SIX-DEGREES-OF-FREEDOM OBJECT PERCEPTION SYSTEM AND METHOD

FIELD OF THE INVENTION

[0001] The present disclosure relates to an object perception system and method, and more particularly to a six-degrees-of-freedom (6DoF) object perception system and method.

BACKGROUND OF THE INVENTION

[0002] The object perception system can be used to pick up boxes. After the system completes the box objection perception (including steps such as identifying and locating a box), the picking task of the box is performed by a robot arm or performed manually according to the identification result. Then, the box is picked up from the pallet or shelf and is moved to the designated places. However, since the position and size of the box may not be fixed, it is difficult for the system to obtain the complete information of the position and size of the box. Therefore, the accuracy of the perception system is decreased, causing the failure of the picking task of the box.

[0003] Existing technology can utilize a 3D camera to take 3D photos of the box, or a 3D scanner to scan the box, thereby identifying the position and size of the box and improving the accuracy of the identification system. However, the existing technology relies on high-resolution 3D cameras or scanners, which costs high. If a general camera is utilized to take photos of the box for price consideration, the photos taken only have the information of one plane of the box, which is unhelpful for improving the ability for identification.

[0004] Therefore, there is a need of providing a six-degrees-of-freedom (6DoF) object perception system and method to obviate the drawbacks encountered from the prior arts.

SUMMARY OF THE INVENTIOND

[0005] It is an object of the present disclosure to provide a 6DoF object perception system and method. In the 6DoF object perception system and method, depth data of the RGBD image and the RGBD image of the object are utilized to estimate the object plane and the corners respectively after the contour of the object has been generated, so as to generate a 6DoF of the object with both orientation and coordinate information. Thus, the position for picking the object can be determined accordingly. By obtaining the 6DoF of the object, the accuracy of the object perception system can be improved. In addition, since the present disclosure obtain the 6DoF of the object through any RGBD image instead of a high-resolution 3D camera or 3D scanner, the cost of the perception system is reduced.

[0006] In accordance with an aspect of the present disclosure, there is provided a 6DoF object perception system. The 6DoF object perception system includes an object determination device, contour detection device and a pose estimation device. The object determination device is configured for receiving an RGBD image and determining whether there is an object in the RGBD image, and the RGBD image includes a depth data. The contour detection device is connected to the object determination device, the contour detection device performs a contour detection on the RGBD image to obtain a contour of the object. The pose estimation device is connected to the contour detection device. The pose estimation device detects a plurality of corners of the object based on the contour of

the object to obtain a coordinate information of the plurality of corners. The pose estimation device extracts a plane of the object based on the depth data to obtain an orientation information of the plane. The pose estimation device generates a 6DoF of the object according to the coordinate information and the orientation information. The 6DoF of the object is used to determine at least one picking position for picking the object.

[0007] In accordance with an aspect of the present disclosure, there is provided a 6DoF object perception method including steps of: (a) obtaining a RGBD image, wherein the RGBD image includes a depth data; (b) determining whether there is an object in the RGBD image through an object determination device, wherein a step (c) is performed if the determination result of the step (b) is positive, and the step (a) is performed again if the determination result of the step (b) is negative; (c) performing a contour detection on the RGBD image to obtain a contour of the object through a contour detection device; (d) detecting a plurality of corners of the object based on the contour of the object to obtain a coordinate information of the plurality of corners through a pose estimation device; (e) extracting a plane of the object based on the depth data to obtain an orientation information of the plane through the pose estimation device; (f) generating a 6DoF of the object according to the coordinate information and the orientation information through the pose estimation device; and (g) utilizing the 6DoF of the object to determine a picking up position through the pose estimation device, wherein the picking up position is utilized for picking up the object.

[0008] The above contents of the present invention will become more readily apparent to those ordinarily skilled in the art after reviewing the following detailed description and accompanying drawings, in which:

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] FIG. 1 is a schematic system diagram illustrating a 6DoF object perception system 1 according to an embodiment of the present disclosure;

[0010] FIG. 2 schematically shows an RGBD image taken by the camera 2;

[0011] FIG. 3 is a schematic system diagram illustrating a 6DoF object perception system 1 according to another embodiment of the present disclosure;

[0012] FIG. 4 is a schematic system diagram illustrating a 6DoF object perception system 1 according to another embodiment of the present disclosure; and

[0013] FIG. 5 is a schematic flow chart illustrating a 6DoF object perception method according to an embodiment of the present disclosure.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENT

[0014] The present disclosure will now be described more specifically with reference to the following embodiments. It is to be noted that the following descriptions of preferred embodiments of this disclosure are presented herein for purpose of illustration and description only. It is not intended to be exhaustive or to be limited to the precise form disclosed.

[0015] FIG. 1 is a schematic system diagram illustrating a 6DoF object perception system 1 according to an embodiment of the present disclosure. FIG. 2 schematically shows an RGBD image taken by the camera 2. As shown in FIGs. 1 and 2, the 6DoF object perception system 1 includes an object determination device 3, a contour detection device 4 and a pose estimation device 5. In an embodiment, a camera 2 is connected to the 6DoF object perception system 1, and the camera 2 is

configured for generating and providing a RGBD image to the object determination device 3. The camera 2 can be externally connected to the 6DoF object perception system 1. Alternatively, in another embodiment, the camera 2 can be one component of the 6DoF object perception system 1. The RGBD image includes a depth data. Specifically, the RGBD image can be image frame pair which includes a monocular image frame (e.g., RGB image) and a respective aligned depth image frame.

[0016] The object determination device 3 is connected to the camera 2 and is configured for receiving the RGBD image and determining whether there is an object 6 in the RGBD image. In an embodiment, the object determination device 3 can detect the object 6 by utilizing only a part of the RGBD image (e.g., the RGB image). The object 6 is for example but not limited to a box.

[0017] The contour detection device 4 is connected to the object determination device 3, and the contour detection device 4 is configured to perform a contour detection on the RGBD image to obtain a contour of the object 6 under the situation that the object 6 exists in the RGBD image. In an embodiment, the contour detection device 4 can apply a model trained and built by a machine learning algorithm to estimate the contour of the object 6, where the model can be stored in the contour detection device 4.

[0018] The pose estimation device 5 is connected to the contour detection device 4. The pose estimation device 5 is configured to detect a plurality of corners of the object 6 based on the contour of the object 6 to obtain a coordinate information of the plurality of corners. The pose estimation device 5 is configured to extract a plane of the object 6 based on the RGBD image with depth data to obtain an orientation information of the plane. The pose estimation device 5 is configured to generate a 6DoF

of the object 6 according to the coordinate information and the orientation information. To be noted that the 6DoF of the object 6 can be used to determine at least one picking position for picking the object 6.

[0019] Please refer to FIG. 3, in an embodiment, the pose estimation device 5 includes a corner detection device 51, a plane recognition device 52 and a calculation device 53. The corner detection device 51 is connected to the contour detection device 4 and performs a corner detection on the object 6, based on the contour of the object 6, to detect the plurality of corners of the object 6. The plane recognition device 52 is connected to the contour detection device 4 and performs a plane recognition on the object 6 based on the depth data and the contour of the object 6 to extract the plane of the object 6. The calculation device 53 is connected to the corner detection device 51 and the plane recognition device 52. The calculation device 53 calculates the coordinate information based on the plurality of corners of the object 6, calculates the orientation information based on the plane of the object 6, and generates the 6DoF of the object 6 according to the coordinate information and the orientation information. The 6DoF include three translational degrees of freedom and three rotational degrees of freedom. The three translational degrees of freedom include forward and backward movement along the Z axis, left and right movement along the X axis, and up and down movement along the Y axis. The three rotational degrees of freedom include rotation around the X axis, rotation around the Y axis, and rotation around the Z axis. The calculation device 53 utilizes the 6DoF of the object 6 to determine a picking position that can be utilized for picking the object 6. In an embodiment, the calculation device 53 calculates a reference center point according to the coordinate information.

[0020] Accordingly, the present disclosure can utilize the RGB image (or RGB data) of the RGBD image to recognize the object 6 and obtain the corner features of the object 6, and can further utilize the depth data corresponding to the RGB image to extract the plane features of the object 6. The corner features can be used to estimate the translation in 6DoF, and the plane features can be used to estimate the other three, i.e., the rotation, in 6DoF, and finally the 6DoF of the object 6 can be obtained. By obtaining the 6DoF of the object 6, the accuracy of the object perception system 1 is improved. In addition, since the present disclosure can obtain the 6DoF of the object through any RGBD image instead of a high-resolution 3D camera or 3D scanner, the cost of the perception system is reduced. Furthermore, compared to the existing technology that requires preset template for identifying the position and size of the box, the present disclosure can obtain the 6DoF of the object through the RGBD image without preset template.

[0021] In an embodiment, the object determination device 3 can use a machine learning algorithm for determining whether there is the object 6 in the RGBD image, and the machine learning algorithm trains a way of determining the object 6 stored in the object determination device 3. The object determination device 3 determines whether there is an object 6 in the RGBD image based on the way of determining the object 6 trained by the machine learning algorithm.

[0022] In an embodiment, the object 6 is disposed on a pallet or a shelf. After the picking position of the object 6 is obtained by the mentioned 6DoF object perception system 1, the user or the robot arm then picks the object 6 according to the picking position and moves the object 6 from the pallet or shelf to a designated place. In an embodiment, the 6DoF object perception system 1 is disposed in the robot arm, and the robot

arm has a movement device. After the robot arm picks the object 6, the robot arm is moved to the designated place through the movement device, and the robot arm puts the object 6 on the designated pallet or shelf.

[0023] In an embodiment, the RGBD image includes a plurality of objects 6. After the calculation device 53 determines the picking up positions of the plurality of objects 6, the calculation device 53 numbers or encodes the plurality of objects 6 respectively, so as to identify different objects 6.

[0024] In an embodiment, the 6DoF object perception system 1 is connected to a chat generation pretrained transformer (CHATGPT), the chat generation pretrained transformer receives a picking up command and transforms the picking up command to the designated number of the object 6 required to be picked up so that the robot arm picks up the corresponding object 6.

[0025] In an embodiment, as shown in FIG. 3, the 6DoF object perception system 1 further includes a first equipment 71 and a second equipment 72. The object determination device 3 is disposed in a first equipment 71, and the contour detection device 4, the corner detection device 51, the plane recognition device 52 and the calculation device 53 are disposed in a second equipment 72. In this embodiment, the first equipment 71 can apply the machine learning algorithm, so the object determination device 3 in the first equipment 71 may determine whether there is an object 6 in the RGBD image based on the way of determining the object 6 trained by the machine learning algorithm. In another embodiment shown in FIG. 4, compared to the embodiment of FIG. 3, the contour detection device 4 is disposed in the first equipment 71 rather than the second equipment 72. Under this circumstance, the contour detection device 4 in the first equipment 71 is allowed to perform the contour

detection on the RGBD image based on the way of detecting the contour trained by the machine learning algorithm.

[0026] FIG. 5 is a schematic flow chart illustrating a 6DoF object perception method according to an embodiment of the present disclosure. The 6DoF object perception method of the present disclosure is applicable for the 6DoF object perception system 1 stated above. Please refer to FIG. 2. The 6DoF object perception method of the present disclosure includes steps S1, S2, S3, S4, S5, S6 and S7. In the step S1, the RGBD image is obtained, and the RGBD image includes a depth data. In the step S2, whether there is an object 6 in the RGBD image is determined through the object determination device 3, the step S3 is performed if the determination result of the step S2 is positive, and the step S1 is performed again if the determination result of the step S2 is negative. In the step S3, the contour detection is performed on the RGBD image to obtain the contour of the object 6 through the contour detection device 4. In the step S4, a plurality of corners of the object 6 is detected based on the contour of the object 6 to obtain a coordinate information of the plurality of corners through a pose estimation device 5. In the step S5, a plane of the object 6 is extracted based on the depth data of the RGBD image to obtain an orientation information of the plane through the pose estimation device 5. In the step S6, a 6DoF of the object 6 is generated according to the coordinate information and the orientation information through the pose estimation device 5. In the step S7, the 6DoF of the object 6 is utilized to determine the picking up position through the pose estimation device 5, and the picking up position is utilized for picking up the object 6.

[0027] In an embodiment, the 6DoF object perception method includes a step of: utilizing a chat generation pretrained transformer to

receive a picking up command and to transform the picking up command to the designated number of the object 6 required to be picked up.

[0028] From the above descriptions, the present disclosure provides a 6DoF object perception system and method. By obtaining the 6DoF of the object, the accuracy of the object perception system is improved. In addition, since the present disclosure can obtain the 6DoF of the object through any RGBD image instead of a high-resolution 3D camera or 3D scanner, the cost of the perception system is reduced.

[0029] While the disclosure has been described in terms of what is presently considered to be the most practical and preferred embodiments, it is to be understood that the disclosure need not be limited to the disclosed embodiment. On the contrary, it is intended to cover various modifications and similar arrangements included within the spirit and scope of the appended claims which are to be accorded with the broadest interpretation so as to encompass all such modifications and similar structures.

WHAT IS CLAIMED IS:

1. A 6DoF object perception system, comprising:

an object determination device, configured for receiving an RGBD image and determining whether there is an object in the RGBD image, wherein the RGBD image comprises a depth data;

a contour detection device, connected to the object determination device, wherein the contour detection device performs a contour detection on the RGBD image to obtain a contour of the object; and

a pose estimation device, connected to the contour detection device, wherein the pose estimation device detects a plurality of corners of the object based on the contour of the object to obtain a coordinate information of the plurality of corners, the pose estimation device extracts a plane of the object based on the depth data to obtain an orientation information of the plane, and the pose estimation device generates a 6DoF of the object according to the coordinate information and the orientation information,

wherein the 6DoF of the object is used to determine at least one picking position for picking the object.

2. The 6DoF object perception system according to claim 1, wherein the pose estimation device comprises:

a corner detection device, connected to the contour detection device, wherein the corner detection device performs a corner detection on the object based on the contour of the object to detect the plurality of corners of the object;

a plane recognition device, connected to the contour detection device, wherein the plane recognition device performs a plane recognition on the object based on the depth data and the contour of the

object to extract the plane of the object; and

a calculation device, connected to the corner detection device and the plane recognition device, wherein the calculation device calculates the coordinate information based on the plurality of corners of the object, calculates the orientation information based on the plane of the object, and generates the 6DoF of the object according to the coordinate information and the orientation information.

3. The 6DoF object perception system according to claim 2, wherein the calculation device further calculates a reference center point according to the coordinate information.
4. The 6DoF object perception system according to claim 1, further comprising:

a camera, connected to the object determination device and configured for generating and providing the RGBD image to the object determination device.

5. The 6DoF object perception system according to claim 1, wherein the contour detection device applies a model trained and built by a machine learning algorithm to estimate the contour of the object, the model is stored in the contour detection device.
6. The 6DoF object perception system according to claim 1, wherein the object determination device uses a machine learning algorithm for determining whether there is the object in the RGBD image, and the machine learning algorithm trains a way of determining the object stored in the object determination device, and the object determination device determines whether there is the object in the RGBD image based on the way of determining the object trained by the machine learning algorithm.

7. The 6DoF object perception system according to claim 1, wherein the RGBD image comprises a plurality of objects, after the calculation device determines the picking up positions of the plurality of objects, the calculation device numbers or encodes the plurality of objects respectively.
8. The 6DoF object perception system according to claim 7, wherein the 6DoF object perception system is connected to a chat generation pretrained transformer, the chat generation pretrained transformer receives a pick command and analyzes the pick command to obtain an object to be picked based on the numbered or decoded objects.
9. A 6DoF object perception method comprising steps of:

(a) obtaining a RGBD image, wherein the RGBD image comprises a depth data;

(b) determining whether there is an object in the RGBD image through an object determination device, wherein a step (c) is performed if the determination result of the step (b) is positive, and the step (a) is performed again if the determination result of the step (b) is negative;

(c) performing a contour detection on the RGBD image to obtain a contour of the object through a contour detection device;

(d) detecting a plurality of corners of the object based on the contour of the object to obtain a coordinate information of the plurality of corners through a pose estimation device;

(e) extracting a plane of the object based on the depth data to obtain an orientation information of the plane through the pose estimation device;

(f) generating a 6DoF of the object according to the coordinate information and the orientation information through the pose

estimation device; and

(g) utilizing the 6DoF of the object to determine a picking up position through the pose estimation device, wherein the picking up position is utilized for picking up the object.

10. The perception method according to claim 9, wherein the contour detection device applies a model trained and built by a machine learning algorithm to estimate the contour of the object, the model is stored in the contour detection device.
11. The perception method according to claim 9, wherein the object determination device uses a machine learning algorithm for determining whether there is the object in the RGBD image, and the machine learning algorithm trains a way of determining the object stored in the object determination device, and the object determination device determines whether there is the object in the RGBD image based on the way of determining the object trained by the machine learning algorithm.
12. The perception method according to claim 9, wherein the RGBD image comprises a plurality of objects, after the calculation device determines the picking up positions of the plurality of objects, the calculation device numbers or encodes the plurality of objects respectively.
13. The perception method according to claim 12, further comprising a step of: utilizing a chat generation pretrained transformer to receive a picking up command and to transform the picking up command to the designated number of the object required to be picked up.