Research on Floating Object Ranging and Positioning Based on UAV Binocular System

Peng Guo¹², Nengwen Zhang³, Bing Jiang¹², Fatao Qin³ and Zhiheng Zhou²

¹School of Electrical & Information Engineering, Wanjiang University of Technology, Ma’anshan 243011, Anhui, China
²College of The IOT Engineering, Hohai University, Changzhou 213022, Jiangsu, China
³Jiangsu Xindaoge Automatic Control Solutions Co., Ltd, Jiangsu 214432, China

Abstract. Aiming at the ranging and positioning problem of floating objects polluted by water surface, the binocular system of UAV is designed to obtain the parallax and distance information of target objects in the scene. Firstly, Zhang Zhengyou calibration method is used to calibrate binocular camera; Secondly, stereo correction is carried out on the image to obtain a standard binocular stereo vision model; Finally, the depth information of the target object is obtained by stereo matching algorithm. The simulation results show that the depth error is less than 5%, and the actual physical coordinate position of the floating object is obtained by combining the GPS information of the UAV itself.

Keywords. Binocular vision; floating object ranging and positioning; stereo matching algorithm.

1. Introduction

According to the general requirements of the country, after the implementation of the river length system, the protection and treatment of water resources has become a national challenge at present. At present, urban rivers are seriously polluted, and the water pollution incidents caused by this have attracted more and more attention from the society. Urban inland river is the “green lung” of a city, and the existence of floating objects on the water surface will affect water quality, aquaculture and ecological environment, etc. At the same time, the plastic components in floating objects of water pollution garbage are difficult to decompose for a long time, which leads to the destruction of the living environment of aquatic organisms. Binocular vision is based on computer vision, which uses binocular ranging and positioning principle to study the positioning and ranging of floating objects. Binocular vision is like human eyes, and the relative position of floating objects in UAV and pictures taken by UAV is measured according to the parallax between left and right eyes, so as to estimate the specific position of floating objects. At present, many ranging technologies, such as laser ranging and infrared ranging, have problems of high cost and low accuracy, and binocular ranging technology can solve such problems [1-2].

In recent years, with the wide application of artificial intelligence technology, intelligent processing based on vision technology has been widely concerned and applied to various fields. Binocular vision is a process of obtaining information by simulating human eyes through two cameras, and a method of restoring the same image from two-dimensional images to three-dimensional scenes through two different perspectives. For binocular vision imaging, it involves the transformation of four coordinate
systems, namely pixel coordinate system, image coordinate system, camera coordinate system and world coordinate system.

Li et al. proposed a vision-based AAR (Autonomous Aerial Refueling) technology, which adopted five algorithms of UAV height estimation based on monocular and binocular vision, and verified it in the extended vision system [3]. In 2009, Di et al. used the Winner-Take-All strategy to match feature points of two different views, and proposed a fast stereo matching algorithm. The matching speed increased by an order of magnitude, but the matching accuracy was relatively low [4]. In 2013, according to the principle of binocular stereo vision, Huang and others used a single camera to shoot the workpiece in different directions, and calculated the three-dimensional spatial coordinates of each feature point [5]. In 2018, Fu and others used Zhang Zhengyou’s calibration method for calibration, and proposed a method of stereo matching sub-pixel feature corners to reduce the stereo matching error, and the result has good matching accuracy [6]. In 2019, Kalogeiton et al. used a stereo vision camera to navigate, locate and avoid any possible obstacles in a fully autonomous robot, which can move the mapping area along the optimal path to the maximum extent [7].

In this paper, the UAV binocular ranging system is used to study the ranging and positioning of floating objects, the global search strategy, namely SGBM (Semi-Global Block Matching) algorithm, is selected, and the distance information between floating objects and UAV is obtained by matching the parallax information obtained by the algorithm, finally, the detection, identification and ranging and positioning of floating objects can be realized by combining the GPS information of UAV and the position and category of images in convolutional neural network.

2. Design of Binocular Stereo Vision System Platform for UAV 2

2.1. UAV GPS Information Acquisition

In this chapter, the airborne GPS M8N module is selected for information acquisition of UAV GPS, which has low power consumption, high precision and accurate positioning accuracy, with this module, the latitude, longitude, altitude and attitude information of UAV can be acquired, providing UAV position information for floating object positioning research, The GPS M8N module is shown in figure 1.

![Figure 1. GPS M8N module diagram.](image)

Because the UAV itself can’t get the physical coordinate information accurately, and the GPS module can record the UAV’s position in real time, so that the UAV can get the accurate position through latitude and longitude coordinates and maps during flight. In order to facilitate the subsequent manual salvage of floating objects, it is necessary to obtain the actual physical position of floating objects. Therefore, the UAV binocular system is used to calculate the relative distance between UAV and floating objects, and the relative coordinate information of floating objects is obtained by combining the longitude and latitude information of UAV. Because the distance calculation of binocular vision system is a relative physical coordinate obtained by taking the camera as the coordinate origin, the latitude and longitude and azimuth of this point can be known according to the UAV airborne GPS module, and then the latitude and longitude of the floater can be finally obtained.
through the formulas of trigonometric functions combined with the disparity value output by binocular matching, so as to obtain the actual physical position of the floater.

2.2. Construction of UAV Binocular System
In this paper, the UAV is equipped with binocular camera to complete the detection and positioning of floating objects in river pollution, the image information collected by UAV is acquired by binocular camera, and the image is stored and visually processed by TX2 calculation through transmission module, according to the positioning algorithm in binocular vision ranging system, the relative 3D information of floating objects is calculated, and finally the 3D information is sent to flight control system to realize the floating object positioning function. The system framework is shown in figure 2:

![System framework diagram](image)

**Figure 2.** System framework diagram.

At first, the binocular camera relates to TX2 through USB interface, and then TX2 relates to PixHawk through serial port. Mission Planner ground station runs on PixHawk, after installing the corresponding firmware, choose to start the obstacle avoidance function. After debugging the parameter table, you can view the data at MP ground station. Then, the left camera is placed in the center of the UAV to reduce the experimental error. Floating object ranging system based on binocular stereo vision uses binocular camera and SGBM matching algorithm to extract feature points, so as to obtain parallax information about images. First, the floater is detected and identified by the left camera, and the position of the floater in the image is drawn by a rectangular frame, Then, the distance information between the floater and the UAV can be obtained by the parallax information obtained by binocular system and the position and category of the image in convolutional neural network, combined with the GPS information of UAV, the floater can be detected, identified and located by UAV.

3. Binocular Calibration
Stereo calibration and stereo correction are the prerequisites for binocular stereo matching. Matlab R2016a development platform is used in this experiment environment. According to the method of Zhang Zhengyou’s calibration [8], taking the checkerboard grid as a reference, then taking binocular checkerboard images in different directions and different positions. The internal parameters of the camera are closely related to each part of the camera, and the internal parameters usually represent the conversion relationship between plane coordinates. The calibration process of binocular cameras mainly includes the following two processes: firstly, the left and right monocular cameras are calibrated with MATLAB calibration toolbox to obtain the internal parameters, and then the relative position relationship between the two cameras is solved. 35mm*35mm In this standard experiment, the checkerboard with the size of ф is selected to ensure that the checkerboard is unobstructed in the shooting process. Figure 3 shows the calibration diagram of Matlab.
Figure 3. Calibration chart of MATLAB.

4. Sections, Subsections and Subsubsections

Binocular Correction

According to the principle of binocular vision ranging, the most stringent standard binocular stereo vision model consists of two parallel cameras with identical internal parameters [9]. However, the binocular camera can’t achieve a good effect because of some errors produced in the process installation, so stereo correction is carried out on the binocular image, so that two non-parallel and coplanar binocular images are corrected to be in the same plane and parallel, and search is carried out from two-dimensional to one-dimensional, which not only improves the matching speed and accuracy of stereo matching, but also obtains a standard binocular stereo vision model, so that when the image is imported into stereo matching algorithm for processing, it can meet the standard of ideal model and achieve the effect of accurate ranging [10]. Comparison diagram of binocular system before and after calibration is shown in figure 4:

![Comparison of binocular system before and after calibration](image)

(a) uncorrected binocular system (b) corrected binocular system

Figure 4. Comparison of binocular system before and after calibration.

Before stereo matching, the image needs to be distorted [11-12]. After distortion removal, stereo rectification function provided by OpenCV is used to correct the original binocular image of T265 binocular camera. Figures 5-7 can be obtained by adjusting the parameters such as projection matrix and rotation matrix obtained at the standard timing.
Figure 5. Left and right images before correction.

Figure 6. Left and right images after correction.

Figure 7. Alignment image after correction.

A set of horizontal green lines obtained in figure 7 shows that the images are strictly aligned after stereo correction, which is very important for stereo matching algorithm to quickly search matching points to obtain disparity data.

In order to ensure that the left camera image can find the corresponding matching point in the right camera image in stereo matching, the parallax is calculated to the right from the (numDisparity-1+minDisparity) column of the right camera image as the starting point in OpenCV. Therefore, the obtained parallax image will be smaller than the image range obtained by the camera [13]. Therefore, it is necessary to extend the pictures obtained by the two cameras to the left by numDisparity column pixels, and then carry out stereo matching on the images, so as to minimize the data loss.

5. Footnotes Stereo Matching Algorithm
There are three stereo matching algorithms, namely global stereo matching, local stereo matching and feature matching [14]. Stereo matching algorithm can quickly and stably match binocular images, mainly by means of similarity measurement rules, constraint criteria, matching primitives and matching strategies [15]. Firstly, the image information corresponding to the matching primitives in the left image is extracted, and then the corresponding matching points are found under the constraint of the boundary condition criteria in the right image. The similarity of the matching points is evaluated according to the similarity measurement criteria, so as to find the most accurate matching point. Finally, the depth distance is calculated by the disparity value of the matching points in the left and right images [16].
Considering the needs of floating objects on the water surface in practical application scenarios and the accuracy requirements of UAV for target object recognition in flight, this topic adopts the global search strategy, namely SGBM (Semi-Global Block Matching) stereo matching algorithm [17], to extract the image features of the area to be determined, and use feature points as primitives for stereo matching. SGBM algorithm has been used in open source since OpenCV2.4.6, which is very convenient to use. The algorithm mainly includes the following three steps: selecting matching primitives, constructing the cost function of multi-directional scanning lines, and solving the optimal solution of the cost function. The matching process is:

Step1: Obtain the disparity value of the pixel points of the left and right images to form a disparity map;
Step2: Construct a global energy function to ensure correlation with disparity map;
Step3: minimize the energy function value, so that the disparity value of each pixel is the optimal disparity value.

6. Stereo Matching Test

6.1. Formation of Depth Map

Parallax map is mainly formed by stereo matching algorithm using binocular images, Its principle is that gray value is proportional to pixel value, and the relationship also represents the distance between object and camera. Through the formation of disparity map, the depth map can be obtained according to the conversion formula of disparity and depth. As shown in figure 8, the size of each pixel value in the depth map represents the distance between the actual object and the camera plane.

![Figure 8](image)

6.2. Calculation of the Distance

After the disparity map is obtained, a program is written to calculate the distance information of each corresponding pixel by binocular ranging principle, and calculate and compare the actual distance and the average distance as shown in table 1.

| Target | Actual distance mm | Average sampling point distance mm | Error |
|--------|-------------------|-----------------------------------|-------|
| one    | 1000              | 1028.83                           | 2.88% |
| two    | 1500              | 1552.23                           | 3.47% |
| three  | 1800              | 1869.6                            | 3.86% |
| four   | 2000              | 2080.5                            | 4.02% |
| five   | 2500              | 2601.6                            | 4.06% |

According to the above table data, the distance between the target object and the camera is within 3 meters, and the depth error is within 5%, which proves that the accuracy of the binocular ranging system meets the requirements. When the distance between the binocular camera and the target object is close, the depth error obtained by stereo matching algorithm will be lower. With the distance between the binocular
camera and the target object increasing, the information obtained by the binocular camera will decrease, the error will gradually increase, and the accuracy will also decrease.

6.3. Stereo Matching Test Results
In general, try to avoid extremely fine lines (often called ‘hairline’ thickness) because such lines often do not reproduce well when printed out—your diagrams may lose vital information when downloaded and printed by other researchers. Try to ensure that lines are no thinner than 0.25 pt. Note that some illustrations may reduce line thickness when the graphic is imported and reduced in size (scaled down) inside Microsoft Word. 

According to the principle of polar ray bundle, the right image point corresponding to a certain point in the left picture is also in line. If the "minimum parallax" parameter of stereo matching is set to 0, the regional window of stereo matching method will match from the right image. 

The characteristic points of the first root correspond to each other, and the window covering the left and right images takes $\phi$ as the center, and the extracted gray value uses SAD (Sum of absolute differences) function to find the sum of the absolute values of the corresponding regional gray values in the left image. $Y_n$ then estimate the match error, measure the similarity of the corresponding points of the left and right windows, move the match window in the right picture, and match the pixels one by one from right to left, repeating this process until the end of the designed maximum visual error search range. Finally, the best matching point is determined by finding the local optimal value, and $d = X_n - X_m$ is the stereo matching disparity. 

$z = \frac{fb}{d}$

In this process complex graph, enough matching points can be obtained by the method of matching, and then according to the principle of binocular distance measurement, the depth calculation formula can be used: "three-dimensional spatial information corresponding to objects can be obtained". The stereo matching test results are shown in figure 9.

![Figure 9. Stereo matching test results.](image)

The result of stereo matching shows that because every object exists in three dimensions, the distance from the target object to the camera is 168.389cm, and the relative world coordinates of the output target floater are $X: -268.315$, $Y: 505.425$, $Z: 1573.04$.

7. Conclusion
Research on the accuracy of stereo matching algorithm plays a key role in the excellent ranging effect. Firstly, the platform of UAV binocular system is built. Secondly, in order to eliminate the interference of noise, the image is preprocessed and smoothed, which can reduce the processing workload of stereo matching optimization algorithm and avoid the interference of unnecessary factors. Then, in order to better ensure that the stereo matching algorithm accurately obtains the matching points, stereo correction is carried out on the binocular image to ensure that the standard and feasible binocular image is obtained in the same plane. Finally, the disparity map of the image is obtained and converted into a depth map, through practical verification, the depth error of UAV binocular system is below...
5%. In this chapter, SGBM global matching algorithm is used to calculate the depth and distance information of the image, by performing such operations on the whole image, the pixel coordinate points on the two-dimensional image can be converted into the corresponding physical coordinate points, thus obtaining the distance accurately.

References
[1] Tang W, Liu S, Gao H, et al. 2019 Vision-based target detection algorithm for surface garbage cleaning robot Science, Technology and Engineering 19 (03) 136-141.
[2] Zhu Y Q 2017 Research on accurate distance measurement system based on ultrasonic technology File (26) 228.
[3] Han L and Duan H B 2016 Verification of monocular and binocular pose estimation algorithms in vision-based UAVs autonomous aerial refueling system Science China Technological Sciences 59 (11) 1730-1738.
[4] Di H W, Chai Y and Li K 2009 A fast stereo matching algorithm for binocular vision Journal of Optics 29 (08) 2180-2184.
[5] Huang F S, Liu E F, Fang Y X, et al. 2013 Monocular stereo vision recognition of parts pose based on intelligent CMM Optical Precision Engineering 21 (05) 1326-1332.
[6] Fu P, Chen X D and Wu L 2018 Research on absolute positioning of binocular vision based on corner Journal of Electronic Measurement and Instrument 32 (03) 1-8.
[7] Kalogeiton V S, Ioannidis K, Sirakoulis G C and Kosmatopoulos E B 2019 Real-time active slam and obstacle avoidance for an autonomous robot based on stereo vision Cybernetics and Systems 50 (3) 1-22.
[8] Zheng Y Z 2000 A flexible new technique for camera calibration IEEE Transactions on Pattern Analysis and Machine Intelligence 11 (22) 1330-1334.
[9] Xiong W, Li W L and Yin Z P 2013 Design and implementation of stereo matching method in binocular PMP vision measurement China Mechanical Engineering (019) 2635-2640.
[10] Redmon J, Divvala S, Girshick R, et al. 2016 You only look once: Unified, real-time object detection Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition pp 779-788.
[11] Lin T Y, Dollar P, Girshick R, et al. 2017 Feature pyramid networks for object detection Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition pp 2117-2125.
[12] He K, Zhang X, Ren S, et al. 2016 Deep residual learning for image recognition Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition pp 770-778.
[13] Guo Q W and Song D M 1994 Implicit and explicit camera calibration: Theory and experiments IEEE Transactions on pattern Analysis & Machine Intelligence 16 (5) 469-480.
[14] Deng Z J 2020 Underwater Stereo Matching Technology Based on Binocular Stereo Vision (Zhejiang University).
[15] Bai M, Zhuang Y and Wang W 2008 Research and development of binocular stereo matching algorithm Control and Decision (07) 3-11.
[16] Zhang Z Q and Shi W H 2019 Improved scale invariant feature transformation algorithm for parallel acceleration of binocular ranging system and its implementation Advances in Laser and Optoelectronics 56 (14) 201-210.
[17] Zhang H, An L, Zhang Q, et al. 2016 The analysis and research of SGBM algorithm and BM algorithm Surveying and Mapping and Spatial Geographic Information 39 (10) 214-216.