1 Introduction

Motivation:

Disaster response and search and rescue missions are among the most difficult missions in which an autonomous robot can be deployed [13]. Missions such as these require a robot to autonomously navigate chaotic, unstructured environments. This is normally done with the aid of visual and lidar sensors. However, the quality of state estimates from visual and lidar methods quickly degrades in the presence of smoke, fog, and other visually degraded environments (VDEs) commonly encountered in disaster response missions as shown in figure 1. Thus, it is clear that there exists a need for a reliable and efficient method for state estimation and perception that is robust to VDEs.

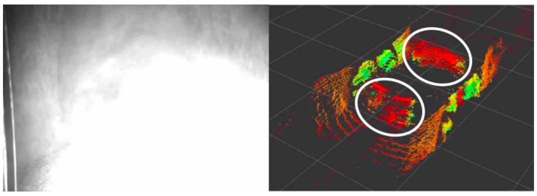

Fig. 1 In the presence of dust, measurements from a monocular camera (left) are saturated. VLP-16 lidar scans (right) are corrupted with unbearable noise (white circles).

Problem Statement:

The problem addressed by this work is that of controlling a small flying robot and avoiding obstacles in dense fog. To this end we present a novel radar-inertial state estimation and mapping technique including two unique features: an improved radar sensor model to more accurately model the sensor's physics and a mapping method incorporating a learned noise filter and a radar-specific sensor model.

Related Work:

Several other methods have been proposed for control and planning in VDEs. Most notably, [7] proposes a navigation system using thermal imaging and inertial measurements. This is a promising area of research but there is another sensing option that is well suited for perception in VDEs: radar.

Radar is well-established in the automotive industry but has largely been ignored in robotics. Only a few works exist using radar as a primary sensor for state estimation and perception. [1] presents methods for radar odometry in challenging conditions; this work is extended in [2]. Additionally [11] and [10] have presented methods for radar-based SLAM using arrays of automotive-grade radar sensors.

Contributions:

Our method represents a significant advance over previous methods in several important ways. First, our method is applicable to 3D environments and is thus usable on flying vehicles. Second, our method fuses radar and inertial measurements; previous radar state estimation methods have used radar as their sole sensor and robotic sensor fusion methods with radar have only been minimally explored. Lastly, our method requires only one single-board radar sensor and an IMU, making sensor package simpler than previous methods which have required mechanically-scanned sensors or arrays of several sensors.

2 Technical Approach

State Estimation

Our approach estimates the global-frame velocity $\mathbf{v}_W$ and orientation $\mathbf{q}_{WS}$ of the sensor platform over a sliding window [12] of K successive radar measurements by minimizing the cost function

$$J(\mathbf{x}) := \underbrace{\sum_{k=1}^{K} \sum_{d \in \mathcal{D}^k} \mathbf{e}_d^T \mathbf{e}_d}_{\text{Doppler term}} + \underbrace{\sum_{k=1}^{K-1} \mathbf{e}_s^{k\,T} \mathbf{W}_s^k \mathbf{e}_s^k}_{\text{inertial term}} \tag{1}$$

where K is the number of past radar measurements, $\mathcal{D}^k$ is the set of targets in the radar scan at time k, $\mathbf{e}_d$ is the Doppler error, and $\mathbf{e}_s$ is the IMU error.

A radar measurement consists of a set of targets $\mathcal{D}$. Each $d \in \mathcal{D}$ consists of $[\tilde{r}_S, \tilde{v}_R, \tilde{\theta}_S, \tilde{\phi}_S]^T$, the range, Doppler, azimuth, and elevation for target d. We assume targets detected by the radar sensor are stationary while the sensor is moving. Doppler measurements are modeled as

$$\tilde{v}_R - e_v = \mathbf{C}_{SW} \mathbf{v}_W \left(\frac{\tilde{\mathbf{r}}_S - \mathbf{e}_r}{\|\tilde{\mathbf{r}}_S - \mathbf{e}_r\|_2} \right)^T \tag{2}$$

where e_v is the error in the Doppler measurement, $\mathbf{C}_{SW}$ is a rotation matrix representing the orientation of the sensor, $\mathbf{v}_w$ is the estimated velocity of the sensor, $\tilde{\mathbf{r}}_S$ is the ray from the sensor to the target, and $\mathbf{e}_r$ is the estimated error in that ray. The full Doppler error is $[e_v \ \mathbf{e}_r^T]^T$, where $\mathbf{e}_r$ is both a state to be estimated and an error to be minimized as in a total least squares problem [3].

The inertial term combines IMU measurements grouped into a pre-integrated factor as described in [4]. This is similar the inertial term used in [8], with the exception that the sensor platform's position is not estimated, only the velocity, orientation, and IMU biases.

Radar Mapping

Using radar scans with common occupancy mapping methods such as octomap [5] creates two problems: 1) the beam sensor model used in octomap does not accurately model the operation of the radar sensor, and 2) radar is particularly prone to false positives from noise and multipath reflections. To address the first of these problems we employ a new, radar-specific sensor model. To combat the noise problem, we developed a learned radar noise filter. This entails training the PointNet segmentation network [9] to label radar points as either true readings or noise using a new, weakly-supervised training method. This allows us to accurately filter out erroneous readings and only add true measurements to our map. Lastly, we have developed a method to detect multipath reflections and condense them into a single, accurate reading. Together these three innovations allow us to accumulate noisy radar returns into a dense, accurate map of a robot's environment.

3 Experiments

We have already done quantitative comparisons of our state estimation and mapping methods against other popular methods. All that remains is to demonstrate our method's use in closed-loop control of a micro-aerial vehicle, and compare radar-based MAV control to popular visual methods in environments with and without visually degraded conditions. The complete list of our experiments is laid out in table 2.

Completed Experiments

We have done quantitative comparisons between our radar-inertial state estimator's performance and that of visual-inertial and lidar methods using a hand-carried sensor rig and a vicon motion capture system as groundtruth. For the visual-inertial benchmark we used the popular Intel Realsense T265 tracking camera, and for lidar we used a velodyne VLP-16 with the LAMP odometry system [6]. These comparisons were done indoors with bright ambient light and no fog or other visual obscurants. We also did qualitative comparisons of radar-inertial, visual-inertial, and lidar state estimators in varying fog density. Fog density was quantified by visibility distance and varied from 12m to less than 6m. The test space is shown in no fog and heavy fog conditions in figure 2.

Fig. 2 The lab space used for evaluation of the state estimation and mapping methods with a handheld sensor rig. Shown with no fog on the left, and high fog on the right.

(a) No fog, unlimited visibility (b) High fog, 6m visibility

Further, we quantitatively compared visual mapping methods with radar using lidar maps as groundtruth in conditions with no fog. Lastly, we made quantitative comparisons between radar, visual, and lidar mapping in varying fog conditions.

Scheduled Experiments

We plan to further evaluate our radar-based state estimation and mapping methods with the following tests.

- Demonstrate the use of radar-based state estimation and mapping for closed-loop control of an autonomous micro-aerial vehicle (MAV)
- Compare accuracy of closed-loop control with radar methods to visual methods both with and without fog.

Through these tests we intend to demonstrate that our radar-based state-estimation and mapping methods allow for autonomous operation of MAVs in conditions that would be prohibitive for the current state of the art in visual methods.

4 Results

Table 1 shows the mean and standard deviation of the RMS error in velocity estimates for our radar-inertial state estimator as well as visual-inertial and lidar odometry methods in a brightly-lit indoor environment with no fog. These results show the accuracy of our radar-inertial state estimator is on par with the state of the art in visual-inertial and lidar-based methods when tested in conditions that are favorable to all three methods.

	Radar-Inertial		Visual-Inertial		Lidar	
	μ	σ	μ	σ	μ	σ
x	0.2576	0.0935	0.2019	0.0898	0.2270	0.1197
y	0.2272	0.0827	0.1830	0.0845	0.2271	0.1325
z	0.1125	0.0268	0.0809	0.0124	0.2054	0.0649

Table 1: Mean (μ) and standard deviation (σ) in m/s of the RMS velocity estimate errors in x, y, and z (over 10 runs of approx 60s each) for each estimator tested in no fog. Ground truth is provided by a motion capture system.

Because fog prevents the use of motion capture, we do not have groundtruth for our state estimator tests with fog. We can make a qualitative evaluation, however. In low to medium fog states estimated by visual-inertial and radar-inertial methods tend to agree while lidar does not. This indicates lidar state estimation is adversely affected by even low fog levels while visual and radar-based methods continue to function mostly normally. As fog levels increase visual and radar-based methods begin to disagree. It is not possible to determine whether the radar or visual method is failing in this case, but the properties of the radar sensor would lead one to believe it should be more robust to fog than vision. This will be definitively demonstrated when we compare vision and radar methods' use in closed-loop control of an autonomous MAV in foggy conditions.

We have also done qualitative comparisons between occupancy mapping using our radar method, lidar, and an Intel Realsense d435 depth camera. As expected, lidar clearly outperforms both radar and vision when mapping in conditions with no fog. In these conditions our radar mapping method's accuracy is quite similar to mapping with the d435, although the d435 does provide denser maps. When low levels of fog are introduced visual mapping is degraded to the point of uselessness. Lidar is somewhat robust to lower fog, but higher fog levels render lidar mapping completely unusable as well. Our radar mapping method, on the other hand, is entirely unaffected by even the most dense fog in which we tested. This demonstrates our radar mapping method's advantages in visually degraded conditions. These comparisons are demonstrated in the attached video.

We have thus far only done qualitative comparisons between mapping methods. In our final submission we will include quantitative tests to rigorously compare the accuracy of each mapping method in terms of both false positives and false negatives.

5 Experimental Insights

This work presents methods for state estimation and perception with a radar-inertial sensor platform. We estimate the sensor platform's velocity and orientation by fusing Doppler velocity measurements from an SoC millimeter wave radar sensor with inertial measurements from an IMU. This work demonstrates other popular robotic perception methods, vision and lidar, are incapacitated by conditions likely to arise in a disaster response scenaroi like dense fog, but radar is unaffected by these conditions. Lastly, the radar-inertial sensor suite is lightweight and has low power requirements making it an attractive alternative for platforms with constraints on their payload and power.

The experiments we have already completed demonstrate the accuracy of the proposed state estimation method in indoor experiments via comparisons with a motion capture system, a commercial VIO system, and a lidar-based system. Our experiments demonstrate that the proposed method is comparable to VIO and lidar methods for state estimation in conditions favorable to VIO and lidar. Further, the our proposed radar-inertial state estimation method's accuracy far exceeds that of VIO and lidar when dense fog is introduced.

We also demonstrate the accuracy of our proposed radar-based mapping method is comparable to mapping with a depth camera in typical, non-DVE scenarios. We also demonstrate that lidar and depth-camera-based mapping methods are incapacitated by fog while our method is completely unaffected. This is well demonstrated by the included video, which shows comparisons between each mapping method in conditions with and without fog. It also demonstrates our full state estimation and mapping pipeline in use with a hand-carried sensor rig.

All of this gives us sufficient reason to believe our radar-based state estimation and mapping methods will perform well when used for autonomous flight of an MAV in dense smoke and fog. Our completed and planned experiments are laid out in table 2. Through these experiments we hope to show that our method's performance is similar to vision and lidar-based methods when conditions are favorable, and that our method is unaffected by conditions in which vision and lidar are unable to operate.

experiment	**planned or completed**
quantitative state estimator tests	completed
qualitative state estimator tests in fog	completed
quantitative mapping comparison	completed
qualitative mapping comparison in fog	completed
radar-inertial MAV control demo	planned
comparison radar and vision-based MAV control	planned

Table 2: Experiments completed and scheduled

References

[1] Cen, S.H., Newman, P.: Precise Ego-Motion Estimation with Millimeter-Wave Radar Under Diverse and Challenging Conditions. In: 2018 IEEE International Conference on Robotics and Automation (ICRA). IEEE, Brisbane, QLD (2018)

[2] Cen, S.H., Newman, P.: Radar-only ego-motion estimation in difficult settings via graph matching. arXiv:1904.11476 [cs] (2019). URL http://arxiv.org/abs/1904.11476

[3] Deming, W.: Statistical Adjustment of Data. J. Wiley & Sons, Incorporated (1943). URL https://books.google.com/books?id=9awgAAAAMAAJ

[4] Forster, C., Carlone, L., Dellaert, F., Scaramuzza, D.: On-manifold preintegration for real-time visual–inertial odometry. Trans. Rob. **33**(1), 1–21 (2017)

[5] Hornung, A., Wurm, K.M., Bennewitz, M., Stachniss, C., Burgard, W.: OctoMap: An efficient probabilistic 3D mapping framework based on octrees. Autonomous Robots (2013). Software available at http://octomap.github.com

[6] Kamak Ebadi Yun Chang, M.P.A.S.A.H.E.H.A.T.B.M.L.C., akbar Aghamohammadi, A.: Lamp: Large-scale autonomous mapping and positioning for exploration of perceptually-degraded subterranean environments. In: 2020 IEEE International Conference on Robotics and Automation (ICRA). Paris, France (2020)

[7] Khattak, S., Papachristos, C., Alexis, K.: Keyframe-based direct thermal-inertial odometry. CoRR **abs/1903.00798** (2019). URL http://arxiv.org/abs/1903.00798

[8] Leutenegger, S., Lynen, S., Bosse, M., Siegwart, R., Furgale, P.: Keyframe-based visual–inertial odometry using nonlinear optimization. The International Journal of Robotics Research **34**(3), 314–334 (2015)

[9] Qi, C.R., Su, H., Mo, K., Guibas, L.J.: PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. arXiv:1612.00593 [cs] (2017). URL http://arxiv.org/abs/1612.00593. ArXiv: 1612.00593

[10] Schuster, F., Keller, C.G., Rapp, M., Haueis, M., Curio, C.: Landmark based radar SLAM using graph optimization. In: 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC), pp. 2559–2564 (2016)

[11] Schuster, F., Wörner, M., Keller, C.G., Haueis, M., Curio, C.: Robust localization based on radar signal clustering. pp. 839–844 (2016)

[12] Sibley, G.: A Sliding Window Filter for SLAM p. 17

[13] Tomic, T., Schmid, K., Lutz, P., Dömel, A., Kassecker, M., Mair, E., Grixa, I., Ruess, F., Suppa, M., Burschka, D.: Toward a fully autonomous uav: Research platform for indoor and outdoor urban search and rescue. Robotics Automation Magazine, IEEE **19**, 46 –56 (2012)