Circle detection and location based on binocular stereo vision

Jing Lin\textsuperscript{1*}, Wenxin Li\textsuperscript{1} and Jiagao Xia\textsuperscript{1}

\textsuperscript{1}Lanzhou Institute of Physics, Lanzhou, Gansu, 730000, China
\textsuperscript{*}qiantang2019@whu.edu.cn, bellneck@163.com

Abstract. In this paper, a binocular stereo vision system is developed for recognition and localization of a spherical object. The system uses two cameras to acquire images of object, shows and processes the images from two cameras simultaneously. It then applies point Hough transform to recognize the object and get real-time coordinates through the two images. Experiments with the proposed method yielded good results, as the system can detect and position a spherical object successfully.

1. Introduction

Rapid improvement of computers’ data processing capability and introduction of new technology makes computer vision (CV) a research hotspot in the field of information technology. The vision-based target location technology is a constant topic of interest in the field of CV. Human beings can observe the three-dimensional structure of the things all around the world. While computers can only obtain the two-dimensional visual information from images. Therefore, we need to acquire three-dimensional position information through two-dimensional images.

Monocular vision positioning and binocular vision positioning are usually used to solve the problem of camera spatial positioning. Part of monocular vision positioning methods requires prior knowledge of certain information of the target, such as the size or special marked point. Other monocular vision positioning methods use one camera to obtain multiple images at different locations, which leads to poor time performance. The characteristic of human vision is emulated in the methods of binocular vision positioning. Two cameras are used to collect images synchronously, then we can get the position information of target from the images. Due to the limitations of monocular vision positioning technology, in this paper, we choose binocular vision positioning technology with better time performance.

In this paper, we design and implement a set of real-time binocular vision system for tracking and locating a spherical object. Firstly, images of the object are synchronously acquired from the binocular cameras. Secondly, image processing techniques and point Hough transform are used to locate the target in the images. Then, the three-dimensional coordinates of the spherical object’s center are calculated by binocular stereo vision.

2. Design principle

Since the sphere target is displayed as a circle on the two-dimensional image, Hough circle transformation is selected for target recognition. For the same sphere, the lines passing through the center of the circle in the two-dimensional image also pass through the center of the sphere, so we only need to detect the position of the center of the circle in the two-dimensional images obtained by the left and right cameras to determine the position of the center of the small target ball.
2.1. Point Hough Transform
The Hough transform is a classical method for finding circles in images. It transforms the image space into the parameter space. Points in the same circle are considered to have the same set of parameters \((a, b, r)\). \((a, b)\) is center coordinates of the circle, \(r\) is Radius of the circle. If the accumulated value of the object edge points that satisfy the parameters is more than the threshold, these points are considered to be from the same circle. To meet the requirements of real-time processing, we choose the point Hough transform which reduces the calculation to detect circle in images.

The principle of the point Hough transformation is shown in Figure 1. \(LM\) and \(LK\) are two non-parallel chords on a circle, and \(l_{LM}\) and \(l_{LK}\) are the perpendicular lines of the two chords. According to the geometric properties of the circle, The perpendicular must intersect at the center point \(O\). Suppose the coordinates of the three points \(L, M, K\) are \((x_L, y_L), (x_M, y_M)\) and \((x_K, y_K)\), then the coordinates \((a, b)\) of the center \(O\) and the radius \(r\) of the circle can be derived as follow:

\[
\begin{align*}
    a &= -(d_{LK} - d_{LM}) (k_{LK} - k_{LM})^{-1} \\
    b &= k_{LK} a + d_{LK} \\
    r &= \sqrt{(x_K - a)^2 + (y_K - b)^2} \\
\end{align*}
\]

(1)

Where

\[
\begin{align*}
    k_{LK} &= -(x_L - x_K)(y_L - y_K)^{-1}, \quad d_{LK} = \frac{1}{2} [y_L + y_K - k_{LK} (x_L + x_K)] \\
    k_{LM} &= -(x_L - x_M)(y_L - y_M)^{-1}, \quad d_{LM} = \frac{1}{2} [y_L + y_M - k_{LM} (x_L + x_M)] \\
\end{align*}
\]

Figure 1. The principle of the point Hough transformation

2.2. Calculation of three-dimensional coordinates
Assume that we have a set of calibrated stereo vision system as shown in Figure 2. Two cameras whose image planes are exactly coplanar with each other, with exactly parallel optical axes that are a known distance apart, and with equal focal lengths \(f_L = f_R = f\). The principal points (the intersection of the optical axis and the image plane) \(c_L\) and \(c_R\) have been calibrated to have the same pixel coordinates in their respective left and right images.

In the figure, \(A\) and \(B\) are the left and right cameras respectively, and their focal length is \(f\). The baseline distance between \(c_L\) and \(c_R\) is \(b\). \(Q\) is a point in the physical world in the left and the right
image views at \( Q_1 \) and \( Q_2 \), which will have the respective horizontal coordinates \( x_l \) and \( x_r \). The vertical distance between \( Q \) and the camera is \( L \). Referring to the figure:

\[
LP^{-1} = \left( L + f \right) \left( P + x_r \right)^{-1}
\]

\[
L(P + b)^{-1} = \left( L + f \right) \left( P + b + x_l \right)^{-1}
\]

We can get the distance from point \( Q \) to the camera

\[
L = bf \left( x_l - x_r \right)^{-1}
\]

And the coordinate of \( c_l \)

\[
\begin{cases}
    c_x = P + b = bx_l \left( x_l - x_r \right)^{-1} \\
    c_y = by_l \left( x_l - x_r \right)^{-1}
\end{cases}
\]

The calibrated stereo coordinate system takes \( c_l \) as the original point, the image plane as the XY plane, the direction perpendicular to the XY plane as the \( Z \) axis, and the direction pointing to the image plane as the positive direction of the \( Z \) axis. As \( x_l \) and \( x_r \) are the coordinates in the respective image plane, the disparity \( d' \) should be written as

\[
d' = x_l - x_r - (c_x - c_x')
\]

Where, \( c_x' \) is the coordinate of \( c_x \) on the right image.

Using the coordinates \((x_l, y_l)\) and \((x_r, y_r)\), the three-dimensional coordinates of \( Q \) in the above coordinate system can be calculated by equation (7).

\[
\begin{cases}
    x = b \left( x_l - c_x \right) / d' \\
    y = b \left( y_l - c_y \right) / d' \\
    z = bf / d'
\end{cases}
\]

3. Design of Binocular Stereo Vision System

The hardware part of the system is composed by two USB cameras and a PC. Images of target are captured by the two cameras. Meanwhile, the PC processes and analyzes the collected images in order to identify, track and locate targets.

The software of the system is written in C++ language and runs under the Windows operating system environment. Intel Open Source Computer Vision Library (OpenCV) is used for image processing. DirectShow is used for camera image acquisition. The system software is responsible for the completion of image pre-processing, target detection and coordinate calculation functions. The image pre-processing work is as follows:

- weighted image grayscale.
- Canny edge detector.
- 3-by-3 median filter.

To acquire three-dimensional space information through two-dimensional images, we need a coordinate transformation between three-dimensional world and two-dimensional images.

To obtain the parameters of the two cameras, Zhang’s camera calibration method has been chosen. It requires the three-dimensional world coordinates and corresponding two-dimensional image
coordinates of a set of feature point. Then we can get a coordinate transformation between three-
dimensional world and two-dimensional images.

The block diagram of the vision system is shown in Figure 3.

Figure 3. Block diagram of the vision system

The interface of the system is shown in Figure 4.

Figure 4. The interface of the system
4. Verification and analysis

The use of sections to divide the text of the paper is optional and left as a decision for the author. Where the author wishes to divide the paper into sections the formatting shown in table 2 should be used.

4.1. Positioning result test

In the previous camera calibration work, the checkerboard is used to obtain feature points set. Several of points were selected to verify the positioning of the binocular stereo vision system. Figure 5 is the left and right images collected by the dual cameras of the system. We can get the two-dimensional coordinates of the corner points 1~5 in the two cameras. By using the system, the three-dimensional coordinates of points are also calculated. Since the side length of the checkerboard is known, we can calculate the distance between adjacent points and evaluate the positioning results. Distances between adjacent points calculated by the positioning system are shown in Table 1.

![Figure 5. Images of checkerboard](image)

| Point Number | Coordinate in left image | Coordinate in right image | Computed coordinate | Measuring distance (mm) | Proportional error |
|--------------|--------------------------|----------------------------|---------------------|------------------------|-------------------|
| 1            | (140,133)                | (212,108)                  | (41,0,508)          | d₁=27.6                | 0.72%             |
| 2            | (137,145)                | (212,119)                  | (35,20,490)         | d₂=27.4                | 1.44%             |
| 3            | (133,158)                | (211,131)                  | (27,40,473)         | d₃=27.0                | 2.88%             |
| 4            | (126,132)                | (197,107)                  | (16,2,518)          | d₄=27.3                | 1.79%             |
| 5            | (123,143)                | (197,118)                  | (9,18,498)          | d₅=26.6                | 3.75%             |

The side length of the checkerboard is 27.8mm. We calculated the proportional error as follow:

$$\text{proportional error} = \frac{\text{actual distance} - \text{measuring distance}}{\text{actual distance}}$$

The calculation result shows that the system can obtain the points’ three-dimensional coordinates, and the proportional error is within 5%.

4.2. Analysis of the measuring range of the system

When the target is too close to the camera, the two cameras cannot collect a complete target image at the same time. As shown in Figure 6, the distance between the two cameras is \(m\), the radius of the target is \(r\), and the camera angle of view is \(2\alpha\). The minimum distance \(d\) that the target should be from the camera can be calculated by equation (8).
\[ d = \frac{r}{\sin \alpha} + \frac{m}{2 \tan \alpha} \]  

(8)

Figure 6. The schematic diagram of system’s minimum distance

The angle of view of the camera we used is 69°, the distance between the two cameras is 140mm, the radius of the ping pong ball is 20mm, and the distance between the camera and the target should be greater than 138mm by calculation.

In addition, if the camera is too far away from the target, and the target takes up a little space in the images, the accumulated value of the object edge points will be very small. If the value cannot reach the judgement threshold of the point Hough transform, the system also failed to detect the target.

4.3. Testing of the visual platform

For demonstrating the system performance, a ping pong ball is used as the spherical target and the images is captures. The object is placed at a fixed distance from the camera. Through this binocular stereo vision system, we can get the coordinates of the target and measure the distance. The experimental results are shown in Table 2.

Table 2. Make use of the binocular stereo vision system to get the coordinates of spherical object and measuring distance.

| Coordinate in left images | Coordinate in right images | Computed coordinate | Actual distance (mm) | Measuring distance (mm) | Proportional error |
|--------------------------|---------------------------|---------------------|----------------------|------------------------|--------------------|
| (262,166)                | (59,188)                  | (-38, 36, -191)     | 198                  | 200                    | 1.0%               |
| (278,151)                | (43,172)                  | (-98, -10, -175)    | 201                  | 200                    | 0.5%               |
| (273,147)                | (67,180)                  | (-109, -10, -201)   | 229                  | 230                    | 0.4%               |
| (265,94)                 | (75,120)                  | (-113, 30, -219)    | 248                  | 250                    | 0.8%               |
| (265,161)                | (90,184)                  | (-121, -22, -238)   | 268                  | 270                    | 0.7%               |
| (257,172)                | (103,193)                 | (-130, -35, -270)   | 302                  | 300                    | 0.7%               |
| (206,167)                | (78,184)                  | (-100, -37, -327)   | 344                  | 350                    | 1.7%               |
| (246,152)                | (127,174)                 | (-158, -22, -357)   | 391                  | 400                    | 2.3%               |
| (234,135)                | (128,163)                 | (-163, -2, -404)    | 436                  | 450                    | 3.1%               |
| (197,115)                | (104,138)                 | (-128, 29, -463)    | 481                  | 500                    | 3.8%               |

According to the experimental data, the spherical object within a certain distance can be well recognized by the binocular vision system, and the distance from the camera to the target can be accurately calculated. Meanwhile, we found that the system can achieve the best measurement results
when the target is 200mm-300mm away from the camera. The proportional error of the measurement is rising as the distance increases.

5. Conclusion
In this paper, a binocular vision platform was established. By using binocular stereo vision, the correspondence between two-dimensional images and three-dimensional space was completed, so that the computer could calculate the coordinates in a specific three-dimensional coordinate system from the coordinates on the two-dimensional image. We used two cameras to obtain the image of a ball in real time, obtained the coordinates of the center of the circle in the two-dimensional image by detecting the circular target in the image using point Hough transform, and then used binocular stereo vision to finally determine the center position of the target ball. Experimental results proved that the vision system detected the target ball well in a certain distance, the relative error of the positioning result was very small, and the system could accurately detect and locate spherical targets.

References
[1] Hou, Z.Q., Han, C.Z. (2006) A Survey of Visual Tracking. J. Acta Automatica Sinica, 32(4): 603–617.
[2] Han, Y.X., Zhang, Z.S., Dai, M. (2011) Monocular vision system for distance measurement based on feature points. J. Optics and Precision Engineering, 19(5):110–117.
[3] Wang, T.Y., Dong, W.B., Wang, Z.Y. (2017) Position and orientation measurement system based on monocular vision and fixed target. J. Infrared and Laser Engineering, 46(4): 0427003-1–0427003-8.
[4] Zhou, F., Yang, C., Wang, C.G, et al. (2013) Circle detection and its number identification in complex condition based on random Hough transform. J. Chinese Journal of Scientific Instrument, 34(3):622–628.
[5] Lin, J.L., Shi, Q.Y. (2003) Circle Recognition Trough a Point Hough Transformation. J. Computer Engineering, 29(11):17–160.
[6] Zhao, X., Yuan, J.Z., Liu, H.Z. (2016) Advances in Vision-based Target Location Technology. J. Computer Science, 43(6): 10–43.
[7] Zhao, X.J., Li, C.J. (2006) Research of Key-tech in a Binocular Real-time Ranging System. J. Laser and Infrared, s36: 874–877.
[8] Zhang, Z.Y. (2000) A flexible new technique for camera calibration. J. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11): 1330–1334.
[9] Bradski, G., Kaebler, A. (2008) Learning OpenCV. O’Reilly Media, The Sebastopol.
[10] Okutomi, M., Kanade, T. (1993) A multiple-baseline stereo. J. IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(4): 353–363.