

Water-MBSL: Underwater Movable Binocular Structured Light-Based High-Precision Dense Reconstruction Framework

Yaming Ou , Junfeng Fan , *Member, IEEE*, Chao Zhou , *Member, IEEE*, Long Cheng , *Fellow, IEEE*, and Min Tan 

Abstract—Structured light systems are widely used in underwater dense reconstruction due to their excellent accuracy. However, the current related methods mainly focus on fixed positions. The reconstruction performance in motion is insufficient. Therefore, we propose an underwater movable binocular structured light (MBSL) based high-precision dense reconstruction framework, named WaterMBSL, to realize the robot reconstruction while moving. Specifically, an onboard binocular structured light system based on mirror-galvanometer is developed first. Then, a simplified underwater point cloud acquisition algorithm is presented to quickly obtain 3-D information of the scene. Besides, a new underwater motion compensation algorithm combining inertial measurement unit and uniform velocity model is proposed. Moreover, the generalized-ICP point cloud registration algorithm is introduced to achieve accurate motion estimation. Finally, an underwater movable reconstruction platform is developed by integrating the self-designed structured light system with the underwater robot BlueROV for validating the performance of our proposed Water-MBSL. Experimental results show that satisfactory motion reconstruction performance can be obtained.

Index Terms—Motion reconstruction, structured light, underwater 3-D reconstruction, underwater exploration.

NOMENCLATURE

Symbol Explanation

\mathcal{W}	World coordinate system.
\mathcal{L}	Left camera coordinate system.
\mathcal{R}	Right camera coordinate system.
\mathcal{I}	IMU coordinate system.
\mathcal{S}	Coordinate system of the scan point cloud ($= \mathcal{L}$).
$p^{\mathcal{A}}$	Representation of p under the coordinate system \mathcal{A} .
$T_{\mathcal{A}}^{\mathcal{B}}$	Transformation matrix from the coordinate system \mathcal{A} to \mathcal{B} .
Scan	Point cloud obtained from each laser line.
Sweep	Point cloud for registration combined by multiple scans.

I. INTRODUCTION

IN RECENT years, underwater robots have been widely used in underwater operations, such as underwater pipeline monitoring [1], defect detection [2], inspection [3], and archaeology [4]. By obtaining 3-D information of underwater scenes, robots can better perform these tasks. Therefore, underwater 3-D reconstruction has important research significance. As a common reconstruction method, vision is mainly divided into two forms: passive vision and active vision. Although there have been some studies based on passive vision to achieve robot navigation [5] and localization [6], the lack of light and the scarcity of scene texture features make passive vision relying on scene image feature matching ineffective for underwater reconstruction. Fortunately, active vision 3-D reconstruction methods represented by structured light systems [7], [8] are highly adaptable to the environment. The line structured light system (LSLS) has been shown to be more suitable for underwater environments [9], [10]. Depending on the scanning method, LSLS can be divided into three categories: single-laser LSLS (SLLSLS) [11], [12], motor-driven LSLS [13], [14], [15], and self-scanning LSLS (SSLSLS) [16], [17].

In order to realize the reconstruction of a larger range of scenes by LSLS, it is usually mounted on an underwater robot. In this case, precise motion estimation of the robot is the key

Manuscript received 10 September 2023; accepted 3 December 2023. Date of publication 29 December 2023; date of current version 4 April 2024. This work was supported in part by Beijing Natural Science Foundation under Grant 4232057, in part by the National Natural Science Foundation of China under Grant 62373354, Grant 62003341, Grant 62333023, and Grant 62025307, 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. Paper no. TII-23-3484. (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, 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; min.tan@ia.ac.cn).

Long Cheng is with the State Key Laboratory of Multimodal Artificial Intelligence Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China (e-mail: long.cheng@ia.ac.cn).

This article has supplementary material provided by the authors and color versions of one or more figures available at <https://doi.org/10.1109/TII.2023.3342899>.

Digital Object Identifier 10.1109/TII.2023.3342899

to accurate reconstruction. Bodenmann et al. [18] designed an SSLSLS-based underwater mapping device, named SeaXerocks 1. By deploying it to the robot Hyper-Dolphin, the mapping of an area of the Pacific Ocean floor was achieved [18]. In a recent study, Hitchcox and Forbes [19] used a commercial SSLSLS reconstruction device, Insight Pro, to achieve a 3-D scan of an underwater shipwreck [19]. Among them, the dead-reckoning (DR) of expensive Doppler velocity log (DVL) was used as the motion estimation of the robot. However, there is a serious cumulative error with long DR. In addition, Bräuer-Burchardt et al. [21] attempted to estimate the robot motion state based on VINS-Mono [20] using a monocular and inertial measurement unit (IMU) [21]. Despite its low cost, this approach still relies on feature extraction from monocular images, which is difficult to implement in harsh underwater environments.

The above methods all rely on external sensors to realize the motion estimation, which have limited reconstruction accuracy due to system calibration errors, cumulative errors, and robustness defects. To achieve self-motion estimation of LSLS, Palomer et al. [22] designed an SSLSLS equipped with an underwater robot to detect a 3-D subsea infrastructure inside a water tank. Among it, a feature-based coarse alignment was performed, which was then refined using iterative closest point (ICP). The results show that the point cloud registration-based reconstruction is superior to that based on DVL's DR. However, to minimize the motion distortion, Palomer et al. [22] applied zero velocity control to the robot at each point cloud acquisition and considered the robot to be completely stationary during this process. On the one hand, the zero velocity control introduced by Palomer et al. [22] prevents reconstruction while moving, which would greatly reduce the reconstruction efficiency and impose more energy consumption on the robot. On the other hand, a completely static state is very difficult to achieve for an underwater robot, and motion distortion will inevitably exist. Castillon et al. [23] verified that motion distortion has a substantial impact on the reconstruction effect of LSLS.

In summary, the existing frameworks for motion reconstruction using LSLS have the following potential problems.

- 1) Reliance on external sensors, which causes insufficient reconstruction accuracy and increases costs.
- 2) Failure to consider motion distortion, which affects the effect of motion estimation and reconstruction.

In order to solve the above problems and achieve more efficient and accurate motion reconstruction in complex underwater environments, which is crucial for various underwater robotic operations including autonomous inspection, collision-free navigation, and seabed exploration, this article offers the following contributions.

- 1) An SSLSLS-based underwater movable high-precision dense reconstruction framework, named Water-movable binocular structured light (MBSL), is proposed, which takes into account both motion distortion and point cloud registration. To the best of authors' knowledge, this is the first work to achieve underwater movable high-precision dense reconstruction without relying on expensive DVL.
- 2) In Water-MBSL, a new underwater motion compensation algorithm combining IMU and uniform velocity model

is proposed to lay the foundation for accurate point cloud registration, which is a rare attempt in the underwater domain. Considering the existence of underwater measurement errors, the generalized-ICP (GICP) point cloud registration algorithm is introduced into the framework to achieve accurate motion estimation, which increased the robustness and reconstructed point cloud quality.

- 3) An underwater movable reconstruction platform is developed by integrating the self-designed SSLSLS with the underwater robot BlueROV to verify the performance of Water-MBSL. Self-stabilizing scanning and motion scanning reconstruction experiments show that the proposed Water-MBSL framework could achieve better motion reconstruction performance compared with state-of-the-art LSLS-based underwater motion reconstruction methods.

II. OVERVIEW OF SYSTEM

To realize the performance verification of Water-MBSL, an underwater movable reconstruction platform is built, as shown in Fig. 1. The system consists of two main parts: the measurement platform and the motion platform. The measurement platform consists of a binocular structured light system, an IMU, an Inter NUC11TNH, an STM32F407, and a power supply unit. In this case, the NUC11TNH acts as the main controller of the system and is responsible for the data processing of all sensors and the control of the robot. The IMU has a frequency of 400 Hz and sends the data to the NUC11TNH through a serial protocol. The binocular structured light system is of SSLSLS-type consisting of binocular camera, laser emitter, and mirror-galvanometer. The scanning frequency of the mirror-galvanometer is controlled by the STM32F407, which receives the signal from the NUC11TNH through the serial port and can control the frequency of the mirror-galvanometer online. Thus, the scanning frequency of the SSLSLS can be controlled in real time. In this experiment, the frequency is set to 1 Hz in order to take into account the number of point clouds scanned in each mirror-galvanometer scan cycle. In total, 25 600 3-D points can be calculated for each period, which is a remarkable number for the field of underwater reconstruction. The motion platform of the system is the commercial robot BlueROV, which is connected to the fathom-X using a twisted pair cable. It is then connected to a micro switch for communication with NUC11TNH. Therefore, the control of the BlueROV can also be implemented onboard and does not necessarily require teleoperation. Specific hardware parameters of the developed platform are given in Table I.

III. FRAMEWORK OF WATER-MBSL

The existing LSLS-based motion reconstruction frameworks either rely on external sensors for motion estimation or fail to consider the effect of motion distortion, both of which have insufficient reconstruction accuracy. Therefore, this article proposes the underwater movable high-precision dense reconstruction framework Water-MBSL in combination with the developed underwater movable reconstruction platform, which is mainly

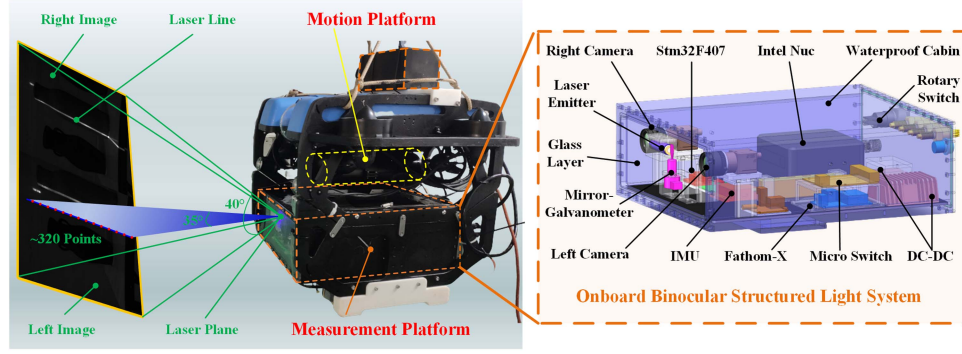


Fig. 1. Hardware design of underwater movable reconstruction platform.

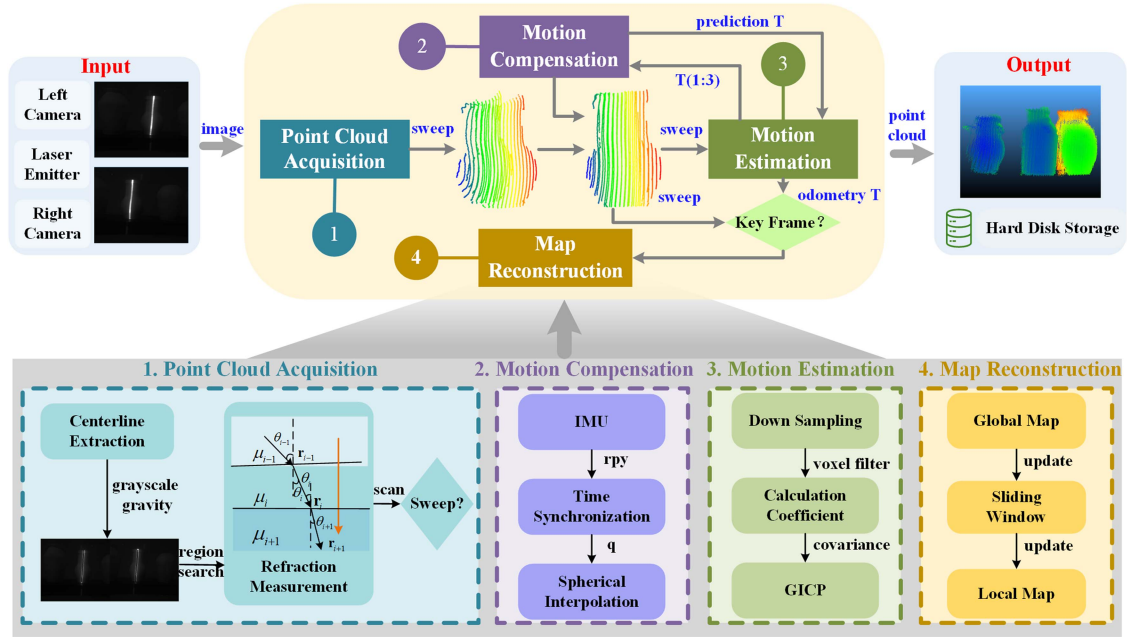


Fig. 2. Framework of underwater movable binocular structured light-based high-precision dense reconstruction, named Water-MBSL.

TABLE I
SYSTEM HARDWARE PARAMETER

Hardware	Parameters	Value
Camera	Resolution	1280×1024 pixels
	Frames	100 fps
	Focal Length	6 mm
	Communication Interface	USB 3.0
IMU	Depth of Field Range	0.4 – 3 m
	Frequency	400 Hz
Laser Emitter	Communication Interface	USART
	Wavelength	450 nm
	Power Consumption	400 MW

aimed at solving the high-precision dense reconstruction problem in underwater robot motion. The specific framework is shown in Fig. 2. It mainly includes four modules: point cloud acquisition, motion compensation, motion estimation, and map reconstruction.

Notations: Some necessary symbols are given in the Nomenclature.

A. Point Cloud Acquisition

It is challenging to acquire dense and accurate point clouds for underwater environments. In our previous work, a binocular SSLSLS was designed to achieve static dense reconstruction in a low-light environment. Based on the previous research, we redesign an onboard underwater binocular SSLSLS for mobile dense reconstruction of underwater robots. Among them, the pixel coordinates of the laser centerline are extracted in the same way that the grayscale center of gravity method is applied

$$U_m = \frac{\sum_{i=0}^M g_i \times u_i}{\sum_{i=0}^M g_i} \quad (1)$$

where M denotes the laser line candidate region width, usually obtained by binarizing the grayscale image. g_i is the candidate

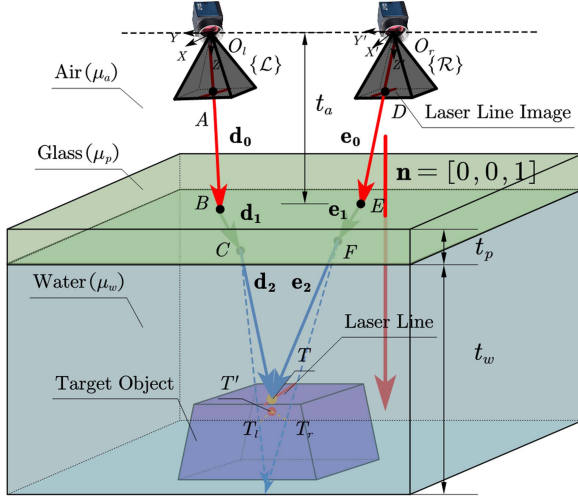


Fig. 3. Simplified binocular refraction measurement model.

region gray value, u_i is the pixel value, and U_m is the computational center coordinate.

In the underwater environment, the camera imaging light will pass through water, glass, and air in turn. A refraction effect will occur during this period, which makes binocular matching based on polar line constraint no longer holds. The refractive indices are μ_w , μ_p , and μ_a , respectively. In order to improve the computational efficiency, a region search matching algorithm is proposed. First, for a certain laser center pixel point p_l in the left image, we use the polar line constraint to obtain the position in the right image as the initial value of the search, and then search a number of pixels up and down in turn. Combined with the idea of reprojection, we believe that the distance between lines CT_l and FT_r should be the closest if the two points are matching points. That is, the distances between p_l and p_{ri} ($i = 0, \dots, l$, l is the search width) reprojected onto the 3-D space are calculated separately each time, and then the p_{ri} corresponding to the minimum value is taken as the matching point.

The traditional triangulation model is also no longer applicable due to underwater refraction. On the basis of existing research, our measurement model is shown in Fig. 3. The difference is that in order to simplify the calibration process and improve the efficiency of the onboard calculation, we assume the left and right cameras are parallel to the glass plane and at the same distance from the plane. This is a reasonable assumption with precise hardware assembly. That is

$$\mathbf{n}^{\mathcal{L}} = (0, 0, 1); t_l = t_r = t_a \quad (2)$$

where \mathbf{n} denotes the glass plane normal vector. t_l and t_r are the distances of the left and right camera optical centers from the glass plane, respectively. t_a is the common representation of the two.

According to the monocular pinhole camera model

$$\mathbf{A}^{\mathcal{L}} = \mathbf{K}_l^{-1} p_l; \mathbf{D}^{\mathcal{R}} = \mathbf{K}_r^{-1} p_r \quad (3)$$

where \mathbf{K}_l and \mathbf{K}_r are the intrinsic matrices of the cameras. $\mathbf{A}^{\mathcal{L}}$ and $\mathbf{D}^{\mathcal{R}}$ are the intersections of the imaging optical paths of

the left and right cameras with their respective imaging planes, respectively.

Suppose the camera extrinsic matrix is $\mathbf{T}_{\mathcal{R}}^{\mathcal{L}}$, then

$$\mathbf{D}^{\mathcal{L}} = \mathbf{T}_{\mathcal{R}}^{\mathcal{L}} \mathbf{D}^{\mathcal{R}}. \quad (4)$$

Therefore, for any point in \mathcal{R} can be converted under the \mathcal{L} using $\mathbf{T}_{\mathcal{R}}^{\mathcal{L}}$. So the following derivation does not distinguish between \mathcal{L} and \mathcal{R} for the time being. All indicates by default that it has been converted to \mathcal{L} . Also

$$\mathbf{d}_0 = \frac{\mathbf{AB}}{|\mathbf{AB}|} = \frac{\mathbf{OA}}{|\mathbf{OA}|}; \mathbf{e}_0 = \frac{\mathbf{DE}}{|\mathbf{DE}|} = \frac{\mathbf{OD}}{|\mathbf{OD}|} \quad (5)$$

where \mathbf{d}_i ($i = 0, 1, 2$) and \mathbf{e}_i ($i = 0, 1, 2$) denote the direction vectors of the left and right optical paths in different media, respectively. \mathbf{B} and \mathbf{E} are the intersections of the left and right optical paths with the air–glass manifold, respectively.

So,

$$\mathbf{B} = \frac{t_a}{\mathbf{d}_{0(3)}} \mathbf{d}_0; \mathbf{E} = \frac{t_a}{\mathbf{e}_{0(3)}} \mathbf{e}_0 \quad (6)$$

where the subscript (3) denotes the third element of the vector. According to the principle of refraction (POR) plane proposed by Agrawal et al. [24]

$$\mathbf{d}_1 = \alpha_{11} \mathbf{d}_0 + [0, 0, \beta_{11}]; \mathbf{e}_1 = \alpha_{12} \mathbf{e}_0 + [0, 0, \beta_{12}]. \quad (7)$$

Among them

$$\begin{cases} \alpha_{11} = \alpha_{12} = \frac{\mu_a}{\mu_p} \\ \beta_{11} = \sqrt{1 - \alpha_{11}^2 [1 - (\mathbf{d}_{0(3)})^2]} - \alpha_{11} \mathbf{d}_{0(3)} \\ \beta_{12} = \sqrt{1 - \alpha_{12}^2 [1 - (\mathbf{e}_{0(3)})^2]} - \alpha_{12} \mathbf{e}_{0(3)}. \end{cases}$$

Similar to (6), the intersections \mathbf{C} and \mathbf{F} of the left and right optical paths with the glass–water manifold can be computed. Namely

$$\begin{aligned} \mathbf{C} &= \mathbf{B} + \frac{t_p}{\mathbf{d}_{1(3)}} \mathbf{d}_1 = \frac{t_a}{\mathbf{d}_{0(3)}} \mathbf{d}_0 + \frac{t_p}{\mathbf{d}_{1(3)}} \mathbf{d}_1 \\ \mathbf{F} &= \mathbf{E} + \frac{t_p}{\mathbf{e}_{1(3)}} \mathbf{e}_1 = \frac{t_a}{\mathbf{e}_{0(3)}} \mathbf{e}_0 + \frac{t_p}{\mathbf{e}_{1(3)}} \mathbf{e}_1. \end{aligned} \quad (8)$$

Furthermore, according to (7), the direction vectors \mathbf{d}_2 and \mathbf{e}_2 can likewise be derived

$$\mathbf{d}_2 = \alpha_{21} \mathbf{d}_1 + [0, 0, \beta_{21}]; \mathbf{e}_2 = \alpha_{22} \mathbf{e}_1 + [0, 0, \beta_{22}]. \quad (9)$$

Among them

$$\begin{cases} \alpha_{21} = \alpha_{22} = \frac{\mu_p}{\mu_w} \\ \beta_{21} = \sqrt{1 - \alpha_{21}^2 [1 - (\mathbf{d}_{1(3)})^2]} - \alpha_{21} \mathbf{d}_{1(3)} \\ \beta_{22} = \sqrt{1 - \alpha_{22}^2 [1 - (\mathbf{e}_{1(3)})^2]} - \alpha_{22} \mathbf{e}_{1(3)}. \end{cases}$$

Due to the existence of systematic errors, usually line $(\mathbf{C}, \mathbf{d}_2)$ and line $(\mathbf{F}, \mathbf{e}_2)$ do not intersect. In order to determine the location of the target point \mathbf{T} , the following derivation is implemented.

First, according to the POR principle [24], \mathbf{T} satisfies two constraints. Specifically, \mathbf{TA} , \mathbf{d}_0 , and \mathbf{n} are coplanar, and \mathbf{TD} ,

Algorithm 1: Simplified Underwater Point Cloud Acquisition Algorithm Based on Binocular Structure Light.

Input: left camera image I_l , right camera image I_r , laser centerline threshold v , size of search area l

Output: scan point cloud S

```

1:  $I_l', I_r' = \text{binarization}(v, I_l, I_r)$ 
2:  $P_l, P_r = \text{laser\_centerline\_extraction}(I_l', I_r')$ , see (1)
3:  $d_{\min} = \text{INT\_MAX}$ 
4: for  $p_{li}$  in  $P_l$  do
5:   for  $p_{r[i-\frac{1}{2}]}$  to  $p_{r[i+\frac{1}{2}]}$  in  $P_r$  do
6:     calculate  $\mathbf{A}^{\mathcal{L}}, \mathbf{D}^{\mathcal{R}}$ , see (3)
7:     calculate  $\mathbf{d}_0^{\mathcal{L}}, \mathbf{e}_0^{\mathcal{R}}$ , see (5)
8:     calculate  $\mathbf{B}^{\mathcal{L}}, \mathbf{E}^{\mathcal{R}}$ , see (6)
9:     calculate  $\mathbf{d}_1^{\mathcal{L}}, \mathbf{e}_1^{\mathcal{R}}$ , see (7)
10:    calculate  $\mathbf{C}^{\mathcal{L}}, \mathbf{F}^{\mathcal{R}}, \mathbf{d}_2^{\mathcal{L}}, \mathbf{e}_2^{\mathcal{R}}$ , see (8)-(9)
11:    calculate  $\mathbf{D}^{\mathcal{L}}, \mathbf{E}^{\mathcal{L}}, \mathbf{F}^{\mathcal{L}}, \mathbf{e}_0^{\mathcal{L}}, \mathbf{e}_1^{\mathcal{L}}, \mathbf{e}_2^{\mathcal{L}}$ , see (4)
12:     $\mathbf{T}_-, d_- = \text{cal\_3D}(\mathbf{C}^{\mathcal{L}}, \mathbf{F}^{\mathcal{L}}, \mathbf{d}_2^{\mathcal{L}}, \mathbf{e}_2^{\mathcal{L}})$ , see (11)
13:    if  $d_- < d_{\min}$  then
14:       $\mathbf{T}_{li} = \mathbf{T}_-$ 
15:       $d_{\min} = d_-$ 
16:    end if
17:  end for
18:   $S.\text{append}(\mathbf{T}_{li})$ 
19: end for

```

\mathbf{e}_0 , and \mathbf{n} are also coplanar. That is

$$\begin{cases} (0, 0, 1) \cdot (\mathbf{T}\mathbf{A} \times \mathbf{d}_0) = 0 \\ (0, 0, 1) \cdot (\mathbf{T}\mathbf{D} \times \mathbf{e}_0) = 0. \end{cases} \quad (10)$$

Second, we assume that \mathbf{T} is the point that minimizes the sum d_- of the distances to line $(\mathbf{C}, \mathbf{d}_2)$ and line $(\mathbf{F}, \mathbf{e}_2)$. Combining with (10), the final 3-D point calculation model is as follows:

$$\begin{aligned} \mathbf{T}, d_- &= \arg \min_{\mathbf{T}} (|\mathbf{T}\mathbf{C} \times \mathbf{d}_2| + |\mathbf{T}\mathbf{F} \times \mathbf{e}_2|) \\ \text{s.t. } &\begin{cases} (\mathbf{T}\mathbf{A} \times \mathbf{d}_0)_{(3)} = 0 \\ (\mathbf{T}\mathbf{D} \times \mathbf{e}_0)_{(3)} = 0. \end{cases} \end{aligned} \quad (11)$$

Combining the above analysis, a simplified underwater point cloud acquisition algorithm is presented in Algorithm 1.

B. Motion Compensation

1) Problem Formulation: With the previously proposed point cloud acquisition algorithm, we can acquire the 3-D point cloud of the scene utilizing the self-developed scanning structured light system. Only one scan point cloud can be acquired in one computation, and the information acquired is limited. In order to carry out the subsequent point cloud registration process, we need to collect multiple scans to form a sweep point cloud, which retains more features of the scene. However, during the sweep formation, the coordinate system of each scan point cloud in the sweep will change due to the motion of the robot. Most of the existing work ignores this change, but it will undoubtedly bring immeasurable motion distortion to the sweep point cloud. After all, in underwater environments, the robot wobbles very randomly and violently. Therefore, it is worth considering how

to overcome motion distortion and achieve a more accurate stitching of the sweep point cloud. In our experiments, we treat the point cloud obtained by turning the mirror-galvanometer back and forth once as a sweep with about 80 scans. Suppose that for a sweep P , $P = \{s_0, \dots, s_n\}$. Where s_i denotes the i th scan acquired by the current sweep in time order. The task of motion compensation is to convert the i th ($i = 1, \dots, n$) scan in P into the coordinate system corresponding to scan s_0 .

2) Compensation Algorithm: Considering that the IMU can provide angular velocity information at higher frequencies, which is utilized by

$$\mathbf{R}_{\mathcal{T}_{i+1}}^{\mathcal{T}_i} = \text{Exp}((\tilde{\omega}_i^{\mathcal{T}} - \mathbf{b}_i^{\mathcal{W}} - \boldsymbol{\eta}_i^{\mathcal{W}}) \delta t) \quad (12)$$

where $\tilde{\omega}_i$ denotes the angular velocity measurement, \mathbf{b}_i is the bias, and $\boldsymbol{\eta}_i$ is the white noise.

It can be easily obtained the rotation matrix \mathbf{R} information. Generally the nine-axis IMU will have an extra magnetometer information used to correct the orientation information, so we think the \mathbf{R} obtained by the IMU is very accurate. Of course, it is difficult to get just the \mathbf{R} on its scanning moment for a certain s_i . But in the fact that the IMU is much faster than the scan, it can be obtained by interpolation. Since the precise interpolation of the rotation matrix is difficult, while the spherical interpolation of the quaternion is exact. Therefore, for the rotation matrices \mathbf{R}_m and \mathbf{R}_n before and after s_i , they are first transformed into the corresponding quaternions \mathbf{Q}_m and \mathbf{Q}_n by the following derivation.

Suppose that the rotation matrix \mathbf{R} is represented as

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}. \quad (13)$$

Based on the idea in [25], \mathbf{R} is transformed into a \mathbf{K}_3 -matrix

$$\mathbf{K}_3 = \frac{1}{3} \begin{bmatrix} a & r_{21} + r_{12} & r_{31} + r_{13} & r_{23} - r_{32} \\ r_{21} + r_{12} & b & r_{32} + r_{23} & r_{31} - r_{13} \\ r_{31} + r_{13} & r_{32} + r_{23} & c & r_{12} - r_{21} \\ r_{23} - r_{32} & r_{31} - r_{13} & r_{12} - r_{21} & d \end{bmatrix}. \quad (14)$$

Among them

$$\begin{cases} a = r_{11} - r_{22} - r_{33} \\ b = r_{22} - r_{11} - r_{33} \\ c = r_{33} - r_{11} - r_{22} \\ d = r_{11} + r_{22} + r_{33}. \end{cases}$$

Then, compute the eigenvector \mathbf{Q} of \mathbf{K}_3 that belongs to the eigenvalue 1. It is the sought quaternion of \mathbf{R} . For more other details, see [25].

The quaternion \mathbf{Q}_i corresponding to s_i is then obtained by interpolating \mathbf{Q}_m and \mathbf{Q}_n

$$\mathbf{Q}_i = \frac{\sin(1-i)\theta}{\sin\theta} \mathbf{Q}_m + \frac{\sin i\theta}{\sin\theta} \mathbf{Q}_n \quad (15)$$

where $i = \frac{t_{s_i} - t_{\mathbf{Q}_m}}{t_{\mathbf{Q}_n} - t_{\mathbf{Q}_m}}$.

Since the robot is relatively stable in its motion during the underwater reconstruction. Based on the fact that rotation brings abrupt changes in the point cloud position while translation

has little effect, we introduce a uniform velocity model for translation inspired by Zhang and Singh [26]. The translation matrix $\mathbf{t}_{(k+1,i)}^{\mathcal{L}} \in \mathbb{R}^{3 \times 1}$ of s_i is estimated as

$$\mathbf{t}_{(k+1,i)}^{\mathcal{L}} = \frac{t_i - t_0}{T_k} \mathbf{t}_k^{\mathcal{L}} \quad (16)$$

where $\mathbf{t}_k^{\mathcal{L}}$ denotes the translation matrix of the last motion estimate. T_k indicates the last sweep time span. $t_i - t_0$ denotes the time span of the s_i relative to s_0 .

At this time, for s_i ($i = 1, \dots, n$) in P , it can be converted to the s_0 coordinate system by the following formula:

$$s_{im}' = \mathbf{R}_{\mathcal{L}}^{\mathcal{S}}(\mathbf{Q}_0)^{-1} \mathfrak{R}(\mathbf{Q}_i) s_{im} + \mathbf{t}_{(k+1,i)}^{\mathcal{L}} \quad (17)$$

where $s_{im} \in \mathbb{R}^{3 \times 1}$ refers to any point in the scan s_i . $\mathbf{R}_{\mathcal{L}}^{\mathcal{S}} \in \mathbb{R}^{3 \times 3}$ is the rotation matrix corresponding to the coordinate transformation matrix $\mathbf{T}_{\mathcal{L}}^{\mathcal{S}} \in \mathbb{R}^{4 \times 4}$. $\mathfrak{R}(\cdot) \in \mathbb{R}^{3 \times 3}$ denotes the transformation of a quaternion $\mathbf{Q}(q_0, \mathbf{q}_v)$ into the corresponding rotation matrix \mathbf{R} . Specifically

$$\mathbf{R} = (q_0^2 - \|\mathbf{q}_v\|) \mathbf{I}_{3 \times 3} + 2\mathbf{q}_v \mathbf{q}_v^T + 2q_0 [\mathbf{q}_v]_{\times}. \quad (18)$$

C. Motion Estimation

1) Problem Formulation: Castillon et al. [27] in the latest research stated that pure rigid registration, such as ICP (used in [22]), has limitations in the application of LSLS due to the uncertainty of measurement errors. Therefore, the GICP [28] point cloud registration algorithm is introduced in Water-MBSL to achieve accurate motion estimation, which increased the robustness of the framework.

To balance the accuracy of point cloud registration and computational time consumption, the sliding window mechanism is applied to our Water-MBSL. Specifically, we synchronously update a local map composed of historical sweep of a window size as the registration object. In order to improve the computational efficiency and reduce the interference of noise, voxel filtering is used to downsample the point cloud. The downsampled point cloud is assumed to be

$$\begin{aligned} \text{sweep} : P &= \{p_i \in \mathbb{R}^3, i = 1, 2, 3, \dots, N\} \\ \text{local map} : Q &= \{q_i \in \mathbb{R}^3, i = 1, 2, 3, \dots, M\}. \end{aligned} \quad (19)$$

The purpose of motion estimation is to use the currently obtained sweep and historical local map to continuously optimize the transformation matrix between them so as to make the shortest distance between corresponding points, i.e.

$$\min_{\mathbf{T}} \sum_{i=1}^N \|q_i - \mathbf{T}p_i\|. \quad (20)$$

2) Object Function: The traditional ICP default point positions are uniquely determined. However, during the SSLSLS scanning, the measured values of each point in the point cloud actually deviate from the true values \hat{p}_i and \hat{q}_i due to the presence of measurement errors, etc. It is assumed that the Gaussian distribution is obeyed

$$p_i \sim \mathcal{N}(\hat{p}_i, C_i^P); \quad q_i \sim \mathcal{N}(\hat{q}_i, C_i^Q) \quad (21)$$

where C_i^P and C_i^Q are the covariance matrices, respectively.

Define $d_i^T = q_i - \mathbf{T}p_i$, which denotes the directed distance of the i th point pair after doing the transformation using \mathbf{T} for the original point cloud. Combining (21), we get

$$\begin{aligned} d_i^{(\mathbf{T})} &\sim \mathcal{N}(\hat{q}_i, C_i^Q) - \mathbf{T} \mathcal{N}(\hat{p}_i, C_i^P) \\ &= \mathcal{N}(\hat{q}_i - \mathbf{T}\hat{p}_i, C_i^Q + (\mathbf{T})C_i^P(\mathbf{T})^T). \end{aligned} \quad (22)$$

According to [28]

$$\mathbf{T} = \arg \max_{\mathbf{T}} \prod_i p(d_i^{(\mathbf{T})}) = \arg \max_{\mathbf{T}} \sum_i \log(p(d_i^{(\mathbf{T})})). \quad (23)$$

Combining the definition of the multivariate Gaussian function, it is obtained that

$$\begin{aligned} \log(p(d_i^{(\mathbf{T})})) &= \log\left(\frac{1}{\sqrt{(2\pi)^k |C_i^B + \mathbf{T}C_i^A\mathbf{T}^T|}}\right) \\ &\quad - \frac{1}{2}(M)^T (C_i^B + \mathbf{T}C_i^A\mathbf{T}^T)^{-1} M \end{aligned} \quad (24)$$

where $M = d_i^{(\mathbf{T})} - (\hat{q}_i - \mathbf{T}\hat{p}_i)$. For the true values \hat{p}_i and \hat{q}_i , there are $\hat{q}_i = \mathbf{T}^* \hat{p}_i$, where \mathbf{T}^* is the ideal correct rigid transform. In M , $\hat{q}_i - \mathbf{T}^* \hat{p}_i = 0$ when $\mathbf{T} = \mathbf{T}^*$. Equation (24) can be reduced to

$$\begin{aligned} \log(p(d_i^{(\mathbf{T})})) &= \log\left(\frac{1}{\sqrt{(2\pi)^k |C_i^B + \mathbf{T}C_i^A\mathbf{T}^T|}}\right) \\ &\quad - \frac{1}{2} d_i^{(\mathbf{T})^T} (C_i^B + \mathbf{T}C_i^A\mathbf{T}^T)^{-1} d_i^{(\mathbf{T})}. \end{aligned} \quad (25)$$

Since $\det(\mathbf{T}) = 1$, then $\det(\mathbf{T}C_i^A\mathbf{T}^T) = \det(C_i^A)$. Besides, for matrices \mathbf{A} and \mathbf{B} , the following theorem exists:

$$\det(\mathbf{A} + \mathbf{B}) = \det \mathbf{A} + \det \mathbf{A} \sum_{k=1}^{n-1} \mathbf{B}_{ik+1}^{[ik+1]} \dots \mathbf{B}_{in}^{[in]} + \det \mathbf{B}. \quad (26)$$

So $|C_i^B + \mathbf{T}C_i^A\mathbf{T}^T|$ is independent of \mathbf{T} . The first term of (25) is a constant and can be ignored.

In summary, the objective function is

$$\mathbf{T} = \arg \min_{\mathbf{T}} \sum_i d_i^{(\mathbf{T})^T} (C_i^B + \mathbf{T}C_i^A\mathbf{T}^T)^{-1} d_i^{(\mathbf{T})}. \quad (27)$$

3) Optimal Solution: For a better optimization solution, we introduce the Lie algebra for \mathbf{T} to transform the problem into an unconstrained optimization problem. Combined with the Gaussian Newton optimization algorithm [29], the previous object function is optimized to get the best motion estimate.

D. Map Reconstruction

In map reconstruction module, the key frame selection mechanism and sliding window model are deployed in our framework. When the motion estimation is completed, the current sweep is judged to be a key frame or not. After passing it, it will be added

to the global map. The local map is also updated for the next motion estimation. Among them, the key frame selection and local map update details are as follows.

1) **Key Frame Selection:** In order to improve the reconstruction efficiency of the onboard system, the keyframe selection mechanism is applied, which can effectively avoid double calculations. In our Water-MBSL, the keyframe selection mechanism is as follows.

- 1) Whether the number of valid 3-D points in the current sweep is greater than n .
- 2) Whether the translational distance in the motion estimation is greater than d .

In the later experiments, n is chosen as 10 000 and d as 0.05.

2) **Local Map Update:** In order to improve the point cloud registration accuracy, we adopt the sweep to map registration strategy. However, the global map will delay the alignment speed significantly. Therefore, the sliding windows technique is applied, which means that each sweep is used to do motion estimation with the local map inside the window. In addition, to reduce memory consumption, our map data are stored in the hard disk, which greatly reduces the load on the onboard controller.

IV. EXPERIMENTS

To verify the effectiveness of our Water-MBSL, experimental system and scenario are built first. Then, the developed underwater movable reconstruction platform is fixed and static scanning experiments are performed. Furthermore, self-stabilizing scanning reconstruction experiments and motion scanning reconstruction experiments in different scenes and at different speeds are carried out, respectively. In addition, the underwater engineering applications of Water-MBSL in general cases are validated, including ship oar reconstruction and pipe inspection. Finally, the differences between state-of-the-art LSLS-based motion reconstruction frameworks and our framework are discussed.

A. Experiment Environment

Underwater motion reconstruction has a significant research value in marine civilization exploration. Especially for the field of underwater archaeology, high-precision dense maps could provide more potential for the identification and grasping of artifacts. To verify the performance of Water-MBSL, we build a simulated underwater archaeological operation scenario in a $5\text{ m} \times 4\text{ m} \times 1.5\text{ m}$ pool, as shown in Fig. 4(a). Among them, some clay pots are used to mimic underwater artifacts as the reconstruction targets of our underwater movable reconstruction platform. The global camera is used to observe the motion of the platform as a whole. In addition, we introduce a common commercial DVL product WaterLinked A50 for comparison experiments, as shown in Fig. 4(b). It can provide DR results at about 3 Hz, which can be used as the motion estimation of the robot in [19], [22], and [23].

B. Static Scanning Reconstruction

Before performing the motion reconstruction, we verify the reconstruction performance in a static environment. In this case,

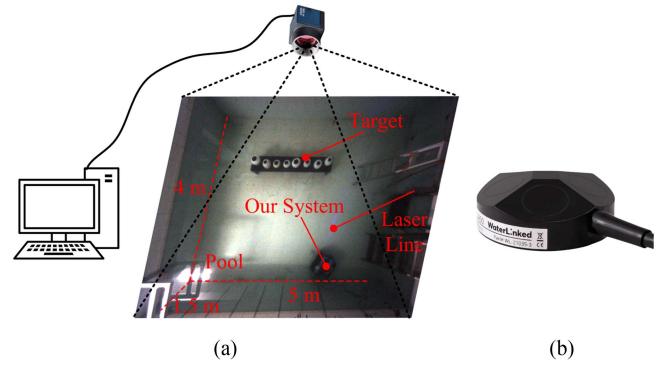


Fig. 4. Experiment environment. (a) Pool. (b) DVL.

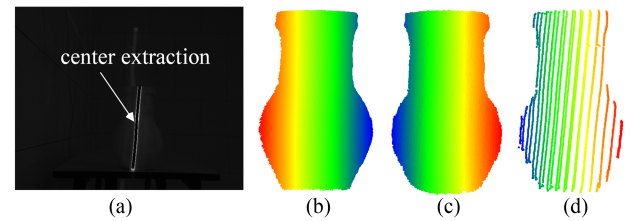


Fig. 5. Static scanning reconstruction results for a clay pot. (a) Laser centerline extraction. (b) Ground truth. (c) Multiple sweeps. (d) One sweep.

the underwater movable reconstruction platform is held in place by a stand and remains completely stationary to scan reconstruction targets. Two sets of experiments are implemented. In the first set of experiments, the target is a clay pot with the aim of verifying the overall reconstruction effect. The low-light scene in this case imitates the real-world dim conditions underwater in nature. Fig. 5(a) shows the results of laser centerline extraction, which verifies the validity of (1) (in Section III).

Fig. 5(b) is the reconstruction result of the clay pot in air by a commercially available high-precision depth camera that is used as the ground truth. The reconstruction results of our self-designed SSLSL system are shown in Fig. 5(c) and (d), where Fig. 5(c) is the result of multiple sweeps and Fig. 5(d) is the reconstruction result of one sweep. In the subsequent moving reconstruction process, we use one sweep information each time. In order to quantitatively compare the accuracy of our system, we make an analysis of Fig. 5(b) and (c) with ICP. The root mean square error (RMSE) between our static scan results and the ground truth is only 0.23 cm, and the average (Ave) distance (Dist) error is 0.13 cm. This shows that our system is sufficiently suited for underwater high-precision dense reconstruction. In addition, the second set of experimental subjects is a hemisphere in order to test the distance measurement error of our SSLSL system. As shown in Fig. 6(a), the true radius R of the hemisphere is 4.96 cm. Its reconstruction result is shown in Fig. 6(b). In this case, the absolute distance between the plane and the top of the sphere is rendered by the color. To compare the results more accurately, we split down the point cloud of the hemisphere and then align it for sphere fitting, as shown in Fig. 6(c). The fitted radius \bar{R} is 5.07 cm, close to the true radius. The error is only 0.11 cm. The static scanning experiments illustrate that our self-developed structured light

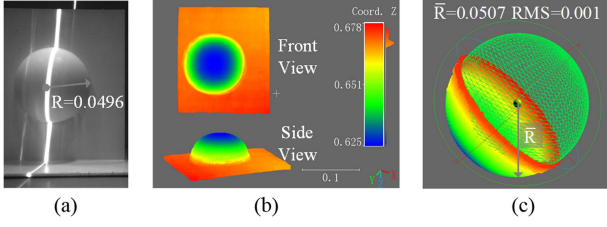


Fig. 6. Static scanning reconstruction results for a hemisphere. (a) Image of the object. (b) Acquired point cloud. (c) Spherical fitting.

system combined with the proposed point cloud acquisition algorithm lays a deep foundation for motion reconstruction.

C. Self-Stabilizing Scanning Reconstruction

For underwater fields, such as underwater archaeology, it is common to scan a scene for a long time in order to obtain dense information from a single viewpoint. Therefore, in this experiment, we make BlueROV turn on the fixed-point self-stabilization mode and scan the underwater clay pot, as shown in Fig. 5(a), for upto 25 s. As can be expected, it is impossible for an underwater robot to remain completely stationary. Using the global camera, we observe that the robot reaches a movement of about 20 cm in 25 s. Four sets of comparison experiments are implemented in order to verify the superiority of our Water-MBSL. In the first set of experiments, we do not use any processing on the point cloud and consider the robot to remain completely stationary as assumed in [22]. The second set of experiments uses the motion state information provided by DVL to directly splice the obtained scan point cloud each time, as in [19]. Using the idea in [23], the motion distortion corrected sweeps are stitched using the motion state information provided by DVL in the third set of experiments. Our Water-MBSL framework is applied in the fourth set of experiments to implement scene dense reconstruction, without any DVL information being used in this process. It is worth stating that a completely consistent replication is difficult due to the differences in experimental environments, equipment, and objectives, so we just apply their partial main ideas for the comparison experiments.

In Fig. 7, we list the reconstruction results at 2, 10, and 25 s. It can be clearly seen that as the scan time increases, the reconstructed point cloud ghosting of other methods, except our Water-MBSL, becomes more and more severe. In order to quantitatively compare the reconstruction results, the reconstructed point clouds at different times are compared with the ground truth in Fig. 5(b) using ICP, respectively. The results of the quantitative comparison are given in Table II, which illustrates the reconstruction performance in terms of RMSE and Ave Dist. In order to show the relationship between reconstruction errors and reconstruction times more intuitively, the error statistics are plotted, as shown in Fig. 8. It can be seen that with the increase of the scan time, the errors of method [22] increase rapidly, with both RMSE and Ave errors reaching the centimeter level at 25 s. It is because that for self-stabilizing motion we cannot simply assume the robot to be completely stationary, which introduces

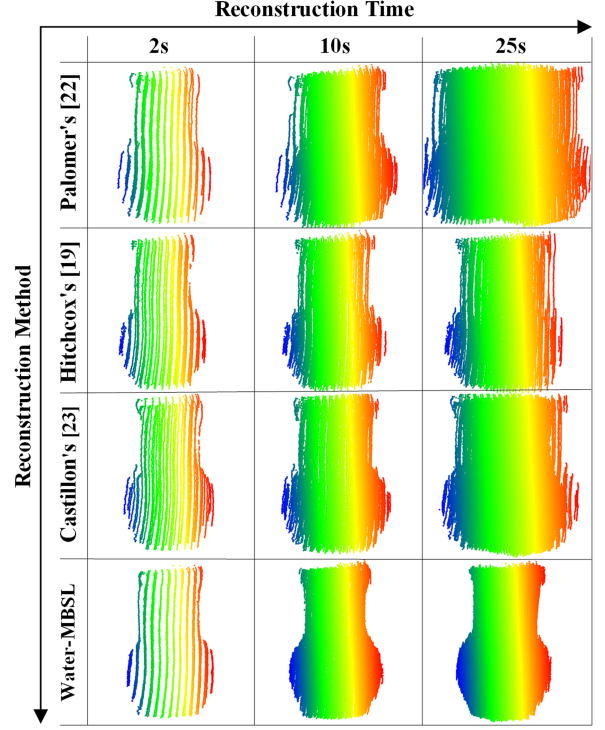


Fig. 7. Self-stabilizing scanning reconstruction results at different times.

TABLE II
QUANTITATIVE COMPARISON OF SELF-STABILIZING SCANNING RECONSTRUCTION RESULTS

Time (s)	Offset Dist (cm)	RMSE (cm)				Ave Error (cm)			
		[22]	[19]	[23]	Water-MBSL	[22]	[19]	[23]	Water-MBSL
2	0.95	0.42	0.25	0.24	0.27	0.24	0.14	0.13	0.14
5	3.87	0.55	0.29	0.32	0.28	0.29	0.17	0.17	0.13
10	8.22	0.70	0.39	0.38	0.33	0.38	0.23	0.22	0.17
15	13.15	1.16	0.45	0.56	0.35	0.68	0.26	0.29	0.18
20	18.35	1.92	0.52	0.61	0.39	1.26	0.29	0.33	0.23
25	23.56	2.40	0.65	0.83	0.37	1.63	0.36	0.45	0.19

The bold entities represent optimal performance metric in each row.

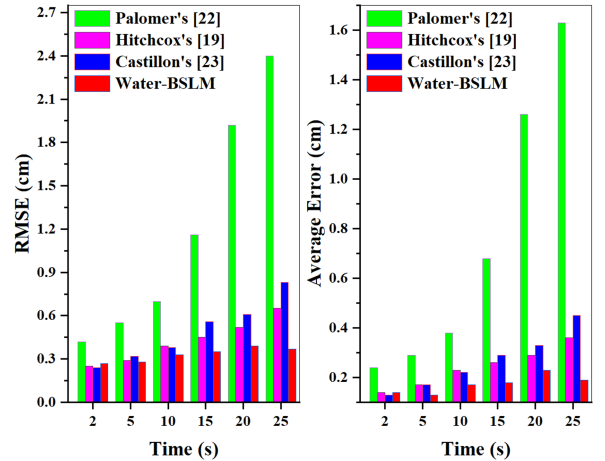


Fig. 8. Self-stabilizing scanning reconstruction errors.

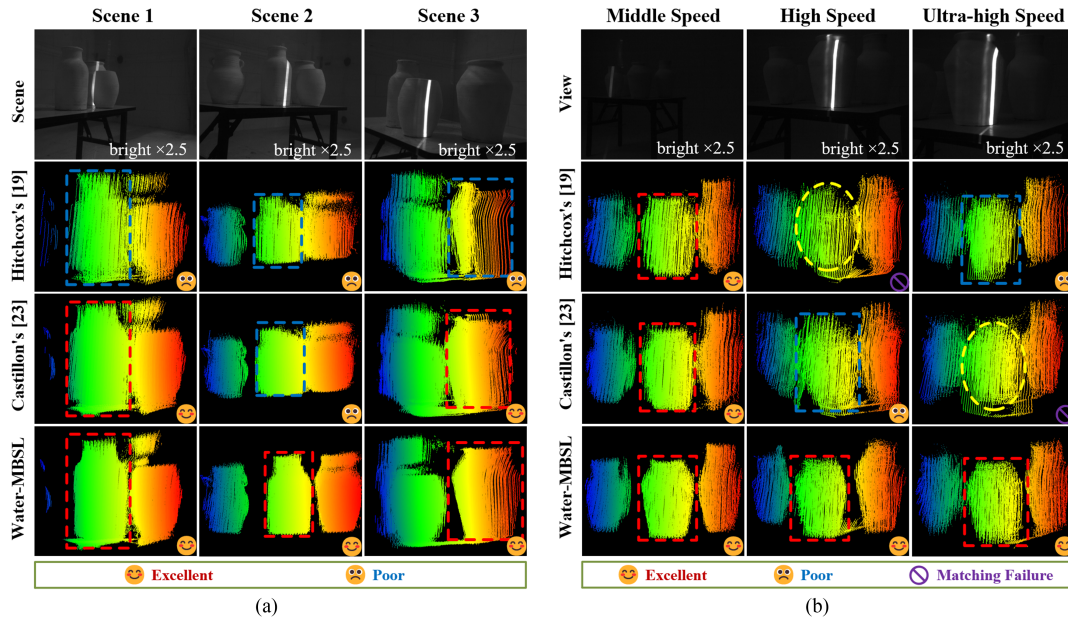


Fig. 9. Moving scanning reconstruction results. (a) Moving scanning reconstruction results in different scenes. (b) Moving scanning reconstruction results at different speeds. Note: The objects in the rectangular box are the ones selected for matching.

immeasurable errors. Methods in [19] and [23] have a slightly slower increase in error, but 25 s is also significantly higher compared with 2 s. The RMSE of the former is around 8 mm and the Ave Dist error is more than 4 mm. The RMSE of the latter reaches 6 mm and the Ave error is more than 3 mm. This suggests that relying solely on motion state estimates provided by external acoustic sensors can suffer from significant error accumulation. Moreover, the limited accuracy of the acoustic sensors does not allow for high-precision reconstruction. In contrast, the reconstruction effect of our Water-MBSL has little effect with increasing time. The difference between the indices at 25 s and 2 s was not significant. Our RMSE is around 3 mm and the Ave error is stable at around 2 mm. The accuracy is remarkable in the underwater field, mainly due to the point cloud matching that improves the consistency of the reconstructed point cloud.

D. Motion Scanning Reconstruction

Underwater motion reconstruction is important for improving the efficiency of underwater operations. Water-MBSL, as an SSLSLS-based motion reconstruction framework, is the first work to achieve underwater movable high-precision dense reconstruction without relying on expensive DVL. Two different groups of variable experiments are deployed to illustrate its performance in different scenes and at different speeds. Note: $\text{bright} \times 2.5$ in all images means that the image brightness is expanded by a factor of 2.5, which is just for a better visual experience and has nothing to do with the experiment itself.

1) *Motion Scanning Reconstruction in Different Scenes:* To test the performance of our Water-MBSL reconstruction while moving, the system is manipulated to reconstruct three scenes shown in the first row of Fig. 9(a). The movement distance of

each scene is around 1 m, which is consistent with the range of realistic underwater archaeological refinement reconstruction. Due to the inability of [22] to achieve reconstruction while moving, no experiments are conducted this time. Three sets of comparison experiments between method in [19], method in [23], and our Water-MBSL are implemented in order to verify the motion reconstruction performance. The reconstruction times for the three scenes are 40, 35, and 35 s, respectively. It is worth noting that the robot is moving at low speed in this group of experiments, roughly 0.02–0.03 m/s, which is the normal speed for underwater robotic motion reconstruction [23]. The reconstruction results of the three scenes are shown in Fig. 9(a). In terms of overall intuition, we find that the point cloud consistency by Water-MBSL is significantly better than the other two. The second row of Fig. 9(a) shows the results of method in [19], where the point cloud has obvious distortion problems and severe ghosting. The point cloud obtained by method in [23] has been significantly improved, but the refinement is still insufficient, as shown in the third row of Fig. 9(a). The fourth row of Fig. 9(a) is the result of our Water-MBSL, which achieves the best performance compared with the first two methods, although no DVL information is introduced. From a localized view, the detailed features of the point clouds in our method are much more obvious, such as at the mouth of the clay pot in ①. Moreover, the point clouds by methods in [19] and [23] have obvious dispersion phenomenon, and the point clouds of two neighboring target objects are almost connected in ② and ③, while our point clouds are clearly distinguishable in ④. In addition, we categorize all the obtained point clouds into Excellent and Poor based on the objective evaluation of the point clouds.

In order to quantitatively compare the effects of the three methods of motion reconstruction, one of the clay pots within the

TABLE III

QUANTITATIVE COMPARISON OF MOVING SCANNING RECONSTRUCTION RESULTS IN DIFFERENCE SCENES

Scene	Move Dist (cm)	Motion Time (s)	RMSE (cm)			Ave Error (cm)		
			[19]	[23]	Water-MBSL	[19]	[23]	Water-MBSL
1	70.8	40	2.46	2.24	0.87	1.94	1.85	0.62
2	102.1	35	2.20	1.74	1.18	1.59	1.22	0.89
3	114.8	35	1.53	1.36	0.66	1.12	0.97	0.46

The bold entities represent optimal performance metric in each row.

three scenes are randomly selected, as shown in the rectangular box in Fig. 9. Similar quantitative experiments-based ICP are conducted as in Section IV-C, and the results are given in Table III. In the motion reconstruction, the effect of motion distortion is serious compared with the self-stabilizing scanning reconstruction. So, the method in [23] considers that the motion distortion is better than the method in [19] from the results. However, the RMSE of both methods are in the centimeter level, with the former exceeding 2 cm and the latter around 1.5 cm. The Ave error is also more than 1 cm. The results of our Water-MBSL reconstruction show that the RMSE is generally less than 1 cm, and the Ave error is even around 5 mm, reaching the current state-of-the-art level of underwater motion reconstruction.

2) *Motion Scanning Reconstruction at Different Speeds:* In previous experiments, it has been demonstrated that the Water-MBSL has excellent motion reconstruction performance at low speed. To test the robustness of our Water-MBSL reconstruction while moving, a total of three sets of experiments at different higher speed levels are conducted, i.e., middle, high, and ultrahigh speeds. The reconstruction scene with three views is shown in the first row of Fig. 9(b). In the middle-speed level experiments, the robot has an average speed of 0.061 m/s and a maximum (Max) speed of 0.079 m/s. In the second set of experiments, the robot is moving at an average speed of 0.116 m/s and a maximum speed of 0.140 m/s, which is considered as a high-speed level. In the ultrahigh speed experiments, the average and maximum speeds of the robot even reach 0.186 and 0.192 m/s, respectively, which are very challenging speeds for dense underwater reconstruction. The reconstruction results of the three speed levels in different methods mentioned in Section IV-D1 are shown in Fig. 9(b).

By analyzing the acquired point cloud, what can be seen is that all three methods are able to reconstruct the scene morphology at middle speeds, but the reconstructions of method in [19] and method in [23] almost fail at high and ultrahigh speeds. This is due to the fact that when the robot is moving underwater at high speed, the shaking will be very violent and frequent, which will affect the accuracy of the DVL data. Our method does not depend on DVL, and can still reconstruct good morphological features at high and ultrahigh speeds. Similar to Section IV-D1, an ICP-based quantitative analysis is performed, as given in Table IV. The analysis results informed that at high speeds, the robot's wobbling caused the performance of the DVL-dependent reconstruction method to degrade drastically. The acquired point cloud fails to match the ground truth. In contrast, the point cloud acquired by our Water-MBSL maintains better features.

TABLE IV

QUANTITATIVE COMPARISON OF MOVING SCANNING RECONSTRUCTION RESULTS AT DIFFERENT SPEEDS

Speed Level	Ave Speed (m/s)	Max Speed (m/s)	RMSE (cm)			Ave Error (cm)		
			[19]	[23]	Water-MBSL	[19]	[23]	Water-MBSL
Middle	0.061	0.079	2.02	1.82	1.32	1.43	1.26	0.98
High	0.116	0.140	×	3.01	1.72	×	2.20	1.34
Ultra High	0.186	0.192	2.36	×	1.98	1.71	×	1.51

×: Point cloud matching failure.

The bold entities represent optimal performance metric in each row.

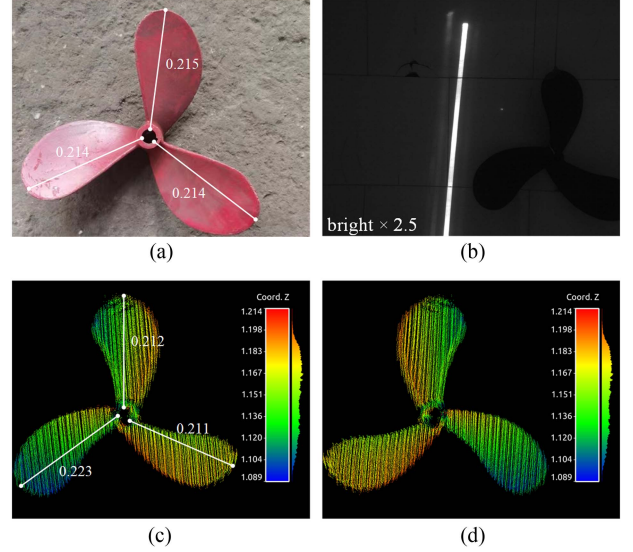


Fig. 10. Moving scanning reconstruction results for ship oar. (a) Target photo. (b) Underwater view of left camera in structured light system. (c) Reconstructed point cloud (front). (d) Reconstructed point cloud (back). Note: All lengths are in meters.

Of course, with the increase in speed, it inevitably brings about a larger error than that in the low-speed motion reconstruction. However, the average error remains stable at no more than 2 cm, which is already rare underwater.

E. Underwater Engineering Applications of Water-MBSL in General Cases

In the previous experiments, the excellent performance of Water-MBSL in artifact reconstruction for underwater archaeology has been verified. To further validate Water-MBSL's capability for underwater engineering applications in other general cases, underwater reconstruction experiments of common man-made infrastructures are deployed, including ship oar reconstruction and pipe inspection.

1) *Engineering Applications in Ship Oar Reconstruction:* Ship oar reconstruction plays an important role in certain underwater wreck salvage operations. As shown in Fig. 10(a), a 12-HP three-bladed ship oar with a diameter of about 0.45 m is used as the Water-MBSL motion reconstruction target in this experiment. The blade lengths of the oar are 0.215, 0.214, and 0.214 m. It is placed underwater with severe light deprivation, as shown in Fig. 10(b). We then reconstruct it using the self-developed movable reconstruction platform. The reconstruction

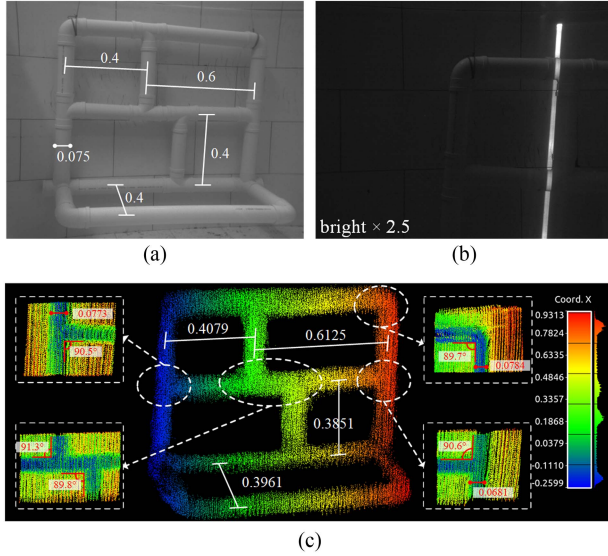


Fig. 11. Moving scanning reconstruction results for pipe. (a) Target photo. (b) Underwater view of left camera in structured light system. (c) Reconstructed point cloud. Note: All lengths are in meters.

results are shown in Fig. 10(c) and (d), where Fig. 10(c) is the front-side point cloud rendering result and Fig. 10(d) is the reverse side point cloud rendering result. What can be seen is that the reconstructed point clouds basically maintain the external features of the oar with a clear spiral phenomenon. Then, using multiple measurements to take the average value, we roughly obtained the measured size of the blades by point clouds, which are 0.212, 0.223, and 0.211 m, respectively. The measured paddle dimensions are very close to the actual dimensions with a maximum error of no more than 0.01 m. The experimental results show that our Water-MBSL can not only reconstruct the oar morphology features well, but also has a very high measurement accuracy.

2) Engineering Applications in Pipe Inspection: As an important route for the transportation of human underwater resources, pipe inspection is the key to maintaining long-term safe operation. In this experiment, Water-MBSL performs motion reconstruction of a 1.08 m × 0.88 m × 0.44 m pipe structure similar to [22], as shown in Fig. 11(a). The diameter of the pipe is 0.075 m, and four key dimensions are measured. Fig. 11(b) shows the scene when it is reconstructed underwater by Water-MBSL. The results of the motion reconstruction are displayed in Fig. 11(c). The overall view of the acquired point cloud shows that the basic morphology of the pipe is consistent with that in Fig. 11(a). Similarly, four critical dimensions shown in Fig. 11(a) are measured in the point cloud. From the results, it can be seen that the measured dimensions are very similar to the real dimensions with an average error of 0.98 cm. In addition, to demonstrate the local structural features of the reconstruction, four regions with vertical structure are selected and the point cloud structure in them is more fully analyzed, as shown in Fig. 11(c)①–④. It is found that the angle measurements are basically around 90°, which indicate that Water-MBSL can maintain the structural relationship between the parts well in

TABLE V
COMPARISON OF STATE-OF-THE-ART METHODS FOR UNDERWATER STRUCTURED LIGHT RECONSTRUCTION

Characteristics	[24]	[20], [21]	[23]	[25]	Water-MBSL
movable reconstruction [†]	x	✓	✓	✓	✓
motion compensation	x	—	—	✓	✓
self-motion estimation [‡]	✓	x	x	x	✓
self-scanning	✓	x	—	✓	✓
low-light	✓	✓	x	✓	✓
low cost	x	x	✓	x	✓

[†]: Reconstruction while moving, without pausing.

[‡]: No reliance on external sensors for motion estimation.

the motion reconstruction. This is of great significance for the morphological inspection of artificial structures. In addition, the average error of diameter measurement for three parts of the pipe is only 0.42 cm. This suggests that compared with static reconstruction, motion reconstruction could cause greater measurement error in the reconstruction results due to uncertainty in motion estimation. However, our Water-MBSL has greatly mitigated this phenomenon.

F. Discussion

The above experiment results indicate that the proposed Water-MBSL achieves mm-level reconstruction accuracy in static scanning reconstruction experiments, long-time self-stabilizing scanning reconstruction experiments, and reconstruction-while-motion experiments. The ship oar reconstruction and pipe inspection experiments demonstrate that our Water-MBSL has significant engineering application capabilities in general cases. The better underwater motion reconstruction performance compared with existing frameworks is mainly due to the fact that our framework takes into account both motion distortion and point cloud registration.

For now, underwater motion dense reconstruction based on LSLS is still a relatively novel field. Here a qualitative comparison is made between our framework and the latest research work in the past five years, as given in Table V. In contrast to the zero-velocity control in [22], our Water-MBSL allows reconstruction of underwater scenes while in motion. Compared with the SLLSLS in [18] and [19], the SLLSLS based on mirror-galvanometer we developed can obtain richer scene information. Bräuer-Burchardt et al. [21] used monocular and IMU to achieve motion estimation, but this approach is very difficult to work in the dim underwater. Our framework enables self-motion estimation without relying on the motion estimation provided by expensive DVL, which means that the deployment cost and cumulative error of our framework is far lower compared with [23]. Overall, our framework has excellent performance in underwater motion reconstruction with significant applications in underwater refinement exploration.

V. CONCLUSION

In this article, we propose an underwater movable binocular structured light-based high-precision dense reconstruction

framework, named Water-MBSL. Moreover, an underwater movable reconstruction platform is developed in parallel to verify its superior performance in motion reconstruction. Compared with existing frameworks, we consider both motion distortion and point cloud alignment. The reconstruction is achieved while the underwater robot moves without relying on other sensors, such as DVL. Our framework is better at reconstructing dense point clouds, but it is suitable for structured scenes with distinct features. In some structurally degraded regions, such as smooth underwater boulder regions, sufficient point cloud features may not be extracted leading to alignment failure. In addition, the structured light-based measurement may not be able to achieve accurate extraction of the laser center in some reflective and transparent areas, thus failing to obtain accuracy 3-D point cloud. Therefore, integrating additional sensors, such as pressure sensors and sonar, is crucial for enhancing the robustness of the framework. In the future, we intend to incorporate more simultaneous localization and mapping techniques to continuously optimize the motion reconstruction performance of Water-MBSL for more underwater operation scenarios, such as uncharted underwater exploration. We sincerely hope that Water-MBSL can serve as a basic framework to bring more researchers with new inspirations to cope with the complex and changing underwater environment.

REFERENCES

- [1] I. Jawhar, N. Mohamed, J. Al-Jaroodi, and S. Zhang, "An architecture for using autonomous underwater vehicles in wireless sensor networks for underwater pipeline monitoring," *IEEE Trans. Ind. Inform.*, vol. 15, no. 3, pp. 1329–1340, Mar. 2019.
- [2] W. Li et al., "Visual and intelligent identification methods for defects in underwater structure using alternating current field measurement technique," *IEEE Trans. Ind. Inform.*, vol. 18, no. 6, pp. 3853–3862, Jun. 2022.
- [3] Y. Wang, R. Wang, S. Wang, M. Tan, and J. Yu, "Underwater bioinspired propulsion: From inspection to manipulation," *IEEE Trans. Ind. Electron.*, vol. 67, no. 9, pp. 7629–7638, Sep. 2020.
- [4] M. Johnson-Roberson et al., "High-resolution underwater robotic vision-based mapping and three-dimensional reconstruction for archaeology," *J. Field Robot.*, vol. 34, no. 4, pp. 625–643, 2017.
- [5] R. Wang, S. Wang, Y. Wang, M. Cai, and M. Tan, "Vision-based autonomous hovering for the biomimetic underwater robot—RobCutt-II," *IEEE Trans. Ind. Electron.*, vol. 66, no. 11, pp. 8578–8588, Nov. 2019.
- [6] Y. Wang et al., "Robust AUV visual loop-closure detection based on variational autoencoder network," *IEEE Trans. Ind. Inform.*, vol. 18, no. 12, pp. 8829–8838, Dec. 2022.
- [7] J. Fan et al., "An initial point alignment and seam-tracking system for narrow weld," *IEEE Trans. Ind. Inform.*, vol. 16, no. 2, pp. 877–886, Feb. 2020.
- [8] W. Chen, X. Liu, C. Ru, and Y. Sun, "Automated exposures selection for high dynamic range structured-light 3D scanning," *IEEE Trans. Ind. Electron.*, vol. 70, no. 7, pp. 7428–7437, Jul. 2023.
- [9] 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, Jan. 9, 2023, Art. no. 5004418, doi: [10.1109/TIM.2023.3235420](https://doi.org/10.1109/TIM.2023.3235420).
- [10] M. Castillón, A. Palomer, J. Forest, and P. Ridao, "State of the art of underwater active optical 3D scanners," *Sensors*, vol. 19, no. 23, 2019, Art. no. 5161.
- [11] G. Matos, P. Buschinelli, and T. Pinto, "Underwater laser triangulation sensor model with flat refractive interfaces," *IEEE J. Ocean. Eng.*, vol. 45, no. 3, pp. 937–945, Jul. 2020.
- [12] X. Wang et al., "An underwater structured light vision calibration method considering unknown refractive index based on Aquila optimizer," *IEEE Trans. Instrum. Meas.*, vol. 72, Dec. 27, 2022, Art. no. 9500812, doi: [10.1109/TIM.2022.3232671](https://doi.org/10.1109/TIM.2022.3232671).
- [13] C. Gu, Y. Cong, and G. Sun, "Three birds, one stone: Unified laser-based 3-D reconstruction across different media," *IEEE Trans. Instrum. Meas.*, vol. 70, Nov. 20, 2020, Art. no. 5004712, doi: [10.1109/TIM.2020.3039641](https://doi.org/10.1109/TIM.2020.3039641).
- [14] C. Gu et al., "MedUCC: Medium-driven underwater camera calibration for refractive 3-D reconstruction," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 52, no. 9, pp. 5937–5948, Sep. 2022.
- [15] Q. Xue, Q. Sun, F. Wang, H. Bai, B. Yang, and Q. Li, "Underwater high-precision 3D reconstruction system based on rotating scanning," *Sensors*, vol. 21, no. 4, 2021, Art. no. 1402.
- [16] A. Palomer, P. Ridao, J. Forest, and D. Ribas, "Underwater laser scanner: Ray-based model and calibration," *IEEE/ASME Trans. Mechatron.*, vol. 24, no. 5, pp. 1986–1997, Oct. 2019.
- [17] M. Castillón, J. Forest, and P. Ridao, "Underwater 3D scanner to counteract refraction: Calibration and experimental results," *IEEE/ASME Trans. Mechatron.*, vol. 27, no. 6, pp. 4974–4982, Dec. 2022.
- [18] 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.
- [19] 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.
- [20] T. Qin, P. Li, and S. Shen, "VINS-Mono: A robust and versatile monocular visual-inertial state estimator," *IEEE Trans. Robot.*, vol. 34, no. 4, pp. 1004–1020, Aug. 2018.
- [21] C. Bräuer-Burchardt et al., "Underwater 3D scanning system for cultural heritage documentation," *Remote Sens.*, vol. 15, no. 7, 2023, Art. no. 1864.
- [22] 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.
- [23] M. Castillón, R. Pi, N. Palomeras, and P. Ridao, "Extrinsic visual-inertial calibration for motion distortion correction of underwater 3D scans," *IEEE Access*, vol. 9, pp. 93384–93398, 2021.
- [24] 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.
- [25] I. Y. Bar-Itzhack, "New method for extracting the quaternion from a rotation matrix," *J. Guid., Control, Dyn.*, vol. 23, no. 6, pp. 1085–1087, 2000.
- [26] J. Zhang and S. Singh, "LOAM: LiDAR odometry and mapping in real-time," in *Proc. Robot.: Sci. Syst.*, 2014, pp. 1–9.
- [27] M. Castillón, P. Ridao, R. Siegwart, and C. Cadena, "Linewise non-rigid point cloud registration," *IEEE Robot. Autom. Mag.*, vol. 7, no. 3, pp. 7044–7051, Jul. 2022.
- [28] A. Segal, D. Haehnel, and S. Thrun, "Generalized-ICP," in *Proc. Robot.: Sci. Syst.*, 2009, Art. no. 435.
- [29] F. D. Foresee and M. T. Hagan, "Gauss-Newton approximation to Bayesian learning," in *Proc. IEEE Int. Conf. Neural Netw.*, 1997, pp. 1930–1935.



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

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 State Key Laboratory of Management and Control 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 State Key Laboratory of Management and Control for Complex Systems, IACAS. His research interests include underwater robot and bionic robot.



Min Tan received the B.S. degree in automation 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 State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. His research interests include advanced robot control, biomimetic

robot, and multirobot systems.



Long Cheng (Fellow, IEEE) received the B.S. degree in control engineering from Nankai University, Tianjin, China, in 2004, and the Ph.D. degree in control theory and control engineering from the Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2009.

He is currently a Full Professor with the Institute of Automation, Chinese Academy of Sciences. He is also an Adjunct Professor with the University of Chinese Academy of Sciences, Beijing. His current research interests include

rehabilitation robots, intelligent control, and neural networks.