Improved Tag-based Indoor Localization of UAVs Using Extended Kalman Filter

N. Kayhani[a,b], A. Heins[b], W.D. Zhao[b], M. Nahangi[a,b], B. McCabe[a], and A.P. Schoellig[b]

[a] Department of Civil and Mineral Engineering, University of Toronto, Canada.
[b] University of Toronto Institute for Aerospace Studies (UTIAS), Canada.
E-mail: navid.kayhani@mail.utoronto.ca, adam.heins@mail.utoronto.ca, wenda.zhao@mail.utoronto.ca, m.nahangi@utoronto.ca, brenda.mccabe@utoronto.ca, schoellig@utias.utoronto.ca

Abstract –

Indoor localization and navigation of unmanned aerial vehicles (UAVs) is a critical function for autonomous flight and automated visual inspection of construction elements in continuously changing construction environments. The key challenge for indoor localization and navigation is that the global positioning system (GPS) signal is not sufficiently reliable for state estimation. Having used the AprilTag markers for indoor localization, we showed a proof-of-concept that a camera-equipped UAV can be localized in a GPS-denied environment; however, the accuracy of the localization was inadequate in some situations. This study presents the implementation and performance assessment of an Extended Kalman Filter (EKF) for improving the estimation process of a previously developed indoor localization framework using AprilTag markers. An experimental set up is used to assess the performance of the updated estimation process in comparison to the previous state estimation method and the ground truth data. Results show that the state estimation and indoor localization are improved substantially using the EKF. To have a more robust estimation, we extract and fuse data from multiple tags. The framework can now be tested in real-world environments given that our continuous localization is sufficiently robust and reliable.

Keywords –

Unmanned aerial vehicles (UAV); Building Information model (BIM); Indoor navigation; Autonomous flight; Visual inspection; Construction automation; Extended Kalman filter (EKF)

1 Introduction

Due to their capabilities, unmanned aerial vehicles (UAVs) are increasingly being employed for various purposes and in many industries, such as the construction industry. UAVs are able to function in unsafe and inaccessible locales that may be hazardous for humans. In addition, they provide maneuverability and perspective advantages over ground robotic platforms, which is an asset, particularly in multi-story buildings. A UAV can be equipped with various sensors to further enhance the data and insight acquisition. Thus, these features can enhance the collection of informative data in congested and dynamically changing construction environments. For instance, the inspection and monitoring of under-construction buildings [1], progress tracking and assessment [2] [3], site surveying [4], quality control [5], civil infrastructure condition assessment [6][7], and safety inspections [8] are among the broad applications of UAVs in the construction industry.

UAVs can be programmed such that they perform their tasks autonomously. To accomplish autonomous missions, the UAV must be able to navigate a collision-free path and be aware of its pose in the operating environment at any point in time. Autonomous robots in outdoor environments may take advantage of robust external sources of localization such as the Global Positioning System (GPS). However, GPS connectivity in highly congested urban areas with a dense distribution of high-rises (i.e., urban canyons) may be challenging even outdoors. Similarly, localization remains a real challenge in indoor environments.

1.1 Indoor Localization

A robust indoor localization system aims to continuously estimate the pose of one or more mobile agents in an indoor environment in real-time, given that the environment's map is available. Such a system has to not only be accurate, but also be financially, energy, space, and time efficient. It also needs to suitably cover the area. Several methods have been investigated to address this problem in the past decade. These state-of-the-art studies have mostly relied on localization systems based on radio-frequency identification (RFID) [9], ultra-wideband (UWB) [10][11], wireless local area networks (WLAN) [12], inertial measurement units (IMU) [13],

and machine vision [14]. In the construction sector particularly, RFID [15][16][17], UWB [18], IMU [19][20], and machine vision [21][22] are among the most commonly employed types of indoor localization systems.

However, not all the aforementioned systems may be suitable in the case of autonomous indoor flight during construction. For instance, due to the limited UWB signal communication range, dense network infrastructure may be needed. Metallic materials, such as piping and wall framing, may interfere with the radio frequency based systems including UWB [23] and RFID [24]. Non-line-of-sight (NLOS) propagation conditions may induce error in UWB [10], while IMU-based estimates may drift [19]. The most important concern in this regard, however, is insufficient localization accuracy of these systems [20] for autonomous UAV navigation. Although machine-vision-based localization methods such as motion tracking systems (e.g., Vicon systems) provide accurate estimates, they are an impractical solution for this application because they impose extra costs, logistics, complexity, and distraction to the construction site.

1.2 Visual Fiducial Markers

Visual fiducial markers, such as ARTag [25], AprilTag [26][27], and CalTag [28], are artificial landmarks consisting of patterns. They are designed to be easily recognized and robustly distinguished from one another and among other features in a natural scene when they are placed in an environment. Fiducial markers have been frequently used in augmented reality applications for camera pose estimation (e.g., ARTag [25]). AprilTag markers are passive square-shaped payload tags with an external black border and a binary internal code. Unlike QR (quick response) codes, AprilTag markers contain a small information payload. Therefore, they can be quickly detected and localized even when they are rotated or located in different lighting conditions. Indeed, AprilTag markers are two-dimensional barcodes that provide 6-DOF camera position estimation, and are cost-effective in that they can be printed on a standard printer [26]. AprilTag markers have also been found to be robust for indoor localization and object tracking in various mobile robotics applications [27]. AprilTag has different marker family types. The difference between the types are twofold: (1) the number of bits and (2) the minimum Hamming distance. The Hamming distance between two tags is defined as the number of positions at which the corresponding bits are different. For instance, an AprilTag of "Tagn^2hm" refers to an $n \times n$ array (n²-bit) marker with a minimum Hamming distance of m between any two markers.

1.3 Extended Kalman Filter (EKF)

Since the position estimates and measurements are imperfect and prone to error, prediction and noise reduction algorithms are critical to enhancing these estimates. By minimizing the mean of the squared error, the Kalman filter [29] estimates the past, present, and even future state of a process using an efficient recursive approach [30]. However, since the Kalman filter (KF) is limited to linear systems, it is not able to solve non-linear localization problems. The EKF linearizes non-linear systems so that it can address non-linear localization problems [30][31].

1.4 Summary

In short, this paper presents the implementation of an EKF for improving the estimation module of the previously developed GPS-denied indoor localization framework using AprilTag markers [22]. In this study, it is shown that a probabilistic approach to the data fusion of multiple markers and the IMU using an EKF can substantially improve the estimation accuracy. Given that the 3D coordinates of the fiducial markers are also identified in the building information model (BIM), a camera-equipped UAV can recognize its pose relative to these artificial fiducial markers and then calculate its global location.

2 Overview of Prior Work

To develop a system for indoor localization of UAVs equipped with an onboard camera, the authors previously proposed a framework [22] using AprilTag markers. As illustrated in Figure 1, the calibration of the on-board camera is first undertaken to ensure that the measurements are accurate. The Robotics Operating System (ROS) package for camera calibration is used to estimate the camera parameters. Next, based on the mission plans and the critical locations that are supposed to be monitored by the UAV, the number and the corresponding location of tags are identified using the BIM model. Tags are then generated and the corresponding coordinates are assigned to them.

With the tags placed in the indoor construction environment, the UAV can be localized relative to the tags. Given that the global locations of the tags are available based on the BIM, the UAV's local coordinates can be translated to global coordinates. The authors reported the results of an experimental study designed to verify and validate the performance of the proposed framework. In this regard, the study investigated the impact of four critical parameters, namely, tag size, tag placement orientation, UAV distance from a single tag, and angle of view.
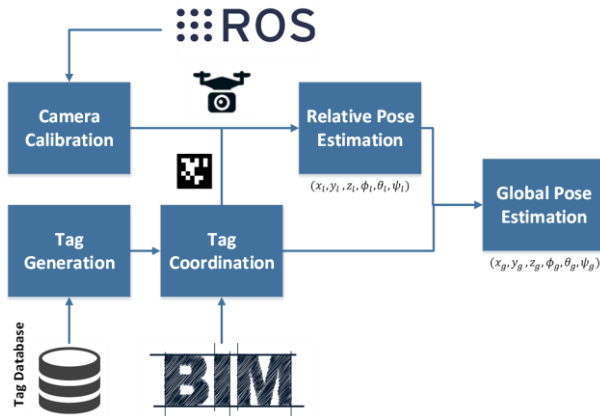
Figure 1. Indoor localization framework overview

That research also identified a number of factors that require further investigation to achieve a more robust indoor localization system. As all sensor measurements contain systematic and random errors, reducing or eliminating these errors is required to improve UAV pose estimation and reduce instability. It is intuitive that data from more sources of information can improve our measurements. Thus, to enhance the estimation process of the previously developed indoor localization framework [22], a standard extension of the KF [29] is used in this study.

In the previously developed method, the UAV used a single-tag detection algorithm. This algorithm only took measurements from a single designated tag at any given time. If this tag was not visible, the algorithm would estimate the UAV's location based solely on the less-reliable IMU measurements. In other words, the previously developed single-tag UAV localization method (hereafter the ST method) produced estimates based on only a single tag or the on-board IMU.

In the EKF method proposed in this work, however, the estimation process accounts for all available measurement data. It can be divided into two steps. The first step is the *prediction step*, which produces a predicted location based on the previous location and IMU data. The prediction is then updated in the *correction step*, where tag data is fused in to produce an overall estimate. This estimate is then used as the previous location during prediction in the next time step, and so on. By so doing, instead of switching back and forth between the IMU and the tag-based localization in different situations, we will combine the information from the IMU and all the visible tags. This multi-tag UAV localization method using EKF is hereafter called the MT method. This process will be investigated further in the next section.

3 The Extended Kalman Filter

Extensions of the KF are able to overcome the linearity limitation of the algorithm and enable the state estimation of nonlinear systems. Generally, the KF algorithm contains two steps: the prediction step and the correction step. In the prediction step, the current state (x_k) is estimated based on the previous estimated state (x_{k-1}) and the system inputs (u_{k-1}) considering a random Gaussian noise (w_{k-1}) when using a prediction model ($f(.)$) (See Eq. 1). The estimates are then updated using the measurement model ($g(.)$), which also has random Gaussian noise (v). (See Eq. 2)

$$x_k = f(x_{k-1}, u_{k-1}, w_{k-1}) \qquad (1)$$

$$y_k = g(x_k, v_k) \qquad (2)$$

A simplified dynamic model is used in which the UAV is modeled as a rigid body double integrator system with discretization time step ΔT. An EKF is then employed to first fuse the captured data from the IMU and AprilTag markers and then estimate the UAV's desired state in terms of position, velocity, and orientation. Thus, the state vector that we intend to estimate is defined as $x := [p^T \quad v^T \quad q^T]^T$, where 3D vectors p, v and quaternion q are representative of the UAV's position, velocity and orientation, respectively.

IMU measurements are treated as the input data for the prediction model, which is responsible for propagating the state from one time step ($k-1$) to the next (k). The process noise of this step is modelled by independent zero-mean Gaussian distributions. The prediction step ends up with a *prior* estimate of x, $\breve{x}_k$.

In the measurement model, because the position and orientation of all the tags are available, the relative position and orientation of the UAV with respect to all visible tags is measured at each time step k. The measurement noise is modeled as an independent zero-mean Gaussian distribution. Based on this model, the *prior* estimation $\breve{x}_k$ is updated. The final result of these two steps will be the *posterior* estimate of the state $\hat{x}_k$. This algorithm is implemented based on the EKF ROS node, *ekf_localization_node*.

By so doing, we are able to continuously fuse data from multiple sources, namely, the on-board IMU and visible tags, providing a more robust system than the previous work.

4 Experimental Design and Setup

To study the MT method proposed in this study, an experimental setup is designed. AprilTag markers of type *36h11* are placed at six locations, which are also identified in the BIM (see Figure 2). Corresponding

coordinates are then easily extracted from the BIM. A Vicon motion capture system is employed as the ground truth to verify and validate the algorithm. By so doing, we are able to compare the results under varying experimental conditions using two methods (i.e., ST and MT). The UAV used in this study is a Parrot Bebop 2 in its default configuration and equipped with a camera and an onboard IMU. The overhead Vicon system is also used for control.
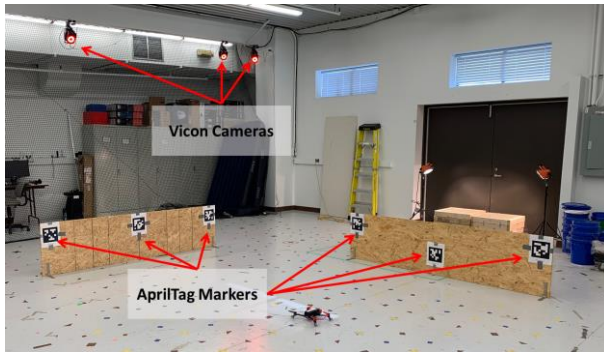

Figure 2. Experimental setup; six AprilTag markers with known coordinates, Parrot Bebop 2 UAV, and overhead Vicon camera system which tracks the motion of the UAV in the laboratory. The Vicon system is considered as the ground truth baseline in this study.

In these experiments, four trajectory scenarios are planned to investigate the dynamic performance of the proposed MT method.

1. *Straight line trajectory*. In the first scenario, a straight line trajectory is designed to investigate the impact of the distance from tags on the pose estimation accuracy. Although the MT method is able to fuse data from multiple sources, previous research [22] showed that the robustness of the tag detection process significantly drops when the

distance increases. Thus, the whole pose estimation process potentially becomes more prone to noise and error. As a result, it is important to investigate the performance of the EKF in this situation. To do so, we run the planned scenario and assess the performance of these two methods relative to the ground truth Vicon measurement.

As illustrated in Figure 3, the UAV starts its motion at a point 4.5m from panel B, which contains three tags. It moves 3m toward the panel at a rate of 0.86 m/s and finishes at a point 1.5m from the panel. In this scenario, all three tags on the panel B are used for the measurement process in the MT method. However, the middle tag is the only tag used in the ST method.

2. *Vertical line trajectory*. To investigate the impact of the angle of view on the MT method, a similar experimental setup is used, as shown in Figure 3. In this case, the UAV takes off from *O(0.0, 0.0, 0.0)* facing panel B, rises to the starting position of *Start(0.0, 0.5, 0.5)* and goes upward to *Finish(0.0, 0.5, 2.0)*.

3. *Horizontal line trajectory*. To study the impact of the angle of view, the UAV takes off from *O(0.0, 0.0, 0.0)* facing panel B, and moves to its starting position of *Start(-1.5, 1.0 ,0.5)*. It then goes sideward to *Finish(1.5, 1.0 ,0.5)*. In this scenario, the ST method uses the nearest tag, and the MT method fuses pose data from of all the visible tags.

4. *A mixed trajectory*. Finally, a combination of the aforementioned trajectories is planned to provide insight into the overall performance of each method in more realistic circumstances. In this scenario, the UAV takes off from *O(0.0, 0.0, 0.0)* facing panel A (see Figure 3). The planned trajectory contains all of the other scenarios and a rotational motion.
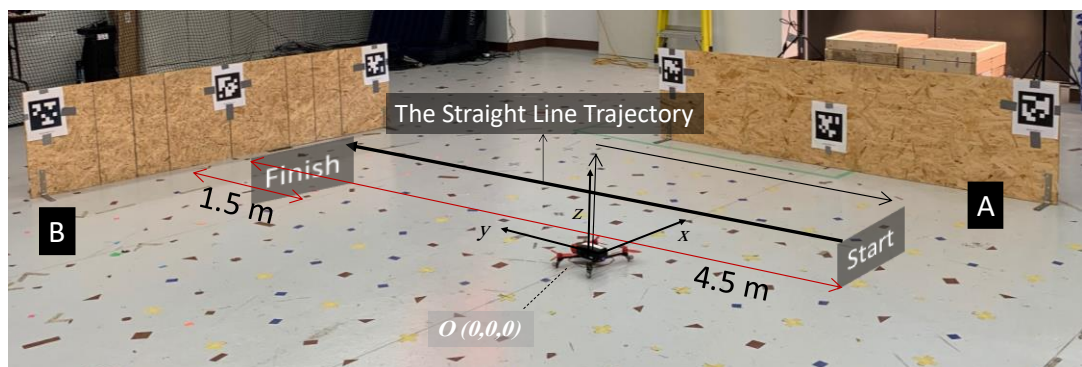

Figure 3. The experimental setup for scenario one; the performance of the methods are investigated at varying distances in the shown straight path. The UAV takes off from the *O(0.0, 0.0, 0.0)* while facing panel B. It flies its path from *Start(0.0, -1.5, 0.0)* to *Finish(0.0, 1.5, 0.0)*.

5   Results and Discussion

The experimental scenarios are designed to evaluate the performance of the MT method in indoor localization of UAVs using AprilTag markers. The aim is to show that this approach makes the prior framework more reliable in real world ever-changing construction environments. To do so, the localization results of the MT and ST methods are compared to the ground truth system. To assess the pose estimation results, the root-mean-square error (RMSE) is used as a quantitative indicator of error reduction. Figures are provided to offer a visual insight of the estimation reliability and show the fluctuations and errors.

First, the impact of distance from tags is investigated while the UAV automatically flies based on the programmed scenario. Figure 4 illustrates in 2D the performance of the UAV in the first scenario. In this scenario, the UAV flies from *Start(0,-1.5,0.5)* to *Finish(0,1.5,0.5)*. The top and bottom graphs in Figure 4 show the performance of estimation using the MT and ST methods, respectively. The Vicon system, as a ground truth system, depicts the actual path taken by the UAV.

In both cases, localization accuracy increases as the UAV gets closer to the tags. However, the estimation error and fluctuation is significantly improved using the EKF in the MT method. In the ST method, where the EKF is not used, localization varies approximately -1m to +0.6m from ground truth whereas the EKF was able to reduce this variation to a range of -0.2m to +0.6m. Thus, it is clear that the MT estimation is smoother and converges toward the ground truth measurements more quickly and at a distance further from the tags. Therefore, it performs more reliably compared to the other method used in the previous study.

To better understand the multidimensional components of the errors, the Euclidean distance between the 3D positions estimated by two methods and the ground truth system also calculated. The results are shown in Figure 5. Here, the MT method performs not only considerably better than the ST method, but also robustly in distances less than 3 meters. Numerically, the RMSE values for the global position estimates using the ST and the MT methods are 0.35 m and 0.20 m respectively. This further supports the finding that MT significantly enhanced our estimations.

Although the proposed framework is not significantly sensitive to the changes in yaw angle and view angle [22], the second and third scenarios are tested to compare the methods' performances. Figure 6 (A) and (B) respectively compare the ground truth measurement of X value in the third scenario with ST and MT estimates. Figure 6 (A) and (B) show that the MT estimates are smoother and more reliable in comparison with the ST estimates.
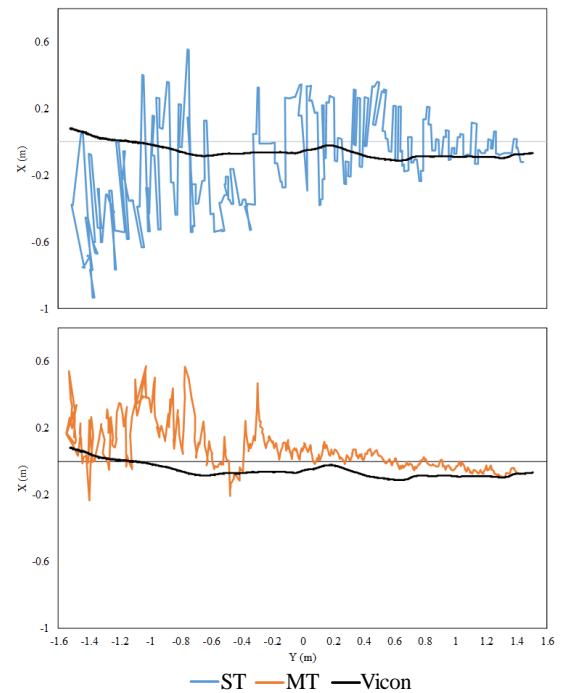

Figure 4. The plan view of the UAV's performance in Scenario 1 and the corresponding estimates: One-dimensional error in the pose estimation of the UAV while it decreases its distance from the tags.
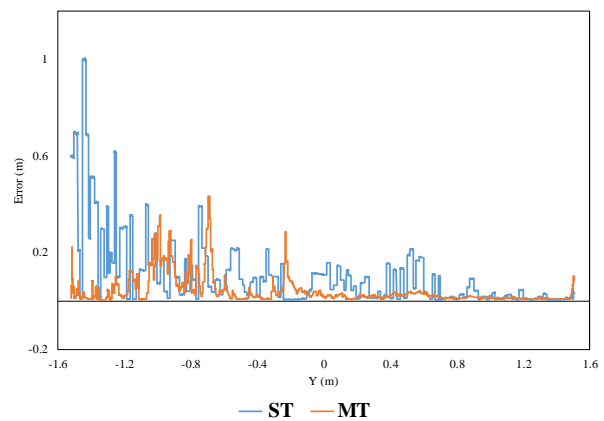

Figure 5. Position estimated error for the ST and the MT methods in Scenario 1
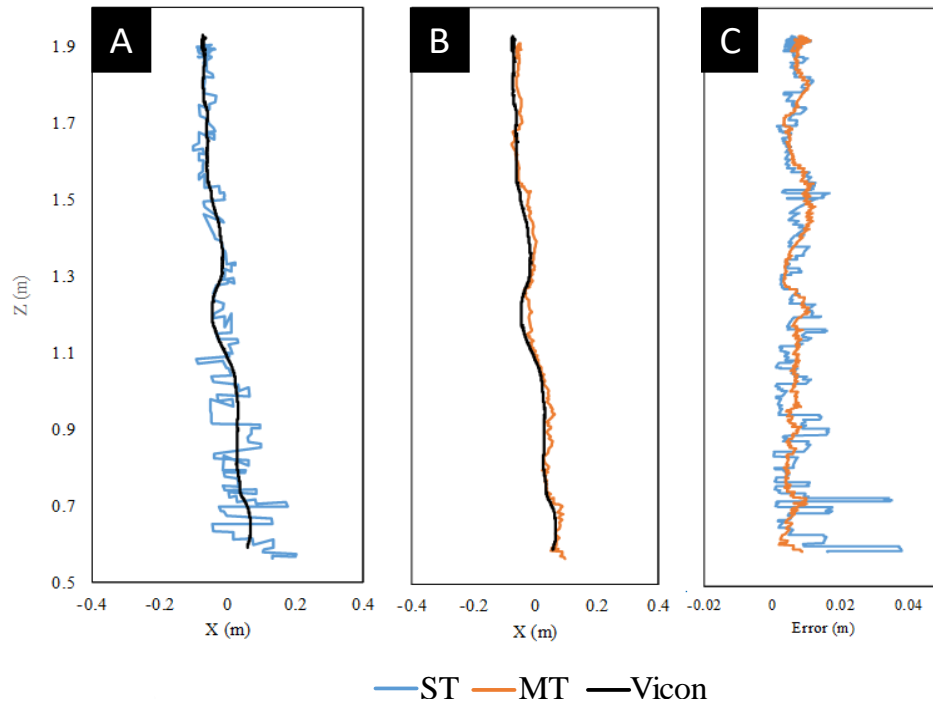
Figure 6. The second scenario: (A) estimated X value using the ST method; (B) estimated X value using the MT method; (C) error for two methods.

The error in Figure 6 (C) is defined as the Euclidian distance between the 3D positions estimated by two methods and the ground truth system. Even though the magnitude of the error is not large, Figure 6 (C) also confirms that the MT method not only performs noticeably more smoothly than the ST method, it also more reliably estimates the UAV's position.

Finally, to evaluate the performance of the proposed improvement in a real indoor construction environment, a full trajectory is tested in the fourth scenario. The estimated and ground truth positions of the UAV within the planned trajectory are plotted in Figure 7. The MT method is able to mitigate localization instability of the UAV by dramatically reducing the estimation error relative to the ST method. This is especially evident when compared in the same graph as shown in Figure 7. In this challenge, the UAV rotates from an orientation facing panel A to one facing panel B (Figure 3). There are some time steps in which the UAV is not able to see any tags, and therefore, the estimate is extremely noisy and unreliable, as evidenced by the ST method estimation performance in area A of the figure. Importantly, the MT method remains smoother and closer to the ground truth in this area.

In short, the overall performance of the MT method is significantly better than the ST method, because it estimates the ground truth position with fewer fluctuations and errors. To measure how these measurement errors are distributed, the RMSE is calculated. The RMSE can provide an overall insight about the performance of these two methods in the global pose estimation of the UAV. The RMSE for the ST method is 0.18m and for the MT method is 0.07m. Thus, it confirms that the MT method successfully improves our pose estimation in all the investigated scenarios.
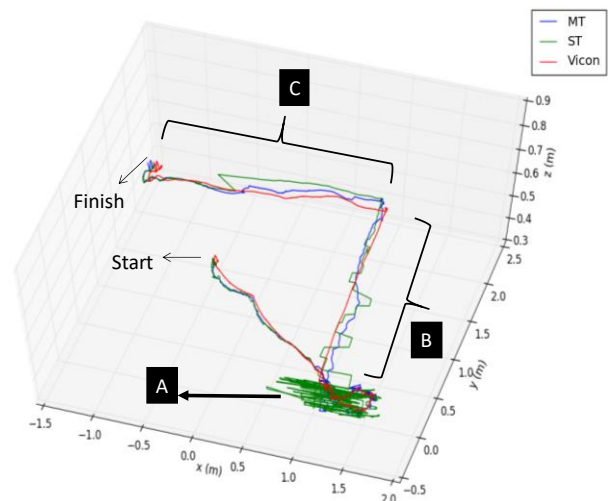

Figure 7. The performance of the methods in the fourth scenario; a planned mixed trajectory

6   Summary and Conclusion

To improve the state estimation process of the previously developed indoor localization framework of UAVs using AprilTag markers (i.e., ST), this study presented the implementation of an EKF (i.e., MT). By using an EKF, not only are we able to improve the estimation process by accounting for uncertainty, but we are also able to fuse data from two sources, namely, multiple tags and the onboard IMU. To evaluate the performance of the MT method within our framework, an experimental study was designed. Four scenarios were implemented, and pose estimation was done using both the ST and MT methods. The experimental results confirmed that the MT method successfully improves our pose estimation in all investigated scenarios. In other words, the estimated path using the MT method is smoother with less errors relative to the ground truth Vicon system. Thus, this approach improves upon our prior framework by making it more reliable for use in real-world and in ever-changing construction environments.

Future work includes using this method to provide state estimates for the control system. To this end, we need to conduct additional analyses and tests and further improve the performance of the state estimation module. For instance, in the proposed MT method, the EKF treats all the visible tags equally. However, localization data extracted from different tags may have different uncertainties (e.g., data associated with closer tags are generally more reliable than farther tags). These improvements should be implemented before any real-world tests are conducted.

Acknowledgment

References

[1]   Roca D, Lagüela S, Díaz-Vilariño L, Armesto J, Arias P. Low-cost aerial unit for outdoor inspection of building façades. *Automation in Construction.* 2013.

[2]   Hamledari H, McCabe B, Davari S, Shahi A, Rezazadeh Azar E, Flager F. Evaluation of Computer Vision- and 4D BIM-Based Construction Progress Tracking on a UAV Platform. *6Th Csce/Asce/Crc Int Constr Spec Conf.*, 2017.

[3]   Lin JJ, Han KK, Golparvar-Fard M. A Framework for Model-Driven Acquisition and Analytics of Visual Data Using UAVs for Automated Construction Progress Monitoring, *Computing in Civil Engineering*, 2015.

[4]   Siebert S, Teizer J. Mobile 3D mapping for surveying earthwork projects using an Unmanned Aerial Vehicle (UAV) system. *Automation in Construction*, 2014.

[5]   WANG F, CUI J-Q, CHEN B-M, LEE Comprehensive UAV Indoor Navigation System Based on Vision Optical Flow and Laser FastSLAM. *Acta Autom Sin.*, 2013.

[6]   Ham Y, Han KK, Lin JJ, Golparvar-Fard M. Visual monitoring of civil infrastructure systems via camera-equipped Unmanned Aerial Vehicles (UAVs): a review of related works. *Vis Eng.*, 2016.

[7]   Ellenberg A, Branco L, Krick A, Bartoli I, Kontsos A. Use of Unmanned Aerial Vehicle for Quantitative Infrastructure Evaluation. *J Infrastruct Syst.,* 2015.

[8]   Gheisari M, Irizarry J, Walker BN. UAS4SAFETY: The Potential of Unmanned Aerial Systems for Construction Safety Applications. In: *Construction Research Congress*, 2014.

[9]   Liu T, Yang L, Lin Q, Guo Y, Liu Y. Anchor-free backscatter positioning for RFID tags with high accuracy. In*: Proceedings - IEEE INFOCOM.*, 2014.

[10]   Witrisal K, Meissner P. Performance bounds for multipath-assisted indoor navigation and tracking (MINT). *IEEE Int Conf Commun.*, 2012.

[11]   Yassin A, Nasser Y, Awad M, et al. Recent Advances in Indoor Localization: A Survey on Theoretical Approaches and Applications. *IEEE Commun Surv Tutorials.,* 2017.

[12]   Deasy TP, Scanlon WG. Stepwise algorithms for improving the accuracy of both deterministic and probabilistic methods in WLAN-based indoor user localisation. *Int J Wirel Inf Network*, 2004.

[13]   Zhang R, Hoeflinger F, Gorgis O, Reindl LM. Indoor localization using inertial sensors and ultrasonic rangefinder. *Int Conf Wirel Commun Signal Process WCSP 2011.* 2011.

[14]   Liang JZ, Corso N, Turner E, Zakhor A. Image based localization in indoor environments. *Proc - 2013 4th Int Conf Comput Geospatial Res Appl COMGeo* 2013.

[15]   Motamedi A, Soltani MM, Hammad A. Localization of RFID-equipped assets during the operation phase of facilities. *Adv Eng Informatics.*, 2013.

[16]   Razavi SN, Moselhi O. Automation in Construction GPS-less indoor construction location sensing. *Automation in Construction*, 2012.

[17]   Costin AM, Teizer J. Fusing passive RFID and BIM for increased accuracy in indoor localization. *Vis*

Eng. 2015.

[18] Shahi A, Aryan A, West JS, Haas CT, Haas RCG. Deterioration of UWB positioning during construction, *Automation in Construction*, 2012.

[19] Taneja S, Akcamete A, Akinci B, Garrett JH, Soibelman L, East EW. Analysis of Three Indoor Localization Technologies for Supporting Operations and Maintenance Field Tasks. *J Computing in Civil Engineering*, 2012.

[20] Ibrahim M, Moselhi O. Inertial measurement unit based indoor localization for construction applications. *Automation in Construction*, 2016.

[21] Baek F, Ha I, Kim H. Augmented reality system for facility management using image-based indoor localization. *Automation in Construction*, 2019;99

[22] Nahangi M, Heins A, McCabe B, Schoellig A. Automated Localization of UAVs in GPS-Denied Indoor Construction Environments Using Fiducial Markers., *Proceedings of the 35th ISARC, Berlin, Germany),* 2018.

[23] Xiao J, Zhou Z, Yi Y, Ni LM. A Survey on Wireless Indoor Localization from the Device Perspective. ACM *Comput Surv.*, 2016.

[24] Mainetti L, Patrono L, Sergi I. A survey on indoor positioning systems. *22nd Int Conf Software, Telecommun Comput Networks, SoftCOM*, 2014.

[25] Mainetti L, Patrono L, Sergi I. A survey on indoor positioning systems., *22nd Int Conf Software, Telecommun Comput Networks, SoftCOM*, 2014.

[26] Fiala M. ARTag , a fiducial marker system using digital techniques., Proceedings - *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005.

[27] Olson E. AprilTag: A robust and flexible visual fiducial system. *Proc - IEEE Int Conf Robot Autom.*, 2011.

[28] Wang J, Olson E. AprilTag 2: Efficient and robust fiducial detection. *IEEE Int Conf Intell Robot Syst.,*2016.

[29] Atcheson B, Heide F, Heidrich W. CALTag : High Precision Fiducial Markers for Camera Calibration., *In VMV*, vol. 10, pp. 41-48. 2010.

[30] Kalman RE. A New Approach to Linear Filtering and Prediction Problems, *Journal of basic Engineering* 82, no. 1 (1960): 35-45, 1960;

[31] Welch G, Bishop G. An Introduction to the Kalman Filter., *University of North Carolina at Chapel Hill, 2001 (Lesson Course).* 2006:1-16.

[32] Moore T, Stouch D. A Generalized Extended Kalman Filter Implementation for the Robot Operating System., *IAS,* 2014.