



A low-cost pipeline surface 3D detection method used on robots

Tianxiang Lan¹ · Guotian Yang¹

Received: 20 November 2023 / Revised: 6 January 2024 / Accepted: 27 January 2024 / Published online: 24 February 2024
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2024

Abstract

This paper presents a low-cost pipeline surface 3D detection method based on line structured light. The method adds 3D detection to the robot at a meager cost and change. The optical flow method is used to derive the motion information of the robot to replace the motion closed loop. Finally, the depth data for the entire surface are generated automatically only from the camera and the line laser projector, without using other devices. In addition, a novel spot centroid extraction algorithm based on the color region of interest and an adaptive threshold is presented. This method can accurately detect the laser centroid in the pipeline surface. We conducted experiments with a quadruped robot and validated algorithms. The experimental results show that the proposed method achieves 3D detection on a trackless robot at a meager cost and is superior to standard depth sensors.

Keywords 3D detection · Line structured light · Optical flow · Laser · Robot

1 Introduction

As pipelines are spread across deep underground tunnels, ensuring the integrity of the pipeline surface is difficult owing to the high temperature and humid underground environment. Regular detection of pipeline surface defect profiles is required to ensure pipeline safety [1–3]. Previously, pipeline detection typically required human resources. Manual detection has security breaches and significant subjectivity, leading to missed and false detections [4], and the safety of the workers cannot be guaranteed due to the high-temperature environment. As robotic technology and optical sensors develop, pipeline surface detection can employ robots installed with optical sensors [5].

Estimating depth from the pipeline surface image using optical sensors has been essential to solving the problem of identifying and locating underground pipeline defects. Therefore, drawing a dense and accurate depth map of the pipeline surface is the first step. The existing technologies for depth maps of object surfaces are divided into the following two main parts: binocular stereo vision and structured light

[6]. Binocular stereo vision has several advantages: simplicity, low cost, and fast detection speed. It is commonly used in autonomous driving, object recognition, and scene depth mapping [7]. Mansour et al. [8] compared the performance of binocular disparity and motion parallax for depth estimation of stationary objects. In short distances, the binocular technique shows greater accuracy. Priya et al. [9] used binocular stereo-vision technology to identify obscured objects. To resolve the self-obscuring problem, an improved geometric mapping technique is proposed for 3D object recognition by using camera self-calibration techniques. In these studies, binocular vision was severely affected by ambient lighting. A change in light can lead to a matching failure, and the image-matching accuracy is lower when the texture is less on the object's surface. Underground tunnels often have insufficient light, and the surface texture of pipelines rarely varies. The surface images cannot be clearly photographed.

In contrast, structured light technology can effectively address these limitations. Structured light technology is an active detection method that uses an artificial light source to project a beam onto the surface of an object. Spatial triangulation measures the depth of the object's surface defect contour. Structured light technology can be applied in different ways, such as dot structured light, coded structured light, and line structured light [10, 11]. The accuracy of point structured light depends on the density of the point cloud data; the denser the point cloud, the higher the precision. The most common

✉ Tianxiang Lan
txlan222@163.com

Guotian Yang
gty@ncepu.edu.cn

¹ School of Control and Engineering, North China Electric Power University, Beijing 102206, China

use of point structured light is Face ID on iPhone. Coded structured light is significantly disturbed by ambient light and typically requires a specific image processing method. Using a unique image processing method may increase the computational complexity and make the detection slower.

Compared with the aforementioned types of structured light, line structured light has strong anti-jamming, improved precision, and fast scanning speed. In addition, it is suitable for the real-time generation of depth maps of object surface contours [12]. For example, Liang et al. [13] built a 3D pavement inspection system based on multigroup line structured light. The three-dimensional data filtering method was used to recover the pavement contour points and eliminate noise. The 3D reconstruction of the pavement contours was achieved using the Delaunay triangulation algorithm. In [14], a 3D shape measurement method was proposed for large objects using line structured light. This method uses a binocular camera to simplify the light plane calibration process. Based on the standard ball-hand-eye calibration, it combines a modified space circle fitting method to achieve line structured light fusion of arbitrary light planes, which is more accurate than the traditional measurement method. Shang et al. [15] proposed and built a gear tooth profile detection system based on incoherent line structured light. The system uses line structured light, which is generated by spherical light from an incoherent light source via a gap and then irradiates the gear surface to form an optical stripe image. This method has high accuracy and is less affected by speckle noise.

The types mentioned above studies used different methods to improve line structured light detection accuracy and achieved significant results. However, the line structured light sensors must be mounted on the device with tracks. The lateral movement information can be obtained to form a depth plane across the object's surface. The confined environment of underground tunnels makes sensor movement tracks impractical. Figure 1 shows the tunnel that needs to be detected in this paper. Depth information can also be captured using binocular cameras and dot matrix sensors in scenes without track. However, both sensors are costly and are sensitive to ambient light and noise. To overcome these limitations, this study presents a novel low-cost detection method using line structured light to detect surface contours. This method eliminates the need for a track or dedicated camera; just add a line laser projector (less than \$15) near an ordinary camera on the robot.

Recognizing spot distortion of line laser can only obtain 1D depth information. Based on the robot's unidirectional motion characteristics, we propose a method that uses optical flow instead of a motion closed loop to scan a surface. The 1D depth information of the laser and the lateral motion information of the image are combined to generate the surface contour information. We also propose an adaptive threshold



Fig. 1 Underground Tunnel with Pipelines

spot centroid extraction algorithm and a de-jitter algorithm to map depth images accurately. This method makes minimal changes to the robot at a meager cost and provides a highly accurate depth image of the pipeline surface.

The remainder of this paper is structured as follows: Sect. 2 describes the line structured light scheme for measuring the longitudinal depth of an object surface, Sect. 3 presents the details of the image-shift measurement based on optical flow, and Sect. 4 describes the de-jittering and deep map generation processes. The experimental results are presented and discussed in Sect. 5. Finally, the conclusions are shown in the last section.

2 Depth detection based on the line structured light method

2.1 Scheme of the Line Structured Light Detection

A line structured depth sensor is a device that uses a camera to recognize spot distortion to determine the 3D structure of an object. The detection mode is divided into 1D and 2D structured light. 2D structured light produces 3D information in one shot, commonly called dot structured light and line array structured light. 1D structured light is line laser projection. As shown in Fig. 2, a relative motion along the direction of the normal line of light is still required to construct the depth information of the object's surface.

The laser normal line is the x -axis, whereas the parallel line of the laser is the y -axis, and the depth direction is the z -axis, as shown in Fig. 2. If a certain distance exists between

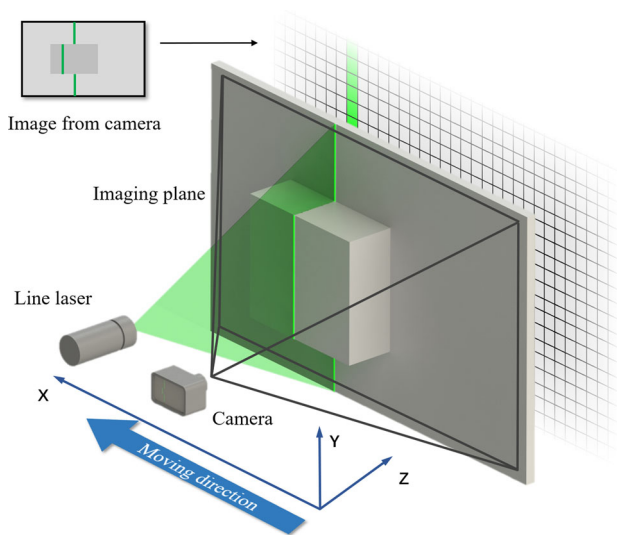


Fig. 2 Principle of 1D Line Structured Light

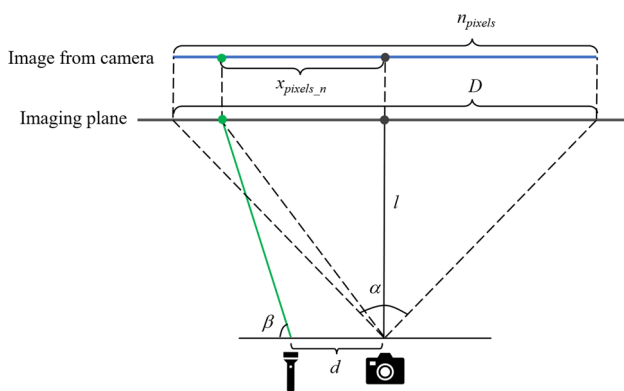


Fig. 3 Quantitative Relationship Between Deep and Laser Line Distortion

the camera and the laser sensors or the laser sensors are at a certain angle to the plane, the distortion of the laser line on the object will be reflected in the imaging plane of the camera. From the vertical view of the structure shown in Fig. 2, the relationship between the laser line deformation and distance to the object can be quantified.

As shown in Fig. 3, the imaging plane has the same proportional relationship as the actual plane.

$$\frac{x_{pixel_n}}{n_{pixels}} = \frac{d \tan \beta + l}{D \tan \beta} \tag{1}$$

where:

$$D = 2l \tan \frac{\alpha}{2} \tag{2}$$

where n_{pixels} are the number of pixels in each row of the image captured by the complementary metal oxide semiconductor (CMOS), x_{pixel_n} is the number of pixels in the n row from

the laser point to the midpoint, d is the distance between the CMOS and the laser, D and l are the length and distance of the plane to be taken, respectively, α is the field of view (FOV) of the lens, and β is the laser offset angle.

Letting

$$C_1 = \frac{n_{pixels}d}{2 \tan \frac{\alpha}{2}} \tag{3}$$

$$C_2 = \frac{n_{pixels}}{2 \tan \frac{\alpha}{2} \tan \beta} \tag{4}$$

Subsequently, according to formula (1) and formula (2), the inverse relations between l and x_{pixel_n} is obtained as follows:

$$l = \frac{C_1}{x_{pixels_n} + C_2} \tag{5}$$

where C_1 determines the proportion between x_{pixel_n} and l , which is the detection accuracy of the line structured light sensor. n_{pixels} depend on the resolution of the CMOS, and α depends on the focal length of the lens. If the camera and lens type is fixed, it is necessary to increase d in the experiment. C_2 can change the offset between x_{pixel_n} and l without affecting the detection accuracy. Therefore, the mounting angle of the laser β is negligible. The depth detection range of the line structured light sensor at $\beta = 90$ is:

$$d \tan \left(90^\circ - \frac{\alpha}{2} \right) < l < \infty \tag{6}$$

2.2 Spot centroid extraction

The shape of the object’s surface can be determined from the positions of the laser lines in the captured image. Because the laser line is directed vertically, each line of the pixels in the image should contain a laser spot without regard to blind spots. The laser line maintained a certain width when it hit the object’s surface. The spot centroid of the laser is the best symbol for the laser position [12]. To achieve color-based fault detection, it is necessary to use RGB cameras rather than monochrome cameras or filters. In addition, for providing the basic data of the optical flow method below, ambient light is also essential.

Color and brightness are essential factors in assessing the laser light position because the laser is a high-intensity light with monochromaticity. Generally, the image’s color is segmented in the HSV color space, where H is the hue, S is the saturation, and V is the value. To filter the color features of the image more intuitively, it is necessary to convert each pixel’s RGB (red, green, blue) values to HSV (hue, saturation, value) values. The commonly used expression for converting RGB

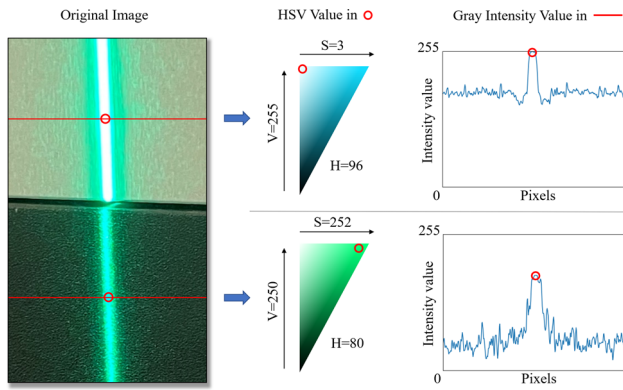


Fig. 4 HSV Value and Gray Intensity Value Using Laser Illuminate on Different Backgrounds

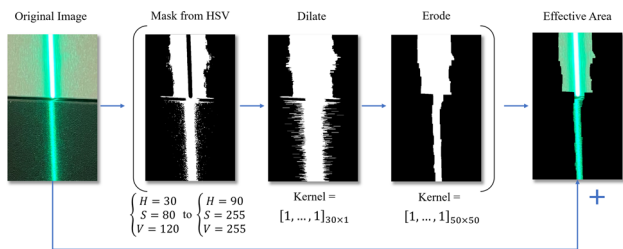


Fig. 5 Process of Color Region of Interest Extraction

to grayscale in OpenCV is as follows:

$$\text{Gray} = R \times 0.299 + G \times 0.587 + B \times 0.114 \quad (7)$$

where the green component has maximum weight; therefore, a green laser was chosen for this study. The color of the object also influences the determination of the laser spot. Figure 4 shows the results of the green laser illuminated on white and black backgrounds.

The HSV values of the spot centroid were 80, 252, and 250 for the green laser light on the black background, which is easily distinguished as the green light. Owing to the saturation limitations of the CMOS camera, the high-brightness green light is reflected as white when the laser is illuminated on a white background. As shown in Fig. 4, the HSV values of the spot were removed from the range of green light, where $H = 96, S = 3$. The luminance value of the spot centroid illuminated on the black background was lower than that on the white background in the grayscale image. It is difficult to obtain the spot centroid of the laser line by relying on a fixed color and luminance thresholds. Therefore, a novel spot centroid extraction algorithm is proposed based on a color region of interest and an adaptive background luminance threshold.

The colored region of the interest extraction process is shown in Fig. 5. First, the green area of the image is extracted in the HSV color space, and the image mask was formed with only one and zero values. Considering the influence

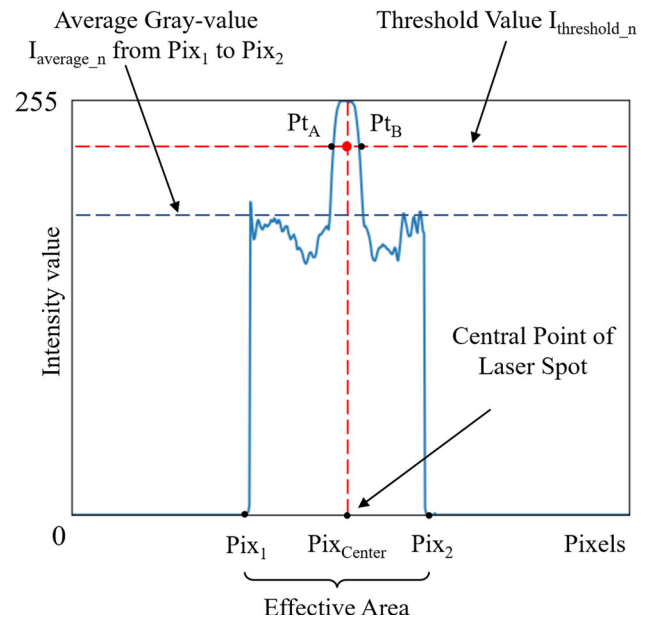


Fig. 6 Spot Centroid Extraction Process

of ambient light and background color, the range of green HSV values is increased. In Fig. 5, the edges of the laser line in the white background are identified as green in the image mask. The central area of the line is defined as white owing to saturation limitations. Second, a 30×1 lateral convolution kernel is used to dilate the image mask to link the lateral masked regions. Third, the mask was corrupted using a 50×50 square convolution kernel to reduce the extent of the green area. Finally, the original image is overlaid with an image mask to obtain the green effective area map.

Extracting the green effective area substantially reduced the misidentification of the other colored light spots. The average gray value of row n in the valid region is calculated as follows:

$$I_{average_n} = \frac{\sum_{x=1}^{pix} I(n, x)}{P_n} \quad (8)$$

where pix is the total number of horizontal pixels in the original image, P_n is the total number of nonzero pixels in row n , and $I(n, x)$ is the gray value of pixel x in row n .

Using the value $I_{average_n}$, the threshold $I_{threshold_n}$ in row n is calculated as follows:

$$I_{threshold_n} = K(I_{max_n} - I_{average_n}) + I_{average_n} \quad (9)$$

where I_{max_n} is the maximum gray value of all pixels in the row n and K is the threshold correction factor. The process of spot centroid extraction is shown in Fig. 6.

The gray value curve of row n intersects with $I_{threshold_n}$ at points Pt_A and Pt_B . The pixel positions at the center of

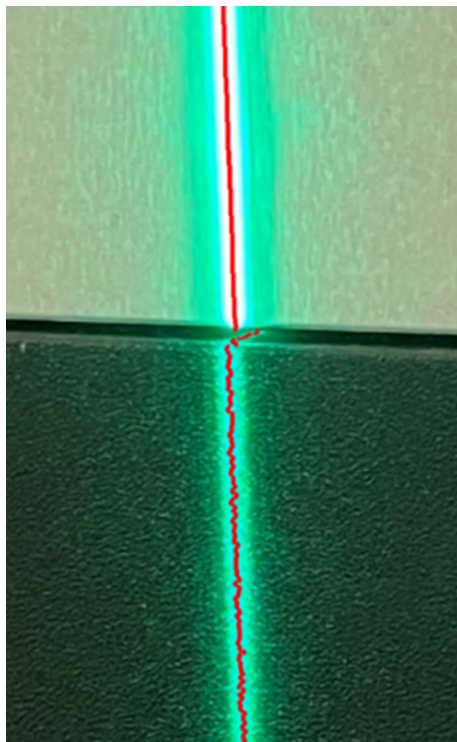


Fig. 7 Results of the Spot Centroid Extraction Algorithm

the Pt_A and Pt_B are the coordinates of the laser spot centroid. The threshold correction factor K is between 0.5 and 1. The smaller values of K capture the more accurate spot centroids, but in lighter reflective materials, other spots are easily caught. The larger values of K can filter out irrelevant spots, but it may cause a spot centroid offset in dark or rough materials. Using extensive testing, the selected value of $K = 0.9$ is suitable for the pipeline surface.

Figure 7 shows the results of the spot centroid extraction algorithm implemented using OpenCV v3.8 and Python v4.2.

The algorithm proposed in this study effectively avoids the background color and ambient illumination effects on the recognition of spot centroids. However, the spot centroid constitutes only 1D depth information. Hence, it was necessary to use the relative motion of the camera in the photographic plane to obtain a complete 2D depth image. This study is oriented toward an application environment without track. Therefore, visible image processing methods can indirectly obtain feedback from the camera motion and plot depth images.

3 Image motion measurement

The essential condition for obtaining plane depth information is to capture the camera's motion along the x -axis. We used the optical flow method to track the feature points, calculate

the moving pixel values in the image, and thus determine the relative motion between the camera and the image.

First, Harris feature corner points were captured for the camera frame [16]. The basic principle is to set a sliding window in the image and calculate the average gray value change within the window. The gray value in the window has a more uniform distribution when there is no change in the gray value. Edge features are contained in the window when the gray value changes in one direction. This window contains corner features when the grayscale value changes in both directions. The change in the gray value was calculated as follows:

$$E(u, v) = \sum_{x, y} w(x, y)[I(x + u, y + v) - I(x, y)] \quad (10)$$

where $E(u, v)$ is the change in gray value along the (u, v) direction, $w(x, y)$ is a window function generally chosen as a 2D Gaussian function, and $I(x, y)$ is the intensity of a grayscale image at location (x, y) . Setting different values for u and v can determine the direction of the movement of the window. In this study, 100 feature points were selected in the area beyond the planned laser line, and the pixels of these feature points were tracked.

The optical flow is a method for calculating the pixel motion vector in an image [17]. The pixel $I(x, y, t)$ is the light intensity of an image at time t . In the next frame, dt is the duration taken to move distance (dx, dy) .

Assuming that the light intensity of the pixel is constant before and after the movement:

$$I(x, y, t) = I(x + dx, y + dy, t + dt) \quad (11)$$

A Taylor expansion in the right-hand side of the formula (11) gives:

$$I(x, y, t) = I(x, y, t) + \frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} dt + \varepsilon \quad (12)$$

where ε is the higher order that is negligible; hence,

$$\frac{\partial I}{\partial x} dx + \frac{\partial I}{\partial y} dy + \frac{\partial I}{\partial t} dt = 0 \quad (13)$$

In the discrete case, ∂I represents the change in the gray level of adjacent pixels in a given direction. The pixel displacements (x, y) along the x - and y -axis are dx and dy , respectively, and the time taken to reach the new pixel position $I(x + dx, y + dy)$ is dt . The pixels in the adjacent frame can be used to calculate dx using formula (13) when dt is a single frame.

The optical flow method was performed on 100 feature points. Not all feature points can be tracked accurately; for example, the similarity and the loss of feature points can

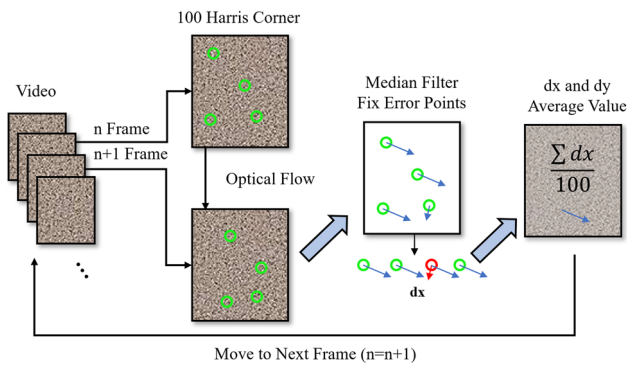


Fig. 8 Process of the Image Motion Measurement Based on Optical Flow Method

cause errors. The dx of 100 feature points are randomly arranged and use median filtered, which a window size is 3, to correct the error points. The mean value is then calculated for 100 dx to obtain the overall dx of the adjacent frame. This process is illustrated in Fig. 8.

4 The image shake and generation depth image

The jitter will produce the motion displacement along the y and z axis (as shown in Fig. 2) when the line structured light sensor is mounted on the detection robot. This becomes more obvious when the ground is uneven or using a quadruped robot. The depth data are distorted owing to the jitter along the y -axis. The jitter along the z -axis causes periodic changes in the depth data.

The longitudinal movement of image dy can be calculated using the method described in Sect. 3, which is similar to dx . The distance dz between the camera and the target plane is reflected in the scaling of the frame and can be confirmed by the relative position of the critical feature points between adjacent frames. The affine matrix between the n frame and the $n + 1$ frame is:

$$M_{affine_n} = \begin{bmatrix} S_{img_n} & 0 & T_{x_n} \\ 0 & S_{img_n} & T_{y_n} \\ 0 & 0 & 1 \end{bmatrix} \tag{14}$$

where T_{x_n} and T_{y_n} are the translations of the screen along the x and y axis, respectively. S_{img_n} is the screen scaling factor. The scaling of the screen S_{img_n} is equal to the scale of the camera distance change S_{dis_n} , as shown in Fig. 3. Figure 9 shows the actual depth data without jitter, calculated using the depth data L of each current frame divided by S_{dis_n} of the frame.

The depth image is presented using the grayscale image construction method when obtaining the camera motion in

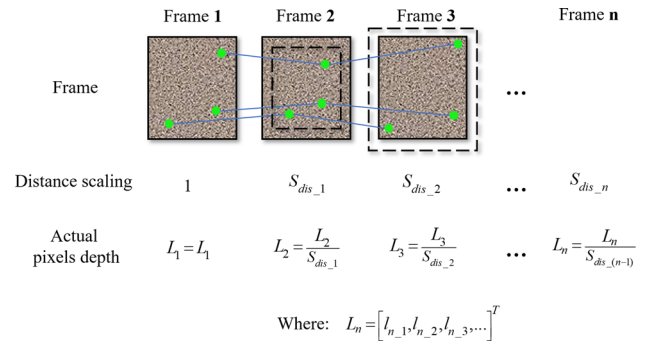


Fig. 9 Calculation Process of Image Depth Data De-jittering

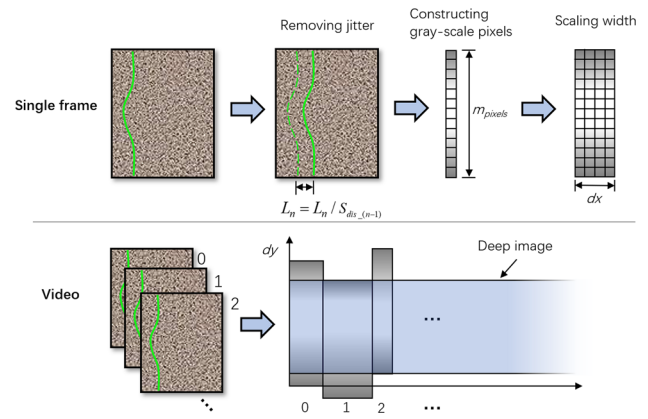


Fig. 10 Depth Image Generation Process

three directions. First, jitter compensation is applied to the full-depth data in each frame. The minimum to maximum distance is then mapped to grayscale values ranging from 0 to 65,535. A grayscale matrix of width dx and height m_{pixels} is constructed. Finally, the gray matrix of all frames in the video is stitched horizontally, and the effects of dy are eliminated vertically. The process of depth-image generation is shown in Fig. 10.

5 Analysis and experiments

5.1 Theoretical accuracy analysis

The detection accuracy of the depth sensors can be defined in terms of the following three aspects: vertical resolution, lateral resolution, and depth resolution. In this study, the vertical resolution of the line structured light sensor depended on the pixel number of the CMOS and the actual size of the frame in the field of view. The relationship between the length D of the subject plane, the object distance I , and field of view angle α in 1D is displayed in Fig. 3. In practice, the distribution of the field of view and lens field of view in a common arrangement such as 1920×1080 pixel CMOS is

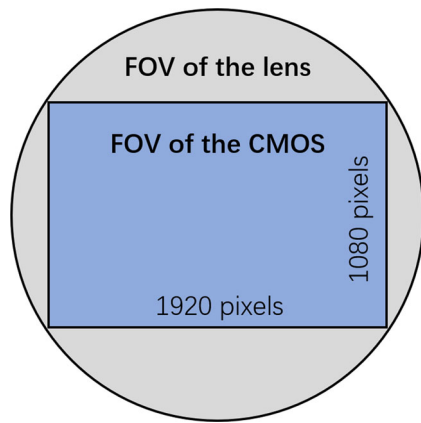


Fig. 11 Distribution of 1920 × 1080 pixel CMOS in the Field of View and Lens Field of View

shown in Fig. 11. By combining formula (2), the actual vertical resolution of the screen in the field of view is calculated as:

$$D_{vertical} = \frac{9}{\sqrt{337}} l \tan \frac{\alpha}{2} \tag{15}$$

The theoretical resolution in the vertical direction is then calculated as follows:

$$R_{vertical} = \frac{D_{vertical}}{1080} \tag{16}$$

where the unit is mm/pixel.

The lateral detection accuracy is mainly influenced by the inter-frame distance of the video. Assuming that the camera moves at a speed of v , the unit is mm/s. The video frame rate is f , and the unit is fps. Therefore, the lateral detection accuracy is defined as follows:

$$R_{lateral} = \frac{v}{f} \tag{17}$$

The depth detection accuracy varies with the distance of the camera from the subject plane, which is the position of the laser spot in the frame. The depth detection accuracy is obtained as follows:

$$R_{deep} = \frac{C_1}{x_{pixels_n} + C_2} - \frac{C_1}{x_{pixels_n} + 1 + C_2} \tag{18}$$

5.2 Experimental results

The line structured light sensor system used in this study is illustrated in Fig. 12. The camera used was a RICOH GR3 [18]. The parameters used in these experiments are listed in Table 1.



Fig. 12 Line Structured Light Sensor Detection System

Table 1 Parameters used in the experiments

Parameter	Value
Camera CMOS vertical resolution	1080pixels
Camera CMOS lateral resolution	1920pixels
Camera CMOS diagonal size	APS-C (29.3mm)
Lens focus	18 mm
Equivalent angle of FOV	84°
Frame rate of camera	50fps
Single-frame exposure	1/640 s
Center distance between laser and CMOS	135 mm

To compare the effectiveness of the line structured light sensor mentioned in this study with other depth sensors, we conducted a comparison test using a RealSense D455 [19] released by Intel in 2020. The RealSense D455 contains a binocular camera and lattice sensor, utilizing advanced algorithms to combine the benefits of both. In the experiment, a line structured light sensor was mounted on the Unitree quadruped walking detection robot. The experimental environment is shown in Fig. 13.

The tested objects were concrete walls and fiberglass pipes, as shown in Fig. 14. The maximum difference between the bulge and the depth of the dent was 38 mm for the concrete walls and 12 mm for the fiberglass pipes.

The robot moved at a speed of 0.2 m/s during the concrete wall test. The average distance between the camera and the subject was 850 mm. With the parameters listed in Table 1, the depth detection resolution of the line structured light sensor is approximately 3.2 mm. The depth detection accuracy of the RealSense D455 was 2% of the distance, indicating a resolution of approximately 17 mm. The process of the detection algorithm is shown in Fig. 15, and the results of the detection experiments are shown in Fig. 16. Under the experimental conditions, the RealSense D455 could only

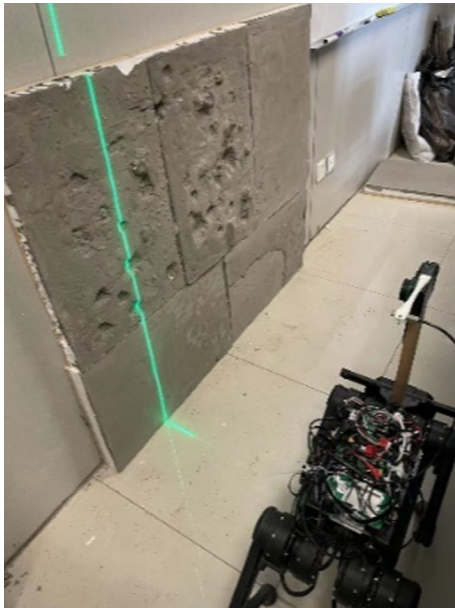
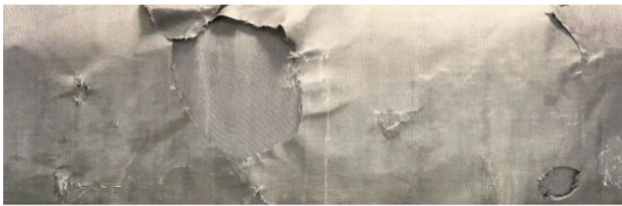


Fig. 13 Virtual Experimental Environment



(a) Concrete Walls



(b) Fiberglass Pipes

Fig. 14 Experimental Test Subjects

detect bulges and depressions and could not detect depths and shapes. The RealSense D455 is more sensitive to noise. For example, more white spots appear in Fig. 16b, whereas the proposed line structured light sensor is virtually immune to noise.

The robot moved at a speed of 0.2 m/s in the fiberglass pipe tests. The average distance between the camera and the subject was 450 mm. The depth detection resolution of the line structured light sensor was approximately 0.7 mm. The experimental results of the line structured light sensor with and without the jitter compensation algorithm are shown in Figs. 17 and 18, respectively. In Fig. 17, the distance between

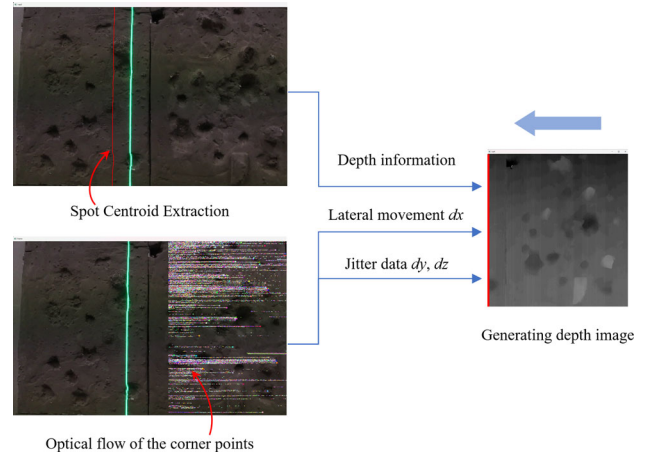
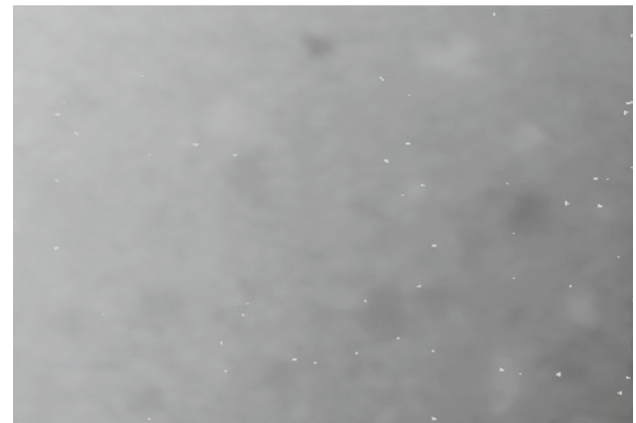


Fig. 15 Process of the detection algorithm



(a) Results from sensors used in this study



(b) Results of the RealSense D455.

Fig. 16 Measurement Results of Two Depth Detection Sensors

the camera and the subject surface changes frequently due to the robot’s movement jitter. A large number of longitudinal stripes appeared in the depth image. With the addition of the jitter-compensation algorithm, a clear depth image was obtained, as shown in Fig. 18.



Fig. 17 Experimental Results without Jitter Compensation



Fig. 18 Experimental Results with the Jitter Compensation Algorithm

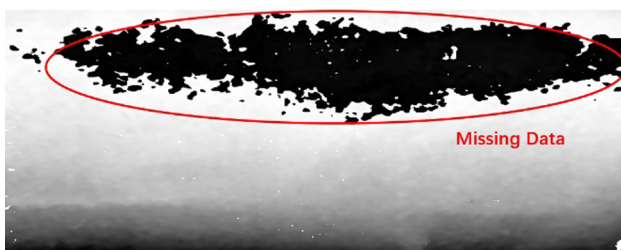


Fig. 19 Depth Detection Result of the RealSense D455

The depth detection results of the RealSense D455 are shown in Fig. 19. The RealSense D455 could not capture point cloud data in highly reflective areas owing to the illumination at the top of the pipe. The binocular camera could not capture contour details, resulting in significant data loss. The line structured light sensor using the spot centroid extraction algorithm proposed in this study can completely exclude interference from the external light.

The surface contour detection method based on the line structured light proposed in this study is superior to the RealSense D455 depth sensor in terms of image resolution and depth resolution. The spot centroid extraction algorithm adapts to the background brightness threshold and removes the effects of the external illumination and the object color. The jitter compensation algorithm compensates for the depth data fluctuations caused by trackless motion.

6 Conclusion

This study presents a low-cost method for pipeline surface contour detection used in quadrupedal robots. In the detection robot with a visible light camera, only one line laser projector with less than \$15 is added; the robot has no additional

modifications or cost. In the process of detection, not only the visible image is retained, but also the high-precision depth image is generated. The experiments showed that the proposed method could detect surface contours more accurately than the commonly used RealSense D455 depth sensor.

Author's contribution T.L. did most of the work and wrote the main manuscript; G.Y. set up the experimental environment and modified the manuscript. All authors reviewed the manuscript.

Funding The authors did not receive support from any organization for the submitted work.

Data availability All data are original for this study, and none has been published elsewhere.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Informed consent No conflict of interest exists in the submission of this manuscript, and all authors approve the manuscript for publication.

References

1. Wang, T., Tan, L., Xie, S., et al.: Development and applications of common utility tunnels in China. *Tunn. Undergr. Space Technol.* **76**, 92–106 (2018)
2. Zhou, R., Fang, W., Wu, J.: A risk assessment model of a sewer pipeline in an underground utility tunnel based on a Bayesian network. *Tunn. Undergr. Space Technol.* **103**, 103473 (2020)
3. Shang, Z., Shen, Z.: Single-pass inline pipeline 3D reconstruction using depth camera array. *Autom. Constr.* **138**, 104231 (2022)
4. Li, Y., Wang, H., Dang, L.M., et al.: A robust instance segmentation framework for underground sewer defect detection. *Measurement* **190**, 110727 (2022)
5. Halfawy, M.R., Hengmeechai, J.: Efficient algorithm for crack detection in sewer images from closed-circuit television inspections. *J. Infrastruct. Syst.* **20**(2), 04013014 (2014)
6. Laga, H., Jospin, L.V., Boussaid, F., et al.: A survey on deep learning techniques for stereo-based depth estimation. *IEEE Trans. Pattern Anal. Mach. Intell.* (2020). <https://doi.org/10.1109/TPAMI.2020.3032602>
7. Poggi, M., Tosi, F., Batsos, K., et al.: On the synergies between machine learning and binocular stereo for depth estimation from images: a survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **44**(9), 5314–5334 (2021)
8. Mansour, M., Davidson, P., Stepanov, O., et al.: Relative importance of binocular disparity and motion parallax for depth estimation: a computer vision approach. *Remote Sens.* **11**(17), 1990 (2019)
9. Priya, L., Anand, S.: Object recognition and 3D reconstruction of occluded objects using binocular stereo. *Cluster Comput.* **21**(1), 29–38 (2018)
10. Van der Jeught, S., Dirckx, J.J.J.: Real-time structured light profilometry: a review. *Opt. Lasers Eng.* **87**, 18–31 (2016)
11. Yang, L., Liu, Y., Peng, J.: Advances techniques of the structured light sensing in intelligent welding robots: a review. *The Int. J. Adv. Manuf. Technol.* **110**(3), 1027–1046 (2020)

12. Xu, X., Fei, Z., Yang, J., et al.: Line structured light calibration method and centerline extraction: a review. *Res. Phys.* **19**, 103637 (2020)
13. Liang, J., Gu, X.: Development and application of a non-destructive pavement testing system based on line structured light three-dimensional measurement. *Constr. Build. Mater.* **260**, 119919 (2020)
14. Yu, H., Huang, Y., Zheng, D., et al.: Three-dimensional shape measurement technique for large-scale objects based on line structured light combined with industrial robot. *Optik* **202**, 163656 (2020)
15. Shang, Z., Wang, J., Zhao, L., et al.: Measurement of gear tooth profiles using incoherent line structured light. *Measurement* **189**, 110450 (2022)
16. Sipiran, I., Bustos, B.: Harris 3D: a robust extension of the Harris operator for interest point detection on 3D meshes. *Visual Comput.* **11**(27), 963–976 (2011)
17. Baker, S., Matthews, I.: Lucas-kanade 20 years on: a unifying framework[J]. *Int. J. Comput. Vision* **56**(3), 221–255 (2004)
18. RICOH.: Imaging performance without compromise. <https://www.ricoh-imaging.co.jp/english/products/gr-3/feature/>, 2019
19. Intel.: Tech Specs, <https://www.intelrealsense.com/depth-camera-d455/>, 2020

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.