Vision Guiding System Based on Strong Positioning and Hand-eye Calibration

Gang Du¹, Shanwen Zhang²
Department of mechanical and electrical technology, Xijing University, Shanxi, China College of Information Engineering, Xijing University, Shanxi, China

¹Corresponding author e-mail: 31513707@qq.com

Abstract. In order to solve the current localization algorithm big background interference affects calibration methods exist damage, physical contact with the workpiece material on the surface of the problem, hand-eye calibration was proposed based on strong orientation and three visual guidance system. First, deepen the Hessian matrix, using Gaussian filter, using the maximum inhibition precisely determine the feature points and direction, in order to construct surf feature points description, finish strong positioning operator based on surf, to remove the background and the ambient light interference, the purpose of accurate positioning center of the target with Angle. And then put forward three points calibration method, calculation scale rotation matrix and translation matrix parameters, complete the Robot world coordinates and binding mapping camera image coordinates. Finally, program implementation algorithm and system, the experiment results showed that compared with the current workpiece material transfer system, this system has higher goal orientation and calibration transfer success rate.

Keywords: Visual guidance; location; three point calibration of hand eye; Hessian matrix; Surf feature; Zoom rotation matrix; Binding map

1. Introduction

With the sublimation and deepening of the concepts of "smart manufacturing" and "Made in China 2025", under the background of rising human costs in the future, enterprises have realized the inevitable trend of using advanced information technology to realize industrial unmanned manufacturing [1-2]. There are a large number of transfer jobs in the manufacturing line. In the past, manpower was used. Now with the promotion of robot applications, many companies have started to use robots to complete the transfer work [3-4]. However, with the deepening of applications, problems have also arisen. For example, the robot needs to rely on physical positioning blocks to complete the grasping and transferring of the product, and the gripping of the physical stop block easily leads to material damage. Therefore, some manufacturers have begun to complete robot transfer based on vision, but two problems need to be solved: (1) The background of the manufacturing environment is complicated and the lighting interference affects the positioning of the visual target. (2) Accurately calculate the mapping between the robot's world coordinates and the camera image coordinates, that is, hand-eye calibration.

In terms of visual localization and calibration, domestic researchers have introduced computer vision technology into this field and researched it. For example, Hu Mingxing [5] proposed a rotating
workpiece target recognition algorithm based on Canny detection and SIFT features. First, the RGB three channels of the collected workpiece image are assigned weights to obtain a grayscale image; and Canny edge detection and Hough line detection are used to process the grayscale image to calculate the workpiece rotation angle; and based on the geometric transformation, an image correction model is defined. To eliminate the rotation angle of the workpiece, reset it, and use template matching to locate the workpiece in the image to achieve the purpose of positioning and calibration. However, this technique relies on accurate contour capture, which often affects visual positioning when complex backgrounds affect edge capture. Du Gang [6] proposed the widespread application of computer-based monocular vision systems for robots, mainly focusing on camera positioning, segmentation calculation of captured images, coordinate recognition of patterns, detection of target distance, and application of vision systems to robots. The use and analysis have achieved the purpose of visual positioning calibration. However, this technique relies on a lot of model experience. When the relative position of the camera and the robot is unpredictable, it is often impossible to accurately bind the robot and the camera coordinate system.

In summary, in order to improve the ability of intelligent robots to complete material transfer based on visual positioning and calibration, a visual guidance system based on strong positioning operator and hand-eye calibration is mentioned in this paper, and algorithms are implemented to verify the system function of this paper.

2. Visual guidance algorithm for workpiece transfer in this paper

The process of vision guidance technology proposed in this paper is shown in Figure 1. This method first performs strong localization processing on the target image under the interference of complex background and ambient light, and then obtains the target center point and angle. The three-point hand-eye calibration method is used to bind robot coordinates and image coordinates to complete hand-eye calibration. The image target coordinates are converted into robot world coordinates. The robot grabs the target material according to the calculated coordinates and completes the transfer work.

The material target and background are shown in Figure 2. It can be seen that the reflection of ambient light is serious, and the image background is complicated, which has a greater impact on the target's visual positioning and calibration.
Figure 1. Visual guidance mechanism of workpiece transfer in this paper

Figure 2. Image to be positioned

2.1 Strong workpiece positioning based on Surf

There are two sources of interference to target positioning in the industrial manufacturing environment: reflective interference caused by ambient light exposure, and target edge positioning difficulties caused by complex backgrounds. Aiming at this problem, a strong localization mechanism based on Surf feature [7-8] is proposed, as shown in Figure 3. In order to analyze the characteristics of each pixel and its surroundings, first establish the basic data of Surf, and calculate the Hessian matrix of each pixel [9] determinant:

\[ H(f(x,y)) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} \]  

In the formula, \((x, y)\) represents any pixel in the image, \(f(x, y)\) is the gray value of the image coordinate \((x, y)\), and the Hessian matrix \(H\) is composed of a function and partial derivatives. Find the H matrix discriminant for:

\[ \det(H) = \frac{\partial^2 f}{\partial x^2} \frac{\partial^2 f}{\partial y^2} - \left( \frac{\partial^2 f}{\partial x \partial y} \right)^2 \]  

Among them, the value of \(\det\) is the eigenvalue of the H matrix, and all points can be classified by using the sign of the determination result, and the value of the discriminant can be used to determine whether the point is an extreme point.

Second-order standard Gaussian functions are used as filters. Calculate the second-order partial derivative by convolution, and calculate the three elements of the H matrix \(L_{xx}, L_{xy}, L_{yy}\) to estimate the H matrix:

\[ H(x,\sigma) = \begin{bmatrix} L_{xx}(x,\sigma) & L_{xy}(x,\sigma) \\ L_{xy}(x,\sigma) & L_{yy}(x,\sigma) \end{bmatrix} \]  

In order to achieve the purpose that the feature points are not related to the scale, perform a Gaussian filtering on the hessian matrix:
Among them, \( L(x,t) \) represents the image representation at different resolutions, which can be achieved by convolution of Gaussian kernel \( G(t) \) and image function \( f(x,t) \) at point \( x \) [10], where:

\[
G(t) = \frac{\partial^2 g(t)}{\partial x^2}
\]  

(5)

Where \( g(x) \) is a Gaussian function and \( t \) is a Gaussian variance.

Through the above process, the characteristic value of h-determinant can be calculated for each pixel in the image. After the Hessian matrix value is obtained, the feature points are located:

(1) Non maximum suppression [11-12] is used to initially determine feature points. The 26 points in the three-dimensional domain of the hazen matrix are compared with the size of each pixel processed. The maximum and minimum values of 26 points are taken out and retained as feature points.

(2) Determine the main direction of the feature point. The main direction of feature points is to calculate the sum of horizontal and vertical Haar wavelet features.

(3) Establish surface feature point description operators. Get a square box around the feature point.

(4) From the feature points, the four top angles of the minimum circumscribed rectangle of the target are selected, and the center point and inclination angle of the target are calculated.

**Figure 3. Strong localization process**

The edge positioning performance of ordinary positioning algorithms is susceptible to complex background interference and ambient light reflection interference. The strong positioning algorithm in
this paper, based on the Gaussian matrix and surf feature point search, overcomes the above interference and completes accurate positioning. As shown in Figure 4 (a), the feature points are marked as shown in Figure 4. As shown in Figure 4 (b), it can be seen that the strong positioning mechanism in this paper can accurately locate the outline of the target material. The algorithm accurately locates the contour of the rotating target material.

2.2 Three-point hand-eye calibration process

After obtaining the target center point and angle, the camera image coordinates need to be converted into Robot world coordinates. This process is hand-eye calibration [13], so that Robot can accurately grasp the target, compensate for the rotation angle, and complete the transfer operation. The purpose of hand-eye calibration is to bind the camera image coordinates and Robot world coordinates to find the mapping relationship. The core is to calculate the zoom rotation matrix and translation matrix, as shown in the following formula:

\[
R = \begin{bmatrix}
a & b \\
a' & b'
\end{bmatrix}
\]  \hspace{1cm} (6)

In the formula, R is a scaling rotation matrix, and a, b, a’, and b’ are parameter elements of the scaling rotation matrix. Then introduce the translation matrix:

\[
M = \begin{bmatrix}
c \\
c'
\end{bmatrix}
\]  \hspace{1cm} (7)
In the formula, M is a translation matrix, and c and c’ are parameter elements of the translation matrix. Simultaneous zoom rotation matrix, translation matrix, image coordinates and robot coordinates:

\[
\begin{bmatrix}
x' \\
y'
\end{bmatrix} = \begin{bmatrix}
x \\
y
\end{bmatrix} + M
\]

(8)

(x, y) represents the image coordinates, and (x’, y’) represents the robot coordinates.

\[
\begin{align*}
x' &= ax + by + c \\
y' &= a'x + b'y + c'
\end{align*}
\]

(9)

Where a, b, c, a’, b’, and c’ are the required parameters, and then determinant:

\[
\begin{bmatrix}
x' \\
y'
\end{bmatrix} = \begin{bmatrix}
a & b & c \\
a' & b' & c'
\end{bmatrix} \begin{bmatrix}
x \\
y
\end{bmatrix}
\]

(10)

According to a, b, c, a’, b’, and c’ in formula (10), a rotation zoom translation matrix is formed. The operating robot walks 3 points on different straight lines, records the corresponding Robot coordinates and image coordinates, and then substitutes into the formula 10 to obtain:

\[
\begin{align*}
x_0' &= ax_0 + by_0 + c \\
x_1' &= ax_1 + by_1 + c \\
x_2' &= ax_2 + by_2 + c \\
y_0' &= a'x_0 + b'y_0 + c' \\
y_1' &= a'x_1 + b'y_1 + c' \\
y_2' &= a'x_2 + b'y_2 + c'
\end{align*}
\]

(11)

(12)

Among them, (x₀, y₀), (x₁, y₁), (x₂, y₂) represent image coordinates; (x₀’, y₀’), (x₁’, y₁’), (x₂’, y₂’) represent robot coordinates.

Then determinize expressions 11 and 12:

\[
\begin{bmatrix}
x_0 & y_0 & 1 & a \\
x_1 & y_1 & 1 & b \\
x_2 & y_2 & 1 & c
\end{bmatrix} = \begin{bmatrix}
x_0' \\
x_1' \\
x_2'
\end{bmatrix}
\]

(13)

\[
\begin{bmatrix}
x_0 & y_0 & 1 & a' \\
x_1 & y_1 & 1 & b' \\
x_2 & y_2 & 1 & c'
\end{bmatrix} = \begin{bmatrix}
y_0' \\
y_1' \\
y_2'
\end{bmatrix}
\]

(14)

Perform determinant calculations according to equations (13) and (14) to obtain a, b, c, a’, b’, c’ to form a rotation-scale translation matrix.

3. Conclusion

For solving the problem of complex background and environment light interference and difficult calibration, this paper proposes a vision guidance system based on strong positioning and hand eye calibration. The Hessian matrix of image is established, the feature points are calculated, and the center points of target image are calculated. Then the three-point calibration method is analyzed, and the scaling rotation translation matrix is calculated to complete the mapping between image
coordinates and robot coordinates. The experimental results show that in the case of serious background light interference, the positioning and calibration method in this paper has higher accuracy and anti-interference than the ordinary technology.

References
[1] Adriana Giret, Emilia Garcia, Vicente Botti. An Engineering Framework for Service-Oriented Intelligent Manufacturing Systems[J]. Computers in Industry, 2016, 81(9): 116-127.
[2] Xu Jie. Moving target location algorithm based on improved weighted centroid and UKF [J]. Journal of Jilin University (Engineering & Technology Edition), 2016, 21 (147): 16-19.
[3] Wang Linbo. Research on Precise Localization of Feature Points of Cooperative Targets Based on Hand-Eye Vision [D]. Beijing: University of Chinese Academy of Sciences, 2014: 33-37.
[4] Wu, Qian, Guan, Haobing, Gu, Jinhui, Giret, Fuyong, Zhu, Xining. Research on Image Preprocessing of Workpiece Recognition and Positioning System[J]. Journal of Convergence Information Technology, 2013, 8 (8): 187-193.
[5] Hu Mingxing. Rotating workpiece target positioning algorithm based on template matching optimization coupled with image correction [J]. Modular Machine Tool & Automatic Manufacturing Technology, 2016,21 (21): 283-287.
[6] Du Gang. Analysis of Calibration and Positioning Algorithms for Robot Binocular Vision System [J]. Computer Engineering, 2016, 16 (4): 181-185.
[7] WeiLong Zhang, LeiBo Liu, ShouYi Yin. An efficient VLSI architecture of speeded-up robust feature extraction for high resolution and high frame rate video[J]. Science China Information Sciences, 2013, 56 (7): 1-14.
[8] Hu, Zhang Fang, Yang, Lin, Luo, Luan. The Research and Application of SURF Algorithm Based on Feature Point Selection Algorithm[J]. Sensors & Transducers, 2014, 169 (4): 67-72.
[9] Ying Wei, Rui Li, Jin-zhu Yang. An algorithm for segmentation of lung ROI by mean-shift clustering combined with multi-scale HESSIAN matrix dot filtering[J]. Journal of Central South University, 2012, 19 (12): 3500-3509.
[10] Zeping Wu, Donghui Wang, Patrick Okolo. N. Unified estimate of Gaussian kernel width for surrogate models[J]. Neurocomputing, 2016, 203(26): 41-51.
[11] Chen Jinhui, Ye Xining. Improvement of non-maximum suppression algorithm in pedestrian detection [J]. Journal of East China University of Science and Technology (Natural Science Edition), 2015, 41 (3): 371-378.
[12] Wang Wentao, Zhu Shun'an, Tang Yan. An Adaptive Non-Maximum Suppression Edge Detection Algorithm and FPGA Design Implementation [J]. Small Microcomputer System, 2014, 35 (8): 1926-1930.
[13] Hu Xiaoping, Zuo Fuyong, Xie Ke. Research on calibration of hand-eye calibration methods for micro-assembly robots [J].