symmetry

MDPI

Article

Design and Implementation of Underwater Robotic Systems for Visual–Inertial Trajectory Estimation and Robust Motion Control

Yangyang Wang [1,*], Tianzhu Gao [1], Yongqiang Zhao [2], Ziyu Liu [1], Hang Yu [2] and Xijun Du [1]

[1] Information Science and Technology College, Dalian Maritime University, Dalian 116026, China; 18341314091brain@gmail.com (T.G.); liuziyu1206@163.com (Z.L.); dxj@dlmu.edu.cn (X.D.)
[2] Marine Engineering College, Dalian Maritime University, Dalian 116026, China; yongqiang_zhao@dlmu.edu.cn (Y.Z.); yu2220240374@dlmu.edu.cn (H.Y.)
* Correspondence: yyw@dlmu.edu.cn

Abstract

Reliable trajectory estimation and precise motion control are the prerequisites for underwater robotic systems to perform complex autonomous tasks, which are essential for enhancing the operational efficiency of intelligent underwater facilities. However, the inherent asymmetry of underwater hydrodynamics, featureless images caused by complex environments, and the lack of high-frequency state feedback significantly hinder stable trajectory tracking and robust autonomous navigation. To address these challenges, this paper proposes an integrated autonomous navigation and robust control scheme for underwater robotic systems. Specifically, we first propose a visual–inertial trajectory estimation method for underwater robotic systems, which effectively overcomes the challenges of featureless images and provides consistent, real-time pose feedback for motion execution. Furthermore, we develop a hierarchical robust motion control strategy for autonomous underwater robots, which integrates model predictive control with incremental nonlinear dynamic inversion to achieve precise positioning performance and reliable operation under environmental disturbances. Finally, we design and implement a customized, highly integrated underwater robotic platform that integrates the proposed trajectory estimation and robust control modules, with its performance validated through extensive field experiments in underwater scenarios. The experimental results demonstrate that the proposed system can effectively achieve high-precision trajectory tracking and maintain operational stability, providing a comprehensive engineering solution for the autonomous navigation of underwater robots in complex environments.

Keywords: underwater robots; visual–inertial trajectory estimation; motion control

Check for updates

Academic Editor: Chengxi Zhang

Received: 11 March 2026
Revised: 29 March 2026
Accepted: 3 April 2026
Published: 6 April 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The exploration and intelligent management of underwater resources have become increasingly dependent on the deployment of autonomous underwater vehicles (AUVs), which serve as indispensable platforms for extending operational capabilities into remote and hazardous underwater domains [1,2]. For these systems to perform complex tasks such as infrastructure inspection and environmental monitoring, the ability to achieve precise motion control and reliable trajectory generation is a fundamental prerequisite [3]. However, the harsh and unpredictable nature of underwater environments presents significant challenges to achieving high-level autonomy. Specifically, the interconnected hurdles of reliable trajectory estimation in featureless scenarios and robust motion execution under hydrodynamic disturbances remain open problems in the field of underwater robotics [4].

Achieving reliable self-localization for AUVs remains a significant challenge because electromagnetic waves attenuate rapidly in water, which renders satellite-based positioning systems inapplicable for submerged operations [5]. While acoustic-based systems offer long-range stability, the practical application of the technology on agile platforms is often constrained by high hardware costs, low update rates, and substantial physical dimensions [6]. Visual sensing has consequently emerged as a promising, lightweight alternative for high-resolution state estimation, offering dense environmental information that is unavailable through acoustic sensors [7,8]. The efficacy of visual methods, however, is frequently compromised by the inherent optical properties of water, such as light scattering and absorption, which result in low-contrast imagery and feature sparsity [9]. In scenarios where visual tracking fails due to featureless regions or turbidity, the integration of an inertial measurement unit (IMU) becomes essential to maintain motion continuity [10]. The fusion of visual and inertial data enables the system to provide high-frequency, consistent trajectory estimation by utilizing the IMU to compensate for short-term visual loss [11]. Overcoming these perceptual limitations to provide reliable state feedback is a critical step toward enabling stable closed-loop control [12,13].

Reliable trajectory estimation establishes the indispensable perceptual basis for autonomous navigation, yet the estimation of the state serves only as a prerequisite for motion and cannot independently guarantee the precise execution of complex maneuvers in dynamic underwater settings [14]. While many existing underwater systems focus exclusively on localization and mapping, such approaches frequently overlook the sophisticated control strategies that are vital for high-precision trajectory tracking [15,16]. Conventional control methods often struggle with the highly non-linear and time-varying hydrodynamic forces, where the asymmetry of vehicle hydrodynamics and the complex coupling with unpredictable external currents lead to a significant degradation in execution accuracy [17]. These challenges are further compounded by the added mass effects and damping variations that are inherent to submerged operations. To bridge the gap between perception and execution, this work develops a unified architecture that tightly couples the aforementioned visual–inertial feedback with a hierarchical robust control module. Within this scheme, model predictive control (MPC) is employed to handle state and input constraints through an optimization-based approach, ensuring that the vehicle operates within safe physical limits. The sensitivity of the MPC algorithm to unmodeled dynamics and external disturbances, however, necessitates an additional layer of robustness to maintain operational stability. By incorporating incremental nonlinear dynamic inversion (INDI) as a high-frequency inner-loop controller, the system can rapidly compensate for hydrodynamic fluctuations and model inaccuracies by utilizing sensor-based acceleration feedback. Furthermore, to ensure that these sophisticated algorithms remain viable for practical field deployments, the entire framework is integrated onto a single, compact embedded system.

To address these challenges, this paper develops a unified trajectory estimation and motion control framework for AUVs. Unlike conventional methods that treat perception and control as isolated modules, our approach emphasizes the synergistic integration of robust state feedback and predictive control to counteract both perceptual uncertainty and dynamic disturbances. By tightly coupling a visual–inertial trajectory estimation method with a hierarchical robust control strategy that combines model predictive control and incremental nonlinear dynamic inversion, the proposed system achieves accurate position control and operational stability in complex underwater environments. The main contributions of this work are summarized as follows:

- We propose a visual–inertial trajectory estimation method for underwater robotic systems, which effectively overcomes the challenges of featureless images and provides consistent, real-time pose feedback for motion execution;

- We develop a hierarchical robust motion control strategy for autonomous underwater robots, which integrates MPC with INDI to achieve precise positioning performance and reliable operation under environmental disturbances;
- We design and implement a customized, highly integrated underwater robotic platform that integrates the proposed trajectory estimation and robust control modules, with its performance validated through extensive underwater experiments.

This article is structured in the following manner. Section 2 provides a brief review of some related works. Section 3 describes the integrated system architecture and the physical design of the underwater robotic platform. Section 4 presents the underwater visual–inertial trajectory estimation method, while Section 5 details the hierarchical robust motion control strategy. Section 6 provides a comprehensive evaluation of the developed framework through a multi-stage validation approach. Finally, Section 7 summarizes the conclusions and identifies potential directions for future research.

2. Related Work

Precise and reliable trajectory estimation is the cornerstone of underwater autonomy, enabling AUVs to navigate safely and execute intricate maneuvers in unstructured environments [18]. Historically, underwater navigation heavily relied on inertial dead reckoning. However, the lack of external reference signals leads to time-dependent drift that quickly compromises the integrity of the estimated path. Currently, the industry standard for bounding these errors involves the integration of acoustic sensors, such as Doppler Velocity Logs (DVL) and Ultra-Short Baseline (USBL) systems [19,20]. Researchers have proposed various fusion frameworks, such as Extended Kalman Filters, to align acoustic range measurements with vehicle kinematics for improved positioning consistency [21]. Despite their long-range reliability, the heavy physical footprint and high power requirements of professional-grade acoustic transducers present significant integration challenges for agile, low-cost AUVs [22]. Furthermore, in shallow or confined water, acoustic signals are frequently plagued by multi-path effects and low refresh rates, which are insufficient for the high-frequency feedback required by modern flight control systems.

The development of autonomous trajectory estimation has also been extensively documented in the context of terrestrial and aerial systems [23,24]. High-precision laser scanning, for instance, has become a standard tool for unmanned platforms to maintain motion consistency [25]. These systems rely on the geometric accuracy of point cloud registration to generate stable trajectory feedback. However, translating these methodologies to the subsea domain faces severe physical hurdles. The rapid absorption of high-intensity coherent light in conductive seawater causes laser energy to dissipate almost immediately, which severely limits the effective sensing range. Furthermore, the excessive power consumption and physical bulk of high-performance laser scanners pose significant integration challenges for compact AUVs. Consequently, optical cameras have emerged as a more viable and lightweight solution for fine-grained trajectory tracking. By leveraging abundant visual information, cameras provide a cost-effective pathway to achieve continuous ego-motion feedback while maintaining a significantly more compact spatial footprint and lower power consumption than laser scanners [26].

Although advanced feature tracking offers significant potential, the scarcity of reliable visual landmarks in open-water scenarios can cause significant drift or total loss of tracking. To address these limitations, researchers have integrated IMUs to provide continuous motion constraints and maintain trajectory integrity when visual information is temporarily degraded. Wang et al. [27] proposed an underwater self-localization method based on pseudo-3D vision–inertial sensing to achieve robust trajectory estimation for autonomous underwater vehicles in challenging subaqueous environments. Bingul et al. [28] introduced

a model-free framework that achieves robust trajectory estimation for autonomous underwater vehicles in the presence of ocean currents and external disturbances. Wang et al. [29] proposed a robust trajectory estimation framework based on stereo vision and inertial sensors that combines point and diagonal features to overcome the challenges of poor visibility in marine environments. Such multi-sensor trajectory estimation strategies, which combine high-frequency acceleration data with the long-term consistency of visual features, have demonstrated superior robustness in complex underwater environments [30]. By incorporating inertial cues to bridge gaps in visual feedback, these integrated frameworks enable more stable and resilient state estimation for autonomous underwater platforms. While such reliable trajectory estimation provides the necessary feedback, the development of robust motion control strategies is equally pivotal for effective task execution.

Early studies extensively employed Proportional–Integral–Derivative (PID) controllers due to their simplicity [28]. However, owing to the strong nonlinearities inherent in the dynamics of underwater robots and the complexity of their motion trajectories, PID-based methods often exhibit unsatisfactory performance. Subsequently, researchers introduced sliding mode control [31] and backstepping techniques [32] to address the nonlinear nature of hydrodynamics. However, their practical application is limited by insufficient disturbance rejection performance and issues such as input chattering. In addition, these approaches do not account for system constraints, which are of critical importance in multi-degree-of-freedom (multi-DoF) motion control. To further satisfy the requirements of autonomous operation, the control schemes of underwater robots must also be capable of executing synchronized multi-DoF motion while handling physical limits. MPC, as an optimal control approach [33], is well-suited to solving such multi-DoF problems and can explicitly account for various system constraints [34]. In a previous study [35], an MPC-based controller was implemented on a fully vectored propulsion underwater robot to improve maneuverability, and the 6-DoF control performance was experimentally validated in an indoor water tank. The effectiveness of MPC depends on the accuracy of the system model [36]. However, in contrast to kinematic models, the dynamics of underwater robots exhibit significant nonlinearity and uncertainty due to interactions between the rigid body and the surrounding fluid [37], as well as poorly characterized environmental disturbances. Such uncertainties can significantly limit the performance of the controller when relying solely on the predictive model.

Er et al. [38] presented a comprehensive survey that categorizes intelligent trajectory tracking and formation control methods for underactuated underwater vehicles. Within the context of these developments, INDI provides a robust alternative by incrementally updating the control input to locally linearize a nonlinear system [39]. This approach reduces the sensitivity of the control architecture to model mismatch and enhances robustness to external disturbances [40], leading to successful applications in the inner-loop control of unmanned aerial vehicles [41]. Moreover, INDI can directly utilize acceleration measurements from onboard sensors, such as low-cost IMUs, to realize feedback control [39], thereby reducing the reliance on external sensors. Despite these advantages, few studies have applied INDI to underwater robot control systems. In [42], INDI is used to accomplish a large-angle pitch-up maneuver of an AUV, representing one of the very few existing studies in this domain. Nevertheless, the positioning performance and trajectory tracking stability under persistent external disturbances have yet to be fully explored.

Unlike these existing studies, our work develops a unified architecture that couples high-frequency visual–inertial trajectory estimation with a cascaded MPC-INDI control module. By integrating real-time state feedback into a robust nonlinear control framework, this approach enables precise 6-DoF trajectory tracking for underwater robots even under significant model uncertainties and hydrodynamic disturbances. The proposed system

effectively bridges the gap between complex underwater perception and constrained motion execution, providing a reliable solution for underwater robotic operations.

3. System Architecture and Robotic Design

The proposed system architecture is engineered to provide a stable experimental robotic platform capable of supporting the real-time computational demands of visual–inertial trajectory estimation and motion control. To achieve this, the design philosophy prioritizes a modular framework that ensures structural integrity while facilitating efficient data transmission across heterogeneous hardware components. This section details the fundamental design choices and the kinematic and dynamic characterizations that define the behavior of the robotic system in underwater environments. By establishing a well-defined physical and theoretical baseline, we create the necessary conditions for implementing robust state estimation and predictive control algorithms.

3.1. Mechanical Structure and Modeling

The prototype employed in this research features a cylindrical hull with a length of 0.5 m and a diameter of 0.3 m, as illustrated in Figure 1. To ensure high maneuverability in complex underwater environments, the robotic platform adopts a fully vectored propulsion layout. This configuration enables each thruster to contribute to all degrees of freedom simultaneously, thereby enhancing propulsion efficiency and operational flexibility compared to traditional decoupled designs. Based on our previous work [35], the external structure has been further refined by optimizing the distance between the center of gravity and the center of buoyancy. This adjustment minimizes unnecessary restoring moments, facilitating more agile trajectory tracking while maintaining sufficient passive stability for the onboard sensing suite.

The coordinate system used to represent the robot motion is illustrated in Figure 1. The robot is modeled as a neutrally buoyant rigid body, with the origin of the body-fixed coordinate system O_b located at the geometric center of the robot, and its motion in the real world represented in the East–North–Up (ENU) coordinate system. To facilitate the design of the controller, the kinematic model describing the relationship between the pose and velocity of the robot is formulated in a decoupled form, as presented in the following equations.

$$\dot{\boldsymbol{\eta}}_t = \boldsymbol{J}_t(\boldsymbol{\eta}_r)\boldsymbol{\nu}_t \tag{1}$$

$$\dot{\boldsymbol{\eta}}_r = \boldsymbol{J}_r(\boldsymbol{\eta}_r)\boldsymbol{\nu}_r \tag{2}$$

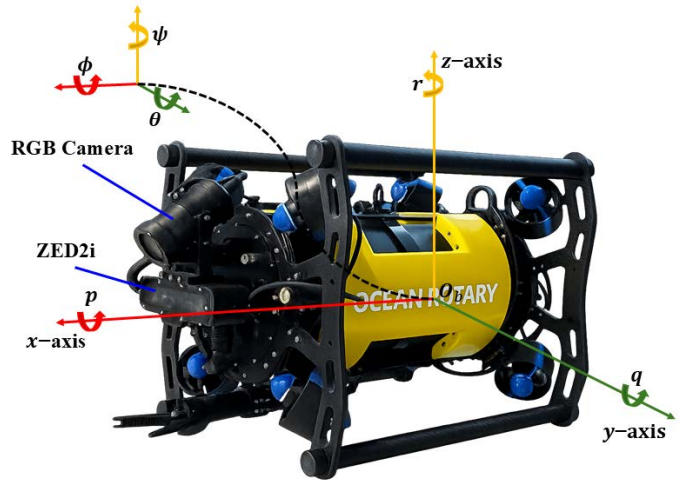
where $\boldsymbol{\eta}_t = [x, y, z]^\mathsf{T}$ and $\boldsymbol{\eta}_r = [q_w, q_x, q_y, q_z]^\mathsf{T}$ represent the position and attitude of the robot in the ENU coordinate system, $\boldsymbol{\nu}_t = [u, v, w]^\mathsf{T}$ and $\boldsymbol{\nu}_r = [p, q, r]^\mathsf{T}$ represent the linear and angular velocities of the robot in the body-fixed coordinate system. To avoid singularities, the attitude of the robot $[\phi, \theta, \psi]^\mathsf{T}$ is represented using unit quaternions $[q_w, q_x, q_y, q_z]^\mathsf{T}$. $\boldsymbol{J}_t(\boldsymbol{\eta}_r)$ and $\boldsymbol{J}_r(\boldsymbol{\eta}_r)$ denote the rotation matrices for the linear and angular velocities.

For the dynamics of the underwater robot, since the designed robot has closely located centers of gravity and buoyancy and typically operates in a low-speed regime [37], the restoring forces, as well as the Coriolis and centripetal forces, can be neglected. In addition, owing to the robot featuring three planes of symmetry and following the same assumptions as in [35], the inertia and added mass matrices, as well as the hydrodynamic damping coefficient matrix, are simplified as diagonal matrices. Furthermore, a simple linear relationship is assumed between the hydrodynamic damping and both the linear and angular velocities of the robot. Based on these considerations, the decoupled dynamic model of the underwater robot is described according to the formulation of Fossen [37] as follows.

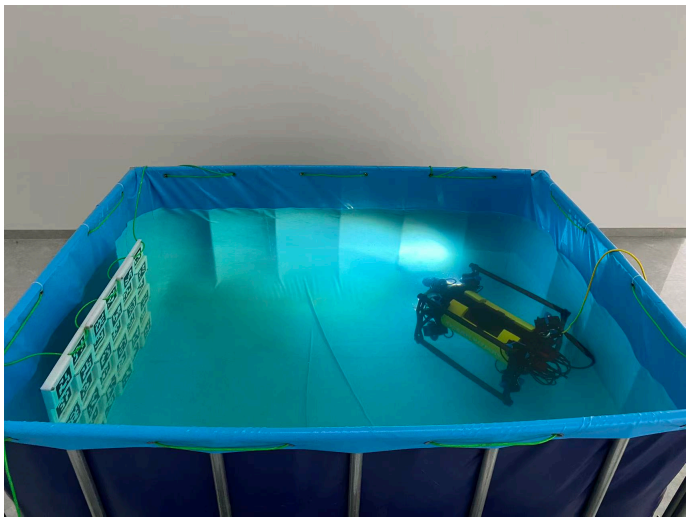
$$\boldsymbol{M}_t\dot{\boldsymbol{\nu}}_t + \boldsymbol{M}_t^A\dot{\boldsymbol{\nu}}_t + \boldsymbol{D}_t(\boldsymbol{\nu})\boldsymbol{\nu}_t = \boldsymbol{\tau}_t \tag{3}$$

$$\boldsymbol{M}_r\dot{\boldsymbol{\nu}}_r + \boldsymbol{M}_r^A\dot{\boldsymbol{\nu}}_r + \boldsymbol{D}_r(\boldsymbol{\nu})\boldsymbol{\nu}_r = \boldsymbol{\tau}_r \tag{4}$$

where $\boldsymbol{M}_t = \text{diag}(m, m, m)$ and $\boldsymbol{M}_r = \text{diag}(I_{xx}, I_{yy}, I_{zz})$ represent the mass matrix and the rotational inertia matrix of the robot, respectively. $\boldsymbol{M}_t^A = \text{diag}(X_{\dot{u}}, Y_{\dot{v}}, Z_{\dot{w}})$ and $\boldsymbol{M}_r^A = \text{diag}(K_{\dot{p}}, M_{\dot{q}}, N_{\dot{r}})$ denote the added mass matrices. $\boldsymbol{D}_t(\boldsymbol{\nu}) = \text{diag}(X_u, Y_v, Z_w)$ and $\boldsymbol{D}_r(\boldsymbol{\nu}) = \text{diag}(K_p, M_q, N_r)$ are the hydrodynamic damping coefficient matrices. $\boldsymbol{\tau}_t = [\tau_u, \tau_v, \tau_w]^\mathsf{T}$ and $\boldsymbol{\tau}_r = [\tau_p, \tau_q, \tau_r]^\mathsf{T}$ represent the control forces and torques acting along the corresponding axes. As in our previous work [35], the developed underwater robot is modeled as a standard cylinder. Accordingly, the physical parameters of the dynamic model in (3) and (4) can be obtained, as shown in Table 1.



(**a**)



(**b**)

Figure 1. The customized underwater robot testbed. (**a**) The fully vectored propulsion underwater robot and coordinate systems. (**b**) Real-world testing in a water tank.

Table 1. Parameters of our underwater robot.

Parameters	Nomenclature	Value (Unit)
Mass	m	25 kg
Rotational inertia, x-axis	I_{xx}	0.64 kg·m^2
Rotational inertia, y-axis	I_{yy}	1.34 kg·m^2

Table 1. *Cont.*

Parameters	Nomenclature	Value (Unit)
Rotational inertia, z-axis	I_{zz}	1.38 kg·m^2
Added mass, x-axis	$X_{\dot{u}}$	2.5 kg
Added mass, y-axis	$Y_{\dot{v}}$	27.90 kg
Added mass, z-axis	$Z_{\dot{w}}$	27.90 kg
Added mass, ϕ-axis	$K_{\dot{p}}$	0 kg·m^2
Added mass, θ-axis	$M_{\dot{q}}$	0.6 kg·m^2
Added mass, ψ-axis	$N_{\dot{r}}$	0.6 kg·m^2
Hydrodynamic damping, x-axis	X_u	27.36 kg/s
Hydrodynamic damping, y-axis	Y_v	67.32 kg/s
Hydrodynamic damping, z-axis	Z_w	67.32 kg/s
Hydrodynamic damping, ϕ-axis	K_p	0 kg·m^2/s
Hydrodynamic damping, θ-axis	M_q	0.28 kg·m^2/s
Hydrodynamic damping, ψ-axis	N_r	0.28 kg·m^2/s

3.2. Hardware Configuration and Software Framework

The hardware architecture of the underwater robot, as illustrated in Figure 2, consists of a sensor unit, a processor unit, an actuator unit, a communication unit, and a power unit. Compared with our previous work [35], the sensor unit has been redesigned and includes a ZED2i stereo camera (Stereolabs Inc., San Francisco, CA, USA), a monocular camera, and an Xsens MTi 630R AHRS (Xsens Technologies BV, Enschede, The Netherlands). The monocular camera is used to capture images of the environment surrounding the robot, while the ZED2i stereo camera provides depth images required for various operational tasks. In addition, the AHRS is installed inside the main control cabin to measure the attitude of the robot, encompassing roll, pitch, and yaw, along with the corresponding angular velocity and linear acceleration. The software framework is implemented in C++11 and Python 3.10 within the Robot Operating System (ROS) Humble middleware, with all onboard computations executed on an NVIDIA Jetson Orin NX 16 GB module.

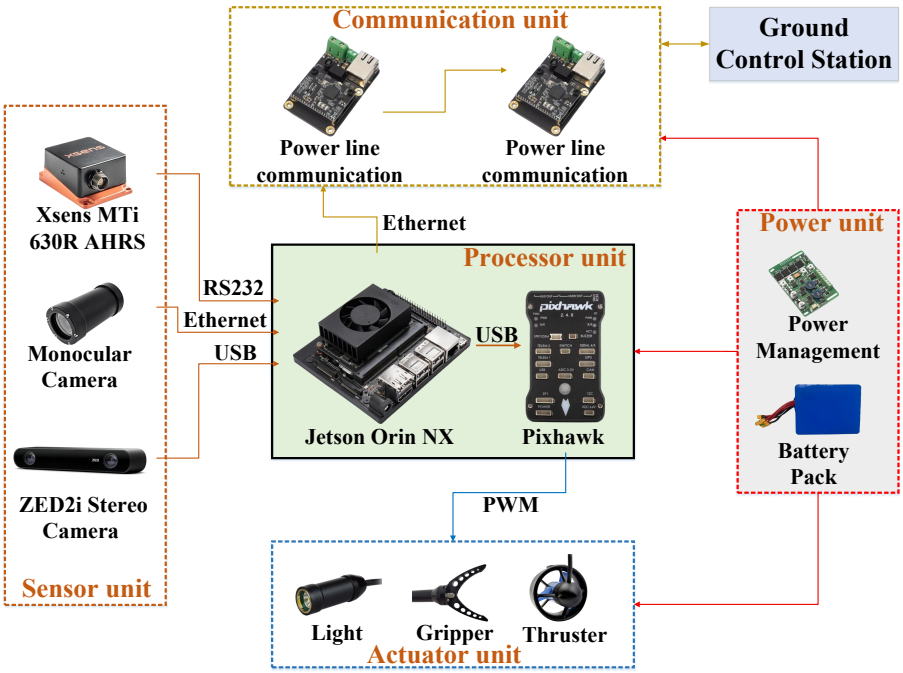


Figure 2. The fully vectored propulsion underwater robot and coordinate systems.

4. Underwater Visual Trajectory Estimation

The task of trajectory estimation is formulated as a factor graph optimization problem to provide high-frequency state feedback for the robust control architecture. In contrast to standard Euclidean parameterizations, the global state $\boldsymbol{\chi}$ is structured as an element of a product manifold. This formulation facilitates the simultaneous estimation of the temporal sequence of navigation nodes and the spatial landmark distribution. The state space is decomposed into three distinct functional subsets comprising the trajectory component $\mathbf{S}$, the extrinsic calibration $\mathbf{T}_{calib}$, and the landmark ensemble $\boldsymbol{\Lambda}$. This configuration is formally established through the following definition.

$$\boldsymbol{\chi} \triangleq \{\mathbf{S}, \mathbf{T}_{calib}, \boldsymbol{\Lambda}\}, \quad \text{where} \quad \mathbf{S} = \{\mathbf{s}_k\}_{k \in \mathcal{W}} \tag{5}$$

where $\mathcal{W}$ denotes the set of state indices within the current sliding window. Each navigation node $\mathbf{s}_k$ encompasses the essential kinematic variables required for the motion control law, as defined in the following vector.

$$\mathbf{s}_k = \left[(\boldsymbol{\eta}_{t,k}^{w})^{\mathsf{T}}, (\boldsymbol{\eta}_{r,k}^{w})^{\mathsf{T}}, (\boldsymbol{\nu}_{t,k}^{b})^{\mathsf{T}}, (\boldsymbol{\nu}_{r,k}^{b})^{\mathsf{T}}, \mathbf{b}_a^{\mathsf{T}}, \mathbf{b}_g^{\mathsf{T}} \right]^{\mathsf{T}} \tag{6}$$

where $\boldsymbol{\eta}_{t,k}^{w} \in \mathbb{R}^3$ and $\boldsymbol{\eta}_{r,k}^{w} \in S^3$ represent the position and unit quaternion orientation in the world frame w, while $\boldsymbol{\nu}_{t,k}^{b}$ and $\boldsymbol{\nu}_{r,k}^{b}$ denote the linear and angular velocities in the body frame b. The terms $\mathbf{b}_a$ and $\mathbf{b}_g$ correspond to the accelerometer and gyroscope biases. The calibration block $\mathbf{T}_{calib}$ aligns the camera optical center with the IMU coordinate system, and the landmark ensemble $\boldsymbol{\Lambda}$ contains the inverse depth of the tracked visual features.

The optimal trajectory is recovered by minimizing a total cost functional $\mathcal{J}(\boldsymbol{\chi})$ that integrates observations from multiple modalities. To differentiate the environmental impact on visual measurements, the optimization adopts a reliability-aware information scaling strategy [29]. The objective function is formulated through the following summation of squared Mahalanobis distances.

$$\mathcal{J}(\boldsymbol{\chi}) = \|\mathbf{r}_p\|_{\boldsymbol{\Sigma}_p}^2 + \sum_{i \in \mathcal{I}} \left\| \mathbf{r}_{imu}^{i,i+1}(\boldsymbol{\chi}) \right\|_{\boldsymbol{\Sigma}_i}^2 + \sum_{m \in \mathcal{M}} \rho\left(\|\mathbf{r}_{vis}^{m}(\boldsymbol{\chi})\|_{\boldsymbol{\Omega}_m(\lambda)}^2 \right) \tag{7}$$

where $\mathbf{r}_p$ represents the prior information from the marginalization of past states. The term $\mathbf{r}_{imu}^{i,i+1}$ captures the inertial residual between consecutive keyframes i and $i+1$. The visual constraint is defined over the set of observations $\mathcal{M}$, where the information matrix $\boldsymbol{\Omega}_m(\lambda)$ is dynamically scaled by a reliability parameter λ. This parameter accounts for the optical attenuation and scattering effects inherent in the underwater environment. To enhance the robustness against dynamic outliers, a Geman–McClure kernel $\rho(\cdot)$ is applied during the iterative optimization process. By executing this framework on the embedded platform, the system generates a stable stream of position and velocity estimates to drive the cascaded control architecture.

5. Robust Motion Control Strategy

The motion control task is formulated as a hierarchical tracking problem to map high-frequency state estimates onto optimal actuation commands. In alignment with the manifold-based estimation module, the control architecture is partitioned into two cascaded functional subsets comprising an outer-loop predictive guidance module and an inner-loop dynamic inversion module. The overall framework, illustrated in Figure 3, decouples the guidance and stabilization tasks to manage nonlinear hydrodynamics across different time scales.

5.1. Augmented MPC for Outer-Loop Control

For the position and attitude tracking of the underwater robot, the control objective is defined as $\boldsymbol{\eta_d} = [x, y, z, q_w, q_x, q_y, q_z]^\mathsf{T}$. In this study, the proposed hierarchical control architecture is applied exclusively to the position control of the robot, while the attitude control is still implemented using a conventional MPC approach. The kinematic Equations (1) and (2), as well as the dynamic Equation (4), are discretized using the forward Euler method. By introducing an augmented state, the model is then transformed into an incremental form. Consequently, the discrete-time state-space equation can be expressed as

$$\boldsymbol{x}_{k+1} = f(\boldsymbol{x}_k, \boldsymbol{u}_k, \delta t) \tag{8}$$

where $\boldsymbol{x} = [\Delta x, \Delta y, \Delta z, x, y, z, q_w, q_x, q_y, q_z, p, q, r]^\mathsf{T}$ denotes the augmented state vector of the robot at time k. Here, $\Delta x = x_{k+1} - x_k$, $\Delta y = y_{k+1} - y_k$ and $\Delta z = z_{k+1} - z_k$ denote the position increments. The augmented control input vector is defined as $\boldsymbol{u} = [\Delta u, \Delta v, \Delta w, \tau_p, \tau_q, \tau_r]^\mathsf{T}$ at time k, where $\Delta u = u_{k+1} - u_k$, $\Delta v = v_{k+1} - v_k$, and $\Delta w = w_{k+1} - w_k$ represent the velocity increments. Similarly, the control objective $\boldsymbol{\eta_d}$ is augmented to $\boldsymbol{x}_d$. At each time step, the augmented state $\boldsymbol{x}$ is predicted forward over N time steps, with each step size of δt. The proposed augmented MPC is then constructed in a quadratic form and the constraints $\boldsymbol{u}_{min}$ and $\boldsymbol{u}_{max}$ are imposed on the control inputs.

$$\begin{aligned} &\min_{\boldsymbol{u}} (\boldsymbol{x}_N - \boldsymbol{x}_d)^\mathsf{T} \boldsymbol{Q} (\boldsymbol{x}_N - \boldsymbol{x}_d) \\ &\quad + \sum_{k=0}^{N} \left\{ (\boldsymbol{x}_k - \boldsymbol{x}_d)^\mathsf{T} \boldsymbol{Q} (\boldsymbol{x}_k - \boldsymbol{x}_d) + \boldsymbol{u}_k^\mathsf{T} \boldsymbol{R} \boldsymbol{u}_k \right\} \\ &\text{s.t.} \begin{cases} \boldsymbol{x}_{k+1} = f(\boldsymbol{x}_k, \boldsymbol{u}_k, \delta t), \\ \boldsymbol{x}_0 = \boldsymbol{x}_{init}, \\ \boldsymbol{u}_{min} \leq \boldsymbol{u}_k \leq \boldsymbol{u}_{max}. \end{cases} \end{aligned} \tag{9}$$

where Q and R represent the weight matrix of the cost function. The augmented MPC is implemented using the open-source frameworks Acados [43] and CasADi [44], and operates at a control frequency of 20 Hz. The resulting velocity increments $[\Delta u, \Delta v, \Delta w]^\mathsf{T}$ are subsequently transformed into acceleration commands in the x, y, and z axes, which are then passed to the INDI controller for execution.

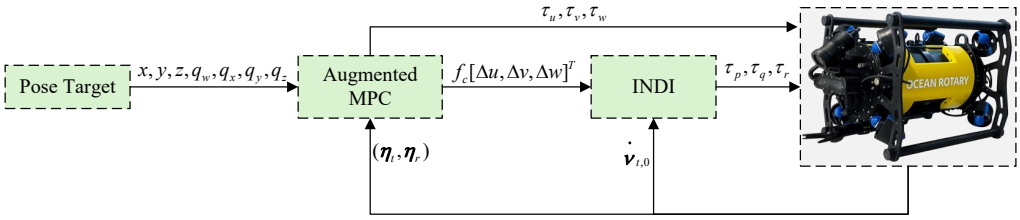


Figure 3. The hierarchical control architecture for the autonomous underwater robot.

5.2. INDI for Inner-Loop Control

INDI is an extension of nonlinear dynamic inversion (NDI). It regulates nonlinear systems through iterative updates of control input increments, thereby reducing reliance on an accurate system model. For the translational dynamics of the robot in (3), the acceleration can be expressed as follows:

$$\dot{\boldsymbol{\nu}}_t = (\boldsymbol{M}_t + \boldsymbol{M}_t^A)^{-1} (\boldsymbol{\tau}_t - \boldsymbol{D}_t(\boldsymbol{\nu}) \boldsymbol{\nu}_t) \tag{10}$$

To derive the control law of INDI, a first-order Taylor expansion is performed on the Equation (10). The linear acceleration of the robot is subsequently formulated as

$$\begin{aligned}\dot{\nu}_t \approx & (\boldsymbol{M}_t + \boldsymbol{M}_t^A)^{-1}[\boldsymbol{\tau}_{t,0} - \boldsymbol{D}_t(\boldsymbol{\nu_t})\boldsymbol{\nu}_{t,0}] \\ & + \frac{\partial}{\partial \boldsymbol{\tau}_t}\left[(\boldsymbol{M}_t + \boldsymbol{M}_t^A)^{-1}\boldsymbol{\tau}_t\right]\Big|_{\boldsymbol{\tau}_t = \boldsymbol{\tau}_{t,0}} (\boldsymbol{\tau}_t - \boldsymbol{\tau}_{t,0}) \\ & - \frac{\partial}{\partial \boldsymbol{\nu_t}}\left[(\boldsymbol{M}_t + \boldsymbol{M}_t^A)^{-1}\boldsymbol{D}_t(\boldsymbol{\nu_t})\boldsymbol{\nu}_t\right]\Big|_{\boldsymbol{\nu}_t = \boldsymbol{\nu}_{t,0}} (\boldsymbol{\nu}_t - \boldsymbol{\nu}_{t,0})\end{aligned} \tag{11}$$

where $\boldsymbol{\nu}_{t,0}$ and $\boldsymbol{\tau}_{t,0}$ denote the linear velocity and the control input vector acting along the translational directions at the current time step. Assuming that the thrust generated by the propulsion system is significantly greater than the hydrodynamic damping forces during the motion of the platform, Equation (11) can be simplified as follows:

$$\begin{aligned}\dot{\nu}_t \approx & (\boldsymbol{M}_t + \boldsymbol{M}_t^A)^{-1}[\boldsymbol{\tau}_{t,0} - \boldsymbol{D}_t(\boldsymbol{\nu_t})\boldsymbol{\nu}_{t,0}] \\ & + (\boldsymbol{M}_t + \boldsymbol{M}_t^A)^{-1}\boldsymbol{\tau}_t(\boldsymbol{\tau}_t - \boldsymbol{\tau}_{t,0})\end{aligned} \tag{12}$$

According to (10), $(\boldsymbol{M}_t + \boldsymbol{M}_t^A)^{-1}[\boldsymbol{\tau}_{t,0} - \boldsymbol{D}_t(\boldsymbol{\nu_t})\boldsymbol{\nu}_{t,0}]$ represents the current linear acceleration, denoted as $\dot{\nu}_{t,0}$. Equation (12) can then be written as

$$\dot{\nu}_t \approx \dot{\nu}_{t,0} + (\boldsymbol{M}_t + \boldsymbol{M}_t^A)^{-1}(\boldsymbol{\tau}_t - \boldsymbol{\tau}_{t,0}) \tag{13}$$

Accordingly, the control input increment $\delta\boldsymbol{\tau} = \boldsymbol{\tau}_t - \boldsymbol{\tau}_{t,0}$ can be determined based on the current acceleration $\dot{\boldsymbol{\nu}}_{t,0}$ and desired acceleration $\dot{\boldsymbol{\nu}}_t$ of the robot as follows:

$$\delta\tau_t \approx (\boldsymbol{M}_t + \boldsymbol{M}_t^A)(\dot{\boldsymbol{\nu}}_t - \dot{\boldsymbol{\nu}}_{t,0}) \tag{14}$$

Finally, the control inputs acting along the x, y, and z directions of the robot are obtained as $\boldsymbol{\tau}_t = \boldsymbol{\tau}_{t,0} + \delta\boldsymbol{\tau}$.

6. Experiments and Analysis

This section presents a comprehensive evaluation of the developed underwater robotic system, spanning from trajectory estimation accuracy to motion control robustness. The experimental validation is organized to demonstrate the integrated performance of the system, utilizing public datasets for standardized benchmarking, an indoor water tank for real-world functionality testing, and high-fidelity numerical simulations for rigorous quantitative control analysis.

6.1. Evaluation of Trajectory Estimation

This subsection focuses on validating the visual–inertial trajectory estimation module through a structured approach. The evaluation transitions from quantitative benchmarking against ground-truth data to qualitative verification of feature-tracking stability within a representative underwater tank environment to ensure the reliability of the algorithm in practical operations.

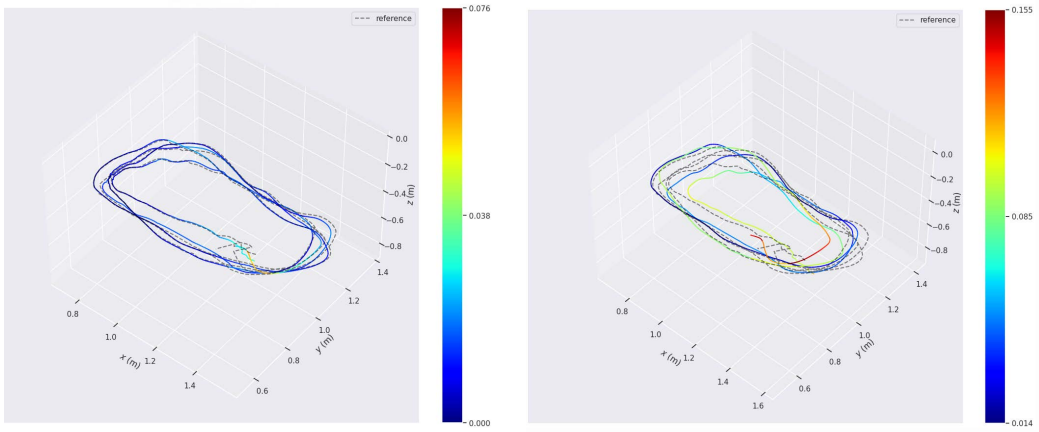
6.1.1. Estimation Evaluation Framework

The performance of the proposed trajectory estimation module is assessed through a combination of quantitative benchmarking on the HAUD [45] and AQUALOC datasets [46], as well as qualitative verification in a real-world water tank. The Absolute Pose Error (APE) is employed as the primary quantitative metric to evaluate the global consistency of the estimated trajectory. The root mean square error (RMSE) of the APE is calculated to quantify the overall localization accuracy. These sequences incorporate challenging underwater conditions such as suspended particles and low-texture substrates, which serve to evaluate the integration of IMU pre-integration and visual features for suppressing cumulative drift. To ensure the results reflect the real-time operational constraints of the robotic system,

the algorithms are executed on the NVIDIA Jetson Orin NX platform. Furthermore, the trajectory estimation is tested within an experimental water tank to evaluate the practical reliability of the system in a real-world underwater environment. While the absence of ground-truth infrastructure in the tank limits numerical benchmarking, this stage serves as a critical verification of the capability of the system to generate consistent and stable trajectory estimation under authentic light attenuation and scattering effects.

6.1.2. Results and Discussion

The quantitative performance of the trajectory estimation is illustrated in Figure 4. To further evaluate the precision and stability of the system, the statistical characteristics of the APE for the proposed method are summarized in Table 2. The experimental results confirm that the proposed framework maintains consistent and continuous trajectory estimation throughout the tested HAUD sequences. The quantitative evaluation of the APE for the different methods is summarized in Table 3. Furthermore, the generalization of the proposed framework is evaluated through extensive quantitative benchmarking on the AQUALOC dataset, as illustrated in Figure 5. These results demonstrate that the integration of inertial pre-integration with reliability-aware visual observations allows the system to adapt to challenging underwater scenarios. Finally, runtime analysis on the NVIDIA Jetson Orin NX platform shows that the system maintains a consistent processing frequency of 30 Hz with an average CPU utilization of 44.2%. Consequently, the proposed method ensures reliable and real-time trajectory estimation even when the robotic platform performs complex maneuvers within these diverse and degraded scenarios.



(**a**) (**b**)

Figure 4. Comparison of the estimated trajectories and ground truth for sequences in the HAUD dataset, illustrating the APE performance. (**a**) Our method with loop closure detection. (**b**) Our method without loop closure detection.

Table 2. Statistical evaluation of the trajectory errors for the proposed method.

Trajectory Errors	RMSE (m)	Mean APE (m)	Max APE (m)	Min APE (m)
With Loop	0.0163	0.0136	0.0760	0.0005
Without Loop	0.0773	0.0695	0.1553	0.0143

Table 3. Comparison of the trajectory error.

Method	VINS-Fusion [7]	ORB-SLAM2 [47]	ORB-SLAM3 [48]	Ours
RMSE (m)	0.050	0.032	0.019	0.016

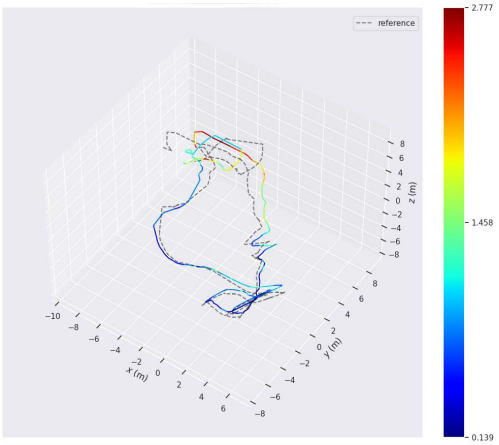
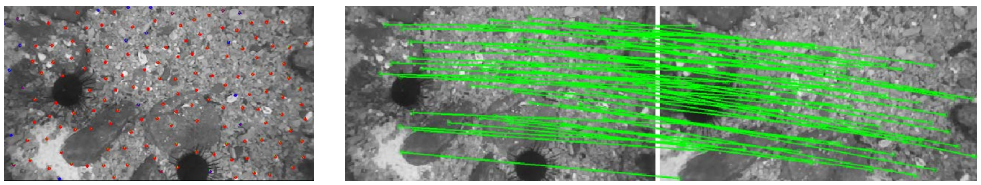


Figure 5. Comparison of the estimated trajectories and ground truth for sequences in the AQUALOC dataset, illustrating the APE performance.

To evaluate the practical effectiveness of the proposed system in a real-world underwater environment, field trials are conducted in a large-scale experimental tank with dimensions of 23 × 8 × 2 m^3, as shown in Figure 6. This experimental environment is specifically designed as a validation platform where the robustness of the visual–inertial trajectory estimation under diverse maneuvers is assessed. The trials focus on the capability of the algorithm to maintain continuous trajectory updates while navigating through authentic light attenuation and scattering effects. As illustrated in Figure 7, the performance of visual feature processing is evaluated through the features tracked by the camera and the descriptor matching for robust feature tracking. This robust data association ensures that the trajectory estimation module remains stable even when visual features become sparse or distorted by the water medium.



Figure 6. The large-scale experimental underwater tank.



(**a**) (**b**)

Figure 7. Performance of visual feature processing in the underwater environment. (**a**) Features tracked by the camera. (**b**) Descriptor matching for robust feature tracking, where the green lines represent the temporal matching relationship of landmarks between adjacent image frames.

The qualitative results of the trajectory estimation in the experimental tank are illustrated in Figure 8. The results confirm that the proposed framework maintains a smooth and continuous trajectory throughout the entire underwater maneuver. The integration of inertial measurements and visual features prevents the occurrence of sudden jumps or interruptions in the estimated path, despite the challenging acoustic and optical conditions within the tank environment. This consistency in the generated trajectory proves the practical reliability of the system for the navigation of the robotic platform in real-world underwater scenarios.

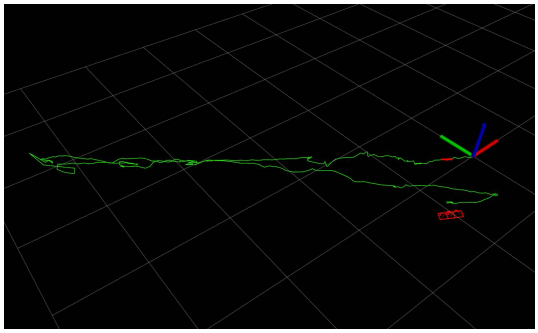


Figure 8. Qualitative results of trajectory estimation in the experimental tank. The green curve represents the AUV trajectory, while the red indicators denote the camera orientations.

6.2. Results and Analysis of Motion Control

This subsection focuses on evaluating the motion control system through high-fidelity numerical simulations. These simulations provide a rigorous platform for the quantitative benchmarking of tracking errors and control effort under controlled conditions. This evaluation confirms the stability of the control laws under the complex hydrodynamic conditions encountered in the underwater environment.

6.2.1. Control Validation Framework

Considering that open-water tank environments often lack the ground-truth infrastructure required for precise quantitative benchmarking, numerical simulations are conducted to serve as a rigorous validation tool for evaluating the controller performance under controlled conditions. This simulation-based approach allows for a precise quantification of tracking errors and control effort, which complements the subsequent qualitative observations from real-world testing. The simulation environment is implemented using Python and ROS on a laptop equipped with an Intel (R) Core (TM) i5-10210U CPU and 16 GB of RAM, with the robotic system's physical parameters consistent with those provided in Table 1. In the simulation studies, we assume that the robot dynamics are perfectly known. Accordingly, the model used for MPC prediction and optimization is identical to that employed in the simulation environment. To demonstrate the capability of the proposed controller in achieving 6-DoF motion control and its robustness under environmental uncertainties, two distinct scenarios are designed to compare its performance with a conventional MPC approach [35]. First, a disturbance-free scenario is considered to verify the fundamental feasibility of the control law. Subsequently, a disturbance-included scenario is introduced to evaluate the disturbance rejection capability, where zero-mean Gaussian noise with a standard deviation equal to 30% of the maximum control input is applied to the robotic system at each sampling instant to emulate the complex water currents inherent in underwater operations.

6.2.2. Results and Discussion

Figures 9–11 present the simulation results of the proposed control strategy, along with a comparison to the conventional MPC approach. In this simulation, the robot starts from the initial position $[x_0, y_0, z_0] = [0, 0, 0]$ m and tracks a helical trajectory defined by $[x_d, y_d, z_d] = [10\cos(\frac{2\pi}{60}t), 10\sin(\frac{2\pi}{60}t), -0.2t]$ m. Meanwhile, the initial attitude $[\phi_0, \theta_0, \psi_0] = [0, 0, 0]$ rad is regulated to the desired attitude, defined as $[\phi_d, \theta_d, \psi_d] = [\frac{\pi}{3}, \frac{\pi}{3}, 0]$ rad. In the proposed control strategy, the parameters of the outer-loop augmented MPC are set as $Q = [0.01, 0.01, 0.01, 1, 1, 1, 1, 0.5, 0.5, 0.5, 0.01, 0.01, 0.01]$ and $R = [0.1, 0.1, 0.1, 0.1, 0.1, 0.1]$, with a prediction horizon of $N = 20$. To ensure that the control inputs do not exceed the operational capabilities of the robot, the constraints on each DoF are set to 60% of the maximum allowable input. It is worth noting that all control inputs are normalized for computational convenience and to enable a fair performance comparison among different methods.

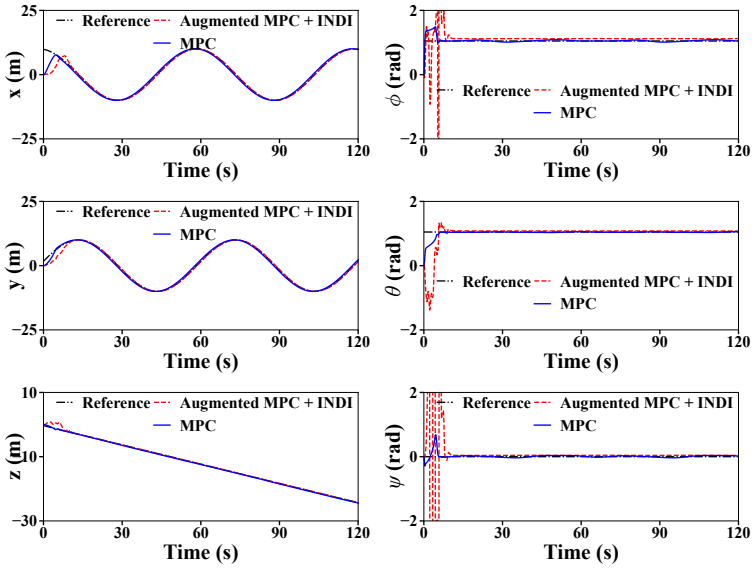


Figure 9. 6-DoF tracking results of underwater robot without disturbances.

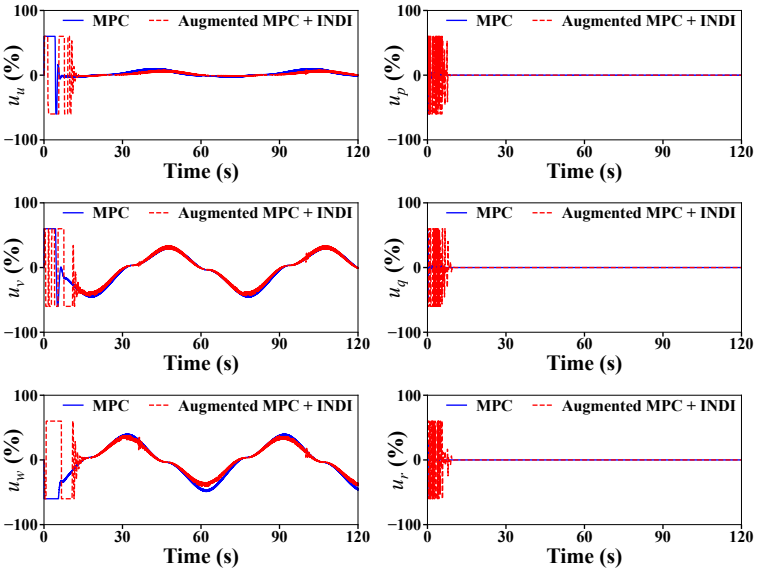


Figure 10. Percentage of force and torque inputs on each DoF without disturbances.

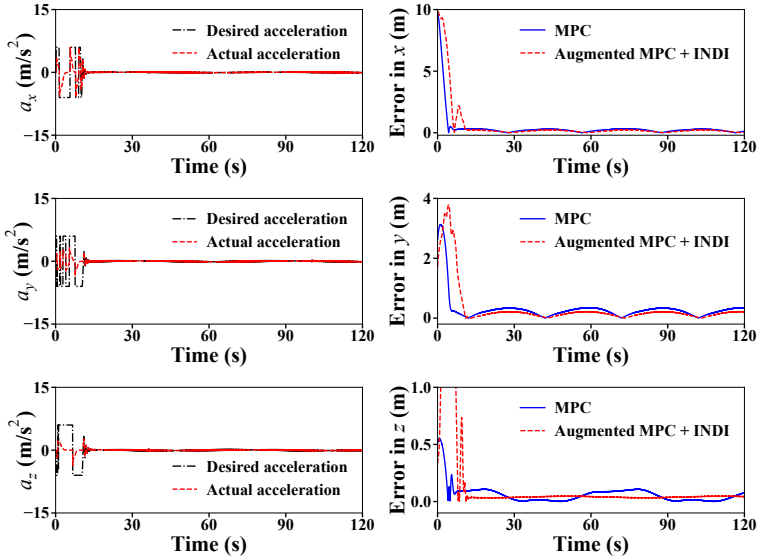


Figure 11. Acceleration tracking results of inner loop and tracking errors in the x, y, and z directions without disturbances.

From Figure 9, the position and attitude responses of the robot during the simulation can be observed. After a brief oscillation, the robot is able to track the desired position and attitude within 15 s and maintain stability. Figure 10 presents the percentage of maximum force and torque applied in each DoF. In the proposed hierarchical control architecture, the force inputs in the x, y, and z directions are provided by the inner-loop INDI controller, whereas in the comparative method they are directly obtained from the MPC controller. It can be seen that, under the constraint that the control inputs in each DoF do not exceed 60% of the maximum allowable values, the proposed control strategy is able to effectively track the reference trajectory.

The RMSE of the proposed control method and the conventional MPC method was then calculated. After 15 s of tracking, the RMSE values for the proposed controller using augmented MPC and INDI are $[0.14, 0.15, 0.04]$ m and $[4.08^\circ, 1.92^\circ, 2.40^\circ]$ in the 6-DoF, respectively. In comparison, the conventional MPC controller yields RMSE values of $[0.21, 0.24, 0.05]$ m and $[1.29^\circ, 0.50^\circ, 1.46^\circ]$, respectively. This indicates that the proposed controller reduces the position error in the x, y, and z directions by 33.4% relative to the conventional MPC. The inner-loop acceleration tracking results based on INDI, along with the tracking errors in the x, y, and z directions, are shown in Figure 11. It is worth noting that the proposed hierarchical architecture is applied only to the x, y, and z directions of the underwater robot, while the motions in the roll, pitch, and yaw directions are still controlled using the same paradigm as the conventional MPC. Therefore, the numerical simulation results demonstrate a certain advantage in position tracking accuracy compared with the conventional MPC method. In addition, the proposed control framework shows a slight degradation in attitude tracking accuracy. The main reason is that, although the translational and rotational dynamics are decoupled in the modeling process, the MPC still uses identical control parameters for the 6-DoF motion of the robot. Consequently, while preserving position tracking performance, a trade-off arises that leads to reduced attitude control performance. This limitation can be alleviated by further decoupling the design of position and attitude controllers.

Figures 12–14 illustrate the simulation results of the proposed control strategy under disturbance conditions. In this simulation, the robot starts from the same initial position $[x_0, y_0, z_0] = [0, 0, 0]$ m and initial attitude $[\phi_0, \theta_0, \psi_0] = [0, 0, 0]$ rad, and tracks a helical trajectory defined by $[x_d, y_d, z_d] = [10\cos(\frac{2\pi}{60}t), 10\sin(\frac{2\pi}{60}t), -0.2t]$ m, while regulating

its attitude to $[\phi_d, \theta_d, \psi_d] = [\frac{\pi}{3}, \frac{\pi}{3}, 0]$ rad. The same control parameters as those used in the disturbance-free simulation are adopted throughout the entire tracking process. The pose tracking results and the control inputs in each DoF under disturbance conditions are shown in Figures 12 and 13. It can be observed that, despite the presence of significant random disturbances, under the constraint that the force and torque inputs in each DoF do not exceed 60% of their maximum allowable values, the proposed hierarchical control architecture is still able to track the desired pose within 15 s and maintain stability.

The inner-loop INDI acceleration tracking results under disturbance conditions, as well as the tracking errors in the x, y, and z directions of the proposed control strategy, are illustrated in Figure 14. Similarly, the RMSE is calculated after 15 s from the start of tracking. Under disturbance conditions, the proposed controller achieves RMSE values of $[0.15, 0.15, 0.04]$ m and $[8.68^\circ, 3.41^\circ, 7.12^\circ]$, compared with $[0.20, 0.24, 0.05]$ m and $[4.29^\circ, 1.89^\circ, 4.18^\circ]$ obtained by the conventional MPC controller. The proposed control strategy reduces the position tracking error by 28.9%, demonstrating its robustness.

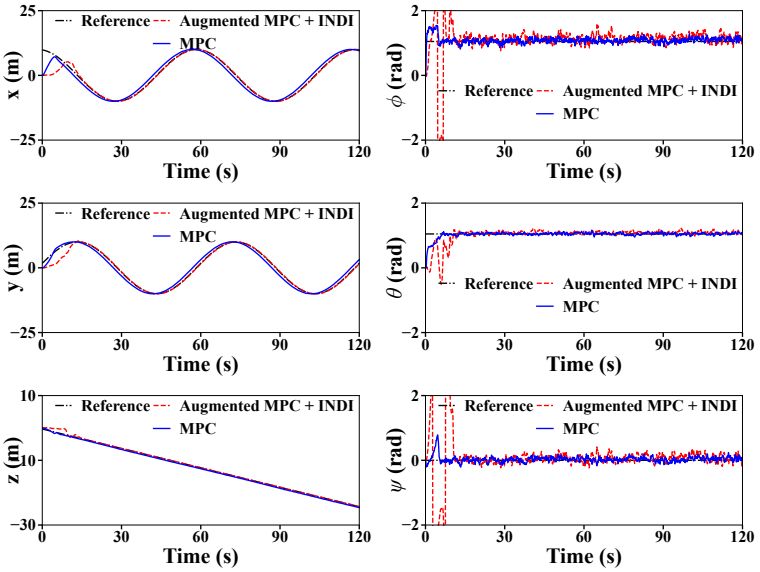


Figure 12. 6-DoF tracking results of underwater robot with disturbances.

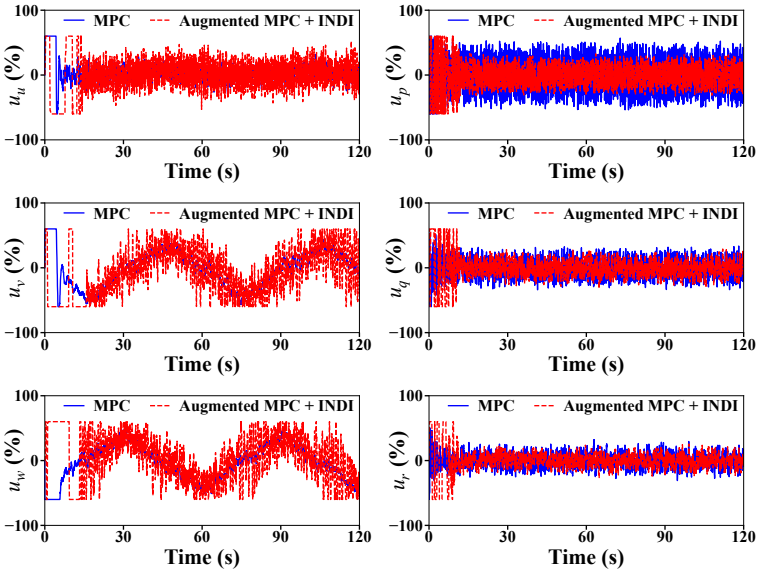


Figure 13. Percentage of force and torque inputs on each DoF with disturbances.

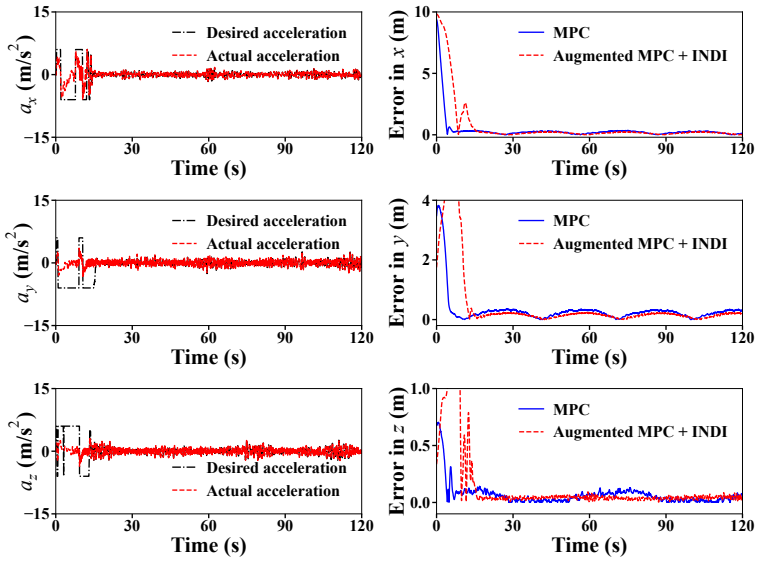


Figure 14. Acceleration tracking results of inner loop and tracking errors in the x, y, and z directions with disturbances.

7. Conclusions

This paper presents an integrated autonomous navigation and robust control scheme designed to address the challenges of featureless images and the asymmetry of underwater hydrodynamics. A significant contribution of this work is the development of a customized, highly integrated underwater robotic platform, which serves as a versatile hardware foundation for the proposed algorithms. The framework combines a visual–inertial trajectory estimation method with a hierarchical robust motion control strategy, integrating model predictive control and incremental nonlinear dynamic inversion. The trajectory estimation module effectively overcomes the limitations of complex underwater scenes to provide consistent and real-time pose feedback, achieving an RMSE of 0.016m in the underwater experiments. Extensive field experiments validate the performance of the complete system, demonstrating that the proposed engineering solution successfully maintains operational stability and robust navigation for intelligent underwater facilities.

Despite these contributions, the experimental validation also highlights areas for future improvement. The stability of the trajectory estimation module remains susceptible to extreme visual degradation and dynamic disturbances, which can occasionally lead to data association challenges. Furthermore, the performance of the hierarchical control law is constrained by the simplified modeling of complex thruster dynamics and the unpredictable nature of strong external currents. Future work will focus on the incorporation of adaptive control algorithms to further enhance the dynamic resilience of the controller against wind, waves, and currents, while optimizing the visual frontend to support more robust feature tracking in highly turbid water. In this context, exploring advanced optimization techniques represents a promising direction for further enhancing the learning and adaptive capabilities of complex underwater robotic systems operating under significant uncertainty. These advancements will aim to further refine the autonomy of underwater robotic platforms, providing a more reliable and scalable solution for complex underwater operations.

Author Contributions: Conceptualization, methodology, visualization, funding acquisition, writing—original draft preparation, Y.W.; validation, software, data curation, review and editing, T.G.; formal analysis, Y.Z.; resource, review and editing, Z.L.; data curation, H.Y. and X.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AUV	Autonomous Underwater Vehicle
IMU	Inertial Measurement Units
MPC	Model Predictive Control
INDI	Incremental Nonlinear Dynamic Inversion
DVL	Doppler Velocity Log
USBL	Ultra-Short Baseline
PID	Proportional–Integral–Derivative
DoF	Degree-of-Freedom
ROS	Robot Operating System
RMSE	Root Mean Square Error

References

1. Zhuang, Y.; Wu, C.; Wu, H.; Zhang, Z.; Xu, H.; Jia, Q.; Li, L. Event coverage hole repair algorithm based on multi-AUVs in multi-constrained three-dimensional underwater wireless sensor networks. *Symmetry* **2020**, *12*, 1884. [CrossRef]
2. Chen, G.; Du, G.; Yang, C.; Xu, Y.; Wu, C.; Hu, H.; Dong, F.; Zeng, J. An underwater visual SLAM system with adaptive image enhancement. *Ocean Eng.* **2025**, *326*, 120896. [CrossRef]
3. Rahman, S.; Quattrini Li, A.; Rekleitis, I. SVIn2: A multi-sensor fusion-based underwater SLAM system. *Int. J. Robot. Res.* **2022**, *41*, 1022–1042. [CrossRef]
4. Yan, X.; Chang, S.; Wang, X.; Zhang, L.; Liu, J. A dual-stage coverage path planning method for bathymetric survey using an AUV in graph-based SLAM framework considering positioning uncertainty. *Ocean Eng.* **2024**, *312*, 119252. [CrossRef]
5. Maurelli, F.; Krupiński, S.; Xiang, X.; Hernandez, J.D.; Zhao, S.S. AUV localisation: A review of passive and active techniques. *Int. J. Intell. Robot. Appl.* **2022**, *6*, 246–269. [CrossRef]
6. Abu, A.; Diamant, R. A SLAM approach to combine optical and sonar information from an AUV. *IEEE Trans. Mob. Comput.* **2024**, *23*, 7714–7724. [CrossRef]
7. Qin, T.; Li, P.; Shen, S. VINS-Mono: A robust and versatile monocular visual-inertial state estimator. *IEEE Trans. Robot.* **2018**, *34*, 1004–1020. [CrossRef]
8. Li, M.; He, J.; Wang, Y.; Wang, H. End-to-end RGB-D SLAM with multi-MLPs dense neural implicit representations. *IEEE Robot. Autom. Lett.* **2023**, *8*, 7138–7145. [CrossRef]
9. Yang, D.; Leonard, J.J.; Girdhar, Y. SeaSplat: Representing underwater scenes with 3D Gaussian splatting and a physically grounded image formation model. In Proceedings of the 2025 IEEE International Conference on Robotics and Automation (ICRA), Atlanta, GA, USA, 19–23 May 2025; pp. 7632–7638.
10. Hu, C.; Zhu, S.; Liang, Y.; Song, W. Tightly-coupled visual-inertial-pressure fusion using forward and backward IMU preintegration. *IEEE Robot. Autom. Lett.* **2022**, *7*, 6790–6797. [CrossRef]
11. He, J.; Li, M.; Wang, Y.; Wang, H. PLE-SLAM: A visual-inertial SLAM based on point-line features and efficient IMU initialization. *IEEE Sens. J.* **2025**, *25*, 6801–6811. [CrossRef]
12. Tang, R.; Qi, L.; Ye, S.; Li, C.; Ni, T.; Guo, J.; Liu, H.; Li, Y.; Zuo, D.; Shi, J.; et al. Three-dimensional path planning for AUVs based on interval multi-objective secretary bird optimization algorithm. *Symmetry* **2025**, *17*, 993. [CrossRef]
13. Piao, Z.; Sun, S.; Chen, Y.; Ju, M. Finite-time control for automatic berthing of pod-driven unmanned surface vessel with an event-triggering mechanism. *Symmetry* **2024**, *16*, 1575. [CrossRef]
14. Paull, L.; Saeedi, S.; Seto, M.; Li, H. AUV navigation and localization: A review. *IEEE J. Ocean. Eng.* **2014**, *39*, 131–149. [CrossRef]
15. Vial, P.; Palomeras, N.; Solà, J.; Carreras, M. Underwater pose SLAM using GMM scan matching for a mechanical profiling sonar. *J. Field Robot.* **2024**, *41*, 511–538. [CrossRef]
16. Qi, C.; Ma, T.; Li, Y.; Ling, Y.; Liao, Y.; Jiang, Y. A multi-AUV collaborative mapping system with bathymetric cooperative active SLAM algorithm. *IEEE Internet Things J.* **2025**, *12*, 12441–12452. [CrossRef]
17. Bucci, A.; Franchi, M.; Ridolfi, A.; Secciani, N.; Allotta, B. Evaluation of UKF-based fusion strategies for autonomous underwater vehicles multisensor navigation. *IEEE J. Ocean. Eng.* **2023**, *48*, 1–26. [CrossRef]

18. Xu, S.; Zhang, K.; Wang, S. AQUA-SLAM: Tightly coupled underwater acoustic-visual-inertial SLAM with sensor calibration. *IEEE Trans. Robot.* **2025**, *41*, 2785–2803. [CrossRef]
19. Aparicio, J.; Álvarez, F.J.; Hernández, Á.; Holm, S. A survey on acoustic positioning systems for location-based services. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 8505336. [CrossRef]
20. Chen, X.; Bian, H.; Li, F.; Wang, R.; Hu, Y.; Li, J. Time-varying current estimation method for SINS/DVL integrated navigation based on augmented observation algorithm. *Symmetry* **2025**, *17*, 1881. [CrossRef]
21. Alexandris, C.; Papageorgas, P.; Piromalis, D. Positioning systems for unmanned underwater vehicles: A comprehensive review. *Appl. Sci.* **2024**, *14*, 9671. [CrossRef]
22. Wang, Y.; Ma, X.; Wang, J.; Hou, S.; Dai, J.; Gu, D.; Wang, H. Robust AUV visual loop-closure detection based on variational autoencoder network. *IEEE Trans. Ind. Inform.* **2022**, *18*, 8829–8838. [CrossRef]
23. Huang, R.; Xue, H.; Pagnucco, M.; Salim, F.D.; Song, Y. Vision-based multi-future trajectory prediction: A survey. *IEEE Trans. Neural Netw. Learn. Syst.* **2025**, *36*, 13691–13708. [CrossRef] [PubMed]
24. He, J.; Li, M.; Wang, Y.; Wang, H. OVD-SLAM: An online visual SLAM for dynamic environments. *IEEE Sens. J.* **2023**, *23*, 13210–13219. [CrossRef]
25. Guo, X.; Chen, J.; Zhang, Y.; Hu, Y.; Chen, L.; Xia, C.; Xia, Y.; Wang, Z. Adaptive dynamic measurement, trajectory correction, and error evaluation method in MEMS LiDAR system. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 8508710. [CrossRef]
26. Wang, Y.; Liu, X.; Gu, D.; Wang, J.; Fu, X. Depth-consistent monocular visual trajectory estimation for AUVs. *IEEE Internet Things J.* **2025**, *12*, 14909–14920. [CrossRef]
27. Wang, Y.; Ma, X.; Wang, J.; Wang, H. Pseudo-3D vision-inertia based underwater self-localization for AUVs. *IEEE Trans. Veh. Technol.* **2020**, *69*, 7895–7907. [CrossRef]
28. Bingul, Z.; Gul, K. Intelligent-PID with PD feedforward trajectory tracking control of an autonomous underwater vehicle. *Machines* **2023**, *11*, 300. [CrossRef]
29. Wang, Y.; Gu, D.; Ma, X.; Wang, J.; Wang, H. Robust real-time AUV self-localization based on stereo vision-inertia. *IEEE Trans. Veh. Technol.* **2023**, *72*, 7160–7170. [CrossRef]
30. He, L.; Xie, M.; Zhang, Y. A review of path following, trajectory tracking, and formation control for autonomous underwater vehicles. *Drones* **2025**, *9*, 286. [CrossRef]
31. Zhu, Q.; Shang, H.; Lu, X.; Chen, Y. Adaptive sliding mode tracking control of underwater vehicle-manipulator systems considering dynamic disturbance. *Ocean Eng.* **2024**, *291*, 116300. [CrossRef]
32. Chen, H.; Tang, G.; Wang, S.; Guo, W.; Huang, H. Adaptive fixed-time backstepping control for three-dimensional trajectory tracking of underactuated autonomous underwater vehicles. *Ocean Eng.* **2023**, *275*, 114109. [CrossRef]
33. Grüne, L.; Pannek, J. *Nonlinear Model Predictive Control*; Springer: Cham, Switzerland, 2017.
34. Bhat, S.; Stenius, I. Controlling an underactuated AUV as an inverted pendulum using nonlinear model predictive control and behavior trees. In Proceedings of the 2023 IEEE International Conference on Robotics and Automation (ICRA), London, UK, 29 May–2 June 2023.
35. Gao, T.; Luo, Y.; Lv, C.; Luo, W.; Fu, X.; Zhao, N.; Luo, X.; Shen, Y. Model predictive control for an autonomous underwater robot with fully vectored propulsion. In Proceedings of the 2024 IEEE International Conference on Robotics and Automation (ICRA), Yokohama, Japan, 13–17 May 2024; pp. 10482–10488.
36. Gao, T.; Luo, Y.; Zhao, N.; Wang, J.; Yan, Y.; Fu, X.; Luo, X.; Shen, Y. Data-driven MPC for attitude control of autonomous underwater robot. In Proceedings of the 2025 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Hangzhou, China, 19–25 October 2025; pp. 14798–14804.
37. Fossen, T.I. *Handbook of Marine Craft Hydrodynamics and Motion Control*; John Wiley & Sons Ltd.: Chichester, UK, 2011.
38. Er, M.J.; Gong, H.; Liu, Y.; Liu, T. Intelligent trajectory tracking and formation control of underactuated autonomous underwater vehicles: A critical review. *IEEE Trans. Syst. Man Cybern. Syst.* **2024**, *54*, 543–555. [CrossRef]
39. Sieberling, S.; Chu, Q.P.; Mulder, J.A. Robust flight control using incremental nonlinear dynamic inversion and angular acceleration prediction. *J. Guid. Control Dyn.* **2010**, *33*, 1732–1742. [CrossRef]
40. Wang, X.; Van Kampen, E.-J.; Chu, Q.; Lu, P. Stability analysis for incremental nonlinear dynamic inversion control. *J. Guid. Control Dyn.* **2019**, *42*, 1116–1129. [CrossRef]
41. Sun, S.; Romero, A.; Foehn, P.; Kaufmann, E.; Scaramuzza, D. A Comparative study of nonlinear MPC and differential-flatness-based control for quadrotor agile flight. *IEEE Trans. Robot.* **2022**, *38*, 3357–3373. [CrossRef]
42. Slawik, T.; Vyas, S.; Christensen, L.; Kirchner, F. Attitude control of the hydrobatic intervention AUV cuttlefish using incremental nonlinear dynamic inversion. In Proceedings of the 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Abu Dhabi, United Arab Emirates, 14–18 October 2024; pp. 781–786.
43. Verschueren, R.; Frison, G.; Kouzoupis, D.; Frey, J.; van Duijkeren, N.; Zanelli, A.; Novoselnik, B.; Albin, T.; Quirynen, R.; Diehl, M. Acados—A modular open-source framework for fast embedded optimal control. *Math. Program. Comput.* **2022**, *14*, 147–183. [CrossRef]

44. Andersson, J.A.E.; Gillis, J.; Horn, G.; Rawlings, J.B.; Diehl, M. CasADi: A software framework for nonlinear optimization and optimal control. *Math. Program. Comput.* **2019**, *11*, 1–36. [CrossRef]
45. Song, Y.; Qian, J.; Miao, R.; Xue, W.; Ying, R.; Liu, P. HAUD: A high-accuracy underwater dataset for visual–inertial odometry. In Proceedings of the 2021 IEEE Sensors, Sydney, Australia, 31 October–3 November 2021; pp. 1–4.
46. Ferrera, M.; Creuze, V.; Moras, J.; Trouvé-Peloux, P. AQUALOC: An underwater dataset for visual–inertial–pressure localization. *Int. J. Robot. Res.* **2019**, *38*, 1549–1559. [CrossRef]
47. Mur-Artal, R.; Tardós, J.D. ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras. *IEEE Trans. Robot.* **2017**, *33*, 1255–1262. [CrossRef]
48. Campos, C.; Elvira, R.; Rodríguez, J.J.G.; Montiel, J.M.M.; Tardós, J.D. ORB-SLAM3: An accurate open-source library for visual, visual–inertial, and multimap SLAM. *IEEE Trans. Robot.* **2021**, *37*, 1874–1890. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.