Research on Working Target Detection of Bulldozer Based on Binocular Vision

Mingchun Li 1, Quancheng Dong 1,*, Renpeng Xing 2, Shukun Cao 1, Xuan Sun 1, Changchun Ma 1

1School of Mechanical Engineering, University of Jinan, Shandong Province, China.
2Institute of Automation Shandong Academy of Sciences, Shandong Province, China.

*Corresponding author e-mail: me_dongqc@ujn.edu.cn

Abstract. Aiming at the problem of bad working environment of bulldozer in landfill, this paper presents a target detection method suitable for autonomous working bulldozer. Firstly, the working environment information was obtained by the binocular camera, the maximum inter-class variance method was used to remove the sky background by binarization of the image, and the color component of the ground was extracted from the HSV color space and converted into a binarization image. The absolute value of the difference between the two binarization images was the approximate area of the working target. Then the binocular camera was calibrated and the feature points were detected and matched in the target area. The parallax was obtained by triangulation principle to obtain the depth information of the target. The experimental results show that the method can correctly extract the depth information of the target area and target, and the accuracy is high, which can meet the requirements of the bulldozer in the landfill for target detection.

1. Introduction

Bulldozer is widely used in construction, road, national defense construction and other fields, but it has problems such as low degree of automation and high labor intensity, especially in work places such as landfill and waste disposal, which will pose a threat to the physical and mental health of drivers [1]. Therefore, the digital, intelligent, and unmanned bulldozer has become the main development trend of the current bulldozer.

Bo Hu et al. [2,3] realized the automatic control of the rise and fall of the bulldozer by detecting the engine speed, the oil pressure in the lifting cylinder, the strik-slip rate and other signals to feedback the force situation of the bulldozer. Yong-seok et al. [4,5] designed a detection system for estimating the position of bulldozer shovel, and realized the calculation of the position of bulldozer shovel by combining the real-time data of sensors. Masami Hirayama et al. [6] developed a path planning method suitable for autonomous bulldozers by using machine vision to map the terrain.

At present, many researchers are engaged in the research work of bulldozer autonomous operation, but the research of bulldozer operating target recognition is less. Based on this, this paper puts forward a detection method of operation target for the landfill, which is used to provide the operation target information to the bulldozer and guide the bulldozer to complete the follow-up work.
2. Binocular vision target detection system

2.1. Hardware Scheme Design

Binocular camera is used as the sensor, and the left and right cameras can output images at the same time. The operation target is extracted from the image data and the parallax is calculated, so as to obtain the depth information of the operation target, as shown in Figure 1, which is the parallax ranging principle of the binocular camera [7].

![Figure 1. Principle of parallax ranging for binocular camera](image)

In figure 1, P is the point in space, O_L and O_R were around the camera's optical center, B is the distance between camera optical center around the baseline, f is the camera focal length, T_L and T_R is respectively about the camera's sensor imaging plane, namely P_L (X_L, Y_L) and P_R (X_R, Y_R) are respectively about camera in imaging plane of the imaging points, thus can get around the parallax of the camera as follows:

\[ d = |X_L - X_R| \] (1)

The distance between two imaging points P_L and P_R is:

\[ L = B - (X_L - X_R) \] (2)

According to the similar triangle principle, the distance Z of point P can be calculated as:

\[ Z = \frac{bf}{X_L - X_R} \] (3)

It can be seen from Equation 3 that parallax b is inversely proportional to distance Z, and both baseline b and focal length f can be obtained through calibration tests. The depth information of a point in space can be obtained by only requiring parallax b.

2.2. Software Scheme Design

Based on the development environment of Visual Studio 2015 + OpenCV4.1, the image is preprocessed and the job target region is extracted by the job target detection algorithm. Then SURF feature point detection and FLANN matching were carried out for the target area. Finally, the image is corrected by the internal and external parameters of the calibrated camera, the parallax is obtained and the depth information is calculated.

The calibration of the binocular Camera was carried out using the STEREO Camera Calibrator toolbox in Matlab. In this paper, 15 groups of pictures were shot and imported into the toolbox to conduct calibration experiments for the binocular Camera, and the required internal and external parameters were obtained, as shown in Figure 2.
3. Job target detection algorithm

3.1. Characteristics analysis of job objectives

The operating targets of the bulldozer in the landfill site mainly have the following characteristics: the landfill site is far away from the city, and the background is generally the sky with high brightness; The foreground is the work target and the work site, the work site is usually the land, and the work target is the waste dump.

3.2. Extraction method of job targets

The maximum inter-class variance method is used to perform binarization processing on the images to remove the sky background, and then the interference areas on the ground are extracted through HSV color space. The absolute values of the two images are subtracted to obtain the target area. The specific process is shown in Figure 3.

Figure 3. Flow chart of job target extraction algorithm

(1) Remove sky background. After the image data is obtained, the image is preprocessed and the image is denoised by Gaussian filter. Then the Ostu algorithm is used to transform the image into a binary image and remove the sky background. As shown in Fig. 4 (a) and (b), they are the original and Ostu binarization graphs.
4. Depth detection of operational targets

4.1. Feature point detection of the job target
After the operation target area is extracted, the feature points of the operation target area are detected. In this paper, SURF feature points detection algorithm is adopted, which has high accuracy and stability, as shown in Figure 7.
4.2. Matching the feature points of the job target
Flann matching algorithm was adopted to match the feature points, and RANSAC algorithm was used to eliminate the mismatched points to improve the accuracy of matching. As shown in Fig. 8 (a) and (b), they are the matching graphs of unoptimized feature points and optimized feature points.

![Figure 8. Feature point matching graph](image)

4.3. Calculation of depth information
Based on the internal and external parameters and stereoRectify function of binocular camera, the rotation matrix of binocular camera correction is obtained. Then initundirectifymap function and remap function are used to correct and transform the image so that the two images have the same Y coordinate and the parallax dis calculated. In addition, stereoRectify function calculates the mapping matrix $Q$ based on binocular camera parameters:

$$Q = \begin{pmatrix}
1 & 0 & 0 & -c_x \\
0 & 1 & 0 & -c_y \\
0 & 0 & 0 & f \\
0 & 0 & -\frac{1}{T_x} & \frac{c_x-c_x'}{T_x}
\end{pmatrix}$$

(4)

Where, $(c_x, c_y)$ is the deviation from the origin of the left camera, $c'_x$ is the coordinate of the origin of the right camera, $f$ is the focal length, and $T_x$ is the distance between the optical centers of the left and right cameras. Through the matrix $Q$ and the parallax $D$, we can get:

$$Q \begin{bmatrix}
x \\
y \\
d
\end{bmatrix} = \begin{bmatrix}
x - c_x \\
y - c_y \\
f \\
D + c_x - c_x'
\end{bmatrix} = \begin{bmatrix}
X \\
Y \\
Z \\
W
\end{bmatrix}$$

(5)

In binocular cameras, the left camera is generally the main camera [8], so the final obtained $(X/W, Y/W, Z/W)$ is the three-dimensional coordinate value in the real space.

5. Experimental verification
5.1. Construction of the experimental environment
Due to the restriction of experimental environment and experimental equipment, this paper adopts the method of simulating construction scene. The stacking of the target material pile was simulated on an outdoor open ground, and the state of the binocular camera installed on the bulldozer was simulated...
with the camera bracket. The notebook was connected with the binocular camera to calculate and process the image data.

5.2. Operation target distance detection experiment

In the simulation test of operating target distance detection, a total of 5 groups of operating target images with different distances were collected, and the actual distances were 1.0m, 1.5m, 2.0m, 2.5m and 3.0m respectively. The baseline distance of the detection camera was 60mm and the focal length was 484Pixel. In the test, the accuracy of the experimental results is measured according to the relative error between the actual distance and the detection distance. The calculation formula of the relative error is as follows:

$$\delta = \frac{|X - L|}{L} \times 100\%$$  \hspace{1cm} (6)

Where, $\delta$ is the relative error, $X$ is the actual measured value, and $L$ is the true distance value. The test results are shown in Table 1.

| Serial number | Actual distance(mm) | Detection distance(mm) | Relative error(%) |
|---------------|---------------------|------------------------|-------------------|
| 1             | 1000                | 1021.31                | 2.13%             |
| 2             | 1500                | 1546.32                | 3.08%             |
| 3             | 2000                | 2086.21                | 4.31%             |
| 4             | 2500                | 2604.57                | 4.18%             |
| 5             | 3000                | 3152.73                | 5.09%             |

As can be seen from Table 1, the relative error of distance detection in the simulation test ranged from 2.13% to 5.09%. Due to the close baseline and low accuracy of the binocular camera in the test, the ranging range can only be within 4m, and there are errors.

5.3. Operation target width and height detection test

For the detection of the width and height of the target, the method adopted in this paper is to solve the vertex coordinates A (XA, YA) at the upper left corner of the rectangular box surrounding the target area in the left image and B (XB, YB) at the lower right corner, and calculate the width W and height H of the target through formula 7.

$$\begin{cases} W = X_B - X_A \\ H = Y_B - Y_A \end{cases}$$  \hspace{1cm} (7)

Since these two vertices are not feature points in the actual debugging, the parallax of feature points on the job target is adopted to obtain the coordinate values of the two points in the actual debugging. The actual width of the operating target in the test is 885mm, and the actual height is 325mm. The test data are shown in Table 2.

| Serial number | Actual distance(mm) | Detection width(mm) | Relative error(%) | Detection height(mm) | Relative error(%) |
|---------------|---------------------|---------------------|-------------------|----------------------|-------------------|
| 1             | 1000                | 935.35              | 5.38              | 338.81               | 4.25              |
| 2             | 1500                | 918.27              | 3.76              | 331.94               | 2.14              |
| 3             | 2000                | 961.62              | 8.66              | 347.23               | 6.84              |
| 4             | 2500                | 864.57              | 2.31              | 315.22               | 3.01              |
| 5             | 3000                | 939.76              | 6.19              | 343.15               | 5.58              |

As can be seen from Table 2, the relative error of target width detection is between 2.31%-8.66%, and that of target height detection is between 2.14%-5.58%.
6. Conclusion
The bulldozer uses a binocular camera as a sensor to obtain the operating target and calculate the depth information through image data, which provides a foundation for the future unmanned bulldozer. Through the simulation test, it is verified that this paper can extract the working target area correctly, and the measurement error is within the acceptable range, and its accuracy and calculation speed can better meet the working target detection of bulldozer.

Acknowledgments
This work was financially supported by Key R&D project of Shandong Province (2019JZZY010443) fund.

References
[1] Feifei Fang, Yan Liu, Shicong Zhang. Application of WIFI Wireless Remote Control Technology on Bulldozer [J]. Construction Mechanization, 2019, 40(01): 19-20.
[2] Weikun Xia. Research on Automatic Control System of Bulldozer Working Device [D]. Chang'an University, 2016.
[3] Lingxiao Ban. Research on Computer Control System of Bulldozer Shovel Lifting [D]. Chang'an University, 2007.
[4] Lee Y S, Kim S H, Seo J, et al. Blade control in Cartesian space for leveling work by bulldozer [J]. Automation in Construction, 2020, 118: 103264.
[5] Kim S H, Lee Y S, Sun D I, et al. Development of bulldozer sensor system for estimating the position of blade cutting edge [J]. Automation in construction, 2019, 106(Oct.): 102890.1-102890.11.
[6] Masami Hirayama, Jose Guivant, Jayantha Katupitiya, Mark Whitty. Path planning for autonomous bulldozers [J]. Mechatronics, 2019, 58.
[7] Peng Shu Yang, Hui Liu, Xiaocui Wang, Qiao Wang, Zhijun Meng. Three-dimensional obstacle detection method for crops at seedling stage based on binocular vision [J]. Journal of Agricultural Mechanization Research, 201, 43(04): 11-16.
[8] Haowen Feng. Research on 3D Measurement of Material Stack Based on Vehicle Stereo Vision [D]. Wuhan University of Technology, 2019.