

Robust Autonomous Navigation of Unmanned Aerial Vehicles (UAVs) for Warehouses' Inventory Application

Woong Kwon , Jun Ho Park , Minsu Lee , Jongbeom Her , Sang-Hyeon Kim , and Ja-Won Seo 

Abstract—The inventory inspection using autonomous unmanned aerial vehicles (UAVs) is beneficial in terms of cost, time, and safety of human workers. However, in typical warehouses, it is very challenging for the autonomous UAVs to do inventory task motions safely because aisles are narrow and long, and the illumination is poor. Prior autonomous UAVs are not suitable for such environments, since they suffer from either localization methods prone to disturbance, drift and outliers, or expensive sensors. We present a low-cost sensing system with an extended Kalman filter (EKF)-based multi-sensor fusion framework to achieve practical autonomous navigation of UAVs in warehouse environments. To overcome the inherent drift, outliers, and disturbance problems of naïve UAV localization methods, we suggest 1) exploiting component test of Mahalanobis norm to reject outliers efficiently, 2) introducing pseudo-covariance to incorporate a visual SLAM algorithm, and 3) recognizing floor lanes to get absolute information—as robust data fusion methods. Exemplar results are provided to demonstrate the effectiveness of the methods. The proposed system has been successfully implemented for diverse cyclic inventory inspection tasks in a materials warehouse.

Index Terms—Aerial systems, applications, aerial systems, perception and autonomy, sensor fusion.

I. INTRODUCTION

THE unmanned aerial vehicles (UAVs) which have the autonomous navigation and environment perception capability are rapidly arising as promising robotic applications in the industrial areas such as logistics, transportation, surveillance, etc. Most of current industrial UAVs are limited to outdoor use where GPS data are available. Indoor autonomous navigation of UAVs has been actively studied in robotics community, but there are only a few practical industrial cases.

The inventory inspection using UAVs at warehouses has several advantages over human workers' manual inspection, in that it can save time and cost when compared with using high-reach forklifts, and avoid injury risk of workers doing manual barcode scanning on a fast moving high-reach forklift. However, most

Manuscript received September 9, 2019; accepted November 13, 2019. Date of publication November 21, 2019; date of current version December 6, 2019. This article was recommended for publication by Associate Editor A. Elfes and Editor J. Roberts upon evaluation of the reviewers' comments. (*Corresponding author: Woong Kwon.*)

The authors are with the Factory Automation Technology Team, Global Technology Center, Samsung Electronics Co., Ltd., Suwon 16677, Republic of Korea (e-mail: wkwon@samsung.com; jun45@samsung.com; minsu46.lee@samsung.com; jongbeom.her@gmail.com; k3special@gmail.com; jawon.seo@samsung.com).

Digital Object Identifier 10.1109/LRA.2019.2955003



Fig. 1. Our UAV system does inventory inspection tasks in fully autonomous flights in a warehouse. It is able to fly within a narrow aisle (1.8 m width), move long distance (55 m~60 m length), and reach to the highest floor (12 m altitude) without any collision, while scanning barcodes of all the stock.

of commercial (non-autonomous) UAVs are not appropriate for the use at warehouses because safe remote operation is very difficult even by trained pilots [1]. Some of the drone makers [2], [3] announced to sell drone solutions for inventory tasks, yet no articles are found regarding practical applications at warehouses.

Fully autonomous UAVs without piloted flights are thought to be suitable solutions. Still, safe operation without collision is very challenging because 1) an UAV should fly within very narrow and long aisles between racks as shown in Fig. 1(a) and (b), 2) it should be close to stocks to read the barcode exactly as in Fig. 1(c), and 3) the illumination is commonly very poor in the middle of the long corridors. Several autonomous UAVs targeting warehouses [1], [4], [5], [6] have been presented, which will be explained in detail in the following section. However, none of them are demonstrated to work well in long and narrow (e.g., 1.8 meters wide) aisles.

Towards safe and successful inventory inspection at indoor warehouses, we suggest an Extended Kalman Filter (EKF)-based multi-sensor fusion framework using low-cost sensors: three cameras, a 2D laser scanner, a 1D range sensor, and an

IMU. We use a simple map which consists of pose information of attached tags, and so there is no need to do piloted flights for online map building. Furthermore, we propose three robust data fusion methods – 1) outlier rejection using Mahalanobis norm's component test, 2) addition of visual SLAM to measurement update by introducing ‘pseudo-covariance,’ and 3) recognition of floor lanes – to overcome the problems of drift, outliers, and disturbance of measurements, resulting to enhance the robustness of our multi-sensor fusion framework.

We have been successfully performing regular inventory inspection with our autonomous UAV. Results of doing various inventory task motions are provided to prove the performance of our methods.

To the authors' knowledge, such practical applications of UAVs in a warehouse are the first case in robotics community. The main contributions of this letter are as follows:

- 1) We suggest a low-cost multi-sensor system and EKF-based multi-sensor fusion framework to get optimal pose estimations.
- 2) We propose robust data fusion methods to strengthen the robustness and to remove the risk of collisions guaranteeing successful inventory inspections inside materials warehouse.
- 3) We build light and simple xml-based maps which are composed of only poses of pre-installed tags.
- 4) Our autonomous maneuvering UAV system can be extended to diverse areas requiring autonomous navigation capability.

II. RELATED WORK

Recently, several groups have published their results for the development of autonomous UAV system aimed at warehouse inventory applications. Beul *et al.* [1] presented high-performance autonomous inventory UAV for operations inside warehouses (funded by German Industry 4.0 project). Their sensor setup includes RFID reader, two high-resolution cameras, and an expensive 3D lidar. They performed experiments in an operative warehouse with an external warehouse management system and high-level inspection missions. Their method needs periodic piloted flights to build a 3D feature map, and has the step to align the semantic map with the 3D feature map.

Campos-Macias *et al.* [4] presented an autonomous navigation framework for exploration in unknown 3D cluttered environments. They used an RGBD camera for depth sensing and a tracking camera for visual-inertial odometry. Their demonstration scenarios include warehouse applications for capturing inventory and locating out of place items, while focusing on exploration task to reach a goal.

Welburn *et al.* [5] proposed a navigational system that enables UAVs to operate within a dark and GPS-denied environment(warehouse-like) to improve safety for human personnel. They use two 2D lidars and one camera for marker recognition, which are similar to our setup. However, their UAV system has not realized real autonomous flights yet.

Kalinov *et al.* [6] presented a high-precision UAV localization system interconnecting an unmanned ground robot(UGR)

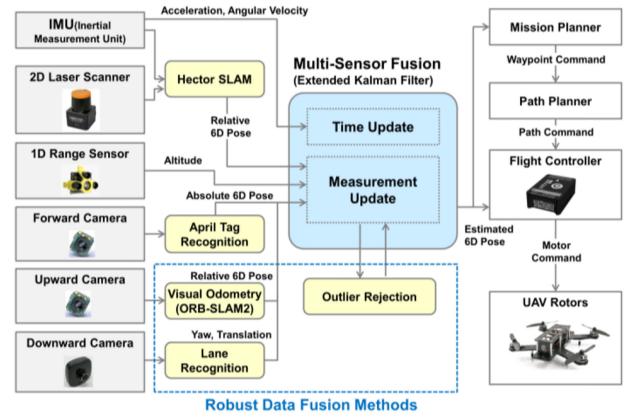


Fig. 2. Overall system architecture of our autonomous UAV.

and UAV, developed for automated inventory management of warehouses. They use a 2D lidar, multiple cameras, ultrasonic sensors, and adaptive active IR markers. Their results are yet limited to accuracy test of the UGR-UAV collaborative localization methods rather than autonomous flights.

III. SYSTEM OVERVIEW

We describe the overall architecture of our system, the hardware setup, the tag map concept and the multi-sensor fusion based localization framework. We also mention why the robust data fusion is necessary for complete autonomous flight.

A. Overall Architecture

Fig. 2 shows our system architecture consisting of the sensors, the software algorithm modules connected to the sensors, EKF-based multi-sensor fusion frameworks, and the other extra components. The proposed robust data fusion methods, which will be clarified in detail in the following section, are shown within the blue dashed rectangle.

Acquired data from the sensors are fed into the measurement update module of the EKF in the form of relative/absolute poses, angles, or distance, via corresponding algorithms. For instance, Hector-SLAM [7] provides 6D relative pose(x, y, z, roll, pitch, and yaw) of the UAV and April Tag [8] recognition algorithm gives its 6D absolute pose. More details of the algorithms including visual odometry(SLAM) and lane recognition will be elaborated in the Section IV.

The EKF-based multi-sensor fusion framework in conjunction with the robust data fusion methods sends optimally estimated UAV poses to the flight controller and the mission planner. The flight controller unit(FCU) outputs low-level motor commands for the quadrotor frame based on the feedback pose data to make stable flights possible. According to the current UAV pose estimation, the mission planner updates the waypoint list and transfers it to the path planner to make path commands updated. Those are used as the flight controller's reference trajectories. Since the mission planner, the path planner, and the flight control [16] is beyond the scope of this letter, we will not mention them anymore.

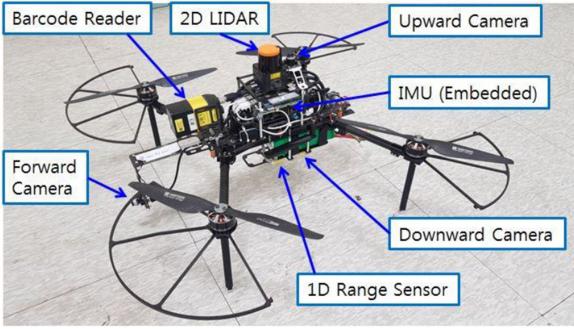


Fig. 3. The sensing system of our autonomous UAV.

B. Hardware Setup

For the purpose of robust indoor autonomous navigation, our UAV's sensing system is composed of a 2D laser scanner, a 1D range sensor, three cameras(forward, upward and downward), and an IMU(Inertial Measurement Unit) as shown in Fig. 3. A barcode reader is equipped for the inventory barcode scanning.

The 2D lidar(laser scanner) is Hokuyo UST-10LX with 270 degrees field of view and 10 meters range, which is much cheaper than 3D lidars. The 1D range sensor(lidar), Terabee TeraRanger One, measures the distance up to 14 meters between the UAV to the ground, to be used as an altitude measurement through UAV posture compensation. The forward and upward cameras, mvBluFox-MLC, are 752×480 monochrome cameras with global shuttering and equipped with flat and wide angle lenses respectively. The downward camera, oCam-1CNG-U, is 1280×960 global shuttering color camera. We use the IMU sensor embedded within Pixhawk3 FCU. All the cameras and the sensors are also low-cost under 400 USD. The barcode reader, Cognex DM363, is an image-based barcode scanner with the maximum range of 500 mm.

All the sensor drivers, software modules and algorithms run on a single board computer, Intel NUC7i7BNH with 16 GB RAM. We use the custom-made quadrotor frames as the flying mechanism. The horizontal size of the UAV is 74 cm \times 71 cm and its weight including battery is 3.2 kg.

C. Concept of Tag Map

The April Tag [8] (abbreviated to "tag" later), which is shown in yellow circles of Fig. 4, was devised to precisely calculate its relative 3D positions and 3D orientations with respect to the camera frame by means of a computer vision algorithm. Conversely, the forward camera and the tag recognition algorithm in Fig. 2 output relative 6D poses (3D positions and 3D orientations) of the camera with respect to the tag frame. Through simple coordinate transformation, we have the absolute 6D pose of the UAV with respect to the world coordinate frame if the poses of the tags in the world frame are provided beforehand.

Prior to the operation, we build an xml-based map storing each pose of tags in the world frame, which are attached to the predetermined locations of the every rack as shown in Fig. 4. Thus, the absolute 6D pose data of the UAV calculated from tag recognitions are used as one of the measurement update of the EKF, once tags are observed.



Fig. 4. April Tags (in yellow circles) attached to the warehouse rack. Their position and orientation information is saved in an xml-based tag map beforehand for the autonomous navigation of the UAV.

Since this tag map is an xml file, its size is very small (under 800 KB per warehouse) when compared with 3D feature map. Also there is no need for piloted flights to build a map periodically in contrast to [1].

D. Multi-Sensor Fusion Framework

We suggest a multi-sensor fusion framework to optimally estimate the 6D pose of the UAV in the indoor(GPS-denied) environments. The framework is based on the EKF where the IMU pose is used as state variables of the time update phase - prediction - and the other sensors' data are processed in the measurement update – correction - phase in a modular way [9] as depicted in Fig. 2.

We assume that the pose of the IMU sensor is equivalent to that of the UAV. The state of the EKF consists of the position vector of the IMU p_w^i in the world frame $\{W\}$, its velocity vector v_w^i , its orientation quaternion q_w^i describing the rotation of IMU with respect to W , gyro bias b_ω and accelerometer bias b_a . The entire state yields a 16-dimensional state vector x :

$$x = [p_w^i \ T \ v_w^i \ T \ q_w^i \ T \ b_w \ T \ b_a \ T]^T. \quad (1)$$

We derive the following differential equations:

$$\dot{p}_w^i = v_w^i \quad (2)$$

$$\dot{v}_w^i = R_w^i(a_m - b_a - n_a) - g \quad (3)$$

$$\dot{q}_w^i = \frac{1}{2}\Omega(\omega_m - b_\omega - n_\omega)q_w^i \quad (4)$$

$$\dot{b}_\omega = n_{b_\omega} \quad (5)$$

$$\dot{b}_a = n_{b_a} \quad (6)$$

where R_w^i is a rotation matrix corresponding to quaternion q_w^i , a_m is a measured acceleration from IMU, $\Omega(\omega)$ is the quaternion-multiplication matrix of ω [10], ω_m is a measured angular velocity from IMU, g is the gravity vector in the world frame, and n_a , n_ω , n_{b_ω} , and n_{b_a} are Gaussian noises. Equations (5) and (6) mean that the bias terms are modeled as random walk. Taking the expectations of differential equations (2)~(6), and defining the error state vector Δx of (1), we have a linearized error state equation

$$\dot{\Delta x} = A\Delta x + Bn \quad (7)$$



Fig. 5. Images from the forward camera showing the recognition of an April Tag while the UAV is flying horizontally. As the tag is seen far from the image center, the possibility of wrong detection (i.e., outlier) increases.

where A and B are linearized matrices of (2)~(6)(See [11] for details). Consequently, we are able to obtain the discrete time update process of the EKF from (7) [11].

The measurement model relates the state vector and the sensor measurement as

$$z = h(x) + n \quad (8)$$

where z is a measurement vector such as position, Euler angle, and pose, $h(x)$ is the corresponding measurement model and n is Gaussian measurement noise. We employ Lynen's [9] modular multi-sensor fusion framework where individual sensors' measurement updates are fused independently to get the optimal estimation of the state. It is advantageous that we are able to add our own more sensors to this framework.

A minimal sensor suite for the valid measurement update of our EKF is a camera which measures the pose of tags attached to the rack frames providing the 6D absolute pose, a 2D laser scanner plus IMU which calculate the 6D relative pose from scan matching techniques like Hector-SLAM [7], and a downward 1D range sensor which estimates the altitude of the UAV.

Theoretically, this concept should guarantee sufficiently consistent results because complete 6D poses are always available by combining the absolute and the relative pose measurements. However, in practice, this is not the case, since there inevitably exist drift, outliers, and disturbance of measurements due to system uncertainty and incomplete sensing data. Moreover, warehouse environments commonly shows very challenging situations for vision-based autonomous navigation, because the illumination is very poor in the middle of the rack, the corridors are only 1.8 meters wide, and the rack is so long (55~60 meters) that a small errors or drifts may cause collisions. Thus, the above sensing system does NOT make practical applications of autonomously maneuvering UAVs possible – requiring extremely high safety – without addressing such problems.

IV. ROBUST DATA FUSION

We suggest three approaches as the robust data fusion methods to increase the robustness of our EKF, in such a way as to overcome the drift, outliers, and disturbance issues. These help the multi-sensor fusion framework to cope with the challenging warehouse environments successfully.

A. Outlier Rejection by Mahalanobis Norm Test

The outliers usually occur from sensor faults, large recognition errors or bad measurement environments. Fig. 5 represents a typical scenario of outliers in tag recognition, where the sequence of images during horizontal flying motion is listed.

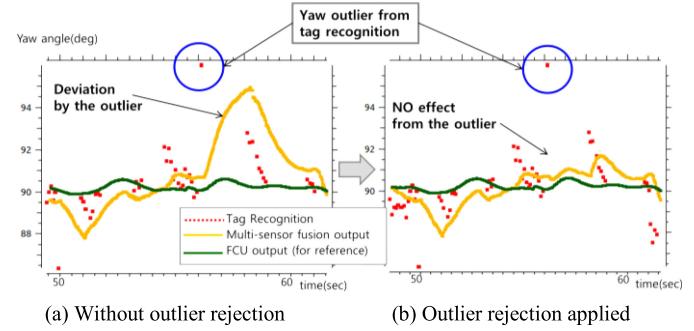


Fig. 6. The effect of rejecting an outlier from tag recognition in yaw angle estimation.

As the tag is detected farther away from the image center, the possibility of occurrence of tag detection outliers increases. We found that in our warehouse outliers of tag recognition also appeared due to bad attachments of the tags and bending of the tag plates. Besides, the 1D range sensor and the 2D laser scanner also generate outliers.

In order to reject such outliers of measurements [9], we calculate and monitor the Mahalanobis norm M of a measurement update

$$M^2 = (z - H\hat{x})^T S^{-1} (z - H\hat{x}) \quad (9)$$

where z is the measurement vector, H is the linearized measurement matrix of $h(x)$ in (8), S is the state covariance matrix, and \hat{x} is the state prediction. We propose that specific components of this norm - for example, yaw/pitch angles from tag recognition or lateral positions from Hector-SLAM - be individually checked during mission flights. If some of them are larger than a predetermined threshold, our EKF framework will reject the corresponding component of the measurement. This component-based Mahalanobis norm test reduces the false rejection ratio by separating a measurement vector into smaller dimensional vectors.

Fig. 6 shows an exemplar result of tag's yaw component outlier rejection. Unless the outlier is rejected, the multi sensor fusion output (yellow line) is deviated largely as in Fig. 6(a) because the absolute pose from "relative" Hector-SLAM data (not shown in this figure) is recalculated based on the recent tag measurements. However, if the outlier is rejected from the measurement update, the deviation of the EKF output due to the outlier disappeared as in Fig. 6(b).

B. Pseudo-Covariance to Utilize Visual SLAM

To minimize the disturbance effects caused by the laser scanner's Hector-SLAM – usually occurring during vertical movement, we add a visual SLAM algorithm (the state-of-the-art is ORB-SLAM2 [12]) implemented by using the upward camera capturing images toward the ceiling. The x , y , and z position; and yaw orientation data from this algorithm are fused into EKF to weaken disturbance effects from Hector-SLAM or other unknown sources.

Since visual SLAM algorithms like ORB-SLAM2 are mostly keypoint-based, they do not provide the covariance of the output,

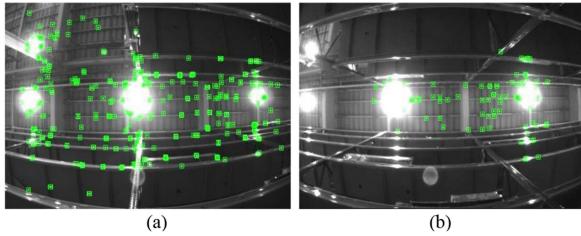


Fig. 7. Images from the upward camera showing the ORB SLAM2 results. The numbers of matching points in motion model are (a) 288 and (b) 47. As the number decreases, the uncertainty of pose estimation tends to increase. Therefore, we use this number to calculate the pseudo-covariance.

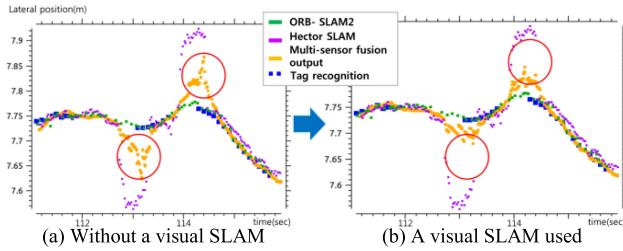


Fig. 8. The effects of applying pseudo-covariance of ORB SLAM2.

which is required for the measurement update phase of EKF. We propose to use the number of matching points in motion model, as depicted in Fig. 7, to extract the uncertainty information of the ORB-SLAM2 - namely ‘pseudo-covariance’. Through repetitive experiments and tuning process, we obtained the best proportional constant and offsets between the number of matching points and the pseudo-covariance that works well.

By exploiting pseudo-covariance of ORB-SLAM2, we can apply its measurement update into EKF to reduce the effects of disturbance. Fig. 8 illustrates the effects of applying ORB-SLAM2. The large deviation of the multi-sensor fusion output in Fig. 8(a) is effectively reduced in Fig. 8(b), as shown within the red circles.

One more thing that should be kept in mind is that the ORB-SLAM2 often loses the tracking if there are not enough matching points. In that case we make the algorithm to try to restart the tracking quickly. Since we use the monocular camera, we should get the scale factor of the ORB-SLAM2 output data after the tracking starts as soon as possible. For the calculation of the scale factor, we employ Maximum Likelihood Estimation(MLE)-based method offered by Engel *et al.* [13] since its calculation considers the uncertainty information and so gives consistent results.

C. Recognition of the Lane on the Floor

To get rid of the drift effects along yaw and lateral direction in the rack aisles, we add a downward color camera to recognize the lane information so as to extract lateral position and yaw orientation as in Fig. 9.

By extracting edge features of the lane through Fast Fourier Transform (FFT) and likelihood computation, we obtain the translational distance between the lane and image center lines, and the twisting angle of the lane. The former can be interpreted as measurements of the lateral position by multiplying the scale

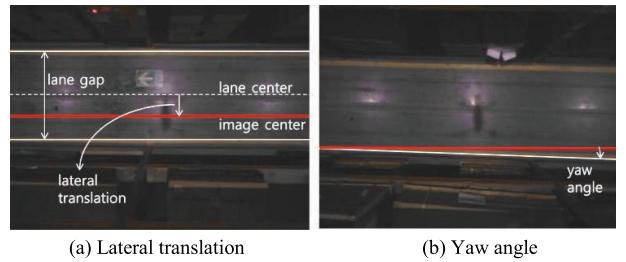


Fig. 9. Images from the downward camera showing the results of lane recognition.

ratio of the physical gap of both lanes to their gap on the image plane. The latter is used as yaw measurements of the UAV. These two absolute measurements are very helpful to remove the influences of drift of the estimation. This lane recognition is especially effective in flights at the 2nd and 3rd floor altitude and in the middle of corridors.

V. RESULTS IN A WAREHOUSE APPLICATION

Using our EKF-based UAV localization including the three methods to enhance robustness, we have been conducting various inventory inspection tasks, actually accomplishing safe autonomous navigation. Regarding the poses from tag recognition as ground truth (approximately), we estimated the localization accuracy of the proposed methods. The average and maximum localization error is 3.12 cm and 25.68 cm, respectively, in a full inspection task of a single rack (6 floors and 55 m long), which are sufficient for secure operations.

The proposed methods prevent the UAV from colliding into near racks or walls by removing the drift of naïve localization methods. Outliers of tag recognition, occurring from bad attachments, bended tag plates, or recognition faults, were clearly rejected to keep the estimation from serious deviation. The bad effects of frequent disturbance of Hector-SLAM appearing mainly in ascending/descending phases were also minimized in the estimation results, consequently showing stable flights. Moreover, the map building process was very simple and cost effective since our map is only composed of marker tags’ pose data. On the contrary, the approaches using 3D feature map, as in [1], typically need big map files and periodic map building steps through manual flights to reflect consecutively changing piled boxes in a warehouse.

Fig. 10 reveals a comparison result to verify the improvements effects of the proposed robust data fusion methods. During an ascending motion between floors, the previous UAV frequently shows large deviating motions approaching to a rack as shown in Fig. 10(a) due to drift or disturbance of the localization. This situation is very dangerous because it may cause collisions into the near rack or stocks. After applying the proposed methods, the UAV is able to maintain the center path safely as in Fig. 10(b). Note that the yellow dotted lines are drawn along the vertical image center line for the easy visual understanding.

Typical inventory inspection tasks are categorized into two scenarios. One is the full inspection of stocks at the whole racks of a warehouse. In our local warehouse, there are 32 racks in total for stock storage. Each rack is 55 meters \sim 60 meters long

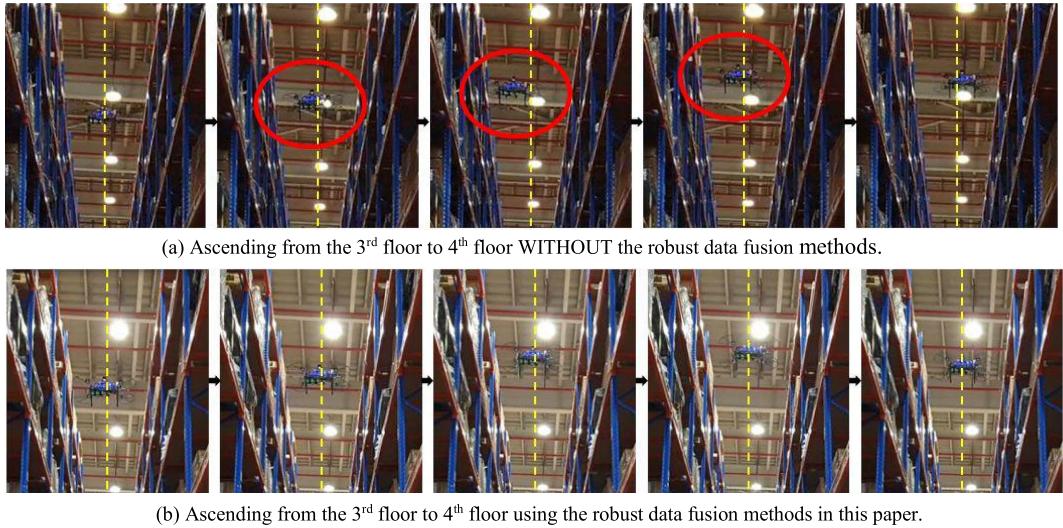


Fig. 10. An exemplar motion demonstrating the improvement of using the robust data fusion methods. While the previous UAV had shown a large disturbance motion approaching to the rack dangerously (the 2nd red circle), the current UAV maintains the center path safely. The yellow dotted lines are drawn along the vertical image center line for the visual understanding.

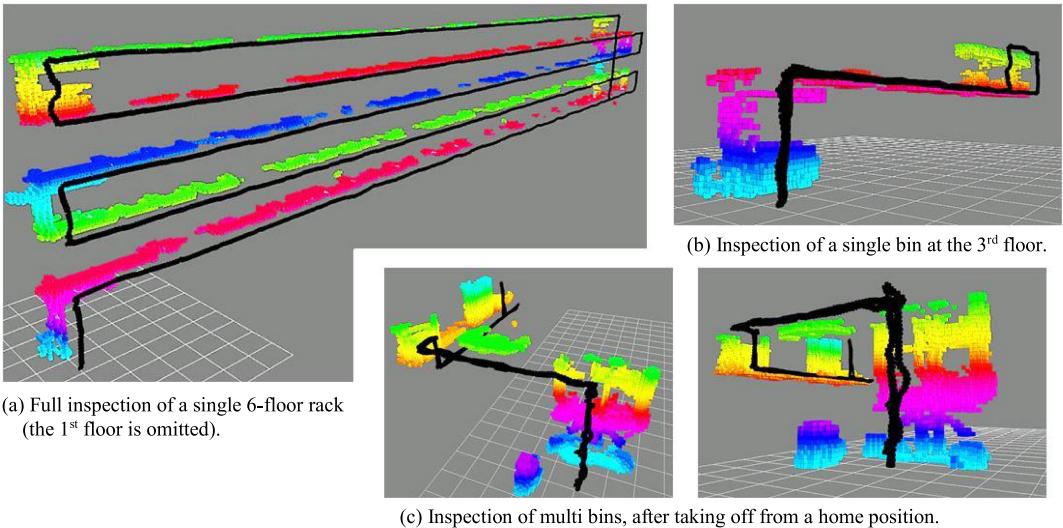


Fig. 11. A variety of tasks for inventory inspection conducted by our UAV system at the warehouse. For the easy visual understanding, we reconstruct the voxel map [14], [15] and the path of the UAV from the data gathered in flights. The colors of the voxels reflect their altitude, and the black trajectory is the UAV's estimated pose. Note that, since a 2D laser scanner is used for environmental mapping, the 3D voxel map is created only within the laser scanner's field of view.

and 6 floors high, and has 180 bins for pallets. Full inspection means that all the bins of the racks are scanned at once. Common warehouses do full inspection periodically for complete inventory management. Fig. 11(a) depicts a full inspection task of a single rack performed by our autonomous UAV. In Fig. 11, for the intuitive understanding, we reconstruct the 3D voxel map [14], [15] and the path of the UAV from the data gathered in real autonomous flights. The colors of the voxels reflect indicate their altitude, and the black trajectory is the estimated pose of the UAV. Since the environmental mapping is done by the 2D laser scanner, the 3D voxel map is created only within the laser scanner's field of view.

The other category is cycle counting which is to count items in specific bins within the warehouse without having to count the entire inventory. Fig. 11(b) and (c) describe cycle counting

task motions done by our autonomous UAV targeting to bins at the 3rd floor. The UAV may start cycle counting tasks by taking off at a place within the corridor in front of the target rack as in Fig. 11(b), and at a home position outside of the aisles to reach multiple racks in turn as shown in Fig. 11(c).

With our autonomous UAV, we have been doing both full inspection and cycle counting tasks successfully in the materials warehouse without any problems for one year, replacing the conventional manual inventory inspection done by human workers. Fig. 12 shows several snapshots captured during a variety of the inventory inspection motions of our autonomous UAVs. Prior manual inspection tasks typically required high-reach forklifts, its licensed driver, and one or two workers on board holding a hand-held barcode scanner. The application of the autonomous UAV enabled the warehouse to save such large equipment costs

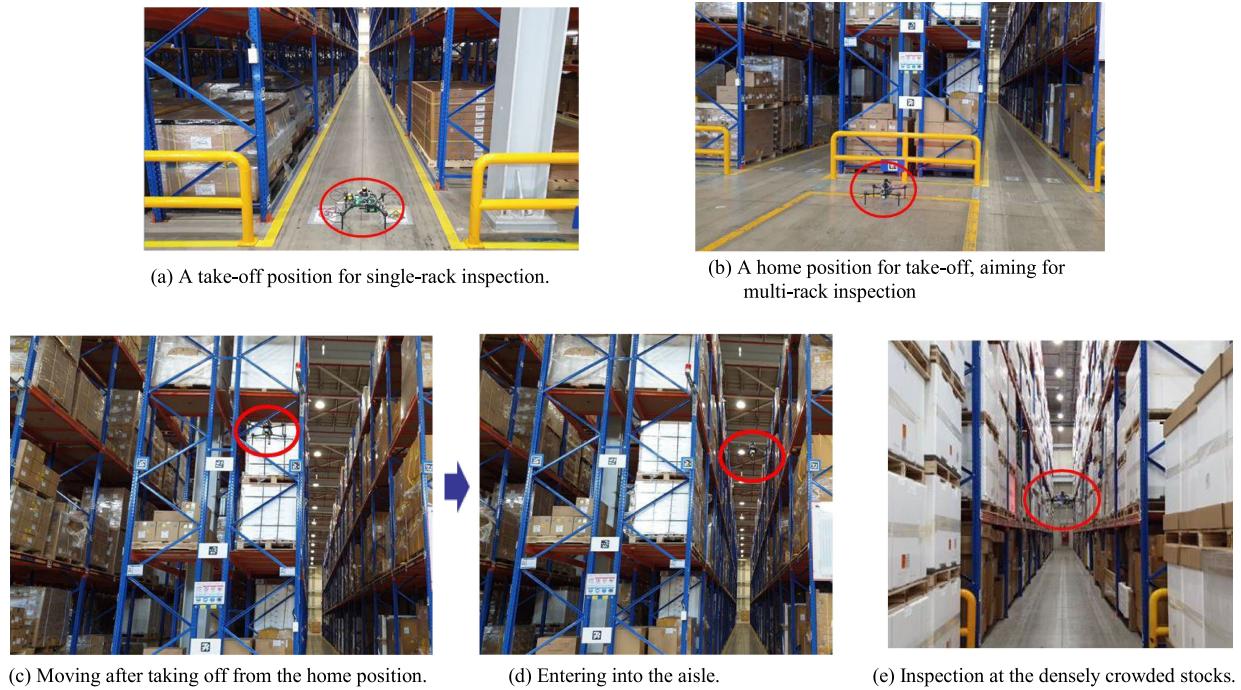


Fig. 12. Photo captures of our UAV system while doing various motions for the inventory inspection tasks at the warehouse.

and human labors efficiently. The inspection time is also reduced since it takes only 9 minutes and 14 seconds to complete a full inspection of a single rack containing 180 bins. One of the most important consequences of our autonomous UAV is to get rid of the risk of human workers' unpredicted fatal injury while doing barcode scanning on a vertically moving high-reach forklift.

VI. CONCLUSION

We proposed an autonomous UAV with a low-cost sensing system and multi-sensor fusion framework to be used effectively for narrow and dark warehouse environments. To address the problems of naïve UAV localization methods, we suggested robust data fusion methods: outlier rejection using component-based Mahalanobis norm test, incorporation of visual SLAM by introducing pseudo-covariance, and recognition of floor lanes for absolute lateral position and yaw measurements. These allow us to perform safe and fully autonomous flights for the cyclic inventory inspections at our materials warehouse. As a fully autonomous system, our method does not need piloted flight, and also reduces the cost, time and the risk of human's high-floor tasks. For further study, we plan to improve the system reducing the cost more, and extend our method to various indoor and outdoor applications.

REFERENCES

- [1] M. Beul, D. Droeschel, M. Nieuwenhuisen, J. Quenzel, S. Houben, and S. Behnke, "Fast autonomous flight in warehouses for inventory applications," *IEEE Robot. Autom. Lett.*, vol 3, no. 4, pp. 3121–3128, Oct. 2018.
- [2] EyeSee, 2019. [Online]. Available: eyesee-drone.com
- [3] DroneScan, 2019. [Online]. Available: www.dronescan.co
- [4] L. Campos-Macias *et al.*, "Autonomous navigation of MAVs in unknown cluttered environments." [Online]. Available: <https://arxiv.org/abs/1906.08839>
- [5] E. Welburn *et al.*, "A navigational system for quadcopter remote inspection of offshore substations," in *Proc. 15th Int. Conf. Autonomic Auton. Syst.*, 2019.
- [6] I. Kalinov, E. Safronov, R. Agishev, M. Kurenkov, and D. Tsetserukou, "High-precision UAV localization system for landing on a mobile collaborative robot based on an IR marker pattern recognition," in *Proc. IEEE 89th Veh. Technol. Conf.*, Kuala Lumpur, Malaysia, 2019, pp. 1–6.
- [7] S. Kohlbrecher *et al.*, "Hector open source modules for autonomous mapping and navigation with rescue robots," in *Robot Soccer World Cup*. Berlin, Germany: Springer, 2013, pp. 623–631.
- [8] J. Wang and E. Olson, "AprilTag 2: Efficient and robust fiducial detection," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2016, pp. 4193–4198.
- [9] S. Lynen, M. W. Achtelik, S. Weiss, M. Chli, and R. Siegwart, "A robust and modular multi-sensor fusion approach applied to MAV navigation," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Tokyo, Japan, 2013, pp. 3923–3929.
- [10] N. Trawny and S. Roumeliotis, "Indirect Kalman filter for 3D attitude estimation," Dept. Comput. Sci. Eng., Univ. Minnesota, Minneapolis, MN, USA, Tech. Rep. 2005-002, 2005.
- [11] S. Weiss, "Vision based navigation for micro helicopters," Ph.D. dissertation, Autonomous Syst. Lab, Dept. Mech. Process Eng., ETH Zürich, Zürich, Switzerland, 2012.
- [12] R. Mur-Artal and J. D. Tardós, "ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras," *IEEE Trans. Robot.*, vol. 33, no. 5, pp. 1255–1262, Oct. 2017.
- [13] J. Engel, J. Sturm, and D. Cremers, "Scale-aware navigation of a low-cost quadrocopter with a monocular camera," *Robot. Auton. Syst.* vol. 62, pp. 1646–1656, 2014.
- [14] H. Oleynikova, Z. Taylor, M. Fehr, R. Siegwart, and J. Nieto, "Voxblox: Incremental 3D euclidean signed distance fields for on-board MAV planning," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Vancouver, BC, Canada, 2017, pp. 1366–1373.
- [15] A. Hornung *et al.*, "OctoMap: An efficient probabilistic 3D mapping framework based on octrees," *Auton. Robots*, vol. 34, pp. 189–206, 2013.
- [16] Pixhawk, 2019. [Online]. Available: pixhawk.org