Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnline CONFERENCE PAPER ARCHIVE

IFAC PapersOnLine 55-31 (2022) 78–84

Vision-based positioning of Unmanned Surface Vehicles using Fiducial Markers for automatic docking

Lars Digerud*, Øystein Volden, Kim A. Christensen*, Sampsa Kohtala*, Martin Steinert***

**Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology, 7491 Trondheim, Norway (e-mail: larsdi@stud.ntnu.no, Kim@fosenregionen.no, sampsa.kohtala@ntnu.no, martin.steinert@ntnu.no)*

*** Department of Engineering Cybernetics, Norwegian University of Science and Technology, 7491 Trondheim, Norway (e-mail: oystein.volden@ntnu.no)*

Abstract: This paper describes a method of using fiducial markers to aid Unmanned Surface Vehicles (USVs) as an additional positioning source during auto-docking. Vision-based techniques allow USVs to localize themselves relative to their surroundings without relying on external communication. This paper shows and evaluates a vision-based strategy to localize USVs to a pier. We used the Global Navigation Satellite System (GNSS) with Real-Time kinematic (RTK) and Inertial Navigation System (INS) with a base station on the pier to validate the vision-based position estimates. The experiment shows that traditional computer vision techniques using fiducial markers can give accurate outdoor position estimates in good conditions. We also highlight some adverse conditions where the performance decreased considerably.

Keywords: Autonomous docking, Fiducial markers, Kalman filter, Object detection, Pose estimation

1. INTRODUCTION

Accurate and reliable positioning is essential for achieving safe auto-docking of Unmanned Surface Vehicles (USV). A common navigation source for USVs is the Global Navigation Satellite System (GNSS), often fused with an onboard Inertial Measurement Unit (IMU) and integrated into an inertial navigation system (INS) to increase positioning accuracy and precision (W. Wang et al., 2019). Real-Time Kinematic (RTK) can be used to increase GNSS accuracy. Combined, RTK-INS can give centimeter precision (Stateczny et al., 2021). Kooij, Colling, and Benson (2018) have projected that the GPS accuracy threshold should be below 0.1 meters during the docking phase before fully autonomous vehicles can become a common reality. Additionally, Det Norske Veritas (DNV, 2018) guidelines have set the required absolute positioning accuracy during auto-docking to be below 0.1 meters, with a 95 % probability for autonomous and remotely operated ships. DNV also addresses that such ships require redundant systems and methods to validate the primary systems during operation and to alert any external operator if the primary system is out of its valid range.

In addition to precise positioning requirements, auto-docking of USVs involves intricate maneuvering and requires efficient motion planning and control algorithms. Current research focuses on fusing INS with proximity sensors such as lidar or ultrasonic to acquire precise positioning estimates that are fed to the motion planning algorithms (Martinsen et al., 2020; W. Wang et al., 2019). Such a setup requires that the dock is recognized, with distinctive features on the pier or in the harbor environment, if the system should give the USV precise and consistent relative position estimates (Esposito & Graves, 2014).

USVs can either dock at a stationary pier or at a floating dock. The latter requires a local reference point to which the USV can localize itself. Additionally, GNSS is an absolute positioning system that requires the exact global coordinates of the target (Volden et al., 2021). Unfortunately, GNSS and RTK-GNSS systems are dependent on external services and can suffer from precision degradation, signal loss in occluded urban or canyon environments, or noise from internal devices such as Wi-Fi modules (Malyuta et al., 2020).

1.1 Related work

Several approaches for vision-based docking of unmanned vehicles have been developed. Malyuta et al. (2020) used AprilTags to make a UAV fly, dock, and charge itself autonomously with high accuracy and precision. Volden et al. 2021) demonstrated a vision-based positioning method using ArUco markers with convolutional neural networks (CNN) that could estimate the position of the USV in a harbor environment. Their approach focused on outdoor performance and demonstrated that vision-based methods could estimate the USV position up to approximately 15 m for mono- and stereo camera configurations. Mateos et al. (2019) used fiducial markers to navigate several USVs relative to each other to latch them to a docking station or each other. As such, (Mateos, 2020) proposes an upgraded AprilTag3D framework of two coupled non-parallel fiducial markers to increase vision-based positioning performance with promising results.

1.2 Main contributions

This paper demonstrates how a single camera can aid USVs in obtaining precise relative position estimates of a pier or a potential floating docking station, using a high precision RTK-INS for validation. The long-term objective is to develop an

10.1016/j.ifacol.2022.10.412

independent vision-based positioning system to increase the redundancy and accuracy of autonomous vehicles' navigation systems during the docking phase. The method uses only computer vision and filtering techniques for computational efficiency. The paper also addresses some adverse weather conditions that a fully developed camera system must overcome. A Kalman filter is added to remove outliers and increase positioning accuracy.

2. DESIGN, ALGORITHMS, AND IMPLEMENTATION

The vision-based system uses a single camera to feed raw images to an AprilTag2 algorithm (Wang and Olson, 2016), where perspective-n-point (PnP) estimates the relative position of the USV to a marker on the pier. A Kalman filter is applied to improve accuracy by removing outliers. An overview of the system is shown in Fig. 1.

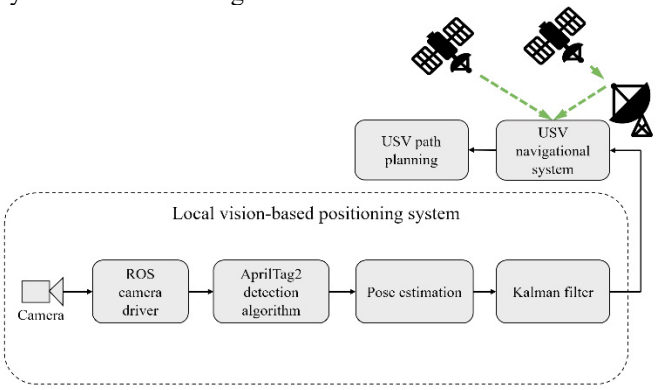

Fig. 1 Block diagram describing our vision-based positioning system. A camera feeds raw images to the AprilTag2 algorithm, where the camera pose relative to the markers is estimated by a PnP solution using a Kalman filter to remove outliers and increase accuracy.

2.1 Fiducial markers

Fiducial markers are commonly used in computer vision applications to detect and localize reference points in a physical space. The markers can be used to estimate the 6-DoF camera-tag pose. As such, fiducial markers can be helpful in dynamic and unstructured environments to obtain global pose estimates.

A fiducial marker system usually consists of a detection algorithm and a coding system. Each type consists of unique patterns which constitute the marker ID. The detection algorithms are often based on traditional image processing techniques such as edge detection, blob detection, and image binarization. For instance, ARTag (Fiala, 2004), ArUco (Garrido-Jurado et al., 2014), and AprilTag (J. Wang & Olson, 2016b) use black-and-white cells in a checker-board system, whose quadrilateral boundary is detected by analyzing the lines. The markers' 6-DoF pose is then estimated using PnP with 2D-3D marker corner correspondences, i.e., the image points and their corresponding 3D projections.

Under certain conditions, the estimated camera pose from a single marker may be subjected to be *ambiguous,* where there can be two unique solutions that can align the correspondences. Resulting in the tag being projected in two different orientations and consequently estimating two different camera locations, which can be observed as the marker position being flipped over an axis. In general, this is not a problem as long as the marker is sufficiently projected or multiple tags are used (Collins & Bartoli, 2014). The accuracy of such markers can be down to a centimeter depending on environmental lighting, camera resolution, distance to marker, and the marker size (Abbas et al., 2019).

2.2 Pose estimation

The pose-estimation algorithm computes a 3×3 homography matrix that projects 2D points to 3D coordinates using a Direct Linear Transform (DLT) algorithm (Hartley & Zisserman, 2004) to compute the marker position and orientation. DLT requires the camera focal length and the physical size of the marker. The homography matrix calculated from the DLT algorithm can be written as a product of the camera matrix, $\boldsymbol{K}$, and the $\boldsymbol{R|t}$, 3×4 joint rotation matrix, which is called a matrix of extrinsic parameters and is used to describe the camera motion around a static scene, or vice versa, the rigid motion of an object in front of a camera. The camera pose can be estimated with the following equation as defined by (Hartley & Zisserman, 2004),

$$s\boldsymbol{m}' = \boldsymbol{K}[\boldsymbol{R}|\boldsymbol{t}]\boldsymbol{M}'. \tag{1}$$

Where s is a projective transformation scale factor, $\boldsymbol{m}'$ is the 2D marker pixel coordinates on the camera sensor, $\boldsymbol{M}'$ is the 3D local coordinates of the marker in the camera frame. Fully extended, (1) can be written as:

$$s\begin{bmatrix}\boldsymbol{u}_x\\ \boldsymbol{v}_y\\ 1\end{bmatrix} = \begin{bmatrix}f_x & 0 & c_x\\ 0 & f_y & c_y\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}r_{11} & r_{12} & r_{13} & t_x\\ r_{21} & r_{22} & r_{23} & t_y\\ r_{31} & r_{32} & r_{33} & t_z\end{bmatrix}\begin{bmatrix}X\\ Y\\ Z\\ 1\end{bmatrix}. \tag{2}$$

Here, $(\boldsymbol{u}_x, \boldsymbol{v}_y)$ is the image point, (f_x, f_y) are the camera focal lengths and (c_x, c_y) are the principal point. The global position of the camera can then be obtained by calculating the inverse of the $\boldsymbol{R|t}$ matrix,

$$\begin{bmatrix}x\\ y\\ z\end{bmatrix} = [\boldsymbol{R}|\boldsymbol{t}]\begin{bmatrix}X\\ Y\\ Z\end{bmatrix}. \tag{3}$$

Where x, y, and z gives the global camera position relative to the markers' X, Y and Z local position in the camera frame. The Euclidean distance between the camera and the marker is then obtained with,

$$|d_{Euclidean}| = \sqrt{x^2 + y^2 + z^2}. \tag{4}$$

2.3 Kalman Filter

Kalman filters are commonly used in navigation to obtain accurate and robust estimates from a set of independent position sources (e.g., GNSS and IMU). The filter was kinematic modeled as a constant velocity model with the following equations as defined by Kim and Bang (2019):

$$\boldsymbol{x}_k = \boldsymbol{F}\boldsymbol{x}_{x-1} + \boldsymbol{B}\boldsymbol{u}_{k-1} + \boldsymbol{w}_{k-1}, \tag{5}$$

Here, $\boldsymbol{x}_k$ is the marker 3D positions, $[x, y, z]^T$. $\boldsymbol{F}$ is defined as the state transition matrix applied to the previous state vector $\boldsymbol{x}_{k-1}$, as

$$\boldsymbol{F} = \begin{bmatrix} 1 & dt & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & dt & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & dt \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \tag{6}$$

With dt set as the time between the frames and $\boldsymbol{B}$ as the control input matrix, and $\boldsymbol{w}_{k-1}$ as the process noise vector. Note that the $\boldsymbol{u}_{k-1}$ term in (5) is assumed zero, as no control inputs were included to improve the vision system performance. The process model is paired with the measurement model that describes the relationship between the state and the measurements at the current time step **k** as,

$$\boldsymbol{z}_k = \boldsymbol{H}\boldsymbol{x}_k + \mathbf{v}_k, \tag{7}$$

where $\boldsymbol{z}_k$ is the measurement vector and **H** is the measurement matrix, $\boldsymbol{I}_{3\times6}$, defined as,

$$\boldsymbol{H} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \tag{8}$$

with the covariance process matrix **Q**. $\mathbf{v}_k$ is the measurement noise vector $\sim \mathcal{N}(0, \mathbf{R})$ with the measurement covariance matrix **R**. The estimation is accomplished in two steps: prediction $(-)$ and update $(+)$. The hat operator $\widehat{\ }$ represents an estimate of a variable. The prediction state estimate and predicted error covariance are subsequently defined,

$$\hat{\boldsymbol{x}}_k^- = \boldsymbol{F}\hat{\boldsymbol{x}}_{k-1}^+, \tag{9}$$

$$\boldsymbol{P}_k^- = \boldsymbol{F}\boldsymbol{P}_{k-1}^+\boldsymbol{F}^T + \boldsymbol{Q}. \tag{10}$$

The update stages are then described as,

$$\tilde{\boldsymbol{y}}_k = \boldsymbol{z}_k - \boldsymbol{H}\hat{\boldsymbol{x}}_k^-, \tag{11}$$

$$\boldsymbol{K}_k = \boldsymbol{P}_k^-\boldsymbol{H}^T(\boldsymbol{R} + \boldsymbol{H}\boldsymbol{P}_k^-\boldsymbol{H}^T)^{-1}, \tag{12}$$

$$\hat{\boldsymbol{x}}_k^+ = \hat{\boldsymbol{x}}_k^- + \boldsymbol{K}_k\tilde{\boldsymbol{y}}, \tag{13}$$

$$\boldsymbol{P}_k^+ = (\boldsymbol{I} - \boldsymbol{K}_k\boldsymbol{H})\boldsymbol{P}_k^-, \tag{14}$$

where (11)-(14) are defined as the measurement residual, Kalman gain, updated state estimate, and updated error covariance. The filter estimates the current measurement by multiplying the predicted state by the measurement matrix. $\tilde{\boldsymbol{y}}_k$ is then multiplied by the Kalman gain, $\boldsymbol{K}_k$, to provide the correction $\boldsymbol{K}_k\tilde{\boldsymbol{y}}$ to the predicted estimate $\hat{\boldsymbol{x}}_k^-$. Once the updated state estimate has been calculated, the error covariance matrix, $\boldsymbol{P}_k^+$ is calculated for the next time step. Finally, the Kalman filter requires an initial estimate of $\boldsymbol{x}_0$ and an initial guess of the error covariance matrix **P** to estimate the next state $\boldsymbol{x}_k$ at timestep k.

3. EXPERIMENTAL SETUP AND TEST SCENARIOS

Our experimental setup consists of an Otter USV developed by Maritime Robotics (Maritime Robotics, 2022) in collaboration with the Norwegian University of Science and Technology (NTNU) as an experimental test platform to conduct sea trials. The USV is equipped with two GNSS receivers in bow and stern, integrated into an INS system that receives RTK corrections from an onshore base station, as shown in Fig. 2. The RTK base station was configured to estimate the phase of the GNSS carrier wave over 20 h before the experiments were conducted—resulting in an absolute GNSS accuracy of 10.48 cm during the experiments.

A wireless radio was used for communication between the USV and the onshore base station. Specific USV and camera specifications can be seen in Table 1 and Table 2, respectively. All actions of the USV during the experiment were remote-controlled, and all data were recorded with the Robotic Operating System (ROS, 2022) on the same computer onboard the Otter USV for post-processing. To evaluate the accuracy of the vision-based system, we compared the estimated positions to RTK INS positions at each timestep using Root Mean Square Error (RMSE). Both systems were post-synchronized, where RTK INS positions were sampled at 25 Hz and camera frames at 15 Hz.

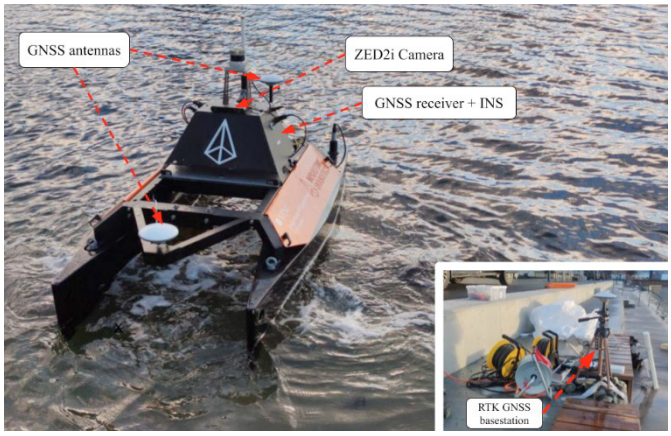

Fig. 2 The fully equipped Otter USV and the onshore RTK GNSS base station at the pier.

Table 1. Otter USV specifications

Dimensions	$2\ m\ x\ 1.08\ m$
Position and heading reference system	Two GNSS receivers in bow and stern
INS	SBG Ellipse 2-D
Sample rate	$25\ Hz$

Table 2. Camera specifications

Model name	ZED 2i
Pixel format	RGB
Resolution	2208×1242
Sample rate	$15\ Hz$
Field of View	120°

The AprilTag markers' physical size was $0.412\ m \times 0.412\ m$ of marker family *36h11*, located on the pier, as seen in Fig. 3. In order to compare the estimated camera positions to RTK INS positions, the North-East-Down (NED) coordinate system was used, denoted as $\eta = [N, E, D]^T$, with the first dimension

pointing North, the second pointing East, and the last pointing towards the earth's center. For convenience, we used the coordinate origin of the NED frame to correspond to the midpoint of the reference tag, m_1, as seen in Fig. 3. As such, we can transform the GNSS measurements into NED coordinates and compare them directly with the tag measurements.

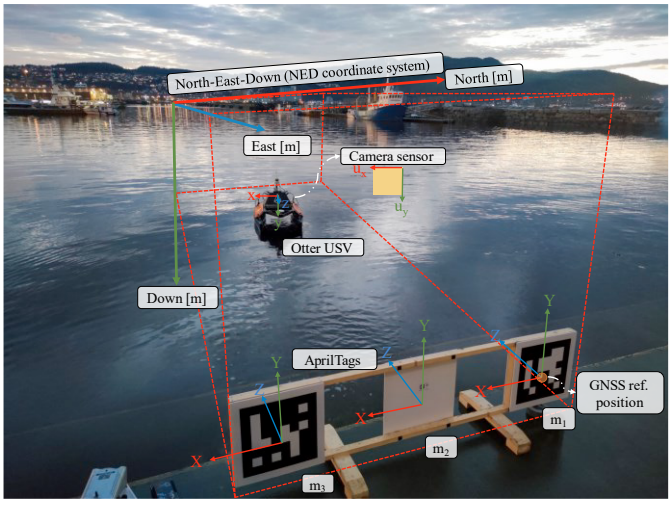

Fig. 3 Experiment setup with the Otter USV and the fiducial markers, m_1, m_2, and m_3, on the pier. Only positions from the tag, m_1, were used in this experiment to estimate the USV position.

3.1 Environmental effects on AprilTag detection

Three scenarios were analyzed to test how the camera system performed in good and adverse weather conditions. The scenarios can be seen in Fig. 4, showing the following conditions: a) optimal light conditions, b) mirrored light from the sun, and c) the harbor in darkness from the Otter USV perspective.

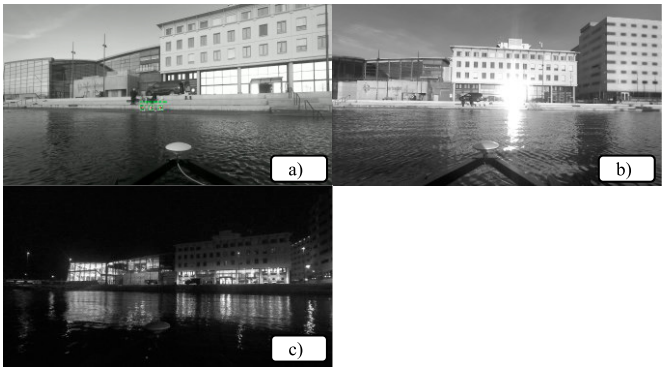

Fig. 4 Camera observations from the Otter USV: a) optimal light conditions, b) mirrored light from the sun, and c) a harbor in darkness.

4. RESULTS AND DISCUSSION

This section presents and discusses the positional accuracy of the vision-based system for each of the three scenarios.

4.1 Scenario 1: Optimal conditions

Fig. 5 shows the top-down view of the Otter approaching the marker in NED coordinates during optimal light conditions. The yellow points represent the Otters' position estimates by the AprilTag system, while the Kalman filtered path can be seen as the blue dotted line and the RTK INS path as the green dotted line. A few erroneous position estimates can be seen in the upper left corner of Fig. 5, which are likely caused by position subjected to ambiguity (marker position predicted incorrectly). However, the figure shows that the Kalman filter can effectively remove these outliers.

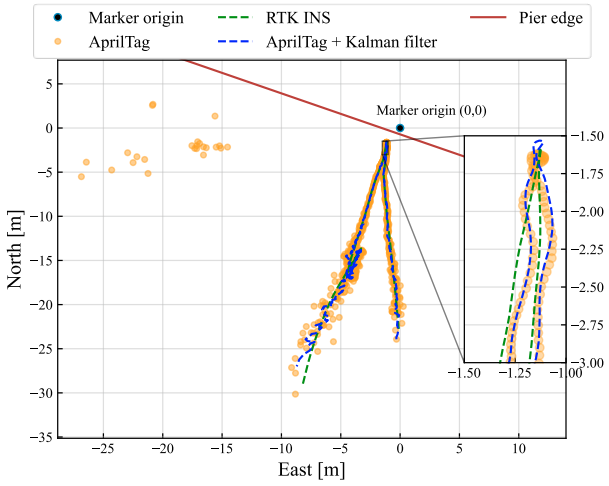

Fig. 5 USV RTK INS path versus raw and Kalman filtered position estimates estimated by the vision system.

The RTK INS and the estimated USV positions during the experiment are shown in Fig. 6, including the Euclidean distance to the marker (a) and the distance along the East, North, and Down axis (b-d). The corresponding RMSE values can be seen to the right in Fig. 6, which are calculated for every one-second interval.

Although some position estimates are flipped symmetrically in the East axis, the corresponding RMSE values show that the Kalman filter effectively removes these outliers and increases the positioning accuracy. In general, the AprilTag and the Kalman filtered positions seem to have a high positive covariance, indicating that the Kalman filter may not significantly increase the positioning accuracy.

Fig. 6 shows overall good performance of the vision-based position estimates, both for the Euclidean position and around the three axes East, North and Down. Additionally, better positioning accuracy can be observed when the USV is close to the marker, which is a critical part of the docking phase. The marker was detected up to a range of 30 m in Euclidean distance, with a corresponding RMSE value of around 1 m at 96 s, as seen in Fig. 6 (a) and (a').

Following the 0.1-meter guideline set by (DNV, 2018) and the GNSS threshold by (Kooij et al., 2018), the RMSE value is below the 0.1 m criteria between 50 to 65 s when the marker is within a range of 5 m to the marker (a')-(d'). Additionally, the results along the Down axis (d') show that the RMSE is below 0.01 m at 56 s and when the USV is within a Euclidean range of 2 m, both for the unfiltered and the Kalman filtered position estimates which indicate high accuracy along this axis.

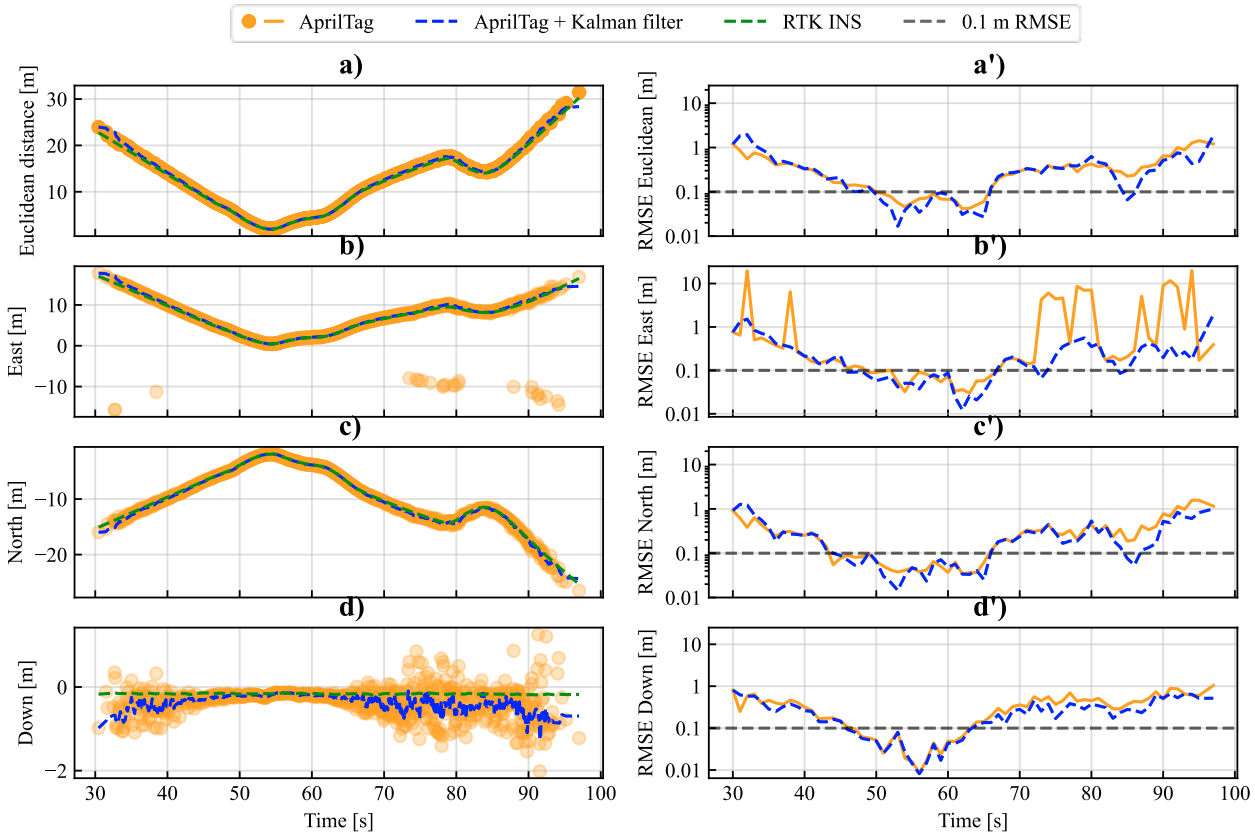

Fig. 6 Positions along Euclidean a), East b), North c), and Down d) versus time. Its corresponding RMSE value can be seen to the right, (a')-(d'), respectively.

4.2 Scenario 2: Mirrored light from the sun

Fig. 7 shows the results of testing with mirrored light from the sun. The figure shows that the camera system estimates contain many outliers compared to the RTK INS and scenario 1. Additionally, the Kalman filter seems to have started with erroneous initial conditions, which may have caused incorrect predictions. However, the Kalman filter accurately predicts the USV position when the distance to the marker is low.

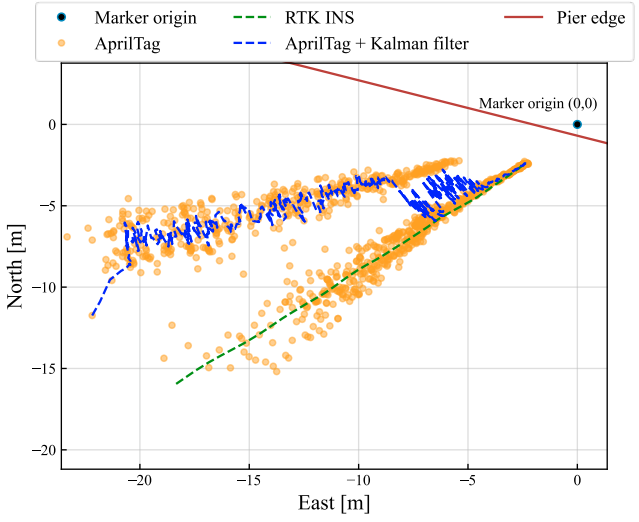

Fig. 7 USV RTK GNSS path with the vision system estimated positions during the mirrored light from the sun scenario.

4.3 Scenario 3: A harbor in darkness

Fig. 8 shows the position estimated when the USV approached the pier at night. The figure shows that the camera system got considerably fewer detections compared to scenarios 1 and 2. In addition, a large portion of the estimates has a significant deviation from the RTK INS, which also might be subjected to ambiguity.

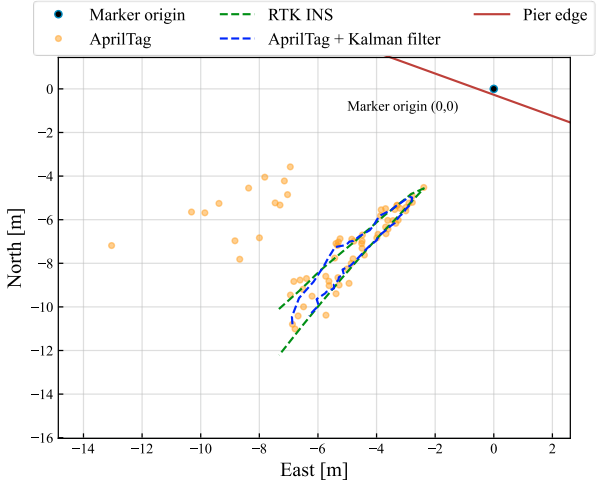

Fig. 8 USV RTK GNSS path with estimated positions by the vision system during the harbor in darkness scenario.

4.4 Discussion

All the results were estimated from one single marker using the AprilTag2 algorithm by Wang and Olson (2016) and validated by RTK INS. In scenario 1, the vision-based system successfully created a local reference point on the pier to estimate the position of the USV. The method accurately estimated the USV position at a range of up to 5 meters, with an RMSE below 0.1 m, demonstrating that the vision-based system can provide sufficient positioning accuracy as an independent positioning system. Furthermore, despite a few outliers, the method is robust in optimal weather conditions. However, the results revealed that the performance degraded significantly in more adverse environmental conditions (e.g., mirrored light from the sun and in a harbor in darkness), most likely due to lighting issues. Using a high dynamic range camera, adjusting camera exposure, and using illuminated backlight markers could increase the detection rate and accuracy in scenarios 2 and 3. Additionally, using more advanced techniques (e.g., CNN) could further increase the detection rate, as found by Volden, Stahl and Fossen (2021).

Estimating the USV position from a marker given a set of four 2D-3D correspondences is the core methodology of our vision-based system. Configuring multiple markers in a known setup may improve positioning accuracy and ensure the vision-based system integrity, as Malyuta et al. (2020) demonstrated by using several AprilTags in a multi-tag configuration to estimate a UAV position more robustly. This approach can be supported by research from Collins and Bartoli (2014), which points out that ambiguity can be solved if certain pre-determined conditions are met. Mateos (2020) suggests a setup of two coupled non-planar markers with promising results (achieved 95% vs. 60% detection rate in outdoor experiments and an improved positioning accuracy). However, such a setup will be ambiguous if only one single marker is detected, as the results from the experiment have shown.

Lastly, the proposed vision-based method has limitations as an independent positioning source because the method requires that the marker is in the camera's field of view and can be identified to provide positioning information. Applying a Kalman filter on the position estimates acquired from the camera system works in optimal conditions. Still, results from the field experiment have demonstrated that ambiguity could cause significantly erroneous position estimates. For practical applications, the Kalman filter can be initialized with GNSS INS positions to improve the vision-based positioning system. To avoid sudden jumps in position, it is also essential to check that the difference in the estimated position between the camera system and the INS is minimal.

5. CONCLUSION

This paper demonstrates how vision-based positioning can be utilized in harbors as an additional positioning source by creating a local reference point between the USV and a pier using a single camera. In terms of positioning accuracy, the results have shown that efficient computer vision algorithms can estimate the position of the USV with reasonable accuracy in optimal weather conditions. However, a vision-based system is limited in adverse conditions, including suboptimal light conditions and darkness. Such scenarios must be further addressed before a vision-based system can serve as an independent positioning system. Lastly, using cameras with a larger dynamic range and greater resolution could increase the vision-based system performance, which, combined with multiple markers in a multi-tag configuration, will form a more robust vision-based positioning system.

6. ACKNOWLEDGEMENTS

We are grateful to the NTNU Center of Autonomous Marine Operations and Systems (AMOS) at the Norwegian University of Science and Technology, who let us use one of their USVs to conduct experimental tests. They have also been accommodating with support and help when preparing the core functionality for the field experiments onboard the Otter USV.

REFERENCES

Abbas, S. M., Aslam, S., Berns, K., & Muhammad, A. (2019). Analysis and Improvements in AprilTag Based State Estimation. *Sensors*, *19*(24), 5480. https://doi.org/10.3390/s19245480

Collins, T., & Bartoli, A. (2014). Infinitesimal Plane-Based Pose Estimation. *International Journal of Computer Vision*, *109*(3), 252–286. https://doi.org/10.1007/s11263-014-0725-5

DNV. (2018). *Class guideline: Autonomous and remotely operated ships*. DNV. https://www.dnv.com/maritime/autonomous-remotely-operated-ships/ class-guideline.html

Esposito, J. M., & Graves, M. (2014). An algorithm to identify docking locations for autonomous surface vessels from 3-D LiDAR scans. *2014 IEEE International Conference on Technologies for Practical Robot Applications (TePRA)*, 1–6. https://doi.org/10.1109/TePRA.2014.6869160

Fiala, M. (2004). *ARTag, An Improved Marker System Based on ARToolkit*. National Research Council Canada. https://doi.org/10.4224/5763247

Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F. J., & Marín-Jiménez, M. J. (2014). Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition*, *47*(6), 2280–2292. https://doi.org/10.1016/j.patcog.2014.01.005

Hartley, R., & Zisserman, A. (2004). *Multiple View Geometry in Computer Vision* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511811685

Kim, Y., & Bang, H. (2019). Introduction to Kalman Filter and Its Applications. In F. Govaers (Ed.), *Introduction and Implementations of the Kalman Filter*. IntechOpen. https://doi.org/10.5772/intechopen.80600

Kooij, C., Colling, A. P., & Benson, C. L. (2018, October 2). *When will autonomous ships arrive? A technological forecasting perspective*. 14th International Naval Engineering Conference and Exhibition, Glasgow, UK. https://doi.org/10.24868/issn.2515-818X.2018.016

Malyuta, D., Brommer, C., Hentzen, D., Stastny, T., Siegwart, R., & Brockers, R. (2020). Long-duration fully autonomous operation of rotorcraft unmanned aerial systems for remote-sensing data acquisition. *Journal of Field Robotics*, *37*(1), 137–157. https://doi.org/10.1002/rob.21898

Maritime Robotics. (2022, May 2). *Maritime Robotics*. Otter Pro. https://www.maritimerobotics.com/otter

Martinsen, A. B., Bitar, G., Lekkas, A. M., & Gros, S. (2020). Optimization-Based Automatic Docking and Berthing of ASVs Using Exteroceptive Sensors: Theory and Experiments. *IEEE Access*, *8*, 204974–204986. https://doi.org/10.1109/ACCESS.2020.3037171

Mateos, L. A. (2020). AprilTags 3D: Dynamic Fiducial Markers for Robust Pose Estimation in Highly Reflective Environments and Indirect Communication in Swarm Robotics. *ArXiv:2001.08622 [Cs]*. http://arxiv.org/abs/2001.08622

Mateos, L. A., Wang, W., Gheneti, B., Duarte, F., Ratti, C., & Rus, D. (2019). Autonomous Latching System for Robotic Boats. *2019 International Conference on Robotics and Automation (ICRA)*, 7933–7939. https://doi.org/10.1109/ICRA.2019.8793525

ROS. (2022). *Robot Operating System (ROS)* (Version 1) [Melodic]. https://www.ros.org/

Stateczny, A., Specht, C., Specht, M., Brčić, D., Jugović, A., Widźgowski, S., Wiśniewska, M., & Lewicka, O. (2021). Study on the Positioning Accuracy of GNSS/INS Systems Supported by DGPS and RTK Receivers for Hydrographic Surveys. *Energies*, *14*(21), 7413. https://doi.org/10.3390/en14217413

Volden, Ø., Stahl, A., & Fossen, T. I. (2021). Vision-based positioning system for auto-docking of unmanned surface vehicles (USVs). *International Journal of Intelligent Robotics and Applications*. https://doi.org/10.1007/s41315-021-00193-0

Wang, J., & Olson, E. (2016a). AprilTag 2: Efficient and robust fiducial detection. *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 4193–4198. https://doi.org/10.1109/IROS.2016.7759617

Wang, J., & Olson, E. (2016b). AprilTag 2: Efficient and robust fiducial detection. *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 4193–4198. https://doi.org/10.1109/IROS.2016.7759617

Wang, W., Gheneti, B., Mateos, L. A., Duarte, F., Ratti, C., & Rus, D. (2019). Roboat: An Autonomous Surface Vehicle for Urban Waterways. *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 6340–6347. https://doi.org/10.1109/IROS40897.2019.8968131