International Journal of Frontiers in Medicine
ISSN 2706-6819 Vol.2, Issue 2: 87-92, DOI: 10.25236/IJFM.2020.020210

Research on object recognition and manipulator grasping strategy based on binocular vision

Haiyan Sheng[1,2,a*], Shimin Wei[1,b] , Yu Yang[1,c] and Xiuli Yu[1,d]

[1]*College of Automation, Beijing University of Posts and Telecommunications, China*
[2]*Century College, Beijing University of Posts and Telecommunications, Beijing, China*
[*]*Corresponding author e-mail: [a]shenghaiyan@ccbupt.cn,*
[b]*wsmly@bupt.edu.cn,*
[c]*yuyang@bupt.edu.cn,*
[d]*yxl@bupt.edu.cn*

ABSTRACT. *Machine vision has been successfully applied in a series of challenging environments, but at the same time, it has encountered problems[1]. In order to solve this problem, this paper presents an object recognition and manipulator grabbing strategies based on binocular vision. Our method first uses the feature point matching algorithm based on binocular vision to recognize the target, then through the detection of the target object in the image, through the extraction and recognition of the target point, the pose estimation and location of the grab point, determines which grab strategy to operate.*

KEYWORDS: *binocular vision, robotic arm, feature point matching, object recognition, grab strategy*

1. Introduction

In recent years, with the development of computer technology, computer vision as an important research field of artificial intelligence, it has been widely used in all walks of life. Vision based robotic arm grabbing is becoming more and more popular, which is a research hotspot in the grab task of robot arm. In general, manual teaching method, such as hand breaking robotic arm, which is adopted to fixed position grabbing [2]. The posture depends on memory and the robot arm has no perception ability. However, when performing tasks, they are vulnerable to many uncertain factors in the external environment. For example, once the position of an object changed, the manipulator will fail to grasp the object. How to use computer vision to solve the grasping problem of manipulator is particularly prominent.

ISSN 2706-6819 Vol.2, Issue 2: 87-92, DOI: 10.25236/IJFM.2020.020210

First, the camera and other acquisition equipment are used to sample the target object, and then combined with pattern recognition, image processing and other methods of analysis; Secondly, the collected image data was processed continually, which could obtain the spatial position of the target object and the attitude and other effective information; Finally, the obtained information was utilized to complete the manipulator computer vision, which enables the machine to "see" in effectively perceive the external environment[3].

2. Feature point matching

In this paper, we study the robot object recognition based on the vision perception system. The vision perception system is based on binocular vision. We mainly study how to establish an image processing system to capture the target object's distance and three-dimensional information for intelligent grabbing applications, which include the characteristics of the object extraction, the accurate coordinates calculation. The depth information can be obtained by fusing the binocular images and observing the parallax between them. The reliable binocular recognition algorithm is also very effective in image segmentation for target recognition and 3D scene for computer graphics.

Feature point stereo matching is an important part and the most critical step of binocular vision [4]. Robot recognition of the target object must recover the three-dimensional information of the target object according to the parallax principle [5], and accurate feature point stereo matching is the premise of calculating the visual difference. The research framework of the whole paper is shown in Figure 1.

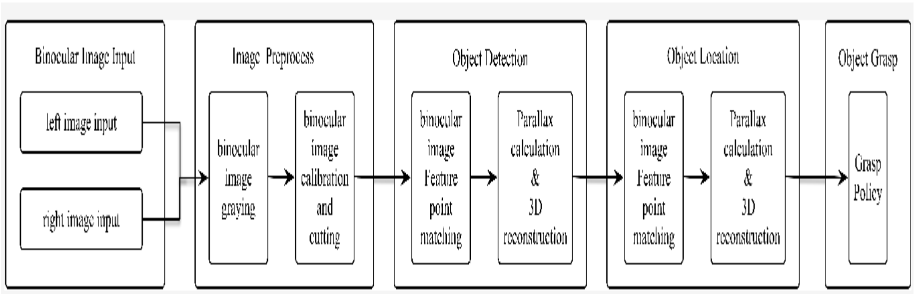

Fig. 1 research framework

In order to improve the disambiguation and calculation efficiency of feature point stereo matching, it is very easy to be interfered by the external environment. The consistency constraint, uniqueness constraint and continuity constraint proposed by Marr are considered as the basic constraints in stereo matching [6]. When stereo matching is realized, these constraints can not only improve the matching accuracy,

ISSN 2706-6819 Vol.2, Issue 2: 87-92, DOI: 10.25236/IJFM.2020.020210

but also accelerate the matching speed. The following briefly describes some common matching criteria.

1) polar constraint

The polar line constraint schematic is shown in Figure 2. In binocular vision system, O_{11} and O_{12} represent the positions of the light centers of two cameras; $P1$ and $P2$ represent the images captured by two cameras. O is any point in three-dimensional space. I_1 and I_2 are the projection points of O in left and right images respectively. The intersection lines between the polar plane where the three points O, O_{11} and O_{12} are located and left and right images are $I1(I1e)$, $I2(I2e')$. No matter how much the O changes, the resulting polar line will pass through the poles e and e', which is the polar constraint. Its algebraic representation is the basic matrix. The method of calculating the basic matrix adopted in this paper is a simple and effective 8-point algorithm [14].

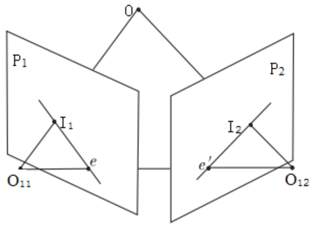

Fig. 2 the polar line constraint schematic

2) uniqueness constraint

Generally, the feature points extracted from the left image have only one corresponding feature point on the right image.

3) disparity continuity constraint

Because of the continuity of the object itself, the image points projected on the image plane are still continuous.

4) sequential consistency constraint

The feature points have the same order on the polar line of the imaging plane of the left and right cameras.

In binocular stereo vision system, stereo matching is the key to realize robot positioning. The purpose of stereo matching is to find the imaging point of a point in the space on the left and right imaging planes by a suitable algorithm [7].

ISSN 2706-6819 Vol.2, Issue 2: 87-92, DOI: 10.25236/IJFM.2020.020210

3. Object recognition strategy

Binocular vision includes two processes. One is fusing the features observed by two cameras, and the other one is reconstructing the three-dimensional original image of these features [8]. Since the original image of the corresponding point appears at the intersection of the imaging point and the corresponding light center ray, the reconstruction of the three-dimensional original image is relatively simple, as shown in Figure 3.

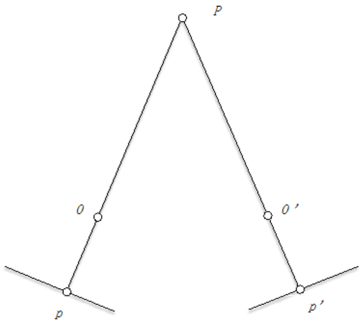

Fig. 3 no ambiguity three-dimensional image reconstruction

At the same time, binocular vision is simple at any time when a single image feature is observed. However, an image often contains hundreds of thousands of pixels, and there are tens of thousands of image features in practical applications, such as edges, corners, etc. Therefore, it is necessary to design some methods for establishing the correct matching to avoid the wrong depth measurement. As shown in Figure 4, the left image plane can take four points arbitrarily, which can correspond to any four points on the right image plane. Only four correspondences are correct, others will lead to incorrect reconstruction.

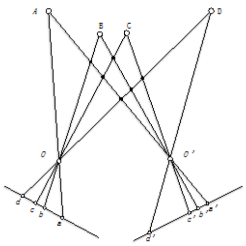

Fig. 4 general three-dimensional image reconstruction

At present, the mainstream object recognition strategy involves inputting a given feature trust set into a multi class classifier, and then using some samples to train

ISSN 2706-6819 Vol.2, Issue 2: 87-92, DOI: 10.25236/IJFM.2020.020210

each object class. All these recognizers are widely used template matchers which can be described (template and matching loss are implied in classifier). Unlike the pure geometry method, their main advantage is that, they can extract more discriminative image texture information, or include color information sometimes. However, an additional segmentation operation is often needed to separate the interested objects from the image background.

These methods need the texture to be discriminative, which is often not satisfied. These methods can’t deal with complex deformation due to that they transform the internal structure in the object category. Finally, the current object recognition methods assume that each entity belongs to only one object classification category.

Another important choice is to use some templates to represent the spatial relationship of objects. These methods should be better template matchers at present, and they inherit the hidden complexity of such methods. Now, such methods can only deal with relatively limited sets; they may not be able to deal with complex deformation; and they prohibit the transformation of internal structure in the object category [9].

The third scheme is to use registration technology. Given the geometric model of an object, the image is registered with the geometric model [10]. If the registration similarity score is very high, the object is considered. If there are many different types of objects, each object has its own geometric model, then each model needs to be registered, and which object is selected according to the registration score.

4. Grab strategy of robotic arm

We mainly study how to extract and identify the target point, how to estimate and locate the position and pose of the grabbing point, and which grabbing strategy to use in this part. In the process of intelligent positioning of the manipulator, the end of the moving manipulator is detected and tracked, the spatial position of the grab point is estimated, and the operation path of the manipulator is corrected according to the position error between the end of the manipulator and the grab point, so that the end of the manipulator is constantly approaching the grab point, so as to realize the location of the grab point. There are four kinds of grabbing strategies in the field of aerospace and industry at home and abroad, which are adaptive grabbing strategy based on Gaussian process, tracking strategy based on PID, intercepting grabbing strategy based on position prediction and dynamic grabbing region equation.

Machine vision is the main method for robot to obtain external information, and robot arm is the main executive part of robot. The robot obtains the position information of the object relative to the camera through the vision system, and then obtains the position information of the object relative to the robot arm through the coordinate conversion between the camera coordinate and the robot arm coordinate. The robot arm motion control system gives the robot arm control command according to the object target position, and controls the robot arm to move and grasp.

ISSN 2706-6819 Vol.2, Issue 2: 87-92, DOI: 10.25236/IJFM.2020.020210

The vision control algorithm mainly includes PID control, image Jacobian pseudo inverse matrix control method, task function method and state space method.

PID control method is a widely used control law in industrial process control, with simple structure, high robustness and good reliability [11]. PID controller is mainly composed of proportional part, integral part and differential part. Through the reasonable deployment of the three parts, the control can achieve fast, stable, accurate and reliable results. How to choose the appropriate proportional, integral and differential coefficients is the main problem in PID controller [12].

5. Conclusion

To sum up, the research content of this paper is challenging and difficult, covering image processing, computer science, mechatronics, control science and robotics, which is a multi-interdisciplinary subject. The next step is to further optimize various algorithms to improve efficiency.

Reference

[1] Wang Xinghua. Research on binocular vision recognition and location technology based on embedded system [D]. Harbin: Harbin University of technology, 2016.
[2] Lin Y, Sun Y. Robot grasp planning based on demonstrated grasp strategies[J]. International Journal of Robotics Research,2015, 34(1): 26-42.
[3] Li Shengli. Research on workpiece recognition and positioning technology based on binocular stereo vision [D]. Harbin: Harbin University of technology, 2016.
[4] Cao Zhile, Yan Zhonghong, Wang Hong.Summay of binocular stereo vision matching technologu [J]. Journal of Chongqing University of Technology: Natural Science,2015,29(2).
[5] Zhang Guangjun.Visual measurement[M].Beijing: Science Press,2008.
[6] Jiang ang. Research and implementation of stereo matching algorithm in binocular vision [D]. Tianjin University of technology, 2016.
[7] Chen Bin. Research and implementation of real-time binocular stereo matching algorithm [D]. Wuhan University of science and technology, 2014.
[8] David A.Forsyth, Jean Ponce. Computer Vision: A Modern Approach, Second Edition[M]. Beijing: Electronic Industry Press,2017.7.
[9] Lin T Y, Dollár P, Girshick R, et al. Feature Pyramid Networks for Object Detection[J]. 2016.
[10] WANG Linyu, JIANG Qiangwei, LI Shuang. Research on Fusion Matching Object Recognition and Location Based on Binocular Image Multifeature Point Fusion[J].Radio Engineering,2018,48(8):628-633.
[11] Shuhuailin. PID neural network and its control system [M]. National Defense Industry Press, 2006.
[12] Kang J, Meng W, Abraham A,et al. An adaptive PID neural network for complex nonlinear system control[J]. Neurocomputing, 2014, 135(8):79-85.