Research on Intelligent Grasping System of Monocular Vision Guided Manipulator

Han Jiabu¹, Chen Saixuan¹, Cui Guohua¹, Cheng Zhenyi¹ and Li Shipei¹

¹School of Mechanical and Automotive Engineering, Shanghai University Of Engineering Science, Shanghai, China

E-mail: 2022653432@qq.com

Abstract. In order to solve the problem of how a cooperative robot can quickly and accurately locate and grasp the specified target work piece and avoid obstacles autonomously under the condition of man-machine cooperation and the variable position of the object to be grasped, a new method of robot arm grasping under the visual guidance of ROS system was proposed. Using ROS framework and tools, the color threshold algorithm in ROS-Opencv is used to realize the rapid recognition and positioning of grasping targets. Use the ROS-Moveit! The kinematics calculation, motion planning and obstacle avoidance of the manipulator are completed the relationship between the image coordinate system and the robot coordinate system is established by hand eye calibration technology. In addition, ROS tools can be used to obtain information such as the position and speed of each joint during the movement of the mechanical arm. Through experimental testing and data analysis, it is shown that the hand-eye system is accurate in calibration, object positioning, and the robot arm can accurately grasp the target object, which has certain practical application value.

1. Introduction

In the current industrial production process, the vast majority of industrial robots rely on the teaching reproduction method to capture and place the workpiece in a fixed position. If the external environment or the state of the workpiece changes, the capture task will be interrupted or fail. In order to solve this problem and improve the flexible production of the factory, the technology of vision-guided robot was produced. The robot uses the visual system to identify and locate the objects in the working space, avoid obstacles autonomously, and at the same time, work in cooperation with human safety. Zhang chi makes use of the VS software and the grasping platform developed by ABB robot[1]. This platform has a long development cycle and great difficulty in development. Xu bofan used the grab system platform of visually guided FANUC Robot Interface developed by FANUC Robot Interface[2]. This platform has poor adaptability and is only suitable for FANUC robots. To solve the above problems, ROS system framework and tools are adopted in this paper to build a gripper system based on UR5 cooperative robot and flexible gripper. Images are collected by the monocular camera, ROS image processing nodes get the image features of the workpiece, and the feature information is converted into positioning information and sent to ros-moveit! The motion planning parameters are sent to the manipulator controller to realize flexible grasping under visual guidance. Improve the code reuse rate of robot development and research, reduce the robot project research and development cycle. And greatly improve the system integration, portability[3,4].
2. System composition

2.1. Hardware system

The hardware of the platform is composed of UR5 cooperative robot, USB camera, flexible gripper, upper computer and lower computer. Hardware composition of the platform is shown in Figure.1.

The camera is installed above the experimental platform and connected to the upper PC through a USB interface. The lower PC communicates with the upper PC through a network port based on the Ethernet protocol. The driving and image processing of the camera and the control of the robot are based on the ROS system in PC. Hardware communication architecture is shown in Figure.2.

2.2. Software system

The whole software system is based on the ROS system[5]. ROS system adopts a distributed computing and communication architecture, providing functions such as hardware abstract description, underlying driver management, inter-program messaging and program distribution package management. This software system mainly includes three modules: camera calibration and hand-eye calibration module, object recognition and positioning module based on ros-opencv and ros-moveit! of the motion planning module. Object feature recognition was developed by ros-opencv computer vision library using C++. After the target object location was obtained, ros-moveit! was used the trajectory planning of the end of the manipulator from the starting point to the target point is completed. The working process of the system is shown in Figure.3.

Camera calibration is carried out to obtain the internal and external parameters of the camera, and the internal and external parameters are imported into the camera driver file and called in the parameter server. All information obtained by the camera are described in the camera coordinate system, through hand-eye calibration technology of the workpiece under the camera coordinate system...
transformation to the robot pose information base coordinates, this experiment adopts the eye-to-hand camera installation, determine the camera coordinate system and robot hand-eye calibration relative relationship between the base coordinate system. After a series of image processing tasks such as filtering, color segmentation and feature extraction, the image acquired by the camera can determine the target object contour to be recognized. Then, according to the central point of the target object contour and the internal and external parameter matrix of the camera, the plane coordinate information of the target object in the camera coordinate system can be obtained. Then, according to the conversion relationship between the camera coordinate system calibrated by hand-eye and the manipulator coordinate system, the spatial position information of the target object can be obtained, and the object information can be transmitted to the ROS system through the ROS communication mechanism for task planning. Ros-moveit! is used according to the planning task. The kinematics library, the kinematics planning library and the obstacle avoidance algorithm are respectively used for inverse kinematics solution, trajectory planning and obstacle avoidance movement. Finally the motion planning of robotic arm motion planning information queue through Ethernet communication mode is passed to the next bit machine and lower machine receives the mechanical arm movement and parses the message queue, at the same time to drive the mechanical arm according to the planning of the path of the movement and scraping, and transfer the real-time position and posture information of the robotic arm to PC, the realization of ultimate target accurately grasping.

3. Principle of visual localization

3.1. Monocular localization model

Inside and outside the camera parameters \((f, u_0, v_0, \alpha_x, \alpha_y, R, T)\) were obtained through the camera calibration experiment[6], and the monocular localization model was established according to the pinhole model[7]. Monocular localization model is shown in Figure 4.

![Figure 4. Monocular localization model](image)

The origin \(o\) of camera coordinate system \(O_X Y Z\) is the optical center of the camera, the height of the camera optical center from the table plane \(XOY\) is fixed, the projection of camera optical center on the table surface coincides with the origin \(o\), so the world coordinate system \(O_X Y Z\) is established with as the origin \(o\), the optical axis intersects the equivalent image plane \(oxy\) at the coordinate value \((u_0, v_0)\) of point \(o\), the distance from \(o\) to \(o\) is the focal length \(f\), point \(P\) is any point on the plane of the carrying object, and its world coordinate value is expressed as \(x, y, z\), \(P\) is its projection position on
the equivalent imaging plane, \( Q, Q \) is the projection of \( P \) and \( P \) on the vertical axis. Frame \( \alpha_j \mu \nu \) is the image pixel coordinate system. \( u \) and \( v \) respectively represent the columns and rows of the image matrix. \( \alpha_z \) and \( \alpha_y \) are the coefficients along the axis \( x \) and \( y \) respectively.

### 3.1.1 Longitudinal positioning

The longitudinal coordinate value \( y_w \), namely the length of \( O_z Q \), of point \( P \) is derived in plane \( O_z O_y \).

\[
y_w = h \tan \angle O_z O_y = h \tan \left( \frac{|OQ|}{f} \right)
\]

In the pixel coordinate system of the image

\[
|OQ| = (v_w - v_o) \alpha_y
\]

The formula can be written as

\[
y_w = h \tan \angle O_z O_y = h \tan \left( \frac{(v_w - v_o) \alpha_y}{f} \right)
\]

### 3.1.2 Lateral positioning

Determine the transverse coordinate of points in plane \( Q \), and get from the proportion relation

\[
x_w = |PQ| \left| \frac{|O_z Q|}{\tan \angle PO_z Q} \right|
\]

\[
|O_z Q| = \frac{h}{\cos \angle O_z O_y} = h / \cos \left[ \arctan \left( \frac{|OQ|}{f} \right) \right]
\]

\[
\tan \angle PO_z Q = \frac{|PQ|}{|O_z Q|} = \frac{|PQ|}{\sqrt{f^2 + (|OQ|)^2}}
\]

Represent \( |OQ| \) and \( |PQ| \) in the image pixel coordinate system

\[
x_w = \frac{h}{\cos \left[ \arctan \left( \frac{(v_w - v_o) \alpha_y}{f} \right) \right]} \cdot \left[ u(P) - u_w \right] \alpha_z \sqrt{f^2 + (v_w - v_o)}
\]

Through the above longitudinal positioning \( y_w \) and lateral positioning \( x_w \) calculation, the coordinates of any point on the plane within the scope of camera vision in the world coordinate system can be obtained, so as to determine the position of the target on the experimental platform.

### 3.2 Hand-eye calibration

According to the application scenarios of the experimental platform[8], the camera of the experimental platform is installed in an eye-to-hand way. The eye-to-hand calibration model is shown in Figure 5.
\[ T = \begin{bmatrix} R & t \end{bmatrix} \] represents the rotational translation matrix in homogeneous coordinates, \( R \) represents the rotation matrix, \( t \) represents the translation vector, \( T_s \) represents the transformation relationship between the calibration plate coordinate system and the camera coordinate system, \( T_c \) represents the transformation relationship between the calibration plate coordinate system and the end coordinate system of the mechanical arm, \( T_e \) represents the transformation relationship between the end coordinate system of the mechanical arm and the coordinate system of the robot base, \( T_r \) represents the transformation relationship between the coordinate system of the robot base and the camera coordinate system, namely the hand-eye transformation matrix to be solved.

According to the transformation relationship between the camera, calibration plate, the end of the mechanical arm and the coordinate system of the robot base in the motion state transition as shown in the figure 6, the following state transition equation is established:

\[
\begin{align*}
(T_s)_{(1)} &= T_c \cdot (T_s)_{(1)} \cdot T_c^\prime \\
(T_s)_{(2)} &= T_c \cdot (T_s)_{(2)} \cdot T_c^\prime \\
&\vdots \\
(T_s)_{(i)} &= T_c \cdot (T_s)_{(i)} \cdot T_c^\prime \\
(T_s)_{(i+1)} &= T_c \cdot (T_s)_{(i+1)} \cdot T_c^\prime \\
&\vdots \\
(T_s)_{(n)} &= T_c \cdot (T_s)_{(n)} \cdot T_c^\prime 
\end{align*}
\]

Then get rid of the \( T_s \)

\[
\begin{align*}
(T_c)_{(i)} \cdot (T_s)^{-1} \cdot T_c^\prime = (T_c)_{(i+1)} \cdot (T_s)_{(i+1)}^{-1} 
\end{align*}
\]

If

\[
\begin{align*}
A &= (T_c)_{(i+1)} \cdot (T_s)^{-1} \\
B &= (T_c)_{(i)} \cdot (T_s)_{(i+1)}^{-1} \\
X &= T_c^\prime
\end{align*}
\]

\( A \) represents the motion relation of the calibration plate coordinate system an \( B \) represents the motion relation of the end coordinate system of the mechanical arm, so solving for is the same thing as solving for \( T_c^\prime \) the unknown \( X \) in equation \( AX = XB \).

According to the formula, to solve the matrix \( X \) at least two relative position transformations of the robot are required, and the rotation axes of these two position changes are not parallel. In order to improve the calibration accuracy and offset the accidental error, this experiment collected the relative pose relationship of 17 groups of mechanical arms. Hand-eye calibration pose transformation of drawing is shown in Figure 7.
In the ROS system, the ROS calibration function package easy_handeye and Opencv are used to write corresponding programs to complete the corner feature extraction of the calibration board and the pose calculation of the camera coordinate system relative to the end-effector coordinate system.

The usb_cam node in the ROS system is used to start the camera and send image messages. Then, the ros-opencv image processing node is used to extract and describe image features. Ros-moveit! is used multiple different positions of the calibration plate at the end of the mechanical arm were planned, and then parameters were calculated based on image feature points through the easy_handeye function package. The position and attitude of the calibration plate relative to the camera coordinate system were obtained, and the hand-eye transformation matrix was obtained to complete the calibration. Hand-eye calibration pose transformation results is shown in Figure 8.

4. Robot motion control

4.1 Target identification and location

We driven camera and get ROS of image information by ROS, using the cv_bridge class into ROS-Opencv image, then Opencv vision library preprocessing, image processing algorithms to use filtering algorithm of median filter was carried out on the original image denoising and according to the experimental platform for fetching in target based on color threshold method for image segmentation.
and edge features based on target for the workpiece position measurement and calculation[9,10]. Object orientation is shown in Figure.9. Finally, the image is converted into a ROS message through the cv_bridge class and the image data is sent to the corresponding ROS topic. Target object contour recognition flow chart is shown in Figure.10.

4.2 Robot grasp control
The UR5 robot motion plan using ROS-Moveit! module[11]. Through Moveit! configuration Assistant (Moveit Setup Assistant Tool) created call URDF mechanical arm model file, and then in turn to collision detection configuration, virtual joint configuration, mechanical arm planning initial position mechanical arm joint group configuration, configuration, mechanical arm end actuator configuration, passive joint configuration, generate Semantic Robot Description. SRDF file, the kinematics and motion planning library configuration file configuration file and the corresponding startup files, in the mission planning to load the corresponding configuration file, the motion planning, visualization and simulation of the robot are completed. Schematic diagram of UR5 grasping results is shown in Figure.11.
According to the working condition information obtained by the camera and the state of the robot arm, the robot arm grabbed the task planning, and loaded the corresponding Moveit! in the process of executing the task configuration assistant to generate configuration files, the use of open source kinematics, motion planning complete library of kinematics calculation and motion planning, and planning of the mechanical arm movement information released to joint server, PC through the Socket communication will further movement information sent to the robot controller and motion information analysis, the driving mechanical arm in accordance with the plan action execution fetching. While executing the planning action, the position sensor on the mechanical arm sends the actual pose of the mechanical arm to the upper computer through the socket.

5. The experimental example
The image pixel coordinates of the center of the upper surface of the object to be captured are obtained through feature extraction of the image of the object to be captured. The pixel coordinates of the center image of the object to be captured are converted into the coordinates under the coordinate system of the robot base by camera calibration and hand-eye calibration technology to complete the object positioning. Ros-Moveit! is used according to the starting point of the end of the robot arm and the coordinate position of the target.

The motion planner is used to obtain the trajectory parameters, which are sent to the robot controller through socket communication and analyzed, and the mechanical arm is driven to complete grasping according to the predetermined action. During the experiment, the object to be grasped was placed in different positions in the photographing area, and the grasping system could successfully complete the visually guided grasping experiment. Visual localization and motion position results is shown in Table 1.

| Serial number | Visual location (x, y) | Motion position (x, y) |
|---------------|------------------------|-----------------------|
| 1             | (-0.335, -0.483)       | (-0.328, -0.480)      |
| 2             | (-0.326, -0.451)       | (-0.315, -0.446)      |
By analyzing the experimental data, it can be concluded that the measurement error of the target position grabbed by the experimental platform is less than 1cm, and the positioning accuracy of the cooperative robot for the target is different in different actuators and working scenes. The flexible gripper grasping mechanism adopted in this experimental platform can complete the grasping action with an error value with-in 1cm. The errors mainly come from camera calibration, hand-eye calibration and repeated positioning of the robot. In the process of camera calibration and hand-eye calibration, there are errors in measurement and installation position of calibration plate. At the same time, the horizontal plane of the object to be grabbed and the base plane of the robot will also affect the positioning accuracy, but the positioning accuracy can meet the application requirements of most intelligent grasping platforms.

6. Conclusion
This paper builds a set of intelligent grasping platform based on ROS system, and using color machine vision algorithm in ROS-Opencv to achieve rapid target recognition. The coordinate transformation of the target is completed through camera calibration and hand-eye calibration, and the precise coordinates of the target in the coordinate system of the robot base are obtained. Using the ROS-Moveit! The trajectory planning of the robot from the initial position to the target position is completed to achieve the accurate capture of the target. This platform has certain portability, only a few configuration files need to be modified, which can be quickly transplanted to different types of ROS supported robots. This experimental platform can accurately locate and grasp the shape rule workpiece, which has certain application value in the actual work scene.

References
[1] Zhang chi, Liao huali. Intelligent grasping system for industrial robot based on monocular vision [J]. Journal of Mechanical & Electrical Engineering.2018,35(3):283-287.
[2] Xu bofan, zhaohuadong. Research on grasping system of FANUC robot based on visual guidance[J]. Modular Machine Tool & Automatic Manufacturing Technique,2018(7):111-114.
[3] ROS/technical overview. http://wiki.ros.org/.
[4] Liu qifan, Xie ming. Industrial robot motion planning based on ROS system[J]. Modular Machine Tool & Automatic Manufacturing Technique,2017(5):36-39.
[5] Lei shi. Vision Guiding System for 6-DoF Robotic Manipulator[D]. Tallinn University of Technology,2017.
[6] Xu de, Tan min, Li yuan. Robot vision measurement and control[M]. Beijing: National Defence Industry Press. 2011.
[7] Wang tianqi, Chen ming. Research on target recognition and location based on monocular vision[J]. Mechanotronics mechanical-electrical integration,2015(11).
[8] Cheng yuli. Hand-eye calibration and object localization for industrial robotic application[D]. Zhejiang University,2016.
[9] ROS/opencv. https://opencv.org/tutorials
[10] Liu xiangde. Research on motion control of Baxter robot based on visual guidance[J]. Journal of Chongqing University of Posts and Telecommunications,2018,30(4):552-557.
[11] Research on robot grasping planning based on RGB-D vision recognition[D]. Harbin Institute of Technology