

Article

Localization Method for Underwater Robot Swarms Based on Enhanced Visual Markers

Qingbo Wei ^{1,2,3} , Yi Yang ^{1,2,*}, Xingqun Zhou ^{1,2}, Chuanzhi Fan ^{1,2}, Quan Zheng ^{1,2} and Zhiqiang Hu ^{1,2}

¹ State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016, China; weiqingbo22@mails.ucas.ac.cn (Q.W.); zhouxingqun@sia.cn (X.Z.); fanchuanzhi@sia.cn (C.F.); zhengquan@sia.cn (Q.Z.); hzq@sia.cn (Z.H.)

² Institutes for Robotics and Intelligent Manufacturing, Chinese Academy of Sciences, Shenyang 110169, China

³ University of Chinese Academy of Sciences, Beijing 100049, China

* Correspondence: yangyi@sia.cn

Abstract: In challenging tasks such as large-scale resource detection, deep-sea exploration, prolonged cruising, extensive topographical mapping, and operations within intricate current regions, AUV swarm technologies play a pivotal role. A core technical challenge within this realm is the precise determination of relative positions among AUVs within the cluster. Given the complexity of underwater environments, this study introduces an integrated and high-precision underwater cluster positioning method, incorporating advanced image restoration algorithms and enhanced underwater visual markers. Utilizing the Hydro-Optical Image Restoration Model (HOIRM) developed in this research, image clarity in underwater settings is significantly improved, thereby expanding the attenuation coefficient range for marker identification and enhancing it by at least 20%. Compared to other markers, the novel underwater visual marker designed in this research elevates positioning accuracy by 1.5 times under optimal water conditions and twice as much under adverse conditions. By synthesizing the aforementioned techniques, this study has successfully developed a comprehensive underwater visual positioning algorithm, amalgamating image restoration, feature detection, geometric code value analysis, and pose resolution. The efficacy of the method has been validated through real-world underwater swarm experiments, providing crucial navigational and operational assurance for AUV clusters.



Citation: Wei, Q.; Yang, Y.; Zhou, X.; Fan, C.; Zheng, Q.; Hu, Z.

Localization Method for Underwater Robot Swarms Based on Enhanced Visual Markers. *Electronics* **2023**, *12*, 4882. <https://doi.org/10.3390/electronics12234882>

Academic Editor: Manohar Das

Received: 1 November 2023

Revised: 24 November 2023

Accepted: 29 November 2023

Published: 4 December 2023



Copyright: © 2023 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 (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: autonomous underwater vehicle; underwater swarm localization; visual localization; image processing; visual markers

1. Introduction

Autonomous Underwater Vehicles (AUVs) have become increasingly vital in military operations, marine resource exploration, and other advanced underwater tasks. Bolstered by advancements in artificial intelligence, their potential applications are perceived to surpass those of Remotely Operated Vehicles (ROVs). However, when tasked with large-scale missions such as extensive resource detection, deep-sea exploration, prolonged patrols, comprehensive topographical mapping, and operations in complex current regions, the capability of individual AUVs is challenged. Consequently, AUV swarm technologies have been thrust into the research spotlight, as depicted in Figure 1, which showcases the collaborative efforts of the TS Mini-AUV from the Shenyang Institute of Automation at the Chinese Academy of Sciences.

In the realm of Autonomous Underwater Vehicle (AUV) swarms, precise inter-vehicular positioning is crucial for effective collective perception and control, representing a fundamental aspect of AUV swarm technology. Established positioning technologies in terrestrial environments, including LiDAR-based Simultaneous Localization and Mapping (SLAM), Real-Time Kinematic GPS (RTK-GPS), and WiFi triangulation, have undergone significant development. Chen et al.'s [1] innovative approach to unstructured scene planning and

control for all-electric vehicles, Meng et al.'s [2] HYDRO-3D object detection and tracking system utilizing 3D LiDAR, and Liu et al.'s [3] work on accurately estimating vehicle sideslip angles exemplify advancements in land vehicle control and detection.



Figure 1. (left) TS Mini-AUV from Shenyang Institute of Automation at the Chinese Academy of Sciences. (right) AUV Swarm from the same institution.

However, the adaptation of these technologies for underwater use faces numerous challenges [4]. The obstruction of GPS signals by water makes GPS technology unsuitable for underwater applications. The substantial scattering and absorption characteristics of water severely diminish LiDAR systems' efficacy. Moreover, the rapid attenuation of wireless signals in aquatic environments renders WiFi-based triangulation methods ineffective for underwater positioning. These limitations underscore the impracticality of directly applying terrestrial positioning technologies to underwater settings. Furthermore, internal sensors such as Inertial Measurement Units (IMUs) are constrained by long-term cumulative errors, compromising their ability to provide stable and reliable navigation data for cluster positioning. While Ullah et al. [5] have demonstrated success in underwater target detection and positioning using acoustic signals, this method still faces challenges in terms of positioning accuracy and the feasibility of deploying equipment in densely packed underwater clusters.

In an underwater environment, the efficiency of visual positioning technology is underscored by its exploitation of light signals' effective propagation for precise, omnidirectional positioning, demonstrating marked efficacy over short ranges. This technology is particularly suited for Autonomous Underwater Vehicles (AUVs), constrained by spatial and structural limitations due to the monocular vision methodology's streamlined structure, compactness, and rapid processing capabilities. Utilizing a monocular camera, this approach captures image data and employs known feature points to align two-dimensional with three-dimensional data, thereby achieving accurate positioning. The criticality of visual markers in this process is evident, as their distinctive geometric structures provide precise coordinates for these feature points within images. By integrating the spatial coordinates of these markers with their corresponding image coordinates, efficient matching of images to three-dimensional models is facilitated. The application of the Perspective-n-Point (PNP) algorithm is instrumental in computing the target's six-axis pose data, culminating in enhanced visual positioning accuracy.

Consequently, this study employs a monocular camera, noted for its simple structure and compact size, to achieve accurate visual positioning in complex underwater environments. Innovations include the development of enhanced AR-coded markers, as illustrated in Figure 2. These markers, strategically placed on Autonomous Underwater Vehicles (AUVs), maintain their streamlined design while optimizing marker visibility. Each AUV is equipped with five high-resolution underwater monocular cameras, positioned to cap-

ture visual marker data from multiple perspectives, thus enabling effective positioning within the swarm's visual range. The focus of this research is on refining the precision and robustness of positioning through the improvement in visual markers and the clarity of underwater imagery. A significant advancement is the introduction of enhanced AR-coded markers, which increase the density of usable feature points within the visual markers, enhancing matching precision and adapting to unique underwater conditions to improve locational accuracy. Furthermore, this paper presents the Hydro-Optic Image Restoration Model (HOIRM), an innovative approach based on the physical model of underwater image degradation. This model applies an inverse degradation process to restore image clarity, markedly enhancing the accuracy and robustness of marker detection in high-turbidity conditions.

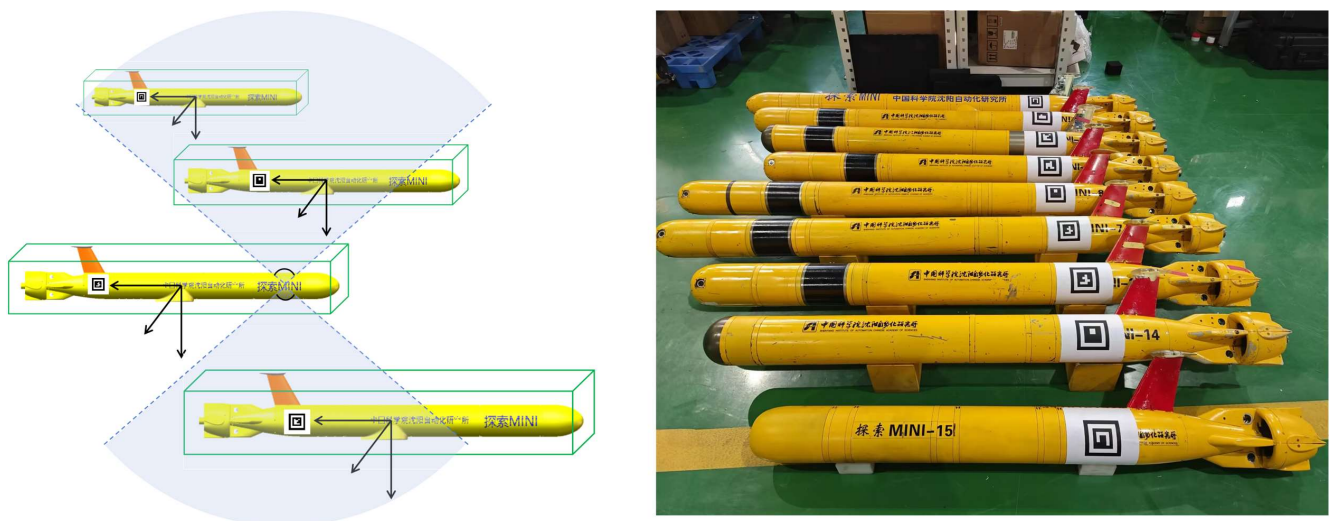


Figure 2. (left) schematic representation of cluster localization featuring the TS Mini-AUV. (right) physical depiction of the TS Mini-AUVs.

The primary contributions of this paper are summarized as follows:

The development and application of an enhanced AR-coded marker for underwater usage are presented. This novel marker demonstrates a notable 1.5-fold increase in visual positioning accuracy compared to existing markers.

A brand-new Hydro-Optic Image Restoration Model is introduced. This model significantly outperforms existing dehazing algorithms, broadening the discernible range of the light attenuation coefficient by 20%, thereby enhancing the quality of underwater imagery.

Our research extends to the creation of supplementary algorithms and the empirical analysis of cluster positioning techniques using enhanced AR-coded markers. These advancements prove to be highly effective in the real-time, stable detection and positioning of AUVs within a cluster, offering a dependable solution for proximal robot positioning in underwater clustering technologies.

The ensuing sections of this paper will methodically address several key areas: firstly, the context of underwater monocular vision positioning; secondly, the conceptualization and design of enhanced AR-coded markers; and thirdly, the development of the Hydro-Optic Image Restoration Model, followed by a detailed discussion of the positioning process leveraging these technologies. This paper will then progress to the experimental framework and an analytical evaluation of the results. Finally, it will conclude with a summary and perspectives for future research endeavors in this field.

2. Related Work

In the domain of underwater robotics, visual markers have emerged as a reliable positioning strategy, offering a novel approach for the positioning of Autonomous Un-

derwater Vehicles (AUVs). Zhang et al. [6] integrated optical beacons with traditional image processing techniques to estimate distance and depth. Feng et al. [7] designed a hybrid positioning strategy wherein long-range positioning utilized optical beacons and short-range positioning switched to AR markers. Meanwhile, Xu et al. [8] opted to deploy multiple ArUco markers on the seabed, thereby advancing underwater visual navigation. Wu et al. [9,10] designed a monocular visual positioning system for manned submersibles reaching depths of up to 7000 m based on cooperative markers. However, the aforementioned visual positioning techniques, employing optical, cooperative, and AR markers, exhibit notable constraints when applied to underwater cluster positioning, as detailed in the following discussion.

Specifically, optical markers employ the centroid of luminous sources in images as feature points for positioning. However, the precision of optical marker-based positioning remains relatively low, is constrained by ambient light, and, when mounted on AUVs, greatly affects the AUV's maneuverability, flexibility, and stealth. Cooperative markers [11–13] primarily rely on specific geometric shapes for recognition, but they are more suitable for individual AUVs rather than clusters. Concurrently, AR markers like ArUco and AprilTag [14–20] possess distinctive encoding schemes, enabling the differentiation of various targets. Yet, they provide a limited number of feature corners. Underwater, the large errors in image coordinate detection greatly affect the positioning results. Additionally, the feature points are prone to loss, leading to the inability to achieve positioning. To address this, previous studies opted to use multiple AR markers to enhance positioning precision [7,8,21], but this method poses challenges for integration on AUVs with rigorous structural constraints. To overcome these constraints, this paper introduces an enhanced AR-coded marker that is applied in underwater cluster positioning.

Moreover, Yang Yi and his team from the Institute of Automation at the Chinese Academy of Sciences [22] proposed an AUV visual positioning solution based on underwater vector laser patterns for dense formations. However, in well-lit environments, the visibility of the vector laser significantly diminishes, making light detection challenging. This method also faces limitations in lateral positioning within cluster carriers. Therefore, there's a need for cluster positioning methods that address these challenges. In recent years, visual positioning methods based on inherent target characteristics and leveraging deep neural networks [23–27] have offered new solutions for underwater cluster positioning. But due to constraints such as lower accuracy, slower computational speed, difficulties in dataset acquisition, and challenges in multi-target positioning, they are currently not ideal for underwater cluster positioning applications.

Ensuring a high success rate for visual marker detection in cluster visual positioning necessitates underwater image enhancement. Traditional enhancement techniques, such as the Dark Channel Prior (DCP) formulated by He et al. [28], histogram equalization [29,30], Retinex-based methods [31,32], and filter-guided techniques [33,34], have not shown ideal performance in actual underwater applications. Research on underwater image enhancement, like Carlevaris-Bianco's wavelength-dependent light attenuation [35], Wang's convolutional neural network color correction [36], and other recent studies [37–41], has shown some positive progress. However, their outcomes in real, complex underwater settings remain suboptimal. Consequently, a highly adaptive cluster positioning approach requires an image enhancement technique suitable for complex underwater environments.

3. Enhanced AR-Coded Marker Design

Visual markers enable the alignment of objects in images with their three-dimensional counterparts through feature points. Notably, passive markers, including cooperative and AR markers, adeptly deliver precise feature point information, yielding higher positioning accuracy. In visual measurement systems, the image coordinate-detection error, identified as the sole irreducible error source [42,43], becomes more evident in challenging underwater environments with complex light propagation. Advancements in sub-pixel-level feature point coordinate detection and optimization of passive visual marker structures, including

increased feature point density, are effective strategies. These enhancements are crucial for improved matching accuracy and precision in underwater positioning.

3.1. Analysis of the Impact of Feature Point Quantity on Positioning Accuracy

In the realm of monocular visual positioning, the primary objective is to address the Perspective-n-Point (PNP) problem. This involves the computation of the pose matrix, denoted as $[R \ T]$.

This process begins with obtaining feature point coordinates from the captured image. Subsequently, these coordinates are employed to transform Equation (1) into a linear system of equations, as illustrated in Equation (2). The image coordinates of these feature points, along with their corresponding world coordinates, are then inserted into the system of equations to compute the matrix. However, aquatic environments in the real world often introduce noise, adversely impacting the accuracy of image coordinate collection and subsequently increasing the detection error associated with these coordinates.

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K [R \ T] \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix} = K \begin{bmatrix} t_1 & t_2 & t_3 & t_4 \\ t_5 & t_6 & t_7 & t_8 \\ t_9 & t_{10} & t_{11} & t_{12} \end{bmatrix} \begin{bmatrix} X_W \\ Y_W \\ Z_W \\ 1 \end{bmatrix}, \quad (1)$$

$$\begin{bmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & -u_1 X_1 & -u_1 Y_1 & -u_1 Z_1 & -u_1 \\ 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -v_1 X_1 & -v_1 Y_1 & -v_1 Z_1 & -v_1 \\ & & & & & & \dots & & & & \\ X_N & Y_N & Z_N & 1 & 0 & 0 & 0 & -u_N X_N & -u_N Y_N & -u_N Z_N & -u_N \\ 0 & 0 & 0 & X_N & Y_N & Z_N & 1 & -v_N X_N & -v_N Y_N & -v_N Z_N & -v_N \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \\ t_6 \\ t_7 \\ t_8 \\ t_9 \\ t_{10} \\ t_{11} \\ t_{12} \end{bmatrix} = AT = 0, \quad (2)$$

The least-squares optimization method is capable of extracting optimal solutions for linear equations, even when inaccuracies in parameters are present. However, this method's iterative results' accuracy is reduced due to the inherent detection error in image coordinates. This reduction in precision becomes more pronounced in underwater environments with elevated noise levels. Under the framework of least-squares optimization, it has been observed that an increase in the number of feature points correlates with a reduction in solution error, thereby leading to a more precise computation of the pose matrix. In environments characterized by high noise, such as underwater settings, this increase in feature points proves particularly beneficial. It effectively mitigates noise interference, resulting in the stabilization and enhancement of the accuracy of positioning solutions.

To conduct a quantitative evaluation, a simulation model for camera imaging and pose computation was established, encompassing actual visual markers and intrinsic camera parameters. Simulations were conducted to evaluate how varying numbers of feature points affect measurement accuracy within the context of different levels of image coordinate detection errors. The intrinsic matrix and distortion matrix derived from actual underwater camera calibrations are defined in Equations (3) and (4), respectively:

$$\begin{bmatrix} 14914.22 & 0 & 1239.18 \\ 0 & 14912.74 & 1088.06 \\ 0 & 0 & 1 \end{bmatrix}, \quad (3)$$

$$[0.33049 \ 4.74831 \ 0 \ 0 \ 0], \quad (4)$$

To simulate varying levels of image coordinate-detection errors under different water quality conditions, Gaussian noise with different standard deviations was introduced into the feature point coordinates. Specifically, five experimental groups were created, each representing a different noise level. Within each group, the number of feature points systematically increased from 4 to 20. For each configuration, 1000 trials of random noise superposition were performed, yielding 1000 measurement errors. These errors were subsequently compiled to determine a cumulative measurement error. The relative reduction in distance error was adopted as the evaluation metric, defined in Equation (5):

$$e_{\text{rot}} = n \cdot \frac{|d_x - d_{x0}| + |d_y - d_{y0}| + |d_z - d_{z0}|}{3} - \text{error}_{20}, \quad (5)$$

where n is the number of measurements and error_{20} represents the cumulative error with 20 feature points at the same noise level.

Having compiled data from 1000 measurements across the five experimental groups, grouped bar charts were utilized to visually present the cumulative errors (Figure 3).

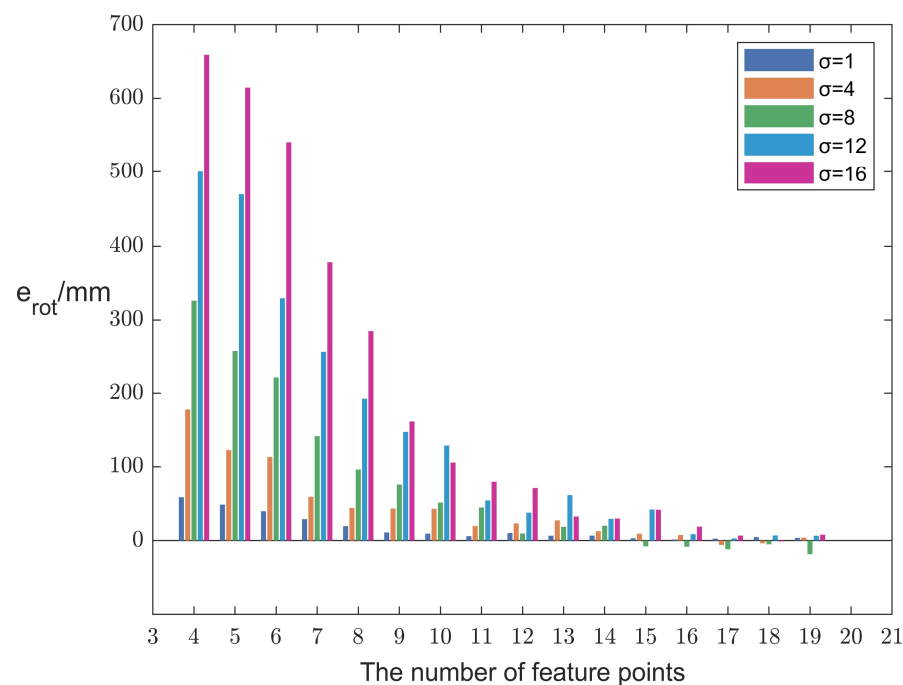


Figure 3. Grouped bar chart of cumulative error comparison.

In Figure 3, bar charts of five distinct colors depict cumulative errors at different noise levels. With a constant number of feature points, cumulative errors exhibit a continual rise as the standard deviation of the Gaussian noise augments. When the noise level remains consistent, the bar charts illustrate a decline in cumulative errors with an increasing number of feature points. The decline is more pronounced at higher noise levels. Specifically, with a noise standard deviation of 16, increasing the feature points from 4 to 12 results in a significant reduction of 588.32 mm in cumulative error. However, beyond 11 feature points, the decline in error becomes less discernible.

By integrating simulation study findings with theoretical insights, it was concluded that increasing the feature points can significantly augment measurement precision, especially in high-noise settings. Nonetheless, it is imperative to recognize that, once the number of feature points surpasses a specific threshold, the reduction in error plateaus. Consequently, optimizing the effective number of feature points per unit area emerges as a crucial factor in achieving precise positioning in high-noise aquatic environments.

3.2. Marker Structure Design

Grounded in foundational theories and robust simulation evidence, it is ascertained that an increase in the number of feature points markedly mitigates the diminution in positioning accuracy arising from image coordinate-detection errors. Simultaneously, visual positioning entails the identification of diverse Autonomous Underwater Vehicles (AUVs) within a swarm. This requires that the visual markers incorporate encoded features, demanding a prudent equilibrium between the expansion of feature points and the optimization of spatial efficiency. Reflecting these prerequisites, the enhanced AR-coded markers conceptualized in this research are illustrated in Figure 4.

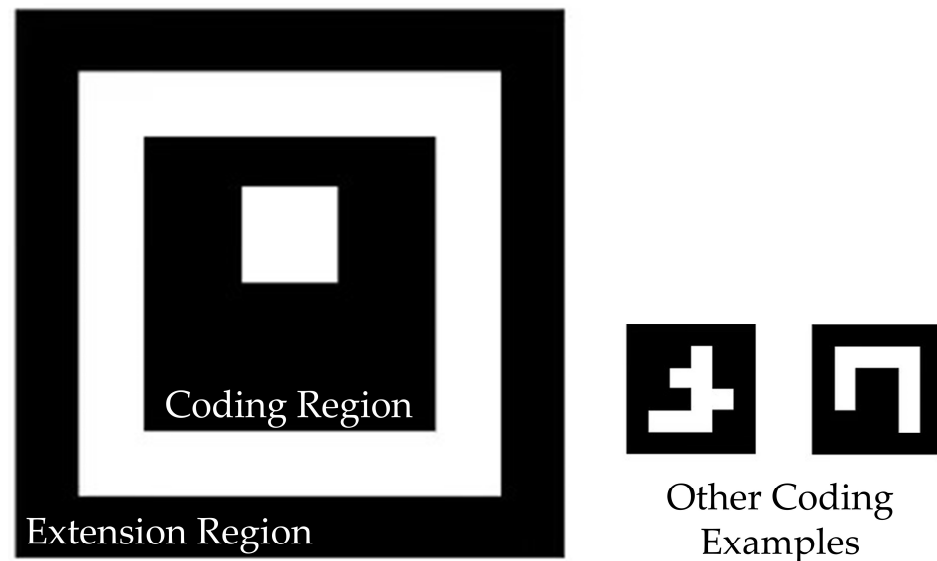


Figure 4. Enhanced AR-coded markers, showcasing a complete example on the left with its internal coding region representing Marker ID 17. To the right are two additional examples of coding regions, with the leftmost indicating Marker ID 15 and the rightmost Marker ID 10.

The enhanced AR-coded marker comprises an internal coding region and an external extension region. The internal coding region is structured as a 6×6 grid square QR code, encompassed by a black square boundary and incorporating black and white modules. At its core lies a 4×4 binary matrix, employed for information encoding and unique code identification, with coding area instances for IDs 10, 15, and 17 illustrated. In detecting the internal coding region, the square boundary is initially discerned within the image, ensuing the interpretation of the internal black and white square layout into binary digits. These digits are then decoded to their respective identifier IDs. Surrounding the central coding area, the external extension region, a black frame, bolsters the marker's contrast and fortification in recognition, concurrently offering additional boundary insights. In the process of marker detection, both the internal and external edges of the extension region, configured as quadrilaterals, are capable of being delineated, contributing an additional eight feature corners. This expands the total to 12 feature corners, a threefold increase from the original 4 provided by the coding region. Drawing upon the foundational theories and corroborating simulation evidence detailed previously, the enhanced AR-coded marker is evidenced to enhance both the stability and accuracy of positioning. The comparative analysis within the experimental section of this study illustrates that, relative to existing visual markers, this enhanced marker yields greater precision in pose data and exhibits lower variability.

4. Hydro-Optic Image Restoration Model

In intricate underwater environments, image degradation is the primary cause of challenges in detecting visual markers and significant errors in image coordinate detection.

Such degradation can eventually lead to large visual positioning errors or even an inability to position at all. The complexity of light propagation underwater presents immense challenges to traditional image restoration techniques. Addressing this issue, this paper introduces a novel algorithm based on the underwater imaging model, harnessing the theory of underwater light propagation to enhance the quality of underwater images.

4.1. Model Building

Historically, the Single-Scattering Atmospheric Model (SSAM) has been employed to describe underwater light propagation. This model has also been widely adopted in previous underwater dehazing studies. The SSAM can be represented as:

$$I(x) = J(x)t(x) + A(1 - t(x)), \quad (6)$$

where x is a pixel; $J(x)$ is the scene radiance in the absence of fog; A signifies the global atmospheric light; and $t(x) = e^{-\beta d(x)}$ is the transmission which represents color or light attenuation due to the scattering medium. The attenuation is dictated by both the scene depth $d(x)$ and the attenuation coefficient β .

However, the SSAM does not adequately account for the significant impact of forward scattering on image blurring, rendering it less effective in mitigating the decline in underwater image quality. To offer a more robust image restoration framework, our study constructs the Aquatic Light Image Restoration Model. This comprehensive model analyzes underwater light scattering, absorption, and ambient light interference using the physical model of underwater imaging. It can be articulated as:

$$I(x) = J(x) * h_{psf}(x) \cdot t(x) + B_L(x) + n(x), \quad (7)$$

$$I(x) = J(x) * h_{psf}(x) \cdot t(x) + B_L(x), \quad (8)$$

where $J(x)$ represents the fog-free image, i.e., the desired recovered image; $h_{psf}(x)$ is the Point Spread Function (PSF); $B_L(x)$ is the ambient light function; and $n(x)$ constitutes various types of noise.

For the purposes of this study, our model (Equation (8)) simplifies ambient light and noise into a single term, $B_L(x)$, encapsulating both elements. This composite term is referred to as ‘ambient light’ for the sake of simplicity.

Given the above articulation, two sub-problems are identified in the image restoration process, addressing different facets of underwater image degradation:

$$I_1(x) = B_L(x), \quad (9)$$

Sub-problem (9) is dedicated to estimating and mitigating the influence of ambient light pre-existing in underwater environments, thereby preliminarily improving image clarity.

$$I_2(x) = J(x) * h_{psf}(x) \cdot t(x), \quad (10)$$

Sub-problem (10) seeks to tackle the reduction in contrast and clarity of underwater images, a challenge predominantly caused by light attenuation and forward scattering.

4.2. HOIRM-Based Image Recovery

The recovery of underwater images can be elucidated via the aforementioned models. This computational process is graphically illustrated in Figure 5 using a representative case where the light attenuation coefficient is $c = 6.12/\text{m}$.

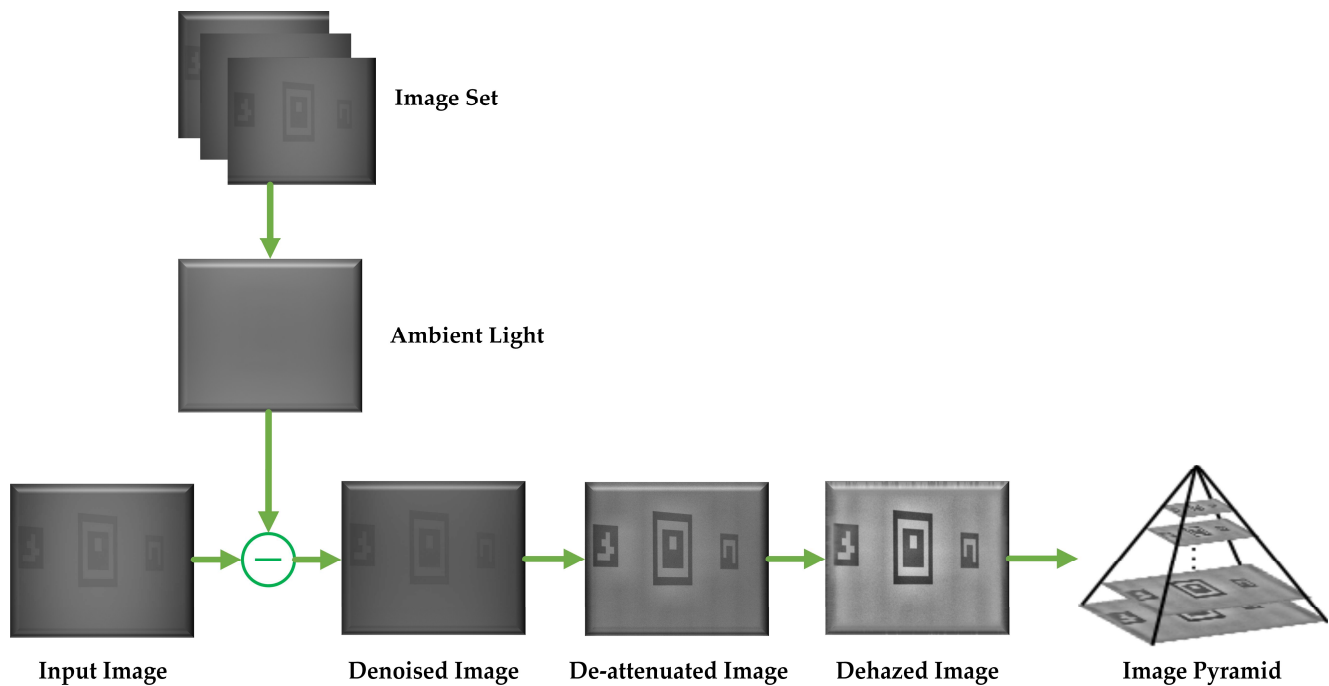


Figure 5. Schematic representation of the underwater image recovery algorithm.

In addressing Sub-problem (9), the estimation and subsequent removal of ambient light are imperative. Ambient light, as defined herein, acts as an environmental parameter that imparts a consistent blurring effect over time. A sequence of images, as depicted in Figure 5, is acquired via continuous sampling and subjected to collective analysis.

Each image is processed through a Gaussian blur, mathematically represented as:

$$Set_n(x) = (Img_n * G(x, \sigma))(x), \quad (11)$$

where Img_n is the n th captured image, $G(x, \sigma)$ is the Gaussian function, σ denotes the standard deviation influencing the extent of blurring, and $*$ denotes the convolution operation.

A synthesized ambient light image $B_L(x)$ is computed by applying weights w_n to each Gaussian-blurred image $Set_n(x)$, with the weights being contingent upon the relative temporal intervals t_n of the images. This process is encapsulated by the equation:

$$B_L(x) = \sum_n w_n \bullet Set_n(x), \quad (12)$$

The weights are subject to the constraint $\sum w_n = \alpha$, where $\alpha < 1$, ensuring the overall intensity remains subdued and can be adjusted in response to environmental variations.

The ambient light image $B_L(x)$, once derived, is subtracted from its corresponding blurred image, yielding the denoised image $I_2(x)$, as follows:

$$I_2(x) = I(x) - B_L(x), \quad (13)$$

This procedure efficiently counteracts the blurring induced by ambient light, consequently augmenting the image's clarity and proficiently resolving Sub-problem (9). With the mitigation of ambient light blur, the image's overall clarity is enhanced, establishing a solid foundation for the ensuing underwater image restoration process.

In the resolution of Sub-problem (10), the focus is directed towards issues related to light attenuation and forward scattering. The formula $t(x) = e^{-\beta d(x)}$, as delineated in Equation (8b), serves to quantify the attenuation of light throughout its underwater transmission, a phenomenon that considerably reduces image contrast. This study incorporates the Contrast Limited Adaptive Histogram Equalization (CLAHE) technique [29], acclaimed for its efficacy in amplifying contrast levels across diverse luminosity ranges in images. The

application of CLAHE facilitates adaptive contrast enhancement in underwater images, significantly alleviating the impacts of light attenuation.

In underwater environments, the forward scattering of light is a phenomenon wherein light rays deviate subtly from their initial trajectories due to particulate interference in the water. This deviation transforms the propagation of light from a singular, linear path to a more intricate pattern of scatter, consequently leading to image blur. The Point Spread Function (PSF), denoted as $h_{\text{psf}}(x)$, encapsulates the physical process of light undergoing forward scatter in aquatic settings. Modeled on a generalized Gaussian distribution [44], this function is defined in the following manner:

$$h_{\text{psf}}(u, v, p, \sigma) = \frac{e^{-\left(\frac{u^2+v^2}{A(p, \sigma)}\right)^{0.5p}}}{2\Gamma\left(1 + \frac{1}{p}\right)A(p, \sigma)}, (u, v) \in \mathbb{R}^2, \quad (14)$$

where p is related to water quality, $\sigma = \frac{1}{2\sqrt{p}}$, and $A(p, \sigma) = \sigma \left(\frac{\Gamma(1/p)}{\Gamma(3/p)} \right)^{0.5}$. The mathematical representation of this function and its associated point spread distribution are depicted in Figure 6.

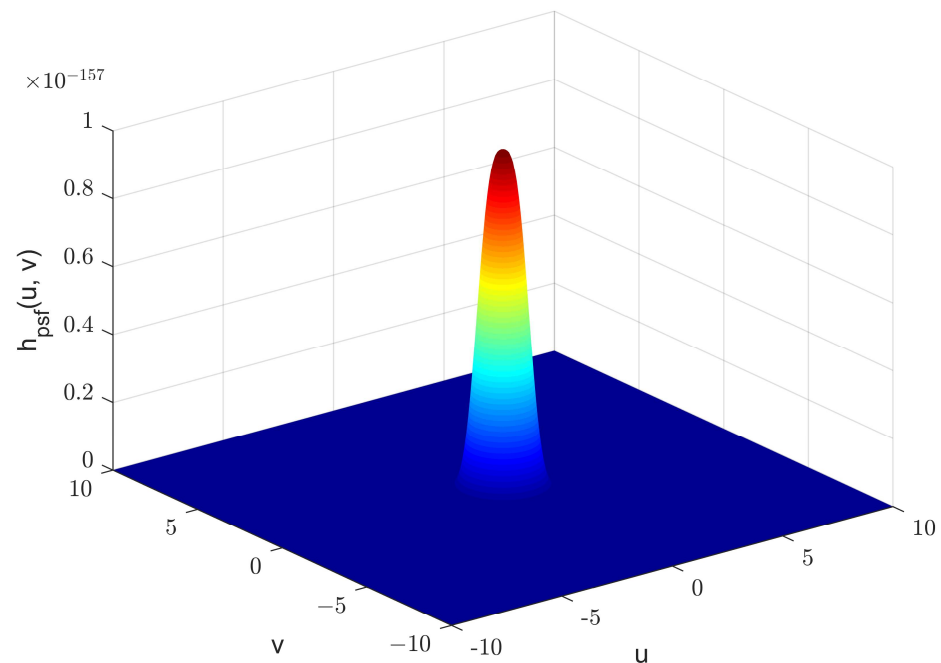


Figure 6. Distribution of Point Spread Function (PSF) models.

The imaging process, wherein each light source point undergoes forward scattering before being captured by the camera, can be characterized through the convolution of an ideal image, $J(x)$, with the Point Spread Function (PSF). This convolution paradigm implies that the deconvolution of a blurred image facilitates the retrieval of the ideal image, thereby enabling effective deblurring. Consequently, integrating this approach with the removal of background illumination and attenuation effects culminates in the restoration of image clarity.

Figure 7 demonstrates the efficacy of our proposed algorithm across four scenarios with varying turbidity levels. The visual enhancement is unmistakable—the augmented contrast and uniform luminance render previously indistinct visual markers distinctly visible, even under high light attenuation coefficients such as 12.20/m.

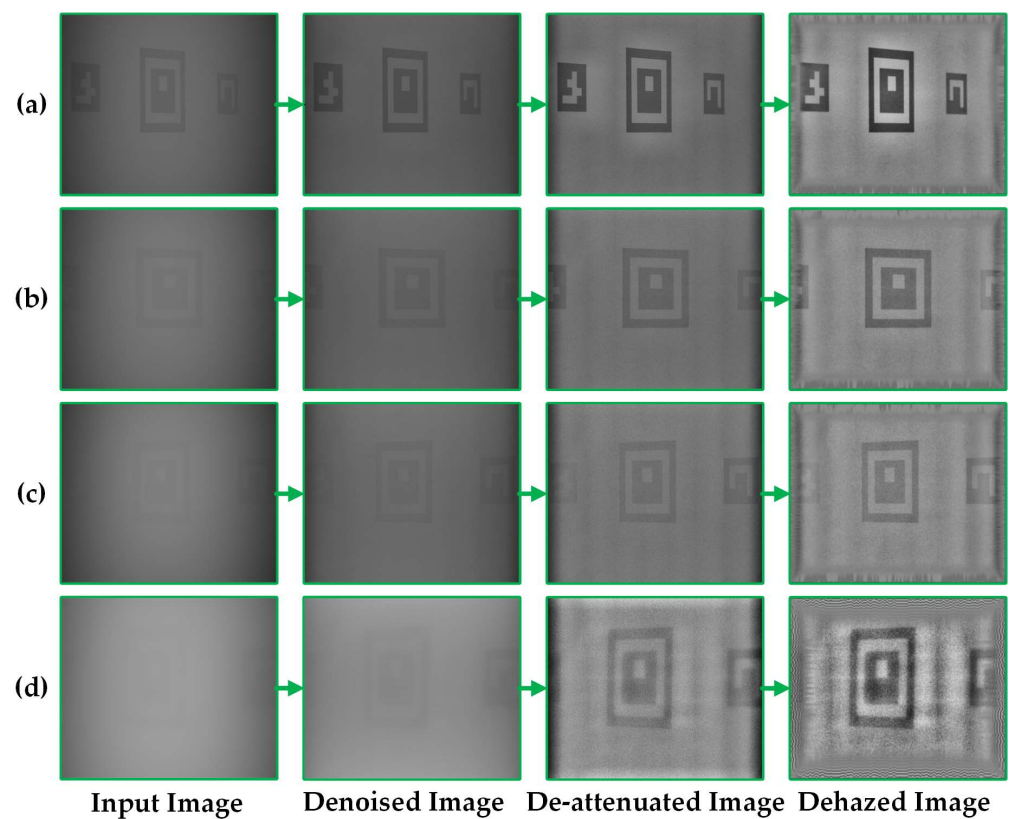


Figure 7. Image restoration outcomes for underwater images at varying light attenuation coefficients. The figure demonstrates the processing sequences of images taken in environments with distinct light attenuation coefficients, specifically 6.12/m (row (a)), 8.22/m (row (b)), 10.13/m (row (c)), and 12.20/m (row (d)). Each column within the figure represents images from different stages of the image restoration process.

Following the enhancement of image clarity, inherently blurred images tend to undergo post-processing distortion. The concluding phase of our image restoration model involves the application of down-sampling techniques to construct an image pyramid. Within this structure, layers better suited for edge detection are identified and processed. This method significantly improves the efficacy of visual marker detection.

5. Underwater Cluster Visual Localization Algorithms

Based on the improvements made to visual markers and the construction of an image restoration model suitable for underwater environments, as discussed in the previous sections, this paper introduces a comprehensive underwater visual positioning algorithm. This algorithm integrates image restoration, feature detection, geometric encoding value analysis, and pose estimation, providing reliable pose data for AUV clusters. Figure 8 provides a detailed structure of the underwater visual positioning algorithm presented in this study.

Initially, the contour features of the restored images are detected. The extracted contours are then evaluated and filtered based on geometric attributes such as shape, size, and edge proximity. The preliminary filtering results can be seen in column “b” in Figure 8.

As illustrated in column “c” in Figure 8, upon isolating the contours that meet the preset criteria, the contours undergo regional segmentation, divided into a 10×10 grid. Within this grid, geometric encoding value detection is performed. Each cell of the grid is coded as “0” or “1” based on the average grayscale value: “1” represents areas with higher grayscale values, while “0” signifies areas with lower grayscale values. These encoded values are then cross-referenced with a set of standardized encoding values to determine

the coding ID and validate each contour's validity. The final contours after filtration are showcased in column "d" in Figure 8.

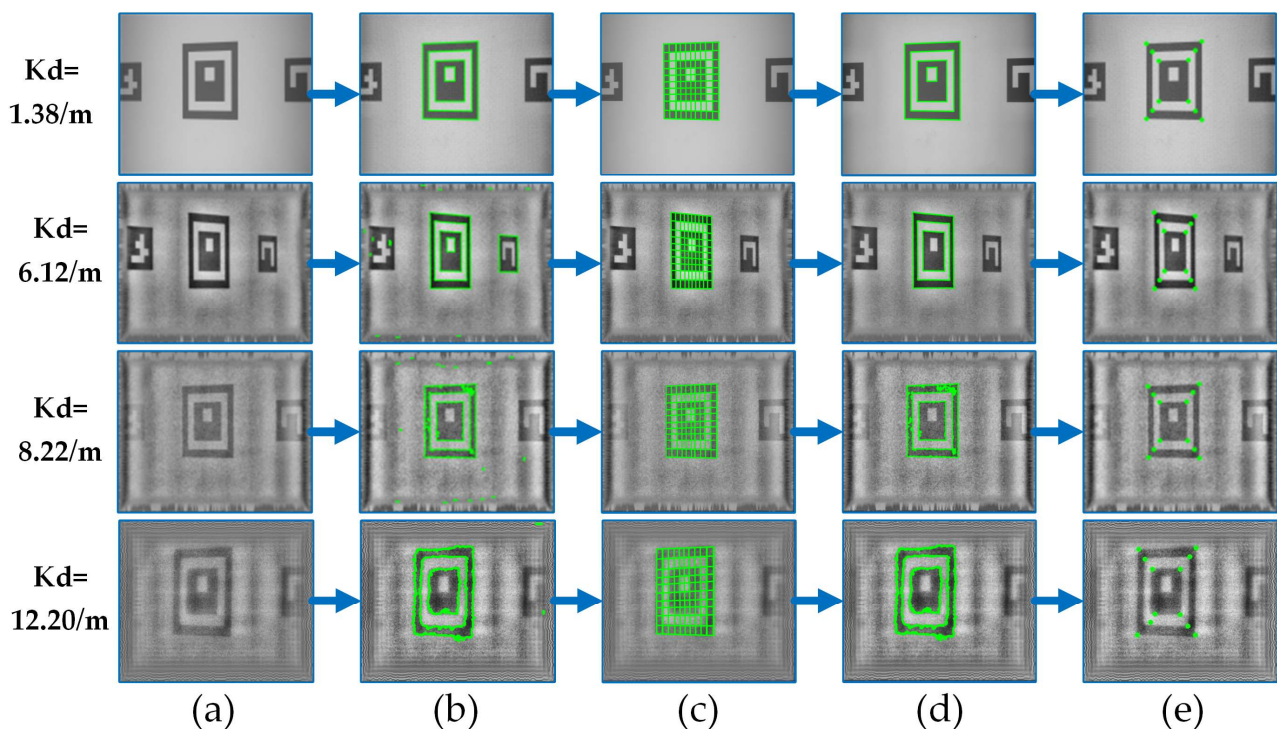


Figure 8. Schematic representation of the underwater visual localization algorithm. Column 'a' displays recovered images across varying water qualities; column 'b' delineates edges after initial screening; column 'c' portrays the geometric coding value detection process; column 'd' illustrates screened contours; and column 'e' exhibits the finalized corner detection results. Each row signifies the process of image localization under varying light attenuation coefficients, denoted as " K_d ".

Utilizing these identified contours, the initial image coordinates of 12 feature corners are determined. Subsequently, a meticulous analysis of the grayscale gradient distribution and grayscale weighted response vector surrounding each corner is undertaken. This allows for the iterative refinement of these coordinates at a subpixel level. This refinement process yields highly precise subpixel feature point coordinates, with the final feature points depicted in column "e" in Figure 8.

In the final stage, the coordinates corresponding to the 12 markers are inputted into the iterative Perspective-n-Point (PNP) algorithm. This algorithm triangulates the spatial positioning of these markers with high precision, enabling accurate underwater visual positioning.

The aforementioned methodology ensures a balance between computational efficiency and positioning precision, offering a comprehensive solution for the implementation of the techniques proposed in this paper.

6. Underwater Cluster Visual Positioning Experiment

For the underwater robot swarm positioning method based on enhanced visual markers presented in this study, a rigorous assessment of the proposed positioning algorithm's underwater performance was deemed imperative. To ascertain its competence in real-world engineering applications, an underwater pose-measurement platform was constructed. This allowed for a quantitative evaluation of underwater image restoration capability and detection positioning accuracy through a series of comprehensive experiments. Subsequent to this, real-water experiments on swarm positioning were conducted. The underwater

pose-measurement platform was specifically designed to overcome challenges encountered in actual water bodies where obtaining precise pose data is elusive and where the control of the light attenuation coefficient is not quantifiable. This design ensures a thorough and reliable quantitative assessment. The subsequent real-water experiments bolster the reliability for actual swarm positioning applications.

Evaluation experiments were conducted on a dedicated high-precision underwater testing platform, as depicted in Figure 9. The forward-facing camera in the AUV's head compartment was utilized to capture image data, with visual markers affixed to the AUV's compartment to test genuine positioning outcomes. The camera was mounted on a high-precision translational platform controlled electronically underwater, boasting a translational error of merely 0.005 mm. The compartment bearing the visual markers was secured to an electronic rotation platform capable of high-precision rotation with a rotation error of just 0.01° . By introducing various turbid solutions, a range of underwater conditions were emulated, and the light attenuation coefficient was measured in real time using a dedicated instrument. The rotation and translational platforms moved according to predefined trajectories, while the camera continuously gathered image data. In total, 11 distinct water quality conditions were established, and image data were captured under each condition for subsequent analysis.

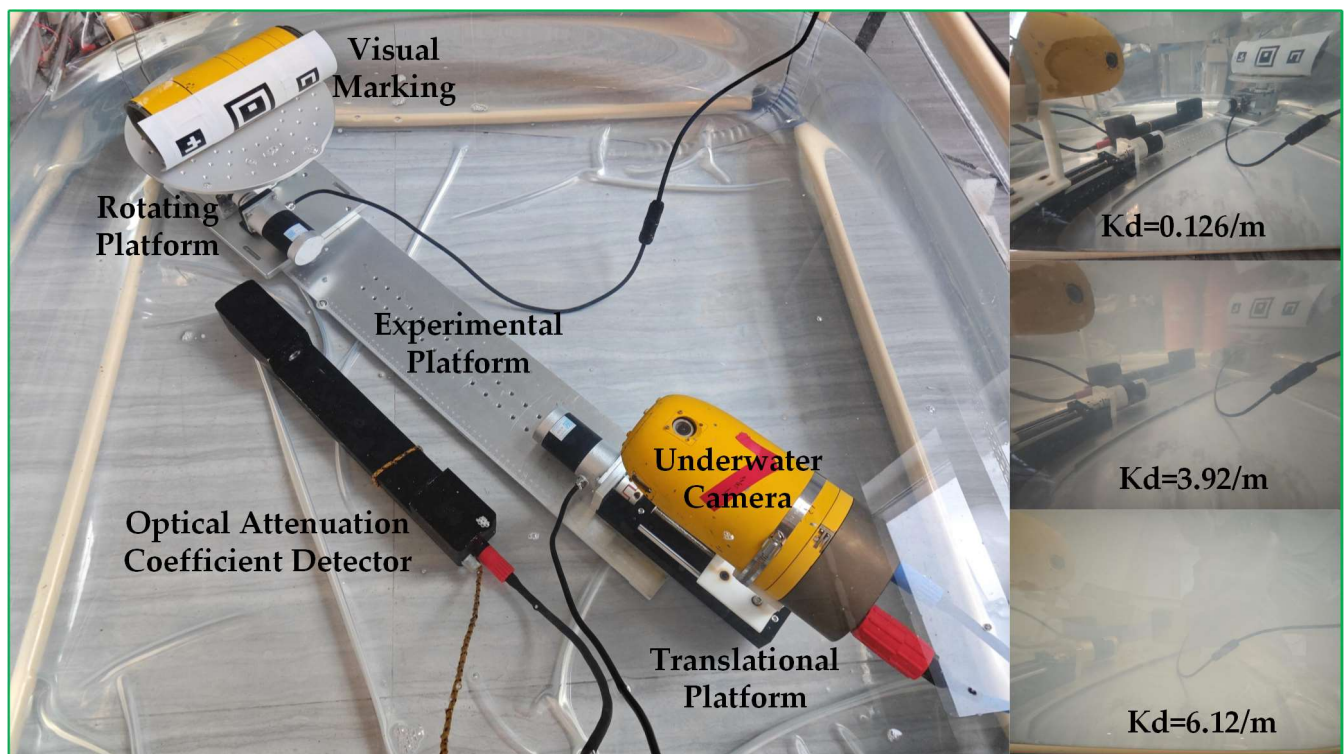


Figure 9. Underwater pose-measurement platform diagram and scenarios under various light attenuation coefficients.

6.1. Analysis of Image Restoration Efficacy

To validate the image restoration capability of the proposed HOIRM algorithm, underwater images taken under various water quality conditions were restored. A comparative analysis followed, contrasting this algorithm with existing image dehazing algorithms. As shown in Figure 10, four scenarios with higher light attenuation coefficients were selected for evaluation.

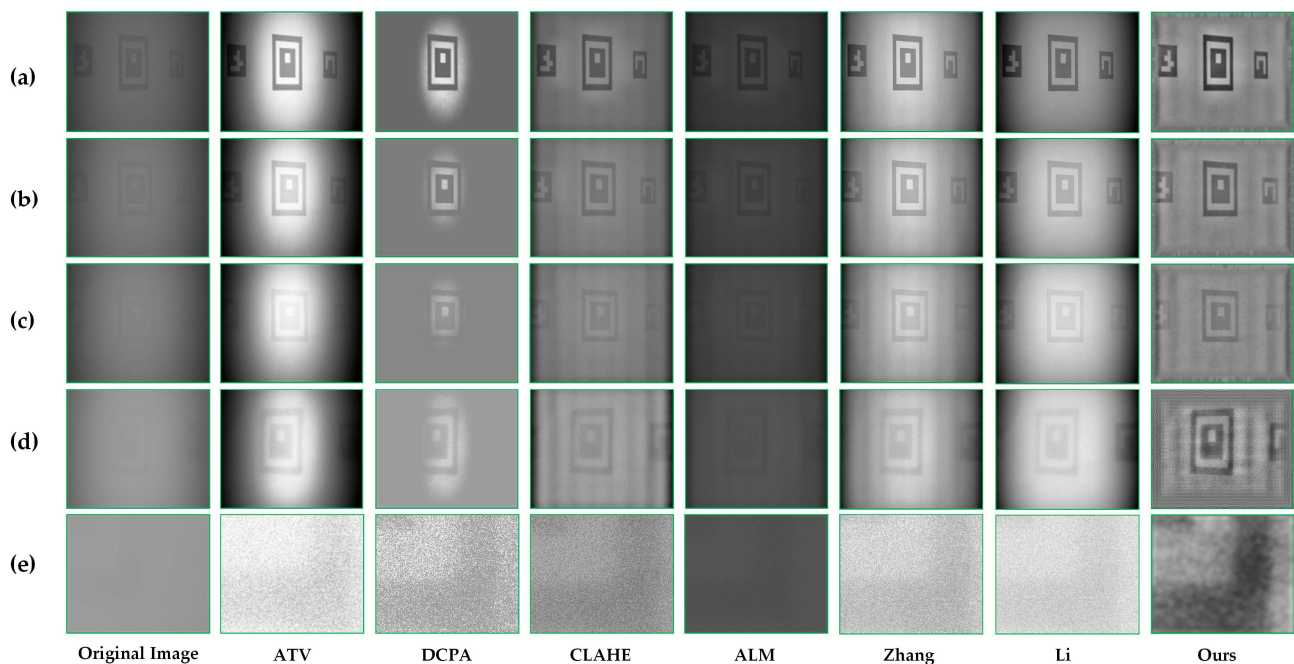


Figure 10. Comparative results of underwater image restoration using different algorithms. Rows (a–d) in the figure represent the processed images under light attenuation coefficients of 6.12/m, 8.22/m, 10.13/m, and 12.20/m, respectively. Row (e) specifically displays detailed views of the images processed by each algorithm, as shown in row (d). Each column represents the results of underwater image processing using the respective algorithms: Original Image, ATV (Author 2012) [34], DCPA (Author 2010) [28], CLAHE (Author 2018) [29], ALM (Author 2017) [38], Zhang (2022) [40], Li (2020) [41], and Ours.

As depicted, the algorithm developed in this study excelled in the qualitative enhancement of image restoration. Compared to other dehazing algorithms, it notably improved underwater image clarity, luminosity uniformity, and contrast. As shown in detail in Figure 10e, under relatively high light attenuation coefficients, the images exhibited extensive irregular salt-and-pepper noise. Additionally, inconsistent gradient distributions near the edge regions resulted in severe edge blurring. While the edge clarity in images restored by other algorithms remains inadequate, the method introduced in this study yields images with distinctively sharper edges, improved contrast between high- and low-grayscale regions, and notably enhanced edge structures, thereby enabling more accurate edge identification.

Subsequent to an initial qualitative evaluation of diverse algorithms, a comprehensive quantitative analysis was conducted. This assessment utilized three established image quality metrics: Structural Similarity Index (SSIM) [45], Peak Signal-to-Noise Ratio (PSNR) [46], and Contrast-to-Noise Ratio (CNR) [46]. These metrics collectively evaluate various dimensions of image quality, encompassing aspects such as color fidelity, reconstruction precision, and overall image clarity. Comparative evaluations were systematically executed across images processed by different algorithms, each subjected to four distinct optical attenuation coefficients. The collated data from this rigorous analysis are presented in Tables 1–4.

Table 1. Image metrics after processing by each algorithm when the optical attenuation coefficient is 12.20/m.

Baselines	SSIM	PSNR	CNR
ATV [34]	0.27564	9.8247	0.058818
DCPA [28]	0.74566	14.4325	−1
CLAHE [29]	0.30476	12.9064	2.6078
ALM [38]	0.5761	7.688	−0.33678
Zhang [40]	0.25082	13.8703	2.8194
Li [41]	0.23306	10.9746	1.7332
Ours	0.7467	14.9633	16.901

Red indicates the best metrics in each column, and blue indicates the second-best metrics in each column.

Table 2. Image metrics after processing by each algorithm when the optical attenuation coefficient is 10.13/m.

Baselines	SSIM	PSNR	CNR
ATV [34]	0.20564	6.4765	−0.44538
DCPA [28]	0.51625	6.3799	−1
CLAHE [29]	0.31598	6.1466	1.4678
ALM [38]	0.41861	3.3054	−0.26999
Zhang [40]	0.2351	6.8945	2.6694
Li [41]	0.20563	6.2143	4.0616
Ours	0.67948	7.9747	12.2967

Red indicates the best metrics in each column, and blue indicates the second-best metrics in each column.

Table 3. Image metrics after processing by each algorithm when the optical attenuation coefficient is 8.22/m.

Baselines	SSIM	PSNR	CNR
ATV [34]	0.48772	10.0439	−0.54203
DCPA [28]	0.70481	13.3453	−1
CLAHE [29]	0.35333	12.6422	2.6573
ALM [38]	0.57505	7.4692	−0.32692
Zhang [40]	0.224	13.526	2.6221
Li [41]	0.18944	12.8116	3.6758
Ours	0.74497	14.2165	15.6027

Red indicates the best metrics in each column, and blue indicates the second-best metrics in each column.

Table 4. Image metrics after processing by each algorithm when the optical attenuation coefficient is 6.12/m.

Baselines	SSIM	PSNR	CNR
ATV [34]	0.45012	10.0744	0.040307
DCPA [28]	0.6372	15.6326	−1
CLAHE [29]	0.18878	14.6112	3.318
ALM [38]	0.6703	8.7573	0.047221
Zhang [40]	0.23141	14.0412	3.6045
Li [41]	0.18166	13.3399	4.8769
Ours	0.81372	14.80951	13.1265

Red indicates the best metrics in each column, and blue indicates the second-best metrics in each column.

Upon examining the data presented in the table, it is noteworthy that the algorithm developed in this research demonstrates a PSNR value less than 5% lower than the CLAHE algorithm solely in the condition where the light attenuation coefficient is 6.12/m. In contrast, across various other water quality environments, our methodology consistently outperforms analogous techniques in all the evaluated metrics. Notably, the CNR value achieved by our approach is at least double that of related algorithms, evidencing a significantly enhanced contrast-to-noise ratio in the visual marker areas of the processed images

compared to the noise levels in adjacent regions, thus facilitating more effective feature identification. Additionally, both the SSIM and PSNR values attained by our method exceed those of other algorithms, indicative of superior image restoration capabilities, particularly in enhancing luminance, contrast, and structural details, aligning more closely with the ideal scenario. These findings compellingly validate the efficacy of our algorithm in underwater image restoration.

The validation of diverse image dehazing algorithms was furthered through comprehensive experimental evaluations, prioritizing image contour-detection success rate as the critical performance indicator. In varying water quality conditions, encompassing six distinct scenarios, a dataset of 100 images was captured at a uniform distance of 600 mm. This study then proceeded to evaluate the success rate of contour detection in images restored by various algorithms, considering different optical attenuation coefficients. Table 5 presents a summary of these results, highlighting a trend where increased optical attenuation correlates with reduced efficiency in contour detection. Notably, the images processed using the proposed algorithm consistently demonstrated superior performance compared to other methods, achieving the highest success rate in contour detection across an optical attenuation coefficient range of 0–12.20/m. Furthermore, in the more challenging attenuation range of 12.20–14.20, the algorithm maintained its effectiveness in contour detection, outperforming other algorithms even when they failed to detect contours.

Table 5. Comparison of the success rate of image contour detection after processing by different algorithms in different water qualities.

Baselines	Profile Detection Success Rate					
	AC: 5.16/m	AC: 6.12/m	AC: 8.22/m	AC: 10.13/m	AC: 12.20/m	AC: 14.20/m
ATV [29]	1	1	0.26	0	0	0
DCPA [23]	1	1	0.39	0	0	0
CLAHE [24]	1	1	0.87	0.61	0	0
ALM [33]	1	1	0.53	0	0	0
Zhang [40]	1	1	0.91	0.63	0.15	0
Li [41]	1	1	0.89	0.29	0	0
Ours	1	1	1	1	0.79	0.22

Red indicates the best metrics in each column, and blue indicates the second-best metrics in each column.

Conclusively, the extensive experiments conducted to assess image restoration capabilities unequivocally established the superior performance of the algorithm introduced in this study. Applied in practical settings, it significantly enhanced the success rate of contour detection, increasing the detectable range of light attenuation coefficients for contours by 20%. This advancement provides a robust assurance of image quality, crucial for swarm visual positioning applications.

6.2. Positioning Accuracy Test

The enhanced AR-encoded visual marker designed in this study is characterized by its high feature point density and high marker matching accuracy. Theoretically, this trait can improve positioning accuracy, especially in aquatic environments with substantial noise. To provide empirical evidence, we conducted comparative analyses of three different AR-encoded markers: the enhanced AR-encoded marker from this study, an ArUco marker, and an AprilTag marker. The evaluation criteria primarily focused on two pivotal metrics: angular positioning accuracy and distance positioning accuracy.

The experiments were executed under four specific water quality conditions. The visual markers and underwater cameras were fixed at a distance of 610 mm, with a yaw angle of 30 degrees. For each environmental condition, datasets comprising 100 image frames were captured for each visual marker. The pose was calculated based on image data

and compared to the actual pose to quantify measurement deviations. Figures 11–14 depict the angular and distance errors for all markers under each water quality scenario.

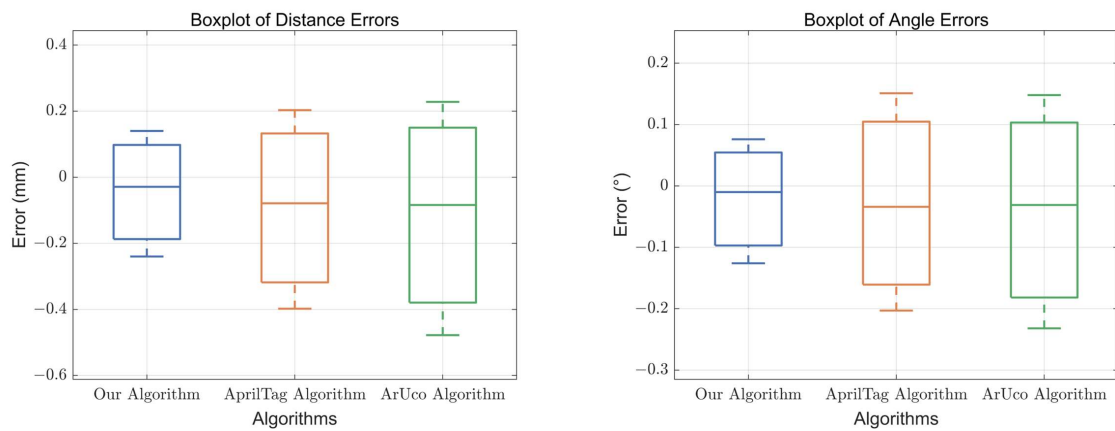


Figure 11. (Left) Distance error boxplot for an optical attenuation coefficient of 0.126/m. (Right) Angle error boxplot for an optical attenuation coefficient of 0.126/m.

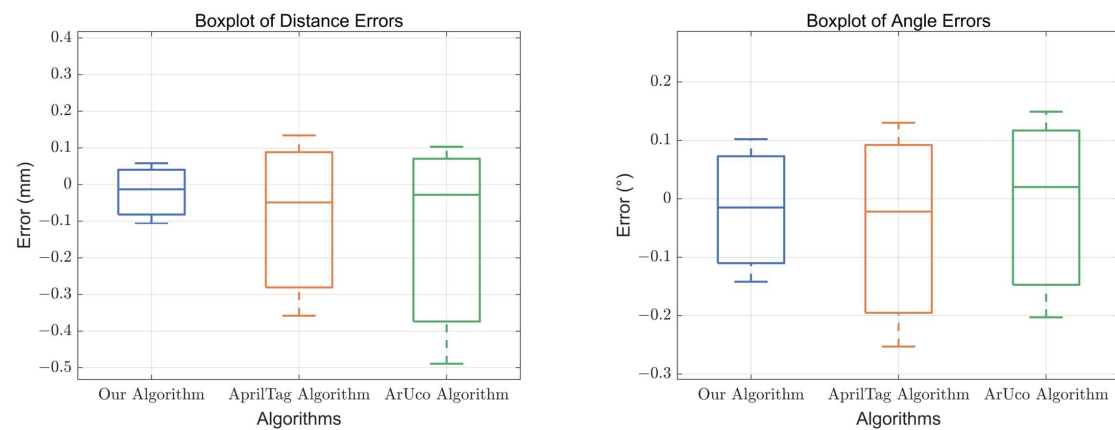


Figure 12. (Left) Distance error boxplot for an optical attenuation coefficient of 1.38/m. (Right) Angle error boxplot for an optical attenuation coefficient of 1.38/m.

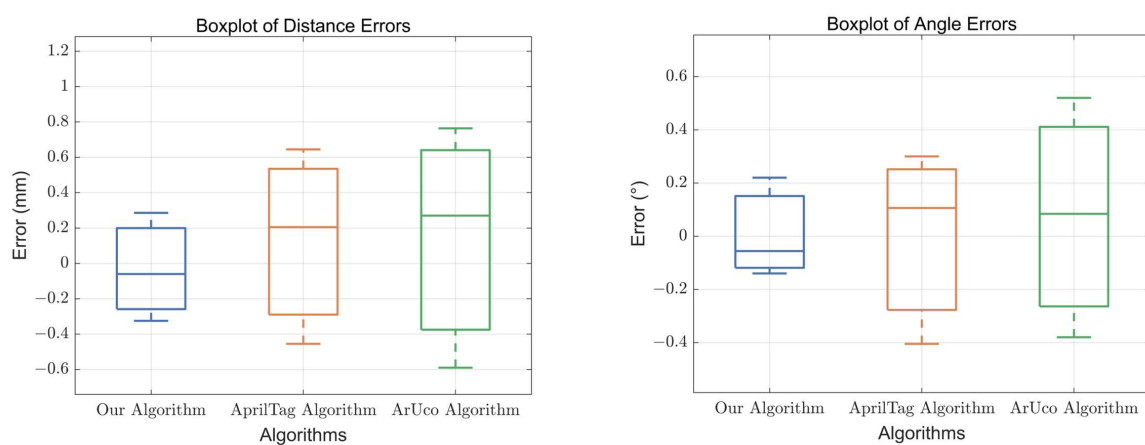


Figure 13. (Left) Distance error boxplot for an optical attenuation coefficient of 5.16/m. (Right) Angle error boxplot for an optical attenuation coefficient of 5.16/m.

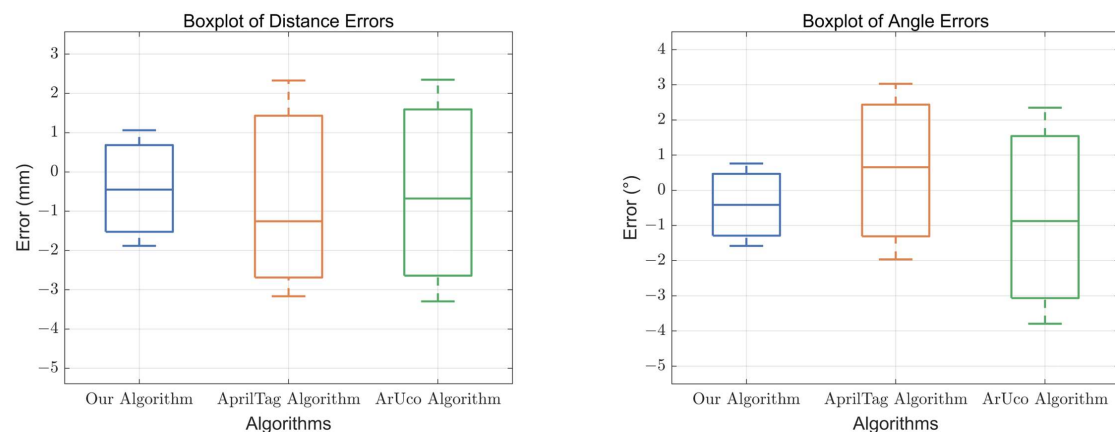


Figure 14. (Left) Distance error boxplot for an optical attenuation coefficient of 8.22/m. (Right) Angle error boxplot for an optical attenuation coefficient of 8.22/m.

As illustrated, in clearer water environments (light attenuation coefficients of 0.126/m and 1.38/m), the median positioning error for our enhanced AR-encoded marker was the lowest. Specifically, the distance error consistently remained below the threshold of 0.03 mm, and the angular error never exceeded 0.02 degrees. In comparison to the ArUco and AprilTag markers, the median positioning error was reduced by 35%. Furthermore, the interquartile ranges of distance error and angular error for our enhanced marker consistently remained below 0.3 mm and 0.2 degrees, respectively, which were over 1.5 times better than the other two markers.

Simultaneously, in murkier water environments with light attenuation coefficients of 5.16/m and 8.22/m, our marker showcased superior efficacy. The median positioning error was approximately 50% lower than the other two markers. Additionally, the interquartile range of positioning error for the other two markers was more than double that of our marker.

To holistically assess marker performance under constant water quality, separate experiments were conducted for each marker in waters with a fixed light attenuation coefficient of 3.92/m. The marker was positioned 600 mm away from the camera. The rotating platform moved in increments of 10 degrees within a ± 50 degree range, capturing 100 images at each angle. The difference between the analyzed visual positioning yaw angle and the actual angle was examined and statistically represented through boxplots, as shown in Figure 15. From the figure, it can be observed that, compared to the other two markers, the pose data derived from the enhanced AR-encoded marker exhibited higher overall positioning accuracy, lower data dispersion, and a more precise and consistent localization.

Lastly, for a comprehensive evaluation, all three markers were placed 600 mm away from the camera at an angle of 20 degrees under nine distinct water quality conditions, capturing 100 frames for each. Subsequently, the pose data derived from the images were compared with the actual data to compute the root-mean-square errors (RMSEs) for both distance and angle measurements. These are graphically presented in Figures 16 and 17. Compared to the other two markers, the enhanced AR-encoded marker consistently demonstrated a lower RMSE across all water conditions. With light attenuation coefficients ranging from 0 to 5.16/m, both distance and angular errors decreased by over 40%. Between 6.12/m and 12.2/m, these errors decreased by more than 50%.

In conclusion, the positioning accuracy test results indicate that the enhanced AR-encoded marker introduced in this study consistently delivers superior visual positioning accuracy across various water quality conditions. It proves especially effective in environments with high light attenuation coefficients, showcasing elevated positioning precision, reduced data dispersion, and markedly enhanced positioning performance. Such capabilities are invaluable in furnishing high-precision pose data for underwater swarm technologies.

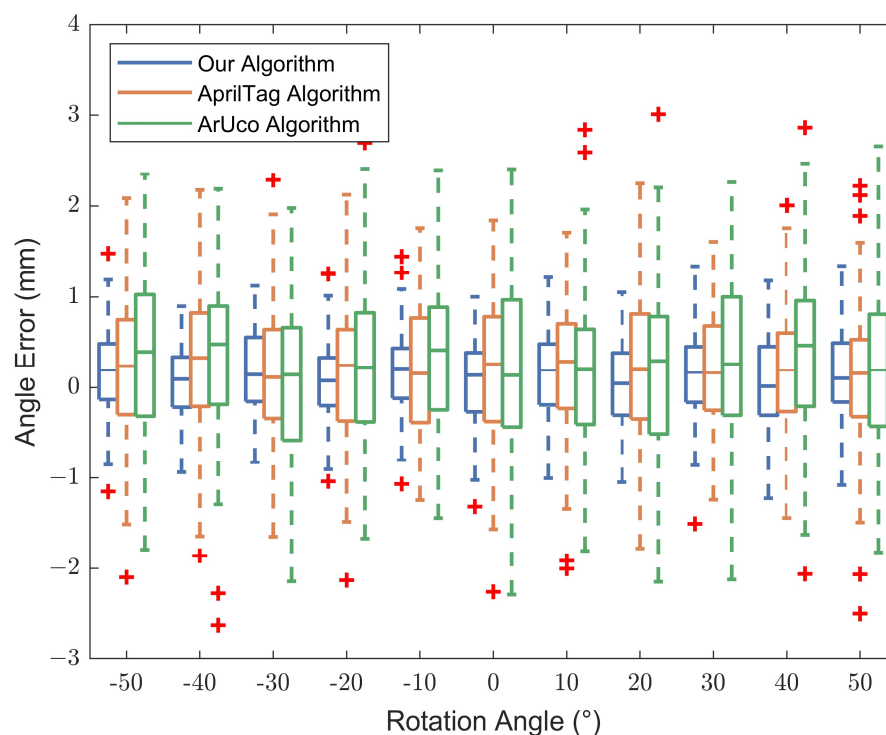


Figure 15. Boxplots of corner error for three markers with optical attenuation coefficient of 5.16. In these plots, the '+' symbol represents outliers.

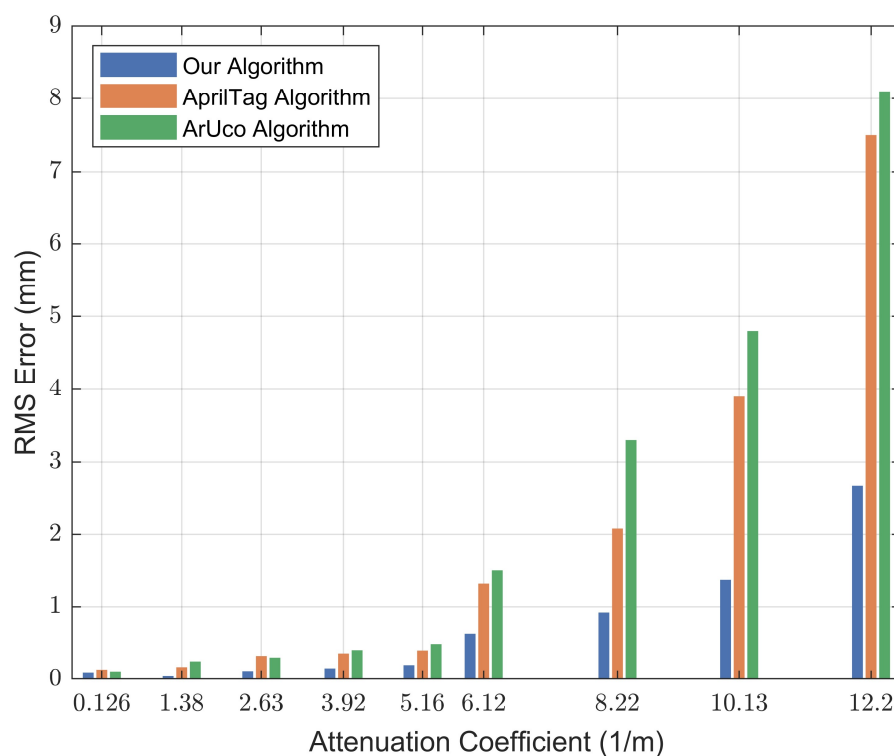


Figure 16. Grouped histogram of root-mean-square error of distance for 9 water qualities.

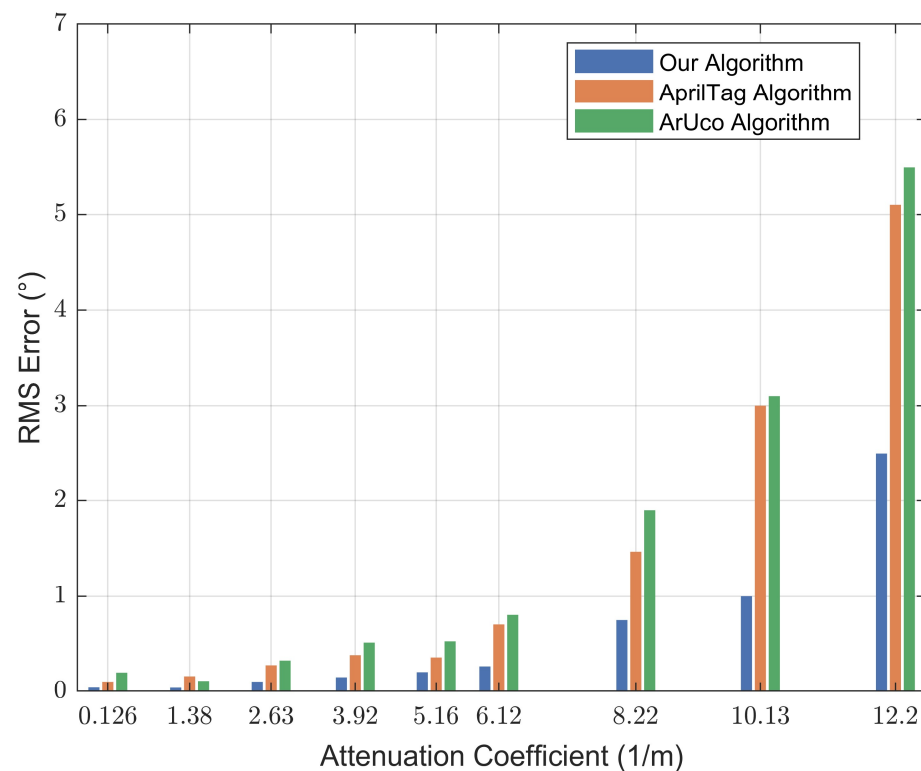


Figure 17. Grouped histograms of root-mean-square error for corners under 9 water qualities.

6.3. Underwater Swarm Localization Experiment

The experiments detailed above offer extensive quantitative validation of the method proposed in this paper, particularly in its image restoration capabilities and positional accuracy, thus ensuring dependable cluster visual positioning. Subsequently, underwater cluster positioning experiments were conducted in real-world aquatic settings, employing the proposed method for detection, recognition, and positioning within an AUV swarm.

In this experiment, conducted in waters with a light attenuation coefficient of 0.5/m, three TS Mini-AUVs served as the experimental platforms, each outfitted with visual markers identified as IDs 10, 15, and 17. The AUVs, measuring 125 mm in diameter, had lengths of 1.5 m, 2.2 m, and 1.8 m, respectively. Each AUV's head was equipped with five monocular cameras, each with a resolution of 1440×1080 ; one camera was oriented forward, with the remaining four positioned in the upward, downward, left, and right directions on the sides. The image data captured by these cameras were processed by the Jetson Orin NX 16GB processing units integrated within each AUV. Equipped with an 8-core processor functioning at 2 GHz and substantial cache capacity (2MB L2 cache and 4 MB L3 cache), these units facilitated efficient image data processing. This setup enabled each AUV to accurately perform detection, identification, and positioning within their visual range. As shown in Figure 18, the method implemented allowed for the identification and positioning of single or multiple AUVs within the field of view. Table 6 presents the pose data of the AUVs within each image, demonstrating detection and identification via the unique ID visual markers. Moreover, the AUVs' poses, including Tx, Ty, Tz, roll, pitch, and yaw angles, were ascertained through the alignment of the visual markers' image coordinates with three-dimensional coordinates.

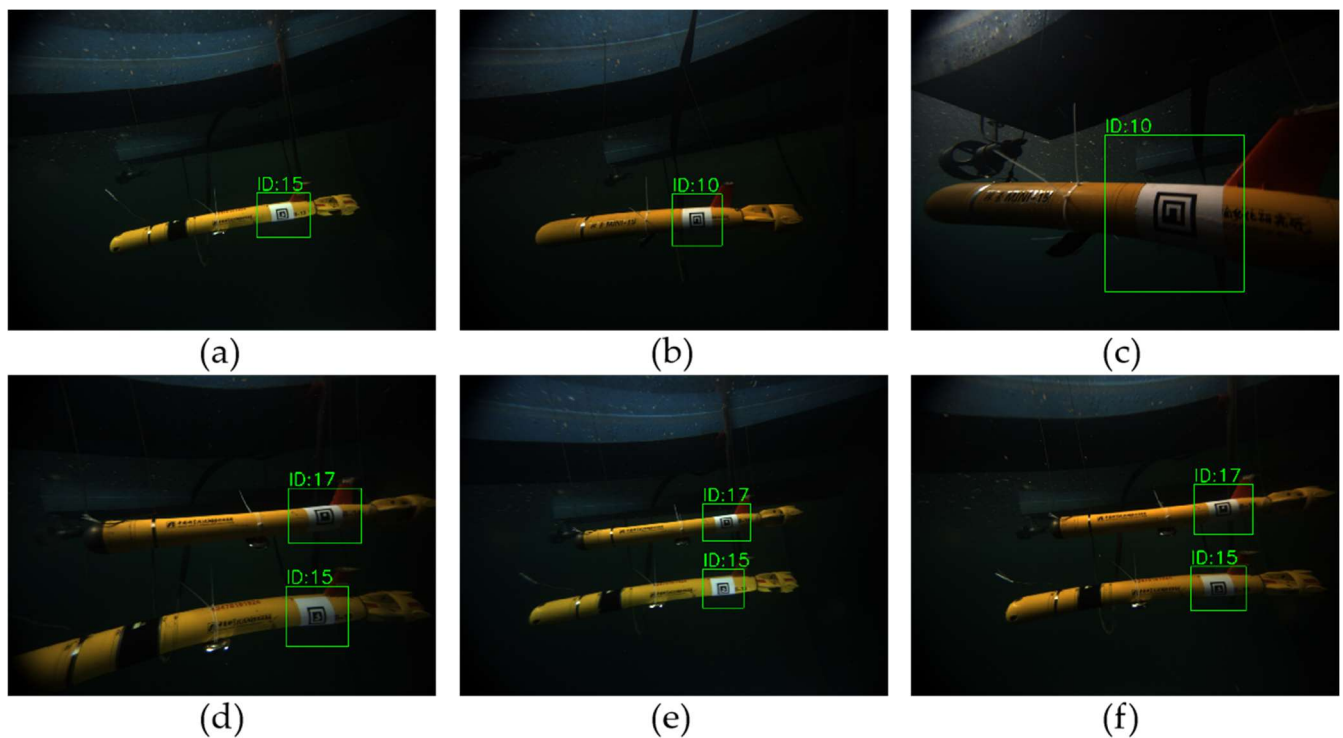


Figure 18. Results of underwater swarm localization experiment (detection and identification). Images (a–c) show the identification results when only one AUV (Autonomous Underwater Vehicle) is present in the field of view, while images (d–f) depict the identification results when two AUVs are within the field of view.

Table 6. Pose data of AUVs in images a–f in Figure 18, encompassing displacement (Tx, Ty, Tz) and attitude angles (roll, pitch, yaw).

Image	ID	Tx (mm)	Ty (m)	Tz (mm)	Roll (°)	Pitch (°)	Yaw (°)
a	15	770.525	574.52	3757.41	−29.2244	16.4337	−13.2625
b	10	200.808	492.161	2935.41	−11.4446	6.87533	−3.5698
c	10	134.463	107.986	979.786	−0.0105203	30.2773	2.01946
d	15	797.536	675.94	2445.91	−7.63387	−5.45088	−1.29422
	17	858.992	−200.079	2437.03	17.6244	25.573	3.44574
e	15	638.725	709.163	3855.67	8.76813	4.66085	−5.34771
	17	680.908	−207.105	3848.43	−19.2831	3.99968	−8.60579
f	15	989.521	560.189	3085.5	9.33964	17.5074	−6.07655
	17	922.165	−292.444	2745.54	21.8831	31.2923	4.13521

To ascertain the real-time positioning proficiency of the method introduced in this study within practical cluster positioning contexts, posture assessments of AUVs were conducted over a span of 100 consecutive image frames. Figures 19 and 20 demonstrate the six-axis pose data for each AUV from the initial to the 100th frame. Figure 19 outlines the relative pose trajectory of an individual AUV identified by ID 15, while Figure 20 presents the trajectories for two AUVs, IDs 15 and 17, within the field of view. These illustrations convey that the relative pose data between AUVs exhibited consistent and stable alterations throughout their relative motion. The analysis of the average time expended on processing, recognizing, and positioning per frame established that the mean duration per frame amounted to 61.28 milliseconds, achieving a positioning rate beyond

16 frames per second. This rate effectively enables high-frequency pose data output among the AUVs in the cluster.

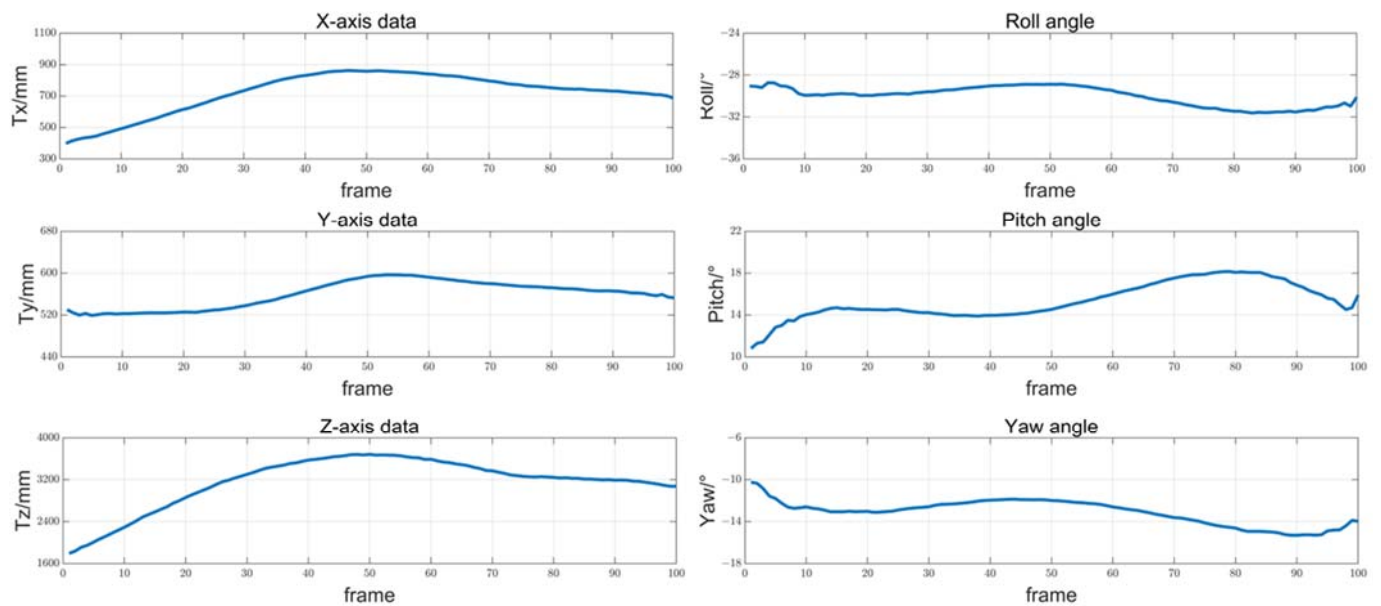


Figure 19. Frame-by-frame relative pose trajectory of the AUV with ID 15, including displacement (x -axis data, y -axis data, z -axis data) and attitude angles (roll angle, pitch angle, yaw angle).

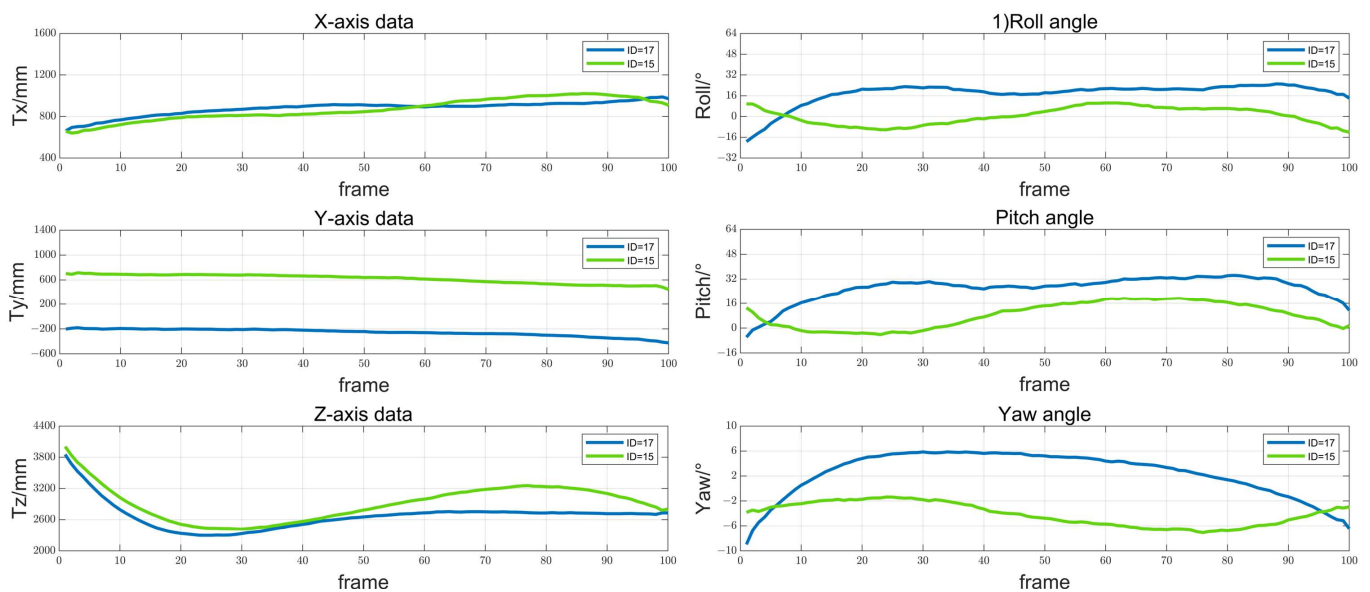


Figure 20. Frame-by-frame relative pose trajectories of two AUVs with IDs 15 and 17, including displacement (x -axis data, y -axis data, z -axis data) and attitude angles (roll angle, pitch angle, yaw angle).

In the analysis of the experimental outcomes, it was discerned that, within underwater environments characterized by a light attenuation coefficient of 0.5, effective positioning could be accomplished at distances up to 4 m. Furthermore, considering the encoding capabilities of the visual markers delineated in this investigation and the field of view pertaining to each Autonomous Underwater Vehicle (AUV), a theoretical framework for the formation of an AUV cluster comprising no less than five units is presented. This potential is further augmented when synergized with advanced swarming algorithms, thereby enhancing the prospects for the assembly of larger-scale AUV clusters.

Ultimately, the experimental evidence robustly substantiates the method's excellence in positioning accuracy and stability. It affirms the method's suitability for real-time, continuous, and stable positioning of Autonomous Underwater Vehicles within real-world aquatic settings, thus solidifying the practicability of its implementation in underwater cluster technology.

7. Conclusions

This study has presented an innovative localization method for Autonomous Underwater Vehicle (AUV) swarms utilizing augmented visual markers. This approach significantly improves the adaptability and precision of visual localization in aquatic environments. Key achievements include the development of a Hydro-Optical Image Restoration Model (HOIRM) and an augmented AR-encoded visual marker, both tailored for underwater use. The HOIRM effectively counters optical blurring and light attenuation, enhancing underwater image clarity and marker recognizability. The AR-encoded visual marker, designed to improve localization accuracy, has demonstrated superior performance in various water quality conditions, with localization precision increasing significantly in optimal and challenging environments.

The integration of these developments into a comprehensive underwater visual localization algorithm has enabled real-time, stable visual detection, recognition, and localization among AUV swarms. This paper's findings are instrumental in advancing underwater AUV swarm technologies, with significant implications for operations requiring high localization accuracy in complex conditions.

Nevertheless, the methodology proposed in this study encounters challenges in accurately localizing AUVs when visual markers are obscured or otherwise unobservable. Furthermore, the effectiveness of this approach is hampered at extended distances due to the inherent limitations in the size of these visual markers. To address these issues, future research should not only focus on enhancing the efficiency of the algorithm and the adaptability of the markers but also investigate leveraging the comprehensive structural features of AUVs for improved visual localization.

This work, therefore, lays a foundation for future advancements in underwater visual localization, aiming to meet the evolving demands of underwater applications.

Author Contributions: Conceptualization, Q.W. and Y.Y.; methodology, Q.W. and X.Z.; software, Q.W.; validation, Y.Y.; formal analysis, X.Z.; investigation, Q.W.; resources, Y.Y. and Q.Z.; writing—original draft preparation, Q.W.; writing—review and editing, Q.W., Y.Y., X.Z. and Z.H.; visualization, C.F.; supervision, Y.Y. and X.Z.; project administration, Y.Y. and Z.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The dataset supporting the conclusions of this article is not available in a public repository due to its classification under a confidential project. This dataset is proprietary and contains information that is subject to privacy and ethical restrictions. Further inquiries about the methodology and limited dataset insights can be directed to the corresponding author via the email address provided at the beginning of this article.

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

References

1. Chen, G.; Hua, M.; Liu, W.; Wang, J.; Song, S.; Liu, C.; Yang, L.; Liao, S.; Xia, X. Planning and tracking control of full drive-by-wire electric vehicles in unstructured scenario. *Proc. Inst. Mech. Eng. Part D J. Automob. Eng.* **2023**. [\[CrossRef\]](#)
2. Meng, Z.; Xia, X.; Xu, R.; Liu, W.; Ma, J. HYDRO-3D: Hybrid Object Detection and Tracking for Cooperative Perception Using 3D LiDAR. *IEEE Trans. Intell. Veh.* **2023**, *8*, 4069–4080. [\[CrossRef\]](#)
3. Liu, W.; Xia, X.; Xiong, L.; Lu, Y.; Gao, L.; Yu, Z. Automated vehicle sideslip angle estimation considering signal measurement characteristic. *IEEE Sens. J.* **2021**, *21*, 21675–21687. [\[CrossRef\]](#)
4. Su, X.; Ullah, I.; Liu, X.; Choi, D. A review of underwater localization techniques, algorithms, and challenges. *J. Sens.* **2020**, *2020*, 6403161. [\[CrossRef\]](#)

5. Ullah, I.; Chen, J.; Su, X.; Esposito, C.; Choi, C. Localization and detection of targets in underwater wireless sensor using distance and angle based algorithms. *IEEE Access* **2019**, *7*, 45693–45704. [\[CrossRef\]](#)
6. Zhang, L.; Li, Y.; Pan, G.; Zhang, Y.; Li, S. Terminal Stage Guidance Method for Underwater Moving Rendezvous and Docking Based on Monocular Vision. In Proceedings of the OCEANS Conference, Marseille, France, 17–20 June 2019.
7. Feng, J.; Yao, Y.; Wang, H.; Jin, H. Multi-AUV terminal guidance method based on underwater visual positioning. In Proceedings of the 2020 IEEE International Conference on Mechatronics and Automation (ICMA), Beijing, China, 13–16 October 2020; pp. 314–319.
8. Xu, Z.; Haroutunian, M.; Murphy, A.J.; Neasham, J.; Norman, R. An underwater visual navigation method based on multiple ArUco markers. *J. Mar. Sci. Eng.* **2021**, *9*, 1432. [\[CrossRef\]](#)
9. Wu, Q.; Li, S.; Hao, Y.; Zhu, F.; Tang, S.; Zhang, Q. Model-Based Visual Hovering Positioning Technology for Underwater Robots. *High Technol. Lett.* **2005**, *15*, 6. (In Chinese)
10. Hao, Y.; Wu, Q.; Zhou, C.; Li, S.; Zhu, F. Hovering Positioning Technology and Implementation of Underwater Robots Based on Monocular Vision. *Robot* **2006**, *28*, 656–661. (In Chinese)
11. Wen, Z.; Wang, Y.; Kuijper, A.; Di, N.; Luo, J.; Zhang, L.; Jin, M. On-orbit real-time robust cooperative target identification in complex background. *Chin. J. Aeronaut.* **2015**, *28*, 1451–1463. [\[CrossRef\]](#)
12. Zhang, Z.; Zhang, S.; Li, Q. Robust and accurate vision-based pose estimation algorithm based on four coplanar feature points. *Sensors* **2016**, *16*, 2173. [\[CrossRef\]](#)
13. Lee, D.; Kim, G.; Kim, D.; Myung, H.; Choi, H.-T. Vision-based object detection and tracking for autonomous navigation of underwater robots. *Ocean Eng.* **2012**, *48*, 59–68. [\[CrossRef\]](#)
14. Garrido-Jurado, S.; Muñoz-Salinas, R.; Madrid-Cuevas, F.J.; Marín-Jiménez, M.J. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognit.* **2014**, *47*, 2280–2292. [\[CrossRef\]](#)
15. Wang, J.; Olson, E. AprilTag 2: Efficient and robust fiducial detection. In Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Daejeon, Republic of Korea, 9–14 October 2016; pp. 4193–4198.
16. Olson, E. AprilTag: A robust and flexible visual fiducial system. In Proceedings of the 2011 IEEE International Conference on Robotics and Automation, Shanghai, China, 9–13 May 2011; pp. 3400–3407.
17. Krogus, M.; Haggemiller, A.; Olson, E. Flexible layouts for fiducial tags. In Proceedings of the 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Macau, China, 3–8 November 2019; pp. 1898–1903.
18. Rijlaarsdam, D.D.; Zwick, M.; Kuiper, J. A novel encoding element for robust pose estimation using planar fiducials. *Front. Robot. AI* **2022**, *9*, 838128. [\[CrossRef\]](#)
19. Ababsa, F.-E.; Mallem, M. Robust camera pose estimation using 2d fiducials tracking for real-time augmented reality systems. In Proceedings of the 2004 ACM SIGGRAPH International Conference on Virtual Reality Continuum and Its Applications in Industry, New York, NY, USA, 16–18 June 2004; pp. 431–435.
20. Romero-Ramirez, F.J.; Muñoz-Salinas, R.; Medina-Carnicer, R. Speeded up detection of squared fiducial markers. *Image Vis. Comput.* **2018**, *76*, 38–47. [\[CrossRef\]](#)
21. Ren, R.; Zhang, L.; Liu, L.; Yuan, Y. Two AUVs guidance method for self-reconfiguration mission based on monocular vision. *IEEE Sens. J.* **2021**, *21*, 10082–10090. [\[CrossRef\]](#)
22. Yang, Y.; Zhou, X.; Hu, Z.; Fan, C.; Wang, Z.; Fu, D.; Zheng, Q. Research on High-Precision Formation Technology for Underwater Robots Based on Visual Positioning with No Communication. *Digit. Ocean. Underw. Attack Def.* **2022**, *5*, 9. (In Chinese)
23. Xiang, Y.; Schmidt, T.; Narayanan, V.; Fox, D. PoseCNN: A Convolutional Neural Network for 6D Object Pose Estimation in Cluttered Scenes. *arXiv* **2017**, arXiv:1711.00199.
24. Li, Z.; Wang, G.; Ji, X. CDPN: Coordinates-Based Disentangled Pose Network for Real-Time RGB-Based 6-DoF Object Pose Estimation. In Proceedings of the International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019.
25. Rozantsev, A.; Salzmann, M.; Fua, P. Beyond Sharing Weights for Deep Domain Adaptation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *41*, 801–814. [\[CrossRef\]](#)
26. Rad, M.; Oberweger, M.; Lepetit, V. Feature Mapping for Learning Fast and Accurate 3D Pose Inference from Synthetic Images. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018.
27. Koreitem, K.; Li, J.; Karp, I.; Manderson, T.; Shkurti, F.; Dudek, G. Synthetically trained 3d visual tracker of underwater vehicles. In Proceedings of the OCEANS 2018 MTS/IEEE Charleston, Charleston, SC, USA, 22–25 October 2018; pp. 1–7.
28. He, K.; Sun, J.; Tang, X. Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *33*, 2341–2353.
29. Garg, D.; Garg, N.K.; Kumar, M. Underwater image enhancement using blending of CLAHE and percentile methodologies. *Multimed. Tools Appl.* **2018**, *77*, 26545–26561. [\[CrossRef\]](#)
30. Iqbal, K.; Salam, R.A.; Osman, A.; Talib, A.Z. Underwater Image Enhancement Using an Integrated Colour Model. *IAENG Int. J. Comput. Sci.* **2007**, *32*, 239–244.
31. Tang, C.; von Lukas, U.F.; Vahl, M.; Wang, S.; Wang, Y.; Tan, M. Efficient underwater image and video enhancement based on Retinex. *Signal Image Video Process* **2019**, *13*, 1011–1018. [\[CrossRef\]](#)
32. Zhang, S.; Wang, T.; Dong, J.; Yu, H. Underwater image enhancement via extended multi-scale Retinex. *Neurocomputing* **2017**, *245*, 1–9. [\[CrossRef\]](#)

33. Hou, G.; Pan, Z.; Huang, B.; Wang, G.; Luan, X. Hue preserving-based approach for underwater colour image enhancement. *IET Image Process* **2018**, *12*, 292–298. [[CrossRef](#)]
34. Jia, D.; Ge, Y. Underwater image de-noising algorithm based on nonsubsampling contourlet transform and total variation. In Proceedings of the 2012 International Conference on Computer Science and Information Processing (CSIP), Xi'an, China, 24–26 August 2012; pp. 76–80.
35. Carlevaris-Bianco, N.; Mohan, A.; Eustice, R.M. Initial results in underwater single image dehazing. In Proceedings of the Oceans 2010 Mts/IEEE Seattle, Seattle, WA, USA, 20–23 September 2010; pp. 1–8.
36. Wang, Y.; Zhang, J.; Cao, Y.; Wang, Z. A deep CNN method for underwater image enhancement. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; pp. 1382–1386.
37. Li, C.; Anwar, S.; Porikli, F. Underwater scene prior inspired deep underwater image and video enhancement. *Pattern Recognit.* **2020**, *98*, 107038. [[CrossRef](#)]
38. Cho, Y.; Kim, A. Visibility enhancement for underwater visual SLAM based on underwater light scattering model. In Proceedings of the 2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore, 29 May–3 June 2017; pp. 710–717.
39. Xie, J.; Hou, G.; Wang, G.; Pan, Z. A variational framework for underwater image dehazing and deblurring. *IEEE Trans. Circuits Syst. Video Technol.* **2021**, *32*, 3514–3526. [[CrossRef](#)]
40. Zhang, W.; Wang, Y.; Li, C. Underwater image enhancement by attenuated color channel correction and detail preserved contrast enhancement. *IEEE J. Ocean. Eng.* **2022**, *47*, 718–735. [[CrossRef](#)]
41. Li, X.; Hou, G.; Tan, L.; Liu, W. A hybrid framework for underwater image enhancement. *IEEE Access* **2020**, *8*, 197448–197462. [[CrossRef](#)]
42. Ho, C.S. Precision of digital vision systems. *IEEE Trans. Pattern Anal. Mach. Intell.* **1983**, *5*, 593–601. [[CrossRef](#)]
43. Moon, C.; McVey, E. Precision measurement techniques using computer vision. In Proceedings of the Conference Record of the 1990 IEEE Industry Applications Society Annual Meeting, Seattle, WA, USA, 7–12 October 1990; pp. 1521–1526.
44. Wang, R.; Li, R.; Sun, H. Haze removal based on multiple scattering model with superpixel algorithm. *Signal Process* **2016**, *127*, 24–36. [[CrossRef](#)]
45. Zhu, Q.; Mai, J.; Shao, L. A fast single image haze removal algorithm using color attenuation prior. *IEEE Trans. Image Process* **2015**, *24*, 3522–3533. [[PubMed](#)]
46. Deng, C.; Ma, L.; Lin, W.; Ngan, K.N. *Visual Signal Quality Assessment*; Springer: Berlin/Heidelberg, Germany, 2016.

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.