Visual Positioning Algorithm Based on Micro Assembly Line

Quandong Ni¹, Dazhai Li¹,* , Yun Chen² and Haiyang Dai²

¹School of Mechanical Engineering and Automation BUAA, Beijing, China
²AVIC Radar and Avionics Institute Wuxi, Jiangsu Province 214063 P.R. China

*Corresponding author email: Lidazhai@buaa.edu.cn

Abstract. To achieve mass production of microelectronic devices, this paper designs a microelectronic device detection scheme based on monocular vision, on the assembly robot platform of micro-electronic devices. By analyzing the principle of camera calibration and distortion and performing internal parameter calibration of the camera with Zhang Zhengyou calibration method, this paper establishes a hand-eye calibration model of the relationship between the robot’s motion and the relative position of the camera, so as to calculate the transformation matrix from the camera model to the gripper. In the end, the paper proposes an accurate positioning method which combined boundary tracking algorithm based on binary image and template matching algorithm, identifying the type of workpieces through barcode.

1. Introduction

With the popularity of electronic devices, the demands for core electronic products have shown an explosive increase, especially for high-density and high-performance microwave components. The large varieties of microwave components, great configuration differences, and high accuracy requirements for processing demand an efficient micro-assembly production lines to achieve its automated processing. At present, the assembly and testing of micro assembly line still rely on manual work, requiring massive manpower and material resources. In order to solve this problem, a target detection technology is proposed to realize the automation of industrial production in this paper.

Target detection technology is widely used in industrial production. In recent years, target detection based on deep learning has excellent accuracy. However, due to the large amount of calculation in deep learning, it needs high-performance GPU as support, and deep learning technology needs large sample data as training set, which is difficult to achieve in many industrial scenarios. In the industrial system, the traditional target detection method is often used. The simple binary processing is applied to the image, the threshold segmentation is made by using the difference of the gray value of the image pixel, and the interested part is selected as the result of the visual detection. However, the industrial image often contains noise, the image background is complex, and the gray value of the image is easily affected by the light conditions, so it is easy to produce errors.

This paper proposes an accurate positioning method which combined boundary tracking algorithm based on binary image and template matching algorithm. The template matching technology based on edge gradient can not be affected by the change of light, and the boundary tracking based on binary image can roughly select the target from the complex environment. In the following passage, the author first introduces the camera model required by target detection, establishing a conversion from pixel coordinate to world coordinate, then presents the target detection algorithm based on boundary tracking contour extraction combined with template matching.
2. Camera Model

2.1. Robot Model
The target detection model of this paper is designed on the basis of the intelligent debugging test robot platform of micro assembly line. As a rectangular coordinate model, the robot has four-degree-of-freedom, with three axes of x, y, z and r rotation degree of freedom, as shown in Figure 1.

![Figure 1. Robot model.](image)

2.2. The Internal Parameter Calibration of Camera
Camera calibration is to set a transformation matrix by index, in order to change the position of the workpiece from pixel in camera to scenes in reality. The matrix can be disassembled into two parts of the internal and external transformation matrix, having internal and external parameters respectively. The principle of camera imaging can be seen below (Figure 2).

![Figure 2. The principle of central imaging.](image)

Camera calibration is to solve the coordinate transformation matrix from pixel coordinate in the picture to world coordinate. As shown in figure 2, camera coordinate system OcXcYcZc only needs to be transformed by matrix transformation to reach the world coordinate system, without deforming the rigid body. This is the external parameter of camera calibration. From the graph, we can get the conversion relation of workpiece from image coordinate system to camera coordinate system. As shown in equation (1). Besides, f in figure represents focal length.

\[
\begin{bmatrix}
    x \\
    y \\
    z
\end{bmatrix} =
\begin{bmatrix}
    0 & 0 & 0 & Xc \\
    0 & f & 0 & Yc \\
    0 & 0 & f & Zc \\
    0 & 0 & 0 & 1
\end{bmatrix}
\]

(1)

In fact, the pixels captured by the camera are on the imaging plane S. And the datum of pixel coordinate is not the same as that of image coordinate. As shown in equation (2), we can get the conversion relationship from pixel coordinates to image coordinates. Besides, 1/dx and 1/dy are the conversion factors.
The projection matrix multiplied by the conversion matrix from pixel coordinates to image coordinates equals the internal parameter matrix in camera calibration. The above two equations on the right can get the transformation equation of workpiece from world to pixel, as shown in Formula (3). Besides, \( f_x = f / dx \), \( f_y = f / dy \). \( f_x, f_y, u_o, v_o \) are the internal parameters of the camera.

\[
\begin{bmatrix}
  u \\
  v \\
  1
\end{bmatrix} =
\begin{bmatrix}
  1 & 0 & u_o \\
  0 & 1 & v_o \\
  0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
  x \\
  y \\
  1
\end{bmatrix}
\]

(2)

In practical applications, there remains errors in the actual projection and theoretical projection on the imaging plane, which is caused by lens distortion. In the camera model of this project, we mainly focus on radial distortion. Assuming that the real imaging point of the workpiece in the imaging plane is \( P(\epsilon, \delta) \), and the ideal imaging point is \( P(x, y) \), the conversion formula could be as follows equation (4).

\[
\begin{bmatrix}
  x' \\
  y'
\end{bmatrix} = (1 + k_1r^2 + k_2r^4 + k_3r^6)x
\]

\[
\begin{bmatrix}
  x' \\
  y'
\end{bmatrix} = (1 + k_1r^2 + k_2r^4 + k_3r^6)y
\]

(4)

In the equation (4), \( r^2 = x^2 + y^2 \), and \( k \) serves as the distortion parameter. Practical applications tend to calculate the first two orders of distortion parameter for an accurate projection image model in real world.

For the internal parameter matrix and distortion model, this project uses the Zhang Zhengyou calibration method[1][2] to calibrate the internal and distortion parameters. Zhang Zhengyou’s calibration method uses a checkerboard of fixed size and shape as a calibration board to establish a world coordinate system. The z axis of the world coordinates is 0. Then a camera with pre-adjusted focus and aperture will take pictures from multiple angles of the fixed checkerboard. With 3 or more pictures, we can calibrate the camera parameters of this project.

| Table 1. Calibration results of camera internal parameter. |
|----------------------------------------------------------|
| camera parameters | \( k_1 \) | \( k_2 \) | \( f_x \) | \( f_y \) | \( u_o \) | \( v_o \) |
| opencv           | -0.1043507 | 0.084831577 | 3616.2 | 3616.4 | 2795.6 | 1809.3 |

2.3. Camera External Parameter Calibration
This chapter discusses the coordinate transformation from camera to gripper. The target detection of this project is to find the position of the workpiece, and then use robot hand to grasp it. To this end, we designed the camera external parameter calibration model below based on the robot model.
As shown in Figure 3, \( h \) is the shooting height and \( \vec{dx} \) is the vector of the workpiece relative to the camera. From the internal parameter matrix and the shooting height, we can find the distance value of gripper movement that needs to be calculated using the calibration parameters after our camera calibration by vector \( \vec{Px} \). As the robot model, the camera and the end gripper in this project are fixed together, vector \( \vec{Rc} \) is invariant. From the formula of \( \vec{Px} = \vec{Rc} + \vec{dx} \), it is clear that the camera external parameter calibration of this project is to calibrate the vector \( \vec{Rc} \) from the gripper to the camera.

When the workpiece \( P \) is captured by the camera, after the internal parameter transformation, its coordinate in the camera coordinate system \( (x_c, y_c, h_c) \) could be calculated. The coordinate \( (c, c, 1) \) is the coordinate of one unit in front of the imaging plane. Moreover, \( \vec{Px} = (x, y, z) \) can be obtained by our teaching of grabbing, by subtracting the gripper coordinate of photographing from the gripper coordinate of grabbing. The formula could be as following.

\[
\begin{align*}
    x_r &= h \cdot x - x \\
    y_r &= h \cdot y - y \\
    z_r &= h - z
\end{align*}
\]

\( (x_r, y_r, z_r) \) is the vector \( \vec{Rc} \); \( (h \cdot x, h \cdot y, h) \) is the vector \( \vec{dx} \); \( (x, y, z) \) is the vector \( \vec{Px} \). From a set of data, we can get 3 equations and 4 unknown numbers. Adding at least one set of data will lead to the increase of one unknown number of \( h \), so we need at least two sets of data to get the vector value from the end of the gripper to the camera. The increase of data will lead to the rise of accuracy in calibration results. This project uses the hand-eye calibration model to calibrate at least twenty sets of data based on linear regression, and obtains the calibration results as the following table.

### Table 2. The parameters of camera external parameter calibration.

| Hand-eye calibration coordinates | \( x_r \) | \( y_r \) | \( z_r \) |
|---------------------------------|--------|--------|--------|
| coordinate values              | 219.065 | -5.448 | 172.7769 |

3. **Target Workpiece Positioning**

Due to the fact that microwave components have the characteristics of small size and high precision requirements, this project chooses to use the algorithm of edge detection combined with feature matching to locate the workpiece. Then it uses the extraction of the two-dimensional code to identify the specific workpiece.

The contour-based edge detection algorithm is to filter and binarize the image, using edge detection to extract the contour, and filtering the extracted contour through a self-designed algorithm. To be accurate, feature matching algorithm is also used for feature extraction to calculate the positioning center of the workpiece.
3.1. Filtering and Contour Extraction

This project uses Daheng Imavision MER camera to photograph, which captures color image, and then converts it into a grayscale image to reduce the calculation time and space. In addition to the workpiece information we need, the image collected by the camera also contains noise and background information. As the workpieces in this project have some color differences with their surroundings, the background information can be used for contour extraction operations. Therefore, we need to filter the image to remove the noise.

As images in industrial environment mainly contain Gaussian noise and salt pepper noise, the mixed denoising method is adopted to combine the linear characteristics of Gaussian filter and the nonlinear characteristics of median filter.

A boundary tracking algorithm based on binary image [3] is used to extract the roughing contour of the workpiece and prepare for template matching. Before contour extraction, the image shall be binarized. To identify multiple workpieces, the Otsu threshold segmentation algorithm is used to extract the binarized image. In order to ensure that all the workpieces are included in the roughing of contour, two expansions and three corrosions are performed after the image binarization.

The boundary tracking algorithm based on binary image is a method of tracking and analyzing the topology of the binary image. It extracts a series of coordinate points and chain codes of the boundary contour, marks and processes the boundary contour, and obtains all the contours of binary image to develop a tree topology. This algorithm has following advantages: 1. It will not miss any part of the workpiece contour; 2. All its contours have a tree-structured relationship, which is convenient to traverse and search for the contour of the target workpiece.

Then the extracted contours are filtered to select the smallest one containing the workpiece. As the workpieces in this project are attached with QR code on the surface, we can filter contours of the workpiece by identifying the number of QR codes. Due to the fact that the contours extracted by the boundary tracking algorithm have a structure of tree topology, it is not necessary to traverse all contours, but contours at the same level. When the contours of the same level do not have a two-dimensional code, its sub-contours are no longer needed to be searched; when the contours of the same level have multiple two-dimensional codes, its sub-contours shall be traversed; when the contours of the same level contain only one two-dimensional code, that contour is exactly the one we want, which is the outline of the smallest contained workpiece. After roughing selection of the workpiece, interference environments could be removed efficiently, which greatly improves the accuracy of the template matching result.
3.2. Template Matching

Template matching\(^4\) is an algorithm to find the position of a specific template in an image. It traverses the image through a sliding window, and calculates the similarity function between the template and the sliding window to obtain a similarity function result matrix. Then by judging the value of the similarity matrix, it could determine the template matching position. This project chooses to use standardized sum of squared differences to calculate the similarity of matches.

\[
R(x, y) = \frac{\sum_{x',y'} (T(x', y') - I(x+x', y+y'))^2}{\sqrt{\sum_{x',y'} T(x', y')^2} \sum_{x',y'} I(x+x', y+y')^2}
\]

(6)

In this formula, \(T(x', y')\) serves as the template; \(I(x, y)\) as the target image; and \(R(x, y)\) is used to describe the similarity.

To cope with the inaccuracy of matching caused by the rotation of the workpiece in the template matching, this paper uses three cross marks for matching. Instead of template matching for the entire workpiece, this paper matches the cross marks precisely processed on the surface of the workpiece. In addition, due to the fact that contours in the roughing selection contains the QR code for workpiece recognition, which will interfere with the matching of cross mark, we choose not to take the best matching point, but to binarize the matching result and select a part of the high matching point.

\[\text{Figure 7. Template image.}\]

\[\text{Figure 8. Template matching results.}\]

\[\text{Figure 9. Binarization.}\]

Binarization processing is performed on the similarity results of template matching, and the positions of the processed three crosses are selected through a specific matching point filtering algorithm.

**Table 3.** Matching point filtering algorithm.

| matching point filtering algorithm |
|-----------------------------------|
| 1. Perform a boundary tracking algorithm on the binary image to extract all contours, then exclude contours with sub-contours; |
| 2. Calculate the centers of all contours; |
| 3. Find the left upper and right lower points of the precisely processed rectangle cross mark; |
| 4. Search for the remaining points. The three points of the cross could be found by locating the points closest to the upper right or lower left. |
The center position of the cross mark selected by the matching point filtering algorithm can be used to calculate the positioning center of the workpiece. Then the actual coordinates of the workpiece could be obtained through the conversion of pixel coordinates to world coordinates.

4. Conclusion

On the intelligent assembly robot platform of microelectronics, this paper designs a target detection scheme based on monocular vision, which establishes a conversion from pixel coordinates to world coordinates for robot model and realizes the positioning of microwave components through vision. It is believed that this scheme could make great contributions to the improvement of the mass production of microwave components, enhancing its efficiency and reducing costs in manpower and material resources significantly.

In this paper, a lot of experiments have been carried out on the robot platform with the target detection method. The recognition accuracy is more than 95%, and the positioning accuracy is 0.2mm. Compared with the target detection algorithm based on deep learning and the traditional threshold segmentation algorithm, it is found that the algorithm is suitable for the small sample and low configuration running environment, and it has a certain adaptability to the changes of light conditions. So it has a certain practical value. The disadvantage is that the recognition rate is poor under the condition of occlusion.

References

[1] Zhang Z. Flexible Camera Calibration by Viewing a Plane from Unknown Orientations[J]. Proc. ICCV, 1999, 1999, 1:666–673 vol.1.
[2] Zhang Z. A Flexible New Technique for Camera Calibration[J]. Tpami, 2000, 22(11):1330–1334.
[3] Suzuki, S. and Abe, K.,Topological Structural Analysis of Digitized Binary Images by Border Following. CVGIP 30 1, pp 32-46 (1985).
[4] J.A.LEESE, C.S.NOVNK, B.B.CLARK. AN AUTOMATED TECHNIQUE FOR OBTAINING CLOUD MOTION FROM GEOSYNCHRONOUS SATELLITE DATA USING CROSS-CORRELATION. JAPPLMETEOROL, 1971, 10: 110-132.