Research on Adaptive Grasping of Robotic Manipulator with Depth Visual Perception

Wenjiang Liang¹, *, Honglin Zhu², HaiSen Zeng², Yongchuan Xiong¹, Chengya Lu¹ and Luoqian Emu²

¹Southwest Petroleum University, School of Information, 637001 Nanchong, China
²Southwest Petroleum University, School of Engineering, 637001 Nanchong, China

*Corresponding author: 201831772135@stu.swpu.edu.cn

Abstract. Aiming at the problem of poor grasping performance and low repetitive teaching efficiency when traditional machine vision-based robots are facing the grasping tasks of objects with unknown heights, an adaptive grasping method of robotic manipulator based on depth visual perception feedback control is proposed. The RGB-D image is acquired through the depth camera, and the embedded computing device is used to process the RGB-D image based on ROS (Robot Operating System) to obtain the object positioning information. Finally, the coordinate information of the object is used to control the joint motion of the robot to perform the grasping task. The experimental results show that the method is robust and requires a small amount of data, and it can grasp the target object with unknown height without teaching.

1. Introduction
Manipulators are the most widely used automation devices, which is widely used in industry, construction, logistics, medical and other fields [1-3]. With the increasing requirements for production automation and the rise of service robots, how to solve the problem of poor grasping adaptability and the need for multiple teachings when grasping objects with high randomness and different placement heights of the manipulator is crucial for the development of industrial robots [4]. Most of the manipulators used at this stage are teaching-type [5]. They need to perform operation demonstrations before production and record working parameters. Their movement methods are simple, and they can only grab parts of a single specification, which cannot meet the needs of future industrial production, let alone adapting to the complexity of future industrial production and its required flexibility [6]. In this regard, this paper proposes a control method for industrial robots combining with deep visual perception technology to improve grasping efficiency, and conducts research under the ROS framework.

2. Overall system structure
The adaptive grasping manipulator system consists of embedded computing device, depth camera, manipulator, manipulator control box, end effector and monitor screen, as shown in Figure 1. As the eyes of the manipulator, the depth camera is mainly responsible for collecting images. The embedded computing device is the main computing unit and is mainly responsible for processing the image data returned by the depth camera. The control box controls the end effector of the manipulator to grab the target object. Using the Inception_v3 model to train data set [7] and image segmentation algorithm [8] determines whether there is a target object in the image and calculates the three-dimensional coordinate information of the geometric center point of the target object in the image. Then use ROS to create a
topic to publish the data and then perform the conversion from the camera coordinate to the base coordinate of the manipulator. When the end effector of the manipulator moves to the position of the target object, the vacuum gripper is controlled to grab the target object. The depth camera is responsible for collecting color images and depth images [9]. The color images can be processed to identify the object category in the image and calculate the two-dimensional coordinates of the object [10]. The depth image can process the depth information of the target object in the image. The control box is the secondary controller, which is mainly responsible for receiving the three-dimensional coordinate data of the target object in the world coordinate system processed by the embedded computing device, and sending this data to the manipulator joint to control the movement of manipulator.

![Diagram of the adaptive grasping manipulator system](image)

**Figure 1.** The adaptive grasping manipulator system diagram

### 3. Grasping process

The workflow of the system is shown in Figure 2.

![System workflow diagram](image)

**Figure 2.** System workflow

The depth camera acquires the image and identifies the object to be grasped in the image and obtains its two-dimensional pixel coordinates \((x, y)\). The motion of the manipulator makes the depth camera
reach directly above the geometric center point of the object to obtain the depth information from the camera to the geometric center point of the object, and obtain the camera’s three-dimensional coordinates \((x_c, y_c, z_c)\). Since the relative spatial position relationship between the camera and the end effector of the manipulator is known, the spatial position relationship between the end effector of the manipulator and the object and the coordinates of the object in the base coordinate system \((x_w, y_w, z_w)\) can be obtained through coordinate conversion. Through \((x_w, y_w, z_w)\) and the posture \((\alpha, \beta, \gamma)\) of the end effector of the manipulator determined according to the demand and environment, we can calculate the rotation angle of each motor of the manipulator joint and control each motor to rotate until the end effector reaches the grasping point, and finally complete the grasp of the target object.

4. Kinematics analysis of manipulator

The six-degree-of-freedom manipulator used in this article is the Xarm manipulator. According to the description of the link coordinate system above, the coordinate system shown in Figure 3 can be established on the link of this manipulator.

![Xarm manipulator coordinate system](image)

**Figure 3.** Xarm manipulator coordinate system

By observing the relationship between the coordinate systems of each link of the Xarm manipulator, the DH parameter table of the six-degree-of-freedom manipulator used in this article can be obtained, as shown in Table 1.

| link \(i\) | \(\alpha_i(\degree)\) | \(a_i(m)\) | \(d_i(m)\) | \(\theta_i(\degree)\) | Angle limit  |
|------------|-----------------|-----------|-----------|-----------------|--------------|
| 1          | 0               | 0         | 0.2670    | \(\theta_1\)    | [-360, 360]  |
| 2          | -90             | 0         | 0         | \(\theta_2\)    | [-124.9, 124.9] |
| 3          | 0               | 0.3055    | 0         | \(\theta_3\)    | [-360, 360]  |
| 4          | 90              | 0.0775    | 0.3425    | \(\theta_4\)    | [-360, 360]  |
| 5          | 90              | 0         | 0         | \(\theta_5\)    | [-100.2, 180]  |
| 6          | -90             | 0.0760    | 0.0970    | \(\theta_6\)    | [-360, 360]  |

According to the data provided in Table 1, it is easy to deduce the homogeneous transformation matrix between two adjacent link coordinate systems in the Xarm manipulator can be obtained as:
In the above formula, \( i = 1, 2, 3, 4, 5, 6 \). After bringing the data in the table 1 into formula (1), transformation matrices \( ^0T_i, ^1T_i, ^2T_i, ^3T_i, ^4T_i, ^5T_i \) can be obtained. Multiply these 6 matrices one by one, and the result of the multiplication is the pose transformation matrix of coordinate system 6 in coordinate system 0, and the pose transformation matrix is the result of the forward kinematics solution of the six-degree-of-freedom manipulator in this laboratory.

\[
^iT_j = \begin{bmatrix}
\cos \theta_i & -\sin \theta_i & 0 & a_{i-1} \\
\sin \theta_i \cos \alpha_{i-1} & \cos \theta_i \cos \alpha_{i-1} & -\sin \alpha_{i-1} & -d_i \sin \alpha_{i-1} \\
\sin \theta_i \sin \alpha_{i-1} & \cos \theta_i \sin \alpha_{i-1} & \cos \alpha_{i-1} & d_i \cos \alpha_{i-1} \\
0 & 0 & 0 & 1
\end{bmatrix}
\] (1)

In Formula (2), the \( n, o, a \) vectors respectively represent the projections of the coordinate axes of the coordinate system 6 on the coordinate axes of the coordinate system 0, and the \( p \) vector refers to the position of the origin of the coordinate system 6 in the coordinate system 0. When the joint angles of the manipulator are known, the pose of the end effector of the manipulator can be uniquely determined by Formula (2).

\[
^0T = ^0T_1^1T_1^2T_2^3T_3^4T_4^5T_5^6T_6
\] (2)

5. Experimental results and analysis
When the manipulator reaches the designated shooting position, the hand depth camera takes pictures to obtain the RGB-D image of the object to be captured, as shown in Figure 4, and the manipulator grasping process is shown in Figure 5. In order to verify the effectiveness of the depth visual perception method, the experiment provides cylinders, cubes and cuboids of different colors (red, green and blue) and different heights to verify the error.

Exporting the depth data of the target geometric center point and using MATLAB for data processing, the comparative analysis curve at different distances is drawn using the actual value to subtract the measured value and take the absolute value, as shown in Figure 6. It can be seen that when the distance is between 35cm and 45cm, the data fitting effect is better. When the distance exceeds 45cm, the difference between the theoretical value and the actual value gradually increases. Therefore, when measuring from 35cm to 45cm, there is no need to eliminate errors in the measured data.
Figure 6. Comparison of actual value and theoretical value measured by depth camera

Robot joint motion information is collected in teaching mode and adaptive grasping mode respectively. We can obtain the real-time angle array of the manipulator joints, and export the topic data to a text file format, and use MATLAB for data import and curve fitting. The joint position curve of each joint over time in the teaching mode and in the depth visual perception mode is shown in Figure 7.

Figure 7. The angle change of each joint is in teaching mode and adaptive grasping mode

It can be seen from the experimental results that when grasping the same object, the manipulator in the teaching mode has a faster grasping speed than the manipulator based on depth visual perception. But in the teaching mode, the noise robustness is poor, the pertinence is weak, and the adaptability is low. The manipulator with visual perception is more adaptable when performing the task of grasping a large number of objects of unknown height, avoiding multiple teaching.

6. Conclusion
In order to solve the problem of poor grasping adaptability and the need for multiple teaching when grasping objects with freely place and different placement heights, this paper proposes an adaptive grasp control method based on depth visual perception. The robot control system is designed with ROS. The
depth camera obtains the position and pose of the grasping target. Through the inverse kinematics solution, the control variables of each joint of the manipulator are obtained to control the robot to perform adaptive grasping operations. Contrast with traditional teaching methods, the experimental results show that the method is simple and effective, and the manipulator grasping system combined with deep visual perception solves the problem of poor grasping adaptability when grasping objects with high random placement and different placement heights and requiring multiple teachings.

Acknowledgments
This work was supported by the Nanchong City and School Science and Technology Strategic Cooperation Project (Grant No.19SXHZ0041, No.18SXHZ0049), Southwest Petroleum University 2020 National College Student Innovation and Entrepreneurship Training Program Project (Grant No. S202010615010), Southwest Petroleum University 2020 Provincial College Student Innovation and Entrepreneurship Training Program Project (Grant No.S202010615131).

References
[1] Canal, G, Alenyà, Guillem, & Torras, C. (2016). Personalization Framework for Adaptive Robotic Feeding Assistance. International Conference on Social Robotics. Springer International Publishing.
[2] Wilson, J. R, Tickle-Degnen, L, & Scheutz, M. (2016). Designing a Social Robot to Assist in Medication Sorting. International Conference on Social Robotics. Springer, Cham.
[3] Z. Pan, Z. Jia, K. Jing, Y. Ding and Q. Liang, "Manipulator Package Sorting and Placing System Based on Computer Vision," 2020 Chinese Control And Decision Conference (CCDC), Hefei, China, 2020, pp. 409-414, doi: 10.1109/CCDC49329.2020.9164071.
[4] Shi, Z (2017). Research on Motion Control of Robotic Manipulator Based on Visual Perception
[5] Sheng, X, (2018). Research and Implementation of Depth Perception Technology Based on Binocular vision.
[6] Y. Han, K. Zhao, Z. Chu and Y. Zhou, "Grasping Control Method of Manipulator Based on Binocular Vision Combining Target Detection and Trajectory Planning," in IEEE Access, vol. 7, pp. 167973-167981, 2019, doi: 10.1109/ACCESS.2019.2954339.
[7] H. Wang et al., "Research on the Relationship between Classic Denavit-Hartenberg and Modified Denavit-Hartenberg," 2014 Seventh International Symposium on Computational Intelligence and Design, Hangzhou, 2014, pp. 26-29, doi: 10.1109/ISCID.2014.56.
[8] Zhou, H, Lin, J, Liu, L, & Tao, M. (2017). Sensor module design and forward and inverse kinematics analysis of 6-dof sorting transferring robot. Iop Conference, 231, 012091.
[9] Fan, S, Gu, Zhang, Y, Jin, M, & Liu, H. (2018). Research on adaptive grasping with object pose uncertainty by multi-fingered robot hand. International Journal of Advanced Robotic Systems, 15(2), 172988141876678.
[10] Chen, Y, Guo, S, Li, C, Yang, H, & Hao, L. (2018). Size recognition and adaptive grasping using an integration of actuating and sensing soft pneumatic gripper. Robotics and Autonomous Systems, 104, 14-24.