Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Outdoor 3D reconstruction method based on multi-line laser and binocular vision ⋆

Xu. Jian [*,**] **Xin. Chen** [*,**] **Wenpeng. He** [*,**] **Xuan. Gong** [*,**]

[*] *School of Automation, China University of Geosciences, Wuhan 430074, China (e-mail: chenxin@cug.edu.cn).*
[**] *Hubei Key Laboratory of Advanced Control and Intelligent Automation for Complex Systems, Wuhan 430074, China*

Abstract: Three-dimensional (3D) reconstruction of substation fittings is of great significance for live working robots. However, the key problem is that active 3D cameras cannot work in outdoor strong light environment, and the passive 3D cameras cannot extract features of weak texture targets. This paper proposes an outdoor 3D reconstruction method based on multi-line laser and binocular vision. To solve the problem that weak texture target has few features, we use multi-line laser to create artificial features. To reduce the interference of natural light on the laser in the images, the frame-difference method is proposed for natural light filtering. Then we use the gray-centroid method to position the multi-line laser accurately. Finally, the binocular vision model is used to complete 3D reconstruction of the target. The experiments show that, compared with traditional 3D reconstruction methods, our 3D reconstruction method can realize 3D reconstruction of outdoor weak texture targets effectively.

Keywords: Live working robot, outdoor 3D reconstruction, binocular vision, frame-difference method, laser positioning.

1. INTRODUCTION

Live working is a kind of operation method for power equipment maintenance without power failure, which plays an important role in smart grid (Wang et al. (2018)). The traditional live working method is manual work, but it poses a huge threat to personnel safety. Therefore, developing live working robot technology is of great significance for industrial production (Jian et al. (2019)). In live working with robots, we focus on the disassembly of the substation fittings, as shown in Fig. 1.

Fig. 1. The substation and fittings

It can be seen from Fig. 1 that the substation fittings are at outdoor high altitude and in strong light environment. What's more, most of the substation fittings are made of metal, whose surfaces are usually smooth and have few textures. In order to get the color and position information of the substation fittings, the visual equipment in the robot needs to complete the three-dimensional (3D) reconstruction of the fittings.

3D reconstruction technology has always been a classic research topic in the field of computer vision, which can get the position and color information of the targets (Zhang et al. (2017), Rezende et al. (2016)). However, the challenge is that traditional 3D cameras are not suitable for 3D reconstruction of substation fittings. For one thing, the structured light camera and TOF (Time of Flight) camera cannot perform well outdoors due to the interference of strong sunlight (Zennaro et al. (2015), Chen et al. (2019)). For another, although the binocular vision can realize outdoor 3D reconstruction by detecting feature points and completing stereo matching, it is not suitable for the objects which have weak textures and repeated areas (Fan et al. (2016)). Therefore, it is necessary to develop new methods for outdoor 3D reconstruction.

In recent years, many outdoor 3D reconstruction methods have been proposed. For example, Chang et al (2018). proposed a novel pyramid stereo matching network (PSM-Net) for 3D reconstruction by using two images from a binocular camera. The PSM-Net can exploit global context information to generate the disparity map and achieve 3D reconstruction. Although PSM-Net makes significant progress compared to conventional reconstruction methods in terms of both accuracy and speed, it's structure is complex and is still difficult to deal with repeated areas, textureless regions, and reflective surfaces. In addition, Cohen (2016) and Nakamura (2017) proposed structure

⋆ This work is supported by the National Natural Science Foundation of China under Grants 61873248, the Hubei Provincial Natural Science Foundation of China under Grant 2017CFA030 and Grant 2015CFA010, the 111 project under Grant B17040,the Science and Technology Project of State Grid Corporation of China under Grant 52153216000R.

from motion (SfM) method for 3D reconstruction, separately. Firstly, they designed the SfM system and got lots of unordered images at different positions of the target. Then, groups of images were used for feature extraction and matching. Finally, other images were added to the SfM system for registration and global optimization. Although the SfM method is very successful for outdoor 3D reconstruction, it takes much time to calculate and is not suitable for small spaces like substations.

Aiming at the outdoor objects with few textures, the reconstruction method with traditional color camera is difficult to achieve good results. Due to the high brightness of the laser, some scholars consider combining laser with the color camera to obtain accurate point cloud of the objects in outdoor environment.

Cui et al. (2017) proposed a method for pavement texture reconstruction based on multi-line laser and binocular vision. In this method, a multi-line laser is used to illuminate the road and a binocular camera is used to position the laser strips and reconstruct the pavement. Since no processing is performed on the laser images, the measured 3D points are few and susceptible to sunlight. Due to the low reconstruction rate, the complex structure or large amount of calculation, the above methods are poor effective for 3D reconstruction of substation fittings. Chen et al. (2019) proposed a 3D reconstruction framework based on multiple sensors for autonomous driving. Firstly, one color camera is used to obtain the color information of target, then one 3D LiDAR (light detection and ranging) is used to get the 3D point clouds around the car. Although the framework can perform a quick 3D reconstruction in the surrounding environment around the car, the 3D reconstruction map is sparse due to the low resolution of the laser radar.

It can be found that the above methods using laser still have poor performance on the 3D reconstruction of outdoor substation fittings. In this case, this paper uses a multi-line laser to create artificial features on the object, then uses the frame-difference method to reduce interference from natural light in the images, as well as uses the gray-centroid method to position laser strips precisely. Finally, we complete the 3D reconstruction of outdoor target with few textures effectively.

The rest of this paper is organized as follows. Section 2 introduces the system components of the proposed method. Section 3 introduces the outdoor 3D reconstruction method based on multi-line laser and binocular vision. Some 3D reconstruction experiments are carried out in the outdoor environment to prove the effectiveness of the proposed method in section 4. Finally, a brief conclusion is provided in section 5.

2. SYSTEM COMPONENTS

As shown in Fig. 2(a), the 3D reconstruction system consists of a binocular camera, a laser emitter, a motor and a computer. In order to balance the speed of 3D reconstruction and the difficulty of stereo matching as much as possible, we choose five-line laser as the laser emitter. To ensure that the laser are captured easily by the binocular camera, the laser emitter is placed directly above the binocular camera.

(a) System equipment (b) System environment

Fig. 2. 3D reconstruction system

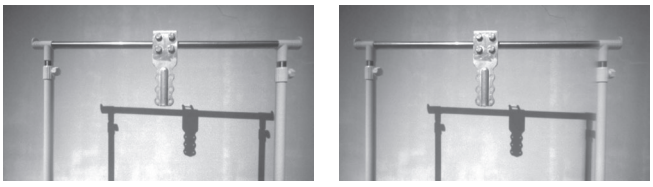
As can be seen from Fig. 2(b), we choose one traditional substation fitting as the experimental target. Because the substation fittings are often around buildings or in high altitude, we set up the experiment background with wall and sun shadow. Since the actual live working robot works near the fitting, we set the distance between the binocular camera in this system and the fitting at about 500 mm. The 3D reconstruction experiment was carried out at 2:00 pm on a sunny day, when the light intensity on the fitting surface was 74230 Lux.

The multi-line laser has a viewing angle of 75°, whose wavelength is 810 nm, and the power is 300 mW. The key parameters of the binocular camera are shown in Table 1.

Table 1. **Binocular camera parameters**

Parameters	Parameter Values
Resolution (pixel)	$1280 * 720$
Baseline	50 mm
Focal length	1.93 mm
Pixel size	2.04 um
Field of view	H 91.2°; V 65.5°; D 100.6°

The binocular camera consists of two infrared CCDs (charge coupled device), so there are only gray information in the images, which are shown in Fig. 3.



(a) The left image (b) The right image

Fig. 3. Two images from the binocular camera

It can be found from Fig. 1 and Fig. 3(a) that the fitting is in a strong light environment, and the fitting has weak texture features, which can be shown as: the surface has few lines or corners. In this case, the active 3D camera cannot work well because the structured light projected by itself is covered by sunlight. In addition, traditional binocular camera can not extract much feature information and thus can not complete 3D reconstruction.

3. 3D RECONSTRUCTION METHOD

3.1 Method overview

This paper proposes a 3D reconstruction method based on multi-line laser and binocular vision, whose procedure is shown in Fig. 4.

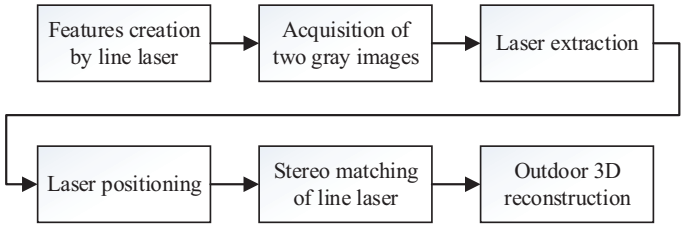


Fig. 4. The procedure of outdoor 3D reconstruction

In the first two steps, the line laser is used to create artificial texture on the surface of the fitting, so that the binocular camera can obtain the gray image with laser features.

The key step of above procedure is laser extraction. In other reconstruction methods, infrared filters are usually used to filter natural light. But the problem is that the same wavelength light is difficult to be removed and can interfere with the laser extraction. However, this paper uses the frame-difference method to complete the filtering of background light and only retains the laser information. Then the gray-centroid method is used to complete the precise positioning of the laser strips. Next, the epipolar constraint is used to accomplish the stereo matching of line laser. Finally, the outdoor 3D reconstruction can be achieve effectively.

3.2 Laser extraction with frame-difference method

Since the fitting is in an outdoor bright light environment, the images from binocular camera contain much background information and the laser is easily interfered by ambient light. In this case, in order to extract the laser effectively, this paper proposes the frame-difference method, whose steps are as follows:

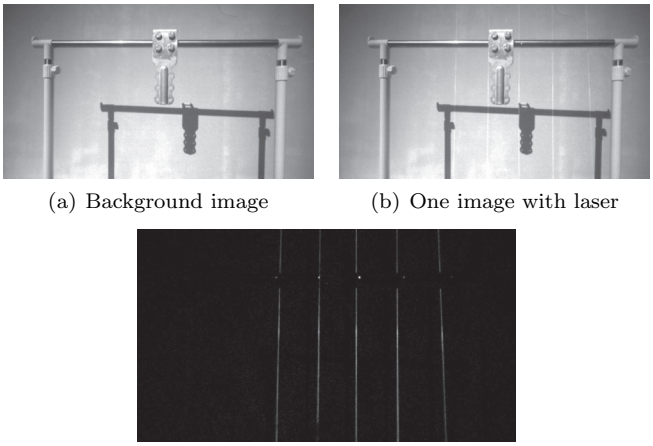
(1) Save one background image M_0 which does not contain any laser;
(2) Use a motor to rotate five-line laser emitter horizontally in front of the fitting and save a new image M_1;
(3) Assume that the gray value of a point (u, v) in M_0 is $V_0(u, v)$, and the gray value of the same position in M_1 is $V_1(u, v)$, then the following processing will be performed:

$$V_1(u, v) = V_1(u, v) - V_0(u, v)$$

(4) With the above subtraction step, we can finally get the laser image M_1 without any background.

In order to verify the feasibility of the frame-difference method, we carried out the laser extraction experiment in outdoor natural light. The laser extraction results are shown in Fig. 5.

As shown in Fig. 5, the proposed frame-difference method can complete the extraction of multi-line laser in an outdoor environment well. Compared to the background image Fig. 5(a), the Fig. 5(b) only has more brightness information of the laser. Thus, the background can be filtered out and the laser strips can be extracted with the frame-difference method, as shown in Fig. 5(c).



(a) Background image (b) One image with laser

(c) Laser extraction results

Fig. 5. The effect of frame-difference method

3.3 Positioning of the laser strip center

The precise positioning of the laser strip center is the basis for stereo matching in binocular vision, and its positioning accuracy will directly affect the accuracy of 3D reconstruction. This section first analyzes the characteristics of the laser stripe, then introduces the positioning of laser center.

When the line-laser emitter generates a laser stripe on the surface of the target, the brightness of the laser stripe is gradually weakened from the stripe center to edge, and usually satisfies gaussian distribution (Hu et al. (2016)). The gray distribution of the laser stripe cross-section is shown in Fig. 6.

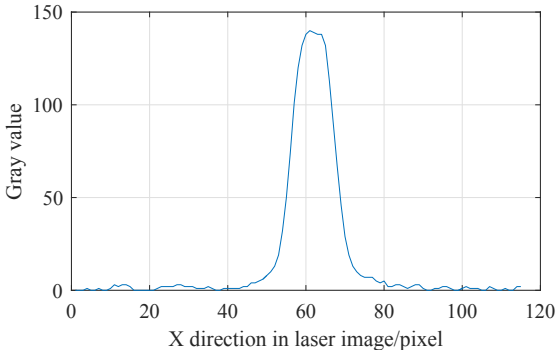


Fig. 6. The gray distribution of the laser stripe

In order to achieve sub-pixel level positioning of the laser stripe, many scholars have proposed their own methods on the positioning of laser stripe center, which can be mainly divided into three types (Li et al. (2013)): the gray threshold method, the gray-centroid method and the direction template method. Compared to other two methods, the gray-centroid method is faster and can reach sub-pixel. Thus it is selected to position the center of the laser stripe.

The principle of the gray-centroid method is similar to the average of the weighting algorithm. Along the cross-section direction of the laser stripe, the stripe coordinate is assumed to be (x_i, y), and the corresponding gray value is $f(x_i, y)$, where $i = 1, ..., n$, n represents the width of laser

stripe. Then the center of laser stripe x_k (k represents the index of laser stripe) is expressed as follows.

$$x_k = \frac{\sum_{i=1}^{n} f(x_i, y) \cdot x_i}{\sum_{i=1}^{n} f(x_i, y)} \quad (k = 1, 2, ..., 5) \tag{1}$$

According to Eq. (1), the center coordinates of the laser stripe are calculated precisely. Then the center of each laser stripe in Fig. 5(c) is positioned, as shown in Fig. 7.

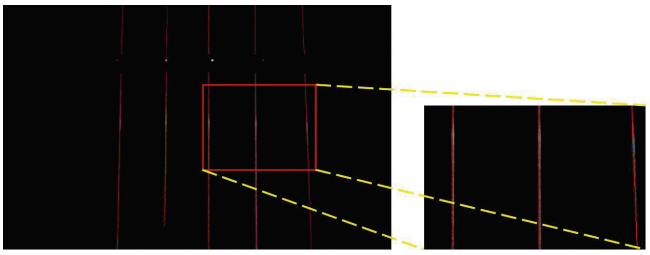


Fig. 7. The center positioning result of laser stripe (best viewed in color)

The experiment shows that the gray-centroid method is good at positioning the center of laser stripe. In the next step, the center coordinates of the laser stripe are used for stereo matching and achieve final 3D reconstruction.

3.4 Stereo matching and 3D reconstruction

After obtaining the sub-pixel coordinates of laser stripe center, we need to achieve stereo matching to realize the final reconstruction. That means, the positions of laser stripes in left and right image need to be matched accurately, which can be shown in Fig. 8.

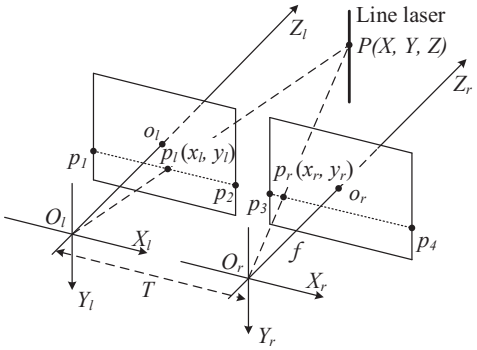


Fig. 8. Binocular camera mathematical model

When the line laser is projected on the surface of the fitting, the binocular camera can observe a 3D point $P(X, Y, Z)$ of laser stripe in space, the P will be imaged at $p_l(x_l, y_l)$ in left image (origin is o_l), and $p_r(x_r, y_r)$ in right image (origin is o_r), separately.

To ensure the accuracy and speed of the 3D reconstruction, the epipolar constraint and ordering constraint are used for stereo matching of line laser. It is assumed that plane PO_lO_r intersects the left image plane to form the pole line p_1p_2, intersects the right image plane to form the pole line p_3p_4. According to the epipolar constraint, it is obvious that the point p_l in left image and the point p_r in right image are in line p_1p_2 and p_3p_4, separately. This constraint can help to reduce the matching area and increase matching speed. And more, according to the ordering constraint, the order of the projection points on the left pole line corresponds to the matching points on the right pole line. With two constraints above, we can complete the stereo matching of the line laser.

Further more, we assume that the camera focus is f, the disparity is d ($d = x_l - x_r$), the baseline is T. Thus, the 3D coordinates of point P on the laser stripe can be obtained with Eq. (2) (Long et al. (2018)).

$$\begin{cases} X = \dfrac{x_l * T}{d} \\ Y = \dfrac{y_l * T}{d} \\ Z = \dfrac{f * T}{d} \end{cases} \tag{2}$$

It can be seen from Eq. (2) that the accuracy of the reconstruction results mainly depends on the calculation accuracy of disparity d. After obtaining groups of matching points and calculating groups of 3D points of the target, we can achieve dense 3D reconstruction of the fitting.

4. SYSTEM TESTING AND RESULTS ANALYSIS

4.1 3D reconstruction experiment

To prove the effectiveness of the proposed method, we perform the 3D reconstruction experiment by using the equipments in section 2. With the scanning of the multi-line laser by motor, we can continuously obtain the disparity information of the scanning area, and finally get the dense disparity map, as shown in Fig. 9.

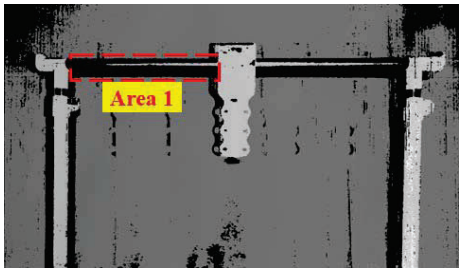


Fig. 9. Disparity map of the fitting

Then we calculate the point cloud of fitting and used the PCL (Point Cloud Library) to display it. The results are shown in Fig. 10.

(a) Front view (b) Top view

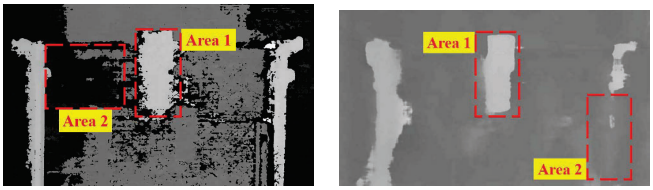
Fig. 10. 3D reconstruction results

As can be seen from Fig. 9 and Fig. 10, the depth information around the fitting can be obtained effectively. The contour of the fitting is very clear, and the most part of the fitting is well reconstructed. The point clouds distributions of the fitting and the background are consistent with the actual experimental environment.

At the same time, we can find that the information in the red rectangle (area 1 in Fig. 9) of the fitting shelf is none. The reason can be explained as follows: this area of fitting shelf reflects light, so the laser stripe is difficult to be detected there.

4.2 Comparison of related methods

To further prove the effectiveness of the proposed method, we use classic SGM (Semi-Global Matching) algorithm (Hirschmuller et al. (2007)) and PSM-Net (Chang et al. (2018))to process the left and right images from binocular camera, and generate the disparity maps of the fitting, which are shown in Fig. 11.



(a) SGM result (b) PSM-Net result

Fig. 11. Disparity maps of other methods

It can be found that the noise of SGM algorithm is more obvious than the proposed method, such as area 1 in Fig. 11(a). Furthermore, the SGM algorithm and PSM-Net are not suitable for dealing with weak textures (such as area 2 in two maps).

By counting the valid points in the disparity map from above methods, we obtain the 3D reconstruction results of the three methods, as shown in Table 2.

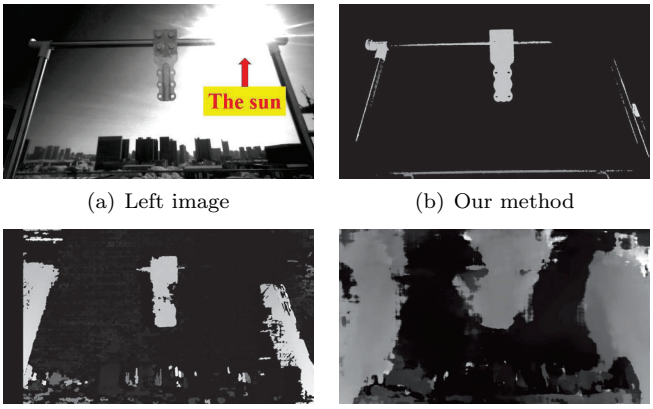
Table 2. **The 3D reconstruction rates**

Method	3D points num	Rate	Evaluation
Our method	743287	80.7%	Edge is clear
SGM	480360	52.1%	A lot of noises
PSM-Net	155182	16.8%	A lot of noises

Since the two CCDs from binocular camera are in different position in 3D space, the common view of the binocular camera is less than 1280∗720. Farther more, due to the reflection of the fitting shelf, some parts of the fitting are not reconstructed. Therefore, the 3D reconstruction rate of the proposed method is about 80%, which is much better than other two methods in the comparison. It is obvious that the method in this paper is good at outdoor 3D reconstruction of weak texture regions.

In order to verify the effectiveness of the proposed method comprehensively, we use the above three methods to carry out 3D reconstruction experiments when the fitting is against to the sun. Then the qualities of these disparity maps are compared, as shown in Fig. 12.

From the results, we can see that the disparity maps obtained from SGM method and PSM-Net method are



(a) Left image (b) Our method

(c) SGM (d) PSM-Net

Fig. 12. The disparity maps when the fitting is against to the sun (light intensity on the fitting is 22480 Lux)

quite messy and sparse. In comparison, the reconstruction result from our method is quite dense and clear. We think it can be explained from the following aspects:

1) The SGM method analyzes the related feature information from left and right images, and calculates the disparity map by searching minimum matching cost. However, the fitting and the background are full of repeated areas, textureless regions, so SGM method has poor performance.
2) Although the spatial pyramid pooling module in PSM-Net can get global context information by aggregating context in different scales and positions, and forms a cost volume, the training images of the network include streets, vehicles and other objects with rich texture, so the network parameters are dependent on the training data set. Therefore, when the test images have few textures and repeated textures, PSM-Net is not robust and has poor performance.

4.3 3D reconstruction accuracy analysis

In this paper, the bolts spacings from the reconstruction result are compared with actual values to analyze the 3D reconstruction accuracy of the method. Firstly, the shape-based template matching method (Jian et al. (2019)) is used to recognize the bolts in the left image. Then the 3D coordinates of four bolts can be calculated from the disparity image. Finally, we get the spacings of four bolts on the substation fitting. The bolt numbers and the recognition result are shown in Fig. 13.

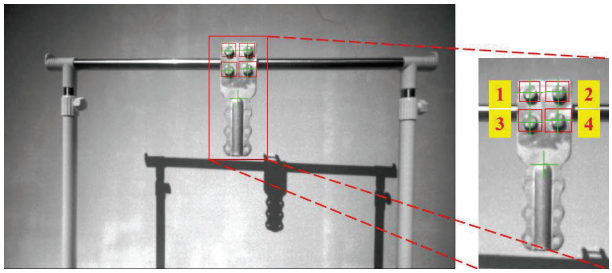


Fig. 13. Bolts recognition results

As can be seen from Fig. 13, all four bolts are well recognized. Then we calculate the 3D coordinates according to

the image coordinates and the corresponding depth values in the left camera coordinate system. The 3D coordinates (unit is mm) of each bolt are extracted multiple times and averaged. The results are shown in Table 3.

Table 3. **The 3D coordinates of bolts**

Bolt number	X data	Y data	Z data
1	-8.30 mm	-184.76 mm	500.0 mm
2	34.00 mm	-182.91 mm	495.0 mm
3	-7.52 mm	-143.65 mm	500.0 mm
4	33.24 mm	-143.75 mm	495.0 mm

It can be seen from Table 3 that the substation fitting is not facing the camera plane. Then we calculate the distance between adjacent bolts according to Table 3 and get the results, as shown in Table 4.

Table 4. **The distance of adjacent bolts**

Group	Real distance	Measured distance	Error	Error rate
1-2	42.3 mm	42.63 mm	0.33 mm	0.78 %
1-3	41.5 mm	41.12 mm	-0.38 mm	-0.92 %
2-4	39.0 mm	39.17 mm	0.17 mm	0.44 %
3-4	40.5 mm	41.06 mm	0.56 mm	1.38 %

As shown in Table 4, the measurement error of the bolt distance is less than 0.6 mm, and the error rate is less than $\pm 1.4\%$, illustrating that the 3D reconstruction of the proposed method is accurate. The reason can be explained as follows: the background light in the two images from the binocular camera is filtered out due to the use of the frame-difference method, so the laser stripe is well extracted. At the same time, the laser stripe is accurately positioned due to the use of the gray-centroid method. Finally, the binocular vision model is used to obtain the 3D coordinates of the bolts precisely.

5. CONCLUSION

To solve the interference of strong light on weak texture targets and achieve the 3D reconstruction of the outdoor substation fitting, a new 3D reconstruction method is proposed in this paper. In our method, the multi-line laser is used to create artificial features on the target, the frame-difference method is used to reduce the interference of natural light on the laser stripes in the images and complete the extraction of laser stripes. The experiments show that the system in the method can complete outdoor 3D reconstruction effectively, and the 3D reconstruction error rate is less than $\pm 1.4\%$. In addition, it is possible for the proposed method to meet the live working requirements of the robot on the maintenance of substation fittings.

REFERENCES

Chang. JR, Chen. YS (2018). Pyramid stereo matching network. *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, 5410–5418.

Chen. C, Xiong. G, Zhu. S (2019). Outdoor 3D environment reconstruction based on multi-sensor fusion for remote control. *Proceedings of the IEEE 8th Joint International Information Technology and Artificial Intelligence Conference*, Chongqing, China, 1753–1757.

Cohen. A, Schonberger. JL, Speciale. P, et al (2016). Indoor-outdoor 3d reconstruction alignment. *Proceedings of the European Conference on Computer Vision*, Springer, 285–300.

Cui. X, Zhou. X, Lou. J, et al (2017). Measurement method of asphalt pavement mean texture depth based on multi-line laser and binocular vision. *Proceedings of International Journal of Pavement Engineering*, Springer, 18(5), 459–471.

Fan. F, Wang. Z, Liu. S, et al (2016). 3D reconstruction of weak texture surface based on binocular vision. *Proceedings of the 3rd International Conference on Mechatronics and Information Technology*. Atlantis Press.

Hirschmuller. H (2007). Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on pattern analysis and machine intelligence*, 30(2), 328–341.

Hu. Y, Ma. Y, Xu. M, et al (2016). Measurement algorithm of mounting holes based on binocular stereo vision. *Proceedings of the IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems*, 489–492.

Jian. X, Chen. X, Xiao. Z, et al (2019). Bolt positioning method based on active binocular vision. *Proceedings of the IEEE Chinese Control Conference*, 7057–7062.

Li. YY, Zhang. ZY (2013). Survey on linear structured light stripe center extraction. *IEEE Laser & Optoelectronics Progress*, 50(10), 100002.

Long. Y, Wang. Y, Zhai. Z, et al (2018). Potato volume measurement based on RGB-D camera. *IFAC-PapersOnLine*, 51(17), 515–520.

Nakamura. K, Mineta. K, Naruse. K. (2016). Investigation of 3D reconstruction from time-series images by towing camera. *IFAC-PapersOnLine*, 50(1), 10317–10322.

Rezende. DJ, Eslami. SMA, Mohamed. S, et al (2016). Unsupervised learning of 3d structure from images. *Advances in Neural Information Processing Systems*, 4996–5004.

Wang. T, Chen. X, Tan. C, et al (2018). Localization of Substation Fittings Based on a Stereo Vision Method. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 22(6), 861–868.

Zennaro. S, Munaro. M, Milani S, et al (2015). Performance evaluation of the 1st and 2nd generation Kinect for multimedia applications. *Proceedings of the IEEE International Conference on Multimedia and Expo*, 1–6.

Zhang. HL, Li. DR, and Zhan. S (2017). Dense 3D reconstruction with an active binocular panoramic vision system. *Proceedings of the IEEE International Conference on Intelligent Robotics and Applications*, 832–837.