Fast target recognition and location based on graph model and point cloud model

Yuntao Zhao1,2*, Jiaming Hu2 and Weigang Li1,2

1 Engineering Research Center for Metallurgical Automation and Measurement Technology of Ministry of Education, Wuhan 430081, China
2 School of Information Science and Engineering, Wuhan University of Science and Technology, Wuhan 430081, China

*Corresponding author’s e-mail: zhyt@wust.edu.cn

Abstract. When using the point cloud model to identify and locate the target, there are some problems, such as low recognition accuracy and slow positioning speed. In order to solve these problems, this paper proposes a fast target recognition and location method based on point cloud model and graph model, which combines the characteristics of graph model processing speed, algorithm efficiency and recognition accuracy. Through the analysis of the characteristics of the target graph model, the pixel coordinates and rotation angle of the target grabbing points are calculated. After the transformation matrix and mapping relationship, the spatial coordinates of the point cloud model are obtained. The efficiency and accuracy of the method are verified by the recognition and positioning experiments on the robot target grabbing platform.

1. Introduction

3D point cloud data is widely used in the production of industrial products because it can realize the three-dimensional digitization of objects. However, in the application of target recognition and positioning, the point cloud data model has not been well developed due to its own accuracy problems, complex algorithm design, slow calculation speed and other shortcomings [1-2].

Therefore, the point cloud model is still a focus of research. The literature [3] uses the CAD model to match the point cloud model to identify and locate the target, but this method is only aimed at the simple target with the CAD model, which is not suitable for the complex structure. The literature [4-5] uses the point cloud model to design some recognition algorithms, but the generality of the algorithm is not high, and the algorithm design is complex; In reference [6-7], the point cloud model is fused with lidar data to realize the target positioning, but this method only uses a large point cloud model, and the accuracy of recognition is related to the accuracy of lidar. Therefore, when using point cloud model to identify and locate, we should not only consider the influence of large point cloud data on the processing speed of the algorithm, but also consider the accuracy and accuracy of these complex algorithms for target recognition.

The algorithm design combined with graph model can effectively avoid these problems. Firstly, the data capacity of the graph model for the same target is much smaller than that of the point cloud model, which can effectively shorten the data reading time and algorithm processing time; secondly, for the plane target, some characteristic parameters of the graph model are much more accurate than those of the point cloud model. Therefore, the plane coordinates of the target point can be calculated by the graph
model first, and then the height can be calculated by the coordinate mapping relationship of the point cloud model, and then the spatial coordinates of the target point can be obtained by combining the two.

2. **Robot target grabbing platform**

The robot target grab platform is mainly composed of three systems: three-dimensional vision system, lighting system and robot execution system, as shown in figure 1. A visual system consists of a fixed 3D structured light camera, an industrial control computer and a material collection table. The lighting system uses a standard parallel light source to obtain clear contour features and avoid uncertainties caused by natural illumination. The robot execution system is composed of 6 kg load Harmo six-axis cooperative robot and robot control cabinet.

The vision system obtains the six degree of freedom information of the target through image processing, and sends the spatial coordinate information to the robot through the industrial PC. the robot plans the path according to the target point to grasp. The image processing process is based on OpenCV open source library [8] and PCL open source library [9]. By using the powerful image processing module in opencv library and the point cloud coordinate conversion module in PCL library, the graph model and the point cloud model are combined to achieve the accurate alignment of 3D objects in Cartesian space. Finally, the empty space of the object relative to the end of the robot is calculated through the conversion matrix obtained by hand eye calibration. The coordinate between the two.

![Figure 1. Robot target grabbing platform](image)

3. **Target recognition and location based on graph model and point cloud model**

The general methods to identify and locate the target by using the point cloud model are as follows: first, the whole point cloud data is obtained by 3D camera shooting, then the target point cloud is extracted from the point cloud by filtering, segmentation and other algorithms. Finally, a series of algorithms are used to obtain the 3D coordinate information of the target capture point. Although this method can finally obtain intuitionistic 3D coordinates, the algorithm of point cloud is difficult to design, has no good adaptability and stability, and the processing speed of the algorithm is slow, especially for the point cloud model with large specifications. The final target coordinates are affected by camera accuracy and algorithm accuracy, so they are not suitable for high precision.

The image model can be applied to two-dimensional image recognition and positioning: the image model of the target is obtained through image graying, filtering, denoising, binarization, segmentation and other preprocessing, and then the two-dimensional coordinates of the grabbing point are calculated according to the contour features of the image model. Using graph model to design recognition algorithm not only has high efficiency, but also greatly improves the expansibility and adaptability of the target,
which makes up for the shortcomings of low efficiency and low stability of the point cloud algorithm. The general flow chart is shown in Figure 2.

![Diagram](image)

**Figure 2. Algorithm processing flowchart**

### 3.1. Image preprocessing

In order to improve the adaptability and scalability of the algorithm, the test artifacts are selected as shown in Figure 3. Before processing the graphic model, the image of the target artifact needs to be preprocessed to obtain an accurate graphic model: first, the image is grayed out, to facilitate the subsequent binarization processing and the calculation of the target contour; then set the image area of interest, which can reduce the working image range and increase the subsequent processing speed; then use the gray value to binarize the interest area to extract the approximate contour area of all the workpieces; finally fill in the area contour to get the full coverage pixel area of each workpiece. The image preprocessing process is shown in Figure 4.

![Image](image)

**Figure 3. Test artifact**
3.2. **Grab point selection**

Considering that the grasping fixture of the cooperative robot is a two finger claw, as shown in Fig. 5 (a), and the opening of the fixture has a certain limit, it is necessary to select the appropriate grasping point for different workpieces. For the test workpieces in the upper section, the grasping points can be roughly divided into two categories: for cylindrical workpieces, as shown in Fig. 5 (b), the grasping point should be selected as the plane center point of the workpiece. For symmetrical workpiece, as shown in Fig. 5 (c), the grabbing point should be the plane center point of any end.
For a cylindrical workpiece, the center point of the pixel region can be obtained by calculating the average value of row coordinates and column coordinates of all pixels in the workpiece region, as shown in Fig. 6 (a). Assuming that the pixel set of the workpiece area is \((Row_i, Column_i), i = 0, 1, \ldots, n\) of which, the pixel coordinates of the cylindrical workpiece grabbing point \((u, v)\) are:

\[
\left( \frac{\sum_{i=0}^{n} Row_i}{n}, \frac{\sum_{i=0}^{n} Column_i}{n} \right)
\]

(1)

For symmetrical workpieces, the algorithm of grabbing point is relatively complex. The method adopted in this paper is to get the central symmetrical rectangle by using the center point of the rectangle outside the region, and then divide the workpiece region into two independent regions through the rectangle, and calculate the center point of the smallest region as the grabbing point, as shown in Figure 6 (b). Suppose that the pixel set of the two regions after segmentation is:

\[
A = \{(Row_j, Column_j) | i = 0, 1, \ldots, m\}, \quad B = \{(Row_j, Column_j) | j = 0, 1, \ldots, n\}
\]

(2)

If \(m < n\), the pixel coordinate of the grabbing point of the symmetrical workpiece \((u, v)\) is:

\[
\left( \frac{\sum_{i=0}^{m} Row_i}{m}, \frac{\sum_{i=0}^{m} Column_i}{m} \right)
\]

(3)

If \(m > n\), the pixel coordinate of the grabbing point of the symmetrical workpiece \((u, v)\) is:

\[
\left( \frac{\sum_{j=0}^{n} Row_j}{n}, \frac{\sum_{j=0}^{n} Column_j}{n} \right)
\]

(4)

The above results are only the two-dimensional pixel coordinates of the target workpiece, which need to be converted into the three-dimensional coordinates of the robot end through the camera internal parameters and hand eye calibration matrix. The camera internal parameters are obtained by Zhang Zhengyou chessboard calibration method [10], as shown in Figure 7 (a), and the hand eye calibration matrix is obtained by Tsai two-step method [11], as shown in Figure 7 (b). The coordinate transformation process is as follows:
Through camera calibration, we can get the camera internal parameters, that is, the conversion matrix from pixel coordinate system to camera coordinate system:

\[
\begin{bmatrix}
X_c \\
Y_c \\
Z_c \\
1
\end{bmatrix} =
\begin{bmatrix}
fx & 0 & u_0 & 0 \\
0 & fy & v_0 & 0 \\
0 & 0 & 1 & 0 \\
1
\end{bmatrix}
\begin{bmatrix}
u \\
v \\
1
\end{bmatrix}
\]

(5)

Among them, \(u, v\) is the row and column coordinates of the pixel, \(fx, fy, u_0, v_0\) is the camera internal parameters, \(X_c, Y_c, Z_c\) is the 3D coordinates under the camera coordinate system, \(Z_c\) is the height of the chessboard plane relative to the camera, and \(c\) is the fixed value. The height \(Z'_c\) of the target workpiece needs to be obtained by using the point cloud coordinate mapping relationship.

Through the two-step Tsai method of hand eye calibration, we can get the transformation matrix of the camera coordinate system to the robot coordinate system:

\[
\begin{bmatrix}
X_b \\
Y_b \\
Z_b \\
1
\end{bmatrix} =
\begin{bmatrix}
R & T \\
0 & 1
\end{bmatrix}
\begin{bmatrix}
X_c \\
Y_c \\
Z_c \\
1
\end{bmatrix}
\]

(6)

Among them, \(X_b, Y_b, Z_b\) is the three-dimensional coordinates of the robot in the terminal coordinate system, \(R\) is the rotation matrix of \(3 \times 3\), and \(T\) is the translation matrix of \(3 \times 1\).
Through the combination of the plane coordinates of the graph model and the height of the point cloud, the three-dimensional coordinates of the target capture point in the camera coordinate system can be obtained. The mapping relationship between the graph model and the point cloud model is shown in Figure 8. Then, the hand eye calibration matrix is used to affine transform the spatial coordinates of the camera coordinate system to the robot end coordinate system, and finally the spatial coordinates of the target grasping point are obtained.

![Figure 8. Mapping graph from graph model to point cloud model](image)

3.3. Euler angle calculation
The three-dimensional coordinates of the target grasping point are obtained from the above section. However, because the pose information captured by the robot is six degrees of freedom, the Euler angle information of the grasping point needs to be calculated to make the robot orient and move relative to the base coordinate system, and finally grasp the target with the desired attitude. Because the end fixture of the robot is a two finger claw, and the grasping plane is parallel to the XOY coordinate system in the robot base coordinate system, the Euler angles RX and RY have no rotation angle, that is zero, only the angle of rotation around the Z axis RZ need to change with the angle of the workpiece.

Take the screw workpiece as an example, first calculate the fitting external rectangle of the screw graph model, and then use the rectangle to calculate the minimum fitting ellipse. The angle between the long axis of the ellipse and the horizontal direction is the rotation angle \( \beta (-180^\circ < \beta < 180^\circ) \) of the graph model, and the sharp side of the long axis is the positive direction. When the sharp conditions on both sides of the long axis are the same (such as the rectangle), the side with small coordinates is the positive direction, as shown in Figure 9 (a).

However, the angle is based on the pixel polar coordinate system, and the level axis of the pixel polar coordinate system and the polar coordinate system of the end gripper of the robot does not coincide, so it can not be directly sent to the robot, and the offset of the angle needs to be considered, as shown in figure 9 (b), assuming that the angle between the pixel polar coordinate system and the polar coordinate system of the end gripper is \( \alpha (0^\circ < \alpha < 180^\circ) \). If the angle of the workpiece in the pixel polar coordinate system is \( \beta (-180^\circ < \beta < 180^\circ) \), the angle of the workpiece in the end claw polar coordinate system is:

\[
R_z = \begin{cases} 
\beta + (-180^\circ + \alpha), & \beta \in (-\alpha, 180-\alpha) \\
\beta + (-360^\circ + \alpha), & \beta \in (-180-\alpha, 180) \\
\beta + \alpha, & \beta \in (-180,-\alpha)
\end{cases}
\] (7)
4. Experimental verification and analysis

In order to verify the accuracy and rapidity of the algorithm, a certain number of target workpieces are randomly placed on the robot target grabbing platform, and the grabbing experiment is carried out by combining the graph model and the point cloud model. In this experiment, 12 data samples are taken, and the effective processing of the algorithm is 12 times, as shown in Figure 10. Table 1 is obtained by counting the effective recognition number of workpieces and the processing time of the algorithm.

![Image: Data calculation results](image)

**Table. 1 Statistical table of algorithm processing time**

| Serial number | Number of workpieces | Number of errors | Processing time/s | Average processing time/s |
|---------------|----------------------|------------------|-------------------|---------------------------|
| 1             | 6                    | 0                | 0.450             | 0.075                     |
| 2             | 6                    | 0                | 0.436             | 0.073                     |
| 3             | 6                    | 0                | 0.448             | 0.075                     |
| 4             | 6                    | 0                | 0.454             | 0.076                     |
| 5             | 6                    | 1                | 0.467             | 0.078                     |
| 6             | 6                    | 0                | 0.443             | 0.074                     |
By carrying out the target grab experiment on 12 groups of randomly placed artifacts, the average processing time of the algorithm for a single workpiece is about 0.075 s, and the accuracy of the effective recognition of a single workpiece is 98.5. Compared with literature [4-6] and experimental data, it can effectively improve the speed and accuracy of target recognition by combining graph model and point cloud model.

5. Conclusion

By combining graph model with point cloud model, and using the efficiency and accuracy of graph model processing and the coordinate mapping relationship of point cloud model, the shortcomings of long time consuming and low accuracy of point cloud model recognition and location processing are avoided, and the target recognition speed and location accuracy are effectively improved. The experimental results show that the method has faster processing speed and higher recognition accuracy in the application of target recognition and location grabbing.

References

[1] Van der Geer, J., Hanraads, J.A.J., Lupton, R.A. (2010) The art of writing a scientific article. J. Sci. Commun., 163: 51–59.
[2] Taryudi, Wang MS. 3D object pose estimation using stereo vision for object manipulation system[C]/International Conference on Applied System Innovation. IEEE, 2017: 1532 – 1535.
[3] Barrois B, Krueger L, Woehler C. Method for matching an object model to a three-dimensional point cloud: US, US8379969[P]. 2013.
[4] Zhao Gang, Guo Xiaokang, Liu Dezheng, et al. Rapid recognition and location of CAD model for point cloud scene of random workpiece [J]. Laser and infrared, 2019, 49 (12): 1490-1496 DOI: 10.3969/j.issn.1001-5078.2019.12.017 (in Chinese).
[5] Qi Yutao. Research and improvement of NDT algorithm based on scale invariant feature of point cloud [D]. Gansu: Lanzhou University, 2020 (in Chinese).
[6] Qiu Yao. Research on disordered sorting technology of industrial robot based on 3D target recognition [D]. Shaanxi: Xi’an University of technology, 2019 (in Chinese).
[7] Zhang Weiwei, Chen Chao, Xu Jun. Positioning and mapping method based on fusion of laser and visual point cloud information [J]. Computer applications and software, 2020, 37 (7): 114-119 DOI: 10.3969/j.issn.1000-386x.2020.07.018 (in Chinese).
[8] Liu Yao, Jiang civil affairs, Yang Hukun. A fast localization method for extracting LIDAR point cloud line features [J]. Sensors and Microsystems, 2020, 39 (3): 34-37, 41 DOI: 10.13873/j.issn.1000-9787 (2020)03-0034-04 (in Chinese).
[9] Gary Bradski, Adrian Kaehler. Learning OpenCV [M]. O’Reilly Media, 2013
[10] Rusu R B, Cousins S. 3d is here: Point cloud library (pcl)[C]//Robotics and automation (ICRA), 2011 IEEE International Conference on. IEEE, 2011: 1-4.
[11] Zhang ZY.. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2000, 22 (11).
[12] Tsai R.. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses[J]. Robotics & Automation, IEEE Journal of, 1987, 3(4). 323-344.