Indoor Drone Positioning: Accuracy and Cost Trade-off for Sensor Fusion

Jono Vanhie-Van Gerwen, Kurt Geebelen, Jia Wan, Wout Joseph, Jeroen Hoebeke, Eli De Poorter

***Abstract*—Indoor drone or Unmanned Aerial Vehicle (UAV) operations, automated or with pilot control, are an upcoming and exciting subset of drone use cases. Automated indoor flights tighten the requirements of stability and localization accuracy in comparison with the classic outdoor use cases which rely primarily on (RTK) GNSS for localization. In this paper the effect of multiple sensors on 3D indoor position accuracy is investigated using the flexible sensor fusion platform OASE. This evaluation is based on real-life drone flights in an industrial lab with mm-accurate ground truth measurements provided by motion capture cameras, allowing the evaluation of the sensors based on their deviation from the ground truth in 2D and 3D. The sensors under consideration for this research are: IMU, sonar, SLAM camera, ArUco markers and Ultra-Wideband (UWB) positioning with up to 6 anchors. The paper demonstrates that using this setup, the achievable 2D (3D) indoor localization error varies between 4.4 cm and 21 cm (4.9 cm and 67.2 cm) depending on the selected set of sensors. Furthermore, cost/accuracy trade-offs are included to indicate the relative importance of different sensor combinations depending on the (engineering) budget and use case. These lab results were validated in a Proof of Concept deployment of an inventory scanning drone with more than 10 flight hours in a 65 000 m^2 warehouse. By combining lab results and real-life deployment experiences, different subsets of sensors are presented as a minimal viable solution for three different indoor use cases considering accuracy and cost: a large drone with little weight- and cost restrictions, one or more medium sized drones, and a swarm of weight and cost restricted nano drones.**

***Index Terms*—drones, Unmanned Aerial Vehicles (UAV), sensor fusion, indoor positioning, Simultaneous Localization And Mapping (SLAM), ArUco markers, Ultra-Wideband (UWB) localization**

I. Introduction

Recent advances in drone manufacturing and research realized safe indoor drone flights in many cases, well beyond the enormous open spaces and human pilots previously required to fly most drones indoor. This opens up a broad spectrum of use cases: from industrial inspection [1], construction [2], security [3] and crowd sensing [4] up to warehouse inventory control [5] and interactive guides for visually impaired visitors [6]. However, due to the lack of a standardized one-size-fits-all indoor positioning solution, similar to the adoption of Global Navigation Satellite Systems (GNSS) in outdoor use cases, it is often a painstaking and error-prone process to identify and select the correct subset of sensors for any given use case. As with outdoor use cases where GNSS is not reliable, a robust sensor fusion system should be designed [7]. This paper aims to facilitate this process by giving a broad overview of different sensor combinations for three different drone classes that span a wide subset of possible indoor drone deployments. Note that these drones and use cases are not categorized by regulatory limits since none of the existing drone specific regulations apply to indoor use. However, the EASA Open Category for civil drones [8] gives a very suitable risk analysis based classification framework in three categories:

1) *Large drones* where equipment weight and cost is only a secondary consideration. These large drones most often span 40-100 cm in diameter and 2-25 kg in weight. They are used for use cases where flight clearances are possible up to several meters in either direction of the drone for safety considerations. These platforms allow for the full sensor set under evaluation to be installed at once. Due to the cost of these platforms, often a limited set of drones is deployed. This decreases the importance of optimizing the cost per drone, especially when accuracy or stability could be compromised.
2) *Medium sized drones* with limited maximum take-off weight and size, from 500 g up to 2 kg. These platforms often use off-the-shelf equipment and allow for a cost effective indoor drone fleet with ample possibilities in sensor selection with consideration of the engineering budgets. For these use cases a careful selection of a subset of the investigated sensors is appropriate.
3) *Small or nano drones* with very limited weight and power provisions are ideal for swarm-like use cases. Due to their limited weight up to maximally 500 g, these drones have lower accuracy requirements since a drone crash can have little to no consequences with a proper hardware design such as a drone cage or propeller guards. However, the platform restriction mandate a carefully investigated engineering budget and sensor selection for optimal operation in terms of flight time and stability.

The results presented in this paper are gathered by recording

Jono Vanhie-Van Gerwen, Jeroen Hoebeke and Eli De Poorter are with the IDLab group - Department of Information Technology, Ghent University, 9000 Gent, Belgium (e-mail: jono.vanhie@ugent.be; jeroen.hoebeke@ugent.be; eli.depoorter@ugent.be)

Kurt Geebelen en Jia Wan are with DecisionS - Flanders Make, 3001 Leuven, Belgium (e-mail: kurt.geebelen@flandersmake.be; jia.wan@flandersmake.be)

Wout Joseph is with the WAVES group, Department of Information Technology, Ghent University-imec, 9052 Ghent, Belgium (e-mail: wout.joseph@ugent.be)

This work was funded by the imec.icon InWareDrones project

TABLE I
COMPARISON OF RELATED INDOOR DRONE POSITIONING RESEARCH

paper	technologies	reported error	positioning rate	technology comparison	cost analysis
[9]	ultrasonic positioning system (5 beacons), external TOF camera	4-8 cm 3D (3 measured positions)	2 Hz	ultrasonic-only (14-35 cm 3D)	no
[10]	UWB (4 TDoA anchors)	11 cm 2D / 26 cm 3D (simulation)	50-70 Hz	no	no
[11]	IMU, external RGB-D camera	10 cm 3D (position losses observed)	10 Hz	no	no
[12]	UWB, visual SLAM	14 cm 3D	32 Hz	UWB-only (20.5 cm 3D)	no
[13]	IMU, Wi-Fi trilateration	111 cm 2D	not reported	no	cost-effective
[14]	5 external IR sensors and 1 camera	2.5 cm 2D (2 positions in 3x4 m area)	not reported	IR-only (4.7 cm 2D), camera-only (6.7 cm 2D)	low
[15]	IMU, Optical flow, UWB (8 TDoA anchors)	13 cm 2D / 26 cm 3D	40 Hz	no	no
[16]	UWB, RGB-D camera	$\leq$22 cm (x, y or z)	not reported	camera-only ($\leq$46 cm) UWB-only ($\leq$22 cm)	no
[17]	IMU, laser SLAM, visual SLAM	30 cm 2D	10-25 Hz	no	no
[18]	IMU, RGB-D camera (Kinect)	8 cm x-axis, 2D/3D not reported	30 Hz	no	no

real-life drone flights in an industrial lab environment. Fig. 1 depicts the followed research methodology, comprising of two main components: data capture and analysis. The measurements of each individual sensor are collected together with ground truth measurements using a Qualisys mm-accurate motion capturing system [19]. Afterwards the raw sensor data is processed by the sensor fusion platform using an Extended Kalman Filter (EKF) [20] without smoothing or backtracking to comply with the near real-time requirements of drone flights. The designed sensor fusion framework, named OASE (for Online Asynchronous State Estimation), supports defining sensor subsets and deals with variable update rates. OASE allows sensor selection during flight or reducing the sampling frequency of a particular sensor. More than 300 sensor configurations were investigated, with different sensor sets and sampling frequencies, with each configuration applying to one or more of the defined drone classes.

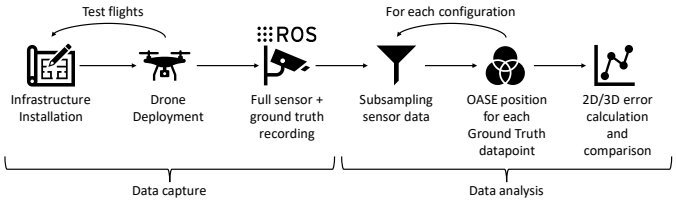


Fig. 1. Research methodology: data capture and data analysis

The main contributions of this paper are as follows:

- Determination of the achievable real-life 2D / 3D localization error with different sensor combinations
- A cost model for sensor configuration & installation is proposed
- Investigation of the accuracy / cost trade-offs for different sensor combinations and drone classes
- The impact of different deployment scenarios with varying numbers of Ultra-Wideband (UWB) anchor nodes and ArUco markers is discussed

The remainder of this paper is organized as follows. Section II discusses relevant related work. Section III describes the first part of our methodology: the testbed, drone and a detailed description of the chosen representative drone flight. The second part of our methodology is covered in Section IV, elaborating on the data analysis: the configuration of the sensor fusion platform, the developed cost model and EKF design for each sensor. Section V presents the results of the full parameter space exploration of all possible sensor combinations in terms of accuracy, stability and cost. Section VI describes a practical application scenario of deploying a Proof of Concept warehouse inventory scanning drone. Finally, Section VII summarizes the relevant results per defined drone class and presents the concluding remarks.

II. RELATED WORK

This research is a continuation of the experimental validation performed in [5] where an autonomous drone was used to evaluate the stability and power efficiency of a wireless backbone driven UWB indoor positioning system. We have expanded the evaluation framework previously used to incorporate the flexible OASE sensor fusion platform and extend the sensor set under evaluation to span more possible use cases and drone classes. Additionally, square fiducial markers are added next to the flight path based on the ArUco library [21] [22].

The OASE sensor fusion platform itself is designed by Flanders Make[1]. It implements a standard EKF for sensor fusion with separate observation functions for each sensor input. Internally, it deals with varying sampling frequencies and measurements that come with a delay by keeping track of a buffer of the last measurements, and rerunning the Kalman equations when required. It's worth to mention that the purpose of developing such a toolbox is not to re-invent the Kalman Filter, but to provide a user-friendly environment to develop and test different sensor fusion algorithms. This provides us with the required flexibility to adjust the sensor fusion to evaluate different scenarios.

Table I summarizes related research on indoor drone localization, with [17], [18] reviewed in [23] and [15], [16] covered in [24]. Our review shows a maximum of 3 different types of technologies for sensor fusion are considered, and only in a

[1]This work has been supported by the LCM – K2 Center within the framework of the Austrian COMET-K2 program.

few cases is the impact of individual sensors investigated. Cost is only explicitly mentioned in 2 papers, but never quantified. We go beyond these papers by considering up to 5 different sensors for sensor fusion, but also compare different subsets of these sensors and different sizes of UWB infrastructure. Finally, adding a cost breakdown of not only the sensor hardware, but also estimating the infrastructure costs is a novel approach for offering a more quantifiable cost-performance trade-off for sensor fusion.

III. Data Capture

All experiments presented in this paper were performed in the Industrial Internet of Things (IIoT) lab of the IDLab research group [25]. All flight and sensor data recorded are the result of real-life experiments with no simulated data present in the dataset. All flights were performed by a trained and licensed drone pilot due to the complexity in terms of testbed layout and limited flight height. By relying on an experienced pilot and manual drone control no a priori assumptions had to be made on which sensors would be reliable for drone control.

A. Testbed

The IIoT lab contains a 240m^2 representative warehouse environment, where the experiments presented in this paper were performed in the open space subsection of 6m (l) x 11m (w) x 2.5m (h). In this space, 8 Qualisys Miqus M3 motion capture (mocap) cameras are installed to provide mm-accurate ground truth measurements of any device under test. The drone itself is fitted with 10 passive infrared reflective mocap markers to allow for Six Degrees of Freedom (6DoF) tracking (x, y, z, yaw, pitch, roll) of the entire drone body even with multiple occlusions. Furthermore, 6 ArUco markers are placed around the flight space and equipped with three mocap markers so their position and rotation is also fully known in 6DoF. Finally, 6 UWB anchors are also placed at the edges of the area with a single mocap marker on top of each antenna providing an accurate position in Three Degrees of Freedom (3DoF), but no orientation since this has no relevance for multilateration using omnidirectional antennas. Fig. 2 details the full flight path, UWB anchor and ArUco marker positions.

B. Drone

All experiments are performed using a DJI M100 drone [26], allowing for a stable and expandable flight platform carrying all the required sensors under test. The stock flight controller is accessed by a ROS [27] interface on the on-board additional processing unit: an Intel NUC (NUC10i7FNKP). This powerful processing unit aggregates all the external sensors, together with the stock sensors on the DJI M100 such as the inertial measurement unit (IMU), magnetometer, barometer and GNSS. Since the magnetometer, barometer and GNSS are unreliable in an indoor context [28] [29], these are not included in the sensor fusion. Only the IMU data from the drone is used as the control input for the generic 6DoF rigid body motion equation. Additional sensors under test are the following:

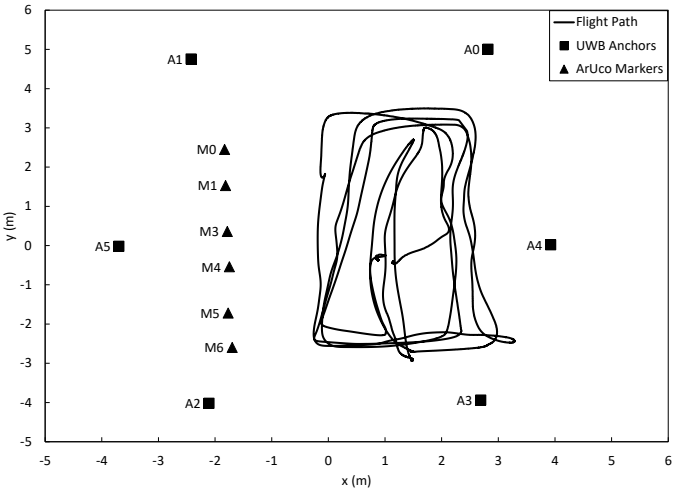


Fig. 2. Testbed layout of UWB anchors, ArUco markers and the flight path

1) *Sonar* at 50 Hz for height estimation as a cost-effective solution for stable indoor flight height estimation [30] since a barometer cannot be reliably used indoor. These sensors can be sensitive to dampening ground covering such as carpet, but this is less an issue in an industrial context. For more challenging environments a 1D LiDAR provides sub-centimeter accuracy, but this was not considered for this research since the sonar provides a more challenging sensor fusion case.
2) *Ultra-Wideband tag* using Two-Way-Ranging at 75 Hz between maximally 6 UWB anchors in the testbed. This sensor provides range information using the Time of Flight of the UWB signal between tag and each anchor. Through sensor fusion, combining the IMU data and these range measurements will result in 6DoF rigid body motion information (position, speed and orientation). The Ultra-Wideband devices are battery powered and use a lightweight wireless infrastructure [31], reducing the power consumption on the drone and allowing UWB anchors to be placed without any cabling considerations.
3) *Simultaneous localization and mapping (SLAM) camera*: The Intel Realsense T265 SLAM camera has an on-board processor for local visual SLAM processing and is used as a standalone sensor providing: position (x, y, z), orientation (yaw, pitch, roll) and speed (in the x-, y-, z-axis) of the SLAM camera at 200 Hz. This data can be used as a measurement of the drone's motion states, with the consideration of an installation offset of the SLAM camera on the drone, and a slow varying random drift of the position estimated by the sensor fusion.
4) *Cameras* with a global shutter for ArUco marker detection [32]. These cameras need to be calibrated carefully, meaning both its intrinsic and extrinsic parameters are known, and provide images at 30 Hz with little motion blur. The raw images are processed by an on-board marker detection algorithm. The sensor fusion platform can estimate the pose of the drone using a camera model and the 4 detected corners of each ArUco marker visible in the camera frame.

C. Measurement Campaign

More than 6 hours of drone flights were performed in the IIoT lab using the DJI M100, where the onboard processing unit collected all raw sensor data using ROS subscribers, together with 200 Hz ground truth updates including measurement delay from QTM [33]. All data stored in the on-board ROS bag was timestamped by the same system. We focus on a representative flight from these experiments that can be used for our three broad classes of drone platforms and related use cases. The annotated dataset and all relevant configuration parameters are made available through [34]. The flight at the basis of our analysis has a total duration of 215 seconds, with a stable flight height and multiple accelerations on the horizontal plane to investigate the sensor fusion's robustness to sudden position changes and high acceleration. For safety reasons indoor flight speeds did not exceed 1 m/s, with a target speed of 0.6 m/s. Fig. 3 gives an overview of the measured velocities during the flight and Table II gives more details on the dataset contents totalling 571979 data points.

TABLE II
DETAILED DATASET CONTENTS

sensor	**# data**	**data type**
ArUco markers	16783	4 corners positions with ArUco ID when detected
ArUco camera data	12962	Camera metadata (distortion model, D, K, R, P)
SLAM acceleration	13523	Linear acceleration x, y, z (+ covariance)
SLAM Gyroscope	43119	Angular velocity x, y, z (+ covariance)
SLAM odometer	43118	Odometry data (6DoF Pose, Twist + covariances)
UWB Ranges	2703	Array of UWB ranges to all 6 UWB anchors
DJI Flight attitude	21602	Quaternion rotation data (x, y, z, w)
DJI M100 IMU	21602	Orientation, angular velocity, linear acceleration
Ground Truth: ArUco	234215	6DoF body position
Ground Truth: drone	40434	6DoF body position

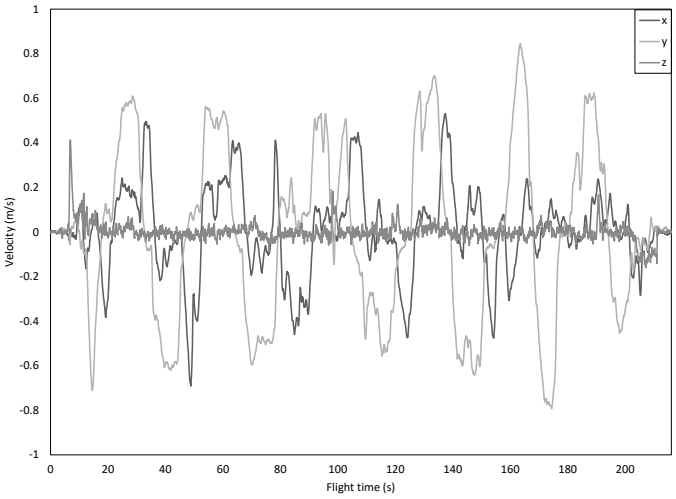


Fig. 3. Measured velocity on the x,y,z plane in m/s

IV. Data Analysis

After capturing both flight data and ground truth via motion capturing cameras, the configuration of the sensor fusion platform is performed offline and the results of the EKF are compared to the ground truth measurements for each configuration. A single configuration comprises the selected sensors, the update rate of the sensors under test and the model settings of the sensor fusion platform. For experiments using the UWB sensor the subset of UWB anchors used for ranging is also defined up to 6 active anchors. This methodology allows the evaluation of many different sensor fusion configurations based on realistic measurement data and noise of a single drone flight. For this research a total of 337 configurations were analyzed, corresponding to more than 20 flight hours.

A. Configurations

Based on the real-life measured data a wide selection of drone configurations are created, together with an estimated cost for each configuration. The costs taken into consideration are indications of material cost following industry standards, but depending on the use case this cost can vary based on the quality requirements. For example, the cameras taken into consideration for ArUco detection are global shutter cameras to limit motion blur due to drone movement and frame vibrations. If the ArUco positions aren't essential for drone stability a less expensive camera can be used when another source is used as primary positioning sensor, such as UWB.

Table III lists an overview of the most interesting configurations to be investigated further and to which drone classes these configurations can apply. All of these configurations also use the IMU sensor, which is not included in the table for brevity.

TABLE III
SENSOR CONFIGURATIONS AND APPLICABLE DRONE CLASSES

UWB	sonar	SLAM				ArUco	drone classes
		altitude	yaw	x,y	speed		
✓							1,2,3
✓	✓						1,2,3
✓		✓					1,2
✓			✓				1,2
✓		✓	✓				1,2
✓				✓			1,2
✓		✓	✓	✓	✓		1,2
✓						✓	1,2
✓	✓					✓	1,2
✓			✓			✓	1
✓	✓		✓			✓	1
✓		✓	✓			✓	1
✓		✓	✓	✓	✓	✓	1
			✓	✓		✓	1,2
		✓	✓	✓	✓		1,2
		✓	✓	✓	✓	✓	1,2

For each subset of sensor inputs under consideration a configuration file for the sensor fusion platform OASE was created, which then was run in offline mode and tasked to give a position and orientation update for each data point of the ground truth measurements during the flight (40434 points of comparison over 215 seconds of flight time).

B. Drone and infrastructure cost

Cost calculations that cover the entirety of a drone deployment are strongly use case specific and are dependant on the amount of drones deployed, the sensor set of each drone and the infrastructure costs. However cost is an important factor in drone and sensor selection, together with accuracy and stability. For this reason we've chosen to use a simple, but representative cost model based on industry standards to try and model both the drone and infrastructure costs. An overview of the additional drone costs per sensor and the typical ranges of quantities installed is listed in Table IV. We'll use the actual sensor cost for easier comparison in the rest of this paper, based on a practical use case of a warehouse inventory scanning drone.

TABLE IV
SENSOR COST

symbol	sensor	weight	power	cost	quantity
S_I	IMU	1 g	0.01 W	€ 10	$[1,2]Q_I$
S_{So}	sonar	10 g	0.1 W	€ 25	$[0,6]Q_{So}$
S_U	UWB	12 g	0.4 W	€ 100	$[0,2]Q_U$
S_{Sl}	SLAM	60 g	1.5 W	€ 200	$[0,1]Q_{Sl}$
S_A	ArUco	36 g	3 W	€ 800	$[0,2]Q_A$

This leads to the following straightforward cost calculation of a single drone (C_D), starting from the basic drone chassis cost (C_C) and adding the quantities of the selected sensors. This result should be extrapolated by the number of drones required for a specific use case:

$$C_D = C_C + Q_I \times S_I + Q_{So} \times S_{So} \times Q_U \times S_U + Q_{Sl} \times S_{Sl} + Q_A \times S_A \quad (1)$$

The infrastructure cost is harder to quantify since it depends heavily on the specific use case and the flight areas available. For example, a large open space of a 1000 square meters would only require 4 UWB anchors in the corners to position the drone adequately, while hundreds of ArUco markers would need to be placed to cover the entire area, and their positions need to be calibrated accurately, leading to excessive personnel costs. Alternatively a collection of racks can easily be fitted with ArUco markers on known locations (e.g., on the crossings of the horizontal and vertical beams that make up the rack structure, whose location is known), while every hallway would require UWB anchors due to the electromagnetic disruption of the metal racks and Non-Line of Sight signals, requiring additional intelligent processing and anchor node selection [35].

For a rack-based flying environment, such as a large warehouse, we have quantified the infrastructure cost as an example of the minimum and maximum cost differential between ArUco markers and UWB anchors based on a real-life use case at a large logistics company. Table V lists the variables taken into account and their default values used in the remainder of this paper.

The personnel cost of applying ArUco markers at one location was determined to be 2 minutes (T_A) over 4 heights. A maximal marker density would be every 0.8 m (A) for each EUR-pallet, with additional calibration for ArUco positioning, which takes 3 minutes per location (T_c) for a total personnel cost of 5 minutes per location. When performing sensor fusion of the SLAM camera with ArUco markers the density of the ArUco markers could be reduced to every 3.2 m and individual ArUco marker calibration was not required, leading to 2 minutes of work every 3.2 m of rack space.

TABLE V
INFRASTRUCTURE COST VARIABLES

sym	description	default		unit
		with SLAM	without SLAM	
T_A	ArUco marker installation time	2	2	min
T_c	ArUco calibration time	0	3	min
A	ArUco density	3.2	0.8	m
U	UWB anchor density per 100m	1	4	-
T_U	UWB calibration time per anchor	20	15	min
C_p	Personnel cost	50	50	€/h
C_U	UWB anchor cost	100	100	€
D_h	Dimensions: hallway width	3.6	3.6	m
D_r	Dimensions: rack depth	1.2	1.2	m

When considering UWB anchor deployment a maximal density of 4 anchors (U) for 100 m of rack space leads to € 400 of equipment cost, with 15 minutes of installation and calibration time per anchor (T_U). When combined with a SLAM camera the UWB anchor density can be reduced to only 1 anchor every 100 m and 20 minutes of personnel cost, only measuring the distance traversed between the racks. This leads to the following calculation to determine the infrastructure cost of ArUco and UWB deployments.

ArUco deployment cost per meter of hallway:

$$C_m^A = \frac{T_A + T_c}{60min} \times \frac{C_p}{A} \quad (2)$$

ArUco deployment cost per square meter of warehouse space:

$$C_{m^2}^A = \frac{C_m^A}{D_h + 2 \times D_r} \quad (3)$$

UWB deployment cost per meter of hallway:

$$C_m^U = \frac{U \times (C_U + T_U)}{100m} \quad (4)$$

UWB deployment cost per square meter of warehouse space:

$$C_{m^2}^U = \frac{C_m^U}{D_h + 2 \times D_r} \quad (5)$$

In Table VI these costs are detailed per meter of rack length and square footage of warehouse space assuming a personnel cost C_p of € 50/hour, UWB anchor cost C_U of € 100, a hallway of 3.6 m wide with rack depth of 1.2 m.

TABLE VI
WAREHOUSE INFRASTRUCTURE COST FOR ARUCO AND UWB

sensors	cost per m	cost per m^2
ArUco	€ 5.21	€ 0.87
ArUco + SLAM	€ 0.52	€ 0.09
UWB	€ 4.5	€ 0.75
UWB + SLAM	€ 1.17	€ 0.19

C. Sensor fusion platform OASE

The OASE (Online Asynchronous State Estimation) Toolbox is used to allow for a fast and flexible configuration of sensor fusion with variable rate sensor inputs. Its modular design enables an easy way to test different sensor configurations, with the most notable features being:

- *Low calculation time* by using CasADi [36] to analytically derive the system model, observation functions and their Jacobian matrices and generate efficient C code for these functions.
- *Complex data scheme handling*: by storing a buffer of the last N timestamped measurements, processing and data transmission delays are compensated in the toolbox, allowing asynchronous sensor inputs with different sample rates and delays. The measurement function of the system is split into separate observation functions for each sensor, such that the respective sensor inputs can be fed to the filter separately. Each sensor measurement is inserted into the buffer, which is sorted by the timestamp of the measurement. Measurements with a later timestamp than the latest measurement added to the buffer are reprocessed.
- *Easy to use* thanks to the Python front-end allowing it to work cross-platform with easily extensible configuration scripts and code, and automatic Jacobian calculation using CasADi. Additionally the toolbox contains commonly used vehicle and sensor models that can be used as-is or as the basis to add new customized models. Finally it contains several utility modules to help with post-processing and data analysis.

D. Sensor fusion design

A traditional EKF is implemented using the OASE toolbox for the drone states estimation. Since the goal of this paper is to compare different sensor combinations in terms of cost and accuracy, the choice for the EKF was made in favor of alternatives such as particle filters [37], Unscented Kalman Filters [38], moving horizon estimation [39], etc. due to its simplicity. The potentially lower estimation accuracy of the EKF compared to other sensor fusion techniques is not expected to have an impact on the relative comparison between different sensor configurations.

The system states include 6DoF rigid body motion states (X_b), as well as the initial position and yaw angle of the SLAM camera reference frame (X_s), as:

$$X_b = [x, y, z, u, v, w, \Psi, \Theta, \Phi]^T \tag{6}$$

$$X_s = [\Psi_s^b, x_s^b, y_s^b]^T \tag{7}$$

Where: x, y and z are the drone location coordinates among the x-, y- and z-axis. u, v and w denote the drone speed in the body reference frame among the x-, y- and z-axis. Ψ, Θ and Φ is the yaw, pitch and roll angle of the drone. Ψ_s^b is the yaw angle shift between the SLAM reference frame and the local frame. x_s^b and y_s^b are the x- and y-positions of the SLAM reference frame in the local frame.

The X_s is required to use the SLAM camera position and yaw data since the SLAM location reference frame is located at the position where the SLAM camera is initialized. By adding those states, the drone can start flying at an arbitrary position.

The 6Dof rigid body kinematic equation is used to calculate the $\dot{X}_b$ [40], while using the IMU data as input, i.e., three linear accelerations and three angular speeds. X_s is assumed to be constant ($\dot{X}_s = 0$).

For this system, observation functions are defined for each sensor mentioned in Table III.

1) UWB: The observation function for the UWB sensor is a vector of Euclidean distances between the tag (on the drone) and all anchors in the environment [41].

2) Sonar: The sonar provides the distance between the sensor and the ground where it points to. The observation function can be derived as:

$$Y_{sonar} = \frac{-z}{\cos\Phi\cos\Theta} - z_{sonar} \tag{8}$$

where z_{sonar} is the z-coordinate of the sensor installation position in drone body frame.

3) SLAM altitude: The SLAM altitude is related to its initialization position, which normally does not correspond to the ground reference. A constant shift of the altitude should be taken into account for the observation function, as:

$$Y_S^h = -z - h_{s0} \tag{9}$$

where h_{s0} is the altitude of the drone at SLAM camera initialization.

4) SLAM yaw: The observation function for SLAM yaw data is:

$$Y_S^{yaw} = \Psi - \Psi_S^b \tag{10}$$

5) SLAM speed: As the SLAM camera is not mounted at the origin point of the drone frame, the linear motion of the SLAM camera is a combination of the linear motion and the rotation of the drone. However, considering that drone motion in indoor drone use cases will not be aggressive, and the SLAM camera is relatively close to the origin point, the effect of the rotation of the drone can be neglected in the observation function, i.e., the speed of the SLAM camera is considered to be the same as the speed of the drone. The observation function is:

$$Y_S^{spd} = [u, v, w]^T \tag{11}$$

6) SLAM x-,y-position: To calculate the 2D position of the drone (x-,y-coordinate) in the SLAM frame ($x_d^{(SLAM)}, y_d^{(SLAM)}$), the initial location and the yaw angle of the SLAM frame related to local frame is required. The observation function is:

$$Y_S^{xy} = \begin{bmatrix} x_d^{(SLAM)} \\ y_d^{(SLAM)} \end{bmatrix} \tag{12}$$

$$= R_{g2SLAM} \times \left(\begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} x_s^b \\ y_s^b \end{bmatrix} \right) \tag{13}$$

where R_{g2SLAM} is the rotation matrix between ground frame and the SLAM frame, which depends on Ψ_s^b.

7) ArUco Marker position: Using a marker detection algorithm, the location of an ArUco marker in camera images can be detected. The output of the detection is a vector of pixel coordinates for the four corners of a marker.

To derive the observation function for the marker detection, the classic pinhole camera model is used. The camera needs to be calibrated for both the intrinsic and extrinsic parameters, including:

- P: the projection matrix of the camera
- d: the distortion parameters of the camera
- R_{d2c}: the homogeneous rotation and transformation matrix from the drone body frame to the camera frame, i.e., the extrinsic parameters.

Meanwhile, the following information for each marker in the environment should also be known: the marker position (x_m, y_m, z_m), orientation (Ψ_m, Θ_m, Φ_m), and marker size (d_m).

The coordinates of all four corners in the marker frame are known, which can be written in matrix form as

$$\frac{d_m}{2}\begin{bmatrix}-1 & 1 & 1 & -1\\ 1 & 1 & -1 & -1\\ 0 & 0 & 0 & 0\end{bmatrix}$$

where each column represents one corner location ($P_{co}^{(m)}$).

Using the pose of the marker and the pose of the drone, the transformation matrix from the marker frame to the local ground frame (R_{m2g}) and the matrix from local ground frame to drone body frame (R_{g2d}) can be obtained.

Combining those transformations, the position of one marker corner in the camera frame can be calculated as:

$$\begin{bmatrix}P_{co}^{(c)}\\ 1\end{bmatrix} = R_{d2c} \times R_{g2d} \times R_{m2g} \times \begin{bmatrix}P_{co}^{(m)}\\ 1\end{bmatrix} \tag{14}$$

With the known corner position in the camera frame and camera intrinsic parameters, the projection model [42] can be applied to obtain the pixel coordinates of all four corners.

V. Results

In this section the results of the different sensor configurations are presented and analyzed. We'll first present a broad overview of the possible sensor configurations that allow for a stable and reliable flight within our parameters. Afterwards, we consider the configuration costs and which trade-offs can be made for each drone class. Finally, the UWB results are further investigated on flight accuracy and stability when a subset of the 6 UWB anchors is used.

Note that the results are based on the deviation of the calculated position in comparison to the ground truth measurement by the mocap system. Although mocap systems can be used for drone positioning and control, the scalability of such systems in an industrial context with Non-line-of-sight (NLOS) conditions is prohibitively expensive. However, for use cases where mocap control is feasible the following results can be interpreted to show the potential gain in accuracy when switching to mocap positioning and control versus the evaluated sensor configurations.

A. General results

Fig. 4 shows the average 2D and 3D error for the configurations selected in Table III together with the respective standard deviations over the course of the flight. A full parameter space exploration of all possible sensor configurations was performed, but only the results of the sensor fusion algorithm that remain within the flying area boundary will be considered. The following sensor configurations did not lead to stable flight results: IMU only flight, SLAM without altitude, ArUco only, Sonar + ArUco, SLAM (yaw only) + ArUco, SLAM (altitude and yaw only) + ArUco.

A first general conclusion is that most stable sensor configurations deliver an error of 10 cm or better in the 2D plane, while only 2 configurations exceed a 20 cm average 2D error due to using ArUco markers without an accurate yaw source. This leads to inaccuracies in the marker detection algorithm. For some use cases this 2D error allows a broad range of sensor configurations in case a relatively accurate altitude sensor can be introduced to the drone. The importance of height estimation for drones is clear when considering the presented 3D error. With the UWB-only configuration, an average 3D error of 67 cm is achieved since the UWB anchors in the test area cannot be differentiated in the z (height) plane due to the limited ceiling height. Adding a relatively inexpensive altitude sonar to the drone immediately improves the average 3D error notably to 10.71 cm and the 2D error slightly to 9.25 cm (from 67.24 cm 3D and 12.54 cm 2D).

1) UWB synergies with other sensors: Comparable results to a sonar height estimation are achieved when using the altitude measurements from the SLAM camera together with UWB. However, other sensor sources are also observed to indirectly affect the 3D error of a UWB setup due to interesting complementary effects:

- SLAM xy readings reduce the 3D error from 67.24 cm to 45.29 cm, a 33% increase of accuracy by greatly reducing the 2D error by 42% and thus improving the implicit multilateration process of the UWB localization by reducing the intersections of the measured distance spheres.
- adding ArUco markers around the flight path reduces the 3D error to 13.17 cm with a larger standard deviation of 14.85 cm due to the drift of the IMU sensor between ArUco marker detections. However, this configuration needs an accurate yaw sensor, since otherwise the ArUco marker detection algorithm can introduce significant errors that negate the improvements made to the 3D position error. In industrial environments the magnetometer and IMU acceleration measurements are clearly not accurate enough to be a reliable yaw source. This is shown in the UWB + ArUco configuration that reduces the 3D error by 59% to 27.91 cm but also drastically increases the 2D error by 68% to 21.02 cm with a standard deviation of 31.57 cm. Even when adding an accurate height estimation through sonar the instability in the 2D plane leads to marginally better results in the 3D plane as well, only improving by 6.76 cm to a 21.15 cm average 3D error. It becomes clear that more fused

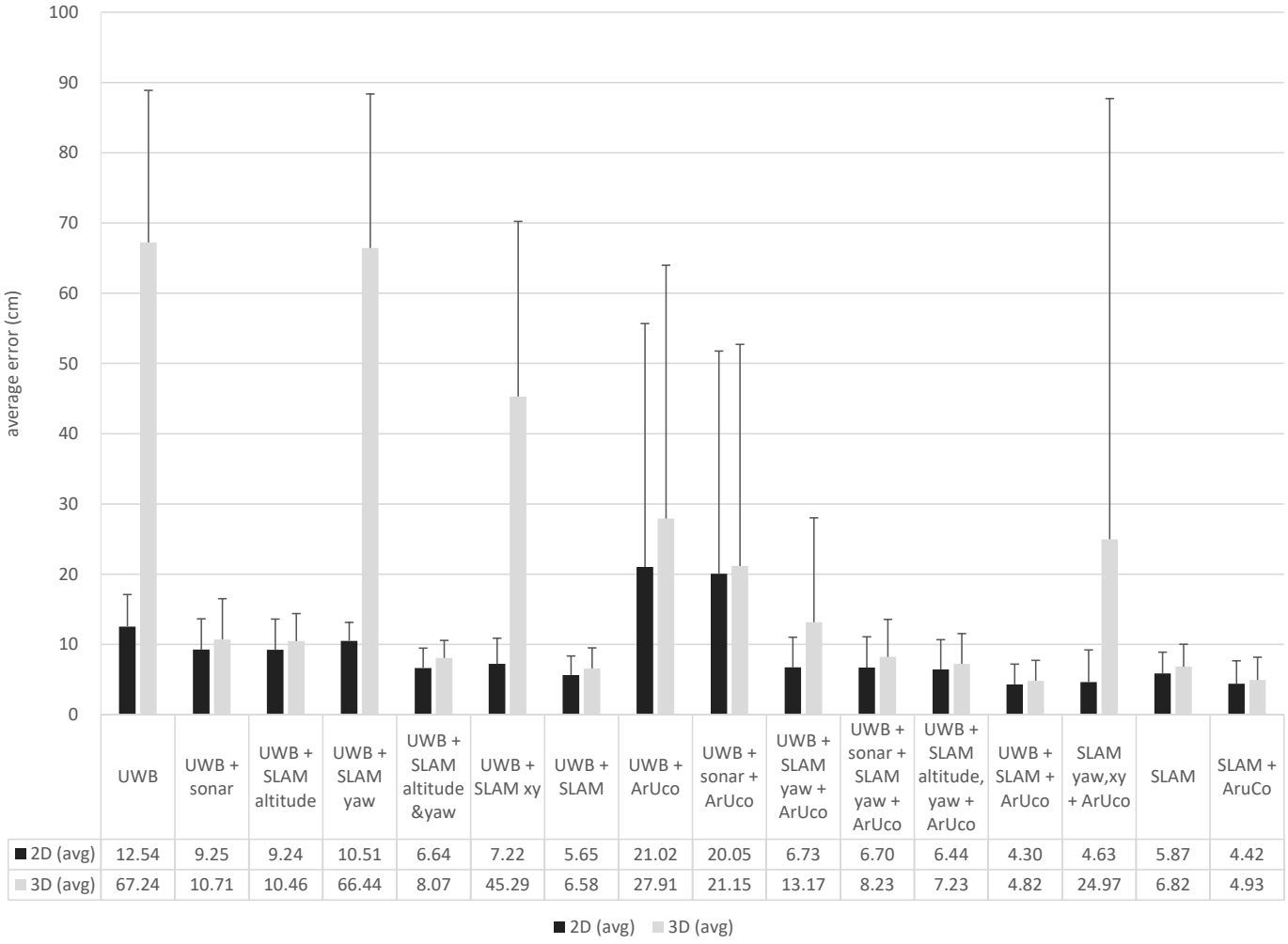


Fig. 4. average 2D and 3D error of different sensor fusion configurations (UWB = Ultra-WideBand, SLAM = Simultaneous localization and mapping)

sensors do not always lead to better results as simply adding a sonar to a UWB setup without ArUco markers leads to a reduction of the average 3D error from 56.53 cm to 10.71 cm.

2) SLAM camera synergies: A similar effect of indirect height estimation is observed when combining ArUco markers and the SLAM xy+yaw sensor. Here, the average 3D error is 24.97 cm, but with a large standard deviation of 62.47 cm. This leads to a very unreliable flight height, which would only be acceptable at sufficient flight heights of multiple meters. This is comparable to the inaccuracies of a pure UWB setup, with the ArUco markers providing absolute positioning information, while the SLAM camera retains relative positioning stability. This complementary effect of a stable, but drifting relative positioning sensor and an intermittent, but absolute positioning sensor shows the real strengths of a good sensor selection for fusion.

When adding altitude measurements to the SLAM and ArUco sensor combinations the average positioning error drops well below 10 cm on both the 2D and 3D metrics. With an average 3D error of 4.93 cm it is only 0.1 cm less accurate than the full sensor set solution of 6 UWB anchors, SLAM and ArUco markers. In flight areas where reliable updates of the ArUco markers are guaranteed, this solution can be a valid alternative for UWB positioning. However, the starting position of the drone flight should be fixed or at least well-defined since the absolute positioning of UWB is unavailable. This is even more important for the SLAM only flight, while the average accuracy seems very comparable with a 6.82 cm error, this sensor configuration works only for relatively short flights and with a very accurately defined starting position in 6DoF (x, y, z, yaw, pitch, roll).

TABLE VII
EFFECT OF ACCURATE STARTING POSITION ON FLIGHT ACCURACY

configuration	3D error		
	aligned	misaligned	relative effect
UWB + SLAM	6.58 cm	8.90 cm	35%
UWB + SLAM + ArUco	4.82 cm	6.65 cm	38%
SLAM	6.82 cm	111.40 cm	1533%
SLAM + ArUco	4.93 cm	8.63 cm	75%

The importance of a well-defined starting position when using a SLAM camera as part of the sensor fusion configuration is shown in Table VII. Here we compare the average 3D error of a flight with an accurate starting position versus a misaligned starting position by 100 cm in the x,y plane.

When only using a SLAM camera for position estimation, this starting error propagates through the entire flight as expected. The effect is also noticeable when adding absolute positioning sensors to the configuration although it does not propagate through the entire flight. From the results it is clear that adding UWB positioning to the fused sensors provides the best robustness to misaligned starting positions due to the constant position updates of similar accuracy. Using ArUco markers as absolute positioning sensors leads to similar positioning errors when starting from a misaligned position, but a larger relative drop in accuracy since the ArUco marker position updates are more sparse than the constant rate of UWB position updates.

B. Adding cost to the equation

As previously mentioned when considering sensor configurations that include a height sensor, the average 3D localization error is reduced to 10 cm or lower, with many configurations reaching comparable results. However, the cost of the used sensors can differ vastly and depending on the use case this cost needs to be scaled to the entire drone fleet in operation.

Fig. 5 shows the distribution of 3D positioning versus cost. Note the logarithmic y-axis scale and the omission of the SLAM-only configuration. When no absolute positioning is required, but simple flight stability and/or relative positioning is sufficient a single SLAM camera would suffice as sensor at €200. As long as the environment is sufficiently distinguishable for the SLAM algorithm and the flight is short enough so that positioning drift has little effect. However, the operational requirements increase by requiring a well-defined take-off position and there are no safety guarantees since a single sensor failure would crash the drone. Therefore, this configuration is not immediately applicable to any of our predefined use cases since the SLAM camera would not fit the small/nano drone class that comes with lessened reliability requirements and is not further taken into account in this cost accuracy trade-off.

When considering the trade-offs for the absolute positioning configurations the Pareto frontier defines the most interesting setups and leads to the following insights:

- UWB-only positioning has the lowest cost but also the highest positioning error, making this configuration only applicable to the nano drone class. However, even for the smallest drones a reliable height estimation is preferred, so adding a sonar or solid state 1D LiDAR is strongly suggested. This leads to the largest improvement in error for the lowest additional cost. The UWB + sonar configuration is furthermore applicable to all use cases under consideration with an average 3D error of 10.71 cm and €735 cost. However, a single sensor failure can

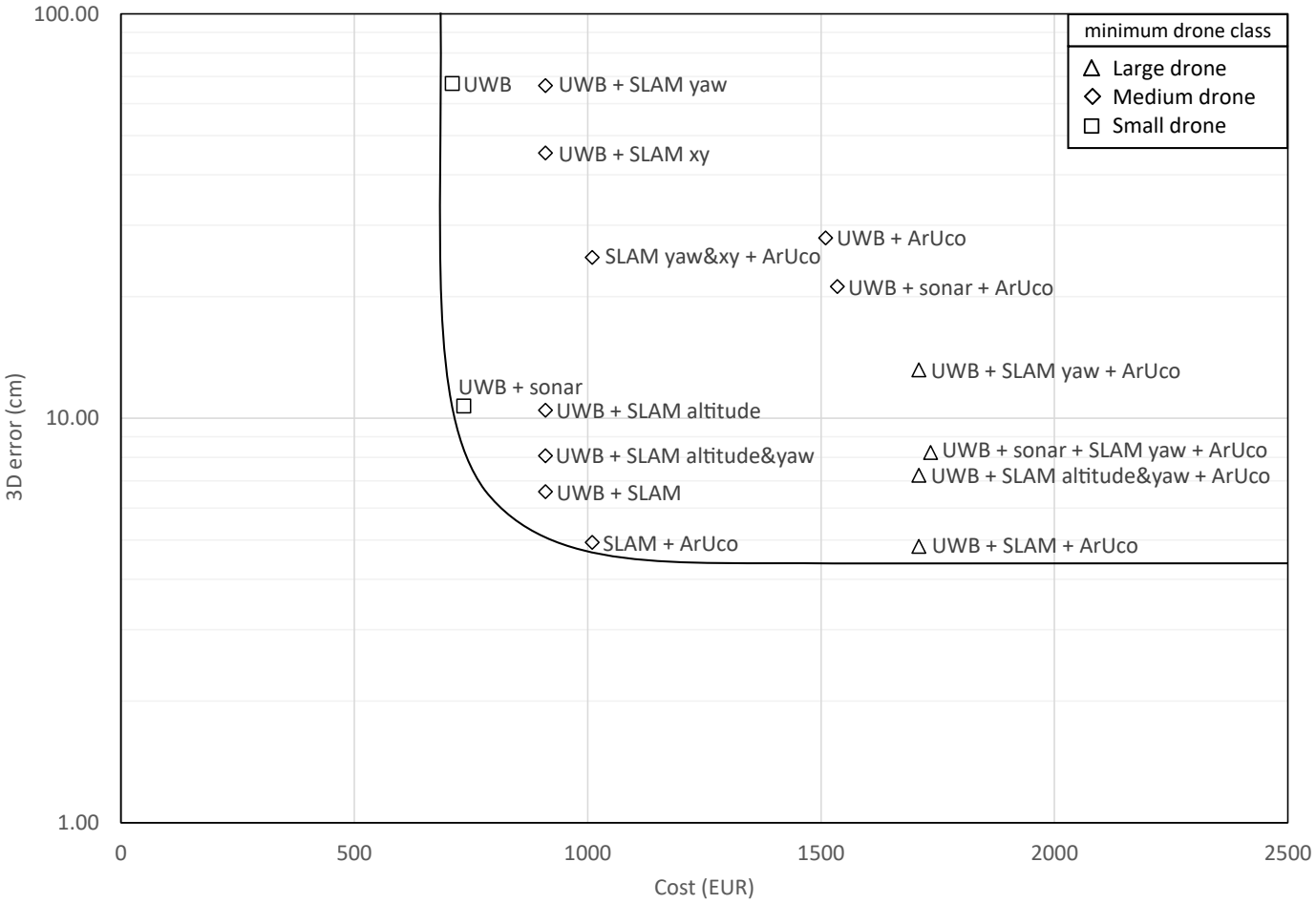


Fig. 5. Sensor set cost versus positioning error trade-off per drone class

still lead to a drone crash and the yaw estimation of the drone is not reliable, which can be an issue for certain use cases.

- Adding a SLAM camera to UWB localization mitigates these concerns, allowing for the SLAM camera or the UWB localization to fail without compromising the stability of the drone, as well as providing a reliable yaw estimation. This is sufficient for video or RFID based asset identification. This configuration reaches a 3D error of 6.58 cm at €910. It is also clear that using a subset of the SLAM camera sensors is not useful in a cost/accuracy trade-off, since a full sensor set from the SLAM camera reliably leads to a lower error at same cost. This configuration is only suitable for medium sized drones and upwards due to the computational and power requirements.
- Dropping the UWB positioning and adding cameras for ArUco markers instead, together with a SLAM camera further reduces the error up to 4.93 cm at a slightly higher cost of €1010. However, the weight is increased as well due to the cameras required for stable ArUco positioning, reducing flight time, as well as an increase in personnel cost to install sufficient ArUco markers to keep the SLAM camera from drifting. Furthermore, this relatively expensive configuration is susceptible to SLAM failures and will quickly crash due to IMU drift and incorrect ArUco position estimations without the SLAM yaw sensor.
- When combining UWB, the SLAM camera and ArUco markers we get the most stable and accurate sensor configuration (4.82 cm avg. 3D error), but also the most expensive at €1735 (including 6 UWB anchors). Due to the power and weight requirements this configuration is most appropriate for large drone platforms that also benefit the most from the additional stability and reliability of this configuration.

When scaling these costs by multiple drones, the UWB anchor cost is not scaled, since the same anchor nodes are used by each drone. This reduces the per drone cost to one UWB tag of €100.

C. UWB Anchor selection

In this section the effect of scaling the amount of UWB anchors per drone is investigated, both on error and cost. The previous results considered a full UWB anchor set of 6 anchors for maximal positioning stability and reliability. However, depending on the use case a smaller set of anchors can be used to reduce costs or comply with wiring and technical constraints of the environment. For this reason a full parameter space exploration of UWB anchor selection was performed, leading to 63 experiments (from 1 UWB anchor to all 6) per considered configuration (UWB + ArUco, UWB + sonar, UWB + sonar + ArUco, UWB + SLAM, UWB + SLAM + ArUco).

The results from these experiments are presented in Fig. 6. Note the logarithmic y-axis scale and the gaps in the graph representing failing configurations.

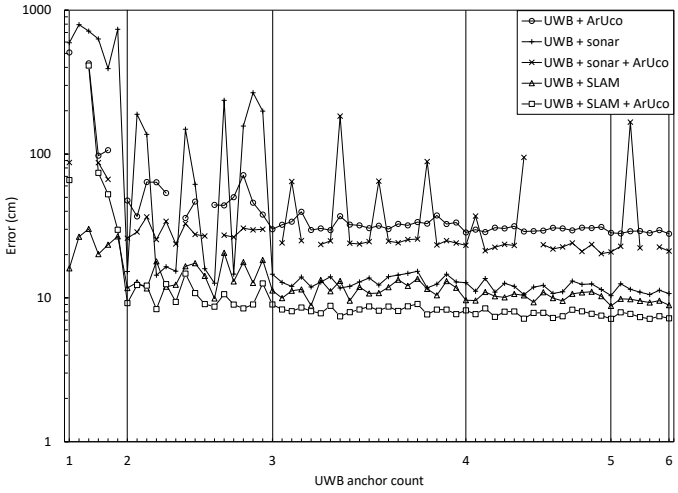
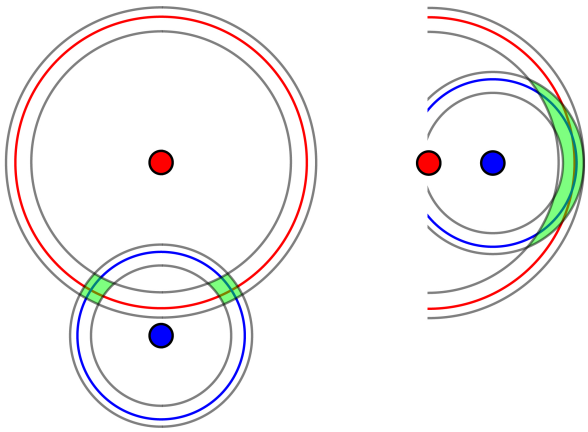


Fig. 6. average 3D error in cm for each anchor selection set

For a UWB + sonar configuration there are no risks of diverging sensor fusion behavior since localization errors do not compound over time. However, a single UWB anchor setup leads to positioning errors upwards of 400 cm, making these setups only applicable for very large free space locations. Interestingly, some 2-anchor setups have a similar error as 3+ anchor setups. typically, 3 anchors are considered a minimum amount for 3D localization, even with an additional height sensor such as a sonar. Furthermore, the 2-anchor selections performing the best are those with the highest Dilution of Precision (DoP). This high DoP leads to one possible larger intersection with low precision, while anchors with a low DoP can lead to 2 intersections with high precision. This leads to increased nonlinear behavior in the EKF, reducing its performance and overall sensor fusion accuracy. This effect is illustrated in Fig. 7.



(a) low DoP: high precision, 2 solutions (b) high DoP: low precision, 1 solution

Fig. 7. Horizontal Dilution of Precision when using only two UWB anchors

From 3 UWB anchors and upwards, the positional stability can be considered reliable with an average error below 15 cm.

In Table VIII these results are further detailed, with the maximum error reduction of an additional anchor as the most appropriate metric, reducing the error by 2 cm on average with each additional anchor from 3 anchors upwards. When considering the best case anchor selection, the accuracy of a 3 anchor setup can rival that of 6 anchors. This confirms the importance of intelligent anchor selection when the environment or cost restriction do not allow for an overdimensioned amount of UWB anchors. When considering equipment failure and possible ranging errors it is appropriate to consider a 4 anchor minimum for ideal drone stability, especially in the higher weight drone classes upwards of 2 kg when no additional absolute positioning source such as ArUco markers is available.

TABLE VIII
3D ERROR IN CM PER AMOUNT OF UWB ANCHORS

UWB anchors	3D error		
	average	minimum	maximum
1	644 cm	392 cm	794 cm
2	100 cm	12.6 cm	266 cm
3	13.3 cm	11.7 cm	15.3 cm
4	11.9 cm	10.5 cm	13.6 cm
5	11.2 cm	10.4 cm	12.5 cm
6	10.7 cm	10.7 cm	10.7 cm

Adding ArUco markers as an additional sensor leads to interesting results. The minimum positioning error is doubled from 10 cm to 20 cm, but a lower error than a UWB only setup is often achieved for 1 and 2 anchor setups with a few exceptions that lead to possible sensor fusion deviation due to an inaccurate yaw angle. Detecting this runaway behavior in a separate routine and falling back on deviation-proof sensors such as UWB and sonar would be an appropriate technique in these cases. The results of ArUco runaway detection on the 3D error of 1-anchor and 2-anchor UWB sensor fusion are detailed in Fig. 8 with the cross hairs representing the runaway detection fallback. This technique would allow even nano drones to operate with high enough positional stability in cases of very sparse UWB anchor deployments if a sufficiently capable camera is included on the drone platform.

For medium sized drones and upwards the most reliable solution is pairing UWB localization with ArUco markers and a SLAM camera or other accurate yaw source. The reliability of each possible anchor subset is drastically improved with stable localization results from 2 anchors upward and a 3D error of 15 cm to 7 cm. When only 1 UWB anchor is available the accuracy improves notably in all cases except the worst case anchor selection, where a runaway position was observed due to divergence of the ArUco marker detection algorithm. However, when looking at the results of the UWB + SLAM configuration, the 1-anchor stability was greatly improved with no runaway positions and an average 3D error of 23.85 cm possibly qualifying for certain use cases with enough flight clearance. When 2 or more UWB anchors are available ArUco markers can safely be added to the UWB + SLAM configuration without chance of instability and an error reduction of 26% on average.

Taking these UWB anchor selection results into account

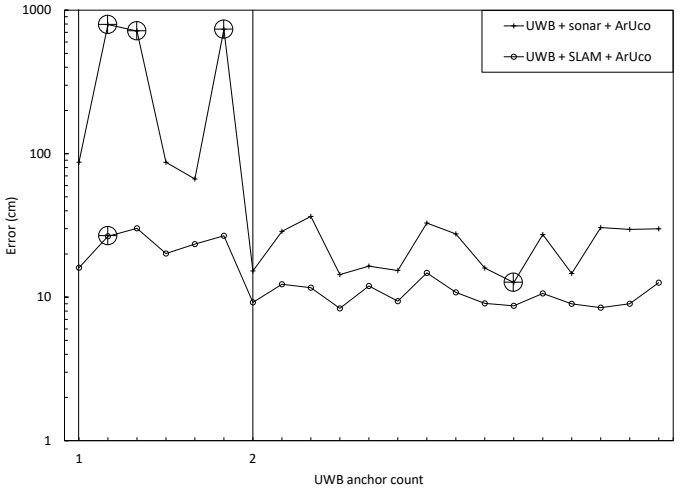


Fig. 8. 1-2 UWB anchor selection 3D error with ArUco runaway detection

gives the following revised cost optimizations for our three drone classes:

1) *Large drone* use cases could reliably cut infrastructure costs when using a full sensor set of UWB, ArUco markers and a SLAM camera or other reliable yaw and altitude source. Using only 2 UWB anchors in Line of Sight of the flight path leads to a 66% reduction of the UWB anchor deployment cost while retaining a positioning error of 15 cm or better. In the example given an average 3D error of 10.5 cm at € 1310 is achieved for 2 anchors versus 7.2 cm at € 1710 for 6 anchors.
2) *Medium sized drones* could benefit from a SLAM camera without ArUco markers to reduce the camera cost and weight of the drone platforms. This leads to a remarkably stable flight configuration, reducing the importance of intelligent anchor selection and delivering an average 3D error of 23.85 cm if only one UWB anchor is available. This can be sufficient for traversing open spaces or wide corridors at a sensor cost of € 410. When 2 UWB anchors are available this average 3D position is further improved to 14.59 cm on average, with a maximum error of 20.56 cm and best of 9.9 cm. This shows the gains of intelligent anchor selection, but this configuration also performs adequately without for use cases where a 20 cm worst case positioning error is acceptable.
3) *Small or nano drones* can rarely rely on a SLAM camera, but can reduce the UWB infrastructure cost in half by requiring only 3 UWB anchors in Line of Sight compared to 6 without intelligent anchor selection, while still retaining an average 3D error of 13.3 cm for a UWB and sonar sensor configuration of € 435. For even sparser UWB anchor deployments of 2 anchors intelligent anchor selection is required leading to an average 3D error of 14.9 cm at € 335 sensor cost in case anchors with a high Dilution of Precision are reliably selected.

VI. Practical Application Scenario

Apart from the IIoT testbed, the different sensor configurations were tested in a portion of a large warehouse of 65 000 m^2. In total hundreds of flights, totaling to more than 10 flight hours, have been performed. Due to the large scale of the environment, the main cost driver is the cost per m^2 depicted in Table VI. From this table, it is clear that the SLAM sensor reduces the cost per m^2 dramatically. The combination with ArUco markers results in the lowest possible cost, especially when considering that the cameras used for detecting the markers may also be used for the inventory scanning process. The accuracy and stability of the combination of SLAM with ArUco markers proved to be sufficient for autonomous flight. For flying in between racks, the proposed solution for large scale warehouses is thus to make use of the SLAM camera and ArUco markers, with a marker placed every 3.2 m on each rack level. For larger, open spaces, the combination of SLAM with UWB is proposed, as in these large areas, fewer UWB anchors are required as there are no obstructions, while ArUco markers will not be detected from a larger distance. This guarantees there's always one absolute positioning method (UWB or ArUco) available to complement the relative positioning of the SLAM camera.

The Proof of Concept deployment showed that a fully autonomous inventory scanning system was possible. Taking into consideration the previous small and large scale experiments led to a full sensor set of UWB + SLAM + ArUco, since the drone in question has a Maximum Take Off Weight (MTOW) of 5.2 kg and falls squarely in the Large drone category. Using our cost model this leads to a sensor cost of € 1280 per drone, with the addition of a 1D LiDAR [43] instead of a sonar to reduce measurement noise.

The infrastructure cost was minimized by using both SLAM and ArUco for positioning between the warehouse racks at € 0.09 per m^2, leading to a total infrastructure cost of € 5850 for the deployment of the ArUco markers. In case a complementary absolute positioning technology would be required between the racks, deploying one UWB anchor every 100m of rack space would increase the infrastructure cost by € 12 350 to a total of € 18 200 to cover the entire 65 000 m^2 warehouse with both ArUco and UWB positioning.

This Practical application scenario was presented as part of the imec.icon InWareDrones project with a demonstration video [44] and validated additional research material such as UWB self-calibration [45] and RFID localization of scanned tags [46].

VII. Conclusion

In this paper a full parameter space exploration was performed of 337 possible sensor configurations of the flexible sensor fusion platform OASE using an IMU, sonar, SLAM camera, ArUco markers and UWB localization with up to 6 UWB anchors. Estimated costs of the sensors under evaluation were also taken into account to show the trade-offs in cost versus accuracy when considering sensor configurations with similar performance and reliability. Additionally a cost model to estimate the infrastructure cost of placing ArUco markers and UWB anchors was developed. This model showed the impact of sensor selection on the infrastructure cost as well, with a significant reduction of ArUco and UWB deployment costs when a SLAM camera was added to the drone.

These performance and cost considerations were applied to three drone classes to differentiate and define a minimally viable and optimal sensor selection set as follows.

- *Large drone* platforms preferably have one sensor redundancy for every possible Degree of Freedom. The suggested sensor configuration contains an IMU, SLAM camera, ArUco markers and UWB tag with 3 anchors along the flight path in constant Line of Sight. This configuration has an average 3D error of 8.3 cm with an estimated sensor cost of € 1410. Additional sonar sensors can be added to the platform for obstacle avoidance, but show no advantage in positional stability when combined with the SLAM camera and ArUco markers.
- *Medium drone* use cases often present a more complex trade-off in the sensor cost and engineering budgets. A minimally viable sensor set was defined as a 1 anchor UWB sensor together with a SLAM camera for accurate height, yaw and x,y positioning redundancy at a cost of € 410. But depending on the reliability requirements up to four UWB anchors might be advised in case a more robust localization is required leading to a total cost of € 710. Additionally, if the use case in question also involves video asset recognition ArUco markers can be added instead of additional UWB anchors to improve reliability at an increased drone cost of € 1210.
- *Small or nano drones* should be able to operate with only UWB localization and a single height sensor such as a sonar or solid state 1D LiDAR for even more weight reduction. Together with intelligent anchor selection, positional error of 15 cm was achieved using 2 anchors with high DoP and a limited sensor set cost of € 335. Adding one additional UWB anchor lowers the onboard processing requirements of intelligent anchor selection and guarantees a stable position with 3DoF of 13.3 cm on average for € 435.

These results show that depending on the use case and engineering budgets there is always a good trade-off to be found between reliability, accuracy and cost for the considered sensor configurations. By incorporating intelligent anchor selection the reliability of limited UWB anchor deployments can be increased drastically, which is a promising future research track. An additional improvement can be found in using the camera feed of the SLAM camera for object and asset recognition with an additional on-board video processing unit, which can also be used for ArUco marker positioning.

References

[1] S. Winkvist, E. Rushforth, and K. Young, "Towards an autonomous indoor aerial inspection vehicle," *Industrial Robot: An International Journal*, vol. 40, no. 3, pp. 196–206, 2013.

[2] B. McCabe, H. Hamledari, A. Shahi, P. Zangeneh, and E. R. Azar, "Roles, benefits, and challenges of using uavs for indoor smart construction applications," in *Computing in Civil Engineering 2017*, 2017, pp. 349–357.

[3] M. Saska, T. Krajnik, and L. Pfeucil, "Cooperative uav-ugv autonomous indoor surveillance," in *International Multi-Conference on Systems, Signals Devices*, 2012, pp. 1–6.

[4] B. Wang, Y. Sun, D. Liu, H. M. Nguyen, and T. Q. Duong, "Social-aware uav-assisted mobile crowd sensing in stochastic and dynamic environments for disaster relief networks," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 1, pp. 1070–1074, 2020.

[5] N. Macoir, J. Bauwens, B. Jooris, B. Van Herbruggen, J. Rossey, J. Hoebeke, and E. De Poorter, "Uwb localization with battery-powered wireless backbone for drone-based inventory management," *Sensors*, vol. 19, no. 3, p. 467, 2019.

[6] M. Avila, M. Funk, and N. Henze, "Dronenavigator: Using drones for navigating visually impaired persons," in *Proceedings of the 17th International ACM SIGACCESS Conference on Computers & Accessibility*, ser. ASSETS '15. New York, NY, USA: Association for Computing Machinery, 2015, pp. 327–328.

[7] C. Luo, S. I. McClean, G. Parr, L. Teacy, and R. De Nardi, "Uav position estimation and collision avoidance using the extended kalman filter," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 6, pp. 2749–2762, 2013.

[8] European Aviation Safety Agency. Easa open category - civil drones. Accessed: 2021-09-09. [Online]. Available: https://www.easa.europa.eu/domains/civil-drones-rpas/open-category-civil-drones

[9] J. A. Paredes, F. J. Álvarez, T. Aguilera, and J. M. Villadangos, "3d indoor positioning of uavs with spread spectrum ultrasound and time-of-flight cameras," *Sensors*, vol. 18, no. 1, p. 89, 2018.

[10] M. Khalaf-Allah, "Particle filtering for three-dimensional tdoa-based positioning using four anchor nodes," *Sensors*, vol. 20, no. 16, p. 4516, 2020.

[11] B. Nenchoo and S. Tantrairatn, "Real-time 3d uav pose estimation by visualization," in *Multidisciplinary Digital Publishing Institute Proceedings*, vol. 39, no. 1, 2020, p. 18.

[12] 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)*. IEEE, 2018, pp. 1–6.

[13] R. I. Marasigan, Y. D. Austria, J. B. Enriquez, L. L. Lacatan, and R. M. Dellosa, "Unmanned aerial vehicle indoor navigation using wi-fi trilateration," in *2020 11th IEEE Control and System Graduate Research Colloquium (ICSGRC)*. IEEE, 2020, pp. 346–351.

[14] E. Martín-Gorostiza, M. A. García-Garrido, D. Pizarro, D. Salido-Monzú, and P. Torres, "An indoor positioning approach based on fusion of cameras and infrared sensors," *Sensors*, vol. 19, no. 11, p. 2519, 2019.

[15] 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.

[16] 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)*. IEEE, 2017, pp. 3495–3502.

[17] D. Magree and E. N. Johnson, "Combined laser and vision-aided inertial navigation for an indoor unmanned aerial vehicle," in *2014 American Control Conference*. IEEE, 2014, pp. 1900–1905.

[18] D. Li, Q. Li, N. Cheng, Q. Wu, J. Song, and L. Tang, "Combined rgbd-inertial based state estimation for mav in gps-denied indoor environments," in *2013 9th Asian control conference (ASCC)*. IEEE, 2013, pp. 1–8.

[19] Qualisys AB. Qualisys motion capture systems. Accessed: 2020-08-03. [Online]. Available: http://www.qualisys.se/

[20] G. Mao, S. Drake, and B. D. Anderson, "Design of an extended kalman filter for uav localization," in *2007 Information, Decision and Control*. IEEE, 2007, pp. 224–229.

[21] F. Romero-Ramirez, R. Muñoz-Salinas, and R. Medina-Carnicer, "Speeded up detection of squared fiducial markers," *Image and Vision Computing*, vol. 76, pp. 38–47, 2018.

[22] S. Garrido-Jurado, R. Muñoz-Salinas, F. Madrid-Cuevas, and R. Medina-Carnicer, "Generation of fiducial marker dictionaries using mixed integer linear programming," *Pattern Recognition*, vol. 51, pp. 481–491, 2015.

[23] G. Balamurugan, J. Valarmathi, and V. P. S. Naidu, "Survey on uav navigation in gps denied environments," in *2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES)*, 2016, pp. 198–204.

[24] M. D. C. Pérez Rubio, D. Gualda Gómez, J. D. Vicente Ranera, J. M. Villadangos Carrizo, J. Ureña Ureña *et al.*, "Review of uav positioning in indoor environments and new proposal based on us measurements," in *IPIN*, 2019, pp. 267–274.

[25] IDLab Ghent University. Industrial iot lab. Accessed: 2020-08-03. [Online]. Available: https://www.ugent.be/ea/idlab/en/research/research-infrastructure/industrial-iot-lab.htm

[26] DJI. Dji matrice 100 quadcopter for developers. Accessed: 2020-08-03. [Online]. Available: https://www.dji.com/be/matrice100

[27] Stanford Artificial Intelligence Laboratory et al. Robotic operating system, kinetic kame. Accessed 2020-11-30. [Online]. Available: http://wiki.ros.org/kinetic

[28] Binghao Li, B. Harvey, and T. Gallagher, "Using barometers to determine the height for indoor positioning," in *International Conference on Indoor Positioning and Indoor Navigation*, 2013, pp. 1–7.

[29] G. Dedes and A. G. Dempster, "Indoor gps positioning-challenges and opportunities," in *VTC-2005-Fall. 2005 IEEE 62nd Vehicular Technology Conference, 2005.*, vol. 1, 2005, pp. 412–415.

[30] M. F. Misnan, N. H. M. Arshad, R. L. A. Shauri, N. A. Razak, N. M. Thamrin, and S. F. Mahmud, "Dynamic height measurement using sonar sensor based on single dimension kalman filter in embedded controller," in *2012 IEEE International Conference on Control System, Computing and Engineering*, 2012, pp. 596–601.

[31] N. Macoir, M. Ridolfi, J. Bauwens, B. Jooris, B. Van Herbruggen, J. Rossey, J. Hoebeke, and E. De Poorter, "Low power, portable and infrastructure light indoor uwb ranging solution," in *Proceedings of the 18th International Conference on Information Processing in Sensor Networks*, 2019, pp. 337–339.

[32] FLIR Systems. Flir blackfly s usb3. Accessed: 2020-08-03. [Online]. Available: https://www.flir.com/products/blackfly-s-usb3?model=BFS-U3-50S5C-C

[33] Qualisys AB. (2020) Qualisys track manager.(qtm). [Online]. Available: https://www.qualisys.com/software/qualisys-track-manager

[34] Drone data set 1. Accessed: 2021-03-17. [Online]. Available: https://gitlab.ilabt.imec.be/jvanhie/drone-dataset-1

[35] S. Kram, M. Stahlke, T. Feigl, J. Seitz, and J. Thielecke, "Uwb channel impulse responses for positioning in complex environments: A detailed feature analysis," *Sensors*, vol. 19, no. 24, p. 5547, 2019.

[36] J. A. Andersson, J. Gillis, G. Horn, J. B. Rawlings, and M. Diehl, "Casadi: a software framework for nonlinear optimization and optimal control," *Mathematical Programming Computation*, vol. 11, no. 1, pp. 1–36, 2019.

[37] G. Raja, S. Suresh, S. Anbalagan, A. Ganapathisubramaniyan, and N. Kumar, "Pfin: An efficient particle filter-based indoor navigation framework for uavs," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 5, pp. 4984–4992, 2021.

[38] W. You, F. Li, L. Liao, and M. Huang, "Data fusion of uwb and imu based on unscented kalman filter for indoor localization of quadrotor uav," *IEEE Access*, vol. 8, pp. 64 971–64 981, 2020.

[39] A. Wenz and T. A. Johansen, "Moving horizon estimation of air data parameters for uavs," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 3, pp. 2101–2121, 2020.

[40] G. D. Padfield, *Helicopter flight dynamics*. Wiley Online Library, 2008.

[41] Y. Xu, Y. S. Shmaliy, C. K. Ahn, G. Tian, and X. Chen, "Robust and accurate uwb-based indoor robot localisation using integrated ekf/efir filtering," *IET radar, sonar & navigation*, vol. 12, no. 7, pp. 750–756, 2018.

[42] P. Sturm and S. Ramalingam, *Camera models and fundamental concepts used in geometric computer vision*. Now Publishers Inc, 2011.

[43] LeddarTech. Leddarone single element sensor module. Accessed: 2021-09-09. [Online]. Available: https://leddartech.com/solutions/leddarone/

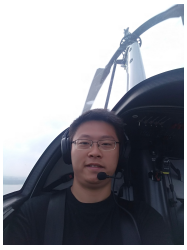
[44] IDLabResearch. (2020) Autonomous indoor drone for inventory management. [Streaming Video] https://youtu.be/3oLRkI7JtIQ. Accessed 2021-09-09.

[45] 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, no. 4, pp. 3007–3023, 2021.

[46] C. Li, E. Tanghe, D. Plets, P. Suanet, J. Hoebeke, E. De Poorter, and W. Joseph, "Reloc: Hybrid rssi- and phase-based relative uhf-rfid tag localization with cots devices," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 10, pp. 8613–8627, 2020.

Jono Vanhie-Van Gerwen received his M. Sc. degree in Computer Science Engineering from Ghent University, Belgium, in July 2009. As of August 2009, he started working as research engineer affiliated with the Department of Information Technology of Ghent University at which he is now a senior research engineer. His research activities have included performance analysis and benchmarking for Wireless Sensor Networks, management and optimization of wireless testbeds, together with enabling them for experimentally driven education. In 2018 he obtained a professional drone pilot license and specializes in drone experimentation and automated operations.

dr. Ir. Kurt Geebelen (male) received the M.Sc. degree in mechanical engineering in 2010 and a PhD in mechanical engineering in 2015, both from KU Leuven, Belgium. His main research focus is on autonomous systems (mobile systems, robots, drones) operating in industrial environments. He has an extensive experience in computer vision, sensor fusion, motion planning, control and decision making for these types of systems.



dr. Ir Jia Wan received the M.Sc degree in aerospace engineering from Beihang University, China, and finished his PhD in aerospace engineering at TU Delft, the Netherlands. His research mainly focuses on unmanned aircraft with extensive experience in dynamic system modeling, flight mechanics, sensor fusion, and system design.

Wout Joseph (M'05) was born in Ostend, Belgium on October 21, 1977. He received the M. Sc. degree in electrical engineering from Ghent University (Belgium) in July 2000. He obtained the Ph. D. degree in March 2005; this work dealt with measuring and modelling of electromagnetic fields around base stations for mobile communications related to the health effects of the exposure to electromagnetic radiation. He was from 2007-2012 a Post-Doctoral Fellow of the FWO-V (Research Foundation – Flanders). Since October 2009 he is professor in the domain of "Experimental Characterization of wireless communication systems." He is IMEC PI since 2017. His professional interests are electromagnetic field exposure assessment, propagation for wireless communication systems, antennas and calibration. Furthermore, he specializes in wireless performance analysis and Quality of Experience.

Jeroen Hoebeke received the Masters degree in Engineering Computer Science from Ghent University in 2002. In 2007, he obtained a PhD in Engineering Computer Science from Ghent University with his research on adaptive ad hoc routing and Virtual Private Ad Hoc Networks. Current, he is an associate professor in the Internet Technology and Data Science Lab of Ghent University and imec. He is conducting and coordinating research on wireless (IoT) connectivity, embedded communication stacks, deterministic wireless communication and wireless network management. He is author or co-author of more than 200 publications in international journals or conference proceedings.

Prof. dr. ir. Eli De Poorter is professor at the IDLab research group from Ghent University and imec (https://idlab.technology). He received his master degree in Computer Science Engineering from Ghent University, Belgium, in 2006, his Ph.D. degree in 2011 at the Department of Information Technology at Ghent University and he became professor at IDLab in 2015. Since 2017, he is also affiliated with the IMEC research institute. His team performs research on wireless communication technologies such as (indoor) localization solutions, wireless IoT solutions and machine learning for wireless systems. He performs both fundamental and applied research. For his fundamental research he is currently the coordinator of several research projects and has over 200 publications in international journals or in the proceedings of international conferences. For his applied research, he collaborates with (Flemish) industry partners to transfer research results to industrial applications as well as to solve challenging industrial research problems. Prof. De Poorter is also co-founder of the lopos spin-off company (https://lopos.be) which offers privacy-aware UWB wearables for safety & social distancing.