AUV Docking Method in a Confined Reservoir with Good Visibility

Mingwei Lin [1] · Canjun Yang [1,2,3]

Abstract

Underwater docking technology enables autonomous underwater vehicles (AUVs) to execute long-term observation missions by periodically recovering and recharging AUVs. The conventional AUV homing and docking operations utilize acoustic and optical sensors at different ranges relative to the docking station. However, this method cannot perform perfectly in confined water regions because of the acoustic reflection and multipath effect. Thus, this paper proposes a novel navigation system, which fuses downward-looking visual odometry and model-based velocity for homing, and recognizes and tracks the light marker for terminal docking, in order to overcome the defects of the conventional navigation method. The reservoir experiment result verifies the effectiveness of the proposed method and shows good potential to extended applications in underwater routine cruise.

Keywords Autonomous underwater vehicle (AUV) · Docking technology · Data fusion · Visual guidance

1 Introduction

Autonomous underwater vehicles (AUVs) are playing an increasingly important role in ocean exploration and observation. With the technological advances, the application scenarios of AUVs are not limited to the open oceans and can also be extended to some confined water regions, such as the pool, reservoir and the aquarium. The power-intensive instrumentation consumes much energy of AUVs and limits their working periods. This means that for nearly all the missions, the AUV should be attended by a surface vessel or shoreside staff, increasing much cost and inconvenience for missions.

Underwater docking technology enables AUVs to be autonomously guided to the docking station, to recharge the batteries, download data, upload a new mission plan, and safely park until the next mission [1, 2]. The navigation process of AUV docking can be divided into homing and docking stages [3]. The former stage refers to the case where AUV is over 10–25 m away from the docking station (DS) and the later stage refers to the case where the distance between the

AUV and the DS is less than 10 m. The ultra-short base line (USBL) positioning system is commonly used for the homing stage because of its error-bounded performance for long-range localization. During the docking stage, the optical guidance [4], camera-based visual guidance [5], and electromagnetic guidance [6] are preferable choices due to their higher localization accuracy. However, in the confined water regions, the performance of the USBL positioning system is remarkably degraded caused by the acoustic reflection and multipath effects, thereby causing intermittent position fixes and many outlier measurements. Additionally, the doppler velocity log (DVL) generates large-noise velocity measurements when the AUV is close to the bottom. Thus, it is not an appropriate choice to use acoustic sensors in confined underwater regions.

Visual simultaneous localization and mapping (SLAM) are popular methods for solving on-land navigation problems [7]. SLAM techniques build a map of an unknown environment firstly and then localize the robot in the map, whereby the localization error of the robot can be bounded within a limited range. One notable defect of the SLAM methods is the large map information that should be stored for loop closure and trajectory optimization. Once the features cannot be matched from two consecutive frames, the map error would increase obviously. The AUV hardly ever revisits the same place underwater because of various requirements of exploration missions. Additionally, it is difficult to accurately guide AUV along a specific path, while the on-land vehicles can do easily by referring to the road-side features. Therefore, the visual SLAM method is rarely used for underwater missions and

✉ Canjun Yang
ycj@zju.edu.cn

1 State Key Laboratory of Fluid Power and Mechatronic Systems, Zhejiang University, Hangzhou 310027, China

2 Pilot National Laboratory for Marine Science and Technology (Qingdao), Qingdao 266000, China

3 Ningbo Research Institute, Zhejiang University, Ningbo 315100, China

the visual odometry (VO) method is much more favored. It is known that underwater features (e.g. leaves) are similar and difficult to be distinguished, especially when the captured image contains much noise. Even if many researchers made significant contributions for underwater VO methods, most of them are operated offline [8, 9].

This paper aims to propose a real-time and robust navigation system to achieve docking operations in confined water regions with good visibility. For the homing stage, an integrated navigation algorithm by fusing visual odometry and kinetic model-based velocity is proposed, and a simple but effective light marker tracking algorithm is used at the docking stage.

The rest of this paper is organized as follows. Section 2 demonstrates the AUV docking system. Section 3 and 4 demonstrate the hydrodynamic model of the refitted AUV and underwater visual odometry method, respectively. Section 5 demonstrates the improved navigation system. Section 6 demonstrates the optical guidance method, Section 7 demonstrates the experiment and Section 8 draws a conclusion and discusses the future work.

2 AUV Docking System

Figure 1 demonstrates the designed AUV, the ultra-wideband (UWB) localization modules/stations, and the docking station. For a standard torpedo-like AUV, its rudders must have a sufficient flow of water over them to provide the necessary lift forces to keep the controllability, i.e., AUV should have a forward velocity as it attempts to dock. While the common AUV speed of 2–3 knot is not fast, it is fast enough to damage the docking station and the vehicle if the trajectory is not correct. For the newly-developed AUV shown in Fig. 1(a), it can be controlled at a specified depth without forward velocity because the AUV uses thrusters to provide lift forces rather

than rudders, i.e., the AUV is able to execute docking missions (at a specific depth) at a relatively slow speed, gathering more localization data in unit time. A light marker is mounted on the docking station to guide the AUV for the terminal docking operations. The UWB localization system is used to provide ground-truth locations of the AUV. The UWB location is acquired by computing distances to three UWB stations, which are deployed around the reservoir in the experiment. The location error of the UWB positioning system is within 10 cm with a maximum update rate of 10 Hz according to product specification. To evaluate the model-based linear velocity of the AUV, the reference velocity is calculated by using the UWB displacement and interval periods.

Figure 2 demonstrates the composition of the AUV. The vehicle is equipped with a downward-looking camera to perform the monocular visual odometry and a forward-looking camera to recognize the navigation lights on the DS. The ten-axis attitude system (TAS) can provide attitude and angular velocity data, and the pressure sensor measures the depth of AUV. The control unit is responsible for commanding AUV motion, sampling from the sensors, and transmitting data. The navigation unit computes the localization result and sends it to the control unit for generating control commands.

3 Hydrodynamic Model-Based Velocity Prediction

The hydrodynamic model is used to predict AUV velocity during the docking mission. Here, the mathematical formulae of the AUV motion are provided according to the SNAME notation. Fig. 3 demonstrates the motion and reference frames of the AUV where $\{b\}$ and $\{n\}$ represent the body-fixed reference frame and the North-East-Down (NED) reference frame, respectively, and $v^b = [u\ v\ w]^T$ and $\omega^b = [p\ q\ r]^T$ denote
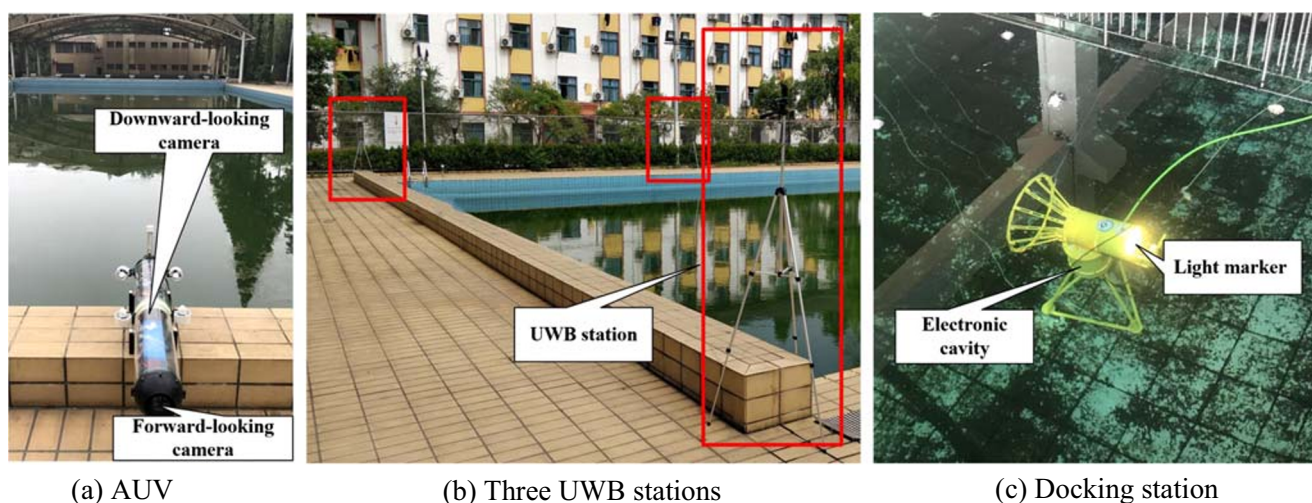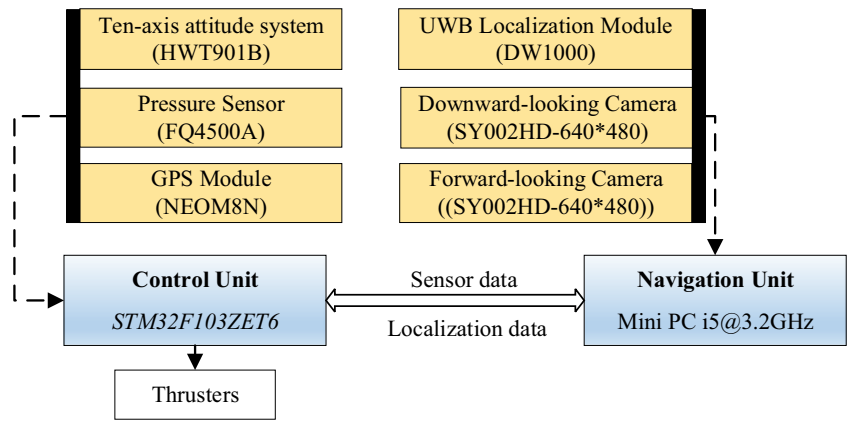

| (a) AUV | (b) Three UWB stations | (c) Docking station |

Fig. 1 The AUV docking system. (a) AUV (b) Three UWB stations (c) Docking station

Fig. 2 Composition of the AUV

the body-fixed velocity vector and angular velocity vector, respectively. Without loss of generality, the homing and docking stages are set at a specific depth. The velocity of the AUV can be computed using a horizontal hydrodynamic model, which is modified based on the model from a standard torpedo-like AUV. According to the control inputs of the AUV, its horizontal model can be formulized as follows.

$$m\left[\dot{u}-vr-y_g\dot{r}-x_g r^2\right] = \frac{1}{2}\rho L^4 X'_{rr}r^2 + \frac{1}{2}\rho L^3\left(X'_{\dot{u}}\dot{u} + X'_{vr}vr\right)$$
$$+ \frac{1}{2}\rho L^2\left(X'_{uu}u^2 + X'_{vv}v^2\right) + a_1 F_{p1} + a_2 F_{p2} \tag{1}$$

$$m\left[\dot{v}+ur-y_g r^2 + x_g\dot{r}\right] = \frac{1}{2}\rho L^4\left(Y'_{\dot{r}}\dot{r} + Y'_{r|r|}r|r|\right)$$
$$+ \frac{1}{2}\rho L^3\left(Y'_{\dot{v}}\dot{v} + Y'_r ur + Y'_{v|r|}v|r|\right) \quad + \frac{1}{2}\rho L^2\left(Y'_{v|v|}v|v| + Y'_v uv\right) \tag{2}$$
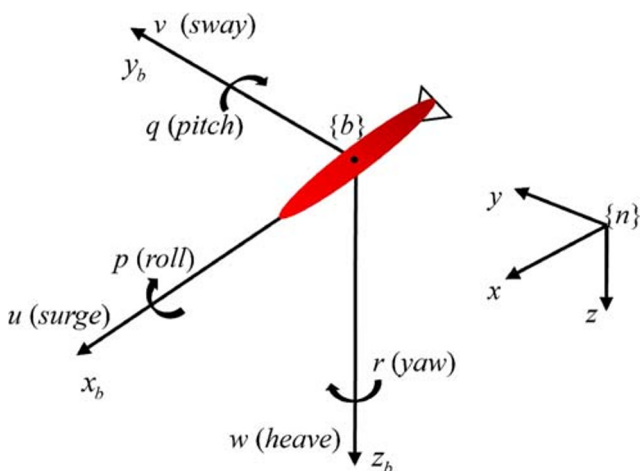

Fig. 3 AUV motion denotations and reference frames

$$J_z\dot{r} + m\left[x_G\left(\dot{v}+ur\right)-y_G\left(\dot{u}-vr\right)\right]$$
$$= \frac{1}{2}\rho L^5\left(N'_v uv + N'_{v|v|}v|v|\right)$$
$$+ \frac{1}{2}\rho L^4\left(N'_{\dot{v}}\dot{v} + N'_r ur + N'_{v|r|}v|r|\right)$$
$$+ \frac{1}{2}\rho L^3\left(N'_{\dot{r}}\dot{r} + N'_{r|r|}r|r|\right) + \left(a_1 F_{p1}-a_2 F_{p2}\right)l \tag{3}$$

J_z denotes the moment of inertia about z_b axis, L denotes the characteristic length of AUV, ρ denotes the density of water, l denotes the distance between two thrusters. Additionally, m and $(x_G\ y_G\ z_G)^T$ denote the mass and the vector of gravity center in the body-fixed coordinate, respectively. F_{p1} and F_{p2} denote the thrust generated by the left and right horizontal thrusters, respectively, and they are calculated by

$$F_{pi} = A_i u^2 + B_i un_{pi} + C_i n_{pi}^2 \tag{4}$$

where i denotes the order of the thruster, n_{pi} denotes the rotation speed of the i-th thruster, and A_i, B_i and C_i are three constants reflecting the characteristics of the i-th thruster and they can be calculated by referring to the work [10]. It is noted that two modification coefficients a_1 and a_2 are added to (1) and (3) to modify the thrust forces because two thrusters cannot generate the same thrust forces when given the same rotational speed due to the error of manufacture. Here, the particle swarm optimization (PSO) algorithm is used to offline identify the modification coefficients [11], and its schematic diagram is illustrated in Fig. 4. Before identification, the AUV was commanded to execute linear and circular motion in order to collect field data including the control input n_{pi} and corresponding velocity measurements (u_t^*, v_t^*, r_t^*). The 19 dimensionless hydrodynamic parameters in (1)–(3) were solved by the method proposed in [12] and their values are provided in Table 1.

In the PSO algorithm, the search ranges of a_1 and a_2 are empirically confined within [0.8,1.2] since the manufacture
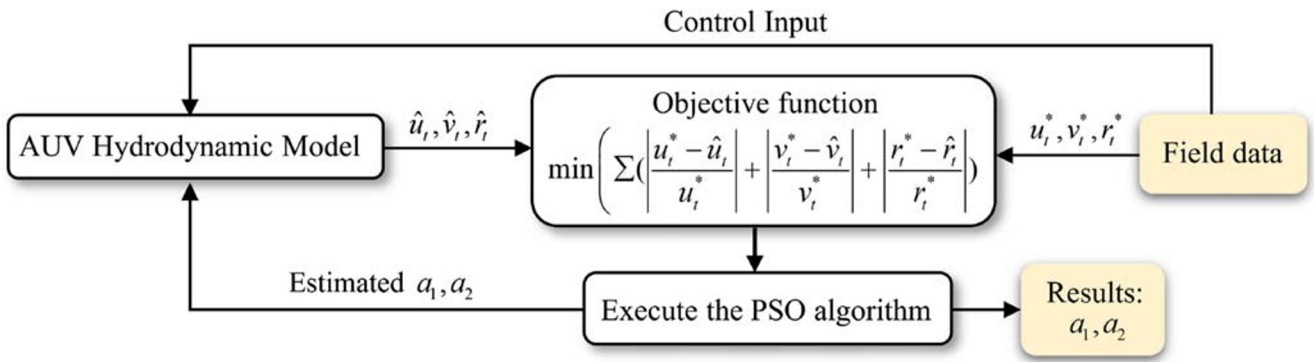
Fig. 4 The schematic diagram of identification of the thrust modification coefficients

consistency of the thrusters would not be too bad based on the test performance in the experiments. The objective function during the identification process is given by:

$$g = \min\left\{ \sum_{t=1}^{T} \left(\left|\frac{u_t^* - \widehat{u}_t}{u_t^*}\right| + \left|\frac{v_t^* - \widehat{v}_t}{v_t^*}\right| + \left|\frac{r_t^* - \widehat{r}_t}{r_t^*}\right| \right) \right\} \qquad (5)$$

where $(\widehat{u}_t, \widehat{v}_t, \widehat{r}_t)$ denotes the model-based velocity computed by the Runge-Kutta method [13] and the T denotes the total time steps of the collected field data. If the sum of the relative error of the model-based velocity converges, it is believed that a_1 and a_2 can modify the deviation of the thrust force well. Figure 5 (a) demonstrates the variation process of the modification coefficients, which are the average result calculated by 50 independent PSO runs. It is observed that a_1 and a_2 converges to 1.095 and 1.093, respectively. Based on this result, the modified model-based velocity can be calculated using the field data, and the corresponding result is shown in Fig. 5 (b)-(c). It is observed that the modified model-based velocity has better accuracy compared to the uncalibrated model-based velocity. More specifically, the modified method improves the computational accuracy by 34.7% and 48.2% for linear and angular velocities, respectively. This result is gained by comparing with the ground-truth velocity (velocity measurements in the figures) calculated from the UWB positioning system.

4 Underwater Monocular Visual Odometry

The bundle adjustment (BA) method is popular for visual navigation as it can optimize the robot pose and map features with a loop closure detection to eliminate the accumulative localization error [14]. However, for the terminal docking operation which only covers dozens of meters, the BA method spends much time matching with the global map. Thus, a simple but effective underwater monocular visual odometry (UMVO) method is proposed to support the original navigation system and maintain the real-time performance. Pyramidal Lucas Kanade (PLK) feature tracker is popular for target tracking problems as it eliminates the need of costly feature description and directly operates on pixel intensities [15]. Here, FAST feature extractor is used to extract the image corners and combined with the PLK tracker to quickly perform the feature registration.

Based on the above-mentioned framework, the proposed UMVO can address two critical problems in homing stages: 1) remove inaccurate feature correspondences; 2) compute the AUV pose. The details are given as follows.

4.1 Remove Inaccurate Feature Correspondences

To compute the AUV pose, the PLK tracker is used to determine the feature correspondences. Although the PLK tracker can achieve good results in most terrestrial applications, it would inevitably encounter feature mismatch problems due to the blurry and similar features in underwater environments.

Table 1 Estimated hydrodynamic parameters and modification coefficients

| Parameter	X'_{rr}	X'_u	X'_{vr}	X'_{uu}	X'_{vv}	$Y'_{\dot{r}}$	$Y'_{r|r|}$
Value	−1.933e−1	-1.304e-1	4.454e-2	−3.062e-1	−5.013e-1	2.906e−2	1.377e-1
Parameter	$Y'_{\dot{v}}$	Y'_r	$Y'_{v|r|}$	$Y'_{v|v|}$	Y'_v	$N'_{\dot{r}}$	$N'_{r|r|}$
Value	−2.490e-1	1.630e-3	4.998e-1	−2.768e-1	-3.519e-1	-2.149e-2	−1.250e-2
Parameter	$N'_{\dot{v}}$	N'_r	$N'_{v|r|}$	$N'_{v|v|}$	N'_v	a_1	a_2
Value	5.941e-4	-1.889e-2	-1.001e-1	3.476e-2	-2.352e-4	1.095	1.093
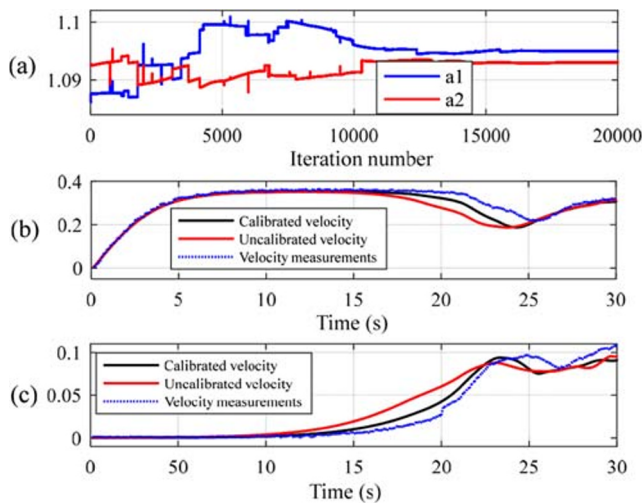
Fig. 5 Test results of the modification method. (a) Variation of modification coefficients. (b) Linear velocity comparison, unit: m/s. (c) Angular velocity comparison, unit: rad/s

Some effective solutions have been proposed to enhance and restore the underwater images, such as the modulation transfer function-based restoration [16], histogram equalization [17], and the dehazing enhancement [18]. However, for a real-time navigation system, these methods are time-consuming and cannot completely avoid wrong feature matching, especially when some suspended solids adhere to the front of the camera. Therefore, we proposed a random sample consensus (RANSAC) based method to remove the outlier feature matchings for UMVO. The proposed method has three steps as given below.

4.1.1 Remove Features around Image Boundary

It is noted that the captured underwater images are inevitably distorted due to refraction of water and the water-proof shell, decreasing the position accuracy of extracted features. Therefore, Zhang's calibration method is used to correct the distorted images [19]. After image calibration, there will be a lot of black pixels around the boundary of images, as shown in Fig. 6. The features located at the boundary region which is confined between two yellow rectangles should be removed
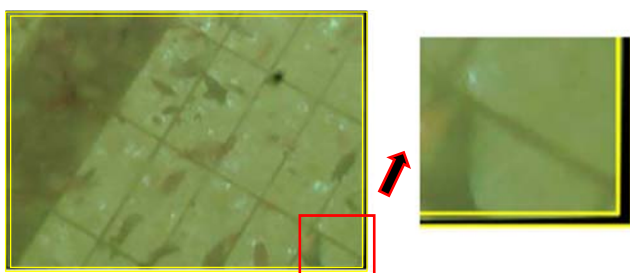
Fig. 6 The calibrated image with an enlarged view

because the extracted corners here may not be the environmental features. This confined region is determined by maximumly selecting the effective image area. The area of the confined region is related to the intrinsic parameters of the camera and calibrated coefficients. Thus, if the camera and its water-proof shell are not changed, the boundary region would not change too.

4.1.2 RANSAC Based Outlier Removal

The RANSAC method can be used to remove outliers based on a mathematics model, which are the essential or homography matrix in the UMVO method. However, it is difficult to determine which model is more suitable to describe the transformation of features because each model is suitable for a specified condition. Clearly, the essential matrix is generally used to describe the epipolar geometry constraint which requires the distance of two consecutive camera positions cannot be too close, while the homography matrix describes the transformation of matched features in a same plane. For real-world scenarios, such as at pool or sea, only little prior information of the environmental characteristics can be acquired. For our UMVO method, it is assumed that the terrain around the docking station would not be very rugged.

Here, a RANSAC based method is proposed to remove incorrect feature correspondences. The models of homography transformation and epipolar geometry constraint are given by:

$$\Gamma_1(p_1, p_2) = p_2 - \mathbf{H}p_1 \tag{6}$$

$$\Gamma_2(p_1, p_2) = p_2^T \underbrace{\mathbf{K}^{-T} \left[t_\times \right] \mathbf{R}\mathbf{K}^{-1}}_{\mathbf{E}} p_1 \tag{7}$$

where p_1 and p_2 denote the coordinates of the matched features in consecutive image I_1 and I_2, respectively, $\mathbf{H}$ and $\mathbf{E}$ denote the homography and essential matrices, respectively, $\mathbf{K}$ denotes the intrinsic matrix of the camera, and $\mathbf{R}$ and t denote the rotation matrix and the displacement vector of the camera from the image I_1 to I_2, respectively. It is noted that $[t_\times]$ is a simplified expression, which is defined as

$$a \times b = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = [a_\times]b \tag{8}$$

where $a = [a_1 \ a_2 \ a_3]^T$ and $b = [b_1 \ b_2 \ b_3]^T$. The rigorous definition of two basic models is that (6) and (7) are equal to 0. However, this condition cannot be satisfied due to the image noise and limited extracted accuracy. Therefore, the inliers and outliers in the RANSAC based method are identified by a predefined threshold as:

$$\|\Gamma_i(p_1, p_2)\| < d_{th}^i, \quad i = 1, 2 \tag{9}$$

where d_{th}^i denotes the threshold for the RANSAC method using $\Gamma_i(p_1, p_2)$, and it is set to 3 based on the recommendation from the OPENCV library.

4.1.3 Fine Removal Using Prior Information

Although the previous steps can remove most outliers, there still remains inaccurate feature correspondences as some identified features are too close and look similar; as a consequence, some different features may also match and satisfy the condition of (9). To remove outliers further, the prior information of terrain near the homing station is used. The detailed execution steps are: a) Triangulate the matched features in image I_1; b) compute the median depth d_{me} of triangulated 3-D feature points. c) Remove the features whose depth have a negative sign to d_{me} or is larger than $3d_{me}$, or smaller than $1/(3d_{me})$.

The optical flow before and after removing the outliers are shown in Fig. 10, in which the white line connects the matched feature positions in image I_1 and I_2, and the white circles denote the feature position in image I_2. Note that the yellow circle 1 in Fig. 7(a) contains a black spot which is a suspended solid adhered to the water-proof shell of the camera, and the yellow circle 2 and 3 mark the position of inaccurate optical flow, which has an obviously different tendency compared to the others. It can be observed that the inaccurate feature correspondences are all removed in Fig. 7(b), proving that the proposed method can effectively improve the correctness of feature matching.

4.2 Compute the AUV Pose

4.2.1 Select the Best Solution for Pose Update

The camera pose can be acquired by decomposing the essential or homography matrix using the singular value decomposition method. However, there are four pose solutions for both essential and homography matrices. To eliminate the incorrect solutions, the matched features are firstly triangulated by the corresponding pose, and then the solutions where the point is in front of both cameras are chosen as the temporarily correct solutions. After these steps, there is only one solution for the essential matrix and two solutions for the homography matrix. Although the prior information believes that the terrain around the docking station would not be very rugged, it is not enough to determine which model is more suitable for computing the AUV pose. Therefore, the following method is proposed to determine and update the AUV pose.

Step 1. Calculate the average parallax of matched features in the forward direction. If the average parallax is smaller than the preset threshold, the homography matrix will be used to describe the feature transformation of the current scenario, and then go step 2. Otherwise, go step 3.

Step 2. Calculate the sum of feature reprojection error for each solution from the homography matrix, and choose the best solution which has the smaller error. Go step 4.

Step 3. Calculate the sum of feature reprojection error for three solutions from both essential and homography matrices, and choose the solution with the smallest error. Go step 4.

Step 4. Compute the AUV trajectory as

$$\mathbf{R}_k = \begin{cases} \mathbf{R}_1, & k = 1 \\ \mathbf{R} \cdot \mathbf{R}_{k-1}, & k \geq 2 \end{cases} \quad (10)$$

$$t_k = \begin{cases} s_k t_1, & k = 1 \\ t_{k-1} + s_k \cdot (\mathbf{R}_k t), & k \geq 2 \end{cases} \quad (11)$$

where t_k and $\mathbf{R}_k$ denote the AUV displacement and rotation matrix at time step k, respectively, and s_k denotes the monocular scale.

4.2.2 Recover the Metric Scale

It is known that the monocular vision-only systems are incapable of recovering the metric scale, thus limiting their usage

Fig. 7 Optical flow results presented in image I_2. (a) Before removing outliers (b) After removing outliers

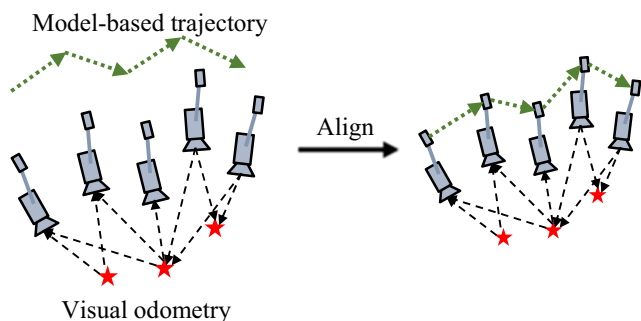(a) Before removing outliers   (b) After removing outliers

Model-based trajectory
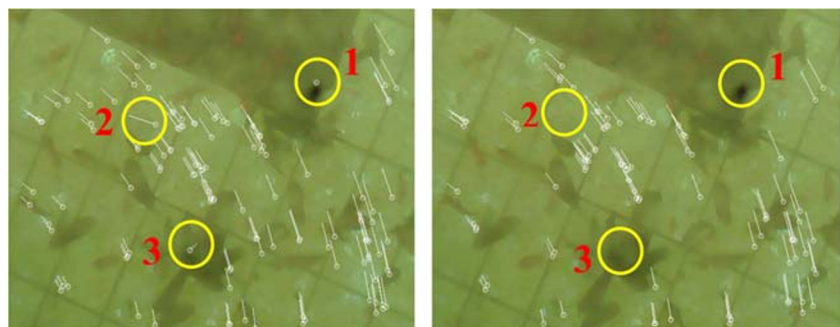
Fig. 8 Recovering the metric scale by the model-based alignment

in real-world applications, such as at pool and sea. In recent years, there is an increasing trend to use a low-cost inertial measurement unit (IMU) to recover the monocular scale, however, this method requires complex and rigorous initialization, such as launching the robot from a known pose, moving slowly or starting missions from a known environment. To address this problem, we propose a scale recovery method by aligning the hydrodynamic model-based motion to UMVO. This method can be formulized as follows:

$$\frac{|V_k| + |V_{k-1}|}{2}(T_k - T_{k-1}) = s_k \|t_k\| \quad (12)$$

where T_k and T_{k-1} denote the timestamp at time step k and k-1, respectively, V_k and V_{k-1} denote the velocity at time step k and k-1, respectively, and s_k denotes the metric scale. The corresponding schematic diagram is shown in Fig. 8.

To evaluate performance of the scale recovery method, the AUV was commanded to move along a square path and collect the images and corresponding control commands at Yuquan pool/reservoir at Zhejiang University (ZJU pool), Hangzhou, China. The collected data are called ZJU dataset in this paper. Figure 9(a) demonstrates the comparison result of the proposed UMVO with the model-based scale and the relative scale [20], respectively. It is observed that the trajec-

tory with the model-based scale is closer to a square, while the trajectory with the scale updated in a relative way has obvious drift, especially when the AUV turned a corner. This is because the model-based scale is essentially an absolute scale which is independent of the previous scale, while the scale calculated in a relative way will accumulate the error and cause trajectory drift. Comparing to the UWB localization module, the proposed method improves the localization accuracy by 8.4%. To further evaluate the proposed visual odometry method, the open-source KITTI dataset (sequence: 00) is used to test the localization accuracy [21] and the test result is shown in Fig. 9 (b). It is observed that the trajectory with the absolute scale outperforms that with the relative scale in long range, and this result is similar to the testing using ZJU pool dataset.

5 ISRT-UKF Based Navigation System

To improve the localization accuracy and robustness of the proposed algorithm, this study fuses the hydrodynamic model-based velocity and UMVO by modifying our previous navigation algorithm and adding an outlier rejection function to prevent filter failures [22]. The modified algorithm is referred to as the improved square root transformed unscented Kalman filter (ISRT-UKF). In the following parts, how to use and fuse each data source will be introduced in detail.

5.1 System and Measurement Models

To apply the ISRT-UKF algorithm, the system and measurement models of the AUV should be firstly established. In our tests, the AUV can move with nearly zero pitch and roll angles during the entire docking mission. Thus, we only consider the horizontal motion of the AUV. The corresponding system and measurement models are provided as:

Fig. 9 The performance test of the proposed visual odometry method. (a) ZJU pool dataset. (b) KITTI dataset (sequence: 00)
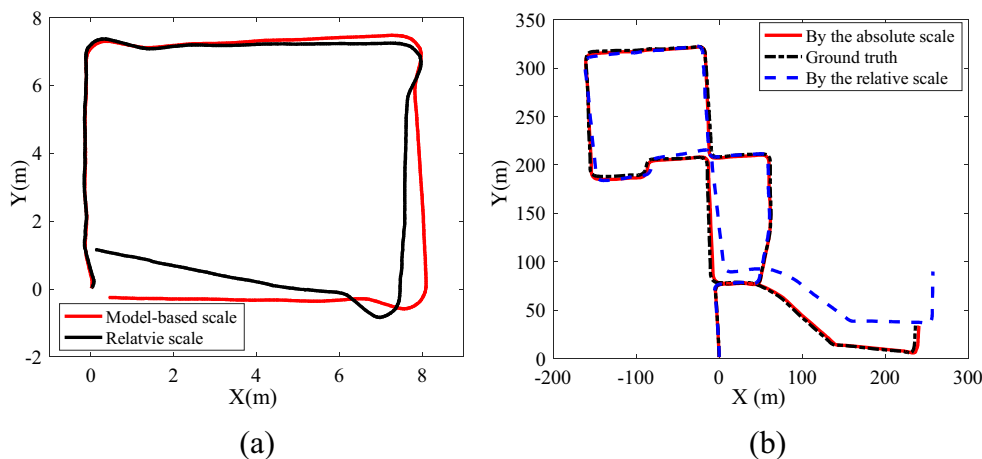
System model:

$$X_k = \underbrace{\begin{bmatrix} x_k \\ y_k \\ \phi_k \end{bmatrix}}_{X_k} = f(X_{k-1}, U_{k-1}, n_{k-1}) = \underbrace{\begin{bmatrix} x_{k-1} \\ y_{k-1} \\ \phi_{k-1} \end{bmatrix}}_{X_{k-1}} + $$

$$\underbrace{\begin{bmatrix} \Delta t \cos\phi_{k-1} & \Delta t \sin\phi_{k-1} & 0 \\ \Delta t \sin\phi_{k-1} & -\Delta t \cos\phi_{k-1} & 0 \\ 0 & 0 & \Delta t \end{bmatrix}}_{\Theta_{k-1}} \underbrace{\begin{bmatrix} u_{k-1} + n_u \\ v_{k-1} + n_v \\ \omega_{k-1} + n_\omega \end{bmatrix}}_{U_{k-1} + n_{k-1}} \quad (13)$$

Measurement model:

$$Z_k = \underbrace{\begin{bmatrix} x_k \\ y_k \\ \phi_k \end{bmatrix}}_{Z_k} = h(X_k) + m_k = \underbrace{\begin{bmatrix} x_k \\ y_k \\ \phi_k \end{bmatrix}}_{h(X_k)} + m_k \quad (14)$$

where Θ_{k-1} denotes the rotation matrix from the body-fixed reference frame to the NED reference frame, Δt denotes the sampling time, X_k and Z_k denote the system and measurement state vectors at time step k, the vector $[x_k \quad y_k \quad \phi_k]^T$ represents the position and heading of the AUV in the NED reference frame, $f(\cdot)$ and $h(\cdot)$ denote the system and measurement model functions respectively, U_{k-1} denotes the control input which contains linear and angular velocity in the body-fixed reference frame, n_{k-1} and m_k denote the noise vectors of the control input and measurements respectively. These two vectors satisfy the Gaussian distribution of $N(0, \mathbf{Q}_{k-1})$ and $N(0, \overline{\mathbf{R}}_k)$

where $\mathbf{Q}_{k-1}$ and $\overline{\mathbf{R}}_k$ denote the system and measurement covariance matrices, respectively.

It is noted that the model-based velocity has accumulative error because it is computed in an iterative way. Therefore, to reduce accumulative error, the AUV is regarded to move in a constant forward speed (when without the rotational motion) when it speeds up to a velocity that is close to the theoretically stable velocity.

The noise statistics of the linear velocity and UMVO are empirically determined by observing performance of the AUV in a constant speed. For the TAS, the noise statistics is tuned according to the specifications of the sensor manufacturers. In the navigation system, the origin of the NED reference frame is set at the center of the docking station entrance. The UMVO results is acquired in the reference frame established at the position regarding the first camera frame. Therefore, to fuse the UMVO and the model-based velocity, a coordinate transformation should be carried out. The zero degree of the heading angle provided by the TAS aligns with the due east direction. Assuming that the heading angle of AUV at the first camera frame is δ_0, the position given by the UMVO should be transformed from the original (x_{vo}^c, y_{vo}^c) to the NED reference frame at $(x_{vo}^{NED}, y_{vo}^{NED})$ in the way:

$$x_{vo}^{NED} = x_{vo}^c \cos\left(\frac{\pi}{2} - \delta_0\right) + y_{vo}^c \sin\left(\frac{\pi}{2} - \delta_0\right) \quad (15)$$

$$y_{vo}^{NED} = y_{vo}^c \cos\left(\frac{\pi}{2} - \delta_0\right) - x_{vo}^c \sin\left(\frac{\pi}{2} - \delta_0\right). \quad (16)$$

5.2 ISRT-UKF Algorithm

The ISRT-UKF algorithm is modified from our previous work [22] whose computational process is provided in Table 2 where $chol()$ denotes the Cholesky decomposition operator, $qr()$ denotes the QR decomposition, and $cholupdate()$ denotes

Table 2 The SRT-UKF algorithm

Prediction	Correction
$\mathbf{M}_{k-1} = chol(\mathbf{P}_{k-1})$	$\overline{\gamma}_k^{[i]} = h\left(\overline{\chi}_{k|k-1}^{[i]}, U_{k-1}\right)$
$\chi_{k-1}^{[i]} = X_{k-1} + (\mathbf{M}_{k-1})^T \zeta^{[i]}, \quad i = 1, 2, ..., 2n_a$	$\widehat{Z}_{k|k-1} = \sum_{i=1}^{2n_a} W\overline{\gamma}_k^{[i]}$
$\overline{\chi}_{k|k-1}^{[i]} = f\left(U_{k-1} + \chi_{k-1}^{u[i]}, \chi_{k-1}^{[i]}\right)$	$\varsigma_i = \sqrt{W}\left(\overline{\gamma}_k^{[i]} - \widehat{Z}_{k|k-1}\right), \quad i = 1, ..., 2n_a$
$W = 1/(2n_a)$	$\overline{\mathbf{S}}_{z_k} = qr\left[\varsigma_1 \; \varsigma_2 \; ... \; \varsigma_{2n_a} \; \sqrt{\overline{\mathbf{R}}_k}\right]$
$\widehat{X}_{k|k-1} = \sum_{i=1}^{2n_a} W\overline{\chi}_{k|k-1}^{[i]}$	$\mathbf{P}_{x_k z_k} = \sum_{i=1}^{2n_a} W\left(\overline{\chi}_{k|k-1}^{[i]} - X_{k|k-1}\right)\left(\overline{\gamma}_k^{[i]} - Z_k\right)^T$
$e_i = \sqrt{W}\left(\overline{\chi}_{k|k-1}^{[i]} - \widehat{X}_{k|k-1}\right), \quad i = 1, ..., 2n_a$	$\mathbf{K}_k = \left(\mathbf{P}_{s_k z_k}/\overline{\mathbf{S}}_{z_k}\right)/\left(\overline{\mathbf{S}}_{z_k}\right)^T$
	$\widehat{X}_{k|k} = \widehat{X}_{k|k-1} + \mathbf{K}_k\left(Z_k - \widehat{Z}_{k|k-1}\right)$
$\mathbf{M}_{k|k-1} = qr[e_1 \; e_2 \; ... \; e_{2n_a}]$	$\mathbf{M}_k = cholupdate\left\{\mathbf{M}_{k|k-1}, \mathbf{K}_k\left(\overline{\mathbf{S}}_{z_k}\right)^T, -1\right\}$

the Cholesky update for the square root factor of the state covariance matrix. The ISRT-UKF contains two steps, and they are prediction and correction. The navigation data for the prediction step come from the model-based velocity and TAS. For the correction step, the data come from the UMVO. The related nomenclature is given at the end of this paper. The primary difference between two algorithms is that the ISRT-UKF has an outlier rejection strategy. The details of additional steps in the filter are given below.

5.3 Initialization and outlier rejection

For most underwater navigation cases, the initial position of the AUV is given by the GPS. The accuracy of the GPS position fixes is generally bounded within 5 m. One of its drawbacks is the localization jumps. This drawback may be negligible for long-range missions at sea, but for a confined water region, an accurate initial position is important for the AUV to enter the light coverage area at the docking stage. Thus, we propose a simple but robust initialization approach before starting the homing mission. To evaluate quality of the GPS position fix, the AUV is commanded to move along a specified heading angle firstly, and then compares the heading angle to the included angle of the line between two successive GPS position fixes. The schematic diagram of the method is shown in Fig. 10, and it can be formulized as follows

$$\left| \pi + \phi_i - \arctan\left(\frac{y_i - y_{i-1}}{x_i - x_{i-1}} \right) \right| \leq \vartheta \qquad (17)$$

where ϕ_i denotes the i-th measurement of the AUV heading angle during the transitory stage, β_i denotes the included angle of the line between successive GPS position fixes (x_{i-1}, y_{i-1}) and (x_i, y_i), and ϑ is an empirical threshold, which is set to $\pi/6$ by trial-and-error procedure. When (17) is satisfied for successiv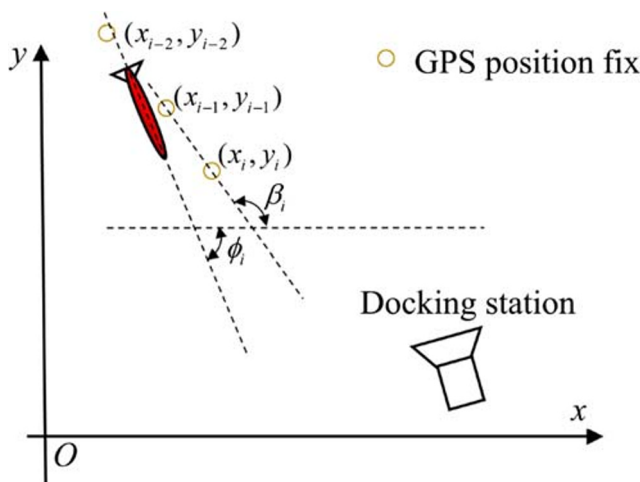e two times, the latest GPS position fix is used to initialize the system state. Although UMVO can provide accurate localization, it may still generate inaccurate visual tracks in scenes of illumination change and textureless areas. To avoid fusing inaccurate information to ISRT-UKF, we propose an outlier rejection method using an uncertainty ellipse method, which is modelled by the Mahalanobis distance. The formulation of the proposed method is demonstrated as follows. First, calculate innovation of the measurement.

$$\widetilde{Z}_k = Z_k - \widehat{Z}_{k|k-1} \qquad (18)$$

where $\widetilde{Z}_k$ denotes the innovation of the measurement, Z_k denotes the actual measurement (UMVO) and $\widehat{Z}_{k|k-1}$ denotes the predicted measurement. In general, $\widetilde{Z}_k$ satisfies the Gaussian distribution of $N(0, \mathbf{T}_k)$ where $\mathbf{T}_k$ denotes the covariance of the innovation. $\mathbf{T}_k$ is calculated by:

$$\mathbf{T}_k = E\left(\widetilde{Z}_k \widetilde{Z}_k^T \right) = \overline{\mathbf{S}}_{z_k}^T \overline{\mathbf{S}}_{z_k} \qquad (19)$$
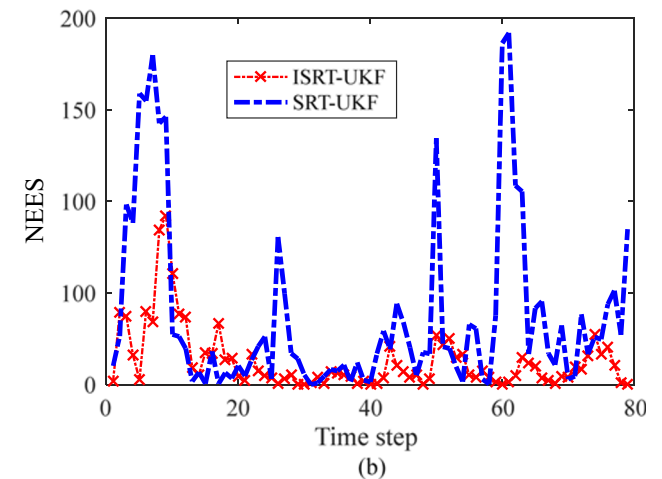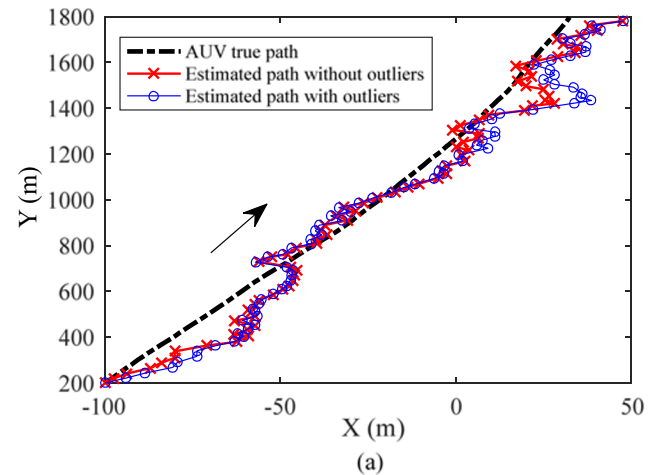

Fig. 11 Simulation test of the outlier rejection method

Fig. 10 Schematic diagram of AUV position initialization

where $\overline{\mathbf{S}}_{z_k}$ denotes the square root factor of the predicted measurement. The outlier can thus be identified:

$$d^2 = \widetilde{Z}_k^T \mathbf{T}_k^{-1} \widetilde{Z}_k < \chi_\alpha^2 \qquad (20)$$

where d^2 denotes the squared Mahalanobis distance and χ_α^2 is the chi square value of confidence. If (20) is not satisfied, the actual measurement is regarded as outlier and ISRT-UKF only executes the prediction step.

Figure 11(a) demonstrates the simulation result of the performance of the outlier rejection method. The AUV is assumed to move along a specified path as the arrow indicates, and the measurements are generated based on the true path. To simulate outliers, some measurements are added random error whose error is 8 times larger than the original error of the measurements. It is observed that the estimated trajectory with the outlier rejection function is closer to the true path. This result reveals that the outlier rejection process can effectively increase estimation accuracy and robustness of the ISRT-UKF algorithm. Figure 11(b) shows the normalized estimation error squared (NEES) in the simulation [23]. A larger NEES indicates that the measurement cannot well match with the prediction results and the filter has the potential to diverge. It is observed that the NEES of ISRT-UKF is smaller than SRT-UKF because the proposed outlier rejection removes the abnormal measurements (remarkable peaks) and maintains good estimation performance of the filter.

6 Docking Operations Using Optical Guidance

To achieve terminal docking operations, we use the optical guidance method which recognizes the light marker on the docking station via a camera housed at the nose of the AUV. Different from the at-sea trails, the experiment on a confined water region, such as a reservoir, would generate virtual objective on the water surface because of the light reflection. This wrong target should not be recognized by the visual algorithm. The complete image processing steps are given as: (1) Convert the RGB image to the grey image; (2)

Binarize the image; (3) Operate the median filter to remove noisy image points; (4) Use the Hough circle recognition algorithm to find the circles in the image and extracts the centers of the circles; (5) Identify the number of circles and remove outlier circles. The processed image sequences are given in Fig. 12. During the docking process, the AUV is commanded to continuously adjust its pose to keep the light spot (extracted center point) in the center of the image.

7 Experiment

The validation experiments were conducted at ZJU pool, which has a size of 50 m (length) × 21 m (width) × 1.5 m (depth). In the experiments, we used the popular proportion-integration-differentiation (PID) controller to command AUV to track predetermined mission points and accomplish docking operations. The forward speed of the AUV is 0.5 m/s and the resulting localization update rate is 6 Hz for homing and 10 Hz for docking. Figure 13 provides the basic schematic diagram of the homing and docking strategies. In the experiments, the AUV was moving at a depth of roughly 0.2 m, enabling that the hull of the AUV was completely submerged by the water but the antennas are able to receive localization information above the water surface.

Figure 14 (a) demonstrates five mission trajectories where the AUV tracks appointed mission points firstly and then go back to the regions around the origin point. In five missions the AUV has an average terminal error of 3.4 m and the average traveling distance is of approximately 110 m. Figure 14(b) shows a comparison result between the ISRT-UKF trajectory and the UMVO-only trajectory. It is observed that the terminal position of the ISRT-UKF trajectory is closer to the origin than that of the UMVO-only method. This is because the heading angle in UMVO is updated in an iterative way, while in ISRT-UKF, the heading angle is updated by fusing TAS and UMVO data so that the error would not obviously accumulated. In five missions, the average terminal error of the UMVO-only method is 5.1 m, i.e., the ISRT-UKF based fusion method improves

Fig. 12 Processed images. (a) raw image (b) gray image (c) binarized image (d) median-filtering image (e) final image with extracted circles
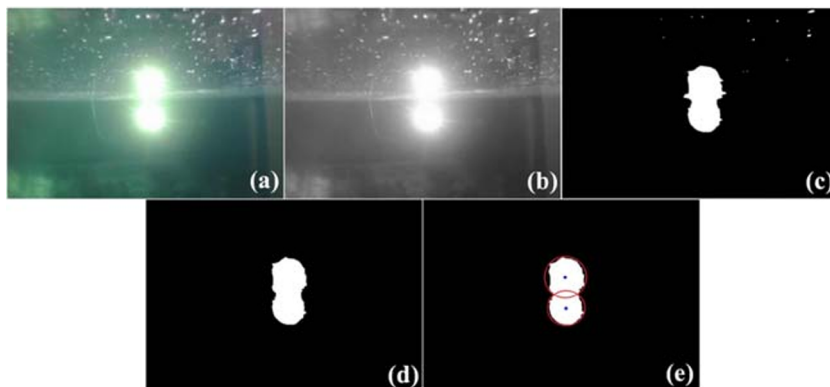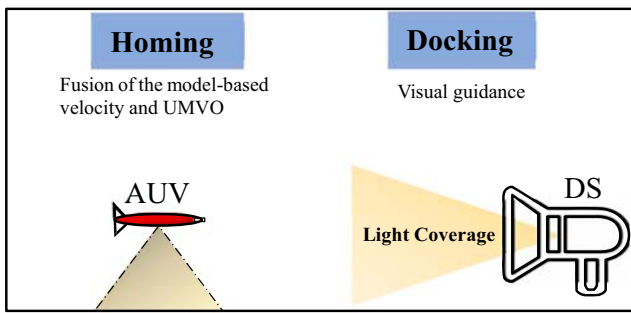
Fig. 13 Schematic diagram of the AUV homing and docking strategy

AUV terminal localization accuracy by 33.3% compared to that of UMVO.

Figure 15(a) shows an AUV docking mission trajectory where the GPS origin point is at the location of the docking station. It is observed that the AUV does not use the first three GPS position fixes to initialize its position because the proposed navigation system requires high-quality initialization data. In addition, from all GPS fixes shown, the GPS
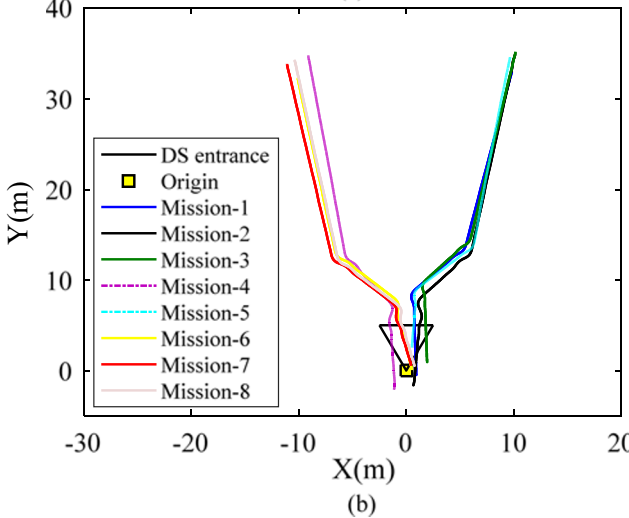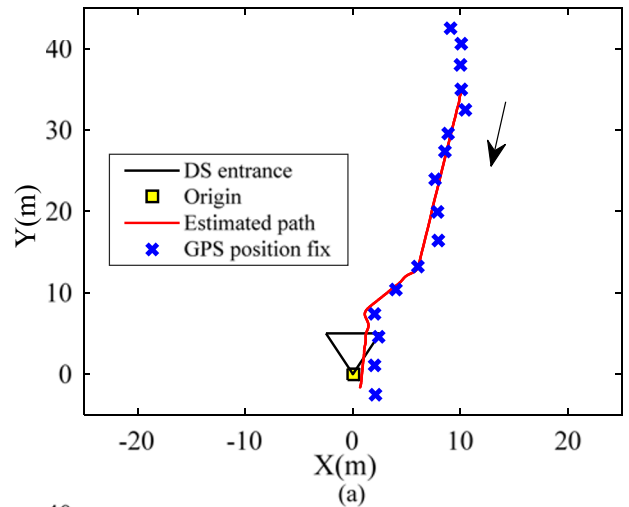

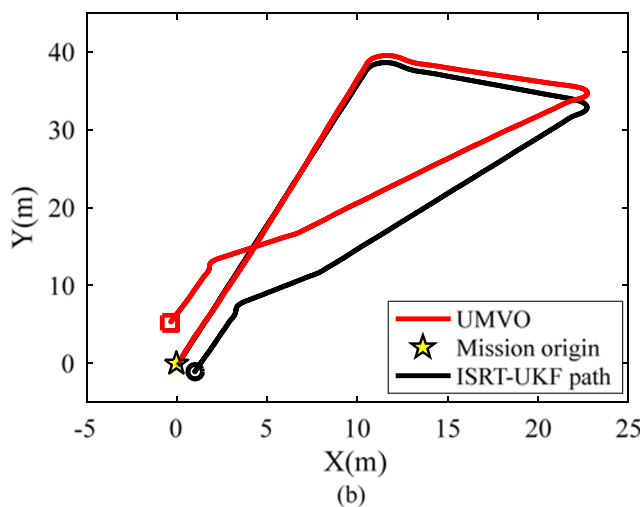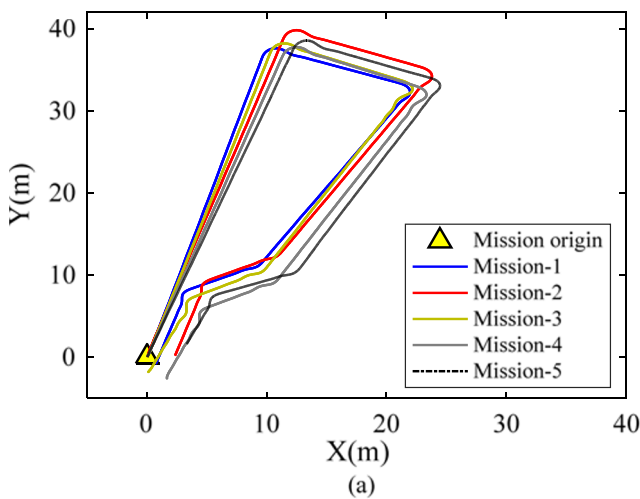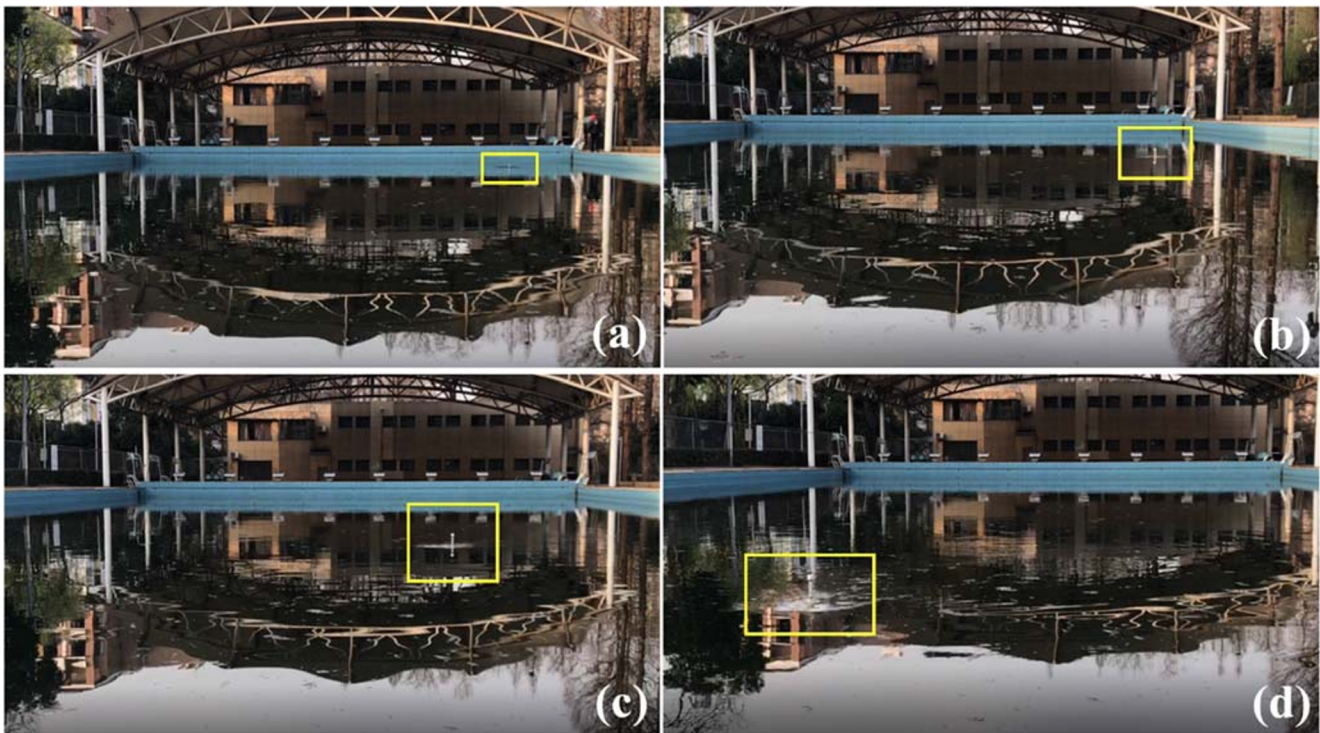**Fig. 15** Tests of AUV homing and docking operations

localization should not be merged into the ISRT-UKF because the jumps have large error (see fixes around the DS entrance), possibly causing diverging of the filter. Even with the outlier rejection in the filter, the resulting algorithm may only execute the prediction step, and thus arises remarkable localization error and makes the AUV miss the light coverage area when it is close to the DS. Figure 15 (b) shows eight AUV trajectories of docking missions. In each docking attempt, AUV was moving along the predefined path and finally guided into the docking station. The homing process is shown in Fig. 16 and the captured image sequences with distance information during the docking process in shown in Fig. 17.

8 Conclusion

This paper proposes an AUV navigation system for docking in a confined water region with good visibility. To enhance the controllability of the conventional AUV (with one thruster and four rudder pieces), a newly-designed AUV is used in

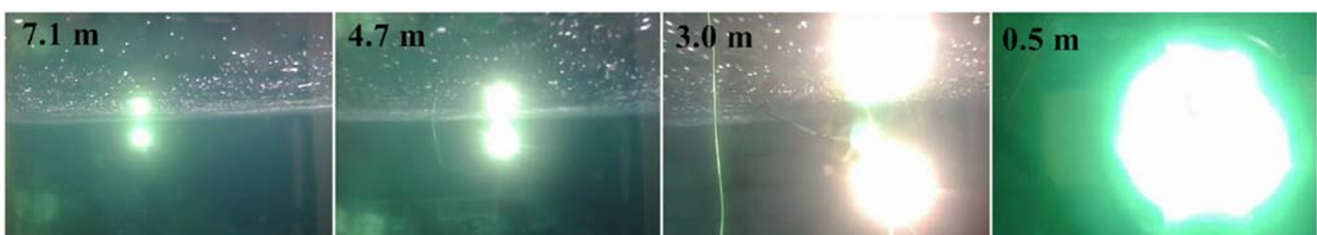**Fig. 14** Testing of AUV Localization performance

Fig. 16 Image sequences of AUV homing operations. (Yellow rectangle marks the position of AUV)

the docking experiments. The simulation validated the advantages of the monocular-scale recovery and the outlier rejection methods, and the pool experiments validated the performance of both homing and docking operations. The methods of model-based velocity prediction and UMVO are low-cost, ease-of-use and promising to AUV routine cruise for observation missions in confined environments. The proposed vision-based optical guidance outperforms most guidance algorithms (electromagnetic and acoustic) in computational complexity and hardware requirements. For an AUV to operate AUV docking operations at night, a floodlight should be equipped for UMVO. In addition to visibility, the primary limitation of the proposed navigation system is the environmental current, which affects the prediction accuracy of the AUV velocity. The current effects can be ignored in confined regions (not open water areas) because such areas are usually surrounded by walls or dam bodies, and the current is very gentle and small. For working in open water regions, such as sea, the proposed navigation system should be modified to equip with a current meter to measure amplitude and direction of the

current, and involves the current items in the hydrodynamic model.

The main contributions of this paper can be summarized as follows.

(1) A horizontal hydrodynamic model of the AUV is derived with two modification coefficients that computed by a PSO-based method. The modified model effectively reduces velocity prediction drift caused by manufacturing error of thrusters. This modification method can also be extended to other underwater vehicles that have two parallel thrusters, and not only for horizontal motion.
(2) The proposed UMVO can remove wrong feature correspondences by a three-layer strategy. The monocular metric scale is recovered by aligning with the model-derived displacement and outperforms the results using the relative scale.
(3) An outlier detection method is proposed in the non-linear filter to remove outliers of measurements. The test results validated the effectiveness of the method.

Fig. 17 Captured image sequences during the AUV docking operation

The future work will focus on designing novel navigation methods to improve AUV localization performance in turbid environments with dynamic features.

Nomenclature X_{k-1}, The system state at time step k-1; $\mathbf{M}_{k-1}$, Square root factor of the state covariance matrix; $\chi^{[i]}_{k-1}$, The i-th sigma point; $\zeta^{[i]}$, The i-th column of the transformed unscented point set; W, Weight of sigma points; n_a, Dimension of the system state; $\overline{\chi}^{[i]}_{k|k-1}$, The i-th propagated sigma point of the system state; $\widehat{X}_{k|k-1}$, Predicted system state; $\mathbf{M}_{k|k-1}$, Predicted square root factor of the system state; $\overline{\gamma}^{[i]}_k$, The i-th propagated sigma point of the measurement; $\overline{Z}_{k|k-1}$, Predicted measurement; $\overline{\mathbf{S}}_{zk}$, Square root factor of the measurement; $\mathbf{P}_{x_k z_k}$, Cross covariance of the system state and measurements; $\mathbf{K}_k$, Kalman gain matrix; $\widehat{X}_{k|k}$, Estimated system state; $\mathbf{M}_k$, Estimated square root factor of the system state

References

1. Yang, C., Lin, M., Li, D.: Improving Steady and Starting Characteristics of Wireless Charging for an AUV Docking System. *IEEE Journal of Oceanic Engineering.* 1–12 (2018)
2. Lin, M., Li, D., Yang, C.: Design of an ICPT system for battery charging applied to underwater docking systems. *Ocean Engineering.* **145**, 373–381 (2017)
3. Palomeras, N., Vallicrosa, G., Mallios, A., Bosch, J., Vidal, E., Hurtos, N., Carreras, M., Ridao, P.: AUV homing and docking for remote operations. *Ocean Engineering.* **154**, 106–120 (2018)
4. Eren, F., Pe'Eri, S., Rzhanov, Y., Thein, M., Celikkol, B.: Optical Detector Array Design for Navigational Feedback Between Unmanned Underwater Vehicles (UUVs). *IEEE Journal of Oceanic Engineering.* **41**, 18–26 (2016)
5. R. P Rez-Alcocer, L. A. Torres-M Ndez, E. Olgu N-D Az, and A. A. Maldonado-Ram Rez, "Vision-Based Autonomous Underwater Vehicle Navigation in Poor Visibility Conditions Using a Model-Free Robust Control," *Journal of Sensors,* pp. 1–16, 2016
6. D. Park, J. Jung, K. Kwak, W. K. Chung, and J. Kim, "3D Underwater Localization using EM Waves Attenuation for UUV Docking," *IEEE Conference on Underwater Technology (UT),* 2017
7. Mur-Artal, R., Tardos, J.D.: ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras. *IEEE Transactions on Robotics.* **33**, 1255–1262 (2017)
8. Nawaf, M.M., Merad, D., Royer, J., Boï, J., Saccone, M., Ellefi, M.B., Drap, P.: Fast Visual Odometry for a Low-Cost Underwater Embedded Stereo System. *Sensors.* **18**, 2313 (2018)
9. Bellavia, F., Fanfani, M., Colombo, C.: Selective visual odometry for accurate AUV localization. *Autonomous Robots.* **41**, 133–143 (2017)
10. D. Li, *Ship motion and modeling*: Chinese Defense Industry Press, 2008
11. J. Kennedy and R. Eberhart, "Particle Swarm Optimization," *Proceedings of International Conference on Neural Networks,* 1995
12. Zhang, T., Li, D., Yang, C.: Study on impact process of AUV underwater docking with a cone-shaped dock. *Ocean Engineering.* **130**, 176–187 (2017)
13. Dormand, J.R., Prince, P.J.: A family of embedded Runge-Kutta formulae. *Journal of Computational and Applied Mathematics.* **6**, 19–26 (1980)
14. P. Lluis Negre, F. Bonin-Font and G. Oliver, "Cluster-Based Loop Closing Detection for Underwater SLAM in Feature-Poor Regions," in *IEEE International Conference on Robotics and Automation ICRA,*
A. Okamura, A. Menciassi, A. Ude, D. Burschka, D. Lee, F. Arrichiello, H. Liu, H. Moon, J. Neira, K. Sycara, K. Yokoi, P. Martinet, P. Oh, P. Valdastri, and V. Krovi, Eds., 2016, pp. 2589–2595
15. J. Bouguet, "Pyramidal Implementation of the Lucas Kanade Feature Tracker Description of the algorithm. intel corporation microprocessor research labs," *OpenCV documents,* 2000
16. W. Hou, A. D. Weidemann, D. J. Gray, and G. R. Fournier, "Imagery-derived modulation transfer function and its applications for underwater imaging,", 2007, pp. 22–28
17. Kim, A., Eustice, R.M.: Real-Time Visual SLAM for Autonomous Underwater Hull Inspection Using Visual Saliency. *IEEE Transactions on Robotics.* **29**, 719–733 (2013)
18. Chong-Yi, L., Ji-Chang, G., Run-Min, C., Yan-Wei, P., Bo, W.: Underwater Image Enhancement by Dehazing With Minimum Information Loss and Histogram Distribution Prior. *IEEE Transactions on Image Processing.* **25**, 5664–5677 (2016)
19. Zhang, Z.: A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* **22**, 1330–1334 (2000)
20. Scaramuzza, D., Fraundorfer, F.: Visual Odometry Part I: The First 30 Years and Fundamentals. *IEEE Robotics & Automation Magazine.* **18**, 80–92 (2011)
21. "The KITTI Vision Benchmark Suite,", 2012
22. Lin, M., Yang, C., Li, D.: An Improved Transformed Unscented FastSLAM With Adaptive Genetic Resampling. *IEEE Transactions on Industrial Electronics.* **66**, 3583–3594 (2019)
23. Liu, D., Duan, J., Shi, H.: A Strong Tracking Square Root Central Difference FastSLAM for Unmanned Intelligent Vehicle with Adaptive Partial Systematic Resampling. *IEEE Transactions on Intelligent Transportation Systems.* **11**, 3110–3120 (2016)

Mingwei Lin received the B.S. degree in mechanical engineering and automation from Fuzhou University, China, in 2014, and the Ph.D. degree in mechatronic engineering from Zhejiang University, China, in 2019, where he is currently a Postdoctoral Research Fellow with State Key Laboratory of Fluid Power and Mechatronic Systems.

His research interests include underwater localization, AUV docking technology, and their applications to ocean observation and exploration.

Canjun Yang received his B.S. and M.S. degrees in mechanical engineering from the Nanjing University of Aeronautics and Astronautics, China, in 1991 and 1994, respectively, and his Ph.D. degree in mechanical engineering from the Zhejiang University, China, in 1997.

Since 1997, he has been a faculty with the State Key Laboratory of Fluid Power and Mechatronic Systems at Zhejiang University. In 2004, he visited Minnesota University as an academic visitor. Since 2004, he has been a professor of institute of Mechatronic and Control Engineering in Zhejiang University. His research interests include navigation, man-machine intelligent system, mechatronic devices and their applications to deep-sea exploitation and exploration.

Mr. Yang was awarded the National Technology Invention Award (Second Prize) in 2009, the University Technology Invention Award (First Prize) of Ministry of Education of P. R. China in 2006, Science and Technology Award of Zhejiang Province (First Prize) in 2006, Zhejiang Youth Science and Technology Award in 2009. He was selected to the Program for New Century Excellent Talents in University (NCET) and New Century 151 Talent Project of Zhejiang Province (First Class). He is currently the member of National Committee Submersible of Standardization and the trustee of China Innovation Strategic Alliance of Rehabilitation Technical Aids.