An Improved Sum of Squared Difference Algorithm for Automated Distance Measurement

Yue Lin, Yixun Gao and Yao Wang*

Guangdong Provincial Key Laboratory of Optical Information Materials and Technology and Institute of Electronic Paper Displays, South China Academy of Advanced Optoelectronics, South China Normal University, Guangzhou, China

In recent years, with improvement of photoelectric conversion efficiency and accuracy, photoelectric sensor was arranged to simulate binocular stereo vision for 3D measurement, and it has become an important distance measurement method. In this paper, an improved sum of squared difference (SSD) algorithm which can use binocular cameras to measure distance of vehicle ahead was proposed. Firstly, consistency matching calibration was performed when images were acquired. Then, Gaussian blur was used to smooth the image, and grayscale transformation was performed. Next, the Sobel operator was used to detect the edge of images. Finally, the improved SSD was used for stereo matching and disparity calculation, and the distance value could be obtained corresponding to each point. Experimental results showed that the improved SSD algorithm had an accuracy rate of 95.06% when stereo matching and disparity calculation were performed. This algorithm fully meets the requirements of distance measurement.

Keywords: sum of squared difference, photoelectric sensor, distance measurement, binocular cameras, stereo matching

INTRODUCTION

A photoelectric sensor is a semiconductor device which can convert light signals into electrical signals based on photoelectric effect. Photoelectric effect means that the energy of photons was absorbed by electrons of material when light is irradiated on materials, and then corresponding electric effect occurs. Photoelectric sensors were widely used in detection field because it has advantages of short response time, long detection distance and high resolution [1].

In recent years, three-dimensional measurement has been carried out by arranging photoelectric sensors to simulate binocular stereo vision. This method has the advantages of long detection distance, recognizable color, and wide application range. It has become an important distance measurement method. Binocular stereo vision is also an application of machine vision. It uses left and right cameras to imitate human left and right eyes. Based on the parallax principle, the imaging device is used to obtain two images of the measured object from different positions, and the three-dimensional geometric information of the object is obtained by calculating the position deviation between corresponding points of an image [2]. Binocular camera sensor ranging method has gradually become the most common stereo ranging method [3, 4]. Therefore, in order to match different parts at the same time, a real-time binocular stereo vision system which used Field Programmable Gate Array (FPGA) for full line design was proposed to improve the processing speed [5]. This algorithm took advantage of parallel computing and fast operation speed in FPGA, but the

versatility was poor and targeted program development was needed. Then, high and low texture scenes were processed separately, a time of flight (TOF) depth camera was used to measure the low texture scene, and a binocular camera was used to measure the high-resolution scene, so as to improve the accuracy of three-dimensional measurement [6]. It needed to equip with TOF sensors, and the hardware design was complicated. In addition, in two sets of binocular vision systems with different accuracy levels, the method of triangulation analysis and spatial plane fitting were proposed to calculate relative poses [7]. This algorithm had a higher accuracy than other methods, but it required four cameras, which had a high cost. And algorithms that recognize objects first was proposed. For example, Canny edge detection was used to extract the target contour in order to determine the three-dimensional coordinates of the target [8]. Likewise, precise position of the target was used to determine robustness of the ellipse fitting strategy after located its approximate area [9]. And Canny edge detection based on binocular vision was used to identify and locate the target [10]. But these algorithms cannot be applied to other occasions. Then, roll angles of the binocular measurement system were calculated to compensate for dynamic measurement errors by installing a tilt angle sensor horizontally [11]. This algorithm required additional sensors to assist in attitude detection, which could increase the complexity of the system. And then a method based on the parallel binocular vision system and similarity judgment function was established by using cluster analysis method, and the distance was calculated by combining features of gradient histogram and cascade classifier [12]. This method had high a computational complexity and a slow operation speed. And combined with polar line matching, the Multi-line centerline detection stereo matching method was proposed for distance estimation [13]. This algorithm required multiple mathematical model conversions, which was not conducive to the real-time performance of the system.

Therefore, we proposed the improved SSD algorithm for automated distance measurement. It could improve the accuracy of automated distance measurement. The algorithm mainly includes five algorithms: gray-scale transformation, Gaussian Blur transformation, edge detection by Sobel operator, model of a binocular stereo camera, improved Sum of Squared Difference. It can accurately perform stereo matching and distance measurement.

SYSTEM DESIGN PRINCIPLE

Gray-Scale Transformation

In order to reduce the amount of calculation and improve the real-time performance, the image data is grayscale transformed. Grayscale transformation is a combination of three channel gray value calculation. According to the importance of three primary colors, three components of RGB are weighted and averaged with different weights, **Eq. 1** can obtain a grayscale image [14], where i and j represent coordinates of horizontal and vertical of images, R

FIGURE 1 | Parallax distance similar triangle calculation model diagram.
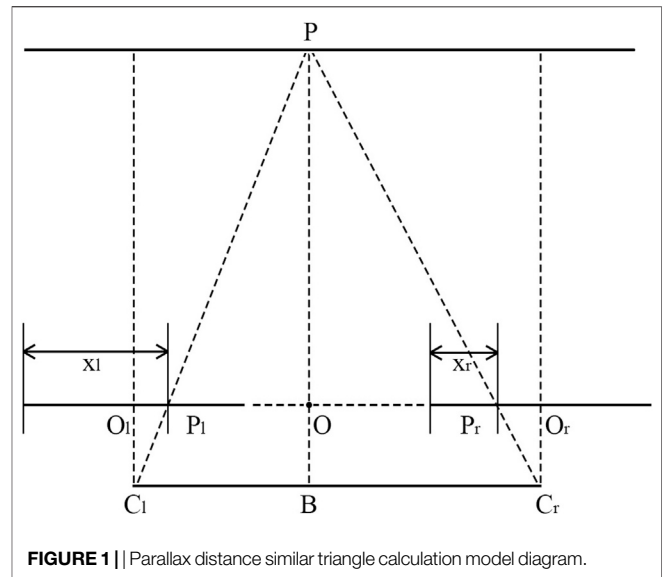
(i, j), $G(i, j)$ and $B(i, j)$ respectively represent components of a points in row i and column j of three primary colors.

$$f(i, j) = 0.299 * R(i, j) + 0.587 * G(i, j) + 0.114 * B(i, j) \quad (1)$$

Gaussian Blur Transformation

Before edge detection, the image is filtered and denoised. It can reduce the interference of the original noise on edge detection. Therefore, Gaussian blur transformation is used to reduce the level of image detail and noise interference, so that images can become smoother and easier to perform stereo matching [15, 16]. Gaussian Blur transformation is defined as **Eq. 2**, where u is abscissa of pixel, v is ordinate of pixel, σ is blur radius, and it is standard deviation of normal distribution.

$$G(u, v) = \frac{1}{2\pi\sigma^2} e^{-\left(u^2 + v^2\right)/\left(2\sigma^2\right)} \quad (2)$$

Edge Detection by Sobel Operator

The surface of some objects is smooth, whose features are not obvious. And stereo matching may be affected seriously. In order to obtain a high-level feature and improve the accuracy of stereo matching, the first derivative Sobel operator can be used to perform gradient extraction operation of spatial convolution. And edge can be extracted from the image to make the feature more obvious [17, 18], the Sobel operator is shown in **Eq. 3**, where I is original image, G_x and G_y are Convolution factor of longitudinal and horizontal axis direction.

$$G = \sqrt{\left(G_x * I\right)^2 + \left(G_y * I\right)^2}$$

$$= \sqrt{\begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix}^2 * I^2 + \begin{bmatrix} -1 & 2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix}^2 * I^2} \quad (3)$$

Model of the Binocular Stereo Camera

The depth of an object refers to the distance between the camera and object. It is measured for distance by ultrasonic, laser, etc. But these methods can only measure a specific point. We measured distance by arranging left and right cameras to simulate parallax of the image, which can be obtained by human binocular. Binocular parallax is a position difference of imaging pixel coordinates in left and right cameras. Therefore, stereo matching can be performed on images which are obtained by the left and right cameras. Cost calculation can obtain three-dimensional information of each point in an image in the real scene [19, 20]. This algorithm uses the principle of similar triangles to construct a parallax distance calculation model. The model is shown in **Figure 1**.

Among them, C_l and C_r are left and right camera sensors respectively, and O_l and O_r are corresponding focal points. P is a point to be measured, P_l and P_r are imaging points on a camera lens, and x_l and x_r are corresponding offset pixel point distances respectively. B is a baseline of left and right cameras, that is, the distance between C_l and C_r. Triangle PP_lO and $C_lO_lP_l$ are similar triangles, so are PP_rO and $C_rO_rP_r$. And triangles PP_lP_r and PC_lC_r are also similar triangles, so **Eq. 4** can be obtained, where Z is the distance from the image sensor to the target object, which is a distance value from point P to baseline B, f is the focal length of the camera, x_l is the abscissa of the image on left, and x_r is the ordinate of the image on left.

$$\frac{Z}{B} = \frac{Z - f}{B - (x_l - x_r)} \tag{4}$$

The available distance between the image sensor to the target object is shown in **Eq. 5**.
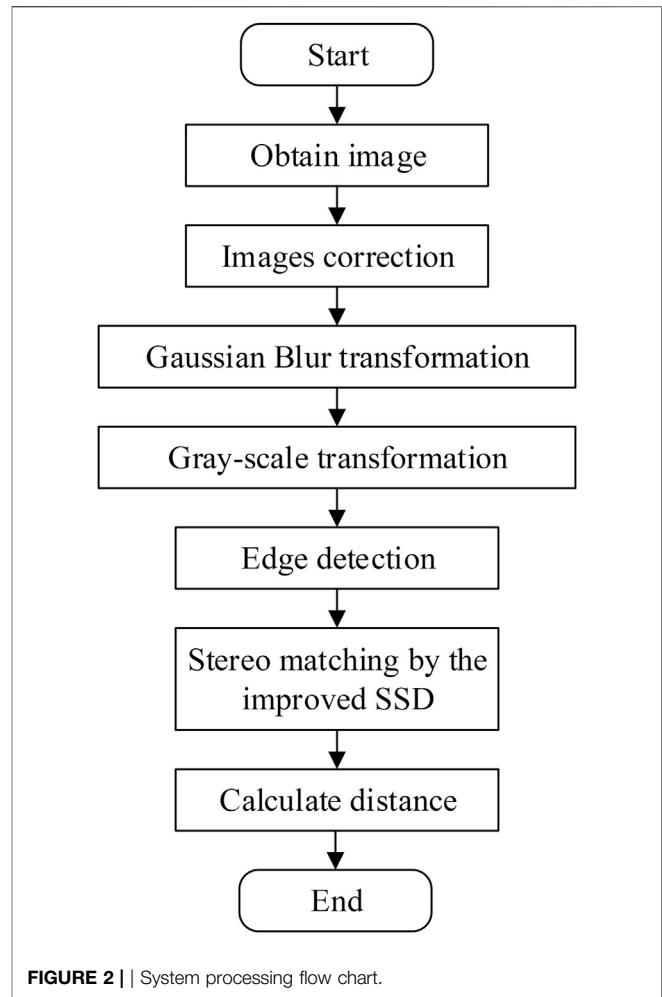
$$Z = \frac{Bf}{x_l - x_r} \tag{5}$$

Improved Sum of Squared Difference

Stereo matching cost calculation can be used to compare similarity between a certain area of an image and a certain area of another image. Common stereo matching algorithms include Sum of Absolute Difference (SAD), Sum of Squared Difference (SSD), Normalization Cross- Correlation (NCC), Census, etc [21–23]. Among these algorithms, SSD has lowest implementation complexity, but the matching accuracy is relatively poor. So we optimized the basis of SSD and proposed an improved SSD, which can improve matching accuracy and maintain original lightweight calculation at the same time.

The improved Sum of Squared Difference is shown in **Eq. 6**, where I_l is left image, I_r is the right image, c is the center area of target window, r is edge area of target window, and d is target window in coordinate difference between left image and right image, x is left horizontal axis, y is coordinate vertical axis, and a is weight coefficient.

$$ISSD(x, y, d) = \sum_{(x,y) \in c} \left[I_l(x, y) - I_r(x - d, y) \right]^2$$
$$+ a * \sum_{(x,y) \in r} \left[I_l(x, y) - I_r(x - d, y) \right]^2 \tag{6}$$
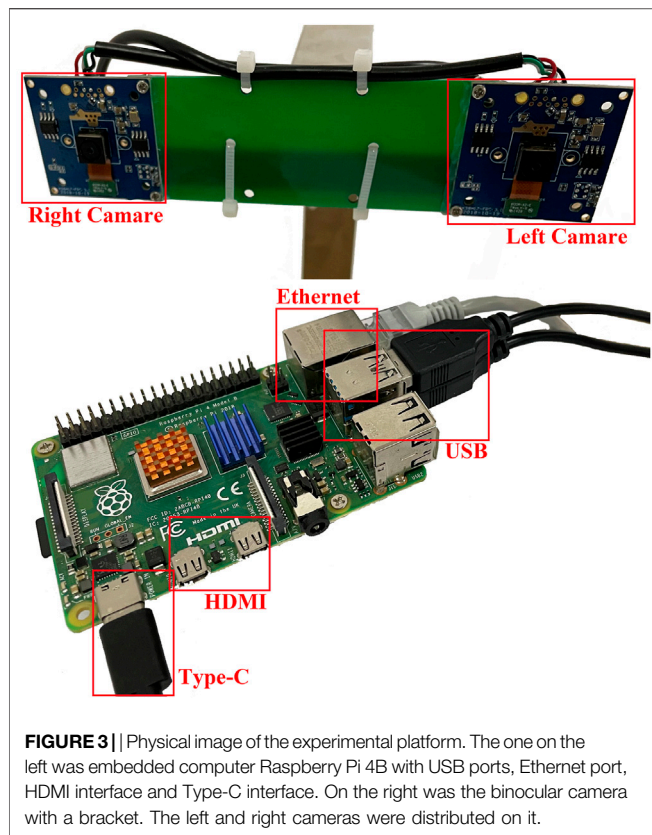

FIGURE 2 | | System processing flow chart.

EXPERIMENT AND DISCUSSION

The platform was equipped with a left camera and a right camera. Before acquiring images, it was necessary to match parameters of the binocular camera, and obtain the correction matrix parameters between left and right cameras to keep them consistent. During the measurement, two frames of images were obtained by each camera at the same time, and parameters of images were calibrated. Then Gaussian blur was used to smooth image, gray scale transformation was performed and Sobel operator was used for edge detection. Finally, improved SSD was used to perform stereo matching and disparity calculation on a target image, and then the distance value could be obtained corresponding to each point, which was combined into a distance value matrix. The specific processing flowchart of system is shown in **Figure 2**.

System Hardware Architecture

The experimental platform is shown in **Figure 3**. Embedded computer Raspberry Pi 4B was used as the main processor to build an experimental platform. An ARM Cortex-A72 quad-core processor with a main frequency of 1.5 GHz was used with

FIGURE 3 | Physical image of the experimental platform. The one on the left was embedded computer Raspberry Pi 4B with USB ports, Ethernet port, HDMI interface and Type-C interface. On the right was the binocular camera with a bracket. The left and right cameras were distributed on it.

Raspberry Pi 4B. It was equipped with 2 GB of LPDDR4 memory, it also had two USB 3.0 ports and two USB 2.0 ports, two external micro HDMI ports and a gigabit Ethernet port. It can be directly powered by a USB Type-C port with 5 V. The operating system was UBUNTU18.04, PYTHON3.7, and OPENCV2.0 environments were deployed. The image sensor was composed of two left and right cameras placed in parallel, and the space was 10 cm between the left and right cameras, which was called a center baseline. The camera was connected to the laptop by USB ports. The camera used a low-sensitivity Sony high-speed CMOS camera with a resolution of 1,920*1,080 pixels and it could obtain 60 frames of images per second. In order to reduce the difficulty of stereo matching, two cameras were guaranteed on the same horizontal line, and the inclination of the horizontal plane was set to be consistent.

Image Preprocessing

Image preprocessing mainly included image correction, Gaussian smoothing, and grayscale transformation. Pre-calibrated differential parameters of the left and right camera were applied for image correction, so that initial coordinates of the left and right image were consistent. The transformation effect is shown in **Figure 4**. We could draw a red straight line on the image. It could be seen from the part in the yellow box, the left car light before calibration was completely above the red line in the left image, but there was still a small part below the red line in the right image. Moreover, the left car light was relatively close to the red line in

the left and right images after calibration. And pixels with the same feature were almost in the same row on the left and right images. That was a prerequisite for stereo matching to improve accuracy, which was conducive to subsequent accurate feature matching.

Then, Gaussian smoothing was performed for the image. All noise must be filtered on the image so that the image level was excessively relaxed, and the influence of noise on subsequent stereo matching was reduced. The license plate area was used as an example, the transformation effect is shown in **Figure 5**. After Gaussian smooth transformation, the gradient of the image was smoother, the sense of hierarchy was reduced, and the feature difference between the left and right cameras, which was caused by different viewing angles, was reduced. Then, it was easier to perform stereo matching and improve matching accuracy.

Finally, the grayscale transformation was performed to reduce the amount of calculation for stereo matching. By mixing RGB channels in proportions, two-dimensional matrix data was generated, the original features of the image were retained as much as possible when the amount of data was reduced.

Edge Extraction

Some objects in an image had a smooth surface. For example, there is a slight reflection on the body surface. Due to the reflection, the image appears to lack texture features. Unobvious features would lead to inaccurate matching results and low stereo matching recognition rate. In order to highlight the details of the smooth part, edge extraction was performed on the image. Gradient operations were performed on the vertical and horizontal directions of the image, which could magnify the small edge changes of inconspicuous texture surface, and the advanced features of smooth part were highlighted on the image. The transformation effect is shown in **Figure 6**. It could be seen that the processed image had obvious features in the areas near the license plate and the lower part of a car body. The original smooth surface had gradient edge features, so the accuracy of stereo matching could be improved with their inconspicuous textures, and the stereo matching recognition rate could be also improved.

Stereoscopic Matching of Left and Right Photos

Stereo matching was the most critical step of this algorithm. The same feature areas of the left and right images were matched. It meant that the feature recognition of the left image was performed on the same horizontal line in the right image. After calculation, it could find out the area with the highest similarity to the left image. According to the horizontal coordinate migration of the same feature, the depth value of the feature area could be calculated. The migration was the distance value from the baseline to objects. The recognition accuracy of stereo matching determined the ranging accuracy. Compared with other stereo matching algorithms, the gray value of the pixel in

FIGURE 4 | | Comparison images of correction. **(A)** is the left image before calibration, and **(B)** is the right image before calibration. **(C)** is the left image after calibration, and **(D)** is the right image after calibration.

FIGURE 5 | | Comparison images of Gaussian Blur Transform in the license plate area. **(A)** was the image before transformation, and **(B)** was the image after transformation.

the corresponding area was distinguished, and then the absolute value of the result was obtained by using the improved SSD. So that the matching of the central region played a more important and dominant role in the whole

FIGURE 6 | | Comparison images of edge detection. **(A)** was the original image, and **(B)** was the effect image after edge detection by using the SOBEL operator.

TABLE 1 | Measurement results.

Distance (m)	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
Ours (m)	0.51	1.03	1.55	2.09	2.63	3.17	3.7	4.24	4.8	5.37
Du Jiang's (m)	0.48	0.95	1.41	1.88	2.34	2.8	3.26	3.72	4.16	4.58
Gap (%)	2.00	2.00	2.67	1.50	1.20	1.00	1.14	1.00	0.89	1.00

The first line was the actual distance between binocular camera and target, the second line was measurement distance by using improved SSD algorithm, the third line was measurement distance of the comparison algorithm, the unit was meter. The fourth line was the actual accuracy difference between our algorithm and the comparison algorithm, the unit was a percentage.

matching results. In order to verify the accuracy and performance of our algorithm, an experimental platform was used to control the binocular camera at a specific distance from the target, and image acquisition was performed. Then, Du Jiang et al. proposed the binocular matching (BM) algorithm which run at the same time to record the data [24]. Measurement results were shown in **Table 1**. It can be seen that the actual measurement accuracy of the improved SSD algorithm was 95.06% on average. It was 1.44% higher than the comparison algorithm. So, it had a higher accuracy and smaller accuracy error fluctuations, which meets actual measurement requirements.

CONCLUSION

In this paper, an algorithm, which could configure the left and right cameras, to obtain images from different angles of the target was proposed. The images were smoothed to reduce noise, and grayscale transformation was performed to reduce stereo matching operations. Then, corresponding area pixels were matched on the same horizontal line of the right image. The absolute value of the degree was made difference and then summed, and the central key area was weighted, and the matching of the central area played a more important role in entire matching results. The algorithm improved the matching accuracy and could obtain measurement results more accurately, which fully met the requirements of distance measurement.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

YW and YL designed this project. YL carried out most of the experiments and data analysis. YG performed part of the experiments and helped with discussions during manuscript preparation. YW and YL contributed to the data analysis and correction and provided helpful discussions on the experimental results. All authors have read and agreed to the published version of the manuscript.

REFERENCES

1. Yi Z-C, Chen Z-B, Peng B, Li S-X, Bai P-F, Shui L-L, et al. Vehicle Lighting Recognition System Based on Erosion Algorithm and Effective Area Separation in 5G Vehicular Communication Networks. *IEEE Access* (2019) 7:111074–83. doi:10.1109/ACCESS.2019.2927731
2. Liu Y, Wang H, Dong C, and Chen Q. A Car-Following Data Collecting Method Based on Binocular Stereo Vision. *IEEE ACCESS* (2020) 8:25350–63. doi:10.1109/ACCESS.2020.2965833
3. Zhou Y, Li Q, Chu L, Ma Y, and Zhang J. A Measurement System Based on Internal Cooperation of Cameras in Binocular Vision. *Meas Sci Technol* (2020) 31(6):065002. doi:10.1088/1361-6501/ab6ecd
4. Wang Q, Wang Z, Yao Z, Forrest J, and Zhou W. An Improved Measurement Model of Binocular Vision Using Geometrical Approximation. *Meas Sci Technol* (2016) 27(12):125013. doi:10.1088/0957-0233/27/12/125013

5. Ma J, Yin W, Zuo C, Feng S, and Chen Q. Real-time Binocular Stereo Vision System Based on FPGA. In: *2018 Optical & Photonic Engineering*. Shanghai: ICOPEN (2018). doi:10.1117/12.2500769
6. Yang Y, Meng X, and Gao M. Vision System of Mobile Robot Combining Binocular and Depth Cameras. *J Sensors* (2017) 2017:1–11. doi:10.1155/2017/4562934
7. Chen Y, Zhou F, Zhou M, Zhang W, and Li X. Pose Measurement Approach Based on Two-Stage Binocular Vision for Docking Large Components. *Meas Sci Technol* (2020) 31(12):125002. doi:10.1088/1361-6501/aba5c7
8. Han Y, Zhao K, Chu Z, and Zhou Y. Grasping Control Method of Manipulator Based on Binocular Vision Combining Target Detection and Trajectory Planning. *IEEE Access* (2019) 7:167973–81. doi:10.1109/ACCESS.2019.2954339
9. Han Y, Chu Z, and Zhao K. Target Positioning Method in Binocular Vision Manipulator Control Based on Improved Canny Operator. *Multimed Tools Appl* (2020) 79(13-14):9599–614. doi:10.1007/s11042-019-08140-9

10. Wan G, Li F, Zhu W, and Wang G. High-precision Six-Degree-Of-freedom Pose Measurement and Grasping System for Large-Size Object Based on Binocular Vision. *Sr* (2020) 40(1):71–80. doi:10.1108/SR-05-2019-0123

11. Cai C, Qiao R, Meng H, and Wang F. A Novel Measurement System Based on Binocular Fisheye Vision and its Application in Dynamic Environment. *IEEE Access* (2019) 7:156443–51. doi:10.1109/ACCESS.2019.2949172

12. Zhang S, Li B, Ren F, and Dong R. High-Precision Measurement of Binocular Telecentric Vision System with Novel Calibration and Matching Methods. *IEEE Access* (2019) 7:54682–92. doi:10.1109/ACCESS.2019.2913181

13. Yu X, Liu S, Pang M, Zhang J, and Yu S. Novel SGH Recognition Algorithm Based Robot Binocular Vision System for Sorting Process. *J Sensors* (2016) 2016:1–8. doi:10.1155/2016/5479152

14. Wen K, Li D, Zhao X, Fan A, Mao Y, and Zheng S. "Lightning Arrester Monitor Pointer Meter and Digits Reading Recognition Based on Image Processing," In Proceedings of 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference; Chongqing. IEEE (2018). doi:10.1109/iaeac.2018.8577545

15. Kostková J, Flusser J, Lébl M, and Pedone M. Handling Gaussian Blur without Deconvolution. *Pattern Recognition* (2020) 103:107264. doi:10.1016/j.patcog.2020.107264

16. Zhao F, Zhao J, Han X, Wang H, and Liu B. Robust Image Reconstruction Enhancement Based on Gaussian Mixture Model Estimation. *J Electron Imaging* (2016) 25(2):023007. doi:10.1117/1.JEI.25.2.023007

17. Yang SH, and Hsiao SJ. H.266/VVC Fast Intra Prediction Using Sobel Edge Features. *Electron Lett* (2020) 57:11–3. doi:10.1049/ell2.12011

18. Joe H, and Kim Y. Compact and Power-Efficient Sobel Edge Detection with Fully Connected Cube-Network-Based Stochastic Computing. *Jsts* (2020) 20(5):436–46. doi:10.5573/JSTS.2020.20.5.436

19. Liu Z, Liu X, Duan G, and Tan J. Precise Pose and Radius Estimation of Circular Target Based on Binocular Vision. *Meas Sci Technol* (2019) 30(2):025006. doi:10.1088/1361-6501/aaf8c8

20. Jiang J, Liu L, Fu R, Yan Y, and Shao W. Non-horizontal Binocular Vision Ranging Method Based on Pixels. *Opt Quant Electron* (2020) 52(4):52. doi:10.1007/s11082-020-02343-3

21. Wang Z, Yue J, Han J, Jin Y, and Li B. Regional Fuzzy Binocular Stereo Matching Algorithm Based on Global Correlation Coding for 3d Measurement of Rail Surface. *Optik* (2020) 207:164488. doi:10.1016/j.ijleo.2020.164488

22. Vazquez-Delgado HD, Perez-Patricio M, Aguilar-Gonzalez A, Arias-Estrada MO, Palacios-Ramos MA, Camas-Anzueto JL, et al. Real-time Multi-Window Stereo Matching Algorithm with Fuzzy Logic. *IET Comp Vis* (2021) 15(3):208–23. doi:10.1049/cvi2.12031

23. Zhang B, and Zhu D. Local Stereo Matching: an Adaptive Weighted Guided Image Filtering-Based Approach. *Int J Patt Recogn Artif Intell* (2021) 35(3):2154010. doi:10.1142/S0218001421540100

24. Jiang D, Zheng Z, Li G, Sun Y, Kong J, Jiang G, et al. Gesture Recognition Based on Binocular Vision. *Cluster Comput* (2019) 22:13261–71. doi:10.1007/s10586-018-1844-5

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.