Tech Science Press

A 3D Measurement Method Based on Coded Image

Jinxing Niu[1,*], Yayun Fu[1], Qingsheng Hu[1], Shaojie Yang[1], Tao Zhang[1] and Sunil Kumar Jha[2]

[1]Institute of Mechanics, North China University of Water Resources and Electric Power, Zhengzhou, 450011, China
[2]IT Fundamentals and Education Technologies Applications, University of Information Technology and Management in Rzeszow, Rzeszow, 100031, Poland
*Corresponding Author: Jinxing Niu. Email: njx.mail@163.com
Received: 11 February 2021; Accepted: 11 April 2021

Abstract: The binocular stereo vision system is often used to reconstruct 3D point clouds of an object. However, it is challenging to find effective matching points in two object images with similar color or less texture. This will lead to mismatching by using the stereo matching algorithm to calculate the disparity map. In this context, the object can't be reconstructed precisely. As a countermeasure, this study proposes to combine the Gray code fringe projection with the binocular camera as well as to generate denser point clouds by projecting an active light source to increase the texture of the object, which greatly reduces the reconstruction error caused by the lack of texture. Due to the limitation of the camera viewing angle, a one-perspective binocular camera can only reconstruct the 2.5D model of an object. To obtain the 3D model of an object, point clouds obtained from multiple-view images are processed by coarse registration using the coarse SAC-IA algorithm and fine registration using the ICP algorithm, which is followed by voxel filtering fusion of the point cloud. To improve the reconstruction quality, a polarizer is mounted in front of the cameras to filter out the redundant reflected light. Eventually, the 3D model and the dimension of a vase are obtained after calibration.

Keywords: 3D reconstruction; structural light; SAC-IA; ICP; voxel filtering

1 Introduction

Artificial intelligence-based computer vision has been replacing human vision in many fields, because of its excellent performance in describing and recognizing the objective world by processing camera-captured images. Specifically, it is expected to perform like the human visual system to process three-dimensional (3D) images, achievement of which can be greatly helpful for research and development in the industry [1–3]. For instance, it can help accurately identify the defective parts of an object with a high efficiency, which greatly reduces the workload that is originally done by humans.

As one of the most representative technologies in the field of optical 3D measurement, the fringe projection technology is favored by various features, including non-contact, high universality, high resolution, high precision, and high speed [4,5]. Meanwhile, the binocular stereo vision

and coded light method has been commonly used to reconstruct and measure 3D objects in many fields (e.g. industrial production, cultural relic protection, and 3D printing) [6–8], and they have the advantages of low equipment cost and high model accuracy. In this study, the fringe projection technology is used to reconstruct and measure a 3D object, which is accompanied by analyses of (1) the principle of binocular camera imaging, calibration, and correction, (2) stereo matching algorithms, such as the block matching (BM) algorithm and the semi-global stereo matching (SGBM) algorithm, (3) performances of the integrated binocular camera and Gray code fringe projection in reconstructing a single perspective model of an object, (4) roles of coarse- and fine-registration algorithms in generating point clouds, and (5) effects of the integration of sample consensus initial alignment (SAC-IA) and iterative closest point (ICP) algorithms. Eventually, a complete 3D model of a vase is established after point cloud fusion processing, followed by the corresponding reconstruction error analysis.

2 Principle

2.1 Binocular Stereo Vision

The principle of stereo vision technology is to reverse the camera imaging process that projects 3D points in the real world to a 2D image plane, which can be simply described by the pinhole imaging model [9], as illustrated in Fig. 1.
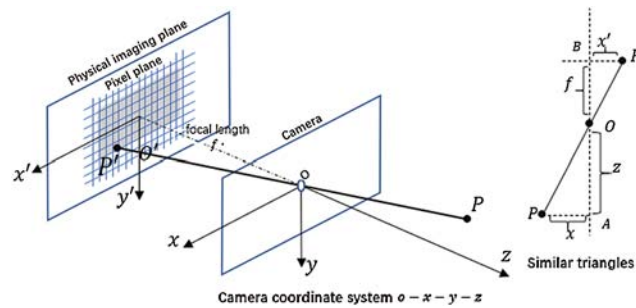

Figure 1: Pinhole imaging model

According to the similar relationship of the triangles, the following relationship exists:

$$\frac{Z}{f} = -\frac{X}{X'} = -\frac{Y}{Y'} \tag{1}$$

where Z is blabla, f is the focal length of the camera, X is blabla, X' is blabla, Y is blabla, Y' is blabla, and the sign in the formula indicates the direction, which is positive from left to right. $P(X, Y, Z)$ and $P'(X', Y', Z')$ are respectively the object space point and the corresponding image point in the camera coordinate system $O - xyz$.

There are three coordinate systems, namely the pixel coordinate system, the camera coordinate system, and the world coordinate system (Fig. 2). Therefore, it is necessary to establish a conversion relationship between the object point and the image point. $P(X, Y, Z)$, $P'(X', Y', Z')$, and $P(u, v)$ are corresponding point coordinates in the world coordinate system, the camera coordinate system, and the pixel coordinate system, respectively. The image becomes upright if the imaging

plane is projected symmetrically in front of the pinhole. In this context, formula (1) can be rewritten as:
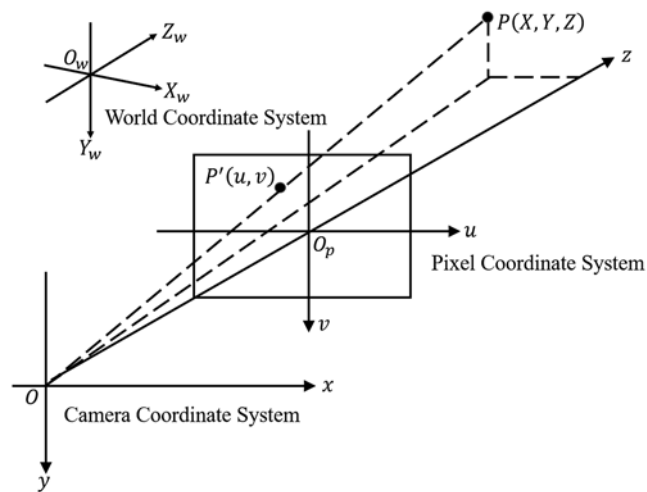
$$\frac{Z}{f} = \frac{X}{X'} = \frac{Y}{Y'}$$

(2)

Figure 2: The relationship between three coordinate systems

2.2 Calibration

Camera calibration is to calculate the internal and external parameters of the imaging system. Specifically, the structured light system calibration essentially determines the relationship between a point in the world coordinate system and the counterpart in the camera coordinate system [9]. A diffuse reflection board is used in this study for calibration (Fig. 3), which has a black and white checkerboard pattern with 12×9 grids. Each grid size is $20m \times 20mm$. The top right vertex is selected as the origin O_w of the world coordinate system, the horizontal and longitudinal sides of the checkerboard are the X_w axis and the Y_w axis, respectively, while the Z_w axis is perpendicular to the checkerboard. The three coordinate axes X_w, Y_w, and Z_w conform to the right-hand rule. The Z-coordinate of each point on the checkerboard plane is 0. The internal and external parameter matrices are obtained according to the chessboard camera calibration method by Zhang et al. [10].

2.3 Stereo Matching

Stereo matching is the key to 3D reconstruction, since it affects directly the accuracy of the reconstructed 3D model. BM and SGBM algorithms are commonly used for local stereo matching [11], and the matching units are generally associated with the image features, such as corners, contour edges, and inflection points. When applying the feature-based stereo matching algorithm, the first step is to extract the feature regions of the left and right images, while the second step is to take the left image as the reference and find the matching points on the corresponding bipolar line in the right image. The sparse disparity map can be obtained after repeating the above process to find all the feature point pairs.
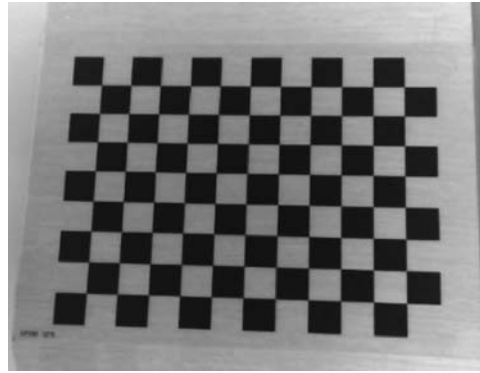
Figure 3: Diffuse reflection calibration board

2.4 Structured Light Coding and Decoding

As an active 3D measurement method [12], the coded structured light method is favored by its high measurement speed, high matching accuracy, and thus suitability for images with few features. It is able to add some texture information to the measured object and obtain better stereo matching by projecting a series of coded patterns to the object.

Binary coding is commonly used in structured light, where black is denoted as 0 and white is denoted as 1. In this way, different stripes can be distinguished in the projected image. Gray code is developed from, and more reliable than, binary code [13,14]. Assuming that a binary code can be represented as $B_{n-1}B_{n-2}\ldots B_2B_1B_0$ and its corresponding Gray code as $G_{n-1}G_{n-2}\ldots G_2G_1G_0$, the transformation relationship between them can be expressed by formula (3), where $G_{n-1}G_{n-2}\ldots G_2G_1G_0$ and $B_{n-1}B_{n-2}\ldots B_2B_1B_0$ have the same highest bit, and the rest are the XOR values between the current bit binary code and the previous bit binary code. The operational symbol $\oplus$ represents the XOR operation.

$$\begin{cases} G_{n-1} = B_{n-1} \\ G_i = B_{i+1} \oplus B_i, \quad i = 0,\ 1,\ 2\ldots\ n-2 \end{cases} \tag{3}$$

Compared with binary code, gray code is advantageous in the aspect that only one code value is different in the corresponding bits between adjacent numbers. As shown in Tab. 1, two bits in the binary code change their values when 1 becomes 2 in decimal; in contrast, only one bit changes in the Gray code. Therefore, Gray code has a certain self-correcting ability in the process of decoding, and correspondingly it has improved error tolerance during coding and decoding.

2.5 Point Cloud Registration and Fusion

In order to reconstruct a complete 3D model, it is necessary to first obtain point clouds from different angles and then splice multiple overlapping point clouds into a complete 3D model through point cloud registration and fusion [15,16]. However, the point cloud data of the object to be constructed are vulnerable to the influences of camera lens distortion, light intensity, and surface texture. Therefore, it is required to first preprocess the point cloud data before registration (e.g. point cloud smoothing and filtering).

Point cloud registration is essentially the process of obtaining the rotation and translation matrices between the source point cloud and the target point cloud.

$$C_T = R \cdot C_S + t \tag{4}$$

Table 1: Binary and gray code comparison

Decimal	Binary	Gray code
0	000	000
1	001	001
2	010	011
3	011	010
4	100	110
5	101	111
6	110	101
7	111	100

where C_T and C_S are the target point cloud and the source point cloud, respectively, R is the rotation matrix, and t is the translation matrix.

Point cloud registration is generally divided into two categories, namely coarse registration and fine registration. Coarse registration is to find an approximate relationship matrix that can render the target point clouds and source point clouds in the same coordinate system, whereas fine registration is to further optimize the relationship matrix when the initial value is known. In this paper, the SAC-IA algorithm is used for the coarse registration of point clouds [17,18], while the ICP algorithm is used for the precise registration of point clouds [19].

An average fusion method named voxel filtering is used to keep the integrity and smoothness of the point cloud model. Specifically, the object space is divided into many 3D voxel grids with small side lengths, where all points will fall inside. The area with high density has more points in one voxel grid, whereas the area with low density has fewer points in one voxel grid. The side length of the voxel grid can be set to ensure that there is approximately one point inside one voxel grid. The voxel grid is also used for filtering, and the average value of all points in one grid is used as the new point value. This method can not only merge the overlapping areas, but also ensure that the density of the point cloud after fusion tends to be consistent.

3 Experiments

The experimental platform consists of a binocular camera with a resolution of 1280×720 pixels, a projector with a maximum resolution of 1980×1080 pixels, a diffuse reflection calibration plate, a rotating turntable, a black background cloth, and a laptop. A small vase (127.0 *mm* high, 72.0 *mm* wide, and 18.0 *mm* bottle mouth) is selected for 3D reconstruction and measurement (Fig. 4). Experiment steps are as follows:

Step 1: Image calibration. The binocular camera calibration algorithm from the OpenCV library is used, and it requires the calibration images to be taken from different angles. A total of 13 images are taken by the left camera to solve the internal and external parameter matrices of the left camera, and the same operation is carried out for the right camera (Fig. 5).

After calibration, the camera parameter matrices K_l and K_r, distortion coefficients D_l and D_r, rotation matrices R_l and R_r and translation vectors t_l and t_r of the left and right cameras are obtained:

$$\begin{cases} K_l = \begin{bmatrix} 1446.20 & 0 & 692.80 \\ 0 & 1447.26 & 400.42 \\ 0 & 0 & 1 \end{bmatrix} \\ K_r = \begin{bmatrix} 1450.53 & 0 & 661.14 \\ 0 & 1451.76 & 386.93 \\ 0 & 0 & 1 \end{bmatrix} \end{cases} \tag{5}$$

$$\begin{cases} D_l = \begin{bmatrix} 0.0645 & -0.0387 & 0.1786 & 6.2745e-04 & -9.5160e-04 \end{bmatrix}^T \\ D_r = \begin{bmatrix} 0.0414 & 0.2708 & -1.2122 & 3.4274e-04 & -6.6006e-05 \end{bmatrix}^T \end{cases} \tag{6}$$

$$\begin{cases} R_l = \begin{bmatrix} 9.964e-01 & -4.599e-02 & 7.093e-02 \\ 2.222e-02 & 9.520e-01 & 3.052e-01 \\ -8.156e-02 & -3.025e-01 & 9.496e-01 \end{bmatrix} \\ R_r = \begin{bmatrix} 9.971e-01 & -4.205e-02 & 6.326e-02 \\ 2.018e-02 & 9.495e-01 & 3.131e-01 \\ -7.323e-02 & -3.109e-01 & 9.476e-01 \end{bmatrix} \end{cases} \tag{7}$$

$$\begin{cases} t_l = \begin{bmatrix} -7.631e+01 & -7.725e+01 & 4.938e+02 \end{bmatrix}^T \\ t_r = \begin{bmatrix} -1.409e+02 & -7.266e+01 & 4.926e+02 \end{bmatrix}^T \end{cases} \tag{8}$$

Two images of non-coplanar alignment are corrected to conform to coplanar alignment. Fig. 6 shows the images of the chessboard before and after the correction.
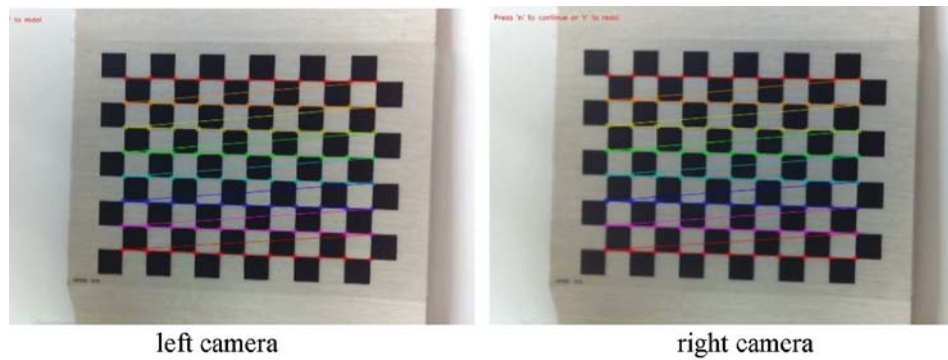

Figure 4: Experiment platform

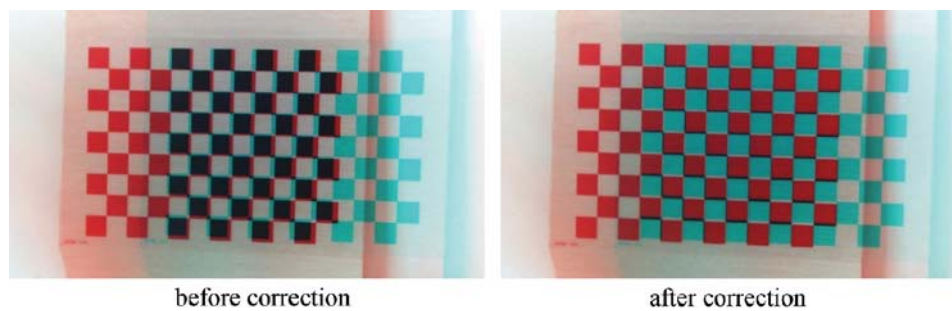**Figure 5:** Calibration process diagram of left and right camera

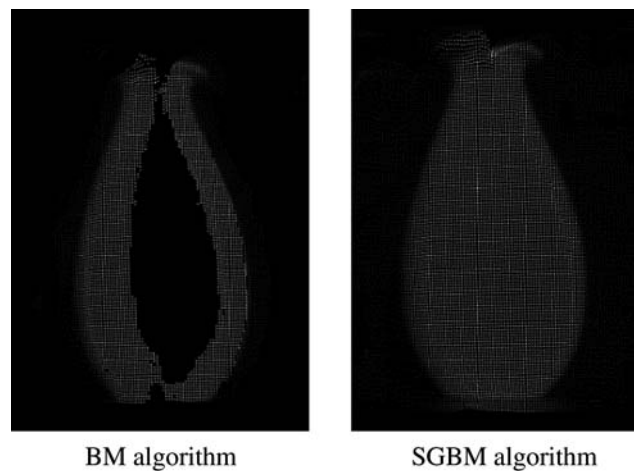**Figure 6:** The images of chessboard before and after correction

Figure 7: Point cloud maps obtained by BM and SGBM algorithms

Step 2: Image matching. The feature areas of the left and right images are first extracted, and then the polar lines are taken in sequence. For each feature point in the left image, the point on the polar line in the right image that conforms to the given matching threshold is identified as the matching point. The above process is repeated to find all the matching points of the binocular image to obtain a sparse disparity map. The experiment proves that the BM algorithm is favored by rapid image processing, but the accuracy is poor. In contrast, the SGBM algorithm takes a

slightly longer time than the BM algorithm, but the accuracy is greatly improved. The disparity map is then transformed into the 3D space, and the point cloud map under one perspective is obtained (Fig. 7).

Step 3: Coding and decoding. A vase is projected by Gray code patterns in vertical and horizontal directions (Fig. 8). It is necessary to first decode the Gray code in two directions separately and then merge the decoded images to obtain a whole image.
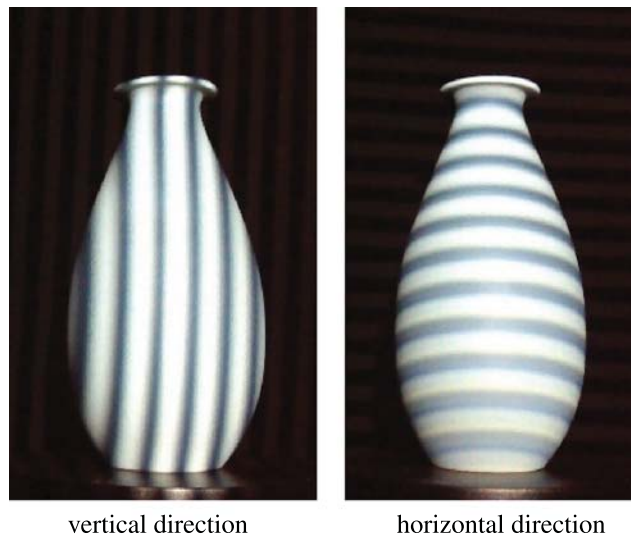

vertical direction   horizontal direction

Figure 8: Gray code pattern projected on the vase

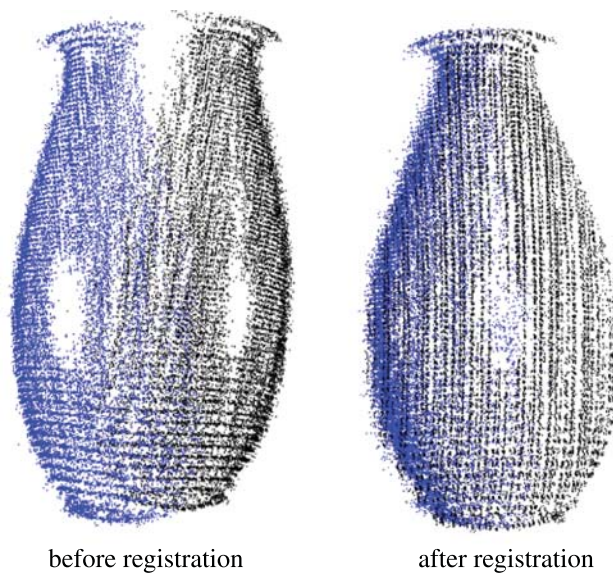before registration   after registration

Figure 9: Point cloud image before and after registration

Step 4: Point cloud registration. The decoding results of the left and right images are used to reconstruct the point clouds. The rotating platform is to get the multi-viewpoint clouds. SAC-IA and ICP algorithms are used to register the point cloud data. Fig. 9 shows the point cloud images before and after registration.

Step 5: Point cloud fusion. Since the distance between the camera and the object is almost constant, the density of the object point cloud is basically the same. Therefore, a fixed side length of the voxel grid, which is 1.0 mm, is suitable for all the point clouds in this experiment, with which the obtained 3D point cloud model has a uniform density (Fig. 10).
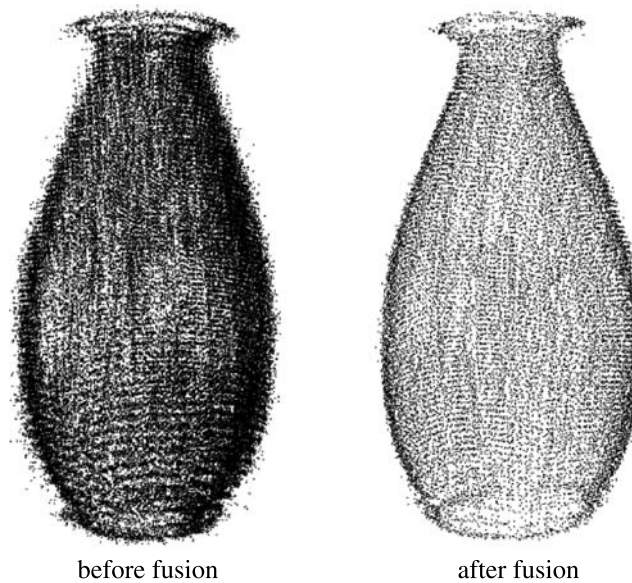

before fusion   after fusion

Figure 10: Comparison of point clouds before and after fusion

4 Results

In order to determine the accuracy of the reconstruction, it is necessary to measure the reconstructed 3D model and compare it with the actual vase. A polarizer is mounted in front of the cameras to filter out the redundant reflected light, which can help improve the reconstruction quality. The measurement results of the 3D model are shown in Fig. 11, and the comparison between the model and the actual vase is shown in Tab. 2. In general, the reconstruction errors for vase height and width are less than 1 mm in all groups no matter whether the polarizer is place. However, the placement of the polarizer improves the reconstruction error of the vase mouth from > 1 mm to <1 mm. The object reconstruction and measurement accuracy can achieve the millimeter level, which should be improved in the next step.
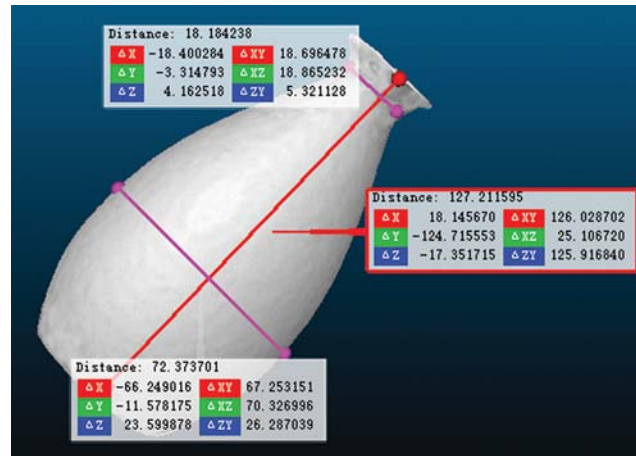
Figure 11: Measurement results of reconstructed vase model

Table 2: Comparison table of measurement results

Measured parameters	Vase height (mm)	Vase width (mm)	Vase mouth diameter (mm)
Actual value	127.00	72.00	18.00
Group 1 without polarizer	127.22	71.27	19.15
Group 2 without polarizer	127.83	72.92	19.24
Group 3 without polarizer	127.68	71.05	18.89
Group 1 with polarizer	127.21	72.37	18.18
Group 2 with polarizer	127.47	72.34	18.26
Group 3 with polarizer	127.34	72.61	18.12

5 Conclusion

In this study, the principle of the binocular stereo vision is first described, followed by the calibration of the binocular camera, the acquisition of internal and external parameter matrices of the cameras, the generation of the disparity map and point cloud map of the vase based on the stereo matching algorithm, and the investigation into the registration and fusion methods for the point cloud. Experiment results show that the reconstructed 3D model of a vase has a satisfactory performance, which can meet the need for rapid measurement of an object.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] S. Gao, "3D reconstruction of small object based on the structured light," M.S. dissertation, University of Electronic Science and Technology of China, 2019.

[2] W. Zhu, X. Li and Y. Shon, "Research on clothing simulation design based on three-dimensional image analysis," *Computers, Materials & Continua*, vol. 65, no. 1, pp. 945–962, 2020.

[3] Z. Zhang, P. Li, S. Zhao, Z. Lv, F. Du *et al.,* "An adaptive vision navigation algorithm in agricultural IoT system for smart agricultural robots," *Computers, Materials & Continua*, vol. 66, no. 1, pp. 1043–1056, 2021.

[4] Y. Zhao, J. Fu, H. Yu, J. Han and D. Zhang, "Defocus projection three-dimensional measurement based on deep learning accurate phase acquisition," *Infrared and Laser Engineering*, vol. 49, no. 7, pp. 161–168, 2020.

[5] Z. Cai, X. Liu, X. Peng and B. Gao, "Ray calibration and phase mapping for structured-light-field 3D reconstruction," *Optics Express*, vol. 26, no. 6, pp. 75–98, 2018.

[6] J. Zhang, K. Yu, Z. Wen, X. Qi and A. K. Paul, "3D reconstruction for motion blurred images using deep learning-based intelligent systems," *Computers, Materials & Continua*, vol. 66, no. 2, pp. 2087–2104, 2021.

[7] W. Sun, X. Zhang, X. He, Y. Jin and X. Zhang, "Two-stage vehicle type recognition method combining the most effective Gabor features," *Computers, Materials & Continua*, vol. 65, no. 3, pp. 2489–2510, 2020.

[8] Z. Xiao, W. Zhang, Y. Zhang and M. Fan, "Fatigue investigations on steel pipeline containing 3d coplanar and non-coplanar cracks," *Computers, Materials & Continua*, vol. 62, no. 1, pp. 267–280, 2020.

[9] S. Zhang, "High-speed 3D shape measurement with structured light methods: A review," *Optics & Lasers in Engineering*, vol. 10, no. 6, pp. 119–131, 2018.

[10] J. Hu, F. Zhang, Z. Li and H. Huang, "Research on indoor three dimensional measurement algorithm based on binocular technology," *Computer Measurement and Control*, vol. 27, no. 9, pp. 66–70, 2019.

[11] H. Zhang, L. An, Q. Zhang, Y. Guo, X. Song *et al.,* "SGBM algorithm and BM algorithm analysis and research," *Geomatics and Spatial Information Technology*, vol. 39, no. 10, pp. 214–216, 2016.

[12] T. Li, F. Hu and Z. Geng, "Structured-light 3D surface imaging technology," *Journal of Network New Media*, vol. 1, no. 1, pp. 22–33, 2012.

[13] J. Geng, "Structured-light 3D surface imaging: A tutorial," *Advances in Optics & Photonics*, vol. 3, no. 2, pp. 128–160, 2011.

[14] K. Herakleous and C. Poullis, "3Dunderworld-sls: An open-source structured-light scanning system for rapid geometry acquisition," 2014. [Online]. Available: https://arxiv.org/abs/1406.6595.

[15] F. Zhao, M. Zhou and J. Wang, "Point cloud registration algorithm based on probability iterative closest point," *Computer & Digital Engineering*, vol. 45, no. 3, pp. 419–422, 2017.

[16] H. Chen, X. Zhang, S. Du and N. Zheng, "A correntropy-based affine iterative closest point algorithm for robust point set registration," *IEEE Journal of Automatica Sinica*, vol. 6, no. 4, pp. 981–991, 2019.

[17] R. B. Rusu, N. Blodow and M. Beetz, "Fast point feature histograms (FPFH) for 3D registration," in *Proc. ICRA*, Kobe, Japan, pp. 1050–4729, 2009.

[18] R. B. Rusu and S. Cousins, "3D is here: Point cloud library (PCL)," in *Proc. ICRA*, Shanghai, China, pp. 427–436, 2011.

[19] H. Zhang, G. Kang, Q. Zhang, W. Zhang and W. Xu, "Laser point cloud coarse registration based on improved SAC-IA algorithm," *Aerospace Control*, vol. 37, no. 5, pp. 67–74, 2019.