

A positioning method for AUV based on guiding light source

Kai Wang^{1,2,a}, Wenyuan Ren^{3,b}, Yaoquan Duan^{1,2,c}, Hongwei Du^{1,2,d}

¹The 713 Research Institute of CSSC ,China

²Henan Key Laboratory of Underwater Intelligent Equipment, China

³Zhengzhou University, China

awangkai24@buaa.edu.cn, bcanye24@qq.com, cyaoquanduan@163.com,

^dhongweidu713@163.com

Abstract. This paper proposes a method for determining the position of underwater vehicles using light vision guidance. In order to solve the problem of limited perception distance and unclear underwater images, the illumination device is installed on the recovery platform as a light source to enhance the feature information of the recovery platform. The collected underwater light source images are enhanced by the dark channel prior defogging algorithm, and then the YOLOv5 algorithm is used to recognize the targets. Through experiments, it is verified that the image enhancement algorithm and target recognition algorithm are effective. Using the recognition results, the positioning accuracy of the binocular camera ranging algorithm is verified, and the results indicate that the proposed method can accurately determine the distance of the AUV relative to the recovery platform.

Keywords: AUV recycling, image enhancement, underwater vision, target detection.

1. Introduction

Autonomous Underwater Vehicle(AUV) can complete a variety of underwater tasks, such as marine scientific research and inspection of underwater equipment^[1]. When the energy is about to run out or the task is completed, the AUV needs to be recycled. Therefore, it is particularly necessary to determine the position distance information of the AUV relative to the recycling platform^[2].

The underwater docking platforms for AUVs mainly include platform-style docking and horn-style docking^[3]. For platform-style docking, the AUV needs to be moved above the docking platform, and after adjusting its horizontal pose with a close-range guidance device, it vertically descends to a certain position on the platform. It is then captured and locked by a capture mechanism or locking mechanism on the platform^[4]. Horn-style docking is similar to platform-style docking, but it moves to the centerline of the recovery platform, and the recovery locking is performed by approaching from the centerline.

This paper will design an AUV position determination method based on light vision guidance for platform based recycling. A visual system for AUV recycling based on guiding lights and binocular cameras is established, and image enhancement and target recognition algorithms under low-light conditions are discussed. By using the binocular vision positioning principle to calculate the position relationship of the camera relative to the light source, the position relationship of the AUV relative to the recycling platform is obtained, and the effectiveness of this method is verified through experiments.

2. Image acquisition

2.1 Binocular camera

This paper selects the ZF-IPC-02B11 deepwater network camera from Weihai Zhifan Marine Equipment Technology Co., Ltd. This camera can achieve 1080P high-definition imaging with a resolution of 1920*1080 and has a waterproof depth of 200 meters. Combine two monocular cameras to form a binocular camera, as shown in Fig. 1.



Fig. 1 Binocular camera

2.2 Guiding light source

When using a binocular camera for distance measurement, only a pair of feature points is needed to obtain the position information of the target. However, when designing a guiding light source for the recycling platform, in order to enhance the reliability of the system, it is necessary to consider the possibility of one camera being blocked during the actual recycling process. According to the P3P principle, when calculating the three-dimensional information of a single camera, at least three known feature points are needed and one verification point is needed to constrain the calculation results. Therefore, four known-location guiding light sources were designed.



Fig. 2 Relative position of guiding light source and binocular camera during pool test

3. image recognition

3.1 image enhancement

The clarity of images captured directly by underwater cameras is low, which affects subsequent recognition and positioning. It is necessary to enhance the images. Trucco et al.^[5] proposed a self-correcting image restoration filter algorithm based on the Jaffe-McGlamery model, which was verified on real-world images collected, proving that the method is suitable for underwater image restoration with finite backscattering. He Kaiming^[6] applied the Jaffe model to foggy image formation and proposed an image defogging algorithm based on the dark channel prior theory. This paper will enhance underwater images based on the theory of the dark channel prior, resulting in underwater light source images with stronger features.

The blind channel prior theory uses the values of the blind pixels in the image directly, Regarding fog and Estimate the transmission information of light, and take the value of the minimum color

channel of a certain pixel point as the dark pixel of the image. When there is no fog, the dark primary color tends to be 0, which can be expressed by the formula:

$$E^{dark}(x) = \min_{y \in \Omega(x)} \left(\min_{c \in \{r, g, b\}} E^c(y) \right) \rightarrow 0 \quad (1)$$

In the formula, c represents a channel among the R, G, B channels of an image, and it is a local region centered at pixel point x on the image and can be moved.

Due to the fact that the imaging principle of underwater images is quite similar to that of images with fog in outdoor environments [7], this paper applies the principle of dark primary color prior to enhance underwater degradation images.

The key to enhancing underwater images with the prior knowledge of dark primary colors lies in the calculation of water transparency and the solution of background light. As shown in Fig. 3, the image is enhanced for underwater photography, and it can be seen that the de-fogging algorithm based on the dark channel has a significant enhancement effect on underwater images. As shown in Fig. 4.



(a) Original image

(b) Enhanced image

Fig. 3 Original image and enhanced image

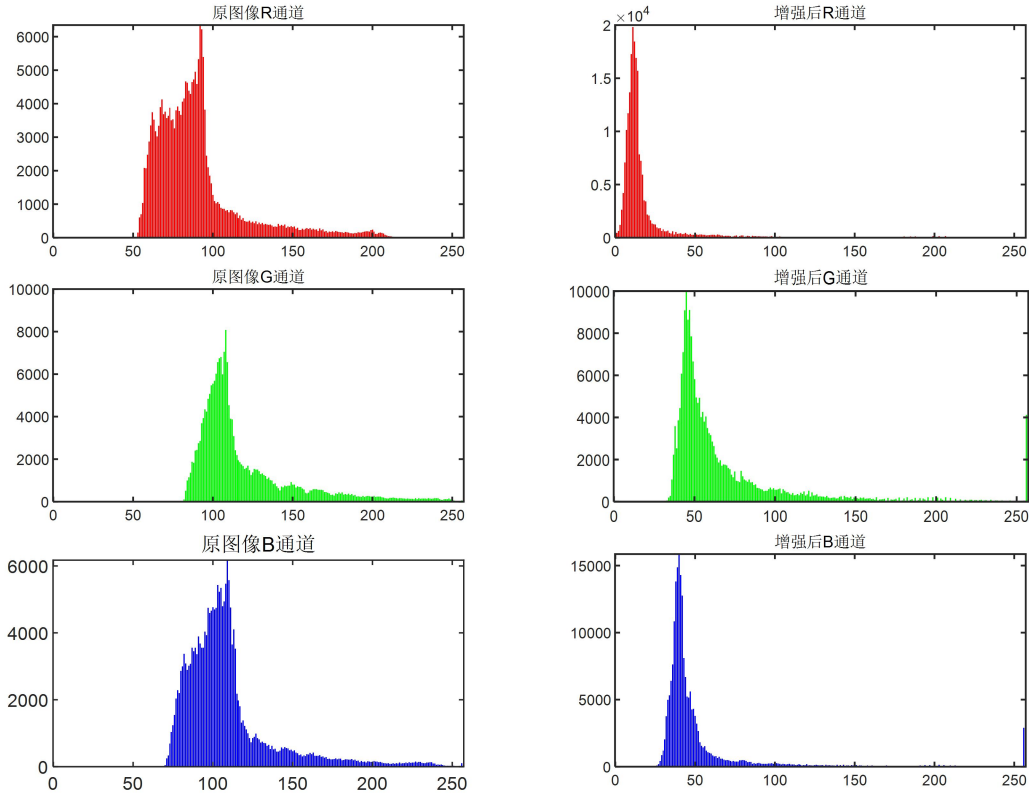


Fig. 4 Original image rgb Channel Histogram and Enhanced Image rgb Channel Histogram

3.2 image recognition

3.2.1 YOLOv5 introduction

In order to calculate the specific position of the camera relative to the guiding light source, it is necessary to accurately obtain the pixel coordinates of the guiding light source in the image. That is, the guiding light source in the image needs to be accurately identified.

Bochkovskiy et al. released yolov4 in 2020^[8], which divides the common detector into four parts. Firstly, the image is transformed at the input end. Then, features are extracted from the image in the backbone segment. Subsequently, methods such as FPN, which is effective for small object detection, are added to the neck. Finally, detection is performed at the head.

Yolov5 has added some new improvement ideas on the basis of Yolov4, which has greatly improved its speed and accuracy^[9].

This paper uses the Yolov5s network structure in the Yolov5 algorithm to train the neural network. The network structure mainly includes four parts: Input, Backbone, Neck, and Prediction. The Input end processes the input image and determines the size of the prediction box. The Backbone is a cross-stage local fusion network that mainly extracts features from the image. The Neck transfers and fuses information to obtain more feature information. The Prediction processes the target detection results.

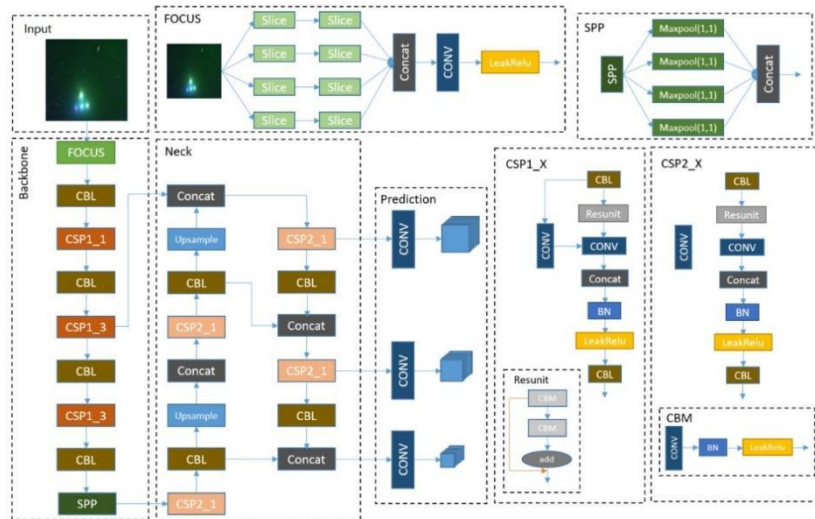


Fig. 5 Network Structure of Yolov5s

3.2.2 Dataset

The objective of this paper is to identify underwater feature sources, and based on the application scenario of AUV recovery, a dataset of underwater sources is generated according to this requirement. Through experiments under different angles, distances, turbidity levels of water quality, changes in illumination, and with different cameras, a total of 4610 images of varying resolutions were acquired. Among them, 3688 images were used as the training set, and 922 images were used as the test set, with a ratio of 4:1.

A dataset consisting of different forms of lights is manually labeled by labeling. There are four labeling categories, namely light, lamp, half_light, and half_lamp. Light represents each individual characteristic light source; lamp represents the overall underwater platform light; half_light represents only seeing part of the characteristic light source; half_lamp represents only seeing part of the underwater platform light. The labeled information is saved in the PASCAL VOC dataset format, which contains the category information of the samples and the location information of the labeled samples in the image. The lights of different shapes are labeled as follows:

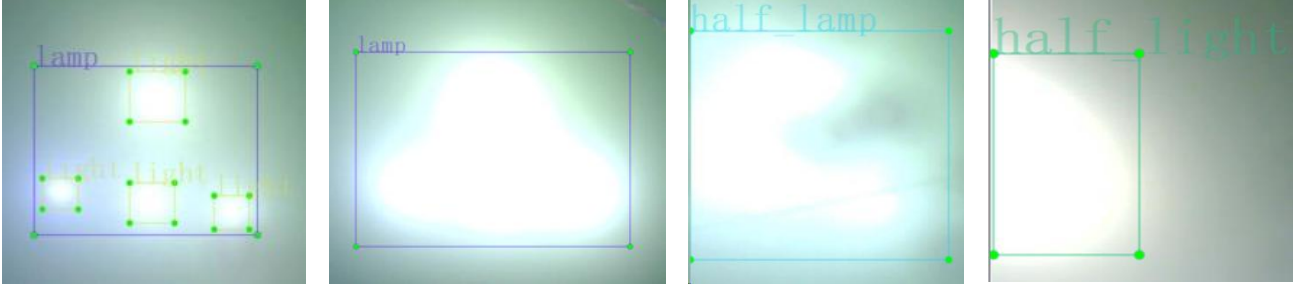


Fig. 6 Annotated image of the dataset

3.2.3 Training process

The experimental environment uses the Ubuntu 18.04 operating system, and the specific environment configuration is shown in Table 1

Table 1. Configuration of experimental environment	
parameter	allocation
GPU	RTX 2070
Memory	8G
System environment	ubuntu18.04
language	Python; PyTorch1.7.0
Accelerated environment	CUDA 11.6

3.2.4 Target detection results

To validate the effectiveness of the image enhancement algorithm on target detection in this paper, we trained 500 rounds using datasets that were not subjected to image enhancement and datasets that were subjected to image enhancement, respectively. We then compared their training metrics and used the trained weight parameters to detect the feature sources, obtaining the location of the source area.

Display the recognition results of light sources under different conditions:

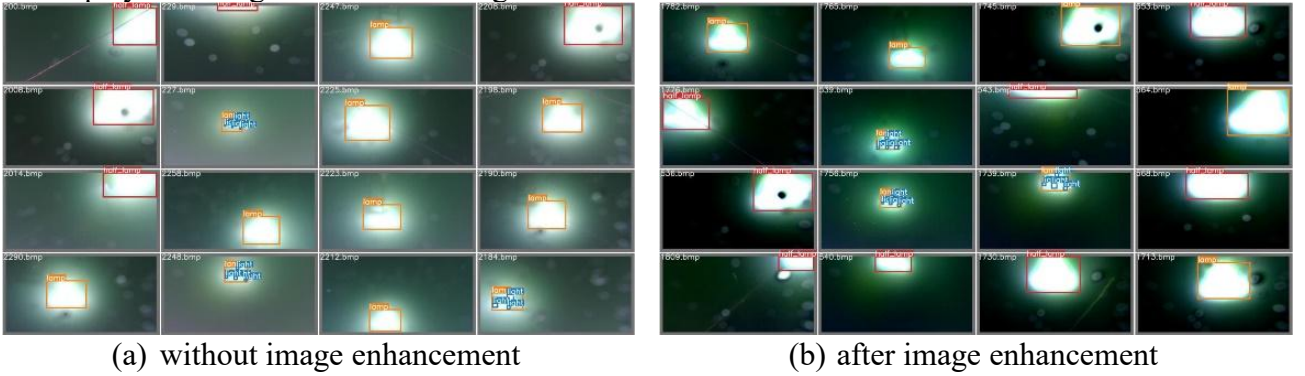


Fig. 7 The results detected by different test sets

As shown in Fig. 7, the detection results of the image without enhancement and the detection results of the image after enhancement are presented. It can be seen that the overall location detection rate has achieved good results.

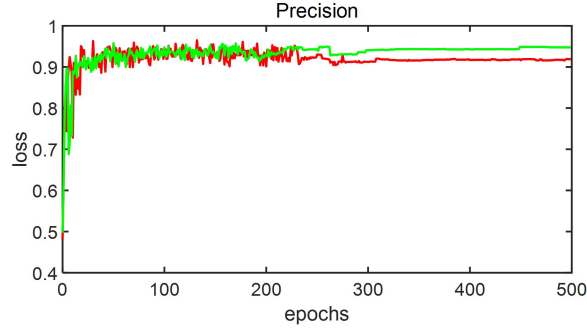


Fig. 8 Precision

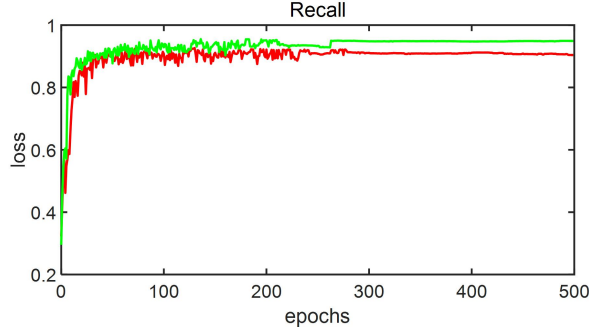


Fig. 9 Recall

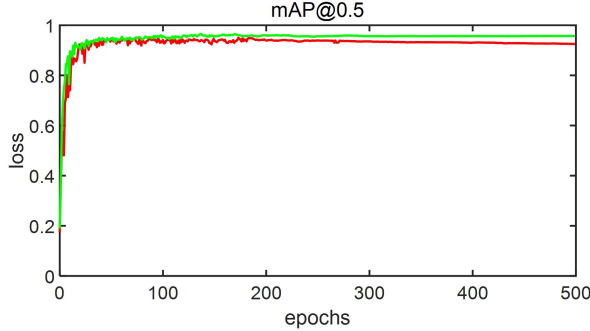


Fig. 10 mAP

Fig. 8, Fig. 9, and Fig. 10 show the values of precision, recall, and average precision for the images before and after image enhancement processing, respectively. The red line represents the training results without image enhancement processing, while the green line indicates the training results after image enhancement processing. After training for approximately 300 rounds, the detection results tend to stabilize. This indicates that more accurate results can be achieved when training the network on images that have been enhanced for image processing.

Table 2. Average accuracy of target recognition based on YOLOv5

name	light	lamp	half light	half lamp
Before enhancement	0.939	0.967	0.864	0.931
After enhancement	0.946	0.983	0.939	0.961

From Table 2, it can be seen that compared with the original image detection, the average detection accuracy of each category has been improved, indicating that the image enhancement algorithm used in this paper has a higher recognition rate for target detection relative to the original image. Therefore, it can be seen that the quality of the image is very important for target detection and recognition.

4. Binocular positioning

When measuring the distance to the target with a binocular camera, based on the binocular principle of visual presentation. In this paper, when detecting the light source, the binocular camera captures images of the feature light source at a relatively far position. Sometimes, the light source cannot be completely present in the image. In this case, only the center point of the entire light source detection area is used for calculation. When the distance is closer, the position of each feature light source can be obtained. By matching each pair of points to calculate the three-dimensional coordinates, the average value of the coordinates is taken as the distance of the target center relative to the camera.

In the laboratory pool, the feature light source was photographed at different distances, and the actual values were compared with the calculated values. Table 3 presents the results of double-eye ranging with the center point of the entire feature light source taken as the feature point at distances ranging from 1 to 9 meters. The pixel points of the target center are basically guaranteed to be on the same straight line, with a maximum pixel difference of 9 pixels, and the average pixel error is guaranteed to be within 3 pixels, with an average error of 4.9182%.

Table 3. Comparison of camera positioning accuracy results

Number	Left Pixel Point	Right Pixel Point	Positioning distance/mm	True distance/mm	error/%
1	(1093, 526)	(908, 523)	1048.0297	1000	4.8030
2	(987, 554)	(881, 550)	1874.5358	2000	6.2732
3	(941, 537)	(878, 534)	3165.7507	3000	5.5250
4	(961, 537)	(912, 535)	4073.5383	4000	1.8385
5	(853, 525)	(814, 524)	5101.4141	5000	2.0283
6	(946, 567)	(914, 566)	6124.1206	6000	2.0687
7	(859, 524)	(832, 522)	7446.8691	7000	6.3838
8	(931, 518)	(909, 516)	8758.9238	8000	9.4865
9	(918, 561)	(897, 569)	8527.1143	9000	5.8568

When the information of the light source can be fully observed, match and calculate the relative distance for each characteristic light source, and take the average value obtained as the final positioning distance. Table 4 presents the positioning results obtained from the average distance, with an average error of 2.7486%, which is lower than that of directly extracting the center of the light source, and the maximum error is 5.0423%.

Table 4. Relative error between underwater platform distance and camera measurement

Number	Positioning distance/mm	True distance/mm	error/%
1	1050.4227	1000	5.0423
2	1915.4426	2000	4.2279
3	3089.1517	3000	2.9717
4	4097.7963	4000	2.4449
5	5103.9460	5000	2.0781
6	6039.6071	6000	0.6601
7	7262.7780	7000	3.7540
8	8279.2993	8000	3.4912
9	9006.0651	9000	0.0674

Compared to calculating the average distance for each feature source, directly calculating the distance from the center of the source has a larger error, but the absolute deviation is still within the acceptable range, meeting the requirements for use.

Therefore, in the actual process of ranging, when the position of each feature source cannot be extracted, one can choose to calculate the distance by treating the entire feature source as a feature point.

5. summary

This paper focuses on the situation encountered by AUV during the recovery process, and designs a positioning scheme based on underwater target recognition using binoculars. A new method of applying image defogging algorithms based on dark channels to underwater applications is proposed. The target detection algorithm based on YOLOv5 is used to recognize feature-source images enhanced by images. The position of the camera relative to the source is calculated using binocular vision positioning principles, which in turn determines the position of the AUV relative to the recovery platform. The experiment was conducted in a pool to verify the expected results.

Acknowledgements

This research was funded by Henan Province Central-led Local Science and Technology Development Fund Project (Z20221343002); Henan Key Research and Development Project (231111212100); Henan Key Research and Development Project (231111221400).

References

- [1] KRONEN D M. Docking the ocean explorer autonomous underwater vehicle using a low-cost acoustic positioning system and a fuzzy logic guidance algorithm[M]. Florida Atlantic University, 1997.
- [2] Pan Guang, Du Xiaoxu. Autonomous Underwater Vehicle Recycling Technology[M]. Science Press, 2021.
- [3] Yuan Xueqing, Diao Jiayu, Li Weimin, et al. Development and Application Status of AUV Underwater Docking [J]. Journal of Ship Science and Technology, 2023, 45(5):1-8.
- [4] T. Kawasaki et al., "'Marine Bird", a new experimental AUV - results of docking and electric power supply tests in sea trials," Oceans '04 MTS/IEEE Techno-Ocean '04 (IEEE Cat. No.04CH37600), Kobe, 2004, pp. 1738-1744 Vol.3.
- [5] Trucco E , Olmos-Antillon A T . Self-Tuning Underwater Image Restoration[J]. IEEE Journal of Oceanic Engineering, 2006, 31:511-519.
- [6] He K , Jian S , Fellow, et al. Single Image Haze Removal Using Dark Channel Prior[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2011, 33(12):2341-2353.
- [7] Lin Sen, Bai Ying, Li Wentao, et al. Underwater image restoration based on the modified model and the prior information of the dark channel [J]. Robot, 2020, 42(4):10.Doi:10.13973/j.cnki.robot.190464.
- [8] Bochkovskiy A , Wang C Y , Liao H Y M .YOLOv4: Optimal Speed and Accuracy of Object Detection[J]. 2020.Doi:10.48550/arXiv.2004.10934.
- [9] Ha Jianming, Research on Turning Vehicle Flow Detection Based on Deep Learning Algorithm [D]. Xi'an: Chang'an University 2022: 20-26.