Application of Machine Vision in Pose Detection of Objects

Yuxi Ying, Qingqing Shao, Xingting Pei and Hongbing Yang
Nanjing Agricultural University, China

Abstract. The application of machine vision technology in pose detection of objects is studied. Subtraction of the morphological operation is used for rough localization and preliminary detection of the target area. Hough transform is used to detect the linear edge of the target. Under the condition of the known target template, the geometric parameters of the target are known, and the pose data of the target is obtained by using simple geometric relations.

1. Introduction
Machine vision refers to using the computer system to imitate the functions of the human vision system to realize the computer's perception and positioning of the target object. With the development of theories of itself, machine vision is applied more and more widely in industry. For example, the hybrid industrial parts are still classified mainly by manual means. This classification method is not only inefficient and high error rate, but also a waste of human resources. If automatic classification of parts can be realized, industrial production efficiency will be greatly improved.

2. Research purposes and its significance
The positions and attitude measurements of space objects are a key issue in the research of real-time navigation and motion target motion tracking. In the large-scale manufacturing system, robot related research, space technology and flexible manufacturing, etc., many insights need to be able to achieve high precision measurement of the posture of non-contact space objects.

As a non-contact measurement method, the technology of visual measurement is characterized by good real-time performance and high degree of data collection. The study of posture measurement system based on monocular vision is of great significance to solve the problems of assembly site and automatic monitoring.

3. Camera calibration
Camera calibration refers to the process of establishing the relationship between the image collected by the camera and the location of the world's midfield spots by measuring various internal and external parameters of the camera.

Traditional calibration methods use 2D or 3D targets. By establishing the spatial coordinate relation between the feature points on the target and the feature points in the image, the internal and external parameters of the camera are calculated. The active visual calibration method makes the camera move according to a running track. The self-calibration method does not need to know the three-dimensional coordinates of the target points of the features. Compared with the others, the calibration method of
active vision is more complex, but the result of self-calibration is not stable, so the traditional calibration method is adopted.

In the camera calibration method based on 2D plane target, it assumes that the internal parameters of the two cameras remain unchanged, and the frame is set in two different directions and angles to keep the camera fixed. We need to change the position and angle of 2D plane target continuously.

Calibration can be done using MATLAB calibration toolbox or Opencv calibration. Generally, the calibration accuracy is higher with MATLAB. Therefore, the calibration software here USES MATLAB 2014b.

4. Image processing
In order to realize the rough localization of the target quickly, the collected image needs to be processed initially. In this paper, we mainly use the image and background image to do the difference and the operation of corrosion and expansion in the morphological operation.

4.1. Background processing
Before collecting the target image, ensure that the lighting conditions remain unchanged and the camera position does not move. Firstly, the background image is collected. The acquired target image is compared with the background image to extract the target area. It can be seen that the target can be initially extracted after the difference is made. These areas are removed by using morphological operations.

4.2. Morphological operations
The edge detection algorithm based on mathematical morphology is not as sensitive to noise as the classical differential algorithm. Expansion and corrosion are the basic operations. Expansion can be used to fill the small holes in the target after segmentation, while corrosion can be used to eliminate small wires of unnecessary and isolated small points.

After binarization of the image removed from the background, the reflective region near the cube can be eliminated by several etching operations. And the interference of high-light region has been removed from the final target, and the contour of the target object has been retained completely.

5. Edge detection and matching

5.1. Edge detection
In the process of image feature detection, edge features are one of the most important ones. For example, Hough transform is used to detect the edge of a straight line. Hough transform maps the line represented by slope and intercept \((k, b)\) in Cartesian coordinate system to the line represented by polar diameter and Angle \((r, \text{the flow path})\) in polar coordinate system. And the opencv-based Hough transform is used to detect the straight line of the target object, and part of the edge of the object can be obtained.

5.2. Linear matching
In order to obtain the three-dimensional space coordinates of the edge of the line, the diagonal points need to be matched. Since the coordinates of the two endpoints of the starting point and the ends have been obtained, the angular point matching of the two endpoints here can add a constraint condition and reduce the matching errors. To reduce the interference of factors such as sampling noise and illumination change, the following three conditions are used as constraints when matching.

1) Polar line constraint
Since there is a certain relation between the two cameras to obtain the two images, the relation of the position can be converted into the position constraint of the points on the image. The polar line constraint is one of the most important constraints for image matching.

2) Gray similarity
The two images of two cameras to gather the value of the angle differently, cause of two images of the same space, the surrounding environment is not the same, but there is still a certain similarity, especially for surface texture with the surrounding environment different, or with a difference of environmental color is bigger, gray similarity still has a certain reference value.

3) Line constraint

The two points that matched should be the starting point or end point of the same edge.

5.2.1. Polar constraints. As shown in figure 1, $p_1$ and $p_r$ are the imaging points of spatial point $p$ on the sampled images of the two cameras, and the plane composed of two optical centers $C_L, C_R$ and $p$ is called polar plane. It can be seen from the observation that there is a constraint on the position of $p_1$ and $p_r$ in the left and right images, what is, both lie on the pole line, and once $p$ is determined, the left and right pole line is unique.

Figure 1. Epipolar constraint

5.2.2. Gray similarity. The region $(2N + 1) \times (2N + 1)$ size is the window template, and the gray similarity of the two points of the feature is calculated.

$$C = \frac{\sum_{i=-N}^{i=N} \sum_{j=-N}^{j=N} [I_L(x_L + i, y_L + j) - I_L] [I_R(x_R + i, y_R + j) - I_R]}{\sqrt{\sum_{i=-N}^{i=N} \sum_{j=-N}^{j=N} [I_L(x_L + i, y_L + j) - I_L]^2 \sum_{i=-N}^{i=N} \sum_{j=-N}^{j=N} [I_R(x_R + i, y_R + j) - I_R]^2}}$$

Among them, $I_L(x_L, y_L)$ and $I_R(x_R, y_R)$ are the gray values of a point in the left and right images, and $I_L$ and $I_R$ are the average gray values of window templates in the left and right images. The range of $C$ is $[-1, 1]$, and the larger the value is, the higher the gray similarity between the two points is.

5.2.3. Linear constraint. Even if two points meet both of the above constraints, there is also the possibility of a false match. Considering that every point that matched is the end of a line, two match points should be on the same line at the same time. If it can also be found and is collinear with the previous matching point, it is considered that the match point on a line is found, and a matching line in the left and right graph is also found.

6. Calculation of coordinate and pose

When the same kind of workpieces are identified in production, the object template information can be used to detect and recognize the object. In the case of the cube, the information needed to know about the object is only the side length. In the experiment, it was 100 mm. There are many ways to determine the exact configuration of a cube. And the positioning is determined by calculating the center of mass. Firstly, we need to look for three line segments that meet the following conditions.

1) The distance between any two line segments shall not be more than twice of the length of the cube side; 2) Three line segments are perpendicular to each other; 3) At least one intersection exists.

The
intersection must be an angle point on the cube, denoted by \( M \). In order to calculate the spatial coordinates of the three line segments, the transformation from image pixel coordinates to world coordinates is realized by camera calibration which get matrix determined by internal and external parameters.

The starting points and ends of line segments \( a, b, c \) are calculated as:

\[
a_1 (130.12, 105.36, 536.77), \quad a_2 (136.24, 12.17, 505.16), \quad b_1 (136.45, 13.05, 508.90), \quad b_2 (220.13, 32.13, 463.81), \quad c_1 (138.44, 7.31, 496.12), \quad c_2 (184.51, -11.05, 588.87);
\]

Where: \( a_1 \) is the starting point of line segments \( a \); \( a_2 \) is the end point of line segments \( a \); \( b_1 \) is the starting point of line segments \( b \); \( b_2 \) is the end point of line segments \( b \); \( c_1 \) is the starting point of line segments \( c \); \( c_2 \) is the end point of line segments \( c \).

Then, pick the intersection \( M (136.45, 13.05, 508.90) \) of three line segments and calculate unit direction vectors of three line segments separately, set point \( M \) as the starting point of unit direction vectors of \( a, b \):

\[
\vec{a}_0 (−0.062, 0.945, 0.321), \quad \vec{b}_0 (0.863, 0.197, −0.465);
\]

Where: \( \vec{a}_0 \) is unit direction vector of \( a \); \( \vec{b}_0 \) is unit direction vector of \( b \). However, the direction of unit vector of \( c \) cannot be determined, it might be:

\[
\vec{c}_{0+} (0.438, −0.175, 0.882) \text{ or } \vec{c}_{0−} (−0.438, 0.175, −0.882);
\]

We can get two possible centroid \( O_1, O_2 \). But \( O_1 \) can be excluded by calculating the distance from it to the starting points and ends of \( a, b, c \). Hence, only \( O_2 \) can meet the geometry condition of a cube and the unique cube can be determined.

7. Conclusion

Algorithm, this paper studied the camera calibration that was conducted in the 2D plane targets on the calibration method of camera calibration. Hough detection methods of straight lines were used to detect the boundary and calculate the matching line of space three-dimensional coordinates. In this paper, the simple geometric relation is used to calculate the pose state of the target, and better results are obtained. However, under the environment of complex background, the effect will decline and affect the accuracy of the final result, which needs further study.

References

[1] Tsai R Y. An efficient and accurate camera calibration technique for 3D machine vision [C]. IEEE Conference on Computer Vision and Pattern Recognition, 1986: 364 - 374.
[2] Lin Yuchi, Cui Yanping, Huang Yinguo. Research on the edge extraction and target recognition under complex background [J]. Optical precision engineering, 2006, 14 (3): 509-514.
[3] Yang Bowen, Zhang Liyan, Ye Nan, Feng Xin, Li Tielin. Camera calibration technology for large field visual measurement [J]. Journal of optics, 2012 (09).