beta, t)$.

[00173] At S502, the initial parametric model is adjusted by minimizing the target energy function so as to obtain the target parametric model.

[00174] In some embodiments, an initial model parameter of the initial parametric model may be adjusted to align 3D key point information of an adjusted parametric model with the first 3D key point information to further obtain the target parametric model more similar to the first mesh model. If a value of the target energy function for adjustment is smaller, it indicates that the obtained target parametric model is more similar to the first mesh model.

[00175] In some embodiments, when an energy value corresponding to the target energy

function is minimum, a parametric model corresponding to a present model parameter is determined as the target parametric model, and the present model parameter is stored as a model parameter of the target parametric model.

[00176] In the embodiment, the target energy function includes the first energy function only. With the construction of the first energy function, 3D key points of the obtained target parametric model may be closer to those of the second mesh model. Furthermore, the model to be reconstructed is further reconstructed from a skeleton angle to obtain a corresponding target parametric model.

[00177] In some embodiments, the target energy function further includes at least one of a second energy function or a third energy function. The second energy function is configured to characterize a difference between the second mesh model corresponding to the initial parametric model and a third mesh model corresponding to the target object. The third energy function is configured to constrain a rotation angle of a 3D key point in the initial parameter model.

[00178] When the target energy function includes the second energy function, referring to FIG. 6, FIG. 6 is an optional flowchart of a model reconstruction method according to an embodiment of the disclosure. Based on any above-mentioned embodiment, such as FIG. 1, FIG. 1 may further include S601, and S103 in FIG. 1 may be updated to S602 to S603. Descriptions will be made in combination with the operations shown in FIG. 6.

[00179] At S601, a third mesh model corresponding to the object to be reconstructed is acquired, the third mesh model being configured to characterize a body shape feature of the object to be reconstructed.

[00180] In some embodiments, the third mesh model is a mesh model obtained by scanning the whole object to be reconstructed through a high-accuracy scanning device. The scanning device may scan a surface point cloud of the object to be reconstructed to further obtain a mesh model of the object to be reconstructed. The scanning device may be an Artec Eva 3D scanner.

[00181] At S602, a target energy function is constructed based on the first 3D key point information and the third mesh model, the target energy function being configured to

determine a difference between the initial parametric model and the real posture of the object to be reconstructed.

[00182] The target energy function at least includes a first energy function and a second energy function. The first energy function is configured to characterize a difference between second 3D key point information corresponding to the initial parametric model and the first 3D key point information. The second energy function is configured to characterize a difference between a second mesh model corresponding to the initial parameter model and the third mesh model corresponding to the target object.

[00183] In some embodiments, when the target energy function includes both the first energy function and the second energy function, a first weight corresponding to the first energy function and a second weight corresponding to the second energy function may be acquired. The target energy function is constructed based on the first weight, the second weight, the first energy function and the second energy function.

[00184] In some embodiments, the construction of the first energy function may be implemented by S6021 to S6023.

[00185] At S6021, the second mesh model corresponding to the initial parametric model is determined based on an initial model parameter of the initial parametric model.

[00186] At S6022, the second 3D key point information corresponding to the initial parametric model is determined based on the second mesh model.

[00187] At S6023, the first energy function is constructed based on a 3D key point distance between the first 3D key point information and the second 3D key point information.

[00188] In some embodiments, the construction of the second energy function may be implemented by S6024 to S6026.

[00189] At S6024, a second model vertex set corresponding to the second mesh model and a third model vertex set corresponding to the third mesh model are acquired, the second model vertex set including 3D coordinates of each second model vertex in the second mesh

model, and the third model vertex set including 3D coordinates of each third model vertex in the third mesh model.

[00190] At S6025, a first distance corresponding to each second model vertex in the third model vertex set and a second distance corresponding to each third model vertex in the second model vertex set are determined based on the 3D coordinates of each third model vertex and the 3D coordinates of each second model vertex.

[00191] At S6026, the second energy function is constructed based on the first distance corresponding to each second model vertex and the second distance corresponding to each third model vertex.

[00192] The second energy function may be constructed through formula (2).

$$[00193] \quad E_S = \frac{1}{|V_H|} \sum_{v_H \in V_H} \min_{v_S \in V_S} \|v_H - v_S\| + \frac{1}{|V_S|} \sum_{v_S \in V_S} \min_{v_H \in V_H} \|v_H - v_S\|$$

Formula (2).

[00194] V_H represents the third model vertex set corresponding to the third mesh model. V_S represents the second model vertex set corresponding to the second mesh model. v_H represents any model vertex in the third model vertex set. v_S represents any model vertex in the second model vertex set. $\sum_{v_H \in V_H} \min_{v_S \in V_S} \|v_H - v_S\|$ represents the determination of the first distance corresponding to each second model vertex in the third model vertex set. $\sum_{v_S \in V_S} \min_{v_H \in V_H} \|v_H - v_S\|$ represents the determination of the second distance corresponding to each third model vertex in the second model vertex set. The second energy function may be obtained through formula (2) based on an average value of the first distances corresponding to the second model vertexes and an average value of the second distances corresponding to the third model vertexes.

[00195] At S603, the initial parametric model is adjusted by minimizing the target energy function so as to obtain the target parametric model.

[00196] In some embodiments, when the target energy function includes the third energy function, the third energy function is constructed by the following operations: a rotation

threshold of at least one 3D key point is acquired, the rotation threshold including a rotation range of the 3D key point corresponding to each rotation axis; and the third energy function is constructed based on the rotation threshold of the at least one 3D key point.

[00197] The third energy function may be constructed through formula (3).

$$[00198] \quad E_a = \frac{1}{J \times 3} \sum_j^{J \times 3} \exp \left(\max(\theta_i - \theta_i^u, 0) + \max(\theta_i^l - \theta_i, 0) \right) - 1$$

Formula (3).

[00199] The rotation range may include an upper limit and lower limit of a rotation angle of the key point. θ_i^u represents the upper limit of the rotation angle. θ_i^l represents the lower limit of the rotation angle. J represents the key point. Rotation about three rotation axes is needed during a rotation of the key point, so each rotation needs to be counted. An energy value of the third energy function obtained through formula (3) when the rotation angle of the key point is larger than a preset rotation threshold is relatively large. Further, a key point angle of the obtained target parametric model is consistent with a physiological condition of a human body. Furthermore, a more real target parametric model may be obtained.

[00200] An exemplary application of the embodiment of the application in a practical application scene will now be described.

[00201] Understanding and modeling a human body is a long-standing issue in the field of computer vision and computer graphics researches as well as a basic technology for numerous applications such as animation, game, augmented reality, and virtual reality. With the development of 3D sensors, the arrangement of depth cameras or low-power laser radars in mobile devices has become popular in recent years.

[00202] A personal data acquisition process is completed by two independent systems, i.e., a human body scanning system and a pose capturing system. The pose capturing system may include scanning devices corresponding to multiple viewing angles. The scanning device includes, but not limited to, a depth camera and a laser radar. The multiple scanning devices are arranged on a preset frame. The pose capturing system may include

multiple Kinect devices capable of providing synchronous RGB images and depth maps (corresponding to the texture maps and depth maps in the scanned images) from different angles for 3D reconstruction. All of the scanning devices are connected to a computer with a control program so as to implement synchronous scanning and data processing. The human body scanning system includes an Artec Eva 3D scanner, a rotational display frame and a computer running a scanning program. Such a hardware combination allows 360-degree full-body scanning. The two systems run respectively to support the simultaneous data acquisition of two persons.

[00203] A camera parameter, i.e., a camera intrinsic parameter and a camera extrinsic parameter, of each scanning device needs to be acquired before model reconstruction. When a target person is located in an acquisition region of the data acquisition apparatus, scanned images of the target person may be acquired from multiple viewing angles through the multiple scanning devices in the data acquisition apparatus. A scanned image sequence of the target person may be acquired based on a preset acquisition frequency so as to complete the complete motion capturing of the target person.

[00204] Taking an acquisition condition of a moment t as an example, the multiple scanning devices may simultaneously acquire scanned images of the target person at the moment t to obtain a set of scanned images corresponding to the moment t . The set of scanned images may provide pose information and texture information of the target person from multiple viewing angles.

[00205] In some embodiments, taking the condition that the scanning device is an RGBD device as an example, an obtained depth map may be an RGBD image, and the RGBD image includes an RGB image and a depth map. The RGB image and the depth map are registered, namely each pixel in the RGB image corresponds to a pixel in the depth map (the pixels are in one-to-one correspondence). For example, taking pixel A1 in the RGB image as an example, the depth map includes pixel A2 corresponding to pixel A1, and a pixel value of pixel A2 characterizes a distance between the scanning device in a real scene and pixel A1 of the target person.

[00206] FIG. 7 is a composition structure diagram of a model reconstruction apparatus according to an embodiment of the disclosure. As shown in FIG. 7, the model

reconstruction apparatus 700 includes an acquisition module 701, a determination module 702, and an adjustment module 703.

[00207] The acquisition module 701 is configured to acquire scanned images of multiple viewing angles corresponding to an object to be reconstructed.

[00208] The determination module 702 is configured to determine first 3D key point information of the object to be reconstructed based on the scanned images of the multiple viewing angles.

[00209] The adjustment module 703 is configured to adjust an initial parametric model based on the first 3D key point information of the object to be reconstructed to obtain a target parametric model characterizing a real posture of the object to be reconstructed.

[00210] In some embodiments, the determination module 702 is further configured to:

[00211] acquire a camera parameter and the texture map that correspond to each of the multiple viewing angles;

[00212] perform 2D key point detection on the texture map of each viewing angle to obtain 2D key point information corresponding to the texture map of each viewing angle;
and

[00213] determine the first 3D key point information of the object to be reconstructed based on the camera parameter corresponding to each viewing angle and the 2D key point information corresponding to the texture map of each viewing angle.

[00214] In some embodiments, the determination module 702 is further configured to:

[00215] determine position information of each 2D key point in the texture map of each viewing angle based on the 2D key point information corresponding to the texture map of each viewing angle; and

[00216] for each 2D key point, determine position information of a 3D key point corresponding to the 2D key point in a world coordinate system based on the position information of the 2D key point in the texture map of each viewing angle and the camera

parameter corresponding to each viewing angle, the first 3D key point information including the position information of the 3D key point corresponding to each 2D key point in the world coordinate system.

[00217] In some embodiments, the model reconstruction apparatus 700 further includes a texture model, configured to:

[00218] construct 3D point cloud data of the object to be reconstructed in a world coordinate system based on the scanned images of the multiple viewing angles;

[00219] construct a first mesh model of the object to be reconstructed based on the 3D point cloud data; and

[00220] determine a surface texture of the first mesh model of the object to be reconstructed based on the first mesh model and the scanned images of the multiple viewing angles.

[00221] In some embodiments, the texture module is further configured to:

[00222] acquire a camera parameter corresponding to each of the multiple viewing angles;

[00223] for each viewing angle, determine texture information and depth information of each of multiple scanned pixels based on the texture map and depth map in the scanned image corresponding to the viewing angle, and generate a cloud point set corresponding to the viewing angle based on the camera parameter corresponding to the viewing angle and the texture information and depth information of each scanned pixel, the point cloud set including position information of a scanning point corresponding to each scanned pixel in the world coordinate system; and

[00224] construct the 3D point cloud data of the object to be reconstructed based on the point cloud set corresponding to each viewing angle.

[00225] In some embodiments, the scanned image includes a texture map. The texture module is further configured to:

- [00226] for any model triangular facet in the first mesh model, project the model triangular facet into the texture map of at least one viewing angle, determine a projected triangular facet corresponding to the model triangular facet in each texture map and texture information of the projected triangular facet, and determine texture information of the model triangular facet based on the texture information of each projected triangular facet; and
- [00227] generate the surface texture of the first mesh model based on the texture information of each model triangular facet in the first mesh model.
- [00228] In some embodiments, the texture module is further configured to:
- [00229] for each viewing angle, determine, in the world coordinate system, projected vertex coordinates of the projected triangular facet corresponding to model vertex coordinates of the model triangular facet in the viewing angle based on the model vertex coordinates of the model triangular facet and a camera parameter corresponding to the viewing angle; and
- [00230] determine, in the texture map corresponding to the viewing angle, the projected triangular facet and the texture information of the projected triangular facet based on the projected vertex coordinates of the projected triangular facet in the viewing angle.
- [00231] In some embodiments, the scanned image further includes a depth map. The texture module is further configured to:
- [00232] generate reconstructed depth information corresponding to each projected pixel in the projected triangular facet based on the model vertex coordinates of the model triangular facet and the camera parameter corresponding to the viewing angle;
- [00233] determine original depth information corresponding to each projected pixel in the projected triangular facet based on depth information of each scanned pixel in the depth map;
- [00234] determine a valid pixel in the projected triangular facet based on the reconstructed depth information and the original depth information corresponding to each

projected pixel in the projected triangular facet, a distance between the reconstructed depth information and the original depth information of the valid pixel being less than a preset depth distance threshold; and

[00235] determine the texture information of the projected triangular facet based on the texture map and the valid pixel in the projected triangular facet.

[00236] In some embodiments, the adjustment module 703 is further configured to:

[00237] construct a target energy function based on the first 3D key point information, the target energy function being configured to determine a difference between the initial parametric model and the real posture of the object to be reconstructed, the target energy function at least including a first energy function, and the first energy function being configured to characterize a difference between second 3D key point information corresponding to the initial parametric model and the first 3D key point information; and

[00238] adjust the initial parametric model by minimizing the target energy function so as to obtain the target parametric model.

[00239] In some embodiments, the adjustment module 703 is further configured to:

[00240] determine a second mesh model corresponding to the initial parametric model based on an initial model parameter of the initial parametric model;

[00241] determine the second 3D key point information corresponding to the initial parametric model based on the second mesh model; and

[00242] construct the first energy function based on a 3D key point distance between the first 3D key point information and the second 3D key point information.

[00243] In some embodiments, the target energy function further includes at least one of a second energy function or a third energy function. The second energy function is configured to characterize a difference between the second mesh model corresponding to the initial parametric model and a third mesh model corresponding to the target object. The third energy function is configured to constrain a rotation angle of a 3D key point in the initial parameter model.

- [00244] In some embodiments, the target energy function includes the second energy function. The adjustment module 703 is further configured to acquire the third mesh model corresponding to the object to be reconstructed, the third mesh model being configured to characterize a body shape feature of the object to be reconstructed.
- [00245] The adjustment module 703 is further configured to:
- [00246] acquire a second model vertex set corresponding to the second mesh model and a third model vertex set corresponding to the third mesh model, the second model vertex set including 3D coordinates of each second model vertex in the second mesh model, and the third model vertex set including 3D coordinates of each third model vertex in the third mesh model;
- [00247] determine a first distance corresponding to each second model vertex in the third model vertex set and a second distance corresponding to each third model vertex in the second model vertex set based on the 3D coordinates of each third model vertex and the 3D coordinates of each second model vertex; and
- [00248] construct the second energy function based on the first distance corresponding to each second model vertex and the second distance corresponding to each third model vertex.
- [00249] In some embodiments, the target energy function includes the third energy function. The adjustment module 703 is further configured to:
- [00250] acquire a rotation threshold of at least one 3D key point, the rotation threshold including a rotation range of the 3D key point corresponding to each rotation axis; and
- [00251] construct the third energy function based on the rotation threshold of the at least one 3D key point.
- [00252] In some embodiments, the scanned image further includes a depth map. The determination module 702 is further configured to:
- [00253] determine position information of each 2D key point in the texture map of each viewing angle based on the 2D key point information corresponding to the texture map of

each viewing angle;

[00254] for each viewing angle, determine intermediate 3D key point information of each 2D key point in the viewing angle in a world coordinate system based on the position information of each 2D key point in the texture map of the viewing angle and based on the depth map and the camera parameter that correspond to the viewing angle; and

[00255] for each 2D key point, determine position information of a 3D key point corresponding to the 2D key point in the world coordinate system based on the intermediate 3D key point information corresponding to the 2D key point at each viewing angle, the first 3D key point information including the position information of the 3D key point corresponding to each 2D key point in the world coordinate system.

[00256] In some embodiments, the determination module 702 is further configured to:

[00257] acquire a preset action set, the action set including multiple preset actions and second 3D key point information corresponding to each preset action; and

[00258] determine a preset action corresponding to target 3D key point information as a real action of the object to be reconstructed, the target 3D key point information matching the first 3D key point information of the object to be reconstructed.

[00259] The descriptions about the above apparatus embodiment are similar to those about the method embodiment, and beneficial effects similar to those of the method embodiment are achieved. Technical details undisclosed in the apparatus embodiment of the disclosure may be understood with reference to those about the method embodiment of the disclosure.

[00260] It is to be noted that, in the embodiments of the disclosure, the model reconstruction method may be stored in a computer-readable storage medium when implemented in form of a software function module and sold or used as an independent product. Based on such an understanding, the technical solutions of the embodiments of the disclosure substantially or parts making contributions to the related art may be embodied in form of a software product. The computer software product is stored in a storage medium, including a plurality of instructions configured to enable a device to

execute all or part of the method in each embodiment of the disclosure. The storage medium includes various media capable of storing program codes such as a U disk, a mobile hard disk, a Read Only Memory (ROM), a magnetic disk, or an optical disk. Therefore, the embodiments of the disclosure are not limited to any target hardware and software combination.

[00261] FIG. 8 is a schematic diagram of a hardware entity of a model reconstruction device according to an embodiment of the disclosure. As shown in FIG. 8, the hardware entity of the model reconstruction device 800 includes a processor 801 and a memory 802. The memory 802 stores a computer program capable of running in the processor 801. The processor 801 executes the program to implement the steps in the method of any above-mentioned embodiment.

[00262] The memory 802 stores the computer program capable of running in the processor 801. The memory 802 is configured to store an instruction and application executable for the processor 801, may further cache data (for example, image data, audio data, voice communication data, and video communication data) to be processed or having been processed by the processor 1201 and each module in the model reconstruction device 800, and may be implemented through a FLASH or a Random Access Memory (RAM).

[00263] The processor 801 executes the program to implement the steps of any above-mentioned model reconstruction method. The processor 801 usually controls overall operations of the model reconstruction device 800.

[00264] An embodiment of the disclosure provides a computer storage medium, which stores one or more programs capable of being executed by one or more processors to implement the steps of the model reconstruction method in any above-mentioned embodiment.

[00265] It is to be pointed out here that the descriptions about the above storage medium and device embodiments are similar to those about the method embodiment, and beneficial effects similar to those of the method embodiment are achieved. Technical details undisclosed in the storage medium and device embodiments of the disclosure are understood with reference to those about the method embodiment of the disclosure.

[00266] The processor may be at least one of an Application Specific Integrated Circuit (ASIC), a Digital Signal Processor (DSP), a Digital Signal Processing Device (DSPD), a Programmable Logic Device (PLD), a Field Programmable Gate Array (FPGA), a Central Processing unit (CPU), a controller, a microcontroller, or a microprocessor. It can be understood that other electronic devices may be configured to realize functions of the processor, and no specific limits are made in the embodiments of the disclosure.

[00267] The computer storage medium/memory may be a memory such as a ROM, a Programmable Read-Only Memory (PROM), an Erasable Programmable Read-Only Memory (EPROM), an Electrically Erasable Programmable Read-Only Memory (EEPROM), a Ferromagnetic Random Access Memory (FRAM), a flash memory, a magnetic surface memory, an optical disk, or a Compact Disc Read-Only Memory (CD-ROM), or may be any terminal including one or any combination of the above-mentioned memories, such as a mobile phone, a computer, a tablet device, and a personal digital assistant.

[00268] It is to be understood that "one embodiment" or "an embodiment" or "the embodiment of the disclosure" or "the above-mentioned embodiment" or "some embodiments" mentioned in the whole specification means that target features, structures or characteristics related to the embodiment are included in at least one embodiment of the disclosure. Therefore, "in one embodiment" or "in an embodiment" or "the embodiment of the disclosure" or "the above-mentioned embodiment" or "some embodiments" appearing anywhere in the whole specification does not always refer to the same embodiment. In addition, these target features, structures or characteristics may be combined in one or more embodiments freely as appropriate. It is to be understood that, in each embodiment of the disclosure, the magnitude of the sequence number of each process does not mean an execution sequence, and the execution sequence of each process should be determined by its function and an internal logic and should not form any limit to an implementation process of the embodiments of the disclosure. The sequence numbers of the embodiments of the disclosure are adopted not to characterize superiority-inferiority of the embodiments but only for description.

[00269] If not specified, when the model reconstruction device executes any step in the

embodiments of the disclosure, the processor of the model reconstruction device executes the step. Unless otherwise specified, the sequence of execution of the following steps by the model reconstruction device is not limited in the embodiments of the disclosure. In addition, the same method or different methods may be used to process data in different embodiments. It is also to be noted that any step in the embodiments of the disclosure may be executed independently by the model reconstruction device, namely the model reconstruction device may execute any step in the above-mentioned embodiments independent of the execution of the other steps.

[00270] In some embodiments provided by the disclosure, it is to be understood that the disclosed device and method may be implemented in another manner. The device embodiment described above is only schematic. For example, the division of the units is only logic function division, and other division manners may be adopted during practical implementation. For example, multiple units or components may be combined or integrated into another system, or some characteristics may be neglected or not executed. In addition, coupling or direct coupling or communication connection between each displayed or discussed component may be indirect coupling or communication connection, implemented through some interfaces, of the device or the units, and may be electrical and mechanical or adopt other forms.

[00271] The units described as separate parts may or may not be physically separated. Parts displayed as units may or may not be physical units, namely located in the same place or distributed to multiple network units. Part of all of the units may be selected according to a practical requirement to achieve the purposes of the solutions of the embodiments.

[00272] In addition, each function unit in each embodiment of the disclosure may be integrated into a processing unit. Alternatively, each unit may serve as an independent unit. Alternatively, two or more than two units may be integrated into a unit. The integrated unit may be implemented in a hardware form and may be implemented in form of hardware and software function unit.

[00273] The methods disclosed in some method embodiments provided in the disclosure may be freely combined without conflicts to obtain new method embodiments.

[00274] The characteristics disclosed in some product embodiments provided in the disclosure may be freely combined without conflicts to obtain new product embodiments.

[00275] The characteristics disclosed in some method or device embodiments provided in the disclosure may be freely combined without conflicts to obtain new method embodiments or device embodiments.

[00276] Those of ordinary skill in the art should know that all or part of the steps of the method embodiment may be implemented by related hardware instructed through a program. The program may be stored in a computer-readable storage medium. When the program is executed, the steps of the method embodiment are executed. The storage medium includes various media capable of storing program codes such as a mobile storage device, a ROM, a magnetic disk, or an optical disc.

[00277] Alternatively, the integrated unit of the disclosure may be stored in a computer-readable storage medium when implemented in form of a software function module and sold or used as an independent product. Based on such an understanding, the technical solutions of the embodiments of the disclosure substantially or parts making contributions to the related art may be embodied in form of a software product. The computer software product is stored in a storage medium, including a plurality of instructions configured to cause a computer device (which may be a personal computer, a model reconstruction device, a network device, etc.) to execute all or part of the method in each embodiment of the disclosure. The storage medium includes various media capable of storing program codes such as a mobile hard disk, a ROM, a magnetic disk, or an optical disc.

[00278] In the embodiments of the disclosure, the descriptions about the same steps and the same contents in different embodiments may refer to those in the other embodiments. In the embodiments of the disclosure, term "and" does not influence the sequence of the steps.

[00279] The above is only the implementation mode of the disclosure and not intended to limit the scope of protection of the disclosure. Any variations or replacements apparent to those skilled in the art within the technical scope disclosed by the disclosure shall fall within the scope of protection of the disclosure. Therefore, the scope of protection of the

disclosure shall be subject to the scope of protection of the claims.

CLAIMS

1. A model reconstruction method, comprising:

acquiring scanned images of multiple viewing angles corresponding to an object to be reconstructed;

determining first Three-Dimensional (3D) key point information of the object to be reconstructed based on the scanned images of the multiple viewing angles; and

adjusting an initial parametric model based on the first 3D key point information of the object to be reconstructed to obtain a target parametric model characterizing a real posture of the object to be reconstructed.

2. The method of claim 1, wherein the scanned image comprises a texture map, and determining the first 3D key point information of the object to be reconstructed based on the scanned images of the multiple viewing angles comprises:

acquiring a camera parameter and the texture map that correspond to each of the multiple viewing angles;

performing Two-Dimensional (2D) key point detection on the texture map of each viewing angle to obtain 2D key point information corresponding to the texture map of each viewing angle; and

determining the first 3D key point information of the object to be reconstructed based on the camera parameter corresponding to each viewing angle and the 2D key point information corresponding to the texture map of each viewing angle.

3. The method of claim 2, wherein determining the first 3D key point information of the object to be reconstructed based on the camera parameter corresponding to each viewing angle and

the 2D key point information corresponding to the texture map of each viewing angle comprises:

determining position information of each 2D key point in the texture map of each viewing angle based on the 2D key point information corresponding to the texture map of each viewing angle; and

for each 2D key point, determining position information of a 3D key point corresponding to the 2D key point in a world coordinate system based on the position information of the 2D key point in the texture map of each viewing angle and the camera parameter corresponding to each viewing angle, the first 3D key point information comprising the position information of the 3D key point corresponding to each 2D key point in the world coordinate system.

4. The method of claim 1, wherein the scanned image comprises a depth map, and the method further comprises:

constructing 3D point cloud data of the object to be reconstructed in a world coordinate system based on the scanned images of the multiple viewing angles;

constructing a first mesh model of the object to be reconstructed based on the 3D point cloud data; and

determining a surface texture of the first mesh model of the object to be reconstructed based on the first mesh model and the scanned images of the multiple viewing angles.

5. The method of claim 4, wherein constructing the 3D point cloud data of the object to be reconstructed in the world coordinate system based on the scanned images of the multiple viewing angles comprises:

acquiring a camera parameter corresponding to each of the multiple viewing angles;

for each viewing angle, determining depth information of each of multiple scanned pixels based on the depth map in the scanned image corresponding to the viewing angle, and

generating a cloud point set corresponding to the viewing angle based on the camera parameter corresponding to the viewing angle and the depth information of each scanned pixel, the point cloud set comprising position information of each scanned pixel in the world coordinate system; and

constructing the 3D point cloud data of the object to be reconstructed based on the point cloud set corresponding to each viewing angle.

6. The method of claim 4, wherein the scanned image further comprises a texture map, and determining the surface texture of the first mesh model of the object to be reconstructed based on the first mesh model and the scanned images of the multiple viewing angles comprises:

for any model triangular facet in the first mesh model, projecting the model triangular facet into the texture map of at least one viewing angle, determining a projected triangular facet corresponding to the model triangular facet in each texture map and texture information of the projected triangular facet, and determining texture information of the model triangular facet based on the texture information of each projected triangular facet; and

generating the surface texture of the first mesh model based on the texture information of each model triangular facet in the first mesh model.

7. The method of claim 6, wherein projecting the model triangular facet into the texture map of at least one viewing angle and determining the projected triangular facet corresponding to the model triangular facet in each texture map and the texture information of the projected triangular facet comprises:

for each viewing angle, determining, in the world coordinate system, projected vertex coordinates of the projected triangular facet corresponding to model vertex coordinates of the model triangular facet in the viewing angle based on the model vertex coordinates of the model triangular facet and a camera parameter corresponding to the viewing angle; and

determining, in the texture map corresponding to the viewing angle, the projected

triangular facet and the texture information of the projected triangular facet based on the projected vertex coordinates of the projected triangular facet in the viewing angle.

8. The method of claim 7, wherein the scanned image further comprises a texture map, and determining, in the texture map corresponding to the viewing angle, the projected triangular facet and the texture information of the projected triangular facet based on the projected vertex coordinates of the projected triangular facet in the viewing angle comprises:

generating reconstructed depth information corresponding to each projected pixel in the projected triangular facet based on the model vertex coordinates of the model triangular facet and the camera parameter corresponding to the viewing angle;

determining original depth information corresponding to each projected pixel in the projected triangular facet based on depth information of each pixel in the depth map;

determining a valid pixel in the projected triangular facet based on the reconstructed depth information and the original depth information corresponding to each projected pixel in the projected triangular facet, a distance between the reconstructed depth information and the original depth information of the valid pixel being less than a preset depth distance threshold; and

determining the texture information of the projected triangular facet based on the texture map and the valid pixel in the projected triangular facet.

9. The method of claim 1, wherein adjusting the initial parametric model based on the first 3D key point information of the object to be reconstructed to obtain the target parametric model characterizing the real posture of the object to be reconstructed comprises:

constructing a target energy function based on the first 3D key point information, the target energy function being configured to determine a difference between the initial parametric model and the real posture of the object to be reconstructed, the target energy function at least comprising a first energy function, and the first energy function being configured to characterize

a difference between second 3D key point information corresponding to the initial parametric model and the first 3D key point information; and

adjusting the initial parametric model by minimizing the target energy function so as to obtain the target parametric model.

10. The method of claim 9, wherein the first energy function is constructed by:

determining a second mesh model corresponding to the initial parametric model based on an initial model parameter of the initial parametric model;

determining the second 3D key point information corresponding to the initial parametric model based on the second mesh model; and

constructing the first energy function based on a 3D key point distance between the first 3D key point information and the second 3D key point information.

11. The method of claim 9 or 10, wherein the target energy function further comprises at least one of a second energy function or a third energy function, the second energy function being configured to characterize a difference between the second mesh model corresponding to the initial parametric model and a third mesh model corresponding to the target object, and the third energy function being configured to constrain a rotation angle of a 3D key point in the initial parameter model.

12. The method of claim 11, wherein the target energy function comprises the second energy function; and the method further comprises: acquiring the third mesh model corresponding to the object to be reconstructed, the third mesh model being configured to characterize a body shape feature of the object to be reconstructed,

wherein the second energy function is constructed by:

acquiring a second model vertex set corresponding to the second mesh model and a third model vertex set corresponding to the third mesh model, the second model vertex set comprising 3D coordinates of each second model vertex in the second mesh model, and the third model vertex set comprising 3D coordinates of each third model vertex in the third mesh model;

determining a first distance corresponding to each second model vertex in the third model vertex set and a second distance corresponding to each third model vertex in the second model vertex set based on the 3D coordinates of each third model vertex and the 3D coordinates of each second model vertex; and

constructing the second energy function based on the first distance corresponding to each second model vertex and the second distance corresponding to each third model vertex.

13. The method of claim 11, wherein the target energy function comprises the third energy function, and the third energy function is constructed by:

acquiring a rotation threshold of at least one 3D key point, the rotation threshold comprising a rotation range of the 3D key point corresponding to each rotation axis; and

constructing the third energy function based on the rotation threshold of the at least one 3D key point.

14. The method of claim 2, wherein the scanned image further comprises a depth map, and determining the first 3D key point information of the object to be reconstructed based on the camera parameter corresponding to each viewing angle and the 2D key point information corresponding to the texture map of each viewing angle comprises:

determining position information of each 2D key point in the texture map of each viewing angle based on the 2D key point information corresponding to the texture map of each viewing angle; and

for each viewing angle, determining intermediate 3D key point information of each 2D key point in the viewing angle in a world coordinate system based on the position information of each 2D key point in the texture map of the viewing angle and based on the depth map and the camera parameter that correspond to the viewing angle; and

for each 2D key point, determining position information of a 3D key point corresponding to the 2D key point in the world coordinate system based on the intermediate 3D key point information corresponding to the 2D key point at each viewing angle, the first 3D key point information comprising the position information of the 3D key point corresponding to each 2D key point in the world coordinate system.

15. The method of any one of claims 1-14, further comprising:

acquiring a preset action set, the action set comprising multiple preset actions and second 3D key point information corresponding to each preset action; and

determining a preset action corresponding to target 3D key point information as a real action of the object to be reconstructed, the target 3D key point information matching the first 3D key point information of the object to be reconstructed.

16. A model reconstruction apparatus, comprising:

an acquisition module, configured to acquire scanned images of multiple viewing angles corresponding to an object to be reconstructed;

a determination module, configured to determine first Three-Dimensional (3D) key point information of the object to be reconstructed based on the scanned images of the multiple viewing angles; and

an adjustment module, configured to adjust an initial parametric model based on the first 3D key point information of the object to be reconstructed to obtain a target parametric model characterizing a real posture of the object to be reconstructed.

17. A model reconstruction device, comprising a memory and a processor, wherein
- the memory stores a computer program capable of running in the processor; and
- the processor executes the computer program to implement the steps in the method of any one of claims 1-15.
18. A computer storage medium, storing one or more programs capable of being executed by one or more processors to implement the steps in the method of any one of claims 1-15.