

UVIO: An UWB-Aided Visual-Inertial Odometry Framework with Bias-Compensated Anchors Initialization

Giulio Delama¹, Farhad Shamsfakhr², Stephan Weiss¹, Daniele Fontanelli² and Alessandro Fornasier¹

Abstract—This paper introduces UVIO, a multi-sensor framework that leverages Ultra Wide Band (UWB) technology and Visual-Inertial Odometry (VIO) to provide robust and low-drift localization. In order to include range measurements in state estimation, the position of the UWB anchors must be known. This study proposes a multi-step initialization procedure to map multiple unknown anchors by an Unmanned Aerial Vehicle (UAV), in a fully autonomous fashion. To address the limitations of initializing UWB anchors via a random trajectory, this paper uses the Geometric Dilution of Precision (GDOP) as a measure of optimality in anchor position estimation, to compute a set of optimal waypoints and synthesize a trajectory that minimizes the mapping uncertainty. After the initialization is complete, the range measurements from multiple anchors, including measurement biases, are tightly integrated into the VIO system. While in range of the initialized anchors, the VIO drift in position and heading is eliminated. The effectiveness of UVIO and our initialization procedure has been validated through a series of simulations and real-world experiments.

I. INTRODUCTION

Autonomous operation of mobile robots and UAVs has seen substantial growth and advancement in recent years. With numerous applications ranging from industrial automation to the exploration of hazardous environments, the requirement for accurate and low-drift state estimation is crucial for ensuring the safe and efficient functioning of these systems. One of the biggest challenges in autonomous navigation is operating in GNSS-denied environments. In such scenarios, the need for robust and reliable alternative localization methods becomes increasingly important. VIO has emerged as a powerful approach for providing real-time state estimation in these scenarios. It uses a combination of visual information from one or multiple cameras and inertial information from Inertial Measurement Units (IMUs) that incorporate accelerometers and gyroscopes. However, despite its successes, VIO suffers from accumulative drift over time due to the inherent limitations of visual-based localization methods. In addition VIO is prone to fail in poor visual conditions. To address this, recent research has explored the integration of UWB range measurements into VIO systems, leveraging the global information provided by these sensors to improve localization performance. In this paper, we introduce UVIO, a multi-sensor framework built upon the well-known OpenVINS [1], that combines the strengths of both

sensor modalities to overcome the limitations of traditional VIO and achieve low-drift localization for mobile robots and UAVs. Together with the automated detection and initialization of the UWB anchors, UVIO provides a complete, easily deployable solution for real-time state estimation in GNSS-denied environments. The initialization procedure starts with a random flight with the objective of obtaining a first coarse estimation of the anchors' positions by solving a linear least squares problem followed by a nonlinear optimization on the same data. This initial estimate is then used in a refinement procedure that consists of computing and flying an optimal trajectory, collecting the data, and solving a nonlinear least squares problem to refine the initial solution. For this purpose, the GDOP, a metric tightly related to the Fisher Information Matrix (FIM), is used as a measure of optimality for deriving a set of waypoints that minimizes the mapping uncertainty. The entire pipeline is carried out by the UAV itself (navigating initially only based on VIO), without the need for any human intervention. By incorporating the proposed initialization procedure, the UVIO framework is able to automatically estimate the position of the UWB anchors, which can subsequently be included in the estimation process to improve the accuracy and reliability of the localization system. This allows the UAV to operate in a completely autonomous manner, without the need for human intervention, even in GNSS-denied environments. The efficacy of the proposed procedure is demonstrated through a series of simulations and real-world experiments, showcasing the robustness and reliability of the UVIO framework even in situations of poor visual conditions where pure VIO is prone to fail.

II. RELATED WORK

Several studies have explored the use of UWB measurements to enhance localization accuracy, which plays a key role in autonomous navigation. This section briefly reviews some of the most relevant works and discusses their significance in relation to our research, with a specific focus on studies that tackle the UWB anchors' initialization problem. Early studies, such as [2], [3], and [4], employed UWB measurements with prior knowledge of anchor positions, resulting in a drift-free estimation of the robot's pose, and showed the potential of this approach to improve current navigation and SLAM systems. These methods required an offline calibration of the UWB anchors' initial position, reducing the autonomy of operation and making them unsuitable for large-scale and dynamic environments.

¹Giulio Delama, Stephan Weiss, and Alessandro Fornasier are with the Control of Networked Systems Group, University of Klagenfurt, Austria. {name.surname}@ieee.org

²Farhad Shamsfakhr and Daniele Fontanelli are with the Department of Industrial Engineering, University of Trento, Italy. {name.surname}@unitn.it

Recent studies, such as [5], [6], [7], [8], [9] and [10], aimed to improve localization accuracy and reduce drift by integrating UWB range measurements and VIO. These studies rely on the assumption that the UWB anchors' positions are known a priori. They adopt graph-optimization methods for state estimation, as opposed to the filter-based method considered in this work. The choice of method depends on the specific requirements of the problem. Graph-based methods for state estimation can often provide more accurate results than filter-based methods, but at a higher computational cost. On the other hand, filter-based methods are typically less computationally intensive and can be more easily implemented in real-time applications where computational resources are limited.

The authors in [11] proposed a multi-sensor fusion framework based on the Extended Kalman Filter (EKF) to estimate anchors and UAV positions jointly. The initialization of the UWB anchors is performed in a single step by solving a linear least-squares problem, which can lead to inaccurate or completely wrong initialization results.

[12] presents an accurate and easy-to-use method for UWB anchor self-localization by utilizing ranging measurements and readings from a low-cost IMU. The locations of the anchors are estimated by freely moving the tag, using a tightly-coupled Error State Kalman Filter (ESKF) to fuse the UWB and inertial measurements. The anchors are first initialized via an Iterative Least-Squares without considering measurement biases.

In [13], a self-calibration algorithm for ultra-wideband positioning systems with active anchor nodes is shown. The algorithm uses iterative gradient descent and error detection from a convolutional neural network to estimate the positions of the fixed anchors with high accuracy, even in non-line-of-sight (NLOS) conditions, but requires the anchors to exchange range messages between each other.

The study in [14] addresses the drift in VIO for indoor localization by incorporating a single UWB anchor into the localization process. The initialization of the anchor is performed by solving a nonlinear Least-Squares problem that requires an initial guess for the position of the anchor that must be provided and known a priori.

The studies of Nguyen et al. in [15] and [16] proposed two tightly-coupled odometry frameworks that combine monocular visual feature observations with distance measurements provided by a single UWB anchor. The first framework uses a variant of the Levenberg-Marquardt non-linear optimization algorithm to simultaneously estimate the scale factor and the anchor position while the second framework leverages the UWB measurements more effectively by addressing the time-offset of each range data and utilizing all available measurements. Later, in [17] and [18], the authors propose a graph-based optimization approach for multi-sensor Simultaneous Localization and Mapping (SLAM) for UAVs. The system fuses data from various sensors, including IMU, cameras, lidars, and UWB range measurements to estimate the state of the vehicle in real time and with long-term consistency.

A recent paper from Jia et al. [19] proposes a Visual-

Inertial-Ranging Odometry (VIRO) approach to reduce localization drift in VIO systems by incorporating UWB ranging measurements from a minimum of 3 anchors. The method utilizes a long-short time-window structure to initialize the anchors' positions and covariance. The paper also analyzes the observability of the VIRO system with unknown UWB anchor positions and shows that there are four unobservable directions in the ideal case, which are reduced by fusing the UWB measurements with a Multi-State Constraint Kalman Filter (MSCKF).

All of the studies presented so far show the potential of including UWB technology for localization and in particular the mitigation of the drift in VIO navigation systems. However, none of the above-cited studies considers the fact that the solution of anchor initialization, which in turn affects the accuracy of the state estimation, is highly dependent on various factors, including the geometry of the disposition of the anchors and the trajectory of the UAV or the mobile robot that collects the initialization data. In a recent study [20], the authors applied the principles of FIM to determine the optimal points within a designated flight volume for collecting data for initializing the position of a single UWB anchor together with its measurement biases. The method consists of a coarse triangulation using random vehicle positions, followed by fine triangulation using FIM-optimized points that provide maximal information for the single anchor.

Our approach extends and improves the ideas in [20] on information-driven UWB anchor initialization and usage by developing a fully automated initialization approach for *simultaneous* initialization and subsequent use of *multiple initially unknown* UWB anchors with measurement biases. We make use of the GDOP as a novel metric to guide the optimization of data collection, initialization, and uncertainty reduction for a scalable set of UWB anchors. More precisely, our contributions are:

- Fully automated procedure to detect, initialize and use initially unknown UWB anchors in a vision-UWB supported estimation approach.
- Novel least squares formulation, dubbed the optimal double method (Sec. IV-A), increasing the matrix condition number for improved coarse anchors initialization.
- Optimal waypoint generation in arbitrarily shaped volumes based on the GDOP metric to enable optimal data collection for precise and *simultaneous* initialization of UWB anchors.
- Validation of the proposed approach in simulations and in real experiments with an autonomously flying UAV (and open sourcing of the code).
- Real-world demonstration of the resilience of our approach against camera dropouts.

III. UWB-AIDED VISUAL INERTIAL ODOMETRY FILTER DESIGN

A. UVIO filter state

Our proposed EKF-based UVIO filter has the following structure. Beside the classical MSCKF-VIO *core state*, ${}^I X$,

including the orientation ${}^I_G \mathbf{q}$, position ${}^G \mathbf{p}_I$, and velocity ${}^G \mathbf{v}_I$ of the IMU frame $\{I\}$ in the global frame $\{G\}$, as well as the gyroscope and accelerometer bias \mathbf{b}_ω , \mathbf{b}_a , the *clones state* ${}^C \mathbf{x}$, including the clones of the IMU pose at different past time steps, the *features state* ${}^F \mathbf{x}$, including the position of N features in the global frame, and the *calibration state* ${}^W \mathbf{x}$, including camera intrinsics and extrinsics for each camera present, the UVIO filter is augmented with the *UWB state* ${}^U \mathbf{x}$. The latter includes the UWB extrinsic parameter, that is the position of the UWB tag with respect to the IMU frame ${}^I \mathbf{p}_U$, the position of M anchors in the global frame ${}^G \mathbf{p}_{A_1}, \dots, {}^G \mathbf{p}_{A_M}$, as well as the bias parameters $\beta_1, \gamma_1, \dots, \beta_M, \gamma_M$ for each anchor. The bias parameters include the so-called *distance bias* and *constant bias*, described in later sections. Therefore, the UVIO state is denoted as

$$\mathbf{x} = ({}^I \mathbf{x}, {}^C \mathbf{x}, {}^F \mathbf{x}, {}^W \mathbf{x}, {}^U \mathbf{x}). \quad (1)$$

In the proposed filter, each UWB anchor can be treated as *fixed*, i.e., their values are kept constant, or can be *refined online* and hence added to the UWB state. With two anchors fixed, we can set a non-drifting global frame (in fact, a non-linear observability analysis shows that fixing one anchor in position and the heading toward another anchor would be sufficient; fixing two anchors in position is done to simplify the implementation in the results section).

B. UVIO range update

For each anchor A , a range measurement $z_A = h(\mathbf{x}) + n$, is described by the following model

$$\begin{aligned} h(\mathbf{x}) &:= \beta d + \gamma = \beta \|{}^G \mathbf{p}_U - {}^G \mathbf{p}_A\| + \gamma, \\ &= \beta \|{}^G \mathbf{p}_I + {}^I_G \mathbf{R}^T {}^I \hat{\mathbf{p}}_U - {}^G \mathbf{p}_A\| + \gamma, \end{aligned} \quad (2)$$

where β and γ , firstly introduced in [20], are respectively a multiplicative scalar that acts as a distance-dependent bias, and a constant bias term.

For each range measurement z_A from anchor A , define

$$\Gamma(\hat{\mathbf{x}}) = \hat{\beta} \frac{\left({}^G \hat{\mathbf{p}}_I + {}^I_G \hat{\mathbf{R}}^T {}^I \hat{\mathbf{p}}_U - {}^G \hat{\mathbf{p}}_A \right)^T}{\|{}^G \hat{\mathbf{p}}_I + {}^I_G \hat{\mathbf{R}}^T {}^I \hat{\mathbf{p}}_U - {}^G \hat{\mathbf{p}}_A\|},$$

then, the measurement Jacobian \mathbf{H}_{z_A} is defined through the following derivatives evaluated at $\hat{\mathbf{x}}$

$$\begin{aligned} \frac{\partial h(\mathbf{x})}{\partial {}^I_G \mathbf{R}} &= \Gamma(\hat{\mathbf{x}}) {}^I_G \hat{\mathbf{R}}^T [{}^I \hat{\mathbf{p}}_U]_\times, & \frac{\partial h(\mathbf{x})}{\partial {}^G \mathbf{p}_I} &= \Gamma(\hat{\mathbf{x}}), \\ \frac{\partial h(\mathbf{x})}{\partial {}^I \mathbf{p}_U} &= \Gamma(\hat{\mathbf{x}}) {}^I_G \hat{\mathbf{R}}^T, & \frac{\partial h(\mathbf{x})}{\partial {}^G \mathbf{p}_A} &= -\Gamma(\hat{\mathbf{x}}), \\ \frac{\partial h(\mathbf{x})}{\partial \beta} &= \|{}^G \hat{\mathbf{p}}_I + {}^I_G \hat{\mathbf{R}}^T {}^I \hat{\mathbf{p}}_U - {}^G \hat{\mathbf{p}}_A\|, & \frac{\partial h(\mathbf{x})}{\partial \gamma} &= 1. \end{aligned}$$

Contrary to [19], we do not interpolate range measurements to align them with the camera measurements, but, to avoid introducing extra source of errors due to interpolation, we rather perform what we called *Delayed Update*. As shown in Fig. 1, this update strategy consists of collecting range measurements in between the latest update at time t_n and

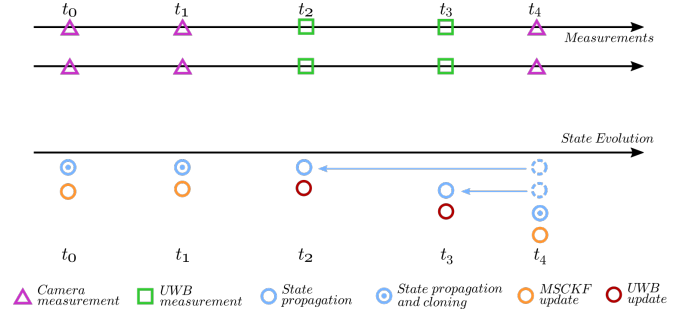


Fig. 1. This figure shows an example of what we described as delayed update strategy. In this example, the actual state estimate is at the time the last camera measurement has been received, thus t_1 . Two UWB measurements are received at time t_2 and t_3 . At the next camera measurement, at time t_4 , the state is propagated and updated through all the UWB measurements collected, and finally, at time t_4 a new clone is added and then the state is updated with the newest camera measurement.

the next camera image at time t_k . Upon receiving a new camera image at time t_k , a series of chained propagation-update from time t_n to time t_k are performed.

IV. UWB ANCHOR INITIALIZATION

In order to be able to exploit UWB measurements in the UVIO filter an initial position of the UWB anchors needs to be provided. To minimize the setup time from the practitioner's perspective, the UVIO framework provides an anchors detection and self-initialization module. The anchors' self-initialization routine only assumes that the robot pose (through VIO in our case), as well as UWB range measurements, are available to the system. Firstly, data is collected from a short random or standard trajectory, which is used to obtain a rough initial solution by solving a linear least squares problem. Subsequently, a nonlinear optimization problem is built on the same data, and it is solved to refine the first coarse anchors positioning. The resulting anchor's position estimate is then used to compute a set of waypoints optimizing our proposed GDOP metric to maximize triangulation information for all involved UWB anchors. The data from these optimal waypoints is then used in a nonlinear optimization problem to refine all UWB anchors positions simultaneously.

A. Coarse anchor positioning: Optimal Double Method

Obtaining a first rough estimation of the anchors' positions is fundamental for having an initial guess for the nonlinear optimization. In [11] and [20], the coarse anchors' initialization is modeled as a least square problem, which provides a fast initial guess of the anchors' positions. However, the least square formulation is in practice often ill-posed yielding poor solutions. In this respect, [11] introduced a measurement inclusion test, where measurements are added to the least square problem only if the condition number of the problem is decreased by adding such a measurement. In this work, we reformulated the least square problem described in [20] with what we called *optimal double method*, aiming at increasing the condition number of the least square problem and minimizing the uncertainty of the solution.

The first choice we made to improve the conditioning of the least square formulation of the anchors' coarse initialization problem is to consider a known $\beta = 1$ in the measurement model in Equ. (2), therefore, for a given range measurement z from anchor A , we write

$$z = \|\mathbf{p}_U - \mathbf{p}_A\| + \gamma + n \quad n \sim \mathcal{N}(0, \sigma_n^2).$$

Ignoring the measurement noise, we can square the previous equation and rewrite it as follows

$$\begin{aligned} z - \gamma &= \|\mathbf{p}_U - \mathbf{p}_A\|, \\ (z - \gamma)^2 &= (\mathbf{p}_U - \mathbf{p}_A)^T (\mathbf{p}_U - \mathbf{p}_A) \\ z^2 &= \|\mathbf{p}_U\|^2 + \|\mathbf{p}_A\|^2 - 2\mathbf{p}_U^T \mathbf{p}_A + 2z\gamma - \gamma^2. \end{aligned} \quad (4)$$

By applying the *double method* [21], it is possible to derive a linear relation with respect to \mathbf{p}_A and γ . More precisely, consider z_i and z_j being two different (not necessarily consecutive) range measurements from the same anchor A , and \mathbf{p}_{U_i} and \mathbf{p}_{U_j} the UWB tag positions at the time the measurements have been collected. Then by subtracting the two equations (4) we obtain

$$\begin{aligned} z_i^2 - z_j^2 &= \|\mathbf{p}_{U_i}\|^2 - \|\mathbf{p}_{U_j}\|^2 - \\ &\quad - 2(\mathbf{p}_{U_i} - \mathbf{p}_{U_j})^T \mathbf{p}_A + 2(z_i - z_j)\gamma, \end{aligned}$$

leading to the following least square formulation:

$$\begin{bmatrix} \vdots \\ \mathbf{A}_k \\ \vdots \end{bmatrix} \mathbf{x} = \begin{bmatrix} \vdots \\ b_k \\ \vdots \end{bmatrix} \Rightarrow \mathbf{A} \mathbf{x} = \mathbf{b},$$

with

$$\begin{aligned} \mathbf{A}_k &:= \begin{bmatrix} -(\mathbf{p}_{U_i} - \mathbf{p}_{U_j})^T & (z_i - z_j) \end{bmatrix}, \\ b_k &:= \frac{1}{2} \left((z_i^2 - z_j^2) - (\|\mathbf{p}_{U_i}\|^2 - \|\mathbf{p}_{U_j}\|^2) \right), \\ \mathbf{x} &:= [\mathbf{p}_A \quad \gamma]^T. \end{aligned}$$

In order to identify the pair z_i and z_j , an optimal *pivot* measurement to be subtracted from any other that minimizes the uncertainty of the solution should be determined. To this end, let $n_k \sim \mathcal{N}(0, \sigma_n^2)$, be the noise associated with the k -th range measurement from anchor A , and $\varepsilon_k \sim \mathcal{N}(0, \Sigma_\varepsilon)$ be the noise associated with the k -th position of the UWB tag. In the analysis that follows, we consider only the vector \mathbf{b} being a stochastic entity while treating the matrix \mathbf{A} as deterministic. Even though this is an approximation, it makes the analysis easier leading to a meaningful criterion for the choice of the optimal pivot.

Let us start by considering $\mathbf{A}_k \mathbf{x} = b_k = \bar{b}_k + \eta_k$, with $\eta_k \sim \mathcal{N}(0, \sigma_{\eta_k}^2)$ and

$$\sigma_{\eta_k}^2 = (z_k^2 + z_p^2) \sigma_n^2 + \mathbf{p}_{U_k} \Sigma_{\varepsilon_k} \mathbf{p}_{U_k}^T + \mathbf{p}_{U_p} \Sigma_{\varepsilon_p} \mathbf{p}_{U_p}^T,$$

where p is the index of the pivot. Define $\Sigma_\eta = \text{diag}[\sigma_{\eta_1}^2, \dots, \sigma_{\eta_M}^2]$ being the covariance of the

vector \mathbf{b} , then the information matrix of the least square solution is written

$$\begin{aligned} \Sigma_{\mathbf{x}}^{-1} &= \mathcal{I}_{\mathbf{x}} = \mathbf{A}^T \Sigma_\eta^{-1} \mathbf{A} = \begin{bmatrix} \mathcal{I}_{11} & \mathcal{I}_{12} \\ \mathcal{I}_{21} & \mathcal{I}_{22} \end{bmatrix}, \\ \mathcal{I}_{11} &= \sum_{k=1}^M \frac{(\mathbf{p}_{U_k} - \mathbf{p}_{U_p}) (\mathbf{p}_{U_k} - \mathbf{p}_{U_p})^T}{\sigma_{\eta_k}^2}, \\ \mathcal{I}_{12} &= - \sum_{k=1}^M (\mathbf{p}_{U_k} - \mathbf{p}_{U_p}) \frac{(z_k - z_p)}{\sigma_{\eta_k}^2}, \\ \mathcal{I}_{21} &= - \sum_{k=1}^M \frac{(z_k - z_p)}{\sigma_{\eta_k}^2} (\mathbf{p}_{U_k} - \mathbf{p}_{U_p})^T, \\ \mathcal{I}_{22} &= \sum_{k=1}^M \frac{(z_k - z_p)^2}{\sigma_{\eta_k}^2}. \end{aligned}$$

Then, it turns out that the optimal pivot index p can be chosen to fulfill the A-optimality criteria, that is such that the associated information matrix $\mathcal{I}_{\mathbf{x}}$ has maximum trace.

B. Nonlinear Refinement

The next step of our anchors' initialization routine is a nonlinear refinement of the first coarse solution based on the Levenberg-Marquardt (LM) algorithm, executed on the same data collected during the first random trajectory. In particular, we seek to minimize the following cost function:

$$\min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{f}(\mathbf{x})\|^2, \quad (5)$$

where \mathbf{x} is the vector of unknown parameters to be estimated, \mathbf{y} is the vector of the measurement outputs, and $\mathbf{f}(\mathbf{x}) = \hat{\mathbf{y}}$ is the vector of estimated outputs. In contrast to the linear solution where only the constant bias γ was estimated for each anchor, in the nonlinear optimization problem we include the distance-dependent bias β for each anchor. The initial guess for the distance bias is set to 1.

In our problem formulation, the nonlinear relation between input and output is described by (2). For each anchor, the vector of parameters is defined to be $\mathbf{x} = [\mathbf{p}_A, \gamma, \beta]$ where \mathbf{p}_A is the anchor position in the global reference frame, γ and β respectively the constant and distant-dependent measurement bias. The LM algorithm iteratively updates the values of \mathbf{x} by linearizing the nonlinear function $\mathbf{f}(\mathbf{x})$ around the current estimate \mathbf{x}_k , and solving the resulting linear system $\mathbf{J}_k^T (\mathbf{y} - \mathbf{f}(\mathbf{x}_k)) \approx \mathbf{0}$ where \mathbf{J}_k is the Jacobian matrix at the current estimate. The algorithm adjusts the step size by introducing a scalar factor λ_k that trades off between the steepest descent direction and the Gauss-Newton direction. The updated estimate is given by $\mathbf{x}_{k+1} = \mathbf{x}_k + \delta \mathbf{x}$ where:

$$\delta \mathbf{x} = (\mathbf{J}_k^T \mathbf{J}_k + \lambda_k \text{diag}(\mathbf{J}_k^T \mathbf{J}_k))^{-1} \mathbf{J}_k^T (\mathbf{y} - \mathbf{f}(\mathbf{x}_k)).$$

The algorithm stops when either the change in the objective function $\|\mathbf{y} - \mathbf{f}(\mathbf{x})\|^2$ is below a certain tolerance or the maximum number of iterations is reached. The estimated parameters covariance matrix of the final solution is:

$$\hat{\Sigma} = \text{MSE}(\mathbf{H} + \lambda_k \text{diag}(\mathbf{H}))^{-1}.$$

C. Optimal Waypoints

Estimating the position of UWB anchors using the ranging measurements collected during a random trajectory has noticeable limitations despite the above introduced *optimal double method*. The optimality of the trajectory with respect to information for anchor initialization can be evaluated based on different measures. In [20], a first estimate of the anchor's position is refined by collecting data navigating through a set of waypoints computed via maximization of the Fisher information matrix. In this study, GDOP [21] is employed as a measure of optimality for estimating the position of multiple anchors *simultaneously*. A trajectory that minimizes the UWB mapping uncertainty is synthesized as follows. Let \mathbf{H} be the Jacobian in terms of the cosine directions of the nonlinear ranging measurements in Equ. (2). Considering one UWB anchor A and n_p waypoints we have

$$\mathbf{H} = \begin{bmatrix} \cos \varphi_1 \cos \gamma_1 & \cos \varphi_1 \sin \gamma_1 & \sin \varphi_1 & 1 \\ \cos \varphi_2 \cos \gamma_2 & \cos \varphi_2 \sin \gamma_2 & \sin \varphi_2 & 1 \\ & \vdots & & \\ \cos \varphi_{n_p} \cos \gamma_{n_p} & \cos \varphi_{n_p} \sin \gamma_{n_p} & \sin \varphi_{n_p} & 1 \end{bmatrix}, \quad (6)$$

where

$$\gamma_i = \arctan \left(\frac{G\mathbf{p}_{A_y} - q_y^{(i)}}{G\mathbf{p}_{A_x} - q_x^{(i)}} \right),$$

$$\varphi_i = \arctan \left(\frac{G\mathbf{p}_{A_z} - q_z^{(i)}}{\sqrt{(G\mathbf{p}_{A_x} - q_x^{(i)})^2 + (G\mathbf{p}_{A_y} - q_y^{(i)})^2}} \right),$$

are, respectively, the Azimuth and Elevation angles of the i -th waypoint position $q^{(i)} = [q_x^{(i)}, q_y^{(i)}, q_z^{(i)}]$ with respect to the UWB anchor $G\mathbf{p}_A$. Now considering a minimum of four target positions, we can calculate the GDOP as follows:

$$\text{GDOP} = \sqrt{\text{Trace}(\mathbf{H}^T \mathbf{H})^{-1}} = \sqrt{\text{Trace} \left[\frac{\text{adj}(\mathbf{H}^T \mathbf{H})}{\det(\mathbf{H}^T \mathbf{H})} \right]}, \quad (7)$$

with

$$\det(\mathbf{H}^T \mathbf{H}) = (h_1 - h_2 + h_3 - h_4)^2, \quad (8)$$

where

$$h_t = -C_{\gamma_t} C_{\varphi_t} (C_{\varphi_k} S_{\gamma_k} \delta s_{ji} + C_{\varphi_j} S_{\gamma_j} \delta s_{ik} + C_{\varphi_i} S_{\gamma_i} \delta s_{kj}),$$

$$\delta s_{ij} = S_{\varphi_i} - S_{\varphi_j}, \quad C_{\gamma_i} = \cos(\gamma_i), \quad S_{\gamma_i} = \sin(\gamma_i),$$

$$C_{\varphi_i} = \cos(\varphi_i), \quad S_{\varphi_i} = \sin(\varphi_i),$$

and $\text{adj}(\mathbf{H}^T \mathbf{H})$ is readily defined by the co-factor of $\mathbf{H}^T \mathbf{H}$. Maximizing (8) is equivalent to minimizing the volume of the estimation uncertainty ellipsoid, which is a function of only the Azimuth and the Elevation angles of the waypoints with respect to the UWB anchor position, thus indicating those angles as the optimization parameters. Despite this remark, multiple anchors should be considered as well. In addition, operational search volumes for the waypoints can have any arbitrary shape (with UWB anchor locations potentially outside of the drone operational volume). As such, the waypoints determination turns into a challenging optimization problem, especially in unstructured environments. Moreover,

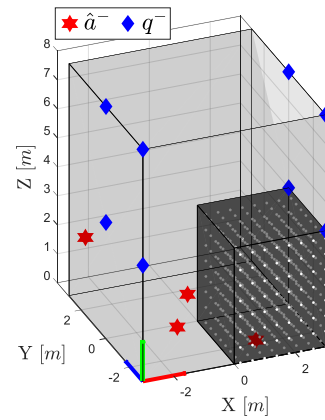


Fig. 2. The sample representation of the drone operational volume with 8 initial waypoints (i.e. q^-). The cube volume is divided into 8 smaller cubes (3D grids), each one indicating all the possible positions (denoted by small white dots) for the corresponding waypoint.

instead of optimizing multiple trajectories separately for each anchor, as in [20], we are interested in generating one optimal trajectory that minimizes the estimation accuracy for all the UWB anchors simultaneously. To this end, instead of angles, we consider the waypoint positions directly as the parameters of the following optimization problem

$$\min_{\mathbf{q}} \bar{g}(\mathbf{q}), \quad \text{s.t. } \mathbf{I}(\mathbf{q}) = 1, \quad (9)$$

where:

$$\bar{g}(\mathbf{q}) = \frac{1}{n_a} \sum_{i=1}^{n_a} \sqrt{\text{Trace}(\mathbf{H}_i^T \mathbf{H}_i)^{-1}},$$

with Jacobian \mathbf{H}_i given in (6) and computed for all the n_a anchors, i.e., $\forall i \in [1, n_a]$, and $\mathbf{I}(\mathbf{q})$ is an indicator function to check the feasible volume of the waypoints (i.e., a 3D polygon [22]) and being $\mathbf{I}(\mathbf{q}) = 1$ if a point is inside the volume, $\mathbf{I}(\mathbf{q}) = 0$ otherwise. To tackle this problem, we applied the grid-based evolutionary algorithm in [23]. The algorithm starts with the drone operational volume considered as a cube with height h , width w and length ℓ , and n_p initial waypoints (i.e. $q^{(1,-)}, \dots, q^{(n_p,-)}$) (see Fig. 2). The cube volume is divided into $n_p = n_{h_p} \times n_{\ell_p} \times n_{w_p}$ smaller cubes, where n_{h_p} , n_{ℓ_p} and n_{w_p} indicate the number of coordinate variations along Z (height), X (length) and Y (width) axis. Each small cube is a 3D grid denoted by $\mathcal{G}_1 \dots \mathcal{G}_{n_p}$, which defines all the possible positions for the corresponding waypoint. The evolutionary Alg. 1 has each chromosome defined as a $3 \times n_p$ matrix whose genes indicate 3 indexes along the three axes of a sample grid point for the corresponding waypoint. Each chromosome represents one particular configuration of the set of waypoints in the discrete space (see Fig. 3). As shown in Alg. 1, the Grid-Based Evolutionary Algorithm has as input the 3D environment configuration parameters and its center c_v , the initial estimates of the UWB anchors (i.e. $G\hat{\mathbf{p}}_A$) and the current position of the UWB tag (i.e. $G\hat{\mathbf{p}}_{U_k}$ at time t_k), and returns the optimal set of waypoints (i.e. $q_1^*, \dots, q_{n_p}^*$) that minimizes (7).

Here, we report the results of applying Alg. 1 on two sample anchor configurations for two different random configurations of UWB anchors, reported in Fig. 4. The algo-

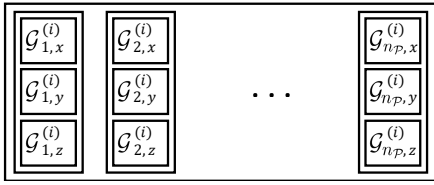


Fig. 3. Schematic representation of a sample chromosome in the i th iteration of the Grid-Based Evolutionary Algorithm. $\mathcal{G}_{j,x}^{(i)}$, $\mathcal{G}_{j,y}^{(i)}$, $\mathcal{G}_{j,z}^{(i)}$ indicate the 3 indexes (xyz) of a sample grid point for the j th waypoint.

Algorithm 1 Grid-Based Evolutionary Algorithm for estimating optimal set of waypoints

Input : $(h, \ell, w, c_v, \hat{G}_{\mathcal{P}_A}, \hat{G}_{\mathcal{P}_{U_k}}, n_p = n_{h_p} \times n_{\ell_p} \times n_{w_p})$
Output : $(q_1^*, \dots, q_{n_p}^*)$
 $\mathcal{G}_{[n_{h_p} \times n_{\ell_p} \times n_{w_p}]} = \text{GenerateCubeGrid}(h, \ell, w, c_v, n_p)$
Require : $n_p, \mathcal{P} = [\mathcal{C}_1, \dots, \mathcal{C}_{n_{pop}}], \forall i \mathcal{C}_i \in \mathcal{G}$
 $\mathcal{P} \leftarrow \text{Initialize}(\mathcal{P})$
while termination criterion not fulfilled **do**
 $\mathcal{F} \leftarrow \text{Fitness_assignment}(\mathcal{P})$ \triangleright based on Equ. (9)
 $\mathcal{P}' \leftarrow \text{Mating_selection}(\mathcal{P})$
 $\mathcal{P} \leftarrow \text{Variation}(\mathcal{P}')$ \triangleright Cross-over, Mutation
end while
 $q^* \leftarrow \min_{\mathcal{C}} \mathcal{F}$

rithm was executed with 2000 maximum evolution of the genetic algorithm, tuned with a population size equal to 40, crossover probability of 0.6, mutation probability 0.3 and elitism probability of 0.1. The initial set of 8 waypoints was used which results in 8 sub-cubes as shown in Fig. 2, and the configuration space of each waypoint has 20^3 different positions. The calculated optimal waypoints for the anchor configuration of Fig. 4 (a) have the cost $\bar{g}(q) = 1.4$, while for the corresponding anchor configuration of Fig. 4 (b) $\bar{g}(q) = 1.18$. In order to have an idea of the repeatability of the solution provided by the Grid-Based Evolutionary Algorithm and of the corresponding execution time, we have carried out Monte Carlo (MC) simulations for the two sample anchor configurations shown in Fig. 4 with 500 trials. For the computation time analysis, we have used the same MC test

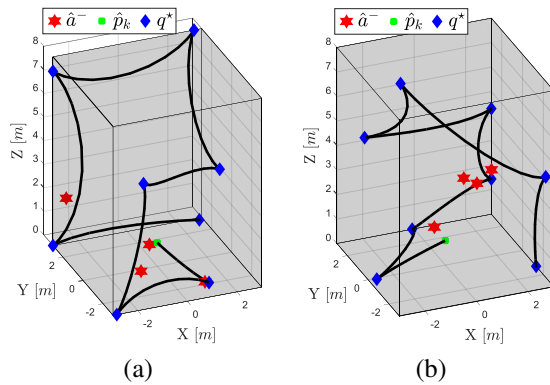


Fig. 4. Two Optimal sets of waypoints estimated as the solutions of the Grid-Based evolutionary algorithm in two different random configurations of UWB anchors. The black curves show a minimum snap trajectory generated for the corresponding waypoints with zero velocity boundary conditions.

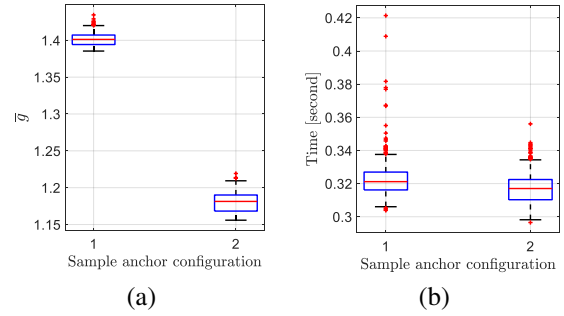


Fig. 5. Average GDOP and time distribution analysis for the two sample configuration cases (1) Fig. 4(a) and (2) Fig. 4(b).

and called the algorithm, which is written as a static C++ library and interfaced with Matlab Mex-Compiler, in each iteration of the MC trial. Fig. 5 shows the average GDOP and time distribution analysis for the two sample configuration cases (1) Fig. 4(a) and (2) Fig. 4(b).

After obtaining the set of optimal waypoints, the UAV can fly through them and collect data that will then be used again in a nonlinear refinement explained in Sec. IV-B on the previously obtained solution. With this last step, we best leverage the information contained in the optimal waypoints previously calculated such that a precise initialization of all UWB anchors in range can be performed *simultaneously*.

V. EXPERIMENTS

A. Simulation

The initialization procedure was simulated and extensively tested in MATLAB, with a comprehensive exploration of multiple scenarios with different anchors configuration, UAV trajectories, parameter values, and noise levels. Particularly, we compared the solutions of the initialization after each step and against random waypoints, i.e., initial coarse least squares solution, nonlinear least squares on the same initialization data, refinement of the nonlinear least squares via random waypoints, and via optimal waypoints. It is important to underline that the refinement via random and optimal waypoints is performed on "new" data with respect to the first two methods. For fairness of comparison, we considered the same amount of samples for all the methods. In all cases, the refinement procedure via optimal waypoints exhibited superior performance compared to other tested methods (Fig. 6) with notably improved minimization of both variance and mean error of the solutions. This outcome underscores the efficacy of incorporating an information-theoretic supported optimization-based approach for refining the initial solution of the anchor's mapping, increasing the reliability of the initialization process even in situations with challenging anchor configurations.

B. Real-world experiments

In this section, we describe the real-world experiments conducted to validate the effectiveness of UVIO and the proposed multi-step initialization procedure. The experiments were carried out using an UAV equipped with an onboard flight computer (Raspberry Pi 4), an IMU, a camera

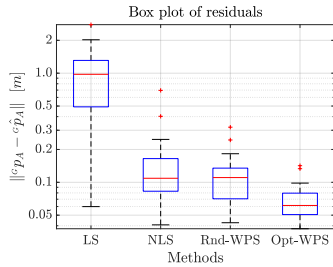


Fig. 6. Average position error for 10 different MC simulations of the full initialization procedure, with 100 realizations each, in a flight volume of $6 \times 6 \times 7 \text{ m}^3$. Each simulation has a different initial trajectory and 4 or 5 randomly placed UWB anchors with different biases. The measurement and position noise values are respectively $\sigma_y = 0.3 \text{ m}$ and $\sigma_p = 0.03 \text{ m}$.

(Matrix Vision BlueFOX), and an UWB transceiver (Qorvo MDEK1001). The vehicle was flown in a large indoor environment within a flight volume of $4 \times 6.5 \times 7 \text{ m}^3$ equipped with an Optitrack motion capture system for recording the ground-truth poses. Four additional UWB modules (Qorvo MDEK1001) were arbitrarily placed in the environment and their ground truth positions were captured. The system-autonomy framework introduced in [24] was used to enable the UAV to autonomously fly through waypoints inside the given flight volume. For the first set of experiments, we deployed the UAV at the center of the flight volume and executed the initialization procedure as described in Sec. IV. For the purpose of evaluating the accuracy of the proposed method, we used the ground-truth position of the UAV to perform the initialization. We performed three experiments with two configurations for the anchors and the corresponding position errors are shown in Tab. I. The increase in error compared to the simulation experiments is explained by the non-ideality of the UWB measurement model that assumes constant values for the biases during the initialization phase, while in reality interference, reflections, and other unwanted physical phenomena may change their values. Moreover, the results of the experiments show that the accuracy of the proposed method is influenced by the

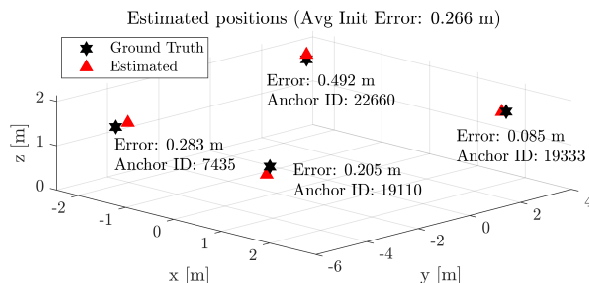


Fig. 7. Experiment 3: initialization

Anc. ID	19333	22660	7435	19110	Avg.
Exp. 1	0.169 m	0.470 m	0.271 m	0.193 m	0.276 m
Exp. 2	0.153 m	0.225 m	0.150 m	0.316 m	0.211 m
Exp. 3	0.085 m	0.492 m	0.283 m	0.205 m	0.266 m

TABLE I
INITIALIZATION ERRORS AND AVERAGES.

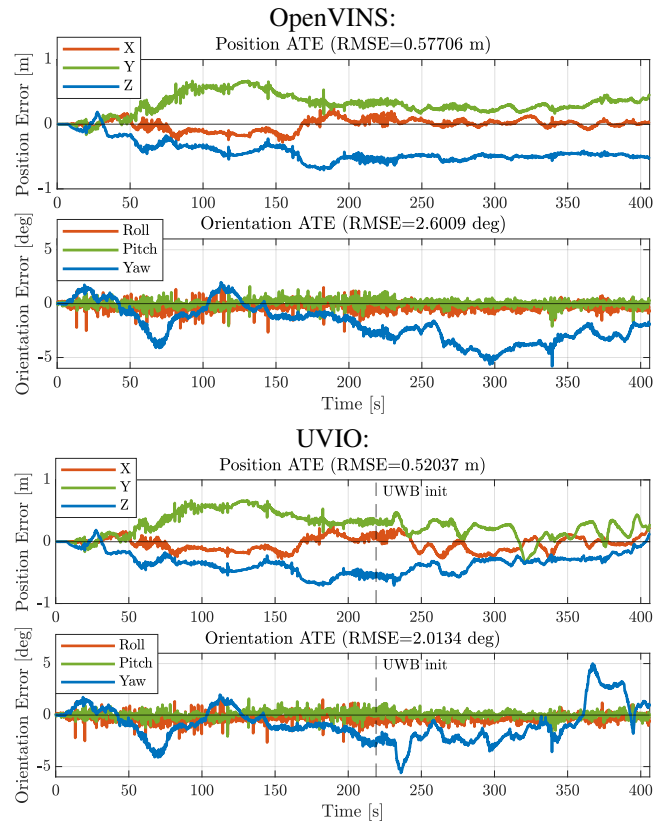


Fig. 8. Real-world experiment: comparison between OpenVINS and UVIO position and orientation Absolute Trajectory Error (ATE). The Root Mean Square Error (RMSE) of the error is reduced of approximately 10% in position and 20% in orientation

placement of the anchors with respect to the flight volume. In particular, anchors that are located outside of the volume and in a corner, result in higher initialization errors. On the other hand, anchors that are placed in favorable locations, such as near the center of a face of the flight volume, produce much lower initialization errors. A comparison with the results from [20] suggests that the proposed approach is superior since it considers multiple anchors simultaneously and is capable of achieving a comparable or even lower initialization error for single anchors that are optimally placed, such as for anchor 19333 in Fig. 7.

A second set of experiments was carried out to evaluate the real-time performance of UVIO. This time, the initialization was performed considering the poses estimated by the VIO framework and stopped with the UAV hovering in place. This was followed by a manual flight of roughly the same duration, allowing for more agile maneuvering compared to the automatic flight through waypoints. For the evaluation of the trajectories, we chose the Absolute Trajectory Error (ATE) as a metric [25] for the assessment of the UVIO performance. In all the experiments, we noticed an overall improvement in localization performance using UVIO compared to a VIO-only. This was reflected in a decrease of the ATE Root Mean Square Error (RMSE) (Fig. 8). To further test and validate the reliability of our framework when compared to pure VIO we simulated a camera dropout of 10 seconds after the initialization of the anchors. As shown in Fig. 9, after an

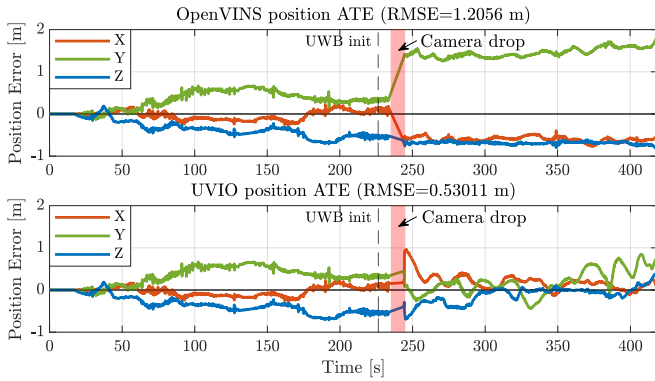


Fig. 9. UVIO and OpenVINS performance comparison with camera faults.

initial transient caused by the VIO settling, UVIO is capable of fully recovering from the dropout, making it a reliable framework also in situations of poor visual conditions or camera faults, where pure VIO would fail.

VI. CONCLUSION

In conclusion, this paper proposes UVIO, a multi-sensor framework that combines UWB technology and VIO to provide reliable and low-drift localization for mobile robots and UAVs operating in GNSS-denied environments. The proposed information-driven initialization procedure enables the automatic detection and mapping of multiple unknown anchors even in situations with challenging anchor configurations. The effectiveness of UVIO and the proposed initialization procedure is validated through simulations and real-world experiments, proving it to be a complete, easily deployable solution for real-time state estimation in challenging scenarios. Future work could focus on further improving the initialization procedure by including a robust outlier detection for received UWB range measurements and expanding the applicability of the proposed approach to other robotic systems.

REFERENCES

- [1] P. Geneva, K. Eickenhoff, W. Lee, Y. Yang, and G. Huang, "Opencvins: A research platform for visual-inertial estimation," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 4666–4672.
- [2] 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)*, 2017, pp. 1602–1609.
- [3] F. J. Perez-Grau, F. Caballero, L. Merino, and A. Viguria, "Multi-modal mapping and localization of unmanned aerial robots based on ultra-wideband and rgb-d sensing," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2017, pp. 3495–3502.
- [4] J. Tiemann, A. Ramsey, and C. Wietfeld, "Enhanced uav indoor navigation through slam-augmented uwb localization," in *2018 IEEE International Conference on Communications Workshops (ICC Workshops)*, 2018, pp. 1–6.
- [5] B. Yang, J. Li, and H. Zhang, "Uvip: Robust uwb aided visual-inertial positioning system for complex indoor environments," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 5454–5460.
- [6] S. Shin, E. Lee, J. Choi, and H. Myung, "Mir-vio: Mutual information residual-based visual inertial odometry with uwb fusion for robust localization," 2021. [Online]. Available: <https://arxiv.org/abs/2109.00747>

- [7] J. P. Queraltá, Q. Li, F. Schiano, and T. Westerlund, "Vio-uwbbased collaborative localization and dense scene reconstruction within heterogeneous multi-robot systems," in *2022 International Conference on Advanced Robotics and Mechatronics (ICARM)*, 2022, pp. 87–94.
- [8] J.-R. Zhan and H.-Y. Lin, "Improving visual inertial odometry with uwb positioning for uav indoor navigation," in *2022 26th International Conference on Pattern Recognition (ICPR)*, 2022, pp. 4189–4195.
- [9] H. Xu, Y. Zhang, B. Zhou, L. Wang, X. Yao, G. Meng, and S. Shen, "Omni-swarm: A decentralized omnidirectional visual-inertial-UWB state estimation system for aerial swarms," *IEEE Transactions on Robotics*, vol. 38, no. 6, pp. 3374–3394, dec 2022. [Online]. Available: <https://doi.org/10.1109/2ftro.2022.3182503>
- [10] H.-Y. Lin and J.-R. Zhan, "Gnss-denied uav indoor navigation with uwb incorporated visual inertial odometry," *Measurement*, vol. 206, p. 112256, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S026322412201452X>
- [11] K. Hausman, S. Weiss, R. Brockers, L. Matthies, and G. S. Sukhatme, "Self-calibrating multi-sensor fusion with probabilistic measurement validation for seamless sensor switching on a uav," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 2016, pp. 4289–4296.
- [12] Q. Shi, S. Zhao, X. Cui, M. Lu, and M. Jia, "Anchor self-localization algorithm based on uwb ranging and inertial measurements," *Tsinghua Science and Technology*, vol. 24, no. 6, pp. 728–737, 2019.
- [13] M. Ridolfi, J. Fontaine, B. Van Herbruggen, W. Joseph, J. Hoebeke, and E. De Poorter, "Uwb anchor nodes self-calibration in nlos conditions: a machine learning and adaptive phy error correction approach," *Wireless Networks*, vol. 27, 05 2021.
- [14] B. Gao, B. Lian, D. Wang, and C. Tang, "Low drift visual inertial odometry with uwb aided for indoor localization," *IET Communications*, vol. 16, no. 10, pp. 1083–1093, 2022. [Online]. Available: <https://ietresearch.onlinelibrary.wiley.com/doi/abs/10.1049/cmu2.12359>
- [15] T. H. Nguyen, T.-M. Nguyen, and L. Xie, "Tightly-coupled single-anchor ultra-wideband-aided monocular visual odometry system," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 2020, pp. 665–671.
- [16] —, "Range-focused fusion of camera-imu-uwb for accurate and drift-reduced localization," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 1678–1685, 2021.
- [17] T.-M. Nguyen, S. Yuan, M. Cao, T. H. Nguyen, and L. Xie, "Viral slam: Tightly coupled camera-imu-uwb-lidar slam," 2021. [Online]. Available: <https://arxiv.org/abs/2105.03296>
- [18] 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," *IEEE Transactions on Robotics*, vol. 38, no. 2, pp. 958–977, 2022.
- [19] S. Jia, Y. Jiao, Z. Zhang, R. Xiong, and Y. Wang, "Fej-viro: A consistent first-estimate jacobian visual-inertial-ranging odometry," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2022, pp. 1336–1343.
- [20] J. Bluemel, A. Fornasier, and S. Weiss, "Bias compensated uwb anchor initialization using information-theoretic supported triangulation points," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 5490–5496.
- [21] D. Fontanelli, F. Shamsfakhr, and L. Palopoli, "Cramer-Rao Lower Bound Attainment in Range-Only Positioning Using Geometry: The G-WLS," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–14, 2021.
- [22] K. Hormann and A. Agathos, "The point in polygon problem for arbitrary polygons," *Computational geometry*, vol. 20, no. 3, pp. 131–144, 2001.
- [23] S. Yang, M. Li, X. Liu, and J. Zheng, "A grid-based evolutionary algorithm for many-objective optimization," *IEEE Transactions on Evolutionary Computation*, vol. 17, no. 5, pp. 721–736, 2013.
- [24] M. Scheiber, A. Fornasier, R. Jung, C. Böhm, R. Dhakate, C. Stewart, J. Steinbrener, S. Weiss, and C. Brommer, "Cns flight stack for reproducible, customizable, and fully autonomous applications," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 11 283–11 290, 2022.
- [25] 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)*, 2018, pp. 7244–7251.