

LIRO: Tightly Coupled Lidar-Inertia-Ranging Odometry

Thien-Minh Nguyen, Muqing Cao, Shenghai Yuan, Yang Lyu, Thien Hoang Nguyen, and Lihua Xie

Abstract—In recent years, thanks to the continuously reduced cost and weight of 3D lidar, the applications of this type of sensor in the community have become increasingly popular. Despite many progresses, estimation drift and tracking loss are still prevalent concerns associated with these systems. However, in theory these issues can be resolved with the use of some observations to fixed landmarks in the operation environments. This motivates us to investigate a sensor fusion scheme of lidar and inertia measurements with Ultra-Wideband (UWB) range measurements to such landmarks, which can be easily deployed in the environments with minimal cost and time. Hence, data from IMU, lidar and UWB are tightly-coupled with the robot's states on a sliding window based on their timestamps. Then, we construct a cost function comprising of factors from UWB, lidar and IMU preintegration measurements. Finally an optimization process is carried out to estimate the robot's position and orientation. It is demonstrated through some real world experiments that the method can effectively resolve the drift issue, while only requiring two or three anchors deployed in the environment.

I. INTRODUCTION

Localization is a crucial task that requires a lot of attention and effort in mobile robotics [1]. Depending on the scenario, one has to factor in the trade-offs among cost, accuracy, robustness, computational load, and ease of deployment to settle on the most appropriate method for the robot at hand. For most applications, selecting the set of sensors is usually the first concern that has to be addressed, especially in GPS-denied environments. In some cases, one can employ some artificially installed sensors such as motion-capture [2], [3], visual markers [4], [5], or Ultra-wideband (UWB) beacons [6]–[10]. On the other hand, to operate in complex and uncontrolled GPS-denied environments, Onboard Self-Localization (OSL) methods such as visual-inertial odometry (VIO) [11]–[18] or Lidar Odometry and Mapping (LOAM) [19]–[22] techniques are often the most viable option for robot localization. Indeed, it has become increasingly clear that OSL systems are going to be the backbone of autonomous navigation for the years to come.

Among OSL methods, 3D lidar based approaches are gaining more and more popularity thanks to the continuously reduced cost and weight of this type of sensor in recent years. Originally weighing over 10 kg and costing tens of thousands of USDs, lidar could only be used in large autonomous cars. In contrast, currently some commercial products only weigh

This work was partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation, under the Wallenberg-NTU Presidential Postdoctoral Fellowship Program.

The authors are with School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, 50 Nanyang Avenue (Corresponding author's e-mail: thienminh.npn@ieee.org)

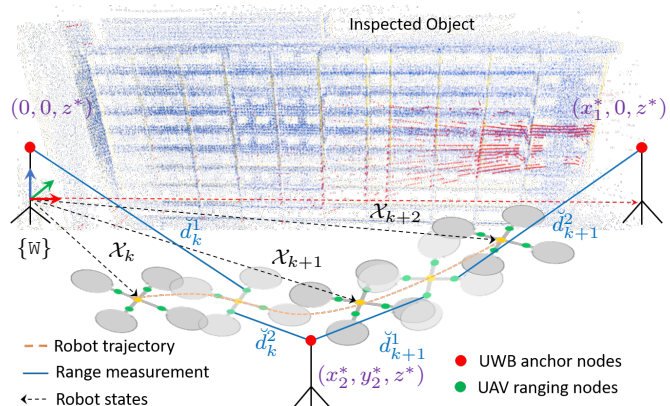


Fig. 1: Illustration of the ranging scheme to fixed landmarks deployed in the environment relative to the object being inspected. The time index k of the distance sample refers to the period $(t_k, t_{k+1}]$ in which it is obtained, not the exact time instance t_k .

for a few hundred grams and cost only a few hundred USDs. Even for Micro/Unmanned Aerial Vehicles (MAVs / UAVs), where the payload capacity is limited, successful uses of lidar have been demonstrated in recent years [19], [23]–[25]. Compared to common camera-based systems, a lidar-based solution offers several advantages. For instance, lidar can directly observe the metric scale features, hence, it can directly provide self-localization information (odometry) for the robot, while simple monocular camera-based solutions only provide odometry data of ambiguous scale. Moreover, even when compared with RGBD or stereo cameras, which are on the same footing with lidar as they can detect metric-scale features, lidar has a much higher sensing range, besides being almost invariant to illumination conditions.

Despite the above advantages, LOAM, being an OSL method, is still prone to estimation drift and loss of tracking due to lack of features in an environment. In addition, under the OSL approach, robots can only estimate their positions relative to their initial location. These issues can be quite inconvenient in applications such as inspection of 3D structures, where predefined trajectories are often desired to be executed in reference to a chosen so-called world frame. One solution can be to fuse GPS poses with LOAM [26], however it can only be effective in wide open areas, since GPS will strongly degrade in clustered and dense urban environments. Another approach could be to train a neural network on recognizing segmented objects from the pointcloud for place recognition and loop closure [27]. Obviously this approach also requires significant effort in

collecting data, labelling and training, not to mention that future changes in the environment can cause the system to lose effectiveness.

In this paper, we posit that by using ranging measurements from the robot to two or three UWB anchors, which can be easily deployed in the environment, long-term drift in both position and orientation can be effectively eliminated. Moreover, the pose (i.e. both position and orientation) estimate can also be referenced in the desired coordinate system defined via the anchor deployment. In this case, we say that our lidar-based odometry estimate is *drift-free* and *global*. The main contributions of this work can be listed as follows:

- We integrate a flexible ranging scheme with a lidar-inertial odometry system to achieve drift-free and global pose estimates.
- We develop an efficient tightly coupled sensor fusion framework to combine the so-called body-offset ranges, IMU preintegration, and lidar-based features for real time estimation of the robot states.
- We demonstrate the efficacy of estimation scheme via real world experiments in practical scenarios, especially on a high-dynamics platform such as UAV.

The remaining of this paper is organized as follows: in Section II, we review some related works to highlight the novelty of our approach; Section III introduces some preliminaries. Section IV presents an overview of the methodology. Section V provides a more detailed description of the cost factors and Section VI presents our experiment results. Section VII concludes our work.

II. RELATED WORKS

Indeed, in recent years, many researchers have employed ranging to UWB anchors to handle the estimation drift of OSL methods. For example, in [28], a loosely coupled sensor fusion scheme was proposed to leverage UWB ranges in correcting a pointcloud-based SLAM system. As the term loosely coupled suggests, this scheme fuses the localization information from two relatively independent systems. As such, if there are only two or three anchors, the UWB-based localization thread may not be realizable. We also note that since there is no observation to connect the consecutive poses in the sliding window, this method employs so-called *smoothness factors*, which are based on the maximum velocity between consecutive ranging times. Thus when the real velocity actually exceeds the assumed maximum velocity, the estimate can exhibit “lagging” behaviour of a low-pass filter. This can be resolved with the use of IMU preintegration, which is one of the key techniques featured in our method.

In [29], a loosely coupled method was developed to correct the scale of a monocular visual odometry (VO) output using the distance measurements from a single UWB anchor. In this method the anchor’s position is also estimated with respect to the coordinate frame of the VO system, whose origin coincides with the initial position of the camera. Hence the estimate is still of local type (it will change when the VO system initiates at a new position). In [30], [31], a tightly coupled scheme was investigated, where UWB

ranging and ORB features are combined into a cost function, which is then optimized to provide a metric-scaled position estimate. This tightly coupled scheme requires a number of anchors less than four. Note that the estimate is still of local type. In [32], the authors proposed a tightly coupled visual-inertial-ranging sensor fusion scheme. Only one anchor is used in this case, but multiple robots can also exchange range to jointly estimate their relative position. However, we note that this system still has to employ smoothness factors, along with being of local type. Finally, in [33], a loosely coupled approach was proposed to combine 2D lidar odometry estimate with UWB ranges to jointly localize the robot and the UWB anchor positions. We consider that this method is of global type, though the loosely coupled approach would require large number of UWB anchors, besides only providing 2D localization estimate. We also note that all of the aforementioned works only focused on the position estimation, while orientation state was ignored. This is expected since the UWBs ranging were only conducted in a body-centered manner.

To the best of our knowledge, our work is the first that investigates the tightly coupled sensor fusion approach of 3D Lidar, IMU and UWB ranges (hence the acronym LIRO for Lidar-Intertia-Ranging Odometry). Moreover, we also employ a novel body-offset ranging model that couples the position, orientation and velocity in the range observations. In addition, by using the communication capability of the UWB, and a very simple anchor deployment scheme, we can straightforwardly estimate the anchor position and start the localization in a user-defined frame, instead of having to estimate the anchor positions in a random frame by using VO data. This makes the LIRO scheme a flexible and accurate solution for localization of robot in inspection tasks.

III. PRELIMINARIES

A. Notations

In this paper we use $(\cdot)^\top$ to denote the transpose of a vector or matrix under (\cdot) . For a vector $\mathbf{x} \in \mathbb{R}^m$, $\|\mathbf{x}\|$ stands for its Euclidean norm, and $\|\mathbf{x}\|_G^2$ is short-hand for $\|\mathbf{x}\|_G^2 = \mathbf{x}^\top \mathbf{G} \mathbf{x}$. For two vectors $\mathbf{v}_1, \mathbf{v}_2$, $\mathbf{v}_1 \times \mathbf{v}_2$ denotes their cross product. In the sequel, we shall denote $\mathbf{R} \in \text{SO}(3)$ as the rotation matrix, and $\mathbf{T} \in \text{SE}(3)$ as the transformation matrix. $\text{Log}(\mathbf{R})$ returns the rotation vector of \mathbf{R} . For a unit quaternion \mathbf{q} , $\mathcal{R}(\mathbf{q})$ is its corresponding rotation matrix, and $\text{vec}(\mathbf{q})$ returns its vector part.

To make explicit that a vector \mathbf{v} , or points in a pointcloud \mathbf{F} , are w.r.t. a coordinate frame $\{\mathbf{A}\}$, we attach a left superscript $^{\mathbf{A}}$ to \mathbf{v} or \mathbf{F} , e.g. $^{\mathbf{A}}\mathbf{v}$, $^{\mathbf{A}}\mathbf{F}$. A rotation matrix and transformation matrix between two reference frames are denoted with the frames attached as the left-hand-side superscript and subscript, e.g. $^{\mathbf{A}}_{\mathbf{B}}\mathbf{R}$ and $^{\mathbf{A}}_{\mathbf{B}}\mathbf{T}$ are the rotation and transform matrices from frame $\{\mathbf{A}\}$ to $\{\mathbf{B}\}$, respectively. When the coordinate frames are the body frame at different time instances, we may also ignore the superscript and subscripts, e.g. ${}^k\mathbf{R}_{k+1} \triangleq {}^{\mathbf{B}_k}_{\mathbf{B}_{k+1}}\mathbf{R}$, or ${}^w\mathbf{T} \triangleq {}^{\mathbf{B}_w}_{\mathbf{B}_m}\mathbf{T}$. For a list of vectors $\mathbf{v}_1, \mathbf{v}_2 \dots, \mathbf{v}_n$ (including scalar) we may write $(\mathbf{v}_1, \mathbf{v}_2 \dots, \mathbf{v}_n)$ as a short-hand for $[\mathbf{v}_1^\top, \mathbf{v}_2^\top \dots, \mathbf{v}_n^\top]^\top$.

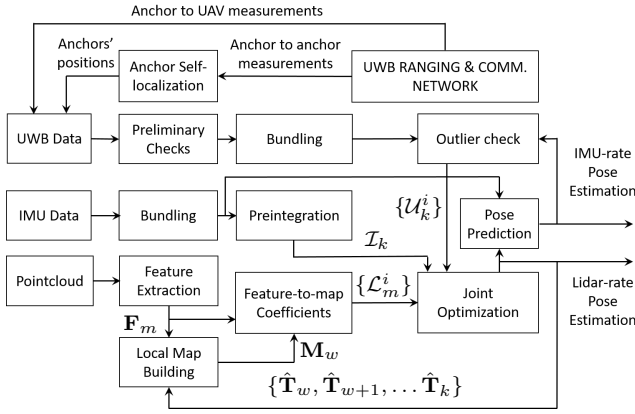


Fig. 2: The Lidar-Inertia-Ranging Odometry framework. A snapshot of the main quantities at time t_{k+1} when a new pointcloud message just arrives is overlaid on the connections.

B. State estimates

We define the robot states to be estimated at time t_k as:

$$\mathcal{X}_k = (\mathbf{q}_k, \mathbf{p}_k, \mathbf{v}_k, \mathbf{b}_k^\omega, \mathbf{b}_k^a), \quad (1)$$

where $\mathbf{q}_k, \mathbf{p}_k, \mathbf{v}_k \in \mathbb{R}^3$ are respectively the orientation quaternion, position and velocity w.r.t. the world frame $\{\mathcal{W}\}$ at time t_k ; $\mathbf{b}_k^a, \mathbf{b}_k^\omega \in \mathbb{R}^3$ are respectively the IMU accelerometer and gyroscope biases. Note that the world frame $\{\mathcal{W}\}$ is defined by the user through the deployment of the anchors, as explained in Section IV-A. We denote the state estimate at each time step k , and the sliding windows as follows:

$$\hat{\mathcal{X}}_k = (\hat{\mathbf{q}}_k, \hat{\mathbf{p}}_k, \hat{\mathbf{v}}_k, \hat{\mathbf{b}}_k^\omega, \hat{\mathbf{b}}_k^a), \quad (2)$$

$$\hat{\mathcal{X}} = (\hat{\mathcal{X}}_k, \hat{\mathcal{X}}_{k+1} \dots \hat{\mathcal{X}}_{k+M}), \quad (3)$$

where $M \in \mathbb{N}$ is the number of steps in the sliding windows. We choose $M = 10$ for all of the experiments in this work.

Note that in this work the extrinsic parameters have been manually calibrated and set as prior. Indeed, we find introducing these extra states to the problem yields inconclusive benefit while also slowing down the convergence rate.

IV. LIDAR-INERTIAL-RANGING ODOMETRY FRAMEWORK

Fig. 2 provides an overview of our framework, where each main process will be described in more details below.

A. UWB ranging and communication network

At the top of Fig. 2, we can see a UWB ranging and communication network, which allows us to measure the distances between the anchor nodes, and between anchor nodes and the ranging nodes on the robot. Fig. 1 illustrates the ranging scheme in more details. First, we have a set of anchors deployed in the area of operation as fixed landmarks. To keep this task simple, we only focus on the scenarios with two or three anchors having the same height from the ground, denoted as z^* . It can be seen that two such anchors are already sufficient to define a coordinate system. Specifically, anchor 0 can be assumed to be at the $(0, 0, z^*)$ position,

while the direction from anchor 0 to anchor 1 shall define the $+x$ direction. Hence, anchor 1's coordinate can be set as $(x_1^*, 0, z^*)$, where x_1^* is obtained by having anchor 1 directly range to anchor 0 multiple times, broadcast these measurements to the robot, and take the average as x_1^* . The position of the third anchor can also be calculated by simple geometry. Note that ranging and communicating capabilities are readily supported by the UWB nodes used in this work.

The next feature in the ranging scheme is the UWB ranging nodes on the robot. In this work, multiple UWB nodes are installed on the robot, whose position in the robot's body frame is known. This body-offset ranging scheme is an important feature that allows the ranges measurements to be coupled with the robot orientation, hence allowing the orientation estimate to be of global type.

B. UWB measurement workflow

After the anchor positions have been determined, the robot-to-anchor ranges can be used for the estimation process. First, they are put through some preliminary checks based on the signal over noise ratio, line-of-sight indicator, rate of change, etc. to remove any obvious unreliable measurements. The measurements are then stored in a buffer. Now, assuming that the system is at the time step t_{k+1} , which corresponds to the arrival of a new pointcloud message and the creation of the state $\hat{\mathcal{X}}_{k+1}$, the UWB buffer is then checked and all measurements that arrived during the period $(t_k, t_{k+1}]$ will be "bundled" together as a set. This set will be checked again using the latest IMU-predicted states to remove the suspected outliers. The final result is the set of $N_{\mathcal{U}}^k$ UWB measurements obtained during $(t_k, t_{k+1}]$, denoted as $\{\mathcal{U}_k^i : i = 1, 2, \dots, N_{\mathcal{U}}^k\}$.

C. IMU measurement workflow

The workflow of the IMU measurement is simpler, where we also extract from the buffer the IMU measurements that arrive during the period $(t_k, t_{k+1}]$ when the time step t_{k+1} elapses. These measurements are then preintegrated to provide a single IMU preintegration measurement \mathcal{I}_k that couples two consecutive states \mathcal{X}_k and \mathcal{X}_{k+1} . Also, IMU measurements are also used to propagate the robot state from the last jointly optimized estimate. For example, at time t_{k+1} where we have just received a new pointcloud message, $\hat{\mathbf{T}}_k$ is the last joint-optimization-based estimate of the robot pose.

D. Lidar pointcloud workflow

The handling of Lidar pointclouds are done similarly to the LIO-mapping framework developed in [20]. Upon receiving a complete sensor scan at time t_{k+1} , the features can be extracted to form a feature pointcloud $\mathbf{F}_{k+1} \triangleq \mathbf{B}_{k+1} \mathbf{F}_{k+1}$ (which is actually a composite of two pointclouds, one consists of plane features, denoted as \mathbf{F}_{k+1}^p , and another of the edge features, denoted as \mathbf{F}_{k+1}^e) using the method in [19], and stored in a buffer. Given the latest $M + 1$ feature pointclouds from \mathbf{F}_w to \mathbf{F}_{k+1} ($w \triangleq k + 1 - M$), we will merge the first M pointclouds from time t_w to t_k to construct the local map $\mathbf{M}_w \triangleq \mathbf{B}_w \mathbf{M}_w = (\mathbf{M}_w^p, \mathbf{M}_w^e)$ by

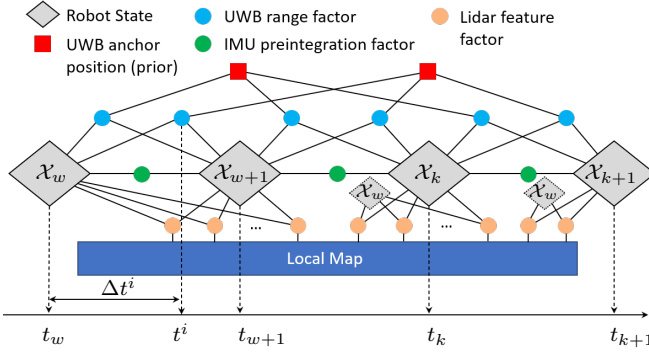


Fig. 3: Example of the factor graph over four time steps ($M = 3$). Note that the time difference between the stamp of a UWB measurement and the latest preceding time step, denoted as Δt_i is also included in the UWB measurement \mathcal{U}_m^i . The factors are described in more details in Section V.

using the latest estimated transforms $\hat{\mathbf{T}}_w, \hat{\mathbf{T}}_{w+1}, \dots, \hat{\mathbf{T}}_k$. Then, for each feature pointcloud \mathbf{F}_m , $m = w, \dots, k+1$, we will compute the set of feature coefficients $\mathcal{F}_m = \{\mathcal{L}_m^i \triangleq (\mathbf{f}^i, \mathbf{n}^i, c^i)\}$ that will be used to construct the Lidar feature factors. More details are discussed in Section V-B.

E. Joint optimization sensor fusion

Fig. 3 illustrates the factor graph of our cost function, where the coupling between each UWB, IMU preintegration, and Lidar feature with the prior and the states is pictorially described. Thus, given all measurements from UWB, IMU and Lidar in the local window, the cost function can be constructed and optimized at time t_{k+M} as

$$f(\hat{\mathcal{X}}) \triangleq \left\{ \begin{aligned} & \sum_{m=k}^{k+M-1} \left\| r_{\mathcal{I}}(\hat{\mathcal{X}}_m, \hat{\mathcal{X}}_{m+1}, \mathcal{I}_m) \right\|_{\mathbf{P}_{\mathcal{I}_m}^{-1}}^2 \\ & + \sum_{m=k}^{k+M-1} \sum_{i=1}^{N_{\mathcal{U}}^m} \left\| r_{\mathcal{U}}(\hat{\mathcal{X}}_m, \hat{\mathcal{X}}_{m+1}, \mathcal{U}_m^i) \right\|_{\mathbf{P}_{\mathcal{U}_m^i}^{-1}}^2 \\ & + \sum_{m=k+1}^{k+M} \sum_{i=1}^{N_{\mathcal{L}}^m} \left\| r_{\mathcal{L}}(\hat{\mathcal{X}}_k, \hat{\mathcal{X}}_m, \mathcal{L}_m^i) \right\|_{\mathbf{P}_{\mathcal{L}_m^i}^{-1}}^2 \end{aligned} \right\}, \quad (4)$$

where $r_{\mathcal{I}}(\cdot)$, $r_{\mathcal{U}}(\cdot)$, $r_{\mathcal{L}}(\cdot)$ are the *residuals* constructed from UWB, IMU and Lidar measurements; $\mathbf{P}_{\mathcal{I}_m}$, $\mathbf{P}_{\mathcal{U}_m^i}$, $\mathbf{P}_{\mathcal{L}_m^i}$ are the covariance matrices of the measurement error; $N_{\mathcal{U}}^m$ is the number of UWB measurements obtained in the period $(t_m, t_{m+1}]$, and $N_{\mathcal{L}}^m = |\mathcal{F}_m|$. In this paper, we use the ceres solver [34] to optimize this cost function.

V. COST FACTORS

A. IMU preintegration factors:

1) *IMU preintegration model*: For a conventional IMU, we can obtain the measurements $\tilde{\omega}$, $\tilde{\mathbf{a}}$, which are respectively the body's angular velocity and acceleration, corrupted by noises and biases. Given $\tilde{\omega}$, $\tilde{\mathbf{a}}$ and some nominal value of the IMU bias, denoted as $\tilde{\mathbf{b}}_k^a$, $\tilde{\mathbf{b}}_k^\omega$, the IMU preintegration

$\mathcal{I}_k \triangleq (\check{\alpha}_{k+1}, \check{\beta}_{k+1}, \check{\gamma}_{k+1})$ can be calculated by:

$$\check{\alpha}_{k+1} \triangleq \int_{t_k}^{t_{k+1}} \int_{t_k}^u {}^k\check{\mathbf{R}}_s (\check{\mathbf{a}}_s - \check{\mathbf{b}}_k^a) ds du, \quad (5)$$

$$\check{\beta}_{k+1} \triangleq \int_{t_k}^{t_{k+1}} {}^k\check{\mathbf{R}}_s (\check{\mathbf{a}}_s - \check{\mathbf{b}}_k^a) ds, \quad (6)$$

$$\check{\gamma}_{k+1} \triangleq \int_{t_k}^{t_{k+1}} \check{\gamma}_s \circ \begin{bmatrix} 0 \\ \check{\omega}_s - \check{\mathbf{b}}_s^\omega \end{bmatrix}, \quad (7)$$

where ${}^k\check{\mathbf{R}}_s \triangleq \mathcal{R}(\check{\gamma}_s)$, and $\int_{t_k}^{t_{k+1}} (\cdot) ds$ denotes the integration of the quaternion derivative. In this work, we employ a standard IMU preintegration procedure in literature. The complete details on the implementation of IMU preintegration technique can be found at [35].

B. Lidar feature factors

Recall that at time step t_{k+1} , we have a series of $M+1$ feature pointclouds $\mathbf{F}_w, \mathbf{F}_{w+1}, \dots, \mathbf{F}_{k+1}$, where the first M are merged into a local map \mathbf{M}_w consisting of points whose coordinates are w.r.t. to the body frame at time t_w ($w = k+1-M$), i.e. the first pose in the local window. Given the feature pointcloud \mathbf{F}^m , we calculate the set of coefficients \mathcal{F}_m following the steps in Algorithm 1, which are indeed the parameters of the 3D planes that the corresponding feature points are supposed to belong to.

Algorithm 1: Find Lidar feature coefficients

Input: $\mathbf{F}_m = (\mathbf{F}_m^p, \mathbf{F}_m^e)$, $\mathbf{M}_w = (\mathbf{M}_w^p, \mathbf{M}_w^e)$, ${}^w\hat{\mathbf{T}}_m$.
Output: $\mathcal{F}_m = \{\mathcal{L}_m^i = (\mathbf{f}^i, \mathbf{n}^i, c^i)\}$

- 1 **for** each $\mathbf{f} \in \mathbf{F}_m$ **do**
- 2 Compute ${}^w\mathbf{f}$ from \mathbf{f} using ${}^w\hat{\mathbf{T}}_m$;
- 3 **if** $\mathbf{f} \in \mathbf{F}_m^p$ **then**
- 4 Find $\mathcal{N}_{\mathbf{f}} = \text{KNN}({}^w\mathbf{f}, \mathbf{M}_w^p)$;
- 5 Find $\bar{\mathbf{n}} = \arg\min_{\mathbf{n}} \sum_{\mathbf{x} \in \mathcal{N}_{\mathbf{f}}} \|\mathbf{n}^\top \mathbf{x} + 1\|^2$;
- 6 Compute: $g = \frac{1}{\|\bar{\mathbf{n}}\|} \left[1 - 0.9 \frac{\bar{\mathbf{n}}^\top ({}^w\mathbf{f} + 1)}{\|\bar{\mathbf{n}}\| \|{}^w\mathbf{f}\|} \right]$;
- 7 Add $(\mathbf{f}, g\bar{\mathbf{n}}, g)$ to \mathcal{F}_m ;
- 8 **else if** $\mathbf{f} \in \mathbf{F}_m^e$ **then**
- 9 Find the set $\mathcal{N}_{\mathbf{f}} = \text{KNN}({}^w\mathbf{f}, \mathbf{M}_w^e)$, and its centroid $\bar{\mathbf{p}} = \frac{1}{|\mathcal{N}_{\mathbf{f}}|} \sum_{\mathbf{x} \in \mathcal{N}_{\mathbf{f}}} \mathbf{x}$;
- 10 Compute: $\mathbf{A} \triangleq \frac{1}{|\mathcal{N}_{\mathbf{f}}|} \sum_{\mathbf{x} \in \mathcal{N}_{\mathbf{f}}} (\mathbf{x} - \bar{\mathbf{p}})(\mathbf{x} - \bar{\mathbf{p}})^\top$;
- 11 Find the eigenvector \mathbf{v}_{\max} corresponding to the largest eigenvalue of \mathbf{A} ;
- 12 Compute: $\mathbf{x}_0 = {}^w\mathbf{f}$, $\mathbf{x}_1 = \bar{\mathbf{p}} + 0.1\mathbf{v}_{\max}$,
 $\mathbf{x}_2 = \bar{\mathbf{p}} - 0.1\mathbf{v}_{\max}$, $\mathbf{x}_{01} = \mathbf{x}_0 - \mathbf{x}_1$,
 $\mathbf{x}_{02} = \mathbf{x}_0 - \mathbf{x}_2$, $\mathbf{x}_{12} = \mathbf{x}_1 - \mathbf{x}_2$;
- 13 Compute: $\mathbf{n}_1 = \mathbf{x}_{12} \times (\mathbf{x}_{10} \times \mathbf{x}_{02})$,
 $\mathbf{n}_1 \leftarrow \mathbf{n}_1 / \|\mathbf{n}_1\|$, $\mathbf{n}_2 = \mathbf{x}_{12} \times \mathbf{n}_1$;
- 14 Compute: ${}^w\mathbf{f}_\perp = {}^w\mathbf{f} - (\mathbf{n}_1 \mathbf{n}_1^\top) \mathbf{x}_{01}$;
- 15 Compute: $c_1 = -\mathbf{n}_1^\top {}^w\mathbf{f}_\perp$ and $c_2 = -\mathbf{n}_2^\top {}^w\mathbf{f}_\perp$;
- 16 Compute: $g = \frac{1}{2} \left[1 - \frac{0.9 \|\mathbf{x}_{01} \times \mathbf{x}_{02}\|}{\|\mathbf{x}_{12}\|} \right]$;
- 17 Add $(\mathbf{f}, g\bar{\mathbf{n}}_1, gc_1)$ and $(\mathbf{f}, g\bar{\mathbf{n}}_2, gc_2)$ to \mathcal{F}_m ;
- 18 **end**

Hence, for each Lidar coefficient $\mathcal{L}_k^i = (\mathbf{f}^i, \mathbf{n}^i, c^i)$, a lidar feature factor can be constructed with the following residual:

$$r_{\mathcal{L}}(\hat{\mathcal{X}}_w, \hat{\mathcal{X}}_k, \mathcal{L}_k^i) = (\mathbf{n}^i)^\top \hat{\mathbf{R}}_w^{-1} [\hat{\mathbf{R}}_k \mathbf{f}^i + \hat{\mathbf{p}}_k - \hat{\mathbf{p}}_w] + c^i.$$

C. UWB range factors

Recall that in Section IV, for each interval $(t_k, t_{k+1}]$ in the local window, we have a bundle of UWB measurements $\{\mathcal{U}_k^i\}_{i=1}^{N_{\mathcal{U}}^k}$. More specifically, \mathcal{U}_k^i is defined as:

$$\mathcal{U}_k^i = (\check{d}^i, \mathbf{x}^i, \mathbf{y}^i, \delta t^i, \Delta t_k), \quad i = 1, 2, \dots, N_{\mathcal{U}}^k, \quad (8)$$

where \check{d}^i is the distance measurement, $\Delta t_k \triangleq t_{k+1} - t_k$ (see Fig. 3), $\delta t^i \triangleq t^i - t_k$, \mathbf{x}^i is the position of the UWB anchor w.r.t. the world frame, and \mathbf{y}^i is the coordinate of the UWB ranging node in the body frame.

If we assume that the velocity and orientation of the robot change at a constant rate from time t_k to t_{k+1} , then at time $t_k + \delta t^i$ the relative position of a UWB ranging node \mathbf{y}^i from an anchor \mathbf{x}^i can be determined as:

$$\begin{aligned} {}^w\mathbf{d}^i &= \mathbf{d}(\mathcal{X}_k, \mathcal{X}_{k+1}, \delta t^i, \Delta t_k) \\ &\triangleq \mathbf{p}_{k+1} + \mathbf{R}_k \text{Exp} \left(s^i \text{Log}(\mathbf{R}_k^{-1} \mathbf{R}_{k+1}) \right) \mathbf{y}^i \\ &\quad - \Delta t_{k+1} \int_{s^i}^1 [\mathbf{v}_k + \tau(\mathbf{v}_{k+1} - \mathbf{v}_k)] d\tau - \mathbf{x}^i \\ &\triangleq \mathbf{p}_{k+1} + \mathbf{R}_k \text{Exp} \left(s^i \text{Log}(\mathbf{R}_k^{-1} \mathbf{R}_{k+1}) \right) \mathbf{y}^i \\ &\quad - a^i \mathbf{v}_{k+1} - b^i \mathbf{v}_k - \mathbf{x}^i, \end{aligned} \quad (9)$$

where $s^i = \frac{\delta t^i}{\Delta t_k}$, $a^i = \frac{\Delta t_k^2 - \delta t^i^2}{2\Delta t_k}$, $b^i = \frac{(\Delta t_k - \delta t^i)^2}{2\Delta t_k}$.

We consider the distance measurement \check{d}^i at time $t_k + \delta t^i$ as the norm of the vector ${}^w\mathbf{d}^i$, corrupted by a zero-mean Gaussian noise $\eta_{\mathcal{U}}^i \sim \mathcal{N}(0, \sigma_{\mathcal{U}}^2)$, i.e. $\check{d}^i = \|{}^w\mathbf{d}^i\| + \eta_{\mathcal{U}}^i$. Thus, the UWB range factor can be defined as:

$$r_{\mathcal{U}}(\hat{\mathcal{X}}_k, \hat{\mathcal{X}}_{k+1}, \mathcal{U}_k^i) \triangleq \|\mathbf{d}(\hat{\mathcal{X}}_k, \hat{\mathcal{X}}_{k+1}, \delta t^i, \Delta t_k)\| - \check{d}^i, \quad (10)$$

Remark 1. A clear novelty of our method can be seen in the model (9), where the ranging model takes into account both the ranging node's offset in the body frame and the time delay. This allows us to increase robustness and applicability of the ranging scheme in cases where such offset and delay cannot be ignored such as the large UAV in our experiments.

VI. EXPERIMENTS

In this section we present the experiment results of the LIRO method on real world datasets. A video recording of the experiments can be viewed at https://youtu.be/Wfp_VcwzNKY, or the in the supplementary materials of this paper. The datasets used for this experiment have been published online [36] and can be downloaded from our website¹.

Fig. 4 presents the hardware setup on a hexacopter in our experiments. Specifically, it consists of one VN-100 IMU, four UWB ranging nodes and two 16-channel OS1 Lidars, where one so-called horizontal Lidar is oriented to scan the surrounding effectively, and the other so-called vertical Lidar

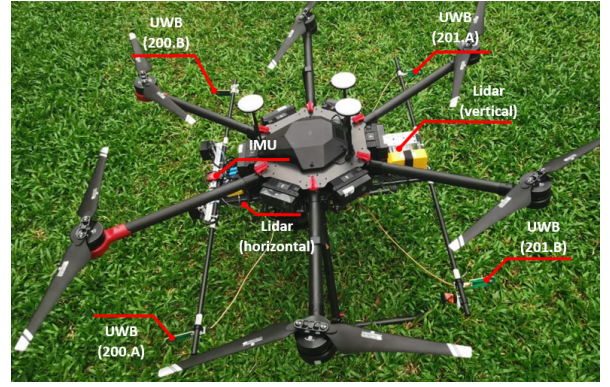


Fig. 4: The hardware setup used in our flight tests.

is set to scan the front, back and the ground effectively. Our intention is to subsequently combine both lidars in a future work. However in this paper, we can show more clearly the effectiveness of UWB in reducing drift when only one lidar is used.

The 3 anchors are deployed as depicted in Fig. 1, where the coordinate x_1^* ranges from 40m to 60m and y_2^* ranges from -15m to -10m, depending on the available space. The four UWB ranging nodes on the hexacopter are positioned on the vertices of a $0.75\text{m} \times 0.55\text{m}$ rectangle around the IMU, which is also the body center. A Leica MS60 station with millimeter-level accuracy is used to provide groundtruth for the experiment. The software algorithms are implemented on the ROS framework².

We conduct three flight tests to scan the facade of a building to collect the data for these experiments. We run the LIRO algorithm on the horizontal and vertical Lidar measurements, and the ranging data being fused are limited to 0, 2, and 3 anchors, respectively, to demonstrate and compare the benefit of the number of landmarks in reducing estimation drift. We refer to the case when range data is not used as LIO and LIRO2 or LIRO3 to the cases where ranging to 2 or 3 anchors are fused, respectively. We use the method in [37] to align the trajectories before calculating the positional root mean square error (RMSE) of the estimated trajectories. Though no groundtruth for orientation is available, we assume the hexacopter's onboard estimation, which fuses in the magnetometer measurement, is accurate and can be used as groundtruth, since no magnetic interference was observed at the test site on the day. This is then used to calculate the rotational RMSE.

The LOAM³ method [38] is also run with the acquired data for comparison. We also attempt to run LIO-SAM [26] with our datasets, however LIO-SAM quickly diverges after the robot takes off. We suspect this is because LIO-SAM requires the roll-pitch-yaw estimate from the IMU's built in estimator, which has significant drift due to high-frequency vibration when the hexacopter is flying. Indeed the experiment of LIRO with 0 anchors can also be considered as substitute for LIO-Mapping [20], since we adopt several software

¹https://ntu-aris.github.io/ntu_viral_dataset/

²<https://www.ros.org/>

³<https://github.com/HKUST-Aerial-Robotics/A-LOAM>

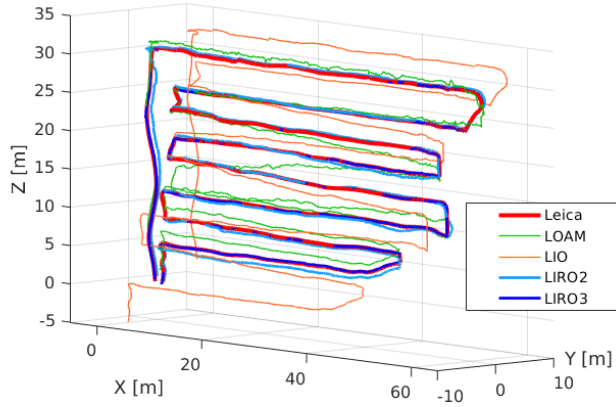


Fig. 5: Trajectories of the estimates from the horizontal Lidar in the first test.

TABLE I: Position and rotation RMSE of odometry estimates from the horizontal Lidar data.

Test	RMSE _{pos} [m]				RMSE _{rot} [deg]			
	LOAM	LIO	LIRO2	LIRO3	LOAM	LIO	LIRO2	LIRO3
01	1.499	4.692	0.527	0.170	2.232	0.693	1.650	1.610
02	5.275	5.242	0.758	0.393	4.239	0.927	2.351	3.180
03	2.549	5.833	0.664	0.204	4.989	0.996	2.597	2.458

components from this work. However, it should be noted that our implementation has been significantly improved to ensure real-time performance, while LIO-Mapping is known to suffer from computation bottlenecks [26]. Tab. I and Tab. II summarize the results of these experiments. Fig. 5 and Fig. 6 show the trajectory and orientation estimation error of some tests for a closer look. More plots and figures can be viewed in the accompanying video.

We can immediately see from Tab. I and Tab. II that the use of ranging factors clearly improve the positioning estimation. While the RMSE of position estimate with only Lidar or Lidar-inertial measurements can be several meters large, by employing ranging to two anchors, the error can be reduced to below 0.8m, and the accuracy when using three anchors can be as small as 0.15m. Moreover, one notable feature that can be observed is that the use of ranging factors can regularize the quality of localization across the experiments. We can see that the RMSE is quite unpredictable in the LOAM and LIO cases, while the accuracy of LIRO is quite consistent across the environments and anchor configurations.

In terms of orientation estimate, we find that LIO estimation appears to be more accurate in most tests. We believe that this is due to LIO estimate being smoother than LIRO (refer to the yaw error in Fig. 6). However it can be seen that these orientation errors are already quite small such that the difference is imperceptible. For example the largest difference between LIRO and LIO is approximately 2.3° or 0.04 rad (test 02 in Tab. I), compared to a difference of several meters in the position errors. Moreover, since the accuracy of groundtruth is not guaranteed, the exact values of the RMSE are not be very important. On the other hand,

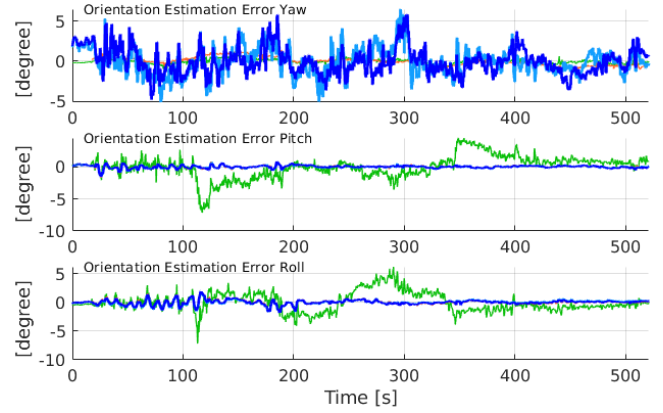


Fig. 6: Orientation estimation error from the horizontal Lidar in the first test (the legend is the same with Fig. 5).

TABLE II: Position and rotation RMSE of odometry estimates from the vertical Lidar.

Test	RMSE _{pos} [m]				RMSE _{rot} [deg]			
	LOAM	LIO	LIRO2	LIRO3	LOAM	LIO	LIRO2	LIRO3
01	13.665	5.421	0.492	0.159	8.421	1.812	1.803	1.639
02	2.710	2.033	0.744	0.394	4.864	1.208	2.463	3.228
03	1.603	2.710	0.643	0.209	4.062	0.801	2.849	2.786

the LIRO's orientation is of global type and the accuracy is consistent throughout the experiments, while LOAM and LIO estimates are relative to the initial pose, and have unpredictable orientation drifts in some experiments.

VII. CONCLUSION

In this paper we have developed a tightly coupled Lidar-inertial-ranging odometry estimation scheme, so-called LIRO, and successfully implemented it on the open-source ceres solver and ROS. Experiments on real-world datasets have been conducted to verify the efficacy and effectiveness of the sensor fusion scheme. We show that via the use of some anchors that can be quickly deployed on the field, estimation drift can be significantly reduced, and the frame of reference for LIRO can be directly determined via the anchor deployment. The results demonstrate that LIRO is an effective and flexible localization solution for robots operating in semi-controlled environments.

For future works, we aim to integrate multiple lidars to improve the system robustness, based on which we can develop a comprehensive SLAM capability with loop closure and bundle adjustment, as current approach only focuses on the local map. Moreover, the framework can be extended to integrate visual (camera-based) factors, or engage in some collaborative localization scheme with other robots or "mobile anchors".

REFERENCES

- [1] S. Yuan, H. Wang, and L. Xie, "Survey on localization systems and algorithms for unmanned systems," *Unmanned Systems*, vol. 9, no. 2, pp. 129–163, 2021.

- [2] A. Kushleyev, D. Mellinger, C. Powers, and V. Kumar, "Towards a swarm of agile micro quadrotors," *Autonomous Robots*, vol. 35, no. 4, pp. 287–300, 2013.
- [3] M. Watterson and V. Kumar, "Control of quadrotors using the hopf fibration on so (3)," in *Robotics Research*. Springer, 2020, pp. 199–215.
- [4] C. Martínez, P. Campoy, I. Mondragón, and M. A. Olivares-Méndez, "Trinocular ground system to control uavs," in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2009, pp. 3361–3367.
- [5] S. Klose, J. Wang, M. Achtelik, G. Panin, F. Holzapfel, and A. Knoll, "Markerless, vision-assisted flight control of a quadcopter," in *2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2010, pp. 5712–5717.
- [6] M. W. Mueller, M. Hamer, and R. D'Andrea, "Fusing ultra-wideband range measurements with accelerometers and rate gyroscopes for quadcopter state estimation," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2015, pp. 1730–1736.
- [7] K. Guo, Z. Qiu, C. Miao, A. H. Zaini, C.-L. Chen, W. Meng, and L. Xie, "Ultra-wideband-based localization for quadcopter navigation," *Unmanned Systems*, vol. 4, no. 01, pp. 23–34, 2016.
- [8] T.-M. Nguyen, A. H. Zaini, K. Guo, and L. Xie, "An ultra-wideband-based multi-uav localization system in gps-denied environments," in *International Micro Air Vehicle Competition and Conference 2016*, Beijing, China, Oct 2016, pp. 56–61.
- [9] J. Tiemann and C. Wietfeld, "Scalable and precise multi-uav indoor navigation using tdoa-based uwb localization," in *2017 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. IEEE, 2017, pp. 1–7.
- [10] J. Paredes, F. Álvarez, T. Aguilera, and J. Villadangos, "3d indoor positioning of uavs with spread spectrum ultrasound and time-of-flight cameras," *Sensors*, vol. 18, no. 1, p. 89, 2018.
- [11] C. Forster, M. Pizzoli, and D. Scaramuzza, "Svo: Fast semi-direct monocular visual odometry," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 15–22.
- [12] S. Shen, N. Michael, and V. Kumar, "Tightly-coupled monocular visual-inertial fusion for autonomous flight of rotorcraft mavs," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2015, pp. 5303–5310.
- [13] T. Qin, P. Li, and S. Shen, "Vins-mono: A robust and versatile monocular visual-inertial state estimator," *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 1004–1020, 2018.
- [14] 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.
- [15] J. Engel, T. Schöps, and D. Cremers, "Lsd-slam: Large-scale direct monocular slam," in *European conference on computer vision*. Springer, 2014, pp. 834–849.
- [16] R. Mur-Artal and J. D. Tardós, "Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras," *IEEE Transactions on Robotics*, vol. 33, no. 5, pp. 1255–1262, 2017.
- [17] G. Loianno, C. Brunner, G. McGrath, and V. Kumar, "Estimation, control, and planning for aggressive flight with a small quadrotor with a single camera and imu," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 404–411, 2017.
- [18] A. Weinstein, A. Cho, G. Loianno, and V. Kumar, "Visual inertial odometry swarm: An autonomous swarm of vision-based quadrotors," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 1801–1807, 2018.
- [19] J. Zhang and S. Singh, "Laser-visual-inertial odometry and mapping with high robustness and low drift," *Journal of Field Robotics*, vol. 35, no. 8, pp. 1242–1264, 2018.
- [20] 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.
- [21] X. Zuo, P. Geneva, W. Lee, Y. Liu, and G. Huang, "Lic-fusion: Lidar-inertial-camera odometry," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 5848–5854.
- [22] S. W. Chen, G. V. Nardari, E. S. Lee, C. Qu, X. Liu, R. A. F. Romero, and V. Kumar, "Sloam: Semantic lidar odometry and mapping for forest inventory," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 612–619, 2020.
- [23] H. Nguyen, F. Mascarich, T. Dang, and K. Alexis, "Autonomous aerial robotic surveying and mapping with application to construction operations," *arXiv preprint arXiv:2005.04335*, 2020.
- [24] M. Dharmadhikari, T. Dang, L. Solanka, J. Loje, H. Nguyen, N. Khedekar, and K. Alexis, "Motion primitives-based path planning for fast and agile exploration using aerial robots," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 179–185.
- [25] R. Reinhart, T. Dang, E. Hand, C. Papachristos, and K. Alexis, "Learning-based path planning for autonomous exploration of subterranean environments," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 1215–1221.
- [26] T. Shan, B. Englot, D. Meyers, W. Wang, C. Ratti, and R. Daniela, "Lio-sam: Tightly-coupled lidar inertial odometry via smoothing and mapping," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020.
- [27] G. Tinchev, A. Penate-Sanchez, and M. Fallon, "Learning to see the wood for the trees: Deep laser localization in urban and natural environments on a cpu," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1327–1334, 2019.
- [28] C. Wang, H. Zhang, T.-M. Nguyen, and L. Xie, "Ultra-Wideband Aided Fast Localization and Mapping System," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017.
- [29] T. H. Nguyen, T.-M. Nguyen, M. Cao, and L. Xie, "Loosely-Coupled Ultra-Wideband-Aided Scale Correction for Monocular Visual Odometry," *Unmanned Systems*, vol. 08, no. 02, pp. 179–190, 2020. [Online]. Available: <https://www.worldscientific.com/doi/10.1142/S2301385020500119>
- [30] T. H. Nguyen, T.-M. Nguyen, and L. Xie, "Tightly-coupled ultra-wideband-aided monocular visual slam with degenerate anchor configurations," *Autonomous Robots*, vol. 44, no. 8, pp. 1519–1534, 2020.
- [31] —, "Tightly-coupled single-anchor ultra-wideband-aided monocular visual odometry system," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 665–671.
- [32] Y. Cao and G. Beltrame, "Vir-slam: Visual, inertial, and ranging slam for single and multi-robot systems," *arXiv preprint arXiv:2006.00420*, 2020.
- [33] Y. Song, M. Guan, W. P. Tay, C. L. Law, and C. Wen, "Uwb/lidar fusion for cooperative range-only slam," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 6568–6574.
- [34] S. Agarwal, K. Mierle, and Others, "Ceres solver," <http://ceres-solver.org>.
- [35] T.-M. Nguyen, M. Cao, S. Yuan, Y. Lyu, T. H. Nguyen, and L. Xie, "Viral-fusion: A visual-inertial-ranging-lidar sensor fusion approach," *Submitted to IEEE Transactions on Robotics*, 2021.
- [36] T.-M. Nguyen, S. Yuan, M. Cao, Y. Lyu, T. H. Nguyen, and L. Xie, "Ntu viral: A visual-inertial-ranging-lidar dataset, from an aerial vehicle viewpoint," *Submitted to IJRR*, 2021.
- [37] Z. Zhang and D. Scaramuzza, "A tutorial on quantitative trajectory evaluation for visual (-inertial) odometry," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 7244–7251.
- [38] J. Zhang and S. Singh, "Low-drift and real-time lidar odometry and mapping," *Autonomous Robots*, vol. 41, no. 2, pp. 401–416, 2017.