Complementary Multi-Modal Sensor Fusion for Resilient Robot Pose Estimation in Subterranean Environments

Conference Paper · September 2020

DOI: 10.1109/ICUAS48674.2020.9213865

Complementary Multi–Modal Sensor Fusion for Resilient Robot Pose Estimation in Subterranean Environments

Shehryar Khattak, Huan Nguyen, Frank Mascarich, Tung Dang, and Kostas Alexis

Abstract— **Resilient pose estimation for autonomous systems, and especially small unmanned aerial robots, is one of the core capabilities required for these robots to perform their assigned tasks in a reliable and efficient manner. Different sensing modalities have been utilized for the robot pose estimation process, particularly in GPS–denied environments. However, as aerial robots are deployed in more complex environments, such as subterranean mines and tunnels, different sensing modalities can become degraded in different parts of the environment due to the diversity of sensor perception challenges presented in terms of both nature and condition of the operational environment. Motivated by this fact, in this work a complementary multi–modal sensor fusion approach is presented that improves the reliability of the pose estimation process for aerial robots by fusing visual–inertial (VIO) and thermal–inertial (TIO) odometry estimates with a LiDAR odometry and mapping solution. In particular, VIO/TIO estimates are utilized for providing robust priors for LiDAR pose estimation as well as for selectively propagating the LiDAR pose estimates when LiDAR pose estimation process becomes degenerate. The proposed approach is experimentally verified in a variety of subterranean environments as well as utilized during the competition run of the tunnel circuit of the DARPA Subterranean Challenge.**

I. INTRODUCTION

Robust robot pose estimation is one of the key capabilities required for the successful execution of a number of complex and critical applications involving aerial robots. In recent years, aerial robots have been increasingly applied in a multitude of application areas, such as search and rescue [1, 2], disaster response [3, 4], security surveillance [5, 6], infrastructure inspection [7, 8], precision agriculture [9, 10], exploration and mapping [11, 12] as well as other commercial and social applications [13, 14]. Due to the increasing application of aerial robots towards more time sensitive and mission critical tasks, it becomes more important to ensure the reliability of robot pose estimation solutions. As these complex applications demand that aerial robots are able to reliably navigate through a variety of diverse environments, while remaining resilient to the challenges and constraints imposed by the operating conditions and geometric nature of the traversed environments. Similarly, these applications also often require aerial robots to estimate their pose fully on–board in real–time without the aid of external pose estimation inputs, such as from GPS. Furthermore, it is also common to encounter sensory degraded conditions during these applications as imposed by the nature, structure and operational conditions of the environment, for example encountering conditions of complete darkness, geometric self similarity and presence of obscurants, such as smoke, fog and dust. In applications such as exploration and inspection of subterranean mines [15] it is typical to encounter such perception conditions, which have also recently become a focus of the robotics community especially due to the DARPA Subterranean (SubT) Challenge [16].

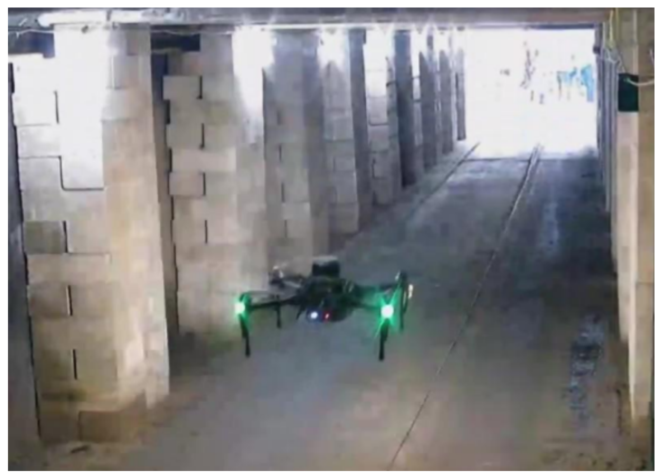

Fig. 1. The image shows an aerial robot autonomously navigating and exploring part of an underground facility during an instance of the competition run of the tunnel circuit of the DARPA Subterranean Challenge.

In GPS–denied environments aerial robots rely on their on–board sensing payload to estimate their pose as they navigate through an environment and perform their assigned tasks. In this regard, a number of sensing modalities have been utilized for the purposes of online robot pose estimation, with each associated for its own strengths and weaknesses. For aerial robots, visual spectrum cameras remain a popular choice as a sensing payload since they provide light-weight, power-efficient and low-cost sensing to enable robot pose estimation in GPS–denied environments. However, in conditions of poor illumination and low texture as well as in the presence of obscurants, such as dust, fog and smoke, data from visual cameras can degrade significantly and cannot be relied upon for robust robot pose estimation. As an alternate to visual cameras, due to their recent miniaturization and cost competitiveness, long range Light Detection And Ranging (LiDAR) units have also become a viable sensing choice to be carried on–board aerial robots. LiDARs provide direct depth measurements and are not

This material is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) under Agreement No. HR00111820045. The presented content and ideas are solely those of the authors.

The authors are with the Autonomous Robots Lab, University of Nevada, Reno, 1664 N. Virginia, 89557, Reno, NV, USA shehryar.khattak@nevada.unr.edu

affected by scene illumination and texture changes, hence enabling online pose estimation and mapping capabilities. However, LiDAR based robot pose estimation methods tend to suffer in structure-less or self similar environments as the underlying pose estimation process solution relies on the environment to provide sufficient geometric constraints for its reliable operation. Furthermore, obscurants such as dust and fog can also significantly degrade LiDAR data by obstructing and absorbing LiDAR beams. In contrast, thermal cameras operating in the Long Wave Infrared (LWIR) spectrum have the ability to maintain perception in challenging visually degraded environments as they are not affected by the conditions of darkness and have the ability to penetrate through most visual obscurants. Recent reductions in the size, weight and cost of thermal cameras also make them a viable sensing choice to be carried on-board an aerial robot for the purposes of robot pose estimation. However, as thermal cameras operate by sensing the thermal emissivity of objects, the information provided by thermal cameras can become scarce in thermally-flat environments with low temperature variations and thus becoming insufficient for the purposes of reliable robot pose estimation.

Motivated by the discussion above, in this paper we propose a loosely–coupled degeneracy–aware multi–modal sensor fusion approach that selectively utilizes visual/thermal and inertial estimates to improve the reliability of LiDAR Odometry and Mapping in degenerate environments. The feasibility of the proposed complementary multi–modal fusion approach is first experimentally demonstrated through a set of aerial robotic missions conducted in structure-less or self similar environments. Furthermore, the real–world application of the proposed approach is demonstrated through an aerial robot field deployment conducted in an active underground mine subject to the conditions of darkness and int the presence of heavy air–borne dust. Finally, the reliability of the proposed approach is demonstrated through its utilization during the tunnel circuit competition run of the DARPA SubT challenge. An instance of the robotic deployment during the scored competition run is shown in Figure 1 along with the full video available at `https://tinyurl.com/darpa-aerial-deployment`.

The rest of the paper is structured as follows: Section II provides an overview of the related work, while Section III details the proposed complementary multi–modal sensor fusion approach. The experimental evaluation studies are presented in Section IV. Finally, conclusions are drawn in Section V.

II. Related Work

Estimation of robot pose in GPS–denied environments under adverse conditions is a particularly challenging task and has been an active area of research as it is one of the most important enabling factors for many potential robotic applications. Robots typically rely on their on–board sensing payload to determine their pose effectively in real–time. Typically utilized sensing modalities have shown excellent results within applications taking place in environments which are visually and geometrically feature–full in nature, however, they remain prone to degradation and unreliability when presented with a particular set of challenging operating conditions. Visible light cameras have been a particularly popular choice for aerial robots due to their low weight, cost and power consumption, and have shown excellent results in a number of applications [17–19]. However, their utility becomes limited in poorly–illuminated, low–texture, and obscurant–filled environments. LiDARs, due to their miniaturization in size, reduction in weight and lowering of cost have also been recently deployed on aerial robots and are a particularly popular choice for large–scale mapping applications [20–22]. Nevertheless, robot pose estimation approaches that utilize direct depth measurements from LiDARs rely on the environment to provide sufficient geometric constraints as otherwise the underlying pose estimation optimization problem can become ill–conditioned or degenerate. Similarly, in the presence of visual obscurants, such as dust and fog, LiDAR scan beams can be obstructed or absorbed making the provided depth data unreliable for robot odometry estimation purposes. In contrast to visible light cameras and LiDARs, LWIR thermal cameras are not effected by the presence of visual obscurants and as shown in [23–25]. can be utilized for the pose estimation of aerial robots. However, in thermally–flat environments containing minimal temperature gradients the reliability of pose estimation using thermal cameras can significantly degrade.

Given the discussion above, it can be understood that each sensing modality can become unreliable under a set of adverse conditions. However, it can also be noted that although that these adverse conditions may degrade a particular sensing modality yet not affect another, for example conditions of darkness can affect the performance of visible light cameras but not affect LiDARs or thermal cameras. Hence, these sensing modalities can be utilized in complementary manner as proposed in [26–28]. Motivated by this fact, in this paper a loosely–coupled multi–modal sensor fusion approach is proposed that utilizes visual/thermal data in a complementary manner with LiDAR data to enable reliable and resilient robot pose in challenging, GPS–denied and perception–degraded environments, especially in LiDAR degenerate environments.

III. Proposed Method

LiDAR based robot pose estimation methods generally operate by minimizing the distance between successive pointcloud scans. In its simplest form this scan–to–scan matching can be be performed by considering all the points in both scans and iteratively determining the transformation that minimizes point–to–point distance between two scans, with this method being generally referred to as Iterative Close Point (ICP) matching. However, considering all points equally for distance minimization is not ideal as all points may not be equally informative and able to produce a valid solution. For example, a point lying on a planar surface is not geometrically constrained in its local neighborhood and can be falsely matched to its neighboring points with low error, hence leading to an incorrect solution for the overall robot pose estimation problem. Similarly, considering all scan points can also significantly increase the computational requirements making the approach not be suitable for real–time applications. Hence, as first mentioned in [29], it is recommended to prioritize the utilization of points that belong to a shape for the estimation of scan–to–scan transformation, as the utilization of a shape model as metric to match points can make the transformation estimation process more robust. Nevertheless, such an approach is only meaningful if shape metrics are known a priori. Therefore, for applications that require robots to navigate through previously unknown environments a generalize-able shape metric must be utilized. In this regard, two of the most commonly used metrics for scan–to–scan matching for robot pose estimation are point–to–line distance [30] and point–to–plane distance [31]. Both of these metrics have been utilized by the state–of–the–art LiDAR Odometry and Mapping (LOAM) method [20], and have demonstrated excellent large–scale localization and mapping results in real–time. However, in the absence of sufficient geometric constraints the underlying scan–to–scan matching problem can become ill–conditioned and lead to incorrect estimation of the robot pose. An demonstrative example is shown in Figure 2, where a robot is traversing through an underground highway underpass. Due to the smooth tube–like structure of the environment the odometry cannot be estimated correctly in the direction of motion along the principal direction of the highway underpass and results in an incorrect map of the environment. Similarly, an incorrect scan–to–map matching can also occur in environments without distinguished geometric features. An example of such a scenario is presented in Figure 3, where an aerial robot is traversing through an urban building corridor and an incorrect map of the environment is built due to the lack of defined geometric features in the environment.

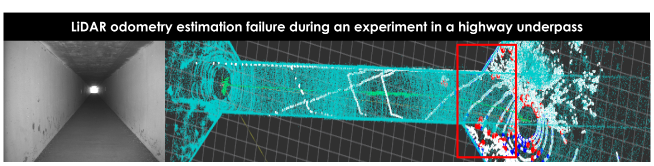

Fig. 2. The figure demonstrates the result of incorrect odometry estimation using LiDAR data during traversal of an underground highway underpass. The camera image on the left shows the lack of structure the present in this environment. The image on the right shows the top-down view of the incorrect map built of the environment due to failure to estimate odometry correctly with the red square highlighting the incorrect stitching of pointclouds at the exit of the highway underpass.

From the presented examples it can be understood that although such self similar and geometrically symmetric environments cannot provide sufficient geometric features for LiDAR odometry estimation to work reliably, yet other sensing modalities may still be used effectively for the estimation of robot pose. For this purpose, in this work we propose to loosely-fuse visual-inertial (VIO) and thermal-inertial (TIO) odometry estimates to improve the robustness of LiDAR methods. We utilize the state-of-the-art visual–inertial method ROVIO [32] to provide for utilizing visual images and modify it to operate on full radiometric thermal imagery, as detailed in [33]. Camera odometry estimates are typically obtained at a higher rate than pointcloud estimates and when a new pointcloud arrives we use camera odometry estimates to calculate the relative transformation between successive pointclouds and provide it as a prior for the scan–to–scan matching process. Provision of priors can improve the convergence rate of the pointcloud alignment process. As LiDAR odometry and mapping estimates are obtained by minimizing the distance between points during scan–to–scan and scan–to–map matching, respectively, iterative optimization processes are typically employed to determine the transformation that minimizes the residual error during the two matching steps. Given two corresponding points, based on an initial estimate of the transformation between two scans or between a scan and a map the distance, labelled as $measurement$ (p_{meas}) and $prediction$ (p_{pred}), where p_{meas} is the $3D$ position of the point in the most recent LiDAR scan and given the old position of the corresponding point (p_{old}), p_{pred} can be given as:

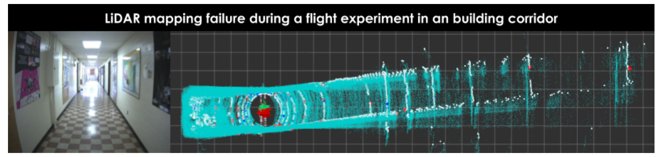

Fig. 3. The figure shows the incorrect LiDAR mapping result during an experiment where an aerial robot traverses through an urban building corridor. The image on the left shows the image from the on–board visual camera showing a straight corridor while the image on the right shows the incorrect LiDAR map with multiple scans combined together incorrectly due to lack of sufficient geometric features in the environment.

$$\mathbf{p}_{pred} = \mathbf{R}(\mathbf{p}_{old}) + \mathbf{t}, \tag{1}$$

where the $\mathbf{R}$ and $\mathbf{t}$ are the initial rotation and translation estimates for the alignment points between pointclouds. These estimates are refined iteratively through an optimization process where the residual distance between p_{pred} and p_{meas} is minimized, and given as:

$$\begin{aligned} \mathbf{r} &= \mathbf{p}_{pred} - \mathbf{p}_{meas}, \\ \mathbf{e}_{res} &= \sum_{i \epsilon P} \|\mathbf{r_i}\|_2, \end{aligned} \tag{2}$$

where $\mathbf{r}$ is the residual between two points and $\mathbf{e}_{res}$ is the residual cost function to be minimized and contains the squared sum of the residual distances for the set of all corresponding points P between two pointclouds. Given the relation above, the Gauss-Newton optimization function can be written as:

$$\mathbf{J^T J}\delta\mathbf{x} = -\mathbf{J^T r}, \tag{3}$$

where $\mathbf{J}$ is the Jacobian of the cost function and $\delta\mathbf{x}$ is the transformation increment. To understand if due to insufficient geometric constraints the underlying scan–to–scan or scan–to–map matching process have become degenerate, we evaluate the eigenvalues of the corresponding $\mathbf{J}^T\mathbf{J}$ matrices and determine if the underlying optimization process has become degenerate similar to the criteria proposed in [34]. In case the scan–to–scan matching fails and the LiDAR odometry process becomes degenerate we propagate the previous LiDAR odometry estimate by using the relative transform calculated from the camera odometry estimates. Furthermore, a separate evaluation is conducted for the scan–to–map matching process and in the event ill–conditioning is detected, the current pointcloud is integrated into the map using the prior LiDAR mapping estimate and the relative transform calculated using the camera odometry estimates. In addition, camera based odometry estimates are evaluated independently as a health check to determine if they are suitable to be utilized as a prior or for the propagation of LiDAR odometry and mapping estimates. For this purpose, the relative growth of the covariance matrix of camera odometry estimates is calculated using the D-Optimality criterion [35], which is given as:

$$D_{Optimality} = \exp(\log(\det(\Sigma)^{1/l})) \tag{4}$$

where Σ is the robot pose covariance matrix with dimensions $l \times l$. Furthermore, the relative translational and rotational pose estimates are individually checked to be within the allowed robot motion bounds and if successful utilized individually for both as a prior for LiDAR pose estimation process and for the propagation of previous LiDAR pose estimates in case of LiDAR degeneracy. An overview of the presented approach is shown in Figure 4.

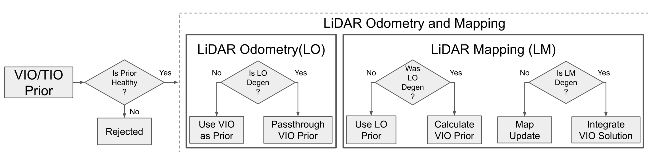

Fig. 4. The presented flowchart shows the process of the integration of Visual–Inertial (VIO) and Thermal–Inertial (TIO) Odometry estimates into the LiDAR Odometry and Mapping process. The relative VIO/TIO estimates are checked for quality before being utilized as priors for the scan–to–scan matching process. Both LiDAR odometry and mapping estimation process are checked individually for ill–conditioning and a VIO/TIO relative transform is utilized to propagate previous LiDAR estimates in case an ill-conditioning is detected

IV. Results

To evaluate the improved performance and reliability of the proposed method, initially an experiment is conducted in a structurally self similar environment presenting a case of an environment where LiDAR only pose estimation and mapping cannot function properly. Furthermore, the real world applications of the proposed method are demonstrated by utilizing the proposed solution during autonomous navigation and exploration missions conducted in an active underground mine as well as during the scored competition run of the tunnel circuit of the DARPA SubT challenge.

A. *Evaluation in a Structurally Self Similar Environment*

To demonstrate the improvement in the performance and reliability of the overall odometry estimation process, an experiment was conducted in a highly self similar environment using an aerial robot. In particular, a quadrotor based on the DJI Matrice M100 platform was utilized. The sensing suite carried on–board for robot pose estimation purposes consisted of a Velodyne PuckLITE LiDAR, providing pointclouds at 10Hz. VectorNav VN–100 IMU, providing inertial measurements at 200Hz, and a FLIR Blackfly camera with shutter–synchronized LEDs, providing images at 20Hz. All integrated sensors were interfaced with an Intel NUC-i7 (NUC7i7BNH) computer carried on–board the robot for performing robot pose estimation operations in real–time fully on-b-oard the robot. During the experiment, the robot traversed through a highway underpass which is highly self similar in terms of its structure, as shown in Figure 2. Due to the nature of the environment as well as the smooth concrete surfaces of the walls, not enough geometric features are provided to properly constraint the underlying LiDAR scan–to–scan matching optimization problem. As shown in Figure 5, LiDAR alone fails to estimate the robot odometry correctly and the built map is incorrect in size. However, it can be noted that with the proposed method integration of suitable VIO priors can aid the optimization process as well as selectively propagating the LiDAR estimates to continue the mapping process can improve the overall robot pose estimation process and result in building a correct map of such a highly self-similar environment.

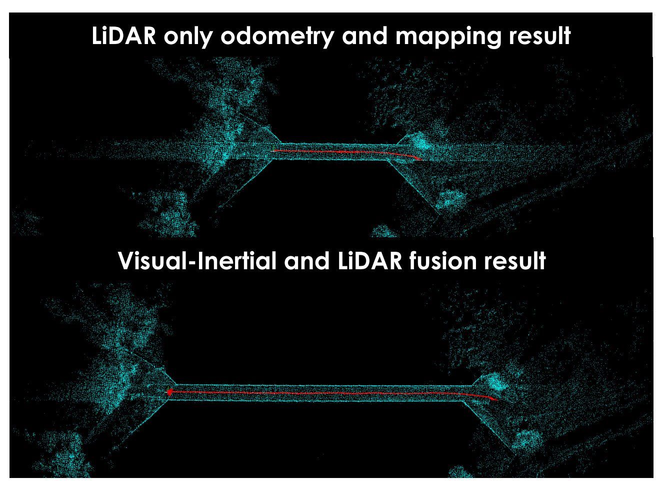

Fig. 5. This figure shows the comparison of performance of a LiDAR-only approach with the proposed approach in a self similar environment. The mapping result of the LiDAR-only approach is shown in the top row and depicts an incorrect map of the environment due to ill–conditioningg of the LiDAR odometry estimation problem. The bottom row shows the improved and correct map built using the proposed scheme.

B. *Exploration and Mapping of an Underground Mine*

To demonstrate the real–world applicability of the proposed approach an underground field deployment was conducted at the Turquoise Ridge Joint Venture (TRJV) underground mine in Winnemucca, Nevada. During this deployment an autonomous aerial exploration mission was conducted in the conditions of complete darkness and in the presence of heavy airborne dust. In this experiment, the aerial robot utilized a 64 beam Ouster OS1-64 LiDAR for laser odometry estimation and mapping purposes by using LOAM [20], aided with thermal–inertial odometry estimation provided by our proposed framework ROTIO [33] utilizing full radiometric imagery provided by a FLIR Tau2 thermal camera and inertial measurements from a VectorNAV VN–100 IMU. For the autonomous exploration mission the previously proposed Graph-Based exploration planning framework (GBPlanner) [36] was utilized. The aerial robot started its mission with an autonomous take–off in a dry part of the underground mine with heavy airborne dust present, it then traversed through a wet and uniform temperature area leading to a dead–end in the exploration. At this point, the robot made a 180 degree turn and proceeded with the exploration along the main mine drift and reaching a machine–shop like area with no admissible paths forward and hence backtracked its path to the initial take–off position and continued exploration of the mine drift until the robot battery was low. At this point the homing behaviour was triggered and the robot returned to its initial take–off position to land autonomously. The total distance traversed during this experiment was approximately 410 meters. In the absence of external ground–truth, the autonomous return to the take–off position is an indicator of minimal drift accumulation in the odometry estimation and mapping solution during the entirety of the mission. Qualitative localization and mapping results from the mission are presented in the Figure 6, with images from external visual camera and on–board thermal camera to provide an understanding of the perception degraded conditions encountered during the experiment.

C. Deployment in the DARPA Subterranean Challenge

The proposed approach was utilized as an on–board localization solution for an aerial autonomous exploration and mapping mission during the DARPA Subterranean Challenge tunnel Circuit held at National Institute for Occupational Safety and Health (NIOSH) - Pittsburgh Research Laboratory in August, 2019. The fully autonomous mission consisted of autonomous take–off, entrance gate detection, autonomous exploration and return to home. The robot navigated through the straight narrow tunnel structure relying on fusion of visual–inertial odometry estimates fused with LiDAR odometry estimation and returned to the same take–off spot at the completion of its exploration mission. The total path length traversed during the mission was approximately 190 meters. In the absence of external ground–truth, the autonomous return to take–off position is an indicator of minimal drift in the odometry estimation and mapping solution. Results from the mission are presented in Figure 7, and show the built map, robot trajectory for the full mission as well as indicative images from the on–board and external cameras to provide an understanding of the operational environment. The full video of the aerial deployment is available at `https://tinyurl.com/darpa-aerial-deployment` .

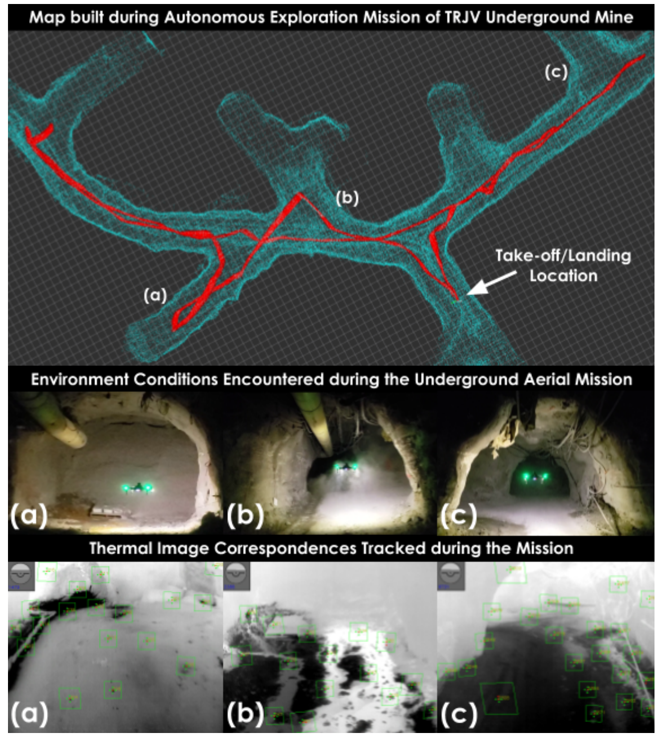

Fig. 6. The top row shows the LiDAR map built along with the path traversed during the autonomous exploration aerial mission at the TRJV underground mine. It can be noted that the take–off and landing positions of the robot are the same indicating minimal accumulated localization drift during the execution of the mission. The labels (a), (b) and (c) highlight instances during the mission, selected to indicate the dark and dust–filled operational conditions encountered during the mission. The middle and bottom row show external visual images of the robot and images from the on–board thermal camera, respectively, related to the selected mission instances.

V. CONCLUSIONS

In this work a complementary multi–modal senor fusion solution to enable resilient pose estimation for aerial robots in complex and challenging environments was presented. The proposed approach fused visual–inertial and thermal–inertial estimates with a LiDAR odometry and mapping solution to improve the overall reliability of the robot pose estimation process, especially in environments where the LiDAR pose estimation process becomes ill-conditioned. Through real–world experiments and robotic field deployments, conducted in diverse and challenging subterranean environments, the efficacy and applicability of the proposed approach was demonstrated.

ACKNOWLEDGMENT

This material is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) under Agreement No. HR00111820045. The presented content and ideas are solely those of the authors.

REFERENCES

[1] T. Tomic *et al.*, "Toward a fully autonomous uav: Research platform for indoor and outdoor urban search and rescue," *IEEE robotics & automation magazine*, vol. 19, no. 3, pp. 46–56, 2012.

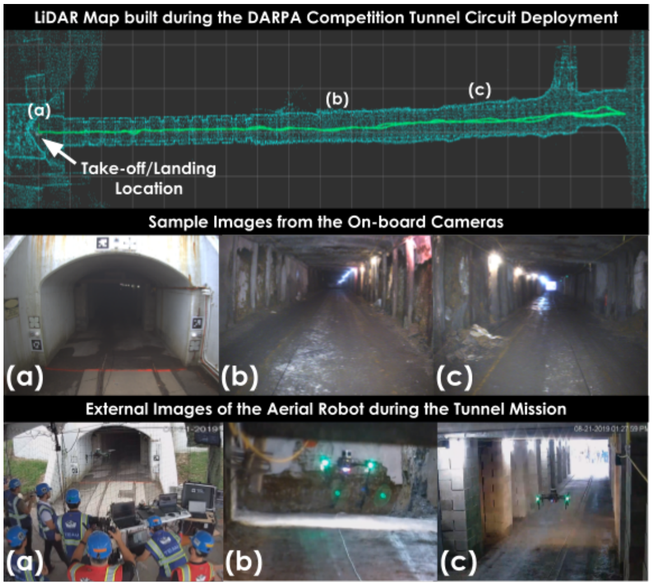

Fig. 7. The top row shows the map built during the autonomous aerial exploration mission during the DARPA Subterranean Challenge tunnel Circuit. It can be seen that once the robot reaches the end of the straight part of the tunnel it traverses back along the same path and autonomously lands at the take–off position. This in turn indicates very low drift in the odometry estimation during the mission. The labels (a),(b) and (c) are indicative instances chosen to present corresponding visual images from the on–board camera (middle row) and the external monitoring camera (bottom row) to provide an understanding of the operational environments and conditions during the mission. The brightness of images have been increased to facilitate better visibility of the environment.

[2] E. T. Alotaibi, S. S. Alqefari, and A. Koubaa, "Lsar: Multi-uav collaboration for search and rescue missions," *IEEE Access*, vol. 7, pp. 55 817–55 832, 2019.

[3] M. Erdelj, E. Natalizio, K. R. Chowdhury, and I. F. Akyildiz, "Help from the sky: Leveraging uavs for disaster management," *IEEE Pervasive Computing*, vol. 16, no. 1, pp. 24–32, 2017.

[4] S. Khattak, C. Papachristos, and K. Alexis, "Marker based thermal-inertial localization for aerial robots in obscurant filled environments," in *International Symposium on Visual Computing*, 2018.

[5] N. H. Motlagh, M. Bagaa, and T. Taleb, "Uav-based iot platform: A crowd surveillance use case," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 128–134, 2017.

[6] B. Grocholsky, J. Keller, V. Kumar, and G. Pappas, "Cooperative air and ground surveillance," *IEEE Robotics & Automation Magazine*, vol. 13, no. 3, pp. 16–25, 2006.

[7] S. Khattak, C. Papachristos, and K. Alexis, "Change detection and object recognition using aerial robots," in *International Symposium on Visual Computing*. Springer, 2016, pp. 582–592.

[8] A. J. Moore *et al.*, "Uav inspection of electrical transmission infrastructure with path conformance autonomy and lidar-based geofences nasa report on utm reference mission flights at southern company flights," 2017.

[9] V. Puri, A. Nayyar, and L. Raja, "Agriculture drones: A modern breakthrough in precision agriculture," *Journal of Statistics and Management Systems*, vol. 20, no. 4, pp. 507–518, 2017.

[10] C. Potena, R. Khanna, J. Nieto, R. Siegwart, D. Nardi, and A. Pretto, "Agricolmap: Aerial-ground collaborative 3d mapping for precision farming," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1085–1092, 2019.

[11] T. Dang, S. Khattak, F. Mascarich, and K. Alexis, "Explore locally, plan globally: A path planning framework for autonomous robotic exploration in subterranean environments," in *2019 19th International Conference on Advanced Robotics (ICAR)*, 2019, pp. 9–16.

[12] S. Khattak, C. Papachristos, and K. Alexis, "Vision-depth landmarks and inertial fusion for navigation in degraded visual environments," in *International Symposium on Visual Computing*. Springer, 2018, pp. 529–540.

[13] B. Rao, A. G. Gopi, and R. Maione, "The societal impact of commercial drones," *Technology in Society*, vol. 45, pp. 83–90, 2016.

[14] H. Shakhatreh, A. H. Sawalmeh, A. Al-Fuqaha, Z. Dou, E. Almaita, I. Khalil, N. S. Othman, A. Khreishah, and M. Guizani, "Unmanned aerial vehicles (uavs): A survey on civil applications and key research challenges," *IEEE Access*, vol. 7, pp. 48 572–48 634, 2019.

[15] C. Papachristos, S. Khattak, F. Mascarich, and K. Alexis, "Autonomous navigation and mapping in underground mines using aerial robots," in *2019 IEEE Aerospace Conference*. IEEE, 2019, pp. 1–8.

[16] "Darpa subterranean (subt) challenge," https://www.darpa.mil/program/darpa-subterranean-challenge, 2018-2021.

[17] C. Kanellakis and G. Nikolakopoulos, "Survey on computer vision for uavs: Current developments and trends," *Journal of Intelligent & Robotic Systems*, vol. 87, no. 1, pp. 141–168, 2017.

[18] T. Qin and S. Shen, "Robust initialization of monocular visual-inertial estimation on aerial robots," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 4225–4232.

[19] C. Papachristos, F. Mascarich, S. Khattak, T. Dang, and K. Alexis, "Localization uncertainty-aware autonomous exploration and mapping with aerial robots using receding horizon path-planning," *Autonomous Robots*, vol. 43, no. 8, pp. 2131–2161, 2019.

[20] J. Zhang and S. Singh, "Loam: Lidar odometry and mapping in real-time," in *Robotics: Science and Systems Conference*, Pittsburgh, PA, July 2014.

[21] T. Shan and B. Englot, "Lego-loam: Lightweight and ground-optimized lidar odometry and mapping on variable terrain," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 4758–4765.

[22] H. Ye, Y. Chen, and M. Liu, "Tightly coupled 3d lidar inertial odometry and mapping," in *2019 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2019.

[23] S. Khattak, C. Papachristos, and K. Alexis, "Visual–thermal landmarks and inertial fusion for navigation in degraded visual environments," in *2019 IEEE Aerospace Conference*. IEEE, 2019.

[24] T. Mouats, N. Aouf, L. Chermak, and M. A. Richardson, "Thermal stereo odometry for uavs," *IEEE Sensors Journal*, vol. 15, no. 11, pp. 6335–6347, 2015.

[25] S. Khattak, C. Papachristos, and K. Alexis, "Keyframe-based thermal–inertial odometry," *Journal of Field Robotics*, vol. 37, no. 4, pp. 552–579, 2020.

[26] T. Lowe, S. Kim, and M. Cox, "Complementary perception for handheld slam," *IEEE Robotics and Automation Letters*, vol. 3, no. 2, pp. 1104–1111, 2018.

[27] X. Zuo, P. Geneva, W. Lee, Y. Liu, and G. Huang, "Lic-fusion: Lidar-inertial-camera odometry," *arXiv preprint arXiv:1909.04102*, 2019.

[28] Y.-S. Shin and A. Kim, "Sparse depth enhanced direct thermal-infrared slam beyond the visible spectrum," *arXiv preprint arXiv:1902.10892*, 2019.

[29] P. J. Besl and N. D. McKay, "Method for registration of 3-d shapes," in *Sensor fusion IV: control paradigms and data structures*, vol. 1611. International Society for Optics and Photonics, 1992, pp. 586–606.

[30] A. Censi, "An icp variant using a point-to-line metric," in *2008 IEEE International Conference on Robotics and Automation*. Ieee, 2008, pp. 19–25.

[31] A. Segal, D. Haehnel, and S. Thrun, "Generalized-icp." in *Robotics: science and systems*, vol. 2, no. 4. Seattle, WA, 2009, p. 435.

[32] M. Bloesch, M. Burri, S. Omari, M. Hutter, and R. Siegwart, "Iterated extended kalman filter based visual-inertial odometry using direct photometric feedback," *The International Journal of Robotics Research*, vol. 36, no. 10, pp. 1053–1072, 2017.

[33] S. Khattak, F. Mascarich, T. Dang, C. Papachristos, and K. Alexis, "Robust thermal-inertial localization for aerial robots: A case for direct methods," in *2019 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 2019, pp. 1061–1068.

[34] J. Zhang, M. Kaess, and S. Singh, "On degeneracy of optimization-based state estimation problems," in *Robotics and Automation (ICRA), 2016 IEEE International Conference on*. IEEE, 2016, pp. 809–816.

[35] H. Carrillo, I. Reid, and J. A. Castellanos, "On the comparison of uncertainty criteria for active slam," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2012.

[36] T. Dang, F. Mascarich, S. Khattak, C. Papachristos, and K. Alexis, "Graph-based path planning for autonomous robotic exploration in

subterranean environments," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019, pp. 3105–3112.