Real-time Reconstruction of Unstructured Scenes Based on Binocular Vision Depth

Yinggang Xie[1,2], Jingyu Xing[1], Guangjun Liu[2], Jiangyu Lan[1], Yaoyao Dong[1]

[1] School of Information & Communication Engineering, Beijing Information Science and Technology University, China
[2] Department of Aerospace Engineering, Ryerson University, Canada
xieyinggang@bistu.edu.cn, 1924713967@qq.com, ljy960716@163.com, gjliu@ryerson.ca, dahai-009@163.com

Abstract

Binocular vision is a method that simulates the principle of human vision and uses a computer to passively perceive distance. By obtaining the depth of field information of the object, the actual distance between the object and the camera can be calculated. Based on the improved SURF (Speeded-Up Robust Features) algorithm, this paper implements image feature extraction from different perspectives. The image fusion based on improved Sobel algorithm is used to achieve image fusion and image feature point matching. The triangulation principle is used to calculate the offset between pixels to obtain the three-dimensional information of the object, reconstruct the three-dimensional coordinates, analyze the actual depth, establish the bad point culling rule based on the numerical relationship of the image sequence, and finally use the visual depth information to construct the unstructured 3D (3 dimensions) real-time scene. The experimental results show that for the target in the actual unstructured scene, the average error of 0.1m to 3m is 2.99%; the average error of 3m to 10m is 5.81%, and the system achieves higher measurement accuracy and better 3D reconstruction effect.

Keywords: Binocular vision, Three-dimensional reconstruction, Visual depth, SURF algorithm

1 Introduction

With the wide application of computer vision, the advantages of non-contact measurements become more and more obvious [1]. on the basis of binocular vision with binocular obtain relatively divided into left and right two pictures, the two pictures with computer by matching algorithm of feature point matching and obtain parallax, through the depth of the parallax can get corresponding to each point information and three-dimensional coordinates, this is the summary of the basic technical basis of binocular vision [2]. The early development of binocular vision is mainly applied to the matching of both eyes, using two plane images to generate stereoscopic depth information [3]. Binocular vision has developed in many aspects to no longer aimed at the simulation binocular vision field, broaden the development gradually [4-5], and is widely used in the fields of Internet of Things, smart cities and artificial intelligence [6-7]. binocular vision in the fields of intelligent robot navigation and fine micro-operation has already had the superiority and precision unmatched by the human eye [8-10].

This paper is a preliminary study on the functional verification and method improvement of binocular depth in the navigation of laboratory robot and obstacle avoidance of underwater robot. The purpose of this study is to provide support for scene reconstruction of our laboratory robot and the ongoing underwater robot project.

The mobile robot needs to use a parallel calibrated binocular camera to directly calculate the depth information of the scene through the principle of triangulation, and it can take into account the close distance and long-distance depth measurement, which is a method of scene reconstruction. The underwater pipe robot needs to check the position (distance) of the butterfly valve in front and valve direction. Due to the influence of the tube wall, the application of ultrasonic and radar is limited, binocular visual depth is an effective method. This study lays a foundation for the related work of our laboratory.

In this paper, the research and implementation of a double video target distance estimation algorithm are completed, and its application effect is studied. In this project, the target area image is acquired by dual cameras, and the distance estimation of the observed target is achieved by comparing and analyzing the images and verified the application effect of the algorithm. Section 2 reviews the research status of binocular vision technology. Section 3 introduces the binocular vision depth system framework. Section 4 discusses the specific process of binocular vision depth realization based on the SURF algorithm. Section 5 introduces the experimental verification and statistical analysis. Section 6 summarizes this paper and Outlines

*Corresponding Author: Yinggang Xie; E-mail: xieyinggang@bistu.edu.cn
DOI: 10.3966/160792642019092005025

future research trends.

2 Relate Work

Depth measurement is the primary problem of scene reconstruction. It can be realized by various devices such as laser ranging, ultrasonic, radar, stereo camera, monocular camera, etc. The method based on laser equipment is too expensive. Based on the multi-sensor data fusion method, the matching time is long, and it is easily interfered by various external factors, and the stability is poor. The vision-based approach is relatively low cost, stable, and easy to implement [11].

In recent years, binocular stereo vision, which uses two cameras to obtain a disparity map and can then be used to calculate depth information and implement 3D reconstruction, has become popular in the field of computer vision [12]. Stereo matching for obtaining disparity maps is a key issue in computer vision applications and one of the most widely studied issues [13-14]. Common matching algorithms include region-based matching, feature-based matching, and phase-based matching algorithms [15]. The region-based matching algorithm is an easy-to-implement algorithm, but the algorithm has a large number of calculations and high hardware requirements. The phase-based matching algorithm uses "phase" as baseband and bandpass filtering to effectively suppress high-frequency noise and obtain dense disparity maps [16]. The feature-based matching algorithm selects the feature part with the highest resolution in the image as the primitive, weakens the overall image matching and reduces the matching time. The algorithm has fast calculation speed and good effect. This paper is based on this method. Li et al. [17] proposed the FAST (Features from Accelerated Segment Test) feature matching algorithm, which can detect feature points without feature description, but it has poor detection effect and is prone to false detection. SIFT [18] (Scale-Invariant Feature Transform) has been proved to be invariant to image scaling, translation, rotation, and also keep good stability to affine transformation, illumination changes, partial occlusion, noise and so on. Therefore, SIFT is selected as the original algorithm in this paper. Bay et al. [19] proposed SURF (Speeded-Up Robust Features) inspired by the SIFT descriptor, uses a Hessian matrix-based measure for the detector and the Haar wavelet response around the point of interest for the descriptor. Zhou et al. [20] proposed A Gauss kernel scale adaptive method is used to simplify the pyramid and reduce the complexity of the original algorithm. In a large number of matching algorithms, SURF feature matching algorithm performs well [21]. Zeng et al. [22] show SURF matching was further improved by combining polar constraints with SURF matching. Yang et al. [23] proposed Exposing Copy-Move Forgery based on improved SIFT descriptor, Yuan et al. [24] proposed Coverless Image Steganography based on SIFT and BOF. Moreover, there are still many other methods [25-31] are proposed to improve the SIFT algorithm in different aspects. However, some improved SIFT algorithms above are for specific situations and needs, and others increase the complexity of the algorithm or reduce the accuracy of the algorithm, which cannot achieve a good balance in speed, complexity, and accuracy.

Based on the improvement of various algorithms, binocular stereo vision (BSV) systems are becoming more and more important in various application fields [32]. Zhao et al. [33] provides a complete binocular stereoscopic vision system for use by underwater vehicles and obtains better measurement accuracy and 3d reconstruction effect. Huo et al. [34] provides a system that can work well within a distance of 2 meters. to improve the spatial location precision, many researchers optimized the camera calibration algorithm [35]. Cui et al. [36] promoted the computational efficiency of camera calibration. To improve the calibration precision, Jia et al. [37] added perpendicularity compensation in camera calibration. Shih et al. [38] also made a lot of researches for the binocular stereo vision system.

In the above study, the literature [15, 22, 25, 34, 39-40] studied the depth measurement of a simple binocular camera, which is similar to the research in this paper. The results of the study are shown in Table 1.

Table 1. Experimental results of existing research

research method	Measurement error	Instructions
Literature [15]	4.01%	This value is calculated based on the error results of multiple measurements within 1 meter given in the literature.
Literature [34]	5.26%	Only the average error calculated at 2 meters in the literature is used.
Literature [25]	--	Focusing on feature point matching, no measurement error information is given.
Literature [40]	4%	The multiple measurement average error of a fixed point of 0.5 m is given in the literature.
Literature [22]	4%	Only the average error of 1.1 meters given in the literature is taken.
Literature [39]	7.48%	This value is based on the average of multiple measurement errors at 2 meters given in the literature.

Note. Remark: (1) indicates that the literature does not provide data; Remark; (2) due to the different requirements and distances of interest of each research, only the error results of the most important measurement data in each literature are given here.

3 Binocular Visual Depth System

3.1 System Framework

In this project, the target area images are acquired by dual cameras, and the distance estimation of the observed targets is achieved through the comparative analysis of the images. This topic mainly uses the improved SIFT algorithm to obtain the estimation of the non-contact measurement of the distance between two objects, and the output depth map. The system block diagram is shown in Figure 1. The frame diagram of the Measurement System is shown in Figure 2. and the principle of Measurement System is shown in Figure 3.

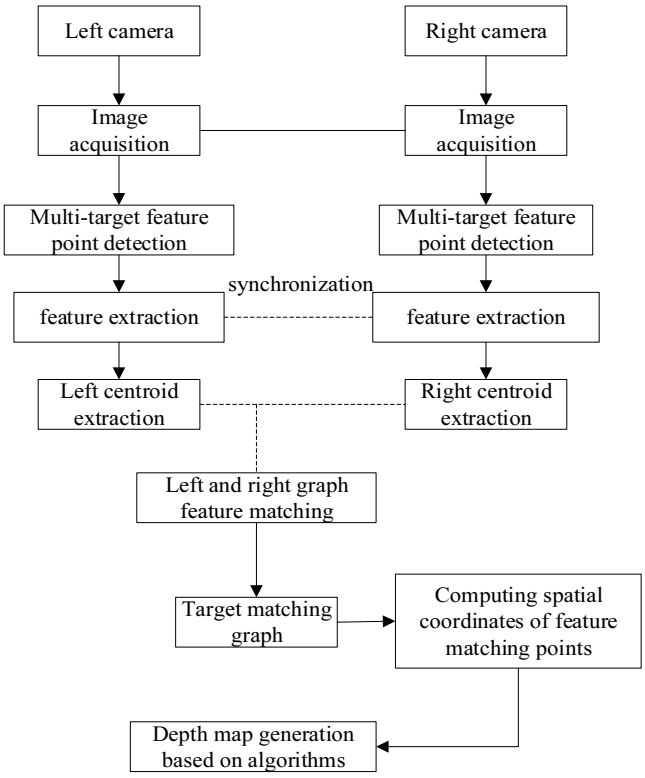


Figure 1. System frame diagram

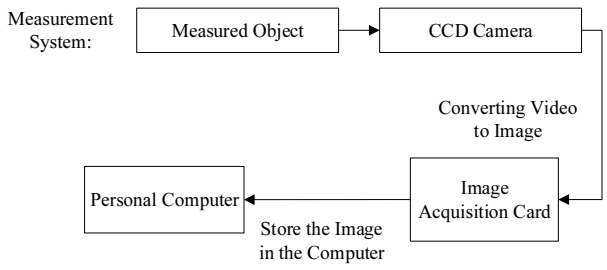


Figure 2. Measurement System diagram of binocular vision analysis

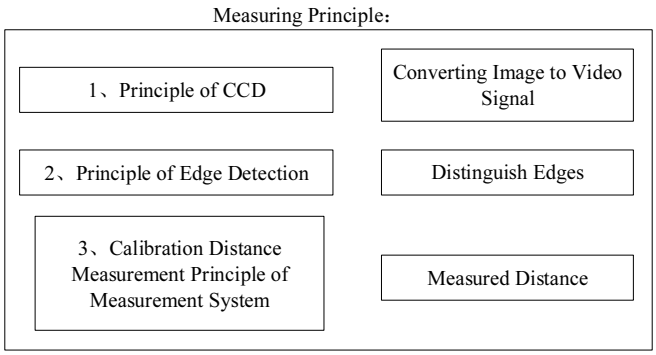


Figure 3. Measuring principle of binocular vision analysis

The point-to-point relationship and depth calculation are determined by the position parameters in the image. The relationship between these positions is determined by the geometric model of camera imaging, whose parameters are camera parameters. These parameters must be determined by experiments and calculations, which is called camera calibration [41]. The digital image processing method is used to analyze and process the images taken by the camera.

In this design, for the part of image processing, it is mainly grayscale transformation and feature extraction of images, as well as the output of the final depth map [42-43]. According to the basic principle of stereo vision, the image of the object is analyzed and processed, so as to carry out 3d reconstruction of the object, restore 3d environment information, and finally realize the 3d coordinate positioning and depth output of the target point [44-45]. In the process of program design, MATLAB is used as the platform, and the Improved Sobel algorithm and Improved SIFT algorithm are used as the basis to implement the later algorithm and system.

3.2 Principle and Theory of Binocular Vision

The specific principle analysis of the camera imaging model has been introduced in detail in the Literature [4, 25, 38]. So this article will not be covered. Based on the above Literature, the relationship between the physical coordinate system and the world coordinate system can be expressed as (1, 2).

$$\begin{aligned} Xw &= \frac{L\cos\alpha\sin\beta}{\sin(\alpha+\beta)} \\ Yw &= \frac{L\sin\beta\tan\theta}{\sin(\alpha+\beta)} \\ Zw &= \frac{L\sin\alpha\sin\beta}{\sin(\alpha+\beta)} \end{aligned} \tag{1}$$

$$\psi(L)=\sqrt{(\frac{\partial Xw}{\partial L})^2+(\frac{\partial Yw}{\partial L})^2+(\frac{\partial Zw}{\partial L})^2}$$
$$=\frac{\sin\beta}{\sin(\alpha+\beta)\cos\theta} \quad (2)$$

In this paper, two cameras are fixed on the stabilizer. After correcting and keeping the camera parallel, the binocular visual depth achieved in this paper is as follows. Binocular ranging based on the principle of human eye imaging, the principle of the range is shown in Figure 4.

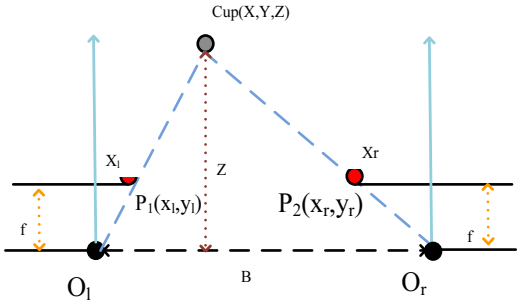


Figure 4. Ranging schematic diagram

As shown in Figure 4, $Cup(x,y,x)$ is the point to be measured on the object, the black point is the position of the two cameras and the distance between the two cameras is measured as B, and $P1$ and $P2$ are respectively the imaging of the two left and right cameras. The camera focal length is also marked from the figure because of this task selects the camera with the same parameters, so the focal length of both cameras are f. Suppose the horizontal channel of the photosensitive chip of two cameras is X, then half is $X/2$, so the distance between $P1$ and $P2$ is:

$$dis=B-(Xr-\frac{X}{2})-(\frac{X}{2}-Xt)$$
$$dis=B+Xt-Xr \quad (3)$$
$$dis=B-(Xr-Xt)$$

Therefore, the final distance between $P1$ and $P2$ is dis, according to the similar triangle principle, thereby obtaining the required depth Z.

$$dis=B-(Xr-Xt)$$
$$\frac{B-(Xr-Xt)}{B}=\frac{Z-f}{Z} \quad (4)$$
$$Z=\frac{fB}{Xr-Xt}$$

Z is the required depth. The focal length of the camera in this project is f=6 mm and B=6 cm. The acquisition of binocular ranging image is also based on the principle of obtaining images through human vision. Images corresponding to the left and right cameras are acquired from the same area by two cameras, parallax is obtained by the algorithm.

4 Algorithm and Implementation

4.1 Feature Matching Based on Improved SIFT Algorithm

The matching of feature points is a crucial step in the depth measurement of two extracted images. In this topic selects the classic feature-based matching algorithm, SIFT, namely the scale-invariant feature transformation algorithm. In some cases, when the shooting conditions are not good, it will inevitably cause some distortion of the image or poor lighting when shooting, in this case, the SIFT algorithm has strong adaptability to these complex conditions, the matching accuracy will be more accurate than other algorithms [46]. In the case of massive data processing, SIDT (Store Interrupt Descriptor Table) has fast reaction speed and computing speed, which can significantly improve its efficiency.

SIFT uses DOG (difference of Gaussians) operator, which is compared with the traditional operator and has higher matching efficiency and precision [47]. The algorithm is divided into the following steps [48-49]:

(1) Construct scale space, detect extremum points and obtain scale invariance;

(2) Filter and precisely locate feature points to eliminate unstable feature points;

(3) At feature points, descriptors are used to assign directional values to feature points;

(4) Generate feature descriptors and use feature descriptions to find matching points;

(5) Calculate change parameters;

In general, the SIFT matching algorithm is mainly divided into two parts, namely, to generate the SIFT feature vector and feature vector matching, so focus on the analysis of these two parts.

4.2 DOG Operator

The scale of space in the present form of expression is the Gaussian pyramid, the Gaussian pyramid generation requires two tasks:

(1) Gaussian blur at different scales of image.

(2) Downsampling the image (alternately sampling)

Perform the detection of the extreme value of the image after generating the Gaussian pyramid, this is also the initial detection of the image feature points. Using the difference of Gaussian pyramid (DOG pyramid) to operate [50]. In the calculation, it's easy to generate the gauss pyramid, Subtract the adjacent upper and lower layers of the gauss pyramid in each group to get the DOG pyramid, detecting the extreme value on the DOG pyramid. The detection of extreme points is done in the adjacent DOG layers within the same group. To determine an extreme point, compare this point with 8 points that are spatially adjacent on the same scale and 9 * 2 = 18 points on both upper and lower scales in space, in total 26 points.

4.3 Improved SIFT Algorithm Based on Improved Sobel Algorithm

The traditional Sobel edge detection is to do 3*3-pixel area through the corresponding template convolution operation, it is concluded that the center pixel gradient values, then compared with preset threshold gradient values if is greater than the threshold criterion for determining the pixel image edge pixels, the opposite is not. The traditional Sobel operator's convolution template has two, one is detection level edge, one is detection vertical edge (Figure 5). Sobel is calculated by taking the maximum value of vertical difference and horizontal difference [51].

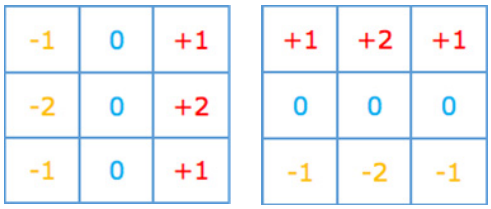


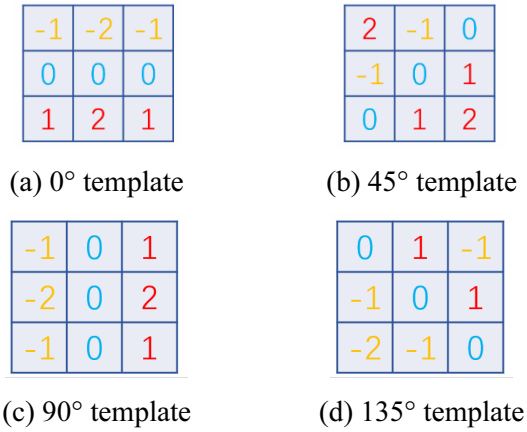
Figure 5. Sobel algorithm filtering template

On the left and right image matching based on SIFT algorithm, SIFT algorithm is to achieve the extraction and matching of image feature points, that is, it is divided into two parts to achieve the extraction of feature points of two images feature points matching, to achieve the new feature point set; The algorithm converts the left and right two images into grayscale images on the premise of collecting them, and then adopts SIFT algorithm on the left and right two images respectively to achieve the final image fusion and image feature point matching, providing key parameters and intuitive performance for the following depth measurement [52].

this paper adopts the method of increasing the direction of gradient detection to solve the problem of misjudgment of the edge of complex images. The improvement idea is 0°, 45°, 90° and 135° four directions, namely with selected pixels around all pixels were compared [53]. The calculation template is shown in Figure 6(a) to Figure 6(d). In addition to the templates, the calculations are the same, and the process of calculating the gradient of the four directions, the synthetic gradient G, then the threshold to determine if it's the edge of the image. Figure 6 is to improve the overall implementation of the Sobel edge algorithm block diagram.

The continuous two-dimensional image function of the operator of 135° is *f(x, y)*, and the expression of the gradient and integral gradient [53] is as follows.

$$\begin{aligned} Gx &= f_x(x,y) \\ &= f(x-1,y+1)+2f(x,y+1)+f(x+1,y+1) \\ &-f(x-1,y-1)-2f(x,y-1)-f(x+1,y-1) \end{aligned} \quad (5)$$



(a) 0° template (b) 45° template

(c) 90° template (d) 135° template

Figure 6. Sobel operator template and pixel template

$$\begin{aligned} Gy &= f_y(x,y) \\ &= f(x+1,y-1)+2f(x+1,y)+f(x+1,y+1) \\ &-f(x-1,y-1)-2f(x-1,y)-f(x-1,y+1) \end{aligned} \quad (6)$$

$$G=\sqrt{G_x^2+G_y^2} \quad (7)$$

The Sobel operator 0°, 45°, 90°, and 135° templates are convolved with *f(x, y)*, respectively [15]:

$$\begin{aligned} L_{0^*} &= (M_{31}+2M_{32}+M_{33})-(M_{11}+2M_{12}+M_{13}) \\ L_{90^*} &= (M_{13}+2M_{23}+M_{33})-(M_{11}+2M_{21}+M_{31}) \\ L_{45^*} &= (M_{23}+2M_{33}+M_{32})-(M_{12}+2M_{11}+M_{21}) \\ L_{135^*} &= (M_{12}+2M_{13}+M_{23})-(M_{21}+2M_{31}+M_{32}) \end{aligned} \quad (8)$$

The gradient magnitude is expressed as an infinite norm as:

$$\begin{aligned} \|L\|_\infty &= \max_1 \sum_{j=1}^{4} |L_{1j}| \\ &= \max\left\{|L_{0^\circ}|,|L_{45^\circ}|,|L_{90^\circ}|,|L_{135^\circ}|\right\} \end{aligned} \quad (9)$$

Overall implementation block diagram of improved Sobel edge algorithm is shown in Figure 7.

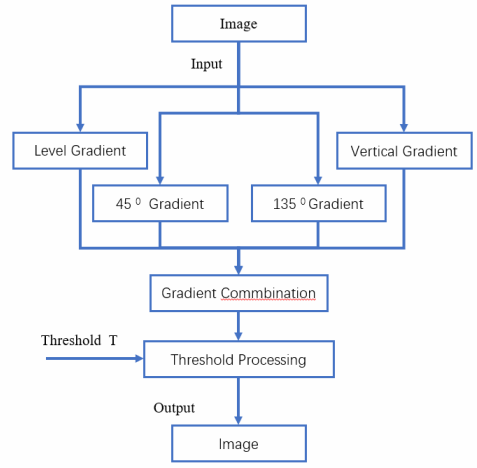


Figure 7. Overall implementation block diagram of improved Sobel edge algorithm

4.4 Improved SIFT Feature Vectors Are Generated

In this project, images captured by cameras can be used to analyze its scale space by using a computer. The scale space is multi-scale features of simulated image data, and the Gaussian convolution sum is the only linear kernel to realize scale transformation [54]. The following is the definition of two-dimensional scale space:

$$L(x,y,\sigma) = G(x,y,\sigma) * I(x,y) \quad \textbf{(10)}$$

Where $G(x,y,\sigma)$ is the Gaussian function with a variable scale, (x,y) is the spatial coordinate, σ is the scale coordinate. In the scale space, the Gaussian pyramid is built through the Gaussian kernel, and the extreme value detection is carried out in it so that the position of feature points can be preliminarily determined.

4.5 Feature Vector Matching

Improved SIFT algorithm is adapted to extract the similarity points of the left and right images captured by the left and right cameras, then the extracted similar points are similarly matching to verify that the scale space is also the same position when the same calibration point is used. In simple terms, the process of feature matching is as follows: by matching the two images to be matched to the feature area respectively, the feature points set of the left and right graphs are generated respectively, and then the two sets are matched by SIFTA (Scale-invariant feature transform advanced) algorithm to generate the new set of matching points Sn. That is, find the similarity points of the corresponding right graph in the left graph. The set of feature points matching each other in the left graph is called Sn, and the offset when the matching similarity reaches the maximum is called parallax [55].

The specific matching conditions corresponding to this test are shown in Figure 8.

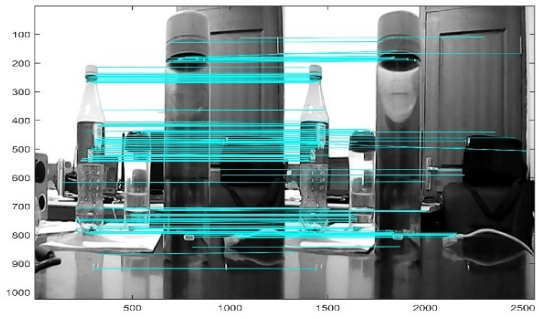


Figure 8. Stereo matching

For this figure, feature points are determined in the feature area of the left figure, and matching points are found through SIFT to generate a new matching set as shown in the figure. This function realizes inputting two pictures, the output is the key point, the round is the feature coordinate of the output key point, first output the left picture column and the left picture line, then output the right picture column and the right picture line.

And in turn the key points of detected into gray image, and in the same coordinate system was achieved by append images statements will be two gray image stack together, calculate the polar of the image, so as to SIFT through matching of key points of two images is realized, the experiment operation output find 134 points (matching point out most of the experimental results has been broken). Corresponding to this design, the core of our implementation is feature matching and ranging. After the above principle of feature matching, now let's analyze the specific results of this experiment.

First, in the feature matching part of the design, two color pictures should be imported. However, the task of feature matching requires the final generation of grayscale images, so the grayscale transformation should be carried out first. After converting to grayscale, you can respectively apply the SIFT algorithm to the left and right two pictures, in order to find two pictures respectively feature point set, finally integrate the two pictures together, and then connect with the polar line, the final output grayscale matching picture. For this figure, feature points are determined in the feature area of the left figure, and matching points are found through SIFT to generate a new matching set as shown in Figure 9.

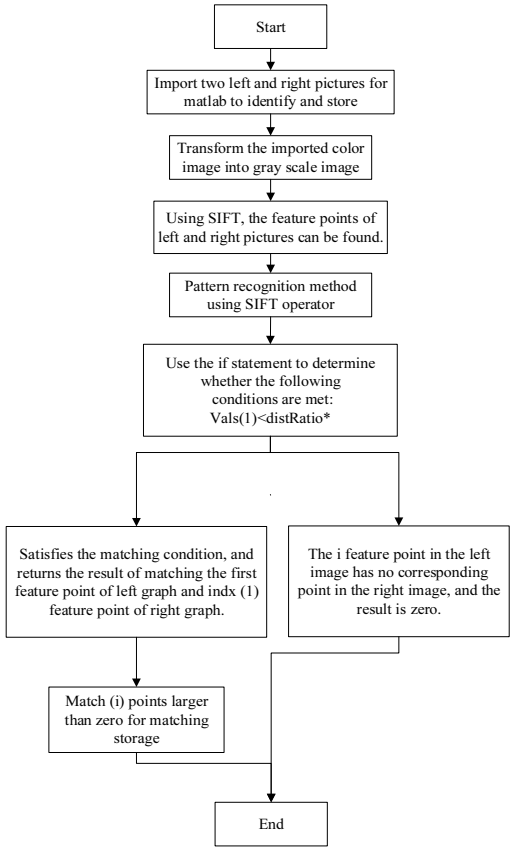


Figure 9. Flow chart of feature matching code design

Design code flow chart as shown in Figure 9 according to the above process.

Because the algorithm does not support color images matching, the task should convert it into a grayscale image, and convert the image of the detected key point into a grayscale image, and in the same coordinate system was achieved by append images statements will be two gray image stack together, calculation of two images of polar, through SIFT algorithm to achieve the two key points of the matching images, The output experiment results are that the left graph finds 2016 feature points, the right graph finds 1884 feature points, and the final output is the new feature point set pairs that the two graphs eventually match, and they are connected by the polar lines.

4.6 Elimination of Bad Points

In the experimental design to achieve the depth determination, the previously calculated feature points need to be calculated algorithmically. However, in this process, this project also needs to consider various situations, such as eliminating bad points. The extreme point that measured in the previous step is in discrete space detection, so in the real sense is not really in the sense of extreme value point, and the point contains many fault points and abandoned points, so these points will affect the real feature points and make large errors in the result [56-57], so the following method is to correct the wrong point and remove the waste point.

Use interpolation to get points close to the extreme. The method of extremum points of continuous space obtained by using known discrete spatial point interpolation is called sub-pixel interpolation. This calculation uses the Taylor function to fit the proportional spatial DOG function to obtain the following formula:

$$D(X) = D + \frac{\partial D^T}{\partial X} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial X^2} x \tag{11}$$

Where $X = (x, y, \sigma)^T$. If the derivative is obtained and the result of the equation is zero, then the offset of the extreme point on the spatial fit curve can be obtained.

$$\hat{X} = -\frac{\partial^2 D^{-1}}{\partial X^2} \frac{\partial D}{\partial X} \tag{12}$$

The corresponding offset extreme point is the detection point of $D(\hat{X})$ through the $\hat{X}$ extreme.

When the offset $\hat{X} = (x, y, \sigma)^T$ is greater than 0.5 in any dimension, this means that the interpolation center has been shifted to its neighboring point, so the position of the current key point must be changed to obtain a new key point.

In addition to the above principles, there are two more intuitive bases for culling bad points. The first one is that when the object is imaged, the sequence on the left must be larger than the sequence on the right, *XL* is greater than *XR*. As long as the left column is smaller than the right column, it will be removed. The second aspect is that the vertical distance of imaging is not very large. Here we specify that it has a range of 10, and if it exceeds 10, it is also considered as a dead point.

In the elimination of bad points, you can write a ranging main code, in the above, this topic has theoretically elaborated on how to achieve ranging. that is, f and B and parallax are the parameters needed for this task. The saved file of the result is shown in Figure 11.

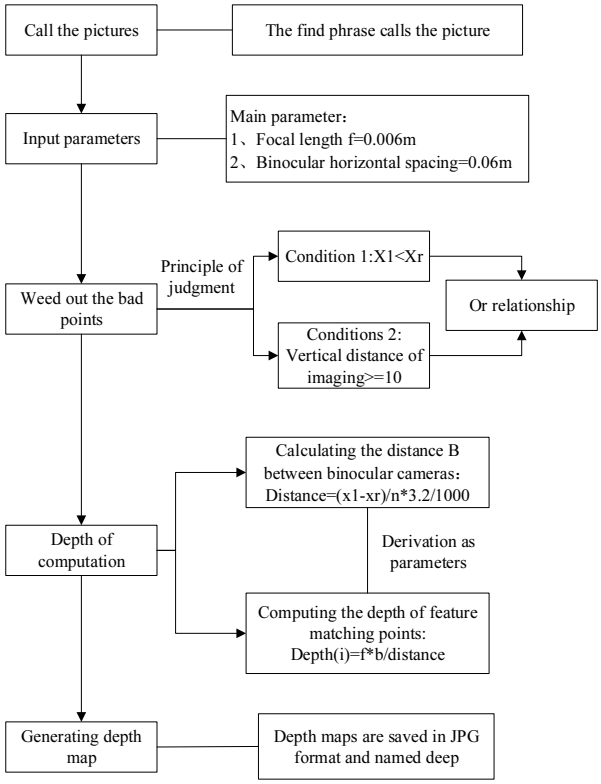
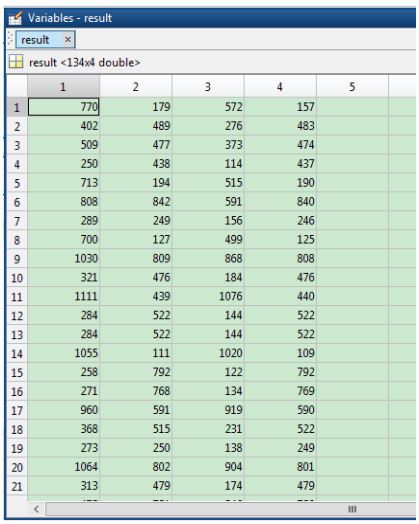


Figure 10. Design flow chart of binocular ranging code



	1	2	3	4	5
1	770	179	572	157	
2	402	489	276	483	
3	509	477	373	474	
4	250	438	114	437	
5	713	194	515	190	
6	808	842	591	840	
7	289	249	156	246	
8	700	127	499	125	
9	1030	809	868	808	
10	321	476	184	476	
11	1111	439	1076	440	
12	284	522	144	522	
13	284	522	144	522	
14	1055	111	1020	109	
15	258	792	122	792	
16	271	768	134	769	
17	960	591	919	590	
18	368	515	231	522	
19	273	250	138	249	
20	1064	802	904	801	
21	313	479	174	479	

Figure 11. Result graph of feature points matching

First, verify the results. For example, we take the data of the first row for verification: on the data of the first row, 1, 2, 3 and 4 respectively represent: the column of the left graph, the row of the left graph, the column of the right graph and the row of the right graph; First, find the corresponding position according to the coordinates of (179,770) in the left figure, and then find the corresponding position points of (157,572) in the right figure, and see whether the last position found is the same position point corresponding to the collected object. It is verified that the last corresponding position is consistent.

Then according to the above-ranging principle, began to find the depth. In the above code, the distance conversion of the distance between two cameras is carried out in this project. In the conversion, the width of the CMOS (Complementary metal–oxide–semiconductor) chip, a sensor, is taken into account, and the parallax calculation on CMOS is finally realized. In the above example, the width of the CMOS is 3.2 mm. When the distance conversion is complete, simple mathematical calculations are required. So, depth of I is equal to f times B/distance, and then you get the final depth (Figure 10). The final depth map results are uniformly marked on the left figure, as shown in Figure 12.

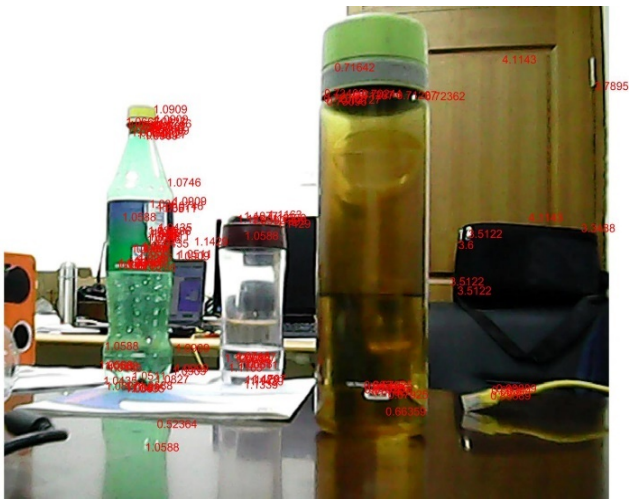


Figure 12. Depth imaging

The above verification is shown in the figure (407,509) on the left and (474,373) on the right. The corresponding points in the figure above are indicated now, which are the points in the box. The following 1.0588 is the depth information indicated, and the actual measured value of this distance is 1.83m, which is relatively accurate, as shown in Figure 13.

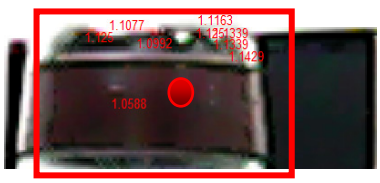


Figure 13. Verification of feature points and distances

5 Experimental and Analysis

5.1 Experimental Preparation Work

The binocular camera used in this experiment is a low-cost binocular camera formed after two ordinary cameras are assembled and fixed. The camera model is a dazzle M800 HD COM camera with 5 million pixels. The focal distance between the two cameras is calculated by the camera parameters and measured by calipers. The actual depth distance of the target was measured by the laser rangefinder. The model of the laser rangefinder was MeiDa high-precision laser rangefinder Y082, with a laser grade of 635NM and an accuracy of 1.0mm. In order to have strong stability of the method, a certain direction is randomly selected in the laboratory for visual depth measurement. In the depth calculation, the specific parameters are: the focal length is f=6 mm, the camera spacing B=6 cm, the width of the sensor CMOS chip is 3.2mm.

5.2 Experimental Results

For the above results, we compared 20 groups of data for result analysis. The corresponding two-dimensional coordinates and distances of the left figure are shown in Table 2.

Table 2. Experimental results

coordinates	The measured distance	The actual distance	Error %
(477,509).	0.724	0.738	1.90%
(522,284).	0.735	0.745	1.34%
(809,1030)	0.679	0.691	1.74%
(842,808).	0.679	0.694	2.16%
(802,1064)	0.900	0.882	2.04%
(194,713).	1.043	1.023	1.96%
(792,258).	1.039	1.012	2.67%
(127,700).	1.036	1.014	2.17%
(515,368).	1.029	1.053	2.28%
(768,271).	1.051	1.077	2.41%
(479,313).	1.125	1.097	2.55%
(250,273).	1.059	1.089	2.75%
(179,770).	1.058	1.083	2.31%
(439,1111)	1.143	1.102	3.72%
(476,321).	1.083	1.126	3.82%
(249,289).	1.091	1.125	3.02%
(438,250).	3.512	3.672	4.36%
(591,960)	3.512	3.678	4.51%
(489,402).	4.114	4.356	5.56%
(111,1055)	4.114	4.331	5.01%

As can be seen from the above table, when the camera distance is constant and the focal length is constant, the distance of the target object is the only factor affecting the error. It can be analyzed from the above table: in the case that the focal length of the camera remains unchanged and the binocular distance remains unchanged, the distance between the target point and the camera can affect the error, as shown in

the figure above the closer the distance, the smaller the error; The farther the distance, the greater the error.

It is likely that the error within one meter is stable at about 3%, and within three meters is stable at about 4%-5%, while the error after three meters is generally more than 5%. It can be clearly seen that as the distance increases, the error becomes an increasing trend. Therefore, we can get that the error within three meters is obviously smaller than the error beyond three meters. In order to get relatively accurate experimental data, we take three meters as the cut-off point to subdivide the distance within three meters and beyond three meters, and finally calculate the average error. Comparison data of 20 groups of measured results and actual values within three meters are shown in Table 3.

Table 3. Experimental results within three meters

coordinates	The measured distance	The actual distance	Error %
(874,559).	0.236	0.238	0.84%
(855,531).	0.387	0.390	0.77%
(884,507).	0.473	0.477	0.84%
(842,514).	0.553	0.565	2.12%
(802,518).	0.679	0.697	2.58%
(769,713).	0.784	0.801	2.12%
(792,801).	1.039	1.011	2.77%
(757,831).	1.082	1.118	3.22%
(770,698).	1.193	1.237	3.56%
(768,915).	1.491	1.543	3.37%
(709,876).	1.947	1.889	3.07%
(677,921).	2.110	2.039	3.48%
(658,774).	2.029	2.111	3.88%
(499,983).	2.425	2.334	3.90%
(476,535).	2.694	2.593	3.90%
(457,633).	2.572	2.672	3.74%
(438,764).	2.947	2.839	3.80%
(406,965).	2.749	2.851	3.58%
(389,660).	3.114	2.988	4.22%
(253,679).	2.917	3.037	3.95%
Average error %		2.99%	

Comparison data of 20 groups of data measurement results and actual values three meters away are shown in Table 4.

Where the average error of 0.1m-3m is 2.99%. The average error between 3m and 10m is 5.81%. In this way, the specific average error within three meters and three meters away from the visual data display. Thus, the linear relationship between the approximate error and distance can be obtained, as shown in Figure 14 and Figure 15.

5.3 Experimental Results Analysis and Discussion

In Table 5, we provide a comparison of our method with other stereo matching methods. Although the research methods are similar, each research focuses on different application objectives, so only the experimental results of interest are given. In the above-

Table 4. Experimental results from three meters away

coordinates	The measured distance	The actual distance	Error %
(496,523).	3.320	3.186	4.21%
(405,531).	3.506	3.701	5.27%
(409,507).	4.356	4.163	4.64%
(511,514).	4.411	4.647	5.08%
(534,518).	4.700	4.967	5.38%
(501,713).	4.902	5.178	5.33%
(517,801).	6.009	5.702	5.38%
(365,821).	5.611	5.932	5.41%
(382,690).	6.505	6.161	5.58%
(404,915).	6.995	6.616	5.73%
(397,836).	6.494	6.908	5.99%
(501,921).	6.528	6.965	6.27%
(323,774).	7.090	7.549	6.08%
(318,983).	8.382	7.903	6.06%
(476,565).	8.873	8.343	6.35%
(294,633).	9.151	8.617	6.20%
(339,758).	8.599	9.185	6.38%
(254,765).	9.042	9.662	6.42%
(281,670).	9.250	9.919	6.74%
(349, 561)	9.722	10.524	7.62%
Average error %		5.81%	

Figure 14. Correspondence of distance and error

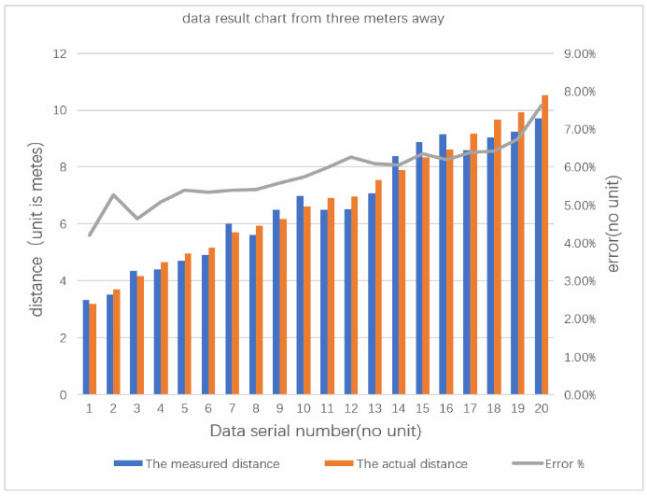


Figure 15. Correspondence diagram of distance and error

Table 5. Comparison of result in this paper with other similar method

Methods	0.1-1 m average error	1-2 m average error	2-3 m average error	3-10 m average error
AD-Census	6.21%	6.89%	-	-
FCVF	6.80%	10.25%	-	-
ARWR	5.75%	3.0%	-	-
SGM	4.46%	11.53%	-	-
Literature [15]	4.01%			
Literature [34]		5.26%		
Literature [40]	4%			
Literature [22]		4%		
Literature [39]			7.48%	
Our approach	1.55	2.2	3.75	5.81%

Note. Remark: (1) m stands for unit meter; Remark: (2) indicates that the literature does not provide data; Remark: (3) the error data of SGM, ARWR, FCVF, and AD-Census are derived from the data provided in Literature [34] and calculated.

mentioned literature participating in the comparison, for example, Literature [40] only gives the error result at 0.5 m, while Literature [22] only gives the error result at 1.5 m. In other studies, when the actual distance is 2 meters, the error of most literature has exceeded 5%. So most literature does not give the results of long-distance experiments. The method proposed in this paper has a wide range of adaptability, with an average error of 1.55% within 1 meter, 2.2% within 2 meters, 3.75% within 3 meters and 5.81% within 10 meters. It can be seen that the measurement error is relatively stable.

5.4 Conclusions and Deficiencies

The experimental results show that, for the target in the actual unstructured scene, the average error of 0.1m to 3m is 2.99%. The average error of 3m-10m is 5.81%. The system achieves higher measurement accuracy, under the same conditions compared with similar methods. Moreover, the algorithm is simple and easy to implement, which can be conveniently embedded in various unstructured environments.

The method in this paper can be realized based on the low-cost binocular camera formed by assembling and fixing two ordinary cameras. The method in this paper aims to reduce the computation and enhance the practicability of the system. It can be realized under low computational load, which is beneficial to the subsequent embedded application. The shortcoming of this paper is that it fails to test the application of the algorithm under extremely complex conditions, such as strong light conditions and very weak light conditions. It fails to explain the algorithm stability of this research method under extreme conditions.

6 Conclusion

In this paper, based on the improved SURF algorithm of image feature extraction and filtering based on Sobel algorithm to realize image fusion and image feature points matching, pixel deviation between calculated through triangulation principle to get three-dimensional information of the object, reconstruction of 3 d coordinate and analyze the actual depth, bad points out based on the numerical relationship between image sequence rules, using visual depth information finally realize real-time 3 d scene reconstruction.

The future will be on the premise of improving the efficiency of the algorithm and adaptive, to reduce the computational complexity practicality as the foundation, enhance the system to improve the algorithm, algorithm also should have adaptability to extreme complex conditions, Even in low light conditions, we can guarantee the stability of the algorithm performance, and achieve a more accurate analysis, the work will be quite a long time in the future practice and continuous optimization.

Acknowledgements

This work is supported by Beijing Natural Science Foundation (Grant No.4192023); Natural Science Foundation of China (Grant No. 61603047); The Qin Xin Talents Cultivation Program of BISTU. (Grant No. QXTCPC201704)

References

[1] K. C. Chen, C. S. Wang, Control System Using Multithread and Stereo Vision to Tracking Target, *Journal of Internet Technology*, Vol. 17, No. 1, pp. 103-107, January, 2016.

[2] Y. F. Wang, C. Shi, Z. F. Xu, L. Cheng, Research on Binocular Visual Feature Points Positioning Based on Improved SURF Matching Algorithm, *Software Engineering*, Vol. 22, No. 3, pp. 5-8, March, 2019.

[3] L. Zhang, D. Z. Li, Research on Mobile Robot Target Recognition and Obstacle Avoidance Based on Vision, *Journal of Internet Technology*, Vol. 19, No. 6, pp. 1879-1892, November, 2018.

[4] X. J. Fan, X. L. Wang, Y. F. Xiao, A Shape-based Stereo Matching Algorithm for Binocular Vision, *Proceedings 2014 IEEE International Conference on Security, Pattern Analysis, and Cybernetics (SPAC)*, Wuhan, China, 2014, pp. 70-74.

[5] P. Angin, B. Bhargava, A Confidence Ranked Co-Occurrence Approach for Accurate Object Recognition in Highly Complex Scenes, *Journal of Internet Technology*, Vol. 14, No. 1, pp. 13-19, January, 2013.

[6] B. Hammi, R. Khatoun, S. Zeadally, A. Fayad, L. Khoukhi, IoT Technologies for Smart Cities, *IET Networks*, Vol. 7, No. 1, pp. 1-13, January, 2018.

[7] J. H. Zhang, Y. Zhu, Scenario-based Trustworthiness

Verification for Systems of Internet of Things, *International Journal of Internet Protocol Technology*, Vol. 11, No. 1, pp. 51-62, January, 2018.

[8] M. Zhou, D. Zhang, T. Zhou, X. Chen, Control of Picking Robot Manipulator Based on Binocular Vision, *Computer and Digital Engineering*, Vol. 46, No. 11, pp. 2355-2359, November, 2018.

[9] Z. Zhang, W. Zhang, H. C. Chao, C. F. Lai, Toward Belief Function-based Cooperative Sensing for Interference Resistant Industrial Wireless Sensor Networks, *IEEE Transactions on Industrial Informatics*, Vol. 12, No. 6, pp. 2115-2126, December, 2016.

[10] W. Zhang, Z. Zhang, D. Qi, Y. Liu, Automatic Crack Detection and Classification Method for Subway Tunnel Safety Monitoring, *Sensors*, Vol. 14, No. 10, pp. 19307-19328, October, 2014.

[11] M. Zhou, *Control System Design of Agricultural Picking Robot Based on Binocular Vision*, Master's Thesis, Qingdao Technological University, Qingdao, China, 2018.

[12] T. Y. Chuang, H. W. Ting, J. J. Jaw, Dense Stereo Matching with Edge-constrained Penalty Tuning, *IEEE Geoscience and Remote Sensing Letters*, Vol. 15, No. 5, pp. 664-668, May, 2018.

[13] S. Kim, D. Min, S. Kim, K. Sohn, Feature Augmentation for Learning Confidence Measure in Stereo Matching, *IEEE Transactions on Image Processing*, Vol. 26, No. 12, pp. 6019-6033, December, 2017.

[14] E. Rosten, R. Porter, T. Drummond, Faster and Better: A Machine Learning Approach to Corner Detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 32, No. 1, pp. 105-119, January, 2010.

[15] R. Palekar, S. U. Parab, D. P. Parikh, V. N. Kamble, Real Time License Plate Detection Using OpenCV and Tesseract, *2017 International Conference on Communication and Signal Processing (ICCSP)*, Chennai, India, 2017, pp. 2111-2115.

[16] X. Zhang, Z. Y. Zhu, H. F. Zhang, Present Situation and Development of Matching Technique in Binocular Stereo Vision, *Metrology & Measurement Technology*, Vol. 37, No. 4, pp. 4-8, April, 2017. DOI: 10.11823/j.issn.1674-5795.2017.04.02

[17] H. Li, Q. Z. Lin, Q. J. Liu, An Automatic Registration Method of Remote Sensing Imagery Based on FAST Corner and SURF Descriptor, *Remote sensing for land & resources*, Vol. 24, No. 2, pp. 28-33, June, 2012.

[18] Z. H. Xu, H. F. Ling, F. H. Zou, Z. D. Lu, P. Li, A Novel Image Copy Detection Scheme Based on the Local Multi-Resolution Histogram Descriptor, *Multimedia Tools and Applications*, Vol. 52, No. 2-3, pp. 445-463, April, 2011.

[19] H. Bay, A. Ess, T. Tuytelaars, L. V. Gool, Speeded-Up Robust Features (SURF), *Computer Vision and Image Understanding*, Vol. 110, No. 3, pp. 346-359, June, 2008.

[20] J. X. Zhou, M. Y. Xu, J. Q. Liu, Image Matching Based on Improved SIFT Algorithm, *Industrial Control Computer*, Vol. 32, No. 5, pp. 97-98, May, 2019.

[21] J. Y. Li, Y. L. Wang, Y. J. Wang, Visual Tracking and Learning Using Speeded up Robust Features, *Pattern Recognition Letters*, Vol. 33, No. 16, pp. 2094-2101, December, 2012.

[22] Y. D. Zeng, H. D. Dai, M. H. Zheng, S. J. Su, Z. X. Wu, X. K. Xia, Z. R. Lin, Q. X. Wu, A 3D Passive Optical Localization System Based on Binocular Infrared Cameras, *2016 IEEE International Conference on Information and Automation (ICIA)*, Ningbo, China, 2016, pp. 368-373.

[23] B. Yang, X. L. Qiu, W. F. Hu, H. L. Guo, C. L. Song, Exposing Copy-Move Forgery Based on Improved SIFT Descriptor, *Journal of Internet Technology*, Vol. 18, No. 2, pp. 417-425, March, 2017.

[24] C. S. Yuan, Z. H. Xia, X. M. Sun, Coverless Image Steganography Based on SIFT and BOF, *Journal of Internet Technology*, Vol. 18, No. 2, pp. 435-442, March, 2017.

[25] S. B. Yang, Y. H. Long, Z. Y. Xiang, J. C. Yao, Research on Binocular Vision Stereo Matching Based on SURF Algorithm, *Journal of Hunan University of Technology*, Vol. 33, No. 3, pp. 75-80, May, 2019.

[26] Y. Zhao, J. G. Jiang, R. H. Hong, An Optimized SIFT Matching Based on RANSAC, *Opto-Electronic Engineering*, Vol. 41, No. 8, pp. 58-65, August, 2014.

[27] S. A. Bakar, M. S. Hitam, W. N. J. H. W. Yussof, Content-Based Image Retrieval Using SIFT for Binary and Greyscale Images, *2013 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, Melaka, Malaysia, 2013, pp. 83-88.

[28] J. Liu, F. Meng, F. C. Mu, Y. Zhang, An Improved Image Retrieval Method Based on SIFT Algorithm and Saliency Map, *2014 11th International Conference on Fuzzy Systems and Knowledge Discovery*, Xiamen, China, 2014, pp. 766-770.

[29] K. Lin, H. F. Yang, J. H. Hsiao, C. S. Chen, Deep Learning of Binary Hash Codes for Fast Image Retrieval, *2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Boston, MA, USA, 2015, pp. 27-35.

[30] F. Zhao, Y. Huang, L. Wang, T. Tan, Deep Semantic Ranking Based Hashing for Multi-label Image Retrieval, *2015 IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, USA, 2015, pp. 1556-1564.

[31] Y. B. Zhang, W. Song, Image Retrieval Method Based on Local Diagonal Color Difference Descriptor, *Computer Engineering*, Vol. 42, No. 12, pp. 248-253, December, 2016.

[32] H. Jiang, T. Zhang, J. P. Wachs, B. S. Duerstock, Enhanced Control of a Wheelchair-mounted Robotic Manipulator Using 3-D Vision and Multimodal Interaction, *Computer Vision and Image Understanding*, Vol. 149, pp. 21-31, August, 2016.

[33] S. Z. Zhao, F. Kang, J. J. Li, Displacement Monitoring for Slope Stability Evaluation Based on Binocular Vision Systems, *Optik*, Vol. 171, pp. 658-671, October, 2018.

[34] G. Huo, Z. Wu, J. Li, S. Li, Underwater Target Detection and 3D Reconstruction System Based on Binocular Vision, *Sensors*, Vol. 18, No. 10, Article No. 3570, October, 2018.

[35] L. Yang, B. Q. Wang, R. H. Zhang, H. Zhou, R. Wang, Analysis on Location Accuracy for the Binocular Stereo Vision System, *IEEE Photonics Journal*, Vol. 10, No. 1, Article No. 7800316, February, 2018.

[36] J. Cui, J. Huo, M. Yang, Novel Method of Calibration with Restrictive Constraints for Stereo-vision System, *Journal of Modern Optics*, Vol. 63, No. 9, pp. 835-846, September, 2016.

[37] Z. Jia, J. Yang, W. Liu, F. Wang, Y. Liu, L. Wang, C. Fan, K. Zhao, Improved Camera Calibration Method Based on Perpendicularity Compensation for Binocular Stereo Vision Measurement System, *Optics Express*, Vol. 23, No. 12, pp. 15205-15223, June, 2015.

[38] S. E. Shih, W. H. Tsai, Optimal Design and Placement of Omni-cameras in Binocular Vision Systems for Accurate 3-D Data Measurement, *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 23, No. 11, pp. 1911-1926, November, 2013.

[39] P. Guo, H. Du, Study on the Robotic Binocular Distance Measurement, *Wireless Internet Technology*, Vol. 15, No. 5, pp. 99-101, March, 2018.

[40] H. L. Zhang, T. Chen, P. Q. Zhuang, Z. S. Zhou, Z. Song, G. H. Jiao, Y. Qiao, A Binocular Stereo Vision System for Underwater 3D Measurement, *Journal of Integration technology*, Vol. 7, No. 3, pp. 1-14, May, 2018.

[41] Z. Liu, C. Q. Xiang, T. Chen, Automated Binocular Vision Measurement of Food Dimensions and Volume for Dietary Evaluation, *Computing in Science & Engineering*, May, 2018. DOI: 10.1109/MCSE.2018.243113429. https://ieeexplore-ieee-org.ezproxy.lib.ryerson.ca/document/8365083.

[42] X. J. Xu, X. N. Zhang, Crack Detection of Reinforced Concrete Bridge Using Video Image, *Journal of Central South University*, Vol. 20, No. 9, pp. 2605-2613, September, 2013.

[43] E. Cui, Y. Wang, T. Zhang, N. Di, Y. Yin, P. Wu, H. Sun, Easy Conductive Calibration Method for Binocular Vision System Based on Collinear Image Transformation, *Optical Engineering*, Vol. 56, No. 10, pp. 103106-103106, October, 2017.

[44] B. Tippetts, D. J. Lee, K. Lillywhite, J. Archibald, Review of Stereo Vision Algorithms and Their Suitability for Resource-Limited Systems, *Journal of Real-Time Image Processing*, Vol. 11, No. 1, pp. 5-25, January, 2016.

[45] Y. H. Lin, S. M. Wang, L. C. Huang, M. C. Fang, Applying the Stereo-vision Detection Technique to the Development of Underwater Inspection Task with PSO-based Dynamic Routing Algorithm for Autonomous Underwater Vehicles, *Ocean Engineering*, Vol. 139, pp. 127-139, July, 2017.

[46] F. Liao, W. Q. Ye, P. C. Wang, J. J. Wu, K. K. Xu, Image Mosaic Algorithm Based on SIFT Feature Matching, *Journal of Hunan University of Technology*, Vol. 28, No. 1, pp. 71-75, January, 2014.

[47] H. Zhang, J. Han, Y. Guan, Y. Zhang, A SIFT Algorithm Based on DOG Operator, *2018 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS)*, Xiamen, China, 2018, pp. 609-612.

[48] T. T. Ma, W. H. Ye, H. Huang, Y. X. Guo, Research on Improved SIFT Algorithm in Vision-based Target Recognition, *Machine Building & Automation*, Vol. 48, No. 2, pp. 188-191, February, 2019. DOI:10.19344 /j.cnki.issn1671-5276.2019.02.048

[49] Y. R. Feng, Application of an Improved SIFT Algorithm in Image Mosaic Algorithm, *Computer Products and Circulation*, No. 6, pp. 86-90, June, 2019.

[50] K. Q. Ren, R. Zhang, Edge Detection Based on Logarithmic Domain Gradient and Improved Sobel Operator, *Chinese Journal of Liquid Crystals and Displays*, Vol. 34, No. 3, pp. 283-290, March, 2019.

[51] Y. P. Chi, C. Liu, An Improved SIFT Algorithm for SAR Image Registration, *Journal of University of Chinese Academy of Sciences*, Vol. 36, No. 2, pp. 259-266, March, 2019.

[52] N. D. Hoang, Q. L. Nguyen, Metaheuristic Optimized Edge Detection for Recognition of Concrete Wall Cracks: A Comparative Study on the Performances of Roberts, Prewitt, Canny, and Sobel Algorithms, *Advances in Civil Engineering*, Vol. 2018, Article ID 7163580, November, 2018.

[53] K. Zhang, Y. M. Zhang, P. Wang, Y. Tian, J. Yang, An Improved Sobel Edge Algorithm and FPGA Implementation, *Procedia Computer Science*, Vol. 131, pp. 243-248, January, 2018.

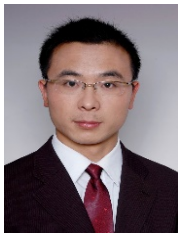
[54] G. M. Sun, C. Y. Wang, B. C. Ma, X. M. Wang, An Improved SIFT Algorithm for Infringement Retrieval, *Multimedia Tools and Applications*, Vol. 77, No. 12, pp. 14745-14765, June, 2018.

[55] S. Zhu, R. Gao, Z. Li, Stereo Matching Algorithm with Guided Filter and Modified Dynamic Programming, *Multimedia Tools and Applications*, Vol. 76, No. 1, pp. 199-216. January, 2017.

[56] B. Lu, Y. Liu, L. Su, Error Analysis of Binocular Stereo Vision System Applied in Small Scale Measurement, *Acta Photonica Sinica*, Vol. 44, No. 10, pp. 1011001-1-1011001-6, October, 2015.

[57] Q. Wang, Y. Yin, W. Zou, D. Xu, Measurement Error Analysis of Binocular Stereo Vision: Effective Guidelines for Bionic Eyes, *IET Science, Measurement & Technology*, Vol. 11, No. 7, pp. 829-838, October, 2017.

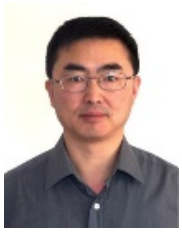
Biographies



Yinggang Xie received the B.Sc. degree in Automatic Control of Engineering from University of Science and Technology Beijing, Beijing, China, in 2001, and the M.E. and Ph.D. degrees in Control theory and control engineering from the University of Science and Technology Beijing, Beijing, China, in 2003 and 2007 respectively. He is currently a professor at the Department of Internet of things, Beijing Information Science and Technology University, China, His current research interests include multiple working modes control design for modular and reconfigurable robots, collaborative robots, Internet of Things. From 2007 to 2018, He has authored or co-authored more than 40 papers in international journals and conference proceedings, and published 2 books, and applied for 15 software copyrights. he has presided over and participated in

more than 10 national-level provincial and ministerial-level projects.

Jingyu Xing received the B.Sc. degree in Bachelor of Engineering from the LiaoNing University, LiaoNing, China, in 2017. She is currently a Master's degree in Beijing Information Science and Technology University, Beijing, China, Her current research interests include real-time reconstruction of unstructured scenes and multiple working modes control design for modular, robotic arm control, target recognition.



Guangjun Liu received the B.E. degree in automatic control from the University of Science and Technology of China, Hefei, China, in 1984; the M.E. degree in intelligent control from the Chinese Academy of Sciences, Shenyang Institute of Automation, Shenyang, China, in 1987; and the Ph.D. degree in mechanical engineering from the University of Toronto, Toronto, ON, Canada, in 1996. He is currently a Professor and the Canada Research Chair in Control Systems and Robotics with the Department of Aerospace Engineering, Ryerson University, Toronto. He has authored or co-authored more than 200 papers in international journals and conference proceedings. His current research interests include control systems and robotics, particularly, in modular and reconfigurable robots, mobile manipulators, and aircraft systems. Dr. Liu is a former Technical Editor of the IEEE/ASME TRANSACTIONS ON MECHATRONICS and a Licensed Member of the Professional Engineers of Ontario.

Jiangyu Lan received the B.Sc. degree in Information & Communication Engineering from the Beijing Information Science and Technology University, Beijing, China, in 2017. He is currently a Master's degree in Beijing Information Science and Technology University, Beijing, China, His current research interests include robot vision and deep learning, reconfigurable robots, robotic arm control, target recognition.

Yaoyao Dong received the B.Sc. degree in Information & Communication Engineering from the Beijing Information Science and Technology University, Beijing, China, in 2017. She is currently a Master's degree in Beijing Information Science and Technology University, Beijing, China, her current research interests include robot vision and reconfigurable robots, binocular vision depth.