Calculation method of Cross Section area of collapsing dangerous Rock based on parallel Binocular Vision

Ping Gan¹, Zhenzhen Ruan¹, Yang Wang¹, Lin Huang¹, Yanyun Li¹ and Yangfan Huang¹

¹School of Microelectronics and Communication Engineering, Chongqing University, Chongqing, China
vigor_gp@163.com, r1041942883@163.com, wy20124448@163.com, 1244358767@qq.com, 815183417@qq.com, hyf@cqu.edu.cn

Abstract. Dangerous rock collapse may cause traffic accidents and other potential traffic safety hazards in mountainous areas. In this paper, aiming at the problem that dangerous rock collapse disaster is difficult to monitor and judge, a method of measuring the cross section area of collapse rock is proposed based on parallel binocular vision, which is convenient for the relevant departments to deal with the emergency ahead of time. This paper first calibrates the binocular camera on the MATLAB platform, and then corrects and matches the collected images on the OpenCV platform. The feasibility of measuring area by pixel method and spatial coordinate method is verified by using regular triangular and rectangular shapes. Then the pixel method is applied to the measurement of cross sectional area of rock collapse, which provides a method for predicting the disaster grade of collapse of dangerous rock.

1. Introduction
The disaster of dangerous rock collapse in mountain area mostly occurs in the remote mountainous area. Because of its complex environment, remoteness, sudden occurrence, and difficult to determine the disaster grade, it is particularly difficult to monitor and warn the disaster of dangerous rock collapse. With the development of computer vision technology, the methods to infer various three-dimensional spatial problems through two-dimensional image information have been widely used in various industries and become one of the hot research fields, but in computer vision, binocular vision has the advantages of high efficiency, suitable precision and low cost. It is suitable for non-contact measurement on the site of highway rock collapse. In this paper, the binocular vision system is applied to the measurement of the cross sectional area of the collapse rock, which facilitates the relevant management departments to release the information in advance, to deal with the emergency and to grasp the field situation in real time. It helps to reduce threats to people's personal and property security.

The realization process of the pixel method in this paper is as follows: Firstly, the distance parameters and 3D coordinates of the object are obtained by binocular vision, and then the regular triangular and rectangular images are preprocessed. Then the object is segmented by the method of inter-class variance, then the pixel area of the object in the image is obtained, and finally the actual area is obtained. The method is applied to measure the cross-sectional area of collapse rock.

2. Binocular stereo vision principle
According to the principle of binocular stereo vision, two cameras are used to obtain the information
of the same object in the three-dimensional world.

2.1 Binocular stereovision model
Compared with the camera arrangement of the intersecting binocular optical axis, in stereo vision, if the left and right cameras meet the conditions: the central line of the left and right cameras is their common X axis and the optical axes of the two cameras are parallel [1], the two cameras are called parallel alignment state.

This paper is based on the parallel optical axis model. The principle of a typical binocular visual imaging system without distortion[2] is shown in figure 1 below:

![Figure 1. Parallel binocular visual model.](image)

The imaging points of the space point P in the image plane of the left and right camera are respectively \( P_L, P_R \). Through the imaging model of the camera, the relationship between the three-dimensional position of the target in space and the imaging point in the image can be determined. The geometric model parameters of camera imaging can be solved by transforming the pixel coordinate system to the world coordinate system [3].

The pixel coordinate system represents the number of rows and columns of a pixel point in the image. The corresponding point of space point \( p \) on the imaging plane is \((u, v)\) in the pixel coordinate system, and the world coordinate of the space point \( P \) is \((x_w, y_w, z_w)\), the transformation relationship between the world coordinate system and the pixel coordinate system can be obtained as follows: (here \( A \) is the inner parameter matrix, \([R \ t]\) is called the outer parameter matrix, \( R \) denotes the rotation of the camera from the origin of the world coordinate, and the translation of \( t \) means the translation.)

\[
s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} [R \ t] \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = A[R \ t] \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}
\]

2.2 Binocular camera calibration
In the calibration process of binocular stereo vision system, the main task is to obtain camera internal parameters, external parameters and distortion parameter vectors. The inner parameters are only related to the camera itself, while the outer parameters (rotation matrix and translation vector) are determined by the relative position and attitude of the camera and the world coordinate system.

In this paper, MATLAB version of the toolbox_Calib(Camera Calibration Toolbox based on MATLAB) calibration toolbox is used to calibrate, and then the subsequent image correction and matching processing are carried out on the Open CV platform. The calibration board is calibrated with a 12 × 9 black and white checkerboard with a high accuracy of 5 × 5 mm for each cell [6]. In the process of collecting the image of calibration board, we try to make the calibration board in different position in the image, different angle, the top and bottom of the screen are covered to a certain extent, the number is about 15-20 groups is more suitable. The 20 groups of pictures taken by the camera in this paper are shown in figure 1.2 below.
2.3. Image correction

After calibrating and getting the inner and outer parameter matrix of the camera, it is necessary to correct the images taken by the left and right cameras.

The plane of the left and right images is coplanar and the line pair is punctual, so that the stereo parallax can be calculated. The correction consists of two steps: one is to eliminate the distortion caused by the lens to the image [7][8][9], which is a transformation for a single camera. Two is binocular parallel correction, is for binocular camera operation. The distortion is divided into radial distortion and tangential distortion. The radial distortion is caused by the technological problems of the lens itself, which is usually much larger than the tangential distortion. Therefore, the radial distortion is only considered in this paper, but the tangential distortion is ignored.

The following (a) is a calibration board image collected directly from the camera, and the edge of the image is obviously distorted, (b) as the corrected image of the calibration board. It can be clearly compared that the edges of the original image with obvious curved edges are well corrected.

In this paper, the stereoRectify() function of Open CV platform is selected for image correction:

It can be seen that the result of image correction is ideal and the right and left image matching points after correction are at the same horizontal line.
3. image preprocessing
It is necessary to preprocess the images collected by the camera before the area measurement, which mainly includes image denoising, image segmentation and so on.

3.1. image denoising
In image processing, before further processing such as image segmentation, the first step is to reduce the noise of the image to a certain extent. In this paper, the median filtering method is used. Its principle is to set the gray value of each pixel point to the median value of all pixel points in a neighborhood window of that point, so as to eliminate the isolated noise points.

The following (a) is the gray image of the original rectangular image, and the figure (b) is the image after the median filter. It can be seen that the noise effect has been filtered out in a certain part.

![Figure 6. Original rectangular image.](image1)
![Figure 7. Image after median filtering.](image2)

3.2. image segmentation
In order to further process the target, it is necessary to separate the object from the background first. In this paper, the method of inter-class variance is used to segment the image. The method divides the image into two parts: target and background according to the gray characteristic of the image. It can obtain satisfactory results when the image quality is better and the background changes steadily. The following (a) is a gray image after denoising, and the graph (b) is a binary image which is segmented by the method of inter-class variance.

![Figure 8. Image after median filtering.](image3)
![Figure 9. The binary image after segmentation.](image4)

4. Area calculation method

4.1. Pixel method
Suppose that the widths of a single pixel in the x direction and y direction in the image plane are $x_p$, $y_p$, respectively, and the actual widths in the x direction and y direction in the plane with depth of field $L$ are $X_p$, $Y_p$ [10], respectively. The focal length of the camera is $f$. Based on the proportional relation of the pinhole imaging model (only the x direction model is drawn, the y direction is the same):
Then the actual area of a single pixel on a plane with a distance of L is [5]:

\[ a = X_p Y_p \]  

(6)

Then calculate the number of pixels of the object after image preprocessing, that is, the pixel area is A, we can calculate the actual area of the object according to the formula:

\[ S = A \cdot a = AX_p Y_p \]  

(7)

After image correction, the SGBM matching algorithm is implemented on the platform of OpenCV. The parallax map can be obtained by matching the left and right images. Through the parallax image, the three-dimensional information of image pixels can be obtained in the command window. By reading out the value of z axis in its 3D coordinate, the depth of field distance data L is obtained, and the actual area of the object is calculated.

4.2. Spatial coordinate method

After matching according to binocular stereo vision, according to parallax, the information of 3D world obtained is as follows:

\[
\begin{aligned}
X_w &= \frac{BX_{\text{left}}}{d} \\
Y_w &= \frac{BY}{d} \\
Z_w &= \frac{f}{d}
\end{aligned}
\]  

(9)

Where d is the parallax of the space point in the X direction of the left and right image plane, B is the baseline distance of the binocular camera, and f is the focal length. In this paper, the reprojectToImage() function of OpenCV platform is used to calculate the three-dimensional coordinates. For regular graphs such as triangles and rectangles, we can calculate the side length according to the spatial coordinates of each vertex, and then calculate the area according to the formula [10] of each area of the regular image.

5. Experimental result

5.1. Measure the area of a regular pattern

1) Pixel method

We use the above method to calculate the area of triangle and rectangle respectively. The experimental data are shown in the following table:

Table 1. Pixel method.

|         | triangle | rectangle |
|---------|----------|-----------|
| Number of pixels | 118616   | 250321    |
| distance(mm)    | 350.912mm| 350.912mm |
| area            | 147.68cm²| 311.66cm²|

The actual area of the triangle is 137.38 cm² and the actual area of the rectangle is 288.86 cm², so the error of the triangle area measured by the pixel method is as follows:
The error of rectangular area measured by pixel method is as follows:

\[ e_1 = \left| \frac{s_1 - s_1^*}{s_1} \right| \times 100\% = 7.5\% \]  

(10)

The error of rectangular area measured by pixel method is as follows:

\[ e_2 = \left| \frac{s_2 - s_2^*}{s_2} \right| \times 100\% = 7.89\% \]  

(11)

2) Space coordinate method

According to the reprojectToImage() function based on the OpenCV platform in this paper, the spatial coordinates of each vertex of the rectangle are obtained respectively (-34.8004, -89.3379, 350.912), (170.375, -88.4611, 350.912), (-35.9695, 52.999, 350.912), (170.375, 52.4144, 350.912). Because there will be some errors, the difference between the left and right coordinates or the average values of the upper and lower coordinates will be taken to calculate the area of the side length. The area of the rectangle can be calculated as 291.37 cm² by the rectangular area formula. On the basis of the experimental data, the area of the triangle is obtained as follows:

|     | triangle | rectangle |
|-----|----------|-----------|
| area | 146.34cm² | 291.37cm² |

The actual area of the triangle used in the measurement experiment is 137.38 cm², the actual area of the rectangle is 288.86 cm², so the error between the triangle area and the actual area measured by the spatial coordinate method is 6.5%. The area error of the rectangle is 0.87%. From the experimental data above, we can see that the error of the spatial coordinate method is smaller than that of the pixel method.

5.2. Measurement of Cross Section area for Slope collapse

Based on the above experimental results, the pixel method calculates the area error within an acceptable range, and the collapsed slope rock is an irregular object, which cannot be calculated by the existing area formula using the space coordinate method. Therefore, this paper applies the pixel method to the measurement experiment of the cross-sectional area of the slope.

Since the cross section of the camera can also be an irregular pattern, taking the image of the collapsed rock collected by the left camera as an example, the gray image of the original image and the image corrected by binocular are as follows:

![Grayscale image of original image](image1)

![Rectified image](image2)

Since the collapsed rock is a three-dimensional object, the depth value of each point calculated by OpenCV is different. In this paper, the average depth of the rock mass surface is used as the actual distance of the rock. The pixel method is then used to calculate the actual cross-sectional area of the collapsed rock.

Before seeking the average depth value of the entire collapsed dangerous rock, we first need to distinguish the depth of the entire collapsed rock from the background[11]. According to the specific experimental environment, we set the depth threshold to [100,400]. The depth values of the pixels within the threshold range are averaged. The following image is a depth map of the collapsed rock and a binary image:
According to the pixel method, the results of the calculation are shown in the following table:

| Number of pixels | Distance | Area   |
|------------------|----------|--------|
| data             | 176923   | 191.91mm | 67.936 cm² |

At the actual site where the dangerous rock collapses, the level of the disaster can be estimated according to the actual measured cross-sectional area of the collapsed rock, and the relevant departments are given an alarm to classify the relevant measures in advance.

6. conclusion
When applied to the actual slope, because the background environment of the site is more complicated, it is necessary to improve the performance of the equipment and the accuracy of image preprocessing. The error of the experiment will be smaller and the prediction accuracy will be better.

References
[1] Ding-Cai, C., Ding-Cheng, W., & Jin-Shui, Z. (2006). Research on measurement of realistic planting leaf area based on machine vision. *Journal of Computer Applications*, 26(5), 1226-1228.
[2] Song, L. M., Wang, M. P., Lu, L., & Jing Huan, H. (2007). High precision camera calibration in vision measurement. *Optics & Laser Technology*, 39(7), 1413-1420.
[3] Gherardi, R., Farenzena, M., & Fusiello, A. (2010). Improving the efficiency of hierarchical structure-and-motion. *Computer Vision & Pattern Recognition*. IEEE.
[4] Okutomi, M., & Kanade, T. (2002). A multiple-baseline stereo. *IEEE Computer Society Conference on Computer Vision & Pattern Recognition*, Cvpr.
[5] Yakimovsky, Y., & Cunningham, R. (1978). A system for extracting three-dimensional measurements from a stereo pair of tv cameras. *Computer Graphics & Image Processing*, 7(2), 195-210.
[6] Bleyer, M., & Gelautz, M. (2007). Graph-cut-based stereo matching using image segmentation with symmetrical treatment of occlusions. *Signal Processing: Image Communication*, 22(2), 127-143.
[7] Shih, S. W., Hung, Y. P., & Lin, W. S. (1992). Efficient and accurate camera calibration technique for 3-d computer vision. *Proceedings of SPIE - The International Society for Optical Engineering*, 1614, 133-145.
[8] Lenz, R. K., & Tsai, R. Y. (2003). Techniques for calibration of the scale factor and image center for high accuracy 3D machine vision metrology. *IEEE International Conference on Robotics & Automation*. IEEE.
[9] Wei, G. Q., & Ma, S. D. (2002). Implicit and explicit camera calibration: theory and experiments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(5), 469-480.
[10] Geiger, D., Ladendorf, B., & Yuille, A. (1995). Occlusions and binocular stereo. *International Journal of Computer Vision*, 14(3), 211-226.
[11] Se, S., Lowe, D., & Little, J. (2001). Vision-based Mobile Robot Localization And Mapping using Scale-Invariant Features. *IEEE International Conference on Robotics & Automation*. IEEE.