

Article | Received 22 May 2026; Revised 18 September 2026; Accepted 24 September 2026; Published 30 September 2026
<https://doi.org/10.55092/interdiscipline20260003>



Low-cost and robust docking angle estimation for miniature USVs through the integration of camera and multi-zone ToF sensor

Junju Huang, Xuanyu Qian, Jiayuan Li and Wei Jiang*

Microelectronics Thrust, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China

* Correspondence author; E-mail: weijiang2@hkust-gz.edu.cn.

Highlights:

- A low-cost and robust dual-sensing strategy is proposed for estimating docking angles of miniature USVs.
- The dual-sensing angle estimation system combines a camera with an 8×8 multi-zone time-of-flight (ToF) sensor.
- Camera and ToF sensor calibration are implemented to enhance the accuracy and stability of short-range angle measurement.
- Experimental results validate the robustness and practical use of the proposed system under varying illumination conditions.

Abstract: Autonomous docking of miniature unmanned surface vehicles (USVs) requires compact, low-cost, and robust sensing for estimating angles within short ranges. To fulfill these requirements, this work presents a dual-sensing strategy that combines a camera and an 8×8 multi-zone time-of-flight (ToF) sensor. In this strategy, the camera-based method can estimate the docking angle through calibrated AprilTag pose estimation, while the ToF-based method can derive angular information from column-wise range gradients. Experiments covering 20–80 cm and 10° – 50° demonstrate that both methods can achieve reliable angle estimation, with an overall mean absolute error (MAE) of angle below 1.5° . The results further show complementary performance of the proposed design: ToF sensing is more stable at small angles while camera-based estimation performs better at larger angles. Moreover, experiments validate the robustness and practical use of the proposed system under varying illumination scenarios.

Keywords: unmanned surface vehicle (USV); low cost; angle estimation; sensor fusion; time-of-flight (ToF); camera



Copyright©2026 by the authors. Published by ELSP. This work is licensed under Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium provided the original work is properly cited.

1. Introduction

Unmanned surface vehicles (USVs) are multifunctional marine platforms capable of performing various tasks either autonomously or under remote control. Benefiting from such operational flexibility, USVs is widely used for coastal inspection and monitoring tasks, including marine environmental sampling, hydrographic surveying, and maritime security surveillance. Compared with conventional manned vessels, USVs offer enhanced deployment flexibility and reduced operational costs, making them highly suitable for repetitive, labor-intensive, and high-risk maritime missions [1–4].

To sustain the maritime missions of USVs, periodic berthing and departure are unavoidable, which makes reliable autonomous docking critical to their continuous operation. Nevertheless, USV terminal docking is performed in constrained short-range maritime environments with limited visibility (e.g., fog or smoke) and varying illumination, which requires precise relative position and angle estimations between the USV and the berth. In such scenarios, Global Navigation Satellite Systems lack sufficient precision required for reliable short-range relative localization and angle estimation. Vision-based perception degrades under maritime interference such as surface glare, water reflections, light scattering, and low-light conditions. Although light detection and ranging (LiDAR) systems can achieve high ranging accuracy and dense three-dimensional mapping, their high cost of the hardware and integration complexity limit their applications on low-cost and cluster-coordinated miniature USVs [5]. These limitations discussed above demonstrate that single-modal sensing is difficult to meet the reliability and cost-efficiency requirements of practical USV docking applications simultaneously. Accordingly, the integration and fusion of complementary sensors become essential for robust short-range positioning and angle estimations during USV terminal docking under challenging maritime conditions [6–8].

To overcome the limitations of single-sensor systems in USV terminal docking, this paper proposes a low-cost docking angle estimation system integrating a monocular camera and a multi-zone time-of-flight (ToF) sensor. This angle estimation system fuses geometric parameters extracted from the Apriltag by the camera and spatial distance distributions acquired by the multi-zone ToF sensor, thus improving the overall accuracy and environmental robustness.

The rest of this paper is organized as follows. Section 2 provides the reviews of fundamental sensing principles and state-of-the-art research progress on USV autonomous docking. Section 3 presents the system methodology and sensor calibration procedures in details. Section 4 describes system integration and comprehensive experimental validation results. Finally, Section 5 concludes this work and discusses potential future research directions.

2. Sensing strategy and recent progress in USV docking

Dynamic disturbances, such as varying ambient light, can significantly degrade the performance of USV docking systems. Therefore, a robust sensing solution is needed for reliable autonomous docking. In this section, we summarize several mainstream sensing modalities for short-range docking, including some feasible sensing strategies from other domains that can be adapted to the current maritime USV applications.

Vision sensors are widely adopted in the USV docking system, due to their low hardware cost and mature pose estimation capability. There are two widely used vision-based strategies. The first strategy utilizes artificial fiducial markers (e.g., ArUco or AprilTag), where predefined marker geometry is used to perform relative pose calculation via the Perspective-n-Point (PnP) algorithm [9]. The second strategy

combines data-driven detection with model-based geometry. And a neural network is firstly used to localize the target region before being processed by geometric algorithms to get relative position and pose [10]. Although the two methods mentioned above have decent performance under ideal illumination, they suffer from significant accuracy degradation under complex interferences. Especially, when the illumination level is extremely low, they may even fail in detecting complete markers, which leads to unreliable docking operations.

To compensate the illumination susceptibility of the vision sensing, active detection and ranging sensors have been introduced to enhance short-range robustness. LiDAR is the most extensively studied active sensor for its high ranging precision and ability to provide direct geometric constraints when approaching the berth. Existing works adopted onboard 3D LiDAR to obtain berth states from point cloud data and integrated LiDAR localization with navigable region detection to improve docking safety in cluttered port environments [11,12]. However, commercial high-performance LiDAR devices suffer from high cost and complex integration, limiting their scalability for low-cost and cluster-deployed USVs [5]. To have a low-cost solution without adding too many weights, miniaturized multi-zone ToF sensors have become attractive alternatives. Previous studies on nano-robot navigation have verified that multi-zone ToF sensors can maintain reliable angle estimation despite a certain portion of measurement zones return invalid data [13]. Such high tolerance to data loss is suitable for maritime scenarios, in which water reflections and wave-induced interferences may frequently cause failed detections.

Generally, a single-modal sensing is insufficient in the complex maritime environments. Therefore, multi-sensor fusion has become an important research direction to improve the USV docking perceptual robustness [14]. Existing sensor fusion schemes include camera–LiDAR fusion and global positioning system (GPS)–inertial measurement unit (IMU) fusion [15]. Infrared cameras and depth sensors have also been validated in some robotic systems to prevent visual degradation. These reported schemes can be easily adapted to maritime docking scenarios [16,17]. Camera–ToF fusion has also been investigated for relative pose estimation in small robotic platforms, indicating its compactness and lightweight [18,19]. However, few studies have developed low-cost fusion frameworks that can achieve precise angle estimation for miniature USVs under varying illumination conditions.

3. Angle estimation method

In this section, we introduce the details of the proposed dual-sensing angle estimation system, which consists of a camera and a ToF sensor. To implement the fast prototyping, selected sensors feature mature software ecosystems and simple interfacing to minimize integration complexity and hardware failure risks during field tests. Furthermore, the entire system adopts a lightweight computational architecture compatible with low-power embedded processors, alleviating the requirement for high-performance onboard computing hardware and improving adaptability to resource-constrained miniature USV platforms.

An OpenMV H7 Plus camera module, which integrates image acquisition and onboard processing in a compact unit, was selected for the miniature-USV prototype. In addition, a VL53L5CX multi-zone ToF sensor was selected for active depth sensing, which uses an 8×8 single-photon avalanche diode (SPAD) array to generate spatial depth maps. The integration of passive visual sensing (*i.e.*, camera) and active ToF ranging constructs a complementary dual-modal hardware core for the proposed docking angle estimation system.

Figure 1 illustrates the USV docking and departure motions relative to the berth and the sensor installation layout. During the short-range docking phase, both sensors are installed on the USV hull side and oriented toward the berth to keep the target within the sensing range during pose adjustment. For data processing, the OpenMV H7 Plus camera utilizes an integrated STM32 microcontroller for onboard image processing, while the VL53L5CX ToF sensor is controlled by an Arduino board. These two sensing pipelines operate in parallel, which helps reduce interference between the modules.

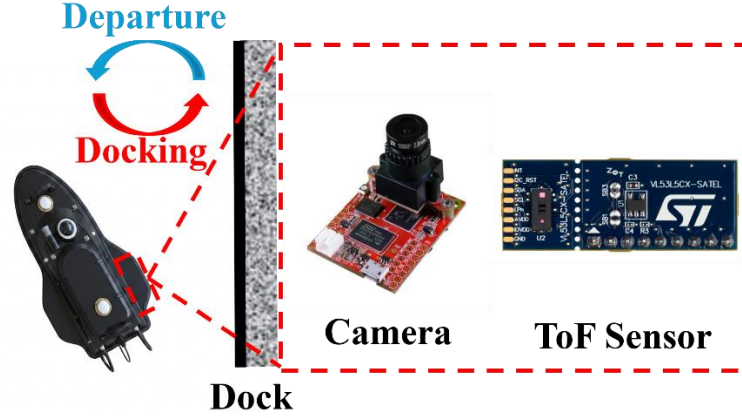


Figure 1. System overview of the proposed sensing setup for USV docking and departure near a dock.

3.1. Principle of angle estimation

As depicted in Figure 2, a yaw misalignment angle β exists between the USV's heading direction and the planar berth surface. Therefore, precise estimation of this angle is critical for accurate autonomous docking. To address this issue, this work proposes a dual-sensor fusion strategy combining AprilTag-based visual method and 45° field-of-view (FoV) multi-zone ToF active ranging, which enhances the environmental adaptability and reliability of docking angle estimation. For visual angle estimation, an AprilTag fiducial marker, which provides robust visual fiducial detection and supports single-image pose estimation [20,21], is deployed on the berth surface as a geometric reference. When the marker lies within the camera FoV, its four image corner points p_i ($i = 1,2,3,4$) are detected and matched to the corresponding known points P_i ($i = 1,2,3,4$) in the AprilTag coordinate frame. Based on these matched corner points and the calibrated camera intrinsics, the relative pose between the camera and the AprilTag can be recovered by solving the PnP problem, formulated as:

$$s_i p_i = K [R \mid t] P_i \quad (1)$$

where s_i is the projection scaling factor, p_i denotes the detected image coordinate of an AprilTag corner, P_i denotes the corresponding known coordinate in the AprilTag local frame, K is the camera intrinsic matrix. The unknowns R and t represent the rotation and translation that transform points from the AprilTag frame into the camera frame, respectively.

To further improve accuracy, the calibrated lens distortion coefficients are applied to correct image-plane distortions prior to pose solving. The estimated rotation matrix R describes the orientation of the dock-mounted AprilTag, and its yaw component is extracted as the camera-based docking angle β_{cam} . However, visual detection reliability deteriorates drastically under low-light or dark conditions, leading to unstable or invalid visual angle estimation. In such scenarios, the multi-zone ToF sensor

serves as a complementary alternative. The ToF sensor generates an 8×8 ranging matrix over a 45° FoV. Each column of the matrix C_i ($i = 0 - 7$) corresponds to a fixed viewing angle α_i ($i = 0 - 7$) relative to the sensor. In this case, the distance d_i ($i = 0 - 7$) for each column is obtained by averaging the zone readings within that column. When the USV is not aligned with the berth ($\beta \neq 0$), one side of the berth is closer to the sensor than the other one. In this case, the measured distance varies across the columns. Once the angle is zero, the distances measured from the 8 columns are equal. This is obtained by assuming that the observed berth surface is approximated as a plane, and the ToF sensor is approximately horizontally installed. Additionally, the short-range pitch and roll variations of the USV are assumed to be negligible. Under these assumptions, the column-wise ranging differences are mainly determined by the relative yaw angle between the sensor and the berth surface. Based on this geometric principle, the USV yaw angle β_{ToF} can be derived from ranging deviations across columns, as defined in the following equation [12].

$$\beta_{ToF} = \arctan \left[\frac{d_r - d_l}{\tan(\alpha_l) \cdot d_l + \tan(\alpha_r) \cdot d_r} \right] \quad (2)$$

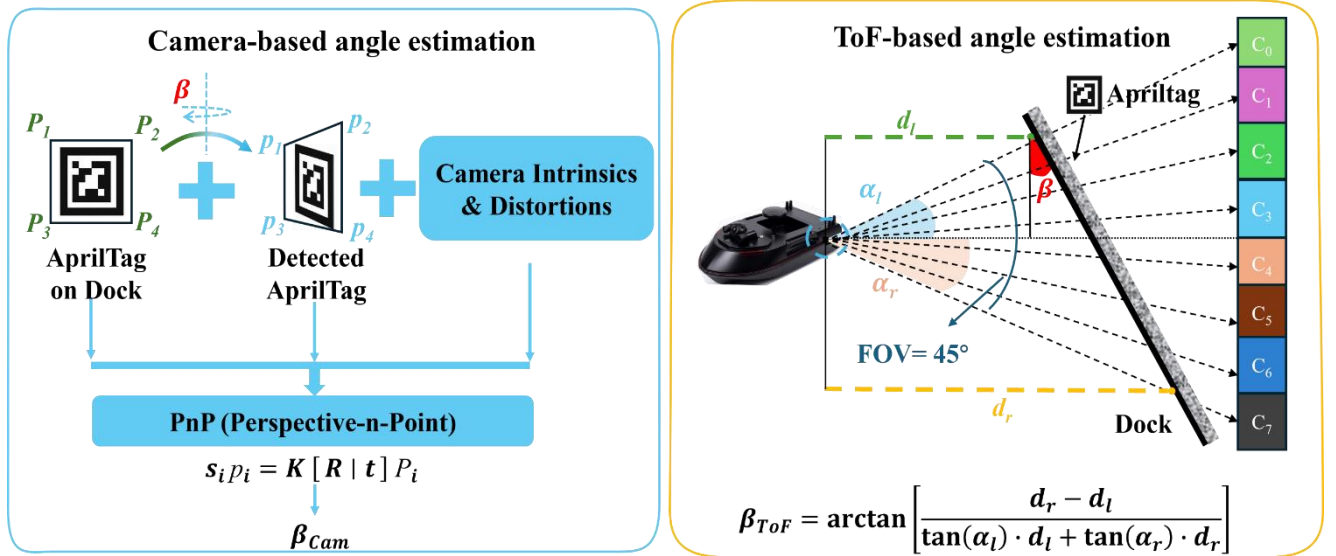


Figure 2. Principle of angle estimation.

3.2. Camera-based angle estimation

3.2.1. Intrinsic calibration of camera

Camera intrinsic calibration was performed using a planar checkerboard. During the calibration, the checkerboard images are captured at various distances and angles to cover the full camera FoV and depth range. By using the checkerboard corner correspondences between 2D image pixels and 3D physical coordinates to fit a distorted pinhole imaging model, two groups of parameters are obtained: intrinsic matrix and distortion coefficients. The intrinsic matrix contains the focal lengths (f_x, f_y) in pixels and the principal point coordinates (c_x, c_y), which together describe the projection geometry of the camera. The distortion coefficients, represented as the vector $[k_1, k_2, p_1, p_2, k_3]$, account for radial distortion (k_1, k_2, k_3) caused by lens curvature and tangential distortion (p_1, p_2) arising from lens misalignment. To illustrate the calibration procedure, Figure 3 presents a representative calibration sample and calibrated

parameters. The calibration quality was further checked using reprojection error. For each image, the detected checkerboard corners were compared with the image locations obtained by projecting the corresponding 3D points. The overall root mean square (RMS) reprojection error was 0.223 pixel. Across all the images, the mean, median, and maximum per-image reprojection error were 0.033, 0.031, and 0.061 pixel, respectively. The low reprojection errors verify the accuracy and stability of calibrated parameters for subsequent visual angle estimation.

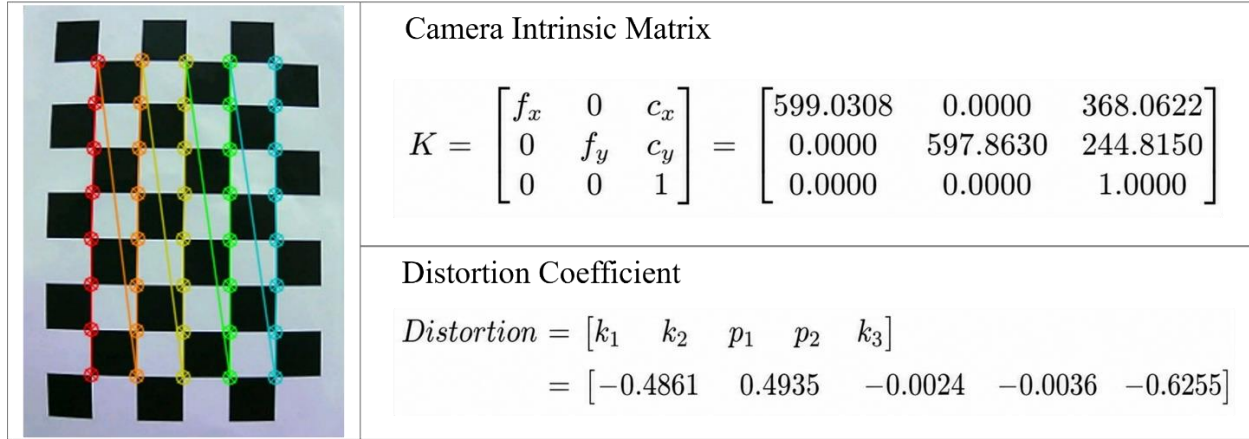


Figure 3. Representative image and results of camera intrinsic calibration.

3.2.2. Performance evaluation of camera-based angle estimation

Prior to AprilTag marker detection, the calibrated intrinsic parameters are applied to the raw images to correct lens distortion and improve pose estimation accuracy. AprilTag features are extracted from rectified images, and docking angles are calculated from the rotation matrix of the estimated pose. As shown in Figure 4a, an AprilTag marker and the camera module are installed on the rotation stages in the experimental setup. To evaluate the accuracy and precision of the camera-based angle estimation, comparative experiments are conducted at distances ranging from 20 cm to 80 cm and incident angles from 10° to 50°. Each test group collects 100 consecutive frames for statistical analysis. The errors (*i.e.*, difference between actual angle and the mean of the measured angle) and standard deviations are summarized in Table 1.

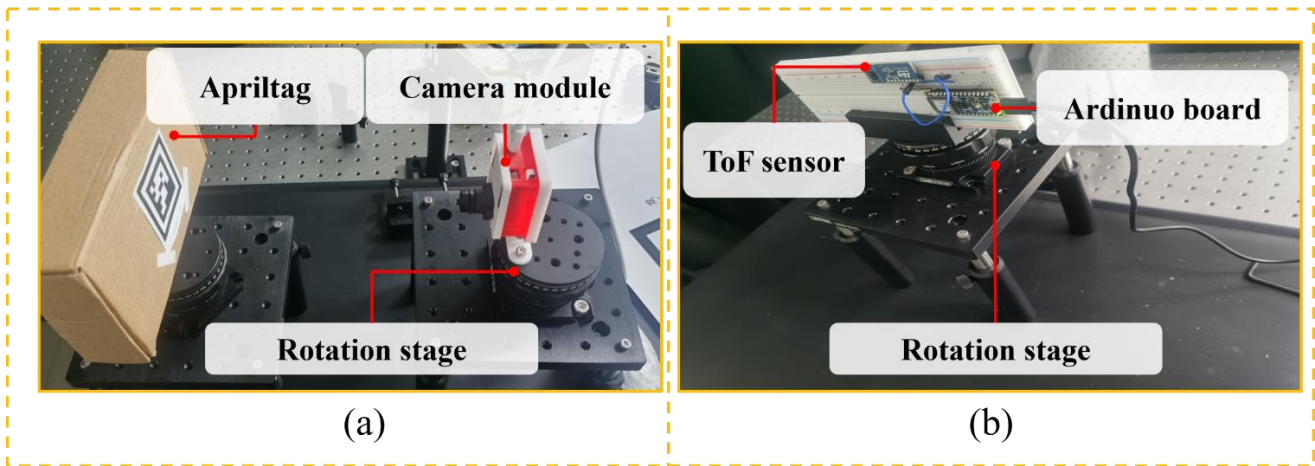


Figure 4. Experimental setups for angle estimation. (a) Camera-based; (b) ToF sensor-based.

Table 1. Camera-based angle estimation results.

Distant (cm)	10°	20°	30°	40°	50°
	Error ± S.D.	Error ± S.D.	Error ± S.D.	Error ± S.D.	Error ± S.D.
20	-1.26 ± 1.12	0.05 ± 0.35	-0.36 ± 0.21	-0.79 ± 0.25	0.03 ± 0.15
40	-3.75 ± 1.45	2.01 ± 0.58	-0.60 ± 0.28	-1.00 ± 0.32	-0.20 ± 0.22
60	-4.35 ± 2.10	-0.78 ± 0.65	-1.13 ± 0.35	1.88 ± 0.45	-0.08 ± 0.28
80	-5.03 ± 2.85	-1.16 ± 0.82	-1.97 ± 0.55	1.05 ± 0.62	-0.34 ± 0.42

1. Error and Standard Deviation (S.D.) for each measurement are calculated based on 100 frames;

2. Error = Mean measured angle–Reference angle.

Two key trends are observed from Table 1. First, the standard deviations of the angles increase with detection distance since the reduced marker pixel occupancy in the image for the longer distance weakens corner localization accuracy and geometric constraint sufficiency. Second, the errors of the angle estimation increase as the incident angles become smaller. For example, at a small incident angle of 10°, the AprilTag appears nearly frontal to the camera and exhibits only limited perspective distortion, which provides weaker constraints for the PnP-based angle estimation and makes the result more sensitive to localization noise. In contrast, as the incident angle increases to 20° or above, the projected marker shape deforms noticeably into a trapezoid, providing robust geometric cues for more accurate angle estimates. Overall, the vision-based method achieves favorable performance, with mean errors below 2° and standard deviations below 1° under most test conditions for the incident angles of 20° or above.

3.3. ToF-based angle estimation

3.3.1. Characterization and calibration of ToF sensor

To characterize the multi-zone ToF sensor, we performed a planar-wall range test with its experimental setup illustrated in Figure 4b. In this setup, the sensor was kept parallel to a flat wall, and measurements were collected at 20 cm intervals from 20 cm to 120 cm. For each distance, 100 frames were acquired to ensure statistical validity. For each pixel, we calculated the mean distance and standard deviation to characterize the spatial bias and temporal stability across the 8×8 array. Pixels with extreme distance values or measurement errors larger than the reference tolerance are defined as invalid. These invalid pixels result from ranging failures caused by the missing of echo signals or out-of-range targets. To mitigate the data corruptions from these invalid pixels, the algorithm calculates column-wise invalid pixel ratios as weighting coefficients to suppress the contribution of low-quality columns during angle estimation.

Box plots in Figure 5 illustrate ToF ranging error distributions across reference distances. Within the short range relevant to this work (up to 120 cm), error distributions remain concentrated with centimeter-level median errors. Increasing detection distance attenuates echo signal strength and reduces signal-to-noise ratios, leading to slightly expanded error ranges and higher invalid pixel ratios, especially beyond 80 cm. Apart from distance dependence, target surface reflectance can introduce systematic ranging bias. Previous studies have shown that reflectance-induced errors can be suppressed through calibration and post-processing, although the error characteristics may still depend on sensor hardware and operating conditions [22].

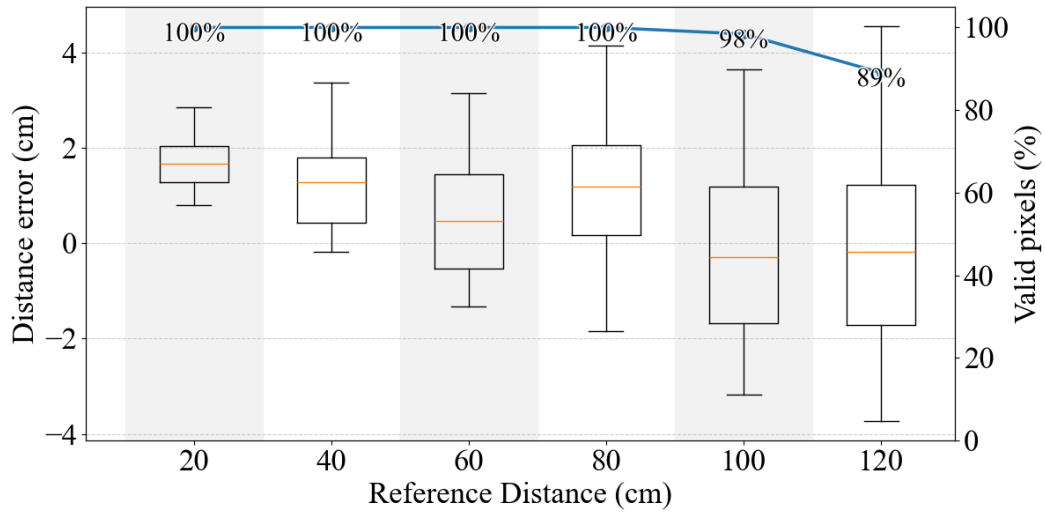


Figure 5. Box plots of distance error and valid-pixel count at different reference distances.

Figure 6 presents the distance offset heatmaps measured at 120 cm and reveals distinct spatial error patterns across the 8×8 sensing array. In Figure 6, the mean distance offset for each pixel is shown in the top panel and the corresponding standard deviation is shown in the bottom panel. For example, the distance offset of the pixel located at row 0, column 0 has a mean value of -215.24 mm and a standard deviation of 5.7 mm. It is worth noting that the edge and corner pixels exhibit larger offsets while central pixels show smaller errors. A consistent systematic gradient is observed: bottom-right pixels generate positive offsets, whereas top-left pixels generate negative offsets. This pattern is mainly attributed to residual geometric misalignment between the sensor plane and the target wall. During this measurement, even a minor sensor mounting tilt relative to the target plane can cause different responses of the pixel array. The maximum offset reaches 21 cm at 120 cm. Therefore, 80 cm is defined as the sensor's maximum stable effective working range, which fully satisfies the dynamic docking adjustment requirements of the tested 53 -cm miniature USV.

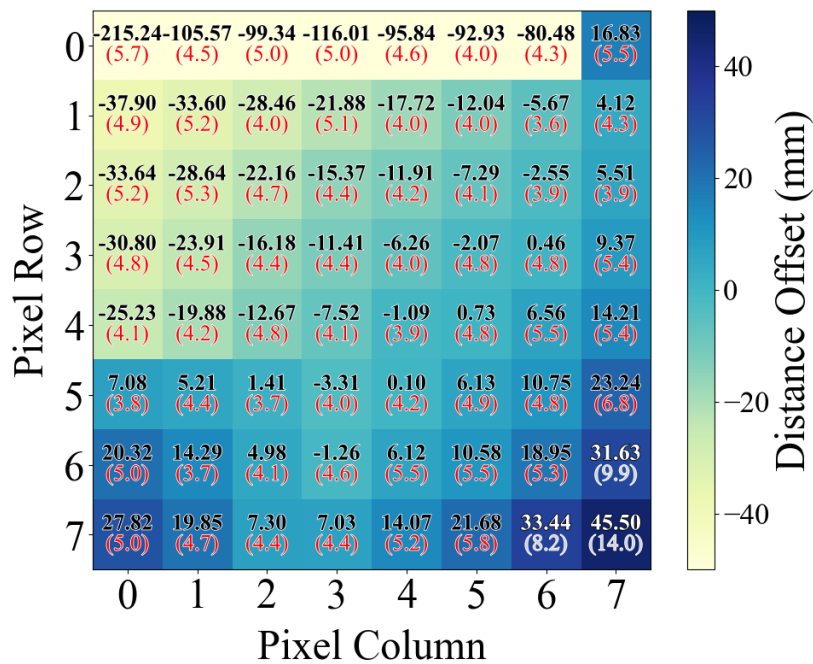


Figure 6. ToF spatial heatmap at 120 cm with per-pixel offset (top) and standard deviation (bottom).

3.3.2. Performance evaluation of ToF-based angle estimation

Following sensor characterization and calibration, ToF-based angle estimation performance is evaluated using the experimental setup in Figure 4b. Tests are performed at distances from 20 to 80 cm with a step size of 20 cm, and at angles from 10° to 50° with a step size of 10°. A multi-stage preprocessing pipeline is implemented to refine raw ToF data prior to angle calculation. First, column-wise Median Absolute Deviation (MAD) filtering is applied independently to each column to identify and remove localized outlier points. Second, columns with fewer than five valid pixels are discarded due to insufficient data credibility. Finally, valid columns are constrained by left-to-right ranging monotonicity derived from planar target geometry, removing abnormal measurements induced by water reflections and environmental obstacles. The preprocessed high-quality dataset is utilized for final angle computation.

Table 2 summarizes angle estimation performance of the ToF sensor. At distances of 60 cm or less and incident angles of 30° or smaller, the ToF sensor achieves stable estimation with maximum angular errors below 3°. However, for the longer distance of 80 cm, the angle estimation accuracy degrades noticeably due to the increased per-pixel noise and invalid data proportion caused by the attenuation of reflected light signals. In addition, large incident angles force partial far-end sensing columns beyond the stable detection range (80 cm), resulting in excessive ranging errors and reducing robustness of the estimation. Overall, the ToF sensor provides reliable angle estimation and can achieve optimal performance at short detection distances.

Table 2. ToF-based angle estimation results.

Distant (cm)	10°	20°	30°	40°	50°
	Error ± S.D.	Error ± S.D.	Error ± S.D.	Error ± S.D.	Error ± S.D.
20	1.06 ± 0.18	0.05 ± 0.20	−2.42 ± 0.21	−3.31 ± 0.22	−3.01 ± 0.29
40	−0.11 ± 0.19	−1.64 ± 0.20	−2.57 ± 0.21	−1.56 ± 0.27	−3.40 ± 0.36
60	0.67 ± 0.21	−0.56 ± 0.23	−0.19 ± 0.25	−1.12 ± 0.27	−3.76 ± 0.61
80	0.84 ± 0.23	−0.16 ± 0.24	−0.39 ± 0.49	−1.47 ± 0.72	−1.52 ± 2.51

1. Error and Standard Deviation (S.D.) for each measurement are calculated based on 100 frames;

2. Error = Mean measured angle − Reference angle.

3.4. Sensor complementarity and fusion evaluation

3.4.1. Complementary performance and weighted fusion

The angle evaluation results show that the camera and multi-zone ToF sensor provide complementary properties in USV docking angle estimation. As shown in Table 1, Table 2, and Figure 7b, the vision-based scheme suffers from larger estimation errors and higher variance at small reference angles, particularly at 10°. However, the ToF-based approach can maintain stable accuracy under the same operating conditions. These results align well with the previous analysis. That is, when the AprilTag is nearly front-parallel to the camera, insufficient perspective deformation weakens geometric constraints for visual pose solving. By contrast, the ToF sensor can still derive reliable angular information from spatial ranging gradient distributions. Conversely, the error tendency reverses at large reference angles such as 50°.

The camera-based method achieves lower errors and smaller variance, as the significantly deformed AprilTag provides strong perspective geometric constraints for robust PnP pose estimation. On the contrary, the ToF sensor yields degraded accuracy at large angles, mainly due to its limited spatial resolution, inherent offset errors at edge sensing zones, randomly occurring invalid pixels, and distance-dependent ranging noise. Therefore, the two sensing modalities exhibit different advantageous operating ranges: ToF sensing is more reliable under near-frontal docking conditions, whereas the camera-based method provides better performance at larger docking angles.

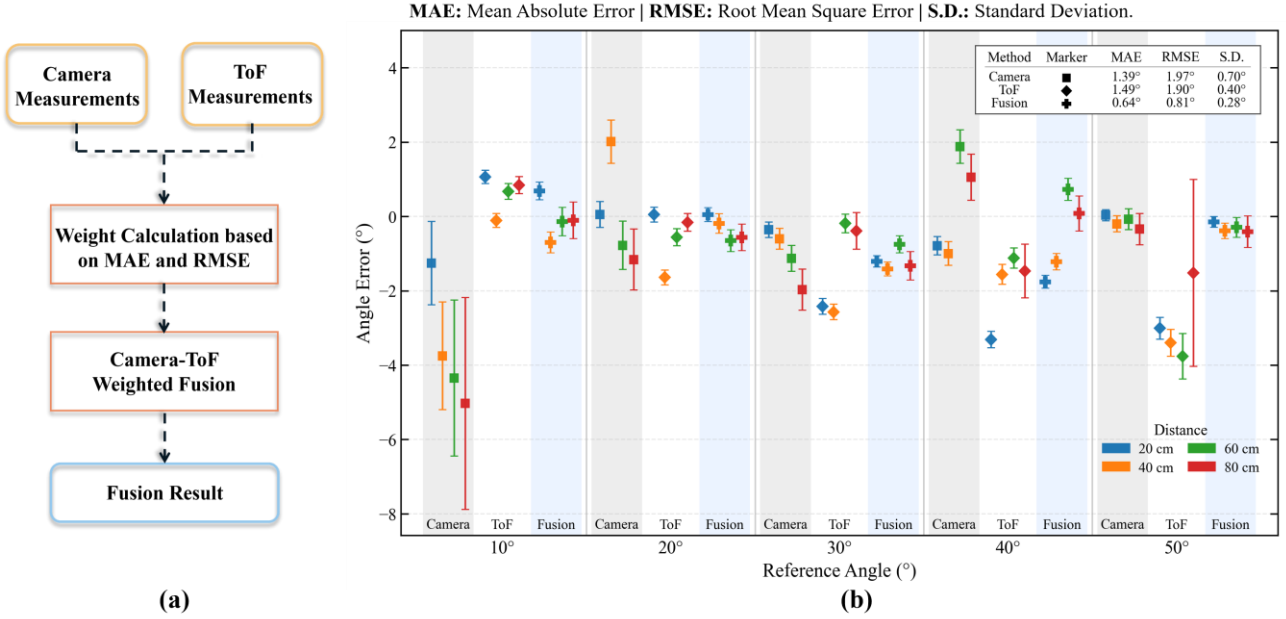


Figure 7. Camera–ToF weighted fusion. **(a)** Data-processing flow; **(b)** Comparison of angle-estimation errors for the Camera, ToF, and Fusion methods. Error bars indicate standard deviations.

To further investigate the fusion potential resulting from this complementarity, a performance-based weighted fusion method was implemented. The corresponding data-processing procedure is illustrated in Figure 7a. Following previous multi-metric evaluation and error-based multi-sensor weighting studies [23–25], a sensing modality with lower estimation error was assigned a larger normalized fusion weight. For each tested reference angle θ_k , the mean absolute error (MAE) and root mean square error (RMSE) of the camera and ToF methods were calculated over the four sensing distances. To account for both the average estimation deviation and relatively large errors, a composite error indicator was defined as

$$Q_s(\theta_k) = \frac{\text{MAE}_s(\theta_k) + \text{RMSE}_s(\theta_k)}{2}, s \in \{c, t\} \quad (3)$$

where the subscripts c and t denote the camera and ToF branches, respectively. The corresponding normalized inverse-error weights were calculated as

$$w_c(\theta_k) = \frac{Q_c^{-1}(\theta_k)}{Q_c^{-1}(\theta_k) + Q_t^{-1}(\theta_k)}, w_t(\theta_k) = \frac{Q_t^{-1}(\theta_k)}{Q_c^{-1}(\theta_k) + Q_t^{-1}(\theta_k)} \quad (4)$$

where $w_c(\theta_k) + w_t(\theta_k) = 1$. Accordingly, the sensing modality with a smaller composite error at a given reference angle receives a larger fusion weight. The resulting Camera/ToF weights were

0.160/0.840, 0.398/0.602, 0.590/0.410, 0.617/0.383, and 0.942/0.058 at reference angles of 10°, 20°, 30°, 40°, and 50°, respectively.

For each sensing distance d and reference angle θ_k , the fused angle was calculated as

$$\hat{\theta}_f(d, \theta_k) = w_c(\theta_k)\hat{\theta}_c(d, \theta_k) + w_t(\theta_k)\hat{\theta}_t(d, \theta_k) \quad (5)$$

where $\hat{\theta}_c$, $\hat{\theta}_t$, and $\hat{\theta}_f$ denote the camera-based, ToF-based, and fused angle estimates, respectively. No additional bias compensation was introduced, ensuring that the fused results were derived directly from the original camera and ToF measurements. Assuming independent random fluctuations between the two sensing modalities, the standard deviation of the fused result was estimated as

$$\sigma_f = \sqrt{[w_c(\theta_k)\sigma_c]^2 + [w_t(\theta_k)\sigma_t]^2} \quad (6)$$

where σ_c and σ_t are the standard deviations of the camera and ToF measurements, respectively.

The comparative results are presented in Figure 7b. Across 20 tested distance–angle configurations, the performance-weighted fusion method achieved an overall MAE of 0.64°, an RMSE of 0.81°, and an estimated mean standard deviation of 0.28°. In comparison, the camera-based method achieved an MAE of 1.39°, an RMSE of 1.97°, and a mean standard deviation of 0.70°, whereas the corresponding values for the ToF-based method were 1.49° for MAE, 1.90° for RMSE, and 0.40° for the mean standard deviation, respectively. As shown in Figure 7b, the errors of fused results generally remain closer to zero and exhibit reduced variations across all the configurations. These results demonstrate that the proposed performance-weighted fusion method effectively converts the complementary characteristics of the two sensing modalities into improved overall angle-estimation accuracy and precision.

The complementary behavior is further reflected under the frontal docking condition at a reference angle of 0°. Representative results at 40 cm show errors of $2.39^\circ \pm 3.15^\circ$ (Error $\pm$ Standard Deviation) and $1.09^\circ \pm 0.22^\circ$ for the camera- and ToF-based methods, respectively, while the corresponding errors at 60 cm are $-2.58^\circ \pm 5.70^\circ$ and $0.95^\circ \pm 0.23^\circ$. Applying the same error-based weighting strategy described above, the camera and ToF weights at 0° were calculated as 0.238 and 0.762, respectively. The resulting fused errors were approximately $1.40^\circ \pm 0.77^\circ$ at 40 cm and $0.11^\circ \pm 1.37^\circ$ at 60 cm. At 60 cm, the fusion reduced the magnitude of the angle error from 2.58° and 0.95° for the individual camera and ToF branches, respectively, to approximately 0.11°, demonstrating the benefit of exploiting their complementary error characteristics under frontal docking conditions.

3.4.2. Partial-obstruction evaluation

To further evaluate the complementary behavior of the camera and multi-zone ToF sensor, a controlled partial-obstruction experiment was conducted using the proposed sensor prototype installed on the miniature USV. As shown in Figure 8a, the target distance and reference angle were fixed at 40 cm and 20°, respectively. A 5 cm $\times$ 9 cm black paper obstacle was placed 20 cm from the sensing system at three positions: in front of the ToF sensor (position A), between the camera and ToF sensor (position B), and in front of the camera (position C). When both sensing modalities provided valid measurements, the weighted-fusion method described in Section 3.4 was applied. When one modality was unavailable, the valid branch was used as the system output.

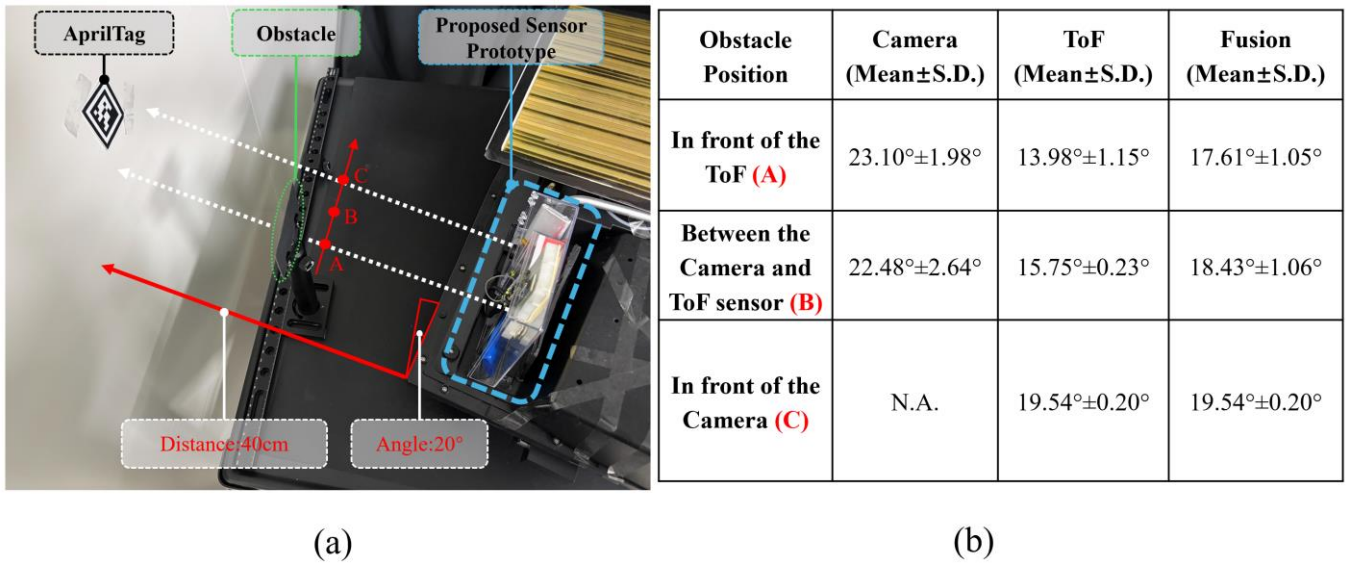


Figure 8. Partial-obstruction test. **(a)** Experimental setup; **(b)** Angle-estimation results.

Figure 8b summarizes the experimental results. At positions A and B, the fused angle estimates were $17.61^{\circ} \pm 1.05^{\circ}$ and $18.43^{\circ} \pm 1.06^{\circ}$, respectively. At position C, direct obstruction of the camera caused the failure of AprilTag detection. However, the ToF sensor still provided a valid estimate of $19.54^{\circ} \pm 0.20^{\circ}$, which was used as the final system output. Overall, the three obstruction cases show the complementary behavior of the two sensing modalities: (1) the camera remained available when the ToF sensing path was locally disturbed; (2) the ToF sensor maintained a valid angle estimate when the camera was directly obstructed; (3) the fused output remained valid when both sensing paths were partially affected. These complementary responses reduce the risk of complete angle-estimation failure caused by a single sensing modality under the tested partial-obstruction conditions.

3.5. Performance comparison with other related work

Table 3 presents a comprehensive summary of angle estimation across a variety of application scenarios using various sensing modalities, including red-green-blue depth (RGB-D) cameras, LiDAR, ultra-wideband (UWB) angle-of-arrival (AoA), global navigation satellite system (GNSS), inertial navigation system (INS) integration, magnetometers, and stereo cameras. The IMU integrated into the Arduino controller can provide low-cost short-term hull attitude measurements with small angular errors and high repeatability, but it requires an initial reference orientation and cannot independently determine the docking angle relative to the berth. Although these angle estimation methods can provide reliable performance in their respective application scenarios, miniature USV docking requires sensing systems with compact structure, low cost and short-range adaptability, which existing solutions cannot fully satisfy. As the core contribution of this work, the proposed sensor prototype integrates a monocular camera with an 8×8 multi-zone ToF sensor to directly estimate the angle relative to the berth target, achieving an optimal balance among angle estimation accuracy, hardware cost and practical deployment capability and making it highly suitable for miniature USV short-range docking tasks.

Table 3. Related work on angle estimation using different sensing modalities.

Reference	Sensing Modality	Application Scenario	Reported Performance	Approximate Cost (USD)
[26]	RGB-D camera	Indoor Orientation Estimation	Absolute rotation error: 0.31° – 0.38° on ICL-NUIM and 1.51° – 2.92° on TUM RGB-D.	NR
[27]	RGB-D camera	Mobile Robot 3D Perception	Average normal deviation: $11.75^{\circ} \pm 3.02^{\circ}$ at VGA, $12.39^{\circ} \pm 3.81^{\circ}$ at QVGA, and $9.32^{\circ} \pm 2.44^{\circ}$ at QQVGA.	NR
[28]	Shipborne 3D LiDAR	Maritime Autonomous Surface Ship Berthing	The overall average relative error is 2.85%, with a calculation frequency of 10 Hz.	~1562–4260
[11]	Shipborne 3D LiDAR	Autonomous Surface Vessel Berthing	Berthing-angle error: $< 0.26^{\circ}$; distance error: < 0.023 m. Robust fitting MAE: 0.194° – 0.257° .	~4400
[29]	Shipborne LiDAR	Marine Target-ship Attitude Estimation	Angle MAE: 0.487° ; angular-rate MAE: $4.2197^{\circ}/s$.	~1325
[30]	LiDAR	Indoor Mobile Robot Docking	Position error: within 1 cm; angle error: within 0.80° .	~240
[31]	LiDAR + reflective marker	Industrial Automated Guided Vehicle Docking	Position accuracy: up to 1 cm; yaw-orientation error: $< 0.05^{\circ}$ in validation; final docking yaw S.D.: $< 1.8^{\circ}$.	~7529
[32]	UWB AoA/PDoA	Indoor Localization	Angular accuracy: approx. 2.4° within the central 90° range; distance precision: centimeter level.	~40
[33]	UWB AoA	Underground Corridor/Tunnel	Azimuth/elevation MAE: $1.00^{\circ}/1.46^{\circ}$ over 5–40 m; overall azimuth MAE: 1.19° over 5–80 m.	NR
[34]	Dual-antenna GNSS + INS + barometer	Ground Vehicle Open/Occluded Environment	Open environment: MAE: 0.4435° , max error: 1.1489° , S.D.: 0.3264° , RMS: 0.5418° . Occluded environment: MAE: 0.5048° , max error: 2.9052° , S.D.: 0.5098° , RMS: 0.6367° .	NR
[35]	Low-cost magnetometer	Small Unmanned Aerial Vehicle Heading Estimation	After calibration, magnetic-heading error: mean 2.02° , S.D.: 1.94° in the third flight test.	NR
[36]	Stereo cameras	Mobile Robot Visual Odometry	Simulated straight route: maximum instantaneous rotation error $< 0.35^{\circ}$ after outlier rejection.	~600
[37]	Stereo cameras	Non-Cooperative Spacecraft Proximity Operations	Simulation: relative rotation-angle error $< 1^{\circ}$; maximum 3D localization error $< 5\%$.	NR
This work	Monocular camera + 8×8 multi-zone ToF	Miniature USV Short-Range Docking	Over 20–80 cm and 10° – 50° , camera branch: MAE: 1.39° , RMSE: 1.97° , mean SD: 0.70° ; ToF branch: MAE: 1.49° , RMSE: 1.90° , mean SD: 0.40° .	93

NR: not reported in the paper.

Approximate costs are representative market prices of the reported sensor hardware.

Abbreviation: RGB-D, Red-Green-Blue Depth camera; LiDAR, Light Detection and Ranging; UWB, Ultra-Wideband; AoA, Angle of Arrival; PDoA, Phase Difference of Arrival; GNSS, Global Navigation Satellite System; INS, Inertial Navigation System; ToF, Time-of-Flight; ICL-NUIM, Imperial College London and National University of Ireland Maynooth dataset; TUM RGB-D, Technical University of Munich RGB-D dataset; VGA, Video Graphics Array; QVGA, Quarter Video Graphics Array; QQVGA, Quarter Quarter Video Graphics Array; MAE, Mean Absolute Error; S.D., Standard Deviation; RMS, Root Mean Square; RMSE, Root Mean Square Error.

4. System integration and field experiments

4.1. System integration

The integrated dual-sensor prototype deployed on the miniature USV is shown in Figure 9a. The camera and multi-zone ToF sensor are mounted on the hull side and encapsulated in transparent waterproof housings for maritime environmental adaptation. The entire sensing pipeline adopts pre-calibrated parameters. In addition, the assembly-induced systematic mounting offsets are corrected via pre-experiment calibrations and data preprocessing. The camera and ToF sensing modalities operated at 8.03 fps and 15.13 Hz on the OpenMV and Arduino platforms, respectively, with estimated processor utilizations of 92.1% and 60.56% and peak heap allocations of 60.1 kB and 3.71 kB. Considering that the final docking stage of a miniature USV involves low-speed and short-range adjustment, the measured update frequencies, including the lower camera update rate of 8.03 fps, can provide sufficiently frequent angle feedback to support successive heading corrections within the present 80 cm effective operating range. For wireless communication shown in Figure 9b, camera image streams are transmitted via an ATWINC1500 Wi-Fi module, while ToF sensor's sampling data are acquired by an Arduino controller board and transmitted via Bluetooth.

4.2. Field experiments

Figure 9c demonstrates docking performance under ambient illumination. The actual USV docking angle is 26° , with camera-based and ToF-based angle estimates of 24.7° and 22.8° , respectively. The relatively consistent results validate cross-modal verification capability under normal ambient light conditions. Figure 9d presents dark-environment docking tests, where insufficient illumination leads to the failure of AprilTag detection and invalid visual perception. In contrast, the ToF sensor operates independently, outputting a valid estimate of 30.2° for an actual docking angle of 33° , verifying reliable active sensing performance using ToF sensor in the dark conditions.

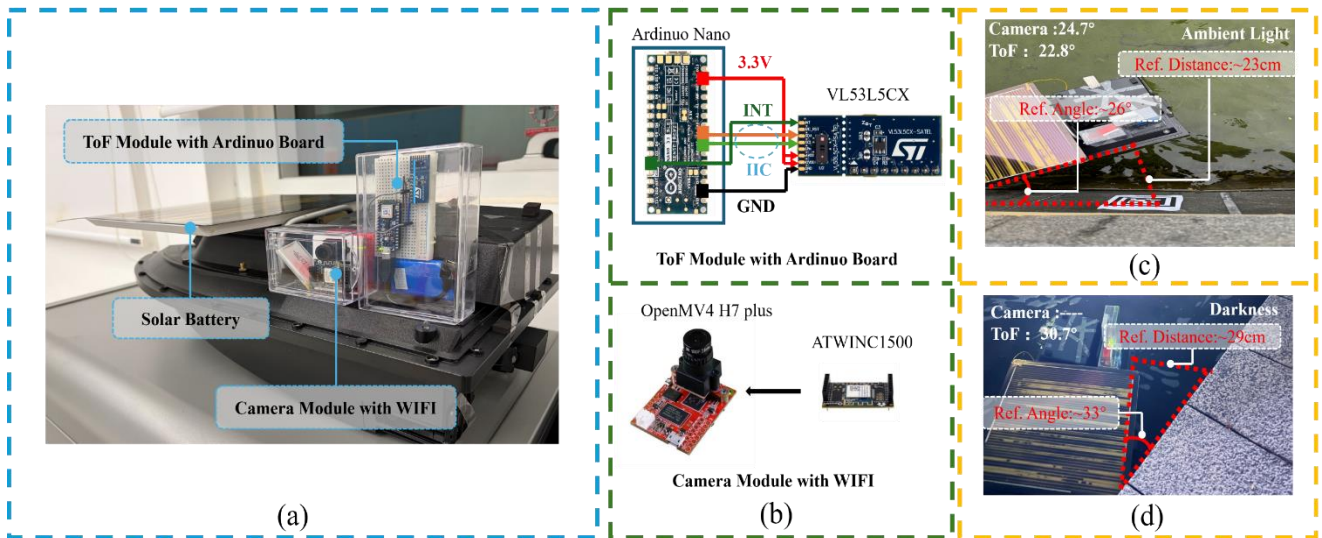


Figure 9. Integrated prototype and representative field results. (a) Side-mounted sensing prototype with transparent waterproof housings for the camera and multi-zone ToF modules; (b) Sensor connections and corresponding microcontroller units (MCUs); (c) Angle estimation results under ambient light; (d) Angle estimation results in darkness (ToF only).

Field experiments validate the complementary strengths and environmental robustness of the proposed camera–ToF sensor fusion system. Under ambient illumination, the cross-verification using the dual sensors improve accuracy and reliability of the angle estimation. Under dark conditions, the active ToF sensing module remains valid perception, realizing illumination-invariant docking angle estimation.

5. Conclusion, limitations and future work

This work presents a low-cost and robust dual-sensing framework that combines a monocular camera with an active multi-zone ToF sensor for short-range docking-angle estimation of miniature USVs. Comprehensive camera intrinsic calibration and systematic ToF sensor characterization are implemented to mitigate inherent hardware biases and guarantee high-quality raw measurement data. Within the established dual-modal sensing scheme, visual orientation angles are calculated based on well-calibrated AprilTag pose solutions, whereas ToF-derived angular results are computed from the preprocessed ranging data after multi-stage validity screening and outlier suppression. The outdoor field experiments fully demonstrate the practicality and environmental robustness of the proposed system. Specifically, the two sensing modalities yield highly consistent and mutually verified angle estimations under normal ambient lighting. Moreover, when visual perception is invalidated in dark environments, the standalone ToF sensor can independently deliver stable and reliable function, enabling reliable docking angle estimation under different illumination scenarios.

The proposed system provides a practical trade-off among hardware cost, compactness, computational demand, and short-range docking performance. However, both sensing modalities are still being affected by hull motion, partial occlusion, mounting errors, and optical distortion introduced by the waterproof housing. Specially, the AprilTag-based method is sensitive to insufficient illumination, motion blur, and corner-localization errors, while the low-cost ToF sensor is constrained by its spatial resolution, invalid ranging zones, target reflectivity, and multipath reflections. In Figure 6, the characterization results indicate a stable effective range of approximately 80 cm and noticeably increased ranging errors at longer distances. These can be improved by optimized optical and waterproof components, post-encapsulation calibration, higher-performance sensors and processors, and enhanced filtering and compensation algorithms. Although this study focuses on yaw-angle estimation using horizontal range gradients, the two-dimensional ToF distance matrix also has the potential to identify angular deviations in other directions using vertical or combined spatial gradients. The proposed system is primarily intended for prepared, short-range docking environments, and reliable estimation cannot be guaranteed when both AprilTag detection and ToF ranging fail simultaneously.

For future work, our efforts will focus on systematic performance optimization and on-site deployment. First, we would carry out post-encapsulation optical recalibration to correct the optical distortion introduced by the waterproof housing. Second, further experimental validation will be conducted under more complex conditions. For example, we will consider interferences from surface-floating obstacles and different berth surface materials. In addition, wave action, rain, fog, water-surface reflections, hull attitude variations, and continuous docking maneuvers will be considered to have a full evaluation of the proposed system under dynamic maritime conditions. Third, the proposed sensing framework will be integrated into the closed-loop docking control system of USV platforms. After that, experiments in larger-scale natural waters and real harbor environments will be performed to further verify the overall performance of the complete autonomous docking system. Finally, we will continue to optimize

hardware configuration, system integration, and software processing to improve its long-term stability, adaptability, and deployment capability in complex environments.

Data availability statement

The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author upon reasonable request.

Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors used generative AI tools only to improve language and readability. Specifically, the authors used ChatGPT for language polishing only for rewriting in limited sections. The authors take full responsibility for the content of the manuscript.

Acknowledgments

This work is supported by Guangdong Provincial Project under Grant 2024QN11X100.

Authors' contribution

Conceptualization, J.H. and W.J.; methodology, J.H., X.Q., J.L. and W.J.; software, J.H.; formal analysis, J.H.; investigation, J.H.; resources, W.J.; data curation, J.H.; writing—original draft preparation, J.H.; writing—review and editing, X.Q., J.L. and W.J.; visualization, J.H.; supervision, W.J.; project administration, W.J.; funding acquisition, W.J. All authors have read and agreed to the published version of the manuscript.

Conflicts of interest

Wei Jiang holds the position of Associate Editor for *Interdiscipline* and has not peer reviewed or made any editorial decisions for this paper. The authors declare that there are no conflicts of interest.

References

- [1] Chen Z, Huang T, Xue Z, Zhu Z, Xu J, *et al.* A novel unmanned surface vehicle with 2D-3D fused perception and obstacle avoidance module. In *Proceedings of the 2021 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Sanya, China, December 27–31, 2021, pp. 1804–1809.
- [2] Kretschmann L, Burmeister HC, Jahn C. Analyzing the economic benefit of unmanned autonomous ships: an exploratory cost-comparison between an autonomous and a conventional bulk carrier. *Res. Transp. Bus. Manag.* 2017, 25:76–86.
- [3] Liu Y, Liu W, Song R, Bucknall R. Predictive navigation of unmanned surface vehicles in a dynamic maritime environment when using the fast marching method. *Int. J. Adapt. Control Signal Process.* 2017, 31(4):464–488.
- [4] Liu Z, Zhang Y, Yu X, Yuan C. Unmanned surface vehicles: an overview of developments and challenges. *Annu. Rev. Control* 2016, 41:71–93.
- [5] Seo S, Jung J, Song J, Kim J, Lee YC. LP-DSG: a LiDAR point-based docking spot generation system for unmanned surface vehicles in berthing environments. *Appl. Sci.* 2025, 15(22):12290.

- [6] Xu C, Chu Y, Gao Q, Wu Z, Wang J, *et al.* Autonomous unmanned surface vehicle docking using large language model guide reinforcement learning. *Ocean Eng.* 2025, 323:120608.
- [7] Cai W, Zhang M, Yang Q, Wang C, Shi J. Long-range UWB positioning-based automatic docking trajectory design for unmanned surface vehicle. *IEEE Trans. Instrum. Meas.* 2023, 72:1–12.
- [8] Shen W, Yang Z, Yang C, Li X. A LiDAR SLAM-assisted fusion positioning method for USVs. *Sensors* 2023, 23(3):1558.
- [9] Lepetit V, Moreno-Noguer F, Fua P. EPnP: an accurate $O(n)$ solution to the PnP problem. *Int. J. Comput. Vis.* 2009, 81(2):155–166.
- [10] Volden Ø, Stahl A, Fossen TI. Vision-based positioning system for auto-docking of unmanned surface vehicles (USVs). *Int. J. Intell. Rob. Appl.* 2022, 6(1):86–103.
- [11] Wang H, Yin Y, Jing Q, Zhang C. Berthing state estimation for autonomous surface vessels using ship-based 3D LiDAR. *J. Mar. Sci. Eng.* 2025, 13(10):1975.
- [12] Wang H, Yin Y, Dong L, Lai H. Integrated LiDAR-based localization and navigable region detection for autonomous berthing of unmanned surface vessels. *J. Mar. Sci. Eng.* 2025, 13(11):2079.
- [13] Niculescu V, Müller H, Ostovar I, Polonelli T, Magno M, *et al.* Towards a multi-pixel time-of-flight indoor navigation system for nano-drone applications. In *Proceedings of 2022 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, Ottawa, Canada, May 16–19, 2022, pp. 1–6.
- [14] Wu G, Li D, Ding H, Shi D, Han B. An overview of developments and challenges for unmanned surface vehicle autonomous berthing. *Complex Intell. Syst.* 2024, 10(1):981–1003.
- [15] Cai J, Chen G, Yin J, Ding C, Suo Y, *et al.* A review of autonomous berthing technology for ships. *J. Mar. Sci. Eng.* 2024, 12(7):1137.
- [16] Pérez J, Nashashibi F, Lefaudeaux B, Resende P, Pollard E. Autonomous docking based on infrared system for electric vehicle charging in urban areas. *Sensors* 2013, 13(2):2645–2663.
- [17] Han Z, Liu L. A 6D object pose estimation algorithm for autonomous docking with improved maximal cliques. *Sensors* 2025, 25(1):283.
- [18] Lion L, Branz F, Caon A, Imperatrice M, Veronese M, *et al.* DOCKS: docking system for microsatellites. In *Proceedings of 75th International Astronautical Congress (IAC)*, Milan, Italy, October 14–18, 2024, pp. 1635–1641.
- [19] Bortotto A, Degli Agli G, Pozzato N, Dignani M, Favotto F, *et al.* Autonomous docking manoeuvre testing in the framework of the ERMES experiment. *Aerotec. Missili Spazio* 2025, 104(6):459–473.
- [20] Olson E. AprilTag: a robust and flexible visual fiducial system. In *Proceedings of 2011 IEEE International Conference on Robotics and Automation (ICRA)*, Shanghai, China, May 9–13, 2011, pp. 3400–3407.
- [21] Wang J, Olson E. AprilTag 2: efficient and robust fiducial detection. In *Proceedings of 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Daejeon, Republic of Korea, October 9–14, 2016, pp. 4193–4198.
- [22] Caroleo G, Albini A, Maiolino P. On the characterisation of the time-of-flight VL53L5CX sensor by STMicroelectronics for indoor robotics applications. *Sensors* 2026, 26(5):1639.
- [23] Zhou J, Lu X, Xiao Y, Tang J, Su J, *et al.* SDWPF: a dataset for spatial dynamic wind power forecasting over a large turbine array. *Sci. Data* 2024, 11(1):649.

- [24] Nemec D, Andel J, Simak V, Hrbcek J. Homogeneous sensor fusion optimization for low-cost inertial sensors. *Sensors* 2023, 23(14):6431.
- [25] Qiu M, Liu B. Multi-sensor adaptive weighted data fusion based on biased estimation. *Sensors* 2024, 24(11):3275.
- [26] Kim P, Coltin B, Kim HJ. Indoor RGB-D compass from a single line and plane. In *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City, USA, June 18–23, 2018, pp. 4673–4680.
- [27] Holz D, Holzer S, Rusu RB, Behnke S. Real-time plane segmentation using RGB-D cameras. In *RoboCup 2011: Robot Soccer World Cup XV*. Berlin: Springer, 2012. pp. 306–317.
- [28] Hu B, Liu X, Jing Q, Lyu H, Yin Y. Estimation of berthing state of maritime autonomous surface ships based on 3D LiDAR. *Ocean Eng.* 2022, 251:111131.
- [29] Wei S, Xiao Y, Yang X, Wang H. Attitude estimation method for target ships based on LiDAR point clouds via an improved RANSAC. *J. Mar. Sci. Eng.* 2023, 11(9):1755.
- [30] Zhao Y, Li X, Lu W, Hu L. A method for autonomous shelf recognition and docking of mobile robots based on 2D LiDAR. In *Proceedings of Artificial Intelligence Technologies and Applications: Proceedings of the 5th International Conference (ICAITA 2023)*, Changchun, China, June 30–July 2, 2024, pp. 104–112.
- [31] Biernacki P, Ziebinski A. The method of reflection-based marker detection and identification to ensure accurate AGV docking. *Sci. Rep.* 2025, 15(1):41343.
- [32] Margiani T, Cortesi S, Keller M, Vogt C, Polonelli T, *et al.* Angle of arrival and centimeter distance estimation on a smart UWB sensor node. *IEEE Trans. Instrum. Meas.* 2023, 72:1–10.
- [33] Li J, Yu B, Cui S, Yang M, Zhao J, *et al.* Multipath credibility selection for robust UWB angle-of-arrival estimation in narrow underground corridors. *Sensors* 2026, 26(6):2002.
- [34] Jiao H, Xu X, Chen S, Guo N, Yu Z. Improving vehicle heading angle accuracy based on dual-antenna GNSS/INS/barometer integration using adaptive Kalman filter. *Sensors* 2024, 24(3):1034.
- [35] Opromolla R. Magnetometer calibration for small unmanned aerial vehicles using cooperative flight data. *Sensors* 2020, 20(2):538.
- [36] Kostavelis I, Boukas E, Nalpantidis L, Gasteratos A. Stereo-based visual odometry for autonomous robot navigation. *Int. J. Adv. Robot. Syst.* 2016, 13(1):21.
- [37] Zhang L, Zhu F, Hao Y, Pan W. Optimization-based non-cooperative spacecraft pose estimation using stereo cameras during proximity operations. *Appl. Opt.* 2017, 56(15):4522–4531.