

# Structured Light-Based Underwater Collision-Free Navigation and Dense Mapping System for Refined Exploration in Unknown Dark Environments

Yaming Ou<sup>ID</sup>, Junfeng Fan<sup>ID</sup>, Member, IEEE, Chao Zhou<sup>ID</sup>, Member, IEEE, Song Kang, Zhuoliang Zhang<sup>ID</sup>, Zeng-Guang Hou<sup>ID</sup>, Fellow, IEEE, and Min Tan<sup>ID</sup>

**Abstract**—Underwater collision-free navigation and dense reconstruction are essential for marine refined exploration. However, existing passive vision-based methods are difficult to apply in low-light and weak-feature underwater environments. In this article, a more adaptable three-dimensional (3-D) dense mapping robotic system based on self-designed scanning binocular structured light (BSL), named ROV-Scanner, is developed to address this challenge. First, the measurement principle based on the refraction model ensures its high accuracy. Second, an underwater 3-D dense mapping algorithm fusing the Doppler velocity log (DVL), inertial measurement unit (IMU), and pressure sensor multifrequency information is proposed to realize dense mapping during robot motion. Then, an air–water two-stage extrinsic calibration algorithm is proposed. In particular, the extrinsic parameters between DVL and camera are innovatively calibrated using BSL, enhancing robustness. Furthermore, for the first time, a framework of BSL-based collision-free navigation is presented to guarantee the safe movement of the system in unknown environments. Experimental results show that our system can simultaneously achieve autonomous collision-free navigation and dense mapping in dark underwater environments, which has great potential for application in marine refined exploration.

**Index Terms**—Refined exploration, structured light vision, underwater collision-free navigation, underwater dense mapping.

Manuscript received 28 December 2023; accepted 19 February 2024. This work was supported in part by the National Natural Science Foundation of China under Grant 62373354 and Grant 62003342; in part by the Beijing Natural Science Foundation under Grant 4232057; in part by the Youth Innovation Promotion Association of CAS under Grant 2022130; in part by the Young Elite Scientists Sponsorship Program of China Association of Science and Technology under Grant YESS20210344; and in part by the Fund of Key Laboratory of Artificial Intelligence, Ministry of Education under Grant AI2022010. This article was recommended by Associate Editor J. Wang. (*Corresponding authors:* Junfeng Fan; Chao Zhou.)

Yaming Ou is with the Laboratory of Cognition and Decision Intelligence for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China, and also with the School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: ouyaming2021@ia.ac.cn).

Junfeng Fan, Chao Zhou, Song Kang, Zhuoliang Zhang, and Min Tan are with the Laboratory of Cognition and Decision Intelligence for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China (e-mail: junfeng.fan@ia.ac.cn; chao.zhou@ia.ac.cn; song.kang@ia.ac.cn; zhuoliang.zhang@ia.ac.cn; min.tan@ia.ac.cn).

Zeng-Guang Hou is with the State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China (e-mail: zengguang.hou@ia.ac.cn).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TSMC.2024.3370917>.

Digital Object Identifier 10.1109/TSMC.2024.3370917

## I. INTRODUCTION

WITH the continuous development of human marine engineering, autonomous underwater robots [1] are playing an increasingly important role, including underwater garbage grabbing [2], structure overhauling [3], and underwater exploration [4]. Among them, for the field of marine refined exploration, collision-free navigation and dense mapping for robots are of great importance [5]. The autonomous navigation ability of a robot will be directly related to its intelligence, which is one of the key foundations for autonomous exploration [6]. It has been a research difficulty to ensure that robots can safely navigate to the target area by avoiding obstacles in unknown underwater environments. In addition, the creation of dense maps is essential for environmental monitoring, including the observation of biological populations [7]. It is indispensable in the refined exploration process. However, the acquisition of dense point clouds has always been challenging for underwater environments. Therefore, the motivation of this article is to simultaneously realize collision-free navigation and dense mapping for robots in unknown environments refined exploration.

Currently, passive vision devices represented by monocular and binocular cameras are widely used in underwater robot collision-free navigation due to the low cost and rich perception information. Lv et al. [8] proposed a collision-free planning and control framework based on binocular, which consists of obstacle avoidance planning, arctangent nonsingularity terminal sliding mode control, and fuzzy inference. The robot's collision-free navigation in dynamic environments was finally realized. Furthermore, to achieve reliable mapless underwater navigation, a low-cost end-to-end navigation system based on a monocular camera and a fixed single-beam echo sounder was introduced by Yang et al. [9]. In addition, Manderson et al. [10] implemented obstacle avoidance and target selection for a six-legged underwater vehicle by processing data acquired by a monocular camera through a convolutional neural network. They found that using real-time visual feedback to make proximity navigation decisions in specific environments, such as coral reef formations, is unrivaled by other sensors. It can be seen that the above work has achieved promising results in the field of robot collision-free navigation based on passive vision. However, as for the underwater environment, due to the absorption of light by the water and the scattering of particles in the

water, the underwater environment is usually characterized by weak illumination. Arain et al. [11] pointed out that, due to visibility, the camera may incorrectly recognize free space and obstacles in the environment, thus giving an incorrect response to the trajectory planner with disastrous consequences. Then, they proposed combining feature-based stereo matching with learning-based segmentation to produce a more robust obstacle map. However, due to the scarcity of underwater features, the maps produced were very sparse and difficult to use for refined exploration. Although Hernández et al. [7] has achieved the dense reconstruction of underwater vehicles in navigation based on monocular depth estimation using a semi-global matching-like multi-image method, it was similarly unable to be used in dark conditions. Therefore, collision-free navigation of robots and dense mapping for refined exploration in unknown dark environments deserve further in-depth research.

Compared to passive vision, underwater structured light systems based on the principle of active vision have been shown to be more suitable in underwater three-dimensional (3-D) perception [12]. Using the high penetration of laser in water, the system is less affected by the water [13]. Palomer et al. [14] developed a structured light scanning system which is possible to produce full 3-D perception by using a mirror-galvanometer to steer a laser line projector and using triangulation with a camera. Ou et al. [15] designed a binocular structured light (BSL) 3-D reconstruction system for low-light underwater environments. The system proved to be more adaptable underwater in different turbidity levels compared to binocular camera. Therefore, it is reasonable to believe that underwater structured light system will bring new inspiration to underwater simultaneous collision-free navigation and dense mapping.

However, most of the current studies fix the structured light system on static platforms and use it to realize 3-D reconstruction of small objects [16]. Only a few works have assembled structured light systems for underwater robots [17]. In order to build a map for a larger scene, Palomer et al. [18] carried a structured light scanner on an autonomous underwater vehicle to achieve the inspection of underwater structures. By fusing Doppler velocity log (DVL), inertial navigation system (INS), and pressure sensor, they realized to build a map of a  $1.4 \times 1.4 \times 1.2$  m artificial water pipe. Bodenmann et al. [19] designed a mapping system consists of a camera, a line laser, LED panels, and a PC in a pressure tight housing, named SeaXerocks 1, which can be mounted on an underwater robot. Combined with the navigation data provided by the robot, they achieved partial mapping of the seafloor at a depth of more than 2000 m in the Pacific Ocean. Hitchcox and Forbes [20] fused a commercially available underwater laser scanner named Insight Pro with the DVL-INS to realize a 3-D scan of an underwater shipwreck in Wiarton, ON, Canada. Using laser-based loop closure, they improved self-consistency in underwater mapping. It can be seen that the few existing works based on structured light systems have achieved better results in underwater mobile dense mapping. However, robots still rely on traditional navigation during map construction, or manual remote control, without utilizing the wealth of

information gained from structured light, which is a great loss in the exploration of unknown dark environments. Moreover, the stitching of single-line point clouds during structured light mapping relies entirely on the results of low-speed DVL-INS dead reckoning (DR), which lead to serious loss of some point cloud information. A better way to merge needs to be utilized.

It is worth mentioning that DVL, as a commonly used underwater speed sensor, is widely used with cameras for robot navigation and mapping. Since the DVL can only be used underwater, the extrinsic parameters between the two must be calibrated underwater, which is made difficult by the presence of water. Therefore, most systems use physical measurements, which will undoubtedly introduce unknown errors. There are few references to calibration in the literature. By posting AprilTag fiducials underwater, camera-DVL calibration was implemented by Westman and Kaess [21]. Xu et al. [22] pointed that this method can reach high calibration accuracy, but is not applicable for open sea. Then, they achieved the calibration between the camera and DVL using feature points extracted from the environment. However, the refraction effect on the camera imaging was not considered. What is worse is that purely visual feature point extraction in some open scene is very difficult, such as turbid water. Considering that BSL is little influenced by water, it is desired to use its information to achieve the extrinsic calibration of DVL in this article.

In conclusion, we note that the following issues still exist for practical applications in the field of underwater refined exploration that are worth studying.

- 1) *Collision-Free Navigation:* a) Rely on passive vision and tend to fail under dark water and b) weak feature areas are difficult to achieve robust obstacle detection.
- 2) *Dense Mapping:* a) Rely on DR data for map stitching, but DR data tends to be slower than point cloud acquisition, with significant loss of information and b) dense map information is not utilized in the navigation framework to enhance robustness.
- 3) *DVL-Camera Extrinsic Calibration:* a) Rely on extraction of corner points of the checkerboard grid, which tends to fail in dark and turbid underwater environments and b) the underwater refraction effect is not considered.

In order to solve the above problems, we utilize the structured light system to realize the underwater collision-free navigation and dense mapping for refined exploration in unknown dark environments. The main contributions are as follows.

- 1) A more adaptable 3-D dense mapping robotic system based on self-designed scanning BSL, named ROV-Scanner, is developed for refined exploration, where the on-board design allows for autonomous mobility and operational capabilities. A more efficient underwater 3-D dense mapping algorithm fusing DVL, inertial measurement unit (IMU) and pressure sensor multifrequency information is proposed to realize dense mapping during robot motion.
- 2) An air–water two-stage underwater multisensor calibration method is presented. In particular, the extrinsic parameters between DVL and camera are innovatively calibrated using BSL based on graph optimization, enhancing robustness.

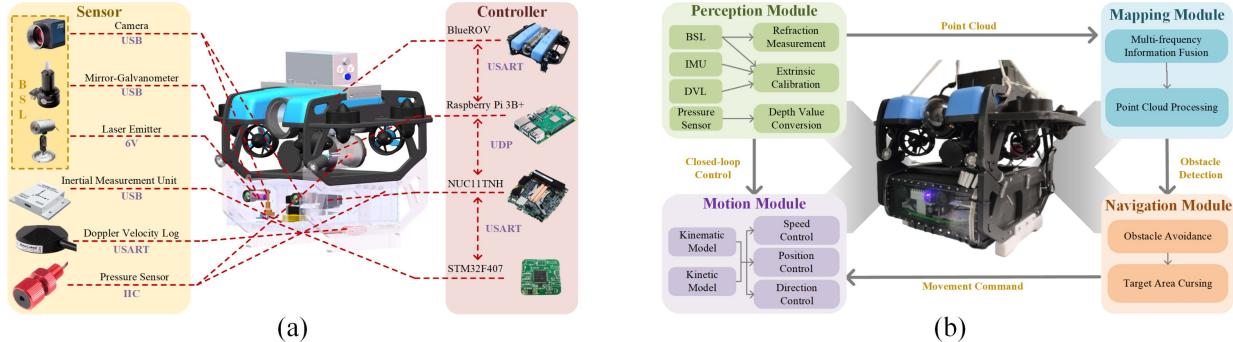


Fig. 1. Architecture of ROV-scanner. (a) 3-D model and key components. (b) Physical prototype and software pipeline.

- 3) A framework of BSL-based collision-free navigation is presented to guarantee the safe movement of the system in unknown dark environments. To the best of our knowledge, this is the first work that can simultaneously realize autonomous collision-free navigation and dense mapping in dark underwater environments by utilizing active structured light vision.

## II. OVERVIEW OF ROV-SCANNER SYSTEM

### A. Hardware Design

As shown in Fig. 1(a), the system mainly consists of BSL system, IMU, DVL, BlueROV, pressure sensor, and onboard controllers. Among them, the BSL contains a binocular camera, a blue line laser emitter, and a mirror-galvanometer. Through the reflection of the mirror, it enables static scanning of the scene. The IMU can provide the linear acceleration and angular velocity information at 400 Hz, and the DVL can provide speed information at 12 Hz. BlueROV is a commercial underwater robot that we use as the motion module of the system. The pressure sensor can measure the absolute water pressure at the location of the robot in 60 Hz, which has a full-projectile relationship with the absolute depth. So it can be used to get the robot depth information. In addition, the controllers involved in the system are STM32F407, NUC11TNH, Raspberry Pi 3B+. The NUC11TNH is the main controller and is responsible for all the algorithmic processing. The STM32F407 and Raspberry Pi 3B+ are responsible for the scanning control of the mirror-galvanometer and the underlying motion control of the system, respectively. Detailed parameters of some hardware are shown in Table I.

### B. Software Pipeline

The software pipeline of ROV-Scanner includes four modules: 1) perception module; 2) motion module; 3) mapping module; and 4) navigation module, as shown in Fig. 1(b). In perception module, binocular laser refraction measurement model is conducted to obtain accurate 3-D data (Section IV-A). Pressure data is converted into depth information by Pascal's Law. In addition, an innovative calibration method is proposed to calibrate the extrinsic parameters between BSL, IMU, and DVL (Section III). The motion module is mainly responsible for the execution of system movement command. In the experiment, the sensor information from the perception module is

TABLE I  
SYSTEM HARDWARE PARAMETERS

Hardware	Parameters	Value
IMU	Frequency	400 Hz
	Communication Interface	USB
DVL	Frequency	12 Hz
	Communication Interface	USART
	Resolution	1280×1024 Pixels
Camera	Frames	100 Fps
	Focal Length	6 mm
	Communication Interface	USB 3.0
Laser Emitter	Wavelength	450 nm
	Power Consumption	400 mW
	Fan Angle	35°
Mirror-Galvanometer	Oscillating Angle Range	±20°
	Control Signal	Sine Wave
Structured Light	Frequency	40 Hz
	No. of 3D points	320
Pressure Sensor	Frequency	60 Hz
	Resolving Power	0.2 mbar

used to realize the closed-loop control of the system to ensure the stable motion. Furthermore, a mobile dense mapping algorithm based on multifrequency information fusion is proposed in the mapping module (Section IV-B). By processing the dense point cloud, the scene obstacles can be perceived in real time. We then use this information to implement collision-free navigation in the navigation module (Section VI). Meanwhile, to achieve efficient underwater exploration, the fusion results in the perception module are utilized to enable the system to cruise the target area. All algorithms are deployed on the onboard NUC11TNH with Docker + Ubuntu 20.04 + Robot Operation System.

## III. AIR-WATER TWO-STAGE EXTRINSIC CALIBRATION

*Notations:* Some necessary notations are introduced in this section.  $\mathcal{W}$  denotes the world coordinate system.  $\mathcal{L}$  denotes the left camera coordinate system.  $\mathcal{R}$  denotes the right camera coordinate system.  $\mathcal{I}$  denotes the IMU coordinate system and  $\mathcal{D}$  denotes the DVL coordinate system. The point cloud obtained from a single computation of BSL is called a Scan. It is worth mentioning that the coordinate system of Scan is  $\mathcal{S}$ , which is consistent with  $\mathcal{L}$  and is not distinguished below. What is more,  $p^A$  denotes the representation of  $p$  under the coordinate system  $\mathcal{A}$ .  $T_A^B$  denotes the transformation matrix from the coordinate system  $\mathcal{A}$  to the coordinate system  $\mathcal{B}$ .

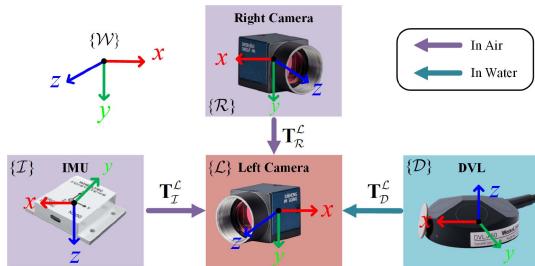


Fig. 2. Schematic of extrinsic calibration.

In order to realize the fusion of information from different sensors, the extrinsic parameters between the binocular camera, IMU and DVL need to be calibrated in ROV-Scanner, as shown in Fig. 2. In general, the binocular camera and IMU are encapsulated in a waterproof chamber with fixed relative positions and can work in air. DVL must be used in water. Existing calibration methods based on passive vision are difficult to apply in dark underwater environments. So we hope to utilize active vision to achieve calibration. In order to avoid the difficulties posed by the water environment for calibration problems as far as possible, the air–water two-stage extrinsic calibration algorithm is proposed: first, the transformation matrix  $\mathbf{T}_L^R$  and  $\mathbf{T}_R^L$  is calibrated in air using visual and inertial information, and then the transformation matrix  $\mathbf{T}_D^L$  between left camera and DVL is calibrated underwater using point cloud from BSL and DR data from DVL, which is more robust. Finally, the transformation matrix between IMU and DVL can be obtained indirectly by  $\mathbf{T}_D^I = \mathbf{T}_L^I \mathbf{T}_D^L$ .

#### A. Air Calibration Between Binocular and IMU

In the air, methods for extrinsic calibration between binocular camera and IMU are well established. There are many open source works that are capable of performing this task. First, the OpenCV library is used to calibrate the extrinsic parameters between the right and left camera. Then, the most commonly available calibration algorithm Kalibr is introduced to get the transformation matrix between IMU and left camera. See [23] for more details.

#### B. Underwater Calibration Between Left Camera and DVL

Given that structured light is more robust than passive vision underwater, we innovatively propose for the first time to utilize the point cloud information obtained from our BSL to realize the extrinsic calibration of DVL. Specifically, utilizing structured light system we can obtain a local point cloud map of the system at two moments during its motion, and the relative state estimation of the system can be obtained through the point cloud alignment registration generalized-ICP (GICP) [24]. In this process, the DR of DVL also realizes the state estimation of the system. As a result, we can get the laser relative state estimation and acoustic relative state estimation corresponding to each other. Subsequently, we construct the factor graph shown in Fig. 3 to realize the optimization of the  $\mathbf{T}_D^L$ .

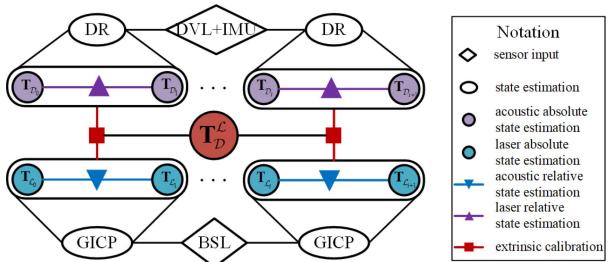


Fig. 3. Factor graph of extrinsic calibration between left camera and DVL.

Suppose the set of matched BSL and DVL transformation matrices  $\{\mathbf{T}_{\mathcal{L}_1}^{\mathcal{L}_0}, \mathbf{T}_{\mathcal{L}_2}^{\mathcal{L}_1}, \dots, \mathbf{T}_{\mathcal{L}_m}^{\mathcal{L}_{m-1}}\}$  and  $\{\mathbf{T}_{\mathcal{D}_1}^{\mathcal{D}_0}, \mathbf{T}_{\mathcal{D}_2}^{\mathcal{D}_1}, \dots, \mathbf{T}_{\mathcal{D}_m}^{\mathcal{D}_{m-1}}\}$  have been obtained. The representation of BSL information in the  $\mathcal{D}$  coordinate system can be obtained using the coordinate transformation, and the transformation relation is as follows:

$$\hat{\mathbf{T}}_{\mathcal{D}_i}^{\mathcal{D}_{i-1}} = (\mathbf{T}_{\mathcal{D}}^L)^{-1} \mathbf{T}_{\mathcal{L}_i}^{\mathcal{L}_{i-1}} \mathbf{T}_{\mathcal{D}}^L. \quad (1)$$

Denote  $\mathbf{T}_{\mathcal{D}}^L$  as

$$\mathbf{T}_{\mathcal{D}}^L = \begin{bmatrix} \mathbf{R}_{\mathcal{D}}^L & \mathbf{p}_{\mathcal{D}}^L \\ \mathbf{0} & 1 \end{bmatrix}. \quad (2)$$

Then, the residual vector  $\mathbf{r}_{\mathcal{D}_i}^{\mathcal{D}_{i-1}}$  of inertial and acoustic information can be expressed as

$$\mathbf{r}_{\mathcal{D}_i}^{\mathcal{D}_{i-1}} = \left[ \log(\mathbf{R}_{\mathcal{D}_i}^{\mathcal{D}_{i-1}} (\hat{\mathbf{R}}_{\mathcal{D}_i}^{\mathcal{D}_{i-1}})^T)^\vee, \mathbf{p}_{\mathcal{D}_i}^{\mathcal{D}_{i-1}} - \hat{\mathbf{p}}_{\mathcal{D}_i}^{\mathcal{D}_{i-1}} \right] \quad (3)$$

where  $\log(\cdot)^\vee$  denotes the map from the 3-D rotation group  $\text{SO}(3)$  to the Lie-Algebra  $\mathfrak{so}(3)$ . Therefore, the process of extrinsic calibration between left camera and DVL can be transformed into the optimization of the following function:

$$\arg \min_{\mathbf{T}_{\mathcal{D}}^L} \sum_m^{i=1} \left\| \mathbf{r}_{\mathcal{D}_i}^{\mathcal{D}_{i-1}} \right\|^2 \sum_i^{\mathcal{D}} \quad (4)$$

where  $\sum_i^{\mathcal{D}}$  denotes the information matrix of the residuals.

## IV. UNDERWATER MOBILE 3-D DENSE MAPPING

#### A. Refraction Measurement Model of BSL

For the BSL, the camera is mounted in a waterproof cabin. When imaging, light passes through water, glass, and air in turn, and then reaches the camera optical center. Refraction effect occurs when it crosses the multilayer medium, as shown in Fig. 4(a). It is pointed out in [25] that the conventional triangulation model does not work in the presence of refraction. Combined with our existing work [26], an underwater binocular laser measurement model based on the refraction model is proposed to realize accurate measurement of our ROV-Scanner system, as shown in Fig. 4(b). Where,  $\mathbf{I}_i$  is the intersection of the medium of the rays, and  $\mathbf{r}_i$  is the direction vector of the rays.  $i = 0, 1, 2$  denote air, glass, and water media, respectively.

In the experiment, the laser centerline is first extracted using the grayscale center of gravity method, and then the calculation

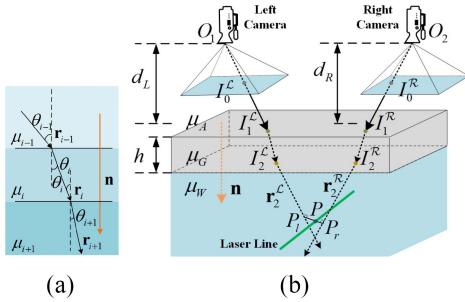


Fig. 4. Underwater refraction model. (a) Flat refractive geometry with multilayer. (b) Binocular laser measurement model of ROV-Scanner system.

of 3-D coordinates begins. The unit vector  $\mathbf{r}_0$  can be obtained from the imaging model of the camera in air as follows:

$$\mathbf{r}_0 = \frac{\overrightarrow{OI_0}}{|\overrightarrow{OI_0}|}, \overrightarrow{OI_0} = \mathbf{K}^{-1} [x \ y \ 1]^T \quad (5)$$

where  $O$  is the optical center of the camera,  $I_0$  is the intersection of the ray with the imaging plane,  $\mathbf{K}$  is the intrinsic camera matrix, and  $(x, y)$  are the pixel coordinates. Using the plane of refraction (POR) constraint [25], we get

$$(\gamma \mathbf{r}_i + \eta \mathbf{r}_{i+1}) \cdot (\mathbf{r}_i \times \mathbf{n}) = 0 \quad i = 0, 1 \quad (6)$$

where  $\gamma$  and  $\eta$  nonzero real numbers. So if  $\mathbf{r}_i$  is known,  $\mathbf{r}_{i+1}$  can be calculated as

$$\begin{cases} \mathbf{r}_{i+1} = \alpha_i \mathbf{r}_i + \beta_i \mathbf{n} & i \in [0, 1] \\ \alpha_i = \frac{\mu_i}{\mu_{i+1}} \\ \beta_i = \sqrt{1 - \left(\frac{\mu_i}{\mu_{i+1}}\right)^2 [1 - (\mathbf{r}_i \cdot \mathbf{n})^2]} - \frac{\mu_i}{\mu_{i+1}} \mathbf{r}_i \cdot \mathbf{n} \end{cases} \quad (7)$$

where  $\mathbf{n}$  is the refractive surface normal vector, and  $\mu_i$  is the refractive index. Both of them can be obtained by measurement or calibration and are treated as known values. So the light propagation direction vectors  $\mathbf{r}_1$  and  $\mathbf{r}_2$  can be calculated by  $\mathbf{r}_0$ , which can be obtained by (5). Furthermore, from the geometric model it is obtained that

$$\begin{cases} \mathbf{I}_1 = \frac{d}{\mathbf{r}_0 \cdot \mathbf{n}} \mathbf{r}_0 \\ \mathbf{I}_2 = \mathbf{I}_1 + \frac{h}{\mathbf{r}_1 \cdot \mathbf{n}} \mathbf{r}_1 \end{cases} \quad (8)$$

where  $d$  and  $h$  are the distance from the optical center of the camera to the glass, and the thickness of the glass, respectively. So the refraction points  $\mathbf{I}_1$  and  $\mathbf{I}_2$  can be calculated. Combining (5)–(8), we can deduce  $\mathbf{I}_2^L$ ,  $\mathbf{I}_2^R$ ,  $\mathbf{r}_2^L$ , and  $\mathbf{r}_2^R$ . Due to the existence of systematic errors, the left and right camera imaging optical paths may not intersect in 3-D space. Suppose the straight line  $l$  is the common vertical line, which intersects  $\mathbf{P}_l$  and  $\mathbf{P}_r$ , respectively. Then,  $\mathbf{P}_l$  and  $\mathbf{P}_r$  coordinates are

$$\begin{cases} \mathbf{P}_l = \mathbf{I}_2^L + \mathbf{r}_2^L \frac{(\mathbf{r}_2^L \cdot \mathbf{r}_2^R)(\mathbf{r}_2^R \cdot \mathbf{i}_2) - (\mathbf{r}_2^L \cdot \mathbf{i}_2)(\mathbf{r}_2^R \cdot \mathbf{r}_2^R)}{(\mathbf{r}_2^L \cdot \mathbf{r}_2^L)(\mathbf{r}_2^R \cdot \mathbf{r}_2^R) - (\mathbf{r}_2^L \cdot \mathbf{r}_2^R)(\mathbf{r}_2^R \cdot \mathbf{r}_2^L)} \\ \mathbf{P}_r = \mathbf{I}_2^R + \mathbf{r}_2^R \frac{(\mathbf{r}_2^R \cdot \mathbf{i}_2)(\mathbf{r}_2^L \cdot \mathbf{r}_2^L) - (\mathbf{r}_2^R \cdot \mathbf{r}_2^L)(\mathbf{r}_2^L \cdot \mathbf{i}_2)}{(\mathbf{r}_2^L \cdot \mathbf{r}_2^L)(\mathbf{r}_2^R \cdot \mathbf{r}_2^R) - (\mathbf{r}_2^L \cdot \mathbf{r}_2^R)(\mathbf{r}_2^R \cdot \mathbf{r}_2^L)}. \end{cases} \quad (9)$$

In which,  $\mathbf{i}_2 = \mathbf{I}_2^L - \mathbf{I}_2^R$ . Taking the midpoint of  $\mathbf{P}_l \mathbf{P}_r$  as the desired target point  $\mathbf{P}$ , the coordinates of the point  $\mathbf{P}$  can be easily obtained by  $\mathbf{P} = (\mathbf{P}_l + \mathbf{P}_r)/2$ . So far, we can calculate the 3-D information of the underwater scene and dense reconstruction with can be realized by our ROV-Scanner.

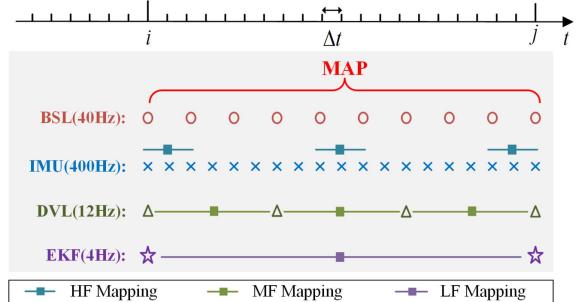


Fig. 5. Different frequency for each sensor in ROV-scanner.

## B. Mobile Dense Mapping Based on Multifrequency Information Fusion

The dense 3-D information of the scene can be acquired using our ROV-Scanner system. It can acquire a Scan point cloud at approximately 40 Hz. A map is defined as the set containing  $m$  Scans  $\{s_0, s_1, \dots, s_m\}$ . The task of mapping is to unify multiple Scans within a map under the coordinate system  $S_0$  of the first Scan. Existing structured light-based underwater dense mapping methods mainly rely on DR. However, the acquisition rate of DR is much lower than the Scan point cloud, resulting in a large amount of data information loss. So we introduce IMU inertial information at high frequencies as well as fully utilize the DVL velocity information at medium frequencies to achieve a denser mapping. The schematic is shown in Fig. 5. It is mainly divided into three mapping periods: 1) high-frequency (HF) inertial information mapping; 2) mid-frequency (MF) acoustic information mapping; and 3) low-frequency (LF) hybrid information mapping from EKF.

1) *HF Mapping*: IMU can provide object linear acceleration information  $\tilde{\mathbf{a}}^T$  and angular velocity information  $\tilde{\omega}^T$  at a HF of 400 Hz, which is much faster than the Scan frequency. So the set  $\{\tilde{\omega}_0^T, \tilde{\mathbf{a}}_0^T, \tilde{\omega}_1^T, \tilde{\mathbf{a}}_1^T, \dots, \tilde{\omega}_k^T, \tilde{\mathbf{a}}_k^T\}$  of inertial information can be obtained during the two scan  $s_{i-1}$  and  $s_i$  generation time intervals. The following is IMU discrete pose integration model in [27]:

$$\begin{aligned} \mathbf{R}_{\mathcal{I}_{i+1}}^{\mathcal{I}_i} &= \text{Exp}\left((\tilde{\omega}_i^T - \mathbf{b}_{gi}^W - \eta_{gi}^W)\delta t\right) \\ \mathbf{v}_{\mathcal{I}_{i+1}}^{\mathcal{I}_i} &= \mathbf{g}^W \delta t + \mathbf{R}_{\mathcal{I}_i}^{\mathcal{I}_i} (\tilde{\mathbf{a}}_i^T - \mathbf{b}_{ai}^W - \eta_{ai}^W) \delta t \\ \mathbf{p}_{\mathcal{I}_{i+1}}^{\mathcal{I}_i} &= \mathbf{v}^{\mathcal{I}_i} \delta t + \frac{1}{2} \left( \mathbf{g}^W \delta t^2 + \mathbf{R}_{\mathcal{I}_i}^{\mathcal{I}_i} (\tilde{\mathbf{a}}_i^T - \mathbf{b}_{ai}^W - \eta_{ai}^W) \delta t^2 \right) \end{aligned} \quad (10)$$

where  $\tilde{\omega}_i$  denotes the angular velocity measurement,  $\mathbf{g}$  denotes the acceleration of gravity,  $\mathbf{b}_{gi}$  and  $\mathbf{b}_{ai}$  are the bias,  $\eta_{gi}$  and  $\eta_{ai}$  are the white noise, and Exp denotes the map from the Lie-Algebra  $\mathfrak{so}(3)$  to the 3-D rotation group  $\text{SO}(3)$ . So the transformation relation  $\mathbf{T}_{\mathcal{I}_{i-1}}^{\mathcal{I}_i}$  of IMU pose can be obtained during the two Scan point cloud acquisitions. In addition

$$\mathbf{T}_{\mathcal{I}_i}^{\mathcal{I}_0} = \mathbf{T}_{\mathcal{I}_1}^{\mathcal{I}_0} \mathbf{T}_{\mathcal{I}_2}^{\mathcal{I}_1} \dots \mathbf{T}_{\mathcal{I}_i}^{\mathcal{I}_{i-1}}. \quad (11)$$

Therefore, the transformation relation of IMU can be known when  $s_i$  and  $s_0$  are acquired. Using  $\mathbf{T}_{\mathcal{I}}^S$ ,  $s_i$  can be transformed to  $s_0$  coordinate system by

$$\mathbf{T}_{s_i}^{s_0} = \mathbf{T}_{\mathcal{I}}^S \mathbf{T}_{\mathcal{I}_i}^{\mathcal{I}_0} \left( \mathbf{T}_{\mathcal{I}}^S \right)^{-1}. \quad (12)$$

2) *MF Mapping*: DVL can achieve stable linear velocity measurements using the acoustic Doppler effect than IMU, but at a LF, usually around 12 Hz, which is lower than the scan frequency. Suppose the DVL velocity information  $v_1$  is obtained at  $s_{i+1}$ . At this time, instead of updating the position information using only inertial information, the velocity information of DVL is fused to update. Thus, the purpose of reducing the cumulative error of IMU is achieved. At this time, the transformation matrix  $T_{s_{i+1}}^{s_0}$  of  $s_{i+1}$  to  $s_0$  coordinate system  $\mathcal{S}_0$  is

$$\begin{bmatrix} \mathbf{R}_{\mathcal{I}}^{\mathcal{S}} \mathbf{R}_{\mathcal{I}_{i+1}}^{\mathcal{I}_0} \mathbf{R}_{\mathcal{I}}^{\mathcal{S}-1} & T_{\mathcal{D}}^{\mathcal{S}} \Xi(\mathbf{p}^{\mathcal{D}_{i+1}}) T_{\mathcal{D}}^{\mathcal{S}-1} \oplus T_{\mathcal{I}}^{\mathcal{S}} \Xi(\mathbf{p}^{\mathcal{I}_{i+1}}) T_{\mathcal{I}}^{\mathcal{S}-1} \\ \mathbf{0} & 1 \end{bmatrix} \quad (13)$$

where  $\Xi$  denotes the mapping of the 3-D vector form of the position information to the homogeneous transformation matrix representation,  $\oplus$  denotes the summation of the position parts of the transformation matrix.  $\mathbf{R}_{\mathcal{I}_{i+1}}^{\mathcal{I}_0}$  is consistent with that in the previous (11).  $\mathbf{p}^{\mathcal{D}_{i+1}}$  is the DVL velocity integral and  $\mathbf{p}^{\mathcal{I}_{i+1}}$  is the IMU acceleration quadratic integral.

3) *LF Mapping*: To achieve a deeper integration of inertial, acoustic and pressure information, the classical EKF framework is used, similar to [28]. Specifically, the pressure and acoustic information are used as the observed value and the inertial information is used as the predicted value. When the pressure or acoustic information arrives, the observation update is started. In practical use, a state update frequency of 4 Hz can be achieved. The transformation matrix obtained by EKF is considered to be more accurate than the previous two approaches. Then, the new matrix  $T_{s_i}^{s_0}$  obtained will be directly used as a new starting point for the first two updates, on which the HF mapping and MF mapping will continue until the arrival of the next new EKF fusion result.

## V. FRAMEWORK OF BINOCULAR STRUCTURED LIGHT-BASED COLLISION-FREE NAVIGATION

In order to realize the refined exploration in unknown underwater environments, real-time obstacle avoidance capability is essential for robotic system safety. As shown in Fig. 6, we propose the framework of BSL-based collision-free navigation in dark underwater environments. It consists of three main parts: 1) obstacle detection; 2) obstacle avoidance strategy; and 3) target area cruising.

### A. Obstacle Detection

In dark environments, traditional binocular-based obstacle detection methods are no longer applicable. Instead, by utilizing BSL, the dense 3-D information of objects in dark scenes can be easily obtained. In order to ensure real-time obstacle detection, we would like to utilize the local map obtained from the ROV-Scanner to realize the obstacle 3-D position evaluation and the best motion direction prediction, as shown in Fig. 6(a). In the experiment, the point cloud composed of about 40 Scans is regarded as a local map  $M$  and used for obstacle avoidance prediction. Namely

$$M = \{s_i | i = 1, 2, 3, \dots, 40\} \quad (14)$$

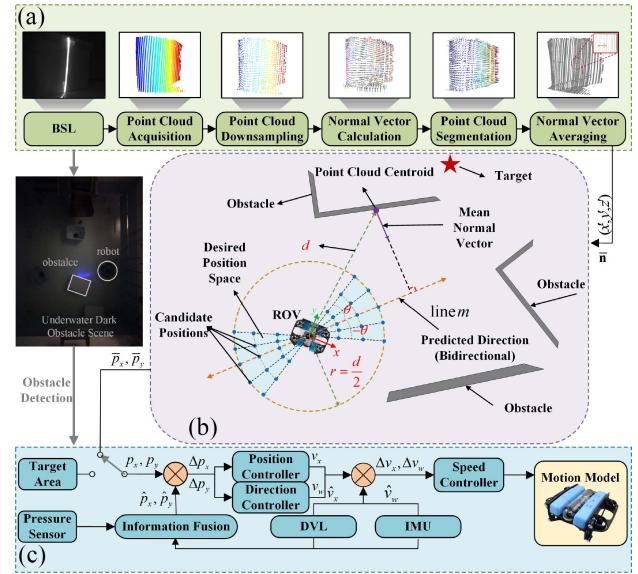


Fig. 6. Framework of BSL-based collision-free navigation in dark underwater environments. (a) Obstacle real-time detection based on binocular structured light. (b) Obstacle avoidance strategy. (c) Target area cruising based on multi-frequency information fusion.

where  $s_i$  denotes the point cloud obtained from a single computation of BSL. To speed up the computation, voxelized downsampling of the point cloud is performed. Assume that the downsampled point cloud  $Q$  is

$$Q = \{\mathbf{q}_i | \mathbf{q}_i \in \mathbb{R}^3, i = 1, 2, 3, \dots, n\}. \quad (15)$$

Let  $\mathbf{q}_i = (x_i, y_i, z_i)$ , then the center of gravity  $\mathbf{q}'$  of the point cloud  $Q$  is found using the following equation:

$$\mathbf{q}' = (x', y', z') = \left( \frac{\sum_{i=1}^n x_i}{n}, \frac{\sum_{i=1}^n y_i}{n}, \frac{\sum_{i=1}^n z_i}{n} \right). \quad (16)$$

Consider  $d = \sqrt{x'^2 + y'^2 + z'^2}$  as the distance from the system to the obstacle. If  $d < \xi$ , we assume that an obstacle is detected and continue the processing below. So  $\xi$  denotes the minimum distance to start obstacle avoidance, which is a positive real number.

In order to get the obstacle direction prediction, first we compute the normal vector for each point  $\mathbf{q}_i$  in the point cloud  $Q$ . Assuming that the set of nearest neighbors of  $\mathbf{q}_i$  is  $N_i$ , then

$$N_i = \{\mathbf{q}_j | \|\mathbf{q}_j - \mathbf{q}_i\|^2 < r, \mathbf{q}_j \in Q\} \quad (17)$$

where  $r$  is the nearest neighbor search radius. The normal vector  $\mathbf{n}_i$  can be computed using the least squares method, so

$$\mathbf{n}_i = \arg \min_n \sum_{q_j \in N_i} \|\mathbf{n} \cdot (\mathbf{q}_j - \mathbf{q}_i)\|^2. \quad (18)$$

The solution to this problem can be obtained by solving the following eigenvalue problem:

$$\lambda_{\min} \cdot \mathbf{n} = (\mathbf{C} - \lambda_{\min} \cdot \mathbf{I}) \cdot \mathbf{n} = 0 \quad (19)$$

where  $\mathbf{I}$  is the unit matrix,  $\lambda_{\min}$  is the minimum eigenvalue in the eigenvalue problem, and  $\mathbf{C}$  is the covariance matrix, defined as

$$\mathbf{C} = \frac{1}{|N_i|} \sum_{q_j \in N_i} (\mathbf{q}_j - \bar{\mathbf{q}}_i) \cdot (\mathbf{q}_j - \bar{\mathbf{q}}_i)^T. \quad (20)$$

In which,  $|N_i|$  is the size of the set of nearest points  $N_i$  and  $\bar{\mathbf{q}}_i$  is the mean value of all point  $\mathbf{q}_i$ . In order to overcome the effect of ground point cloud, normal vector-based point cloud segmentation is applied. Namely, if  $\arccos(\mathbf{n}_i(3)/|\mathbf{n}_i|) < \epsilon$  ( $\epsilon$  is a small positive constant), it is removed from the point cloud  $Q$ . In this article,  $\mathbf{n}(i)$  denotes the  $i$ th element of the vector  $\mathbf{n}$ . After that, the mean normal vector  $\hat{\mathbf{n}}$  of the remaining point cloud  $Q'$  is computed. Projecting it onto the  $xy$ -plane yields the vector  $\mathbf{n}' = (\hat{\mathbf{n}}(1), \hat{\mathbf{n}}(2)) \in \mathbb{R}^2$ . Then, normalize  $\mathbf{n}'$  to obtain the final obstacle direction prediction vector  $\bar{\mathbf{n}} = \mathbf{n}'/|\mathbf{n}'|$ , where  $|\mathbf{n}'|$  is the mold length of  $\mathbf{n}'$ .

### B. Obstacle Avoidance Strategy

For the underwater mapping system, on the one hand, it is hoped that the system can safely traverse the obstacles to reach the destination. On the other hand, a smaller change in the distance between the system and the obstacles is desirable, so that the system's mapping will be more stable. Therefore, the direction perpendicular to the normal vector  $\bar{\mathbf{n}}$  under the current position of the system is regarded as the desired prediction direction, and the linear expression is

$$\text{line } m : \bar{\mathbf{n}}(2)y + \bar{\mathbf{n}}(1)x = 0. \quad (21)$$

The range that has an angle of  $\theta$  with  $m$  and is no more than  $d/2$  away from the system is regarded as the desired position space  $R$ . Let  $(a, b)$  be any point in  $R$ , then  $R$  is

$$\begin{cases} R_1 = (a, b) \mid \sqrt{a^2 + b^2} \leq \frac{d}{2} \\ R_2 = (a, d) \mid |\bar{\mathbf{n}}(1)a + \bar{\mathbf{n}}(2)b| \leq \sin \theta \sqrt{a^2 + b^2} \\ R = R_1 \cap R_2, d > 0, \theta \in [0, \frac{\pi}{2}] \end{cases} \quad (22)$$

As shown in Fig. 6(b), the desired position space is a region of two diagonal sectors centered on the current position of the system. By discretizing it, a series of candidate positions  $P = \{\mathbf{p}_i \mid \mathbf{p}_i = (p_{xi}, p_{yi}) \in \mathbb{R}^2, i = 1, 2, \dots, n\}$  can be obtained. Suppose the set of discrete points of obstacles that have been explored in the current map within a certain range  $r$  from the current system is  $O = \{\mathbf{o}_k \mid \mathbf{o}_i = (p_{oi}, p_{oi}) \in \mathbb{R}^2, k = 1, 2, \dots, m\}$ , and the target point is  $\mathbf{t} = (x_t, y_t)$ . Then the best candidate position  $\mathbf{p}_d$  is selected by the following evaluation function:

$$\begin{aligned} \mathbf{p}_d &= (p_x, p_y) \\ &= \arg \min_{\mathbf{p}_i} \left( \alpha \cdot \sigma(\mathcal{G}(\mathbf{t}, \mathbf{p}_i)) - \beta \cdot \sigma \left( \sum_{k=1}^m \mathcal{Q}(\mathbf{o}_k, \mathbf{p}_i) \right) \right) \end{aligned} \quad (23)$$

where  $\alpha$  and  $\beta$  denote the weights, both non-negative real numbers. The functions  $\mathcal{G}(\mathbf{t}, \mathbf{p}_i)$  and  $\mathcal{Q}(\mathbf{o}_k, \mathbf{p}_i)$  denote the Euclidean distances of the point  $\mathbf{p}_i$  to  $\mathbf{t}$  and  $\mathbf{o}_k$ , respectively.  $\sigma$  is a normalization function that mitigates the nonsmoothing

---

### Algorithm 1 Underwater Real-Time Obstacle Avoidance Algorithm Based on Binocular Structure Light

---

**Input:** local map point cloud  $M$ , target point  $T$ , the set of history obstacle points  $O$ , constant  $\xi, r, \epsilon, \theta, \delta_\theta, \delta_r, \alpha, \beta$   
**Output:** desired candidate position  $\mathbf{p}_d$

- 1: Initialization
- 2: Downsample the point cloud  $M$  to get  $Q$  via voxel filtering
- 3: Calculate the center of gravity  $\mathbf{q}'(x', y', z')$  of  $Q$  (Eq. (16))
- 4: **if**  $d = \sqrt{x'^2 + y'^2 + z'^2} > \xi$  **then**
- 5:   **return** false
- 6: **end if**
- 7: **for**  $\mathbf{q}_i$  in  $Q$  **do**
- 8:   Find nearest neighbor  $N_i$  within  $r$  by KD tree (Eq. (17))
- 9:   Compute the normal vector  $\mathbf{n}_i$  (Eq. (18)-(20))
- 10: **end for**
- 11: **for**  $\mathbf{q}_i$  in  $Q$  **do**
- 12:   **if**  $\arccos(\mathbf{n}_i(z)/|\mathbf{n}_i|) > \epsilon$  **then**
- 13:      $Q'.append(\mathbf{p}_i)$
- 14:   **end if**
- 15: **end for**
- 16: Calculate average normal vector  $\mathbf{n}$  in  $Q'$
- 17: Obtain obstacle direction prediction vector  $\bar{\mathbf{n}}$
- 18: Calculate desired position space  $R$  by  $\theta$  and  $d$  (Eq. (21)-(22))
- 19: Discrete  $R$  by  $\delta_\theta$  and  $\delta_r$  to obtain candidate positions  $P$
- 20:  $cost_{min} = \text{INT\_MAX}$
- 21: **for**  $\mathbf{p}_j$  in  $P$  **do**
- 22:   Calculate evaluation cost  $cost_i$  of  $\mathbf{p}_i$  by  $\alpha$  and  $\beta$
- 23:   **if**  $cost_{min} > cost_i$  **then**
- 24:      $cost_{min} = cost_i, \mathbf{p}_d = \mathbf{p}_j$
- 25:   **end if**
- 26: **end for**
- 27: **return**  $\mathbf{p}_d$

---

effect due to discontinuous sensor information during localized obstacle avoidance. So

$$\begin{aligned} \sigma(\mathcal{G}(\mathbf{t}, \mathbf{p}_i)) &= \frac{\mathcal{G}(\mathbf{t}, \mathbf{p}_i)}{\sum_{i=1}^n \mathcal{G}(\mathbf{t}, \mathbf{p}_i)} \\ \sigma(\mathcal{Q}(\mathbf{o}_k, \mathbf{p}_i)) &= \frac{\mathcal{Q}(\mathbf{o}_k, \mathbf{p}_i)}{\sum_{i=1}^n \mathcal{Q}(\mathbf{o}_k, \mathbf{p}_i)}. \end{aligned} \quad (24)$$

From this, we can get the desired candidate position that the system expects in the next moment in order to avoid obstacles and mapping. The overall process of obstacle avoidance based on BSL is shown in Algorithm 1.

In subsequent practical experiments of Section VI, the constant values  $\xi, r, \epsilon, \theta, \delta_\theta, \delta_r, \alpha, \beta$  in Algorithm 1 are set to  $0.6, 0.05, (\pi/12), (\pi/3), (\pi/12), 0.1, 1, 1$ , respectively. It is worth stating that in the proposed algorithm the constant values are chosen in a wide range and do not require fine tuning.

### C. Target Area Cruising

As a mature commercial robot, BlueROV's underlying control is relatively sophisticated. Both [29] and [30] designed special controllers to ensure its base motion capability. As for

the unknown underwater environment, the traditional position acquisition based on global vision is not satisfied [8]. In the previous Section IV-B, the acoustic, inertial and pressure information are fused for robot state estimation. Therefore, in this article, a target area cruising mechanisms based on acoustic–inertial–pressure fusion is designed, which does not rely on global vision to realize the closed-loop control of the system position, as shown in Fig. 6(c). Our ROV-Scanner performs the task in real-time motion toward the target area based on the results of the LF fusion through the position controller and direction controller. In this process, once an obstacle is detected, it starts to execute the desired candidate position obtained from our obstacle avoidance strategy to realize obstacle avoidance as well as reconstruction of the obstacle. Meanwhile, in order to realize the motion stability of the system, the linear velocity provided by the DVL and the angular velocity provided by the IMU are used to realize the velocity closed-loop control for it. In subsequent practical experiments of Section VI, the classical PID controller is implemented.

## VI. EXPERIMENTS

In order to verify the refined exploration capability of the ROV-Scanner in dark underwater environment, an experimental scenario is first constructed (Section VI-A). Subsequently, the results of the system's extrinsic calibration are given, and the errors in the extrinsic parameters between DVL and left camera obtained by different methods are compared (Section VI-B). Then, Scanner is fixed inside a tank alone to statically reconstruct some regular and irregular objects (Section VI-C). After that, some underwater mobile mapping experiments are implemented (Section VI-D). Furthermore, simultaneous collision-free navigation and dense mapping experiments are carried out in a pool with obstacles (Section VI-E). Finally, the peculiarities of our system are discussed (Section VI-F).

### A. Experiment Environment

Refined exploration in dark underwater environments has always been a challenging dilemma. In order to simulate this scenario, the ROV-Scanner test experiments are conducted in a  $5 \times 4 \times 1.5$  m pool. As shown in Fig. 7, the dark underwater scene is simulated by adjusting the external ambient light. Plastic boxes are placed in the pool to simulate underwater obstacles. Our system is monitored by a global camera during motion. It is worth mentioning that the global camera is only used to record the robot's trajectory as the ground truth of XY-plane and is not involved in any internal control. The Z-direction ground truth of the trajectory is provided by a different pressure sensor.

### B. Extrinsic Calibration Results

For a multisensor system, accurate extrinsic calibration is one of the foundations of data fusion. In our ROV-Scanner, the extrinsic parameters to be calibrated include right camera to left camera  $\mathbf{T}_R^L$ , IMU to left camera  $\mathbf{T}_I^L$ , DVL to left camera  $\mathbf{T}_D^L$ . In order to avoid as much as possible the inconvenience

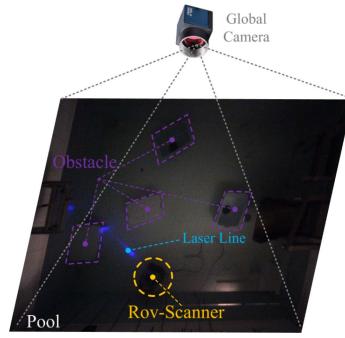


Fig. 7. Experimental pool with obstacles in darkness.

TABLE II  
EXTRINSIC CALIBRATION RESULTS

Extrinsic Parameter	Transformation Matrix			
$\mathbf{T}_R^L$	$\begin{bmatrix} 0.9999 & -0.0014 & -0.0153 & 0.1373 \\ 0.0014 & 1.0000 & -0.0011 & 0.0059 \\ 0.0153 & 0.0011 & 0.9999 & 0.0024 \\ 0 & 0 & 0 & 1 \end{bmatrix}$			
$\mathbf{T}_I^L$	$\begin{bmatrix} 0.9959 & -0.0885 & 0.0195 & 0.0312 \\ 0.0163 & -0.0364 & -0.9992 & 0.0530 \\ 0.0891 & 0.9954 & -0.0348 & -0.0479 \\ 0 & 0 & 0 & 1 \end{bmatrix}$			
$\mathbf{T}_D^L$	$\begin{bmatrix} -0.2092 & 0.9687 & -0.1338 & 0.1582 \\ -0.0672 & 0.1222 & 0.9902 & 0.0222 \\ 0.9756 & 0.2161 & -0.0396 & -0.3546 \\ 0 & 0 & 0 & 1 \end{bmatrix}$			

of the water environment to the calibration process, the air-water two-stage extrinsic calibration algorithm is applied. The calibration results are shown in Table II.

Among them, we innovatively propose the use of structured light to calibrate  $\mathbf{T}_D^L$ . Currently, the extrinsic parameters of DVL are generally based on physical measurement, and there are few methods to realize the calibration between it and the camera. In the experiments, we first use BSL to perform state estimation using the GICP algorithm in some structured scenes, and record the results of DR synchronously. Two sets of data are collected in total. Then, we perform graph-based optimization on one set of data to obtain the extrinsic calibration results. Subsequently, the extrinsic parameters obtained from the physical measurements and the extrinsic parameters obtained from the calibration are each applied to another set of data, de-transformed the DR data to the BSL coordinate system, and then compared the mean errors of the positions. The results are shown in Fig. 8. It can be clearly seen that after our calibration method, the average error has been significantly reduced, from 0.016 m in physical measurement to 0.007 m. Moreover, the standard deviation has also been significantly decreased, with more stable data.

### C. Underwater Static Dense Mapping Experiments

Prior to mobile mapping, in order to individually test the 3-D perception performance of the BSL, the Scanner is first fixed in a small  $1.5 \times 0.4 \times 0.4$  m tank, as shown in Fig. 9(a). Reconstruction experiments are conducted for irregular cup and regular walls, respectively, as shown in Fig. 9(b) and (c). Fig. 9(d) and (e) shows the results of the reconstructed point cloud visualization. It can be seen that the obtained point cloud

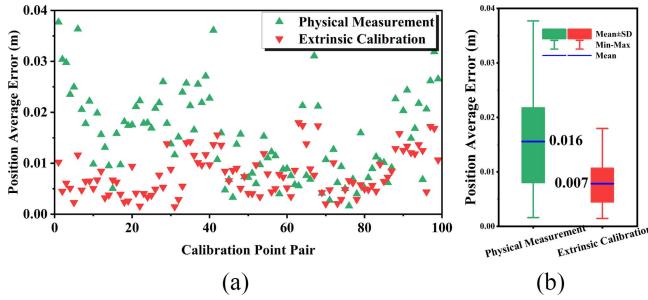


Fig. 8. Position errors between laser relative state estimation and the transformed acoustic relative state estimation using different extrinsic parameters. (a) One-pair error. (b) Average error.

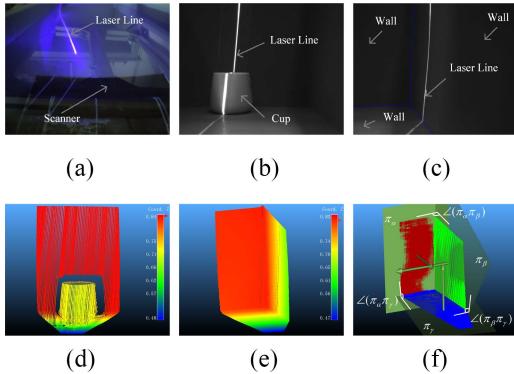


Fig. 9. Static reconstruction results. (a) Dark underwater environment in tank. (b) and (c) Results of underwater regular wall and irregular object camera shots, respectively. (d) and (e) Corresponding reconstruction results. (f) Point cloud segmentation and plane fitting.

TABLE III  
RESULTS OF PLANE FITTING

Plane	RMS(cm)	Normal
$\pi_\alpha$	0.104	(0.058;-0.022;-0.998)
$\pi_\beta$	0.137	(-0.983;-0.037;-0.178)
$\pi_\gamma$	0.074	(-0.003;-0.999;-0.004)

is pretty dense, which is rare in the underwater domain. In addition, the obtained regular wall point clouds are processed in order to quantitatively compare the reconstruction effects. First, the point clouds are segmented in plane using the point cloud segmentation technique. Then, the segmented point clouds are fitted to the plane separately, as shown in Fig. 9(f). The fitting results are shown in Table III. From the data in the table, the root mean square error (RMSE) of the fitting plane is small, reaching millimeter level. This indicates that the reconstruction accuracy of ROV-Scanner is high and has reached the level of the state-of-the-art (SOTA) underwater.

Then, we calculate the normal vectors from the fitting plane, as shown in Table IV. The values of the plane angle measured using physical instruments are used as ground truth to compare with the fitted plane angle values. It can be seen that the errors are quite low. The minimum error reaches 0.66% and the maximum is only 1.37%. This proves the ability of our ROV-Scanner system to reconstruct underwater structural features. It is well suited for high precision and dense reconstructions in underwater environments.

TABLE IV  
COMPARISON OF ANGLE MEASUREMENT RESULTS BETWEEN PLANES

Angle	Ground Truth	Fitting Value	Error Rate
$\angle(\pi_\alpha \pi_\beta)$	84.15°	83.00°	1.37%
$\angle(\pi_\alpha \pi_\gamma)$	89.40°	88.54°	0.96%
$\angle(\pi_\beta \pi_\gamma)$	88.30°	87.70°	0.66%

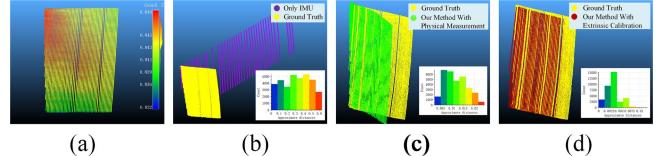


Fig. 10. Mobile local mapping results of plane in linear motion. (a) Static mapping result. (b) Mapping result by only IMU. (c) Mapping result by our method with physical measurement. (d) Mapping result by our method with extrinsic calibration. The histogram in the lower right corner indicates the distribution of the closest distances of the points from the ground truth.

#### D. Underwater Mobile Dense Mapping Experiments

In order to fully validate the ROV-Scanner's dense mapping performance during movement, three groups of comparison experiments are conducted separately. In the first two sets of experiments, local maps are shown with smaller mapping objects to highlight the contrast in effects. After all, the quality of the local map directly affects the obstacle avoidance performance. The system builds a map of a plane in linear motion for the first group. Then, it performs a rotational motion and builds a map of a curved surface for the second group. In the third group, the system moves freely for a long time in a random scene. To further validate the necessity of our multifrequency information fusion, some module comparison experiments are included in each group of experiments.

1) *Mobile Mapping of Plane in Linear Motion:* In this group, four sets of comparison experiments are implemented. First, the ROV-Scanner is fixed on a bracket to perform a static scanning of a plane, and the obtained point cloud is used as the ground truth. The system then performs a linear motion. In the second set of experiments, the system uses only the HF inertial information from the IMU for local mapping. While in the third and fourth set of experiments the system utilizes our proposed multifrequency information fusion method to build map. Among them, in order to verify the validity of our proposed extrinsic calibration method, the extrinsic parameters of the sensors are obtained by physical measurement in the third set, and the extrinsic parameters used in the fourth set are obtained by our calibration. The mapping results are shown in Fig. 10. The point clouds obtained by the different methods are then quantitatively compared with the ground truth, as shown in Table V. What can be seen is that the cumulative error of IMU is more serious than that of DVL, and relying solely on IMU to build a map is ineffective. In contrast, our proposed method reduces the map construction error. Moreover, the introduction of extrinsic calibration makes the mapping accuracy smaller than the usual extrinsic parameter matrix using physical measurement. The average error reaches the centimeter level, and the angle between the fitting plane and the ground truth is only 2.85°, which is meaningful for underwater mobile mapping.

TABLE V  
QUANTITATIVE COMPARISON OF DIFFERENT MOBILE MAPPING METHODS FOR PLANE

Method	Maximum Distance (m)	Average Distance (m)	Fitting Plane Angle
Only IMU	0.179	0.093	65.30°
Our Method with Physical Measurement	0.041	0.015	20.25°
Our Method with Extrinsic Calibration	<b>0.019</b>	<b>0.004</b>	<b>2.85°</b>

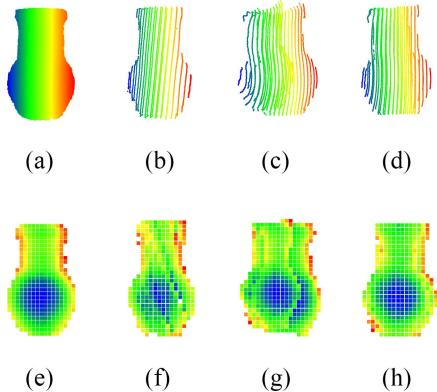


Fig. 11. Mobile local mapping results of curved surface in rotary motion. (a) Long time static mapping result. (b) Static local mapping result. (c) Mapping result by Palomer's [18]. (d) Mapping result by our method. (e)–(h) Projections of the point clouds of (a)–(d) in the z-direction, respectively.

TABLE VI  
QUANTITATIVE COMPARISON OF DIFFERENT MOBILE MAPPING METHODS FOR CURVED SURFACE

Method	Surface (m <sup>2</sup> )	RMSE (cm)	Average Distance (cm)
Ground Truth	0.038	—	—
Palomer's [18]	0.045	0.70	0.43
Our Method	<b>0.037</b>	<b>0.27</b>	<b>0.15</b>

2) *Mobile Mapping of Curved Surface in Rotary Motion:* In this experiment, the mapping target is a curved ceramic bottle. Similarly, the ROV-Scanner is first fixed and held stationary for a long period of time, and the scanning results are used as the ground truth, as shown in Fig. 11(a). For better comparison, Fig. 11(b) shows the local map result for part of the time. The system then performs a rotational motion. Two sets of comparison experiments are carried out. In the first set, Palomer's method [18] is utilized, while the second set utilizes our proposed method. The mapping results are shown in Fig. 11(c) and (d). Fig. 11(e)–(h) are the projections of the point clouds of (a)–(d) in the z-direction, respectively. For quantitative comparison, the iterative closest point (ICP) algorithm is used to evaluate the difference between the mobile mapping and static mapping. The results are shown in Table VI. It can be seen that our mobile mapping method is superior to Palomer's which has a large motion distortion. The projected surface area using our proposed method is closer to the ground truth. The RMS is as low as 0.27 cm and the average distance is only 0.15 cm.

3) *Mobile Mapping of Random Scene in Free Motion:* In the previous mobile mapping on small objects for short time,

the detailed features are quantitatively compared used to verify the effectiveness of our proposed mapping method. In this experiment, a long time free-motion mapping experiment is conducted in a large random scene as shown in Fig. 12(a). Fig. 12(b) shows the dark scene created during the experiment and the results of trajectory tracking by the global camera. The specific trajectory results are presented in Fig. 12(c) and (d), where the system trajectories obtained by DR and our method are compared. The trajectory error results for different directions are shown in Fig. 12(e)–(g). From the results, by introducing additional IMU, our method has more accurate estimation when local orientation changes, such as the enlarged area in Fig. 12(c). Failure in orientation will undoubtedly create fatal problems for the construction of the map compared to the translation accumulation error. Therefore, our method will get better performance in building maps even though the errors of both methods in XY direction are close. The trajectory RMSE of DR is 0.029 m in X-direction and 0.043 m in Y-direction, while that of our method is 0.037 and 0.034 m, respectively. However, our method achieves higher frequency state estimation, which ensures the effective utilization of the point cloud. From Fig. 12(h), it can be seen that the map by LF DR used in [19] and [20] is sparse and loses a lot of point cloud information. In addition, by fusing the pressure sensors, our method has basically no cumulative error in the Z-direction, as shown in Fig. 12(g). In the Z-direction, the trajectory RMSE of our method is only 0.016 m, while that of the DR method reaches 0.089 m, which makes the maps constructed based on the DR method produce great distortion. Fig. 12(i) and (j) shows the mapping results of our method with and without pressure sensor, respectively. Although both are dense, the former has a significant cumulative error in the Z-direction, as shown in the partial zoomed-in region of the figure. The latter has much better map consistency with the pressure sensor observations, which is expected to be used to a variety of underwater engineering applications.

#### E. Underwater Simultaneous Navigation and Dense Mapping

The refined exploration of unknown underwater environment is significant for marine development and is one of the key operations of autonomous underwater robots. Generally speaking, the robot is required to have autonomous cruising, collision-free navigation and dense mapping capabilities underwater. Existing methods are difficult to be put into practical applications in dark conditions, while our ROV-Scanner can realize this task well. In order to verify it, two groups of experiments are conducted. In the first group of experiments, the cruising ability of the system to the target areas and the mapping performance of the object in the process are tested. Then, the collision-free navigation ability and the obstacles mapping performance of the proposed system were tested in the second group of experiments.

1) *Target Areas Cruising and Dense Mapping:* In general, in order to maximize the efficiency in the exploration of underwater unknown areas, some key search areas are often artificially prioritized, and robots are used to carry out operations at the destinations by autonomous movement. Therefore,

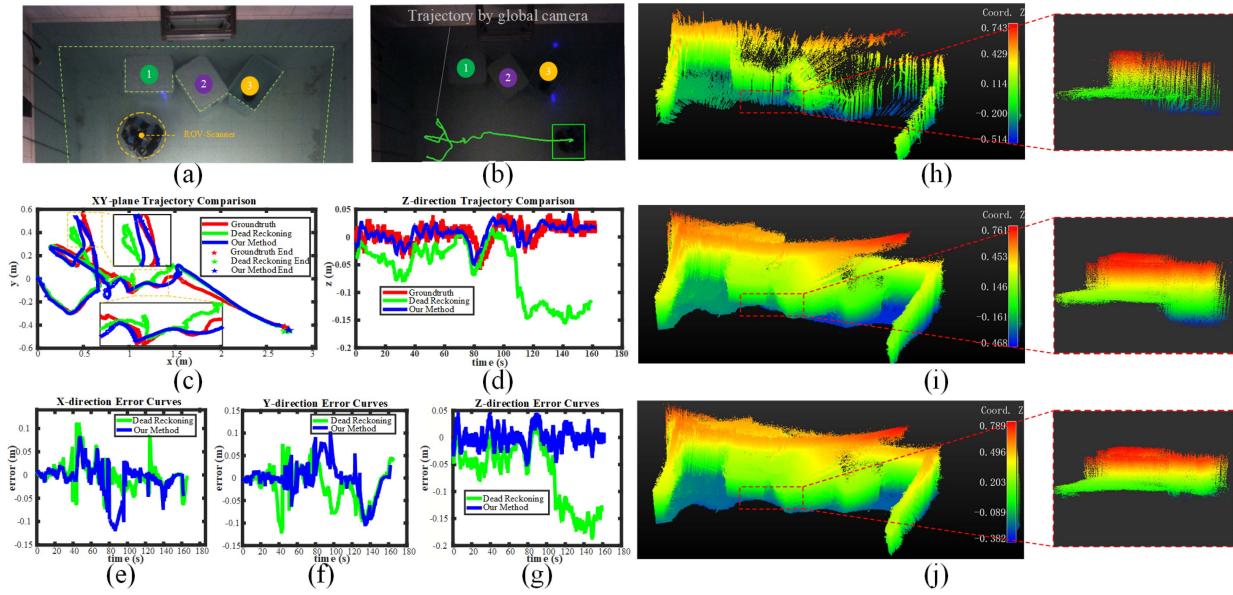


Fig. 12. Mobile mapping results of underwater random scene in free motion. (a) Experiment scene. (b) Trajectory measured by the global camera. (c) XY-plane trajectory comparison results. (d) Z-direction trajectory comparison results. (e)–(g) Error results in different directions, respectively. (h) Mobile mapping by the single LF DR method used in [19] and [20]. (i) Mobile mapping by our method without pressure sensor. (j) Mobile mapping by our method with pressure sensor.

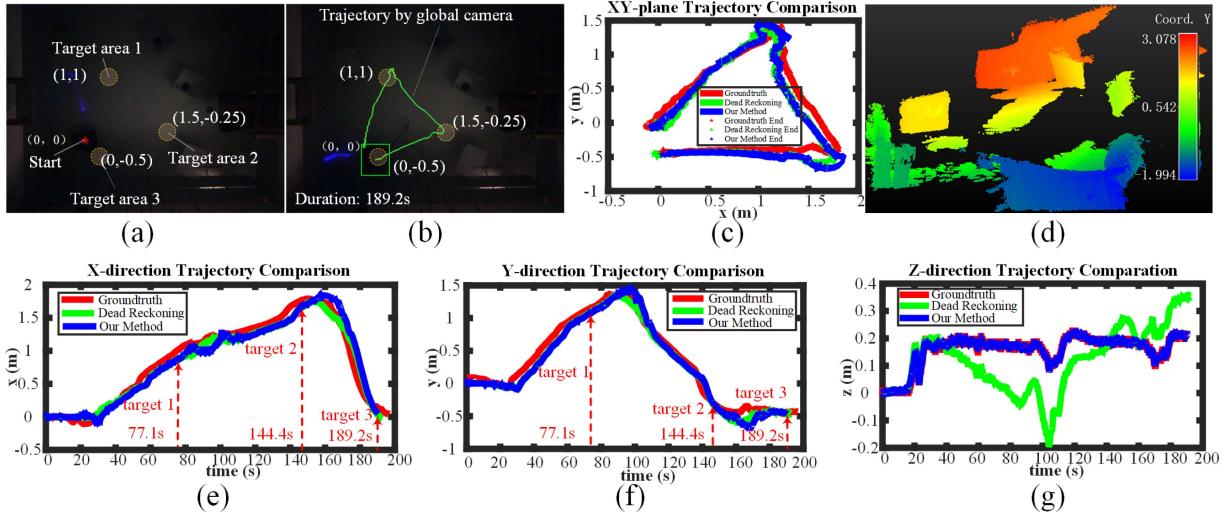


Fig. 13. Target areas cruising and dense mapping. (a) Experiment scene and target areas. (b) Trajectory measured by the global camera. (c) XY-plane trajectory comparison between DR used in [19] and [20] and our method. (d) Dense map built during the cruise. (e)–(g) Different direction trajectory comparison.

area cruising capability is essential. Based on acoustic–inertial–pressure information fusion, our ROV-Scanner realizes this goal. A set of experiment is conducted in dark underwater scene to verify it. As shown in Fig. 13(a), we set three circular cruising target areas for ROV-Scanner, each with a radius of 0.15 m and centers of (1, 1), (1.5, −0.25), (0, −0.5). The system starts from the starting position and moves autonomously toward the target areas sequentially, performing the mapping task in the process. The global camera records the motion trajectory of the system from the beginning to the end of the task, as shown in Fig. 13(b). It can be seen that the ROV-Scanner successfully reached the designated areas. It is worth stating that the cruise mission only needs to reach the target area and ultrahigh control accuracy is not considered in this article. Fig. 13(d) shows the result of mapping, which is

very dense and meets the requirements of refined exploration. Fig. 13(c),(e),(f), and (g) present the system’s trajectories in different directions. We can clearly find that ROV-Scanner reaches target 1 at 77.1 s, reaches target 2 at 144.4 s, and then reaches target 3 at 189.2 s, successfully completing the cruise task. Moreover, our method is significantly more stable in the Z direction than the DR method used in [19] and [20]. The trajectory RMSE of DR’s reaches 0.119 m, while ours is only 0.046 m.

2) *Collision-Free Navigation and Dense Mapping*: For unknown territory, obstacle avoidance capability is critical for safe robot navigation. Existing passive vision-based obstacle avoidance methods are hardly effective in dark and weak texture underwater scenes. In order to verify the collision-free navigation capability of ROV-Scanner based on active vision

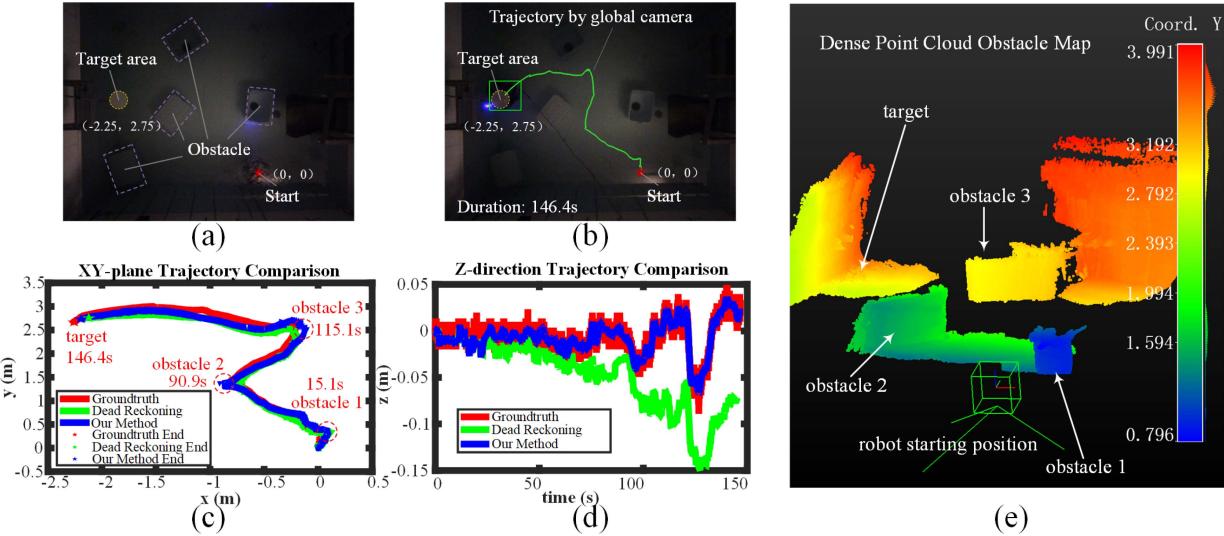


Fig. 14. Collision-free navigation and dense mapping. (a) Experiment scene with obstacles. (b) Trajectory measured by the global camera. (c) and (d) Different direction trajectory comparison between DR used in [19] and [20] and our method. (e) Dense map built during the collision-free navigation.

and the effect of obstacle mapping, a set of obstacle avoidance experiment in dark underwater scene is carried out, as shown in Fig. 14(a). A target area is set first, and the ROV-Scanner moves autonomously at starting point in conjunction with the previously described cruise capability. It is worth stating that the obstacle information in the scene is unknown, and the system uses our proposed underwater real-time obstacle avoidance algorithm based on BSL to realize collision-free navigation during the movement. From the view of global camera in Fig. 14(b), the ROV-Scanner successfully avoids the obstacles and safely moves to the target area. Detailed trajectory results are shown in Fig. 14(c) and (d), where it can be seen that the ROV-Scanner detects the obstacles and successfully avoids them at 15.1, 90.0, and 115.1 s, respectively, and finally successfully reaches the target area at 146.4 s. The conclusions of the evaluation of the trajectory RMSE are consistent with the previous experiments. In this experiment, the RMSE of our method in the  $X$ ,  $Y$ ,  $Z$  directions are 0.035, 0.039, and 0.005 m, respectively, while the DR-based RMSE reaches 0.036, 0.063, and 0.056 m. Fig. 14(e) shows the reconstructed maps of the obstacles information. Similarly, the point cloud map is very dense, and the obstacle information is very rich to observation. We believe this work provides new inspiration for underwater exploration and other related operations.

#### F. Discussion

The above experimental results demonstrate that our ROV-Scanner can well fulfill the task of refined exploration in unknown dark environments, with autonomous cruising, collision-free navigation, and dense mapping capabilities. This is mainly attributed to the accurate calibration method, the efficient multifrequency fusion mapping algorithm, the novel collision-free navigation framework and the self-designed hardware system. Among them, the proposed multifrequency information fusion mapping method ensures the effective utilization of HF structured light information, which makes

TABLE VII  
RMSE RESULTS OF TRAJECTORY FOR DIFFERENT EXPERIMENTS

RMSE (m)	X-axis		Y-axis		Z-axis	
	DR	Our	DR	Our	DR	Our
VI-D3	<b>0.029</b>	0.037	0.043	<b>0.034</b>	0.089	<b>0.016</b>
VI-E1	0.065	<b>0.059</b>	0.055	<b>0.053</b>	0.119	<b>0.046</b>
VI-E2	0.036	<b>0.035</b>	0.063	<b>0.039</b>	0.056	<b>0.005</b>

TABLE VIII  
MAE RESULTS OF TRAJECTORY FOR DIFFERENT EXPERIMENTS

MAE (m)	X-axis		Y-axis		Z-axis	
	DR	Our	DR	Our	DR	Our
VI-D3	<b>0.020</b>	0.023	0.032	<b>0.022</b>	0.069	<b>0.012</b>
VI-E1	0.049	<b>0.047</b>	0.039	<b>0.037</b>	0.095	<b>0.029</b>
VI-E2	0.028	<b>0.025</b>	0.040	<b>0.026</b>	0.043	<b>0.004</b>

the constructed maps denser. Moreover, the fusion of multiple sensors, including inertial information from IMU, acoustic information from DVL, and pressure information from pressure sensors, makes the robot's state estimation more accurate. The RMSE and mean absolute error (MAE) statistics of the trajectories in the above experiments (Sections VI-D3, VI-E1, and VI-E2) are shown in Tables VII and VIII. It can be seen that our method has higher accuracy and greatly reduces the trajectory error especially in the  $Z$ -direction.

At present, underwater exploration has long been a fascinating research hotspot with many interesting studies. We statistically analyze some of the SOTA underwater exploration systems, as shown in Table IX. By using a variety of sensors as well as various algorithms, these systems achieve underwater exploration operations under specific conditions. However, underwater dark conditions and refraction effects affect their robustness. By utilizing active structured light vision, our ROV-Scanner can simultaneously realize autonomous collision-free navigation and dense mapping in dark underwater environments for the first time.

Furthermore, it has always been a challenge to apply underwater exploration systems in real oceans due to the complexity

TABLE IX  
COMPARISON OF SOTA SYSTEMS FOR UNDERWATER EXPLORATION

Characteristics	[4]	[5]	[7]	[19], [20]	[31]	ROV-Scanner
collision-free navigation	✓	—	✓	x	x	✓
target areas cruising	✓	✓	✓	x	✓	✓
dense mapping	x	✓	✓	✓	x	✓
overcome refraction	x	x	x	—	—	✓
work in low-light	✓	x	x	✓	✓	✓

and variability of the underwater environment. But our ROV-Scanner has great application advantages. First, the robots could float up and down affected by irregular underwater currents. Undoubtedly, it would cause more suffering for refined exploration. Compared with the LF DR mapping, our multifrequency information fusion method is more effective when facing this HF changes. In addition, the introduction of pressure sensors allows our system to effectively compensate for depth variations, thus improving the accuracy of the maps. Moreover, our BSL is based on the underwater refraction measurement model, which can effectively compensate the influence of underwater refraction effect on the measurement accuracy. As can be seen from Section IV-A, the only variable in our model that is affected by the environment is the refractive index of water. For different areas, the refractive index of water will change due to different turbidity and salinity. Therefore, only field measurements with a specialized refractometer are needed for the actual operation, and no other additional compensations such as artificial lighting. Furthermore, for some turbid areas, light scattering by particles in the water is so severe that the common passive vision-based mapping method is no longer suitable. ROV-Scanner, thanks to the high penetration of the laser, can still work effectively in these turbid areas. Overall, our system has great potential for marine engineering.

## VII. CONCLUSION

In this article, a structured light-based underwater robotic collision-free navigation and dense mapping system, named ROV-Scanner, is designed for refined exploration in unknown dark environments, where the on-board design allows for autonomous mobility and operational capabilities. Then, an underwater 3-D dense mapping algorithm fusing DVL, IMU, and pressure sensor multifrequency information is proposed to realize dense mapping during robot motion. Moreover, an air-water two-stage underwater multisensor calibration method is presented. In particular, the extrinsic parameters between DVL and camera is innovatively calibrated using BSL by graph-based optimization, enhancing robustness. In addition, a framework of BSL-based collision-free navigation is presented to guarantee the safe movement of the robot in unknown dark environments. Finally, experimental results show that our system can simultaneously realize autonomous collision-free navigation and dense mapping in dark underwater environments, which has great potential for underwater application.

## REFERENCES

- R. Wang, S. Wang, Y. Wang, M. Tan, and J. Yu, "A paradigm for path following control of a ribbon-fin propelled biomimetic underwater vehicle," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 49, no. 3, pp. 482–493, Mar. 2019.
- S. Kong, M. Tian, C. Qiu, Z. Wu, and J. Yu, "IWSCR: An intelligent water surface cleaner robot for collecting floating garbage," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 10, pp. 6358–6368, Oct. 2021.
- P. Xu et al., "Design and implementation of lightweight AUV with multisensor aided for underwater intervention tasks," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 69, no. 12, pp. 5009–5013, Dec. 2022.
- E. Vidal, N. Palomeras, K. Istenič, N. Gracias, and M. Carreras, "Multisensor online 3D view planning for autonomous underwater exploration," *J. Field Robot.*, vol. 37, no. 6, pp. 1123–1147, 2020.
- B. Joshi et al., "Underwater exploration and mapping," in *Proc. IEEE/OES Auton. Underw. Veh. Symp. (AUV)*, 2022, pp. 1–7.
- Z. He, L. Dong, C. Sun, and J. Wang, "Asynchronous multithreading reinforcement-learning-based path planning and tracking for unmanned underwater vehicle," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 52, no. 5, pp. 2757–2769, May 2022.
- J. D. Hernández et al., "Autonomous underwater navigation and optical mapping in unknown natural environments," *Sensors*, vol. 16, no. 8, p. 1174, 2016.
- J. Lv, Y. Wang, S. Wang, X. Bai, R. Wang, and M. Tan, "A collision-free planning and control framework for a biomimetic underwater vehicle in dynamic environments," *IEEE/ASME Trans. Mechatronics*, vol. 28, no. 3, pp. 1415–1424, Jun. 2023.
- P. Yang, H. Liu, M. Roznere, and A. Q. Li, "Monocular camera and single-beam sonar-based underwater collision-free navigation with domain randomization," in *Proc. Int. Symp. Robot. Res.*, 2022, pp. 85–101.
- T. Manderson, J. C. G. Higuera, R. Cheng, and G. Dudek, "Vision-based autonomous underwater swimming in dense coral for combined collision avoidance and target selection," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2018, pp. 1885–1891.
- B. Arain, C. McCool, P. Rigby, D. Cagara, and M. Dunbabin, "Improving underwater obstacle detection using semantic image segmentation," in *Proc. Int. Conf. Robot. Autom. (ICRA)*, 2019, pp. 9271–9277.
- N. Lyu, H. Yu, J. Han, and D. Zheng, "Structured light-based underwater 3-D reconstruction techniques: A comparative study," *Opt. Lasers Eng.*, vol. 161, Feb. 2023, Art. no. 107344.
- J. Fan, X. Wang, C. Zhou, Y. Ou, F. Jing, and Z. Hou, "Development, calibration and image processing of underwater structured light vision system: A survey," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–18, Jan. 2023, doi: [10.1109/TIM.2023.3235420](https://doi.org/10.1109/TIM.2023.3235420).
- A. Palomer, P. Ridao, J. Forest, and D. Ribas, "Underwater laser scanner: Ray-based model and calibration," *IEEE/ASME Trans. Mechatronics*, vol. 24, no. 5, pp. 1986–1997, Oct. 2019.
- Y. Ou, J. Fan, C. Zhou, S. Tian, L. Cheng, and M. Tan, "Binocular structured light 3-D reconstruction system for low-light underwater environments: Design, modeling, and laser-based calibration," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–14, Mar. 2023, doi: [10.1109/TIM.2023.3261941](https://doi.org/10.1109/TIM.2023.3261941).
- X. Li et al., "High-accuracy calibration method for an underwater one-mirror galvanometric laser scanner," *Opt. Exp.*, vol. 31, no. 4, pp. 5973–5989, 2023.
- C. Bräuer-Burchardt et al., "Underwater 3D scanning system for cultural heritage documentation," *Remote Sens.*, vol. 15, no. 7, p. 1864, 2023.
- A. Palomer, P. Ridao, and D. Ribas, "Inspection of an underwater structure using point-cloud SLAM with an AUV and a laser scanner," *J. Field Robot.*, vol. 36, no. 8, pp. 1333–1344, 2019.
- A. Bodenmann, B. Thornton, and T. Ura, "Generation of high-resolution three-dimensional reconstructions of the seafloor in color using a single camera and structured light," *J. Field Robot.*, vol. 34, no. 5, pp. 833–851, 2017.
- T. Hitchcox and J. R. Forbes, "Improving self-consistency in underwater mapping through laser-based loop closure," *IEEE Trans. Robot.*, vol. 39, no. 3, pp. 1873–1892, Jun. 2023.
- E. Westman and M. Kaess, "Underwater AprilTag SLAM and calibration for high precision robot localization," School Comput. Sci., Carnegie Mellon Univ. Pittsburgh, Pittsburgh, PA, USA, Rep. TR-CMU-RI-TR-18-43, 2018.
- S. Xu et al., "Underwater visual acoustic SLAM with extrinsic calibration," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2021, pp. 7647–7652.

- [23] P. Furgale, T. D. Barfoot, and G. Sibley, "Continuous-time batch estimation using temporal basis functions," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2012, pp. 2088–2095.
- [24] A. Segal, D. Haehnel, and S. Thrun, "Generalized-ICP," in *Proc. Robot., Sci. Syst.*, vol. 2, 2009, p. 435.
- [25] A. Agrawal, S. Ramalingam, Y. Taguchi, and V. Chari, "A theory of multi-layer flat refractive geometry," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2012, pp. 3346–3353.
- [26] H. Yang, S. Tian, J. Fan, Y. Han, C. Zhou, and Y. Ou, "A precise calibration method for the underwater binocular camera with unknown refractive index," *IEEE Sensors J.*, vol. 23, no. 2, pp. 1565–1574, Jan. 2023.
- [27] C. Forster, L. Carbone, F. Dellaert, and D. Scaramuzza, "IMU preintegration on manifold for efficient visual-inertial maximum-a-posteriori estimation," in *Proc. Robot., Sci. Syst.*, 2015, pp. 1–10.
- [28] M. Li and A. I. Mourikis, "High-precision, consistent EKF-based visual-inertial odometry," *Int. J. Robot. Res.*, vol. 32, no. 6, pp. 690–711, 2013.
- [29] K. L. Walker et al., "Experimental validation of wave induced disturbances for predictive station keeping of a remotely operated vehicle," *IEEE Robot. Autom. Lett.*, vol. 6, no. 3, pp. 5421–5428, Jul. 2021.
- [30] R. Vekkanandan, D. Chang, and G. A. Hollinger, "Autonomous underwater docking using flow state estimation and model predictive control," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, 2023, pp. 1062–1068.
- [31] Y. Ling, Y. Li, T. Ma, Z. Cong, S. Xu, and Z. Li, "Active bathymetric SLAM for autonomous underwater exploration," *Appl. Ocean Res.*, vol. 130, Jan. 2023, Art. no. 103439.



**Yaming Ou** received the B.E. degree in automation from Southeast University, Nanjing, China, in 2021. He is currently pursuing the Ph.D. degree in control theory and control engineering with the Institute of Automation, Chinese Academy of Sciences, Beijing, China, and also with the School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing.

His research interests include underwater 3-D vision, SLAM, multisensor fusion, and autonomous robot navigation.



**Junfeng Fan** (Member, IEEE) received the B.S. degree in mechanical engineering and automation from the Beijing Institute of Technology, Beijing, China, in 2014, and the Ph.D. degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Sciences (IACAS), Beijing, in 2019.

He is currently an Associate Professor of Control Theory and Control Engineering with the Laboratory of Cognition and Decision Intelligence for Complex Systems, IACAS. His research interests include robot vision and underwater robot.



**Chao Zhou** (Member, IEEE) received the B.S. degree in automation from Southeast University, Nanjing, China, in 2003, and the Ph.D. degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Sciences (IACAS), Beijing, China, in 2008.

He is currently a Professor of Control Theory and Control Engineering with the Laboratory of Cognition and Decision Intelligence for Complex Systems, IACAS. His research interests include underwater robot and bionic robot.



**Song Kang** received the B.S. degree in mechanical design manufacture and automation from Xinjiang Agricultural University, Urumqi, China, in 2011, and the Ph.D. degree in robotics engineering from Beihang University, Beijing, China, in 2022.

He is currently an Assistant Professor with the Laboratory of Cognition and Decision Intelligence for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing. His research interests include underwater robots and intelligent control systems.



**Zhuoliang Zhang** received the B.E. degree in automation from Tongji University, Shanghai, China, in July 2018, and the Ph.D. degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Sciences (IACAS), Beijing, China, in June 2023.

He is currently working as a Postdoctoral Fellow with the Laboratory of Cognition and Decision Intelligence for Complex Systems, IACAS. His research interests include measurements, sensor signal processing, and intelligent control.



**Zeng-Guang Hou** (Fellow, IEEE) received the B.E. and M.E. degrees in electrical engineering from Yanshan University (formerly, North-East Heavy Machinery Institute), Qinhuangdao, China, in 1991 and 1993, respectively, and the Ph.D. degree in electrical engineering from the Beijing Institute of Technology, Beijing, China, in 1997.

He was a Postdoctoral Research Fellow with the Key Laboratory of Systems and Control, Institute of Systems Science, Chinese Academy of Sciences, Beijing, from May 1997 to June 1999 and a Research Assistant with The Hong Kong Polytechnic University, Hong Kong, China, from May 2000 to January 2001. From July 1999 to May 2004, he was an Associate Professor with the Institute of Automation, Chinese Academy of Sciences and has been a Full Professor since June 2004. From September 2003 to October 2004, he was a Visiting Professor with the Intelligent Systems Research Laboratory, College of Engineering, University of Saskatchewan, Saskatoon, SK, Canada. He is a Professor with the State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences. His research interests include neural networks, robotics, and intelligent systems.

Prof. Hou was an Associate Editor of the *IEEE Computational Intelligence Magazine* and the *IEEE TRANSACTIONS ON NEURAL NETWORKS*. He is an Associate Editor of the *IEEE TRANSACTIONS ON CYBERNETICS* and *ACTA Automatica Sinica* and an Editorial Board Member of *Neural Networks*. He was the Chair of Neural Network Technical Committee of Computational Intelligence Society and is currently the Chair of Adaptive Dynamic Programming and Reinforcement Learning Technical Committee of CIS.



**Min Tan** received the B.E. degree from Tsinghua University, Beijing, China, in 1986, and the Ph.D. degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Sciences, Beijing, in 1990.

He is currently a Professor with the Laboratory of Cognition and Decision Intelligence for Complex Systems, Institute of Automation, Chinese Academy of Sciences. His research interests include advanced robot control, biomimetic robot, and multirobot systems.