ISSN 2706-655X Vol.4, Issue 7: 30-34, DOI: 10.25236/IJFET.2022.040707

Research on Adjustable Baseline Binocular Vision Measurement System

Bing Yang[a], Baonan Yang[b], Jindong Liu[c]

School of Computer Science, Xijing University, Xi'an, 710123, P.R. China
[a]934032350@qq.com, [b]yangbaonan@foxmail.com, [c]amosliu@yeah.net

Abstract: *Binocular stereo vision belongs to machine vision. It uses two cameras to imitate human eyes to obtain two images of the object to be measured, and obtains the three-dimensional information of the object by calculating the positional deviation between the corresponding points of the images. The baseline length of the binocular camera determines the size of the observable distance and range. In today's increasingly complex work environment, the baseline of traditional binocular measurement equipment is fixed, and the usage scenarios are too limited. Therefore, this paper proposes a binocular vision measurement system with dynamically adjustable baselines. The system solves the problem of object localization at different depths by changing the baseline length.*

Keywords: *Machine vision, 3D information, Baseline, Target setting*

1. Introduction

With the continuous advancement of science and technology, all walks of life are developing rapidly in the direction of intelligence. Machine vision [1] stands out for its advantages of high efficiency, speed, accuracy and low cost, and plays a pivotal role in many fields such as measurement, positioning, 3D reconstruction, and unmanned driving [2]. The measurement method based on binocular vision [3-4] uses two cameras to collect the image information of the target from different angles, calibrates the binocular camera, and performs stereo matching on the left and right images [5-7] to obtain the disparity map of the target. The three-dimensional coordinate information of the target is calculated according to the disparity map and the calibrated internal and external parameters.

In the application of binocular stereo vision, for the binocular stereo imaging model, the data acquisition method of the binocular baseline is basically used. According to the image information collected by the dual cameras, the triangulation method is used to calculate the position of the target in the environment. However, the binocular camera with a fixed baseline is only practical for targets with small distances and depths. For targets with long distances and large depth differences, it is limited by the length of the baseline, and it is difficult to obtain good position information. The length of the baseline greatly limits the flexibility and environmental applicability of binocular vision. Based on the above background, this paper studies a binocular stereo vision system with a variable baseline, and uses parallel cameras that vary within a range of 40cm from the baseline to locate targets within a range of 1m-120m.

2. Binocular Stereo Vision Localization with Variable Baseline

2.1. Analysis of Binocular Measurement System

The traditional binocular stereo vision system is mainly composed of two cameras with the same parameters, the relative positions of the two cameras are fixed, and the image information of the same target is obtained at the same time. This method fixes the baseline, limits the detection distance, and can only image 3D targets within a certain range. Inferences on how changes in baseline parameters affect accuracy are as follows:

$$D = \frac{B \cdot f}{d} \tag{1}$$

$$D + \Delta D = \frac{B \cdot f}{d + \Delta d} \tag{2}$$

ISSN 2706-655X Vol.4, Issue 7: 30-34, DOI: 10.25236/IJFET.2022.040707

$$\Delta D = \frac{B \cdot f}{d} - \frac{B \cdot f}{d + \Delta d} = D - \frac{1}{\frac{1}{D} + \frac{\Delta d}{B \cdot f}} \tag{3}$$

Where D is the depth distance, ΔD is the depth deviation, d is the parallax, Δd is the parallax deviation, B is the baseline distance, and f is the camera focal length. Assuming the depth distance D of the fixed object, that is, the accuracy is judged at the same depth, because the parallax accuracy is constant, so Δd is also constant. It is not difficult to find that B and f have the same influence on D. The larger B is, the larger f is, and the smaller Δd is, it means that the larger the baseline, the longer the focal length, the higher the depth accuracy, and the depth accuracy is the same as the baseline and focal length. proportional. From the above inference, the schematic diagram of constructing a variable-distance binocular measurement system is shown in Figure 1. The method of variable baseline length is used to move the left and right cameras. This process ensures that other parameters of the camera do not change. After changing once, the binocular camera obtains the target image once. Therefore, in the same disparity map, different baseline parameters can be used for the target at different distances to calculate and obtain the distance information of the target from the short range to the long range. As shown in Figure 1, P, P1, and P2 respectively represent the most suitable object positions for measurement under different baseline lengths.

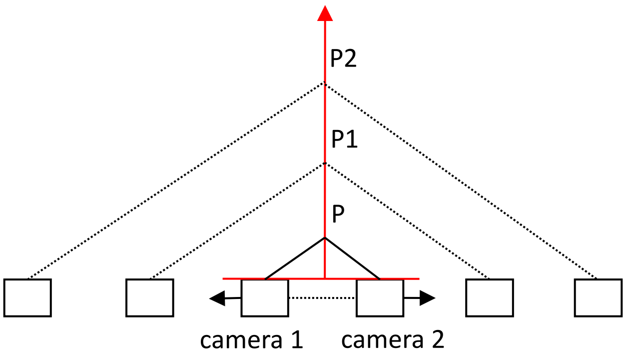

Figure 1: Analysis of binocular measurement system.

2.2. Mathematical model of binocular vision with variable baseline

The imaging principle of the object is the transformation of the coordinate system between the imaging models, Formula (4) is used to describe the relationship between the pixel coordinate system, the image coordinate system, the camera coordinate system and the world coordinate system:

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & O & u_0 \\ O & \frac{1}{dy} & v_0 \\ O & O & 1 \end{bmatrix} \begin{bmatrix} f_x & 0 & 0 \\ 0 & f_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} R & T \\ O^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = M_1 M_2 \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \tag{4}$$

Assuming that the position of the measured target in the camera coordinate system is (Xc, Yc, Zc), the pixel coordinates of the measured target are (u, v, 1), the conversion matrix between the camera coordinates and the pixel coordinates is replaced by A, the camera The transformation matrix between one and camera two is replaced by B. Then there is a conversion relationship between the camera coordinates and the pixel coordinates and the relative coordinates of the two cameras as in formula (5):

$$\begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = [A_1] \begin{bmatrix} \frac{X_{c1}}{Z_{c1}} \\ \frac{Y_{c1}}{Z_{c1}} \\ 1 \end{bmatrix} \quad \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = [A_2] \begin{bmatrix} \frac{X_{c2}}{Z_{c2}} \\ \frac{Y_{c2}}{Z_{c2}} \\ 1 \end{bmatrix} \quad \begin{bmatrix} X_{c1} \\ Y_{c1} \\ Z_{c1} \end{bmatrix} = [B] \begin{bmatrix} X_{c2} \\ Y_{c2} \\ Z_{c2} \end{bmatrix} \tag{5}$$

Seven parameter equations of the position of the measured target in the camera coordinate system are constructed by formula (5), and the need to solve is (X_{c1}, Y_{c1}, Z_{c1}), (X_{c2}, Y_{c2}, Z_{c2}), a total of six unknowns.

ISSN 2706-655X Vol.4, Issue 7: 30-34, DOI: 10.25236/IJFET.2022.040707

2.3. Calibration and parameter datasets

When the binocular baseline changes, the image distortion increases. The binocular internal and external parameters and distortion coefficients are obtained through calibration, and the LM algorithm is used to optimize them. Finally, the distortion-corrected image is calculated. First use the maximum likelihood estimation method to correct the image distortion:

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\left\|m_{ij}-m\left(A,R_i,t_i,M_j\right)\right\|^2 \tag{6}$$

For the resulting nonlinear distortion, the Taylor series around the coordinates of the principal point containing the radial distortion is:

$$\begin{cases}\tilde{u}=u+(u-u_0)[k_1(x^2+y^2)+k_2(x^2+y^2)^2]\\ \tilde{v}=v+(v-v_0)[k_1(x^2+y^2)+k_2(x^2+y^2)^2]\end{cases} \tag{7}$$

Converting the above formula into a matrix representation, we get:

$$\begin{bmatrix}(u-u_0)(x^2+y^2) & (u-u_0)(x^2+y^2)^2\\ (v-v_0)(x^2+y^2) & (v-v_0)(x^2+y^2)^2\end{bmatrix}\begin{bmatrix}K_1\\ K_2\end{bmatrix}=\begin{bmatrix}u-\tilde{u}\\ v-\tilde{v}\end{bmatrix} \tag{8}$$

The three-dimensional coordinates corresponding to the two-dimensional coordinates correspond to two equations for each point, so the m points on the n images will have a total of $2mn$ equations, then:

$$\mathrm{DK}=\mathrm{d} \tag{9}$$

in, $\mathrm{K}=\begin{bmatrix}K_1 & K_2\end{bmatrix}$

The least squares solution to this equation is:

$$\mathrm{K}=(D^TD)^{-1}D^T\mathrm{d} \tag{10}$$

Use the optimization function to calculate the parameters, and the objective function is:

$$\sum_{i=1}^{n}\sum_{j=1}^{m}\left\|m_{ij}-m\left(A,k_1,k_2,R_i,t_i,M_j\right)\right\|^2 \tag{11}$$

In formula (11), $\widetilde{m}\left(A,k_1,k_2,R_i,t_i,M_j\right)$is the coordinates of the j points projected onto the ith image after correction. The LM algorithm is used for optimization, and finally the image after distortion correction is calculated.

When the baseline of the system changes dynamically, the internal parameters of the camera remain unchanged, and the external parameters change accordingly. Aiming at the constant change of parameters caused by changes in the baseline, several sets of parameter data were calibrated and a variable baseline parameter data set was established. as shown in Figure 2.

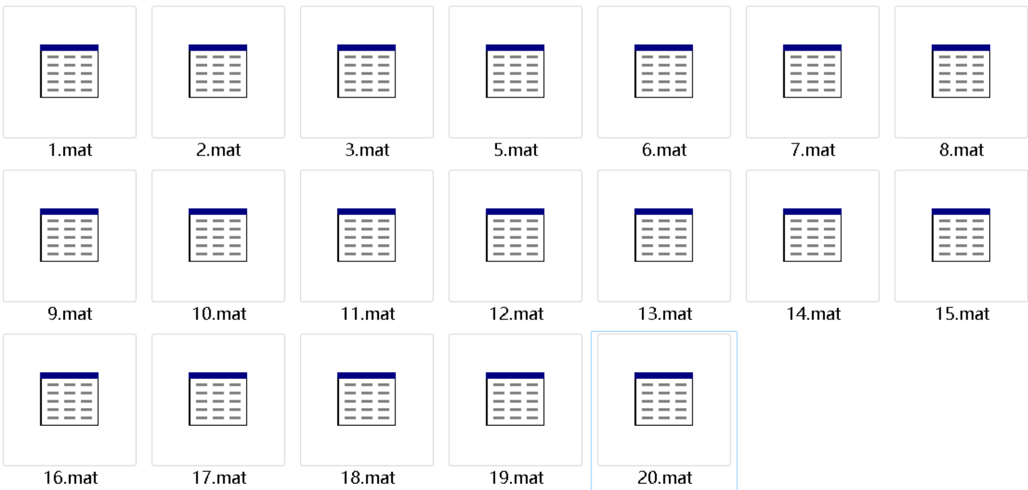

Figure 2: Variable baseline parameter dataset.

ISSN 2706-655X Vol.4, Issue 7: 30-34, DOI: 10.25236/IJFET.2022.040707

2.4. Image Stereo Correction with Unknown Parameters

The purpose of calibration is to obtain the internal and external parameters of the camera. In the variable-baseline binocular positioning system, the change of the baseline affects the extrinsic parameter matrix. If the calibration polar line correction is used in the variable baseline binocular ranging system, it will cause a lot of repeated calibration work. Therefore, this paper adopts the non-calibration normal polar line correction to realize the stereo correction of the image. The non-calibration method of polar line correction is not to obtain the relationship between image pairs through camera calibration, but only through the matching corresponding points of the two images, solve two suitable homography matrices to perform projection transformation, and correct the three-dimensional position information of the images to make the corresponding points. There is no parallax in the Y-axis direction. The purpose of Hartley's algorithm is to find the homography matrix that maps the poles to infinity, while minimizing the computational error between the two stereo images. The realization of this algorithm is achieved by matching the corresponding points between the two images. This method can bypass the calculation of the camera intrinsic parameters of the two cameras, because such intrinsic parameter information is implicit in the matching points. Its stereo correction result is shown in Figure 3.

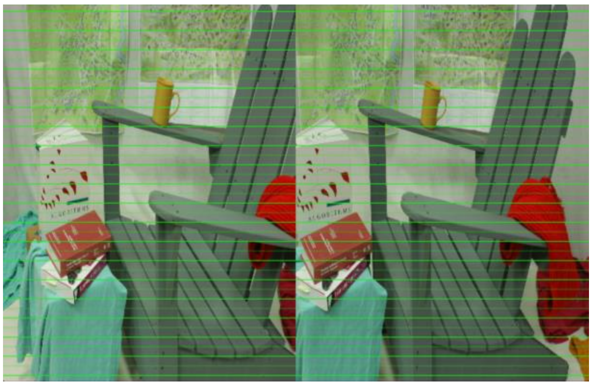

Figure 3: Stereo correction results of unknown parameters.

2.5. Semi-Global Stereo Matching Algorithm

Analyze the stereo matching algorithm and generate a disparity map to meet the depth recovery and 3D positioning of the target by the binocular device. Change the cost function in the SGM algorithm to AD+Census for cost calculation, and the AD cost is the mean of the three-channel difference of the color of the two pixels. The Census transform is to calculate the Hamming distance of the Census transform values of the two pixels corresponding to the left and right images. As shown in formula (12).

$$C_{AD}(\boldsymbol{p},d)=\frac{\sum_{i=R,G,B}\left|I_i^{left}(\boldsymbol{p})-I_i^{right}(\boldsymbol{p}-(d,0))\right|}{3} \tag{12}$$

The real-time test is shown in Figure 4.

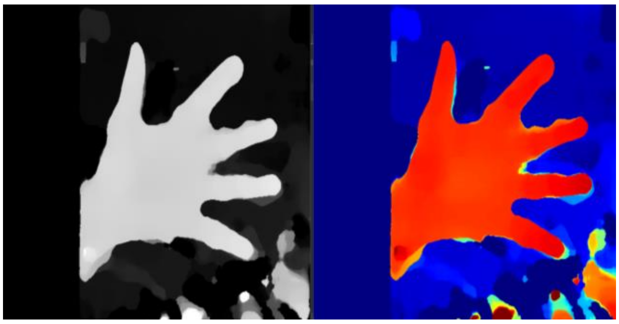

Figure 4: Deep recovery

3. Hardware platform construction

A binocular vision hardware platform with dynamically adjustable baseline is designed and developed,

ISSN 2706-655X Vol.4, Issue 7: 30-34, DOI: 10.25236/IJFET.2022.040707

which can realize the functions of binocular vision baseline adjustment control and remote image acquisition. By displaying the acquired images in real time on the PC, multiple baseline adjustments can be made for precise experiments. The hardware options are as follows:

(1) Screw slide table: unilateral stroke 20CM, bilateral stroke 40CM; load 0-30 kg; accuracy 0.05mm; suitable for 57 motor mounting holes; speed 0-50mm/S.

(2) 24V DC switching power supply: output power 120W; input voltage AC 110V/220V; output voltage DC 24V; output current DC5A.

(3) 128 subdivisions of DM542 high performance driver: power supply DC 20V-50V; output current: 1A-4.2A; subdivision: 0-128; input frequency 0-200KH.

(4) KH-01 single-axis stepper motor controller: power supply: AC 220V (power supply error is not greater than ±15%); single-axis control; maximum output frequency 40KHz; output frequency resolution 1Hz; a total of 14 programming instructions; number of programming 99 strip.

The complete hardware platform is shown in Figure 5.

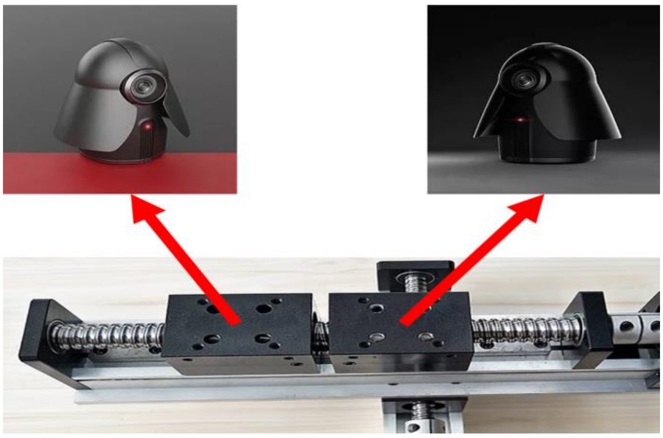

Figure 5: Hardware platform.

4. Conclusion and Analysis

This paper studies the target localization method of the traditional binocular positioning system based on a single fixed baseline, analyzes the principle of binocular positioning, and draws the conclusion that changing the baseline distance can improve the target positioning range. A binocular localization method with dynamically adjustable baselines is proposed to achieve target localization with different distances and fields of view.

References

[1] Zhang Guangjun. Machine Vision [M]. Beijing: Science Press, 2005.He Yikang, Pan Zaiping. Power Electronics Technology [M]. Beijing: Science Press,2019.
[2] Mattoccia S. Stereo Vision: Algorithm and Application[M]. Bologna: Published by I-Tech Education and Publishing, 2015: 5-22.
[3] Puglia, L., Vigliar, M., & Raiconi, G. (2017). Real-Time Low-Power FPGA Architecture for Stereo Vision. IEEE Transactions on Circuits and Systems II: Express Briefs, 64(11), 1307–1311.
[4] Luo Z, Zhang K, Wang Z, et al. 2017. 3D pose estimation of large and complicated workpieces based on binocular stereo vision [J]. Applied Optics, 56(24):92~119.
[5] Y Lu, B Wang, R Zhang. 2017.Analysis on Location Accuracy for Binocular Stereo Vision System[J]. IEEE Photonics Journal, PP(99):1-1.
[6] Muquit M A, Shibahara T, Aoki T. A High-Accuracy Passive 3D Measurement System Using Phase-Based Image Matching[J]. Ieice Transactions on Fundamentals of Electronics Communications & Computer Sciences, 2006, E89-A(3):686-697.
[7] Robotics; Nanjing University of Aeronautics and Astronautics Researchers Describe Findings in Robotics (Design and evaluation of a needle tip measurement system based on binocular vision) [J]. Robotics & Machine Learning, 2020.