Robotic Arm Grasping System Based on 3D Object Matching

Haikuan Du[1,a,*]

[1]*Donghua University, North Renmin Road, Shanghai, China*
a. hanson_du@163.com

Keywords: 3D Object Matching, Object Detection, Robotic Arm Grasping System.

Abstract: Object grabbing by robotic arms is a research hotspot in current human-computer interaction research. This paper first introduces the overall process framework of the robotic arm grasping system, and then begins to introduce the specific operation process or specific implementation method of each part of the system. The method in this paper first uses a depth camera to reconstruct 3D objects. Then, before using 3D point cloud matching to solve the object pose, the current popular and highly accurate 2D object detection method is used to filter scene point clouds. In this way, it makes the speed of subsequent point cloud matching greatly improved, which also improves the speed and accuracy of the entire grasping system.

1. Introduction

Due to the growing demand for applications, in areas such as medical and industrial, people began to try to use the three-dimensional object to handle specific jobs. For three-dimensional object recognition and matching, the sensor is to analyse and extract the objects in the scene, which is then matched to a known model. Once a match is found, the object model according to the current, converts it to calculate the rotation and translation, and ultimately the object recognition task.

The specific research on point cloud matching is broadly divided into two categories. The first type is rough matching of point clouds. This category is mainly feature-based matching methods. Rusu et al. [1] proposed the PFH (Point Feature Histograms) feature, and uses this feature to complete point cloud matching. Then they proposed FPFH (Point Feature Histograms) [2], which simplified the calculation method and improved the speed of feature calculation.

The second category is precise matching of point clouds. Precise matching is to make the three-dimensional point clouds approach each other in an iterative manner to make the error function smaller, and finally obtain the optimal pose transformation. However, the iterative method can easily fall into the local optimal situation, so the precise matching needs to know the initial pose. The initial pose can be obtained by the rough matching method introduced earlier. The most widely used at this stage is the ICP (Iterative Closest Point) algorithm proposed by Besl et al. In 1992 [3].

Since point cloud registration has some problems mentioned above in practical applications, in order to improve the accuracy of point cloud registration, this paper also uses a two-dimensional target detection method to assist point cloud registration.

This paper designs and implements a robotic arm grabbing system based on point cloud registration. The specific processing is divided into the following five aspects. The first is to reconstruct the complete surface point cloud of the object using 3D reconstruction technology; the second is to use the hand-eye calibration technology to obtain the pose relationship between the camera and the robot arm; the third is the two-dimensional object. The fourth is the three-dimensional point cloud registration, which is used to obtain the pose relationship between the object and the camera; the fifth is to use the pose relationship to control the robot arm for object grabbing.

2. Overall Framework

This section will introduce the overall framework of the robotic arm grasping system. The overall flowchart is shown in Figure 1, which is mainly divided into the following five parts.

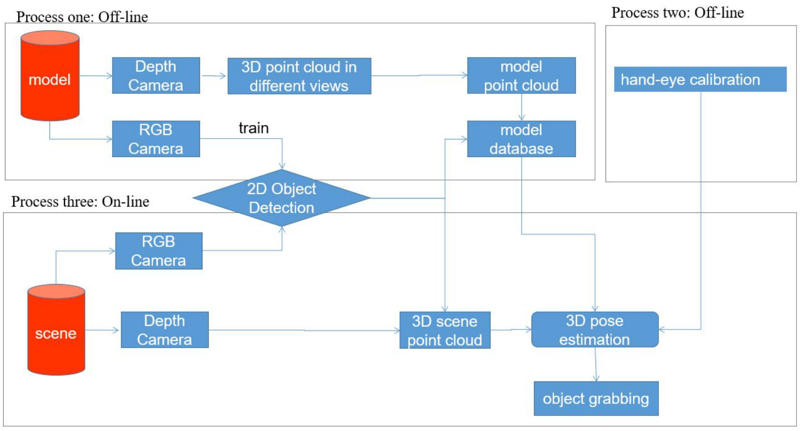

Figure 1: Schematic diagram of the overall process of the grabbing system.

1) Pre-reconstruction based on depth camera. The main purpose of this step is to obtain the complete surface point cloud of the model and to prepare for subsequent 3D point cloud matching.

2) Hand-eye calibration. The purpose of this part is to obtain the posture transformation relationship between the camera and the gripper of the robotic arm.

3) 2D object detection. The purpose of this step is very simple, which is to obtain the bounding box of the object and use this bounding box to cut the point cloud of the actual scene so as to only retain the object point cloud.

4) 3D point cloud matching based on FPFH and ICP. The purpose of this step is to obtain the pose relationship between the object in the actual scene and the camera.

5) Grabbing.

The above is the overall process of the entire robotic arm grasping system. Among them, the training part of 1), 2) and 3) is offline operation, which is the preliminary preparation of the grasping system in this paper. The grasping accuracy is demanding, for the opening range of the electric gripper used in the experiments is 0-70mm. The small opening means that the grasping is difficult.

3. Processing Procedure

3.1. Pre-reconstruction Based on Depth Camera

In this paper, we use Kinect-based 3D point cloud reconstruction method[4], a more mature point cloud reconstruction scheme. Finally, a certain calculation method is used to obtain the point cloud or depth map of the object.

3.2. Hand-Eye Calibration

Hand-eye calibration is the prerequisite for object grabbing. Hand-eye calibration is used to acquire the pose transformation relationship between the camera and the gripper of the robotic arm. Refer to Figure 2 for schematic diagram for hand-eye calibration.

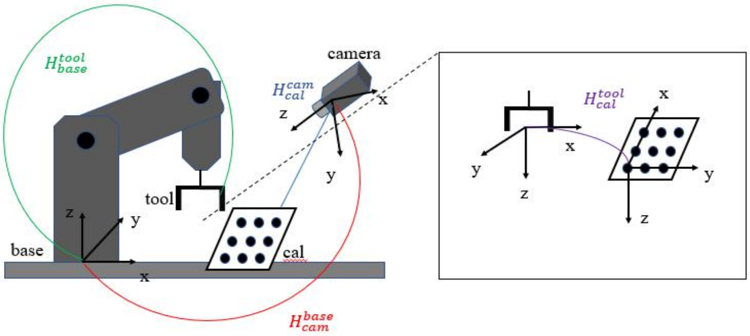

Figure 2: Hand-eye calibration diagram.

3.3. 2D object Detection

This part is to segment the point cloud of the scene, keeping only the point cloud of the object expected to capture. Using 2D object detection method can directly use the position of the object on the 2D image to filter the 3D point cloud and only retain the object point cloud. The experiments in this paper use the SSD[5] object detection framework.

3.4. 3D point Cloud Matching

3D point cloud matching is to get the pose transformation relationship between two point clouds, and overlap the two point clouds after performing the pose transformation.

Point cloud matching is mainly divided into two processes, which are rough matching and fine matching. Among them, the rough matching is performed when the two point clouds do not know the relative position at all. And there are no other known conditions except the two point clouds. It is mainly used to find the approximate rotation translation between the two point clouds. While the precise matching is to further solve a more precise pose relationship when the initial pose relationship is known.

4. Experiment

4.1. Experimental Environment

The robotic arm grasping system environment includes a great deal which will be described in detail below. Due to the use of the 2D detection framework SSD, the GPU used is GTX 1080Ti and the CPU is i7-8700K. The robot arm uses the AUBO i5 robot arm with a payload of 5kg and 6 joint points.

4.2. Experimental Environment

Table 1 shows the recognition rate of SSD network trained in this paper. Considering the actual application process of this article, the precision in this experiment is the IOU of the object detection bounding box and the real object bounding box is greater than 0.5, and the center deviation is no more than 12 pixels.

Table 1: 2D object detection precision.

object	Cans	Yogurt bottle	Skinny toy
SSD	99%	99%	97%

Figure 3 shows the scene where the grabbing system in this paper works. As is seen, the robotic arm grasping system in this paper accurately finds the position of the object, and controls the robotic arm to grasp the object to realize the object grasping function.

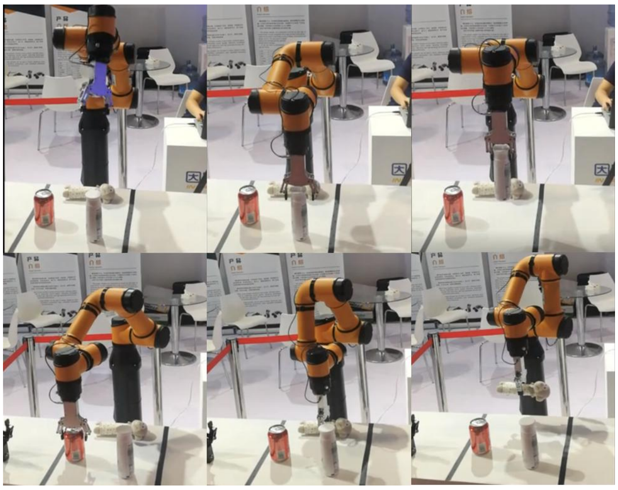

Figure 3: Robot arm object grasping diagram.

5. Conclusions

This paper mainly introduces a 3D matching-based robotic arm grasping system. Through detailed explanation of each part of the system, the entire grasping process is completely displayed.

References

[1] Rusu R B, Marton Z C, Blodow N, et al. Persistent point feature histograms for 3D point clouds[C]//Proc 10th Int Conf Intel Autonomous Syst (IAS-10), Baden-Baden, Germany. 2008: 119-128.

[2] Rusu R B, Blodow N, Beetz M. Fast point feature histograms (FPFH) for 3D registration[C]//2009 IEEE International Conference on Robotics and Automation. IEEE, 2009: 3212-3217.

[3] Besl P J, McKay N D. Method for registration of 3-D shapes[C]//Sensor fusion IV: control paradigms and data structures. International Society for Optics and Photonics, 1992, 1611: 586-606.

[4] Newcombe R A, Izadi S, Hilliges O, et al. Kinectfusion: Real-time dense surface mapping and tracking[C]//ISMAR. 2011, 11(2011): 127-136.

[5] Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector[C]//European conference on computer vision. Springer, Cham, 2016: 21-37.