

INSANE: Cross-Domain UAV Data Sets with Increased Number of Sensors for developing Advanced and Novel Estimators

Christian Brommer¹, Alessandro Fornasier¹, Martin Scheiber¹, Jeff Delaune², Roland Brockers², Jan Steinbrener¹ and Stephan Weiss¹

Abstract—For real-world applications, autonomous mobile robotic platforms must be capable of navigating safely in a multitude of different and dynamic environments with accurate and robust localization being a key prerequisite. To support further research in this domain, we present the INSANE data sets - a collection of versatile Micro Aerial Vehicle (MAV) data sets for cross-environment localization. The data sets provide various scenarios with multiple stages of difficulty for localization methods. These scenarios range from trajectories in the controlled environment of an indoor motion capture facility, to experiments where the vehicle performs an outdoor maneuver and transitions into a building, requiring changes of sensor modalities, up to purely outdoor flight maneuvers in a challenging Mars analog environment to simulate scenarios which current and future Mars helicopters would need to perform. The presented work aims to provide data that reflects real-world scenarios and sensor effects. The extensive sensor suite includes various sensor categories, including multiple Inertial Measurement Units (IMUs) and cameras. Sensor data is made available as raw measurements and each data set provides highly accurate ground truth, including the outdoor experiments where a dual Real-Time Kinematic (RTK) Global Navigation Satellite System (GNSS) setup provides sub-degree and centimeter accuracy (1-sigma). The sensor suite also includes a dedicated high-rate IMU to capture all the vibration dynamics of the vehicle during flight to support research on novel machine learning-based sensor signal enhancement methods for improved localization. The data sets and post-processing tools are available at: <https://sst.aau.at/cns/datasets>

Index Terms—Data set, Long-term Autonomy, State-Estimation, Sensor-Fusion, Autonomous Flight, Mobile robotics, Field Robotics, Planetary Robotics, Exploration, Computer Vision, Cameras, UWB, Laser, GNSS, Benchmarks, Stereo

I. INTRODUCTION

Real-World data sets are an essential part of the research and development process in the field of robotics. When developing new methods, one of the first steps is to test and prove an approach with flawless simulated data sequences, followed by more advanced verification in which the simulated data needs to reflect real-world sensor behaviors such as noise, non-Gaussian signal distributions, and environment-based signal degradation. Modeling realistic sensor signals and their degradation linked to the environment is a difficult, if

not impossible, task. At this stage, real-world data sets with accurate ground truth are a prerequisite to move from the ivory tower to real-world applications.

Existing Unmanned Aerial Vehicle (UAV) data sets focus on isolated topics such as Visual Inertial Odometry (VIO), indoor navigation, or vehicle control with aspects to energy efficiency as presented by [1]. Data sets focusing on outdoor UAV applications are sparse, and the provided ground truth for subsequent algorithm development and evaluation is not of sufficient quality. In addition, research towards multi-environment UAV operations is progressively increasing. One example is the transition of UAVs from outdoor environments to indoor locations and vice versa.

Such operations cause changes to the sensor availability, such as GNSS sensors, which become unavailable in particular phases. It also requires changes in navigation reference frames if the available sensors provide relative navigation, e.g., VIO. Other side effects that can occur during a transition from outdoor to indoor concern changing lighting conditions or changes in the magnetic field due to building structures. Autonomous tasks requiring such scenarios are package delivery applications, automated emergency response, vessel inspection (e.g., the BugWright2 project¹), and long-term environmental surveying, e.g., agricultural applications. Corresponding research for transitioning robotic vehicles includes [2] and long-duration UAV autonomy with possible indoor recharging [3].

Unfortunately, openly available multi-environment data, which is necessary to elevate this field of research and to move from simulated scenarios into the real world, especially for UAVs, do not exist. Simulated environments include the work introduced by [4] and [5] as well as data sets with real sensor data but artificially augmented vision presented by [6]. Yet, using simulated data for this aspect is only suitable for initial development stages as it does not account for realistic environmental effects introduced to a sensor, such as near infrastructure affecting GNSS or Ultra-Wide-Band (UWB) signals during an indoor-outdoor transition.

The presented INSANE data set aims to overcome this limitation and evolve the field of cross-environment robotics by providing a variety of cross-domain and multi-environment flight data sets with highly accurate ground truth for position and orientation (6 DoF) for indoor and outdoor setups. A flight platform with a versatile sensor suite is introduced to gain

¹Control of Networked Systems Group of the University of Klagenfurt, Austria. E-Mail: { [christian.brommer](mailto:christian.brommer@uni-klu.ac.at), [alessandro.fornasier](mailto:alessandro.fornasier@uni-klu.ac.at), [martin.scheiber](mailto:martin.scheiber@uni-klu.ac.at), [jan.steinbrener](mailto:jan.steinbrener@uni-klu.ac.at), [stephan.weiss](mailto:stephan.weiss@uni-klu.ac.at) }@ieee.org

²Jet Propulsion Laboratory, California Institute of Technology. E-Mail: roland.brockers@jpl.nasa.gov, jeff.h.delaune@jpl.nasa.gov

¹<https://www.bugwright2.eu/project/>

sufficient knowledge about the environment. This platform hosts 18 individual sensors, ranging from high-resolution navigation images over high-rate and multi-IMU signals to multi-GNSS and UWB data.

For the generation of ground truth data, special attention and effort was given to the acquisition of high-quality raw measurements provided by the sensors that are required to generate this ground truth. The ground truth provided for the presented data set is given as an absolute entity and is not expressed relative to existing localization algorithms, as done by distinct related work. Thus, comparing localization algorithms to the ground truth provided by this work allows for a definite evaluation of errors without restrictions to specific metrics.

The same platform was used to record 27 data sets with accumulated trajectories of more than 2km while operating in four distinct environmental domains. The presented UAV data set package is tailored to promote the development and robustness of new localization algorithms. It provides the necessary data to validate individual algorithm setups in a controlled environment and gradually increases the difficulty for successive proof of algorithms and methods.

The main features are:

- 6 DoF absolute ground truth with centimeter and sub-degree accuracy 1-sigma for outdoor data sets.
- Indoor trajectories with motion capture ground truth (6 DoF millimeter and sub-degree accuracy) for the initial proof of algorithms.
- Outdoor to indoor transition trajectories with continuous absolute ground truth.
- Trajectories in a Mars analog desert environment for Mars-Helicopter analog setups, including various ground structures, cliff flight over, and cliff-wall traversing trajectories for mapping.
- Vehicle and sensor integrity, including intrinsic information such as static IMU data and RPM correlated vibration data.
- Real-world sensor effects and degradation posed by individual scenarios.
- Initialization sequences for VIO algorithms.
- Inter-sensor calibrations in pre-calculated form and raw calibration data sequences for custom calibration routines.

II. RELATED WORK

This section provides an overview of UAV research data sets and how the presented work is positioned within this ecosystem. The majority of open-source data sets in the robotics community focuses on isolated research aspects and most of them are tailored towards ground vehicles and aspects of autonomous driving. This includes large-scale outdoor data sets such as KITTI [7] (1392x512 images @10Hz and IMU @10Hz) and the Oxford data set [8] (1280x960 images @16Hz, GNSS and INS solution @50Hz and no raw IMU data), with the later addition of sporadically sparse RTK GNSS for ground truth [9]. However, given the time at which the data sets were published, the provided data rates and image resolutions are lower compared to the sensor setup of the presented work. Another ground vehicle data set,

targeted specifically for the SLAM and odometry community, is presented by [10]. The data set focuses on an agricultural environment and provides stereo imagery (672x376 @20Hz), IMU (140Hz), and odometry data. The data set uses an RTK GNSS for positional ground truth but does not provide ground truth for the global orientation.

Several indoor data sets are making use of local ground truth in the form of high-quality SLAM in post-processing. This includes the TUM-LSI large-scale indoor data set [11] and [12] which adds a Leica station with fiducial markers at dedicated locations. Another data set, for a large-scale shopping mall scenery, introduced by Naverlabs [13] makes use of Structure-from-Motion (SfM) in post-processing for the generation of ground truth data. The presented work does provide data to perform the same approach if desired. However, providing post-processed SLAM / SfM data is not within the scope of this publication.

The next category of data sets concerns the indoor to outdoor transition aspect. Such data sets are mainly performed by handheld sensor suites. [14] introduced the PennCOSYVIO data set with a sensor setup that extends a Google Tango platform with additional camera modalities (2x752x480 images @20Hz and IMU @200Hz). The data set features indoor and outdoor locations, but it does not make use of a GNSS sensor. The ground truth for this data set is solely generated by utilizing pre-calibrated fiducial markers placed throughout the experiment area. [15] later introduced the TUM data set, also using a handheld sensor setup for indoor to outdoor trajectories without GNSS information (1024x1024 images @20Hz and IMU @200Hz). For this data set, ground truth information is only provided for the indoor segment using a motion capture system.

Finally, data sets for the development and proof of localization algorithms for UAVs also exist, but the focus does not lie on sensor degradation and sensor switching nor cross-domain operation. The EuROC data set introduced by [16] focuses on various difficulty levels of VIO for indoor scenarios (2x768x480 images @20Hz and IMU @200Hz), with ground truth generated by using a motion capture system and a Leica station. The UZH-FPV drone racing data set [17] features dynamic flights for isolated indoor and outdoor trajectories (640x480 images @30Hz and IMU @500/1000Hz respectively). This data set uses a motion capturing system for the ground truth of the indoor trajectories, and SLAM for the ground truth of the outdoor trajectories. Another interesting approach is the Blackbird data set [6], which performs real-world flights and collects IMU (100Hz) as well as motion capture information (360Hz) to generate a multitude of photo-realistic vision streams, each for the same set of recorded trajectories. UAV data sets for possible future real-world applications also exist; [1] generated a data set for the analysis of energy consumption for a package delivery drone. However, the data set does not provide image streams, and the accuracy of position ground truth is only rated with $\pm 2m$ @10Hz.

Another data set, mainly concerning the Mars analog contribution of the presented work, is the MADMAX data set [18]. This data set provides a sensor suite similar

to the Mars-Rover but is performed in a handheld approach. However, this data set only provides lower rate IMU measurements (100Hz) and lower rate image streams (1032x772 @15Hz and 2064x1544 @4Hz) compared to the presented INSANE data set. This is adequate given that a rover platform does not perform agile trajectories. The ground truth for this data set is generated by fusing measurements from two RTK GNSS units (1Hz) with IMU measurements, resulting in a 100Hz filtered ground truth data stream.

To the best of our knowledge, no UAV data set concerning indoor-outdoor transitioning with aspects to real-world sensor degradation and an extensive sensor suite (see Tab. I) with corresponding redundancies, various frame-rates and resolutions, and continuous global ground truth for outdoor, indoor, and transition areas, such as presented by this work, exists at this point.

III. SYSTEM SETUP

A. Vehicle Configuration

All data sets have been recorded with the flight platform shown by Figure 2. The base of this platform is a commercially available carbon frame equipped with a minimal power unit, rotors, and a PixHawk4 autopilot. The base frame was heavily extended and altered from its original. The final platform setup weighs 3kg and carries a number of additional sensors (see Tab. I).

The small size of the aerial vehicle constrains the amount of additional payload. However, the UAV setup needs to be able to process the sensor data during closed-loop experiments and needs to be able to record the data of the sensor suite in raw format without loss of information. Because of this, the vehicle is equipped with two Raspberry Pi4 companion boards. This allows for computational load balancing, interface bandwidth distribution associated with a specific sensor, and distributed sensor data storage. These three aspects of the vehicle system are shown as a block diagram in Figure 1.

It might be of interest to the reader that the vehicle was running an in-house developed flight stack [19] which is generalized and deployable across many standard computation platforms by utilizing a robust and versatile OS [20]. This flight stack will be published and open-sourced by September 2022 and referenced in this paper for its camera-ready version.

The vehicle includes additional Electromagnetic Interference (EMI) shielding to allow optimal functioning of RF sensitive components such as the GNSS (further detailed in Section III-B5) despite high-frequency data lines. Additional dust shielding and individual cooling appliances were added to the flight platform for optimal operation in the hot and sandy environment posed during the Mars analog data recording sessions.

B. Sensors

This section outlines the setup, calibration procedures and individual aspects concerning the sensor suite. The sensors are summarized in Table I.

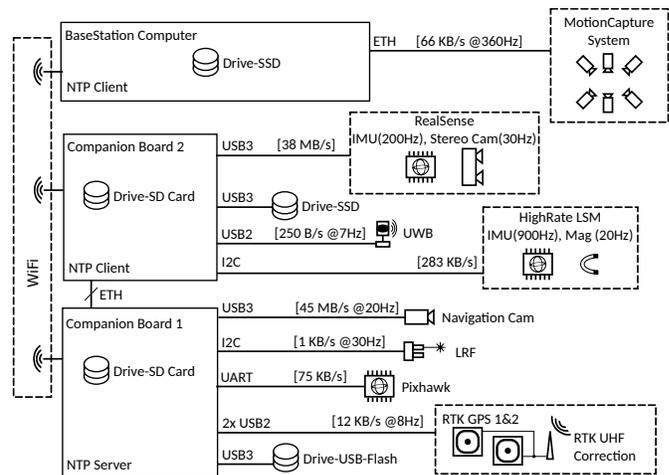


Fig. 1. This diagram shows the embedded companion boards and sensor integration as well as storage and interfaces distribution according to individual sensors data rates and the available bandwidth of the interfaces from the embedded hardware.

1) *IMUs*: In general, the sensor suite includes four Micro-Electromechanical Systems (MEMS) IMUs. Two 200Hz IMUs are part of the Holybro Pixhawk v5 autopilot, namely the BMI055 from Bosch and ICM20689 from TDK. One 200Hz BMI055 IMU is part of the RealSense T256 stereo camera, and one dedicated high-rate 900Hz LSM9DS1 IMU from ST-Electronics is added by itself. The two IMUs of the autopilot can not be distinguished and are actively switched based on a sensor voting scheme by default. To enable a clear association between the IMU calibration and its bias, the BMI055 IMU was deactivated. Another reason is that the RealSense T256 already provides measurements of a BMI055 type IMU. Thus, all active IMUs provided by the data set are from individual manufacturers, providing a good range of different IMU characteristics. In addition, it may be noted that the IMUs of the autopilot are hardware dampened, shown by Figure 3 and further detailed in the vibration analysis Section III-C and IV-D, respectively.

The IMU sensor positioning is shown by Figure 2. The autopilot IMU, which is referred to as the main IMU, is positioned close to the vehicle's center, which should avoid amplified vibrations within the linear acceleration measurements.

The RealSense IMU is positioned forward-facing and tilted down for ideal stereo camera positioning. This IMU has the highest lever arm, does receive the most amplification, and possible resonances in terms of vibration. Specific scenarios also showed undersampling effects. The measurement stream of this IMU can be seen as a challenging scenario for possible research.

The high-rate IMU is rigidly attached (not dampened) and provides measurements at more than 900Hz to support IMU filter applications or machine learning approaches, as shown by [21] for active vibration analysis and noise reduction. The intrinsic calibration of the specific IMUs is done by performing the Allan variance method outlined by [22]. IMU recordings with a length of five hours and the corresponding tools are

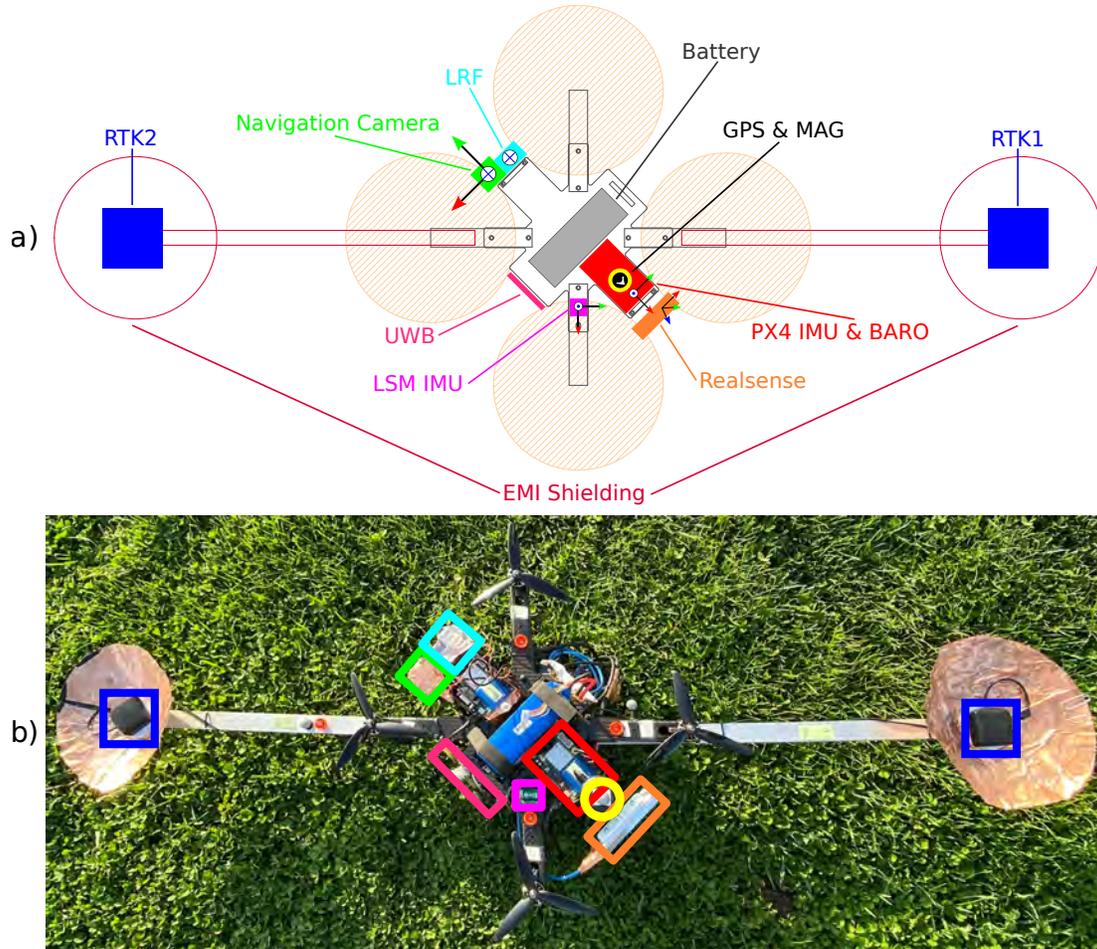


Fig. 2. This illustration shows the vehicle design with coordinate frames for the placement and orientation of the sensors on the experiment platform in **a)**, and the real vehicle with sensor position overlays in **b)**. Accurate calibration of each sensor (extrinsics and intrinsics where applicable) is stored within each data set. The method on how each sensor is calibrated and the intent on the sensor placement is described in Section III-B

open-sourced with the data set.

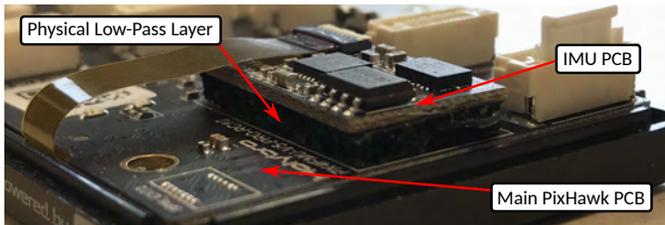


Fig. 3. The sensor suite has three IMUs where each serves a different aspect. This figure shows the assembly of the primary IMU sensor within the PixHawk autopilot, on top of a damping layer that acts as a low-pass filter between the IMU and main PCB board. This hardware filter is beneficial for closed-loop applications. However, in Section IV-D we are carrying out a vibration analysis of the vehicle. This analysis made use of the rigidly attached high-rate IMU. A brief comparison of the dampened IMU and the high-rate IMU with applied filters, designed based on the vibration analysis, is described, and the interested reader is referred to [21] for further use on the vibration data.

2) *Magnetometer*: The helicopter platform hosts two independent magnetometers from different manufacturers for variation in sensor characteristics and redundancy. The first module is included in the PixHawk4 sensor suite (80Hz), and the second module is located within the external LSM9DS1

System on Chip (SoC) (20Hz). For magnetometers, several aspects have to be considered: The intrinsic and extrinsic calibration of the sensor for its specific location on the experiment platform, the magnetic variation which depends on the geolocation, and local magnetic disturbances posed by the environment. The data set includes a default magnetometer calibration for this vehicle and dedicated magnetometer calibration data sets, which allows for different calibration methods by the user.

It is important to note that the magnetic variation at the locations in which the data set was recorded are not negligible. This variation depends on the geolocation and concerns the magnetic inclination (spherical elevation) and the declination (spherical azimuth), which build the local direction and strength of the magnetic field. The pertinent information is available online². Location dependent variables and descriptions for the data sets can be found in Section III-D. Tables II, III and IV specifically address the magnetic variation.

The provided data segments can be used to calculate the intrinsic and extrinsic calibration of a magnetometer with respect

²<https://www.ngdc.noaa.gov/geomag/calculators/magcalc.shtml>

TABLE I

THE TABLE DESCRIBES THE INDIVIDUAL COMPONENTS OF THE SENSOR SUITE FOR THE EXPERIMENT PLATFORM. IN PARTICULAR, AMONG OTHERS, IT FEATURES THREE IMUs, THREE CAMERAS AND THREE GNSS SENSORS AS WELL AS TWO MAGNETOMETER.

Sensor	Type	Rate [Hz]	Description
High Rate IMU	LSM9DS1		
IMU		900	Rigidly attached
Magnetometer		20	
Pixhawk			
IMU	ICM20689	200	Internally dampened
IMU	BMI055	200	(Disabled)
GNSS		5	
Magnetometer	UST8310	80	
Barometer	MS5611	20	
Motorspeeds		100	
RealSense T256			
IMU	BMI055	200	
6DoF Odometry	V-SLAM	200	
Cameras			
Nav Camera	IDS UI-3270LE-M-GL	20	CMOS Mono, 2056x1542, global-shutter; Lens BM4018S118C, FoV(D=126°, H=101°), 3MP, aperture 1.8
Stereo Cam	RealSense T256	30	848x800, global-shutter, 64mm baseline, 163°FoV
External Sensors			
2× RTK GNSS	UBLOX C94-M8P	8	Coordinates and velocity
Laser Range Finder (LRF)	Garmin Lidar Lite v3	30	40m range, 1cm resolution
3× UWB	Decawave TREK1000	7	With additional vehicle marker
125× Fiducial Marker	ArUCO		Rate is the same as Nav Cam
Motion Capture	Optitrack	300	37 camera dronehall setup (TBD accuracy)
Pulse Tachometer	Wachendorff PT99		RPM ground-truth

to an IMU sensor. The intrinsic calibration of a magnetometer concerns hard-/ and soft-iron effects. Both effects are mainly dependent to the mounting location of the sensor. Hard-iron effects cause a fixed offset in the magnetometer readings, while soft-iron effects affect the magnetic field strength and its direction locally. As shown by Figure 4, the red sphere shows raw magnetic measurements, and the blue sphere shows measurements that are compensated for their hard-/ and soft-iron effects. The red sphere is not only shifted off-center due to hard-iron effects but also distorted in its spherical shape due to the mentioned soft-iron effects.

The calibration data was recorded in a non-occluded environment, and the vehicle was rotated around various rotational axes such that the magnetic vector, if projected onto a sphere, covers this sphere sufficiently. At first, an ellipsoid fit for the raw data was performed and the center of mass for this ellipsoid was found. This provides the general offset of the ellipsoid $\mathbf{b}_{ct} \in \mathbb{R}^{3 \times 3}$. Afterwards a transformation was found which transforms the ellipsoid into a sphere $\mathbf{t}_{sp} \in \mathbb{R}^{3 \times 3}$. Knowing both parameters, individual magnetometer measurements can be corrected by applying $\mathbf{m}_{corr} = \mathbf{t}_{sp}(\mathbf{m}_{raw} - \mathbf{b}_{ct})$. Detailed information on this method can be found in [23].

After finding the intrinsic calibration, the extrinsic calibration between IMU and magnetometer as well as the local magnetic inclination can be found. This is done by applying the method outlined in [24]. Individual data sequences were recorded which provide static rotations of the vehicle in all six rotational directions with respect to the gravity vector. This data is processed with the cost function described in

[24] to find the most accurate rotation between the gravity and magnetic vectors, including the magnetic inclination. The data set is published with a default calibration for the vehicle setup and with a set of tools to recompute the calibration with potentially different parameters.

The data sequences for the transition experiments show an additional, real-world effect that we aimed to capture with our data set. Because the indoor area and surrounding elements have metal structures, the magnetic field also changes depending on the location within the environment. Figure 18 shows how the magnetic field changes when entering the indoor area. One possible approach on interacting with changes in the magnetic field is outlined in [25], which describes a method of detecting magnetic field changes with subsequent adaptation to the changing magnetic field without measurement rejection.

The magnetometer measurements for all flight sequences are not compensated nor calibrated. This provides the possibility to include online self-calibration of magnetometer intrinsics, if desired.

3) *Cameras*: The sensor suite includes a RealSense T256 stereo camera (848x800 @30Hz) and an IDS UI-3270LE-M-GL global-shutter 3MP navigation camera (2056x1542 @20Hz). The navigation camera is facing downward and is aligned with an LRF for associated pixel range information (see Sec. III-B4). The stereo camera faces in flight direction and is tilted by 60° towards the ground. The reason for this is threefold. First, the stereo camera observes the horizon, and horizon detection can be used to improve attitude estimation. Second, the forward-facing image stream,

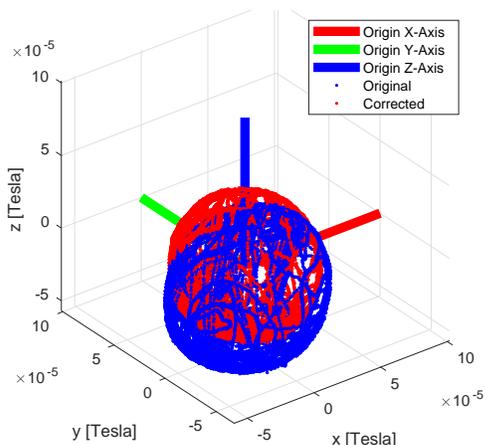


Fig. 4. The image illustrates the offset for the magnetometer measurements from the PX4 platform, as an example on the magnetometer intrinsic calibration for hard- and soft-iron effects. The calibration is provided for all magnetometers.

especially for the desert data sets, can be used to observe and map crater walls. A similar task might be included in future Mars explorations. And third, starting at the height of 2 meters, the stereo camera and the navigation camera have an overlapping field of view which allows for additional feature matching, as shown by Figure 5.

Camera exposure time and gain were chosen such that motion blur is minimized. For the navigation camera, the exposure time was limited to $< 5\text{ms}$, and the gain was automatically adapted within a range that provides balanced brightness throughout an individual trajectory. This especially concerns the transition data sets, which have three lighting conditions: Outdoor with natural light and radiant sun, the transition area with shadow casting and reflecting snow patches, and the indoor area with artificial light providing yet another light intensity. The pure outdoor data sets, including the desert segments, make use of the same settings to reduce gain where flat patches of reflecting areas might raise the brightness of the image over a suitable threshold. The camera settings for the stereo camera are left to default values which provide suitable imagery throughout all scenarios.

The timestamps of the navigation camera are provided by an internal clock of the camera, which is synchronized to the host system on startup. Thus, the timestamps reflect the time at which the image was taken by the camera module.

The data set further provides dedicated recordings for the calibration of both cameras. This includes measurement streams of all IMUs and an image stream for the dedicated camera observing a checker/fiducial marker board. This data can be used to generate the intrinsic and extrinsic calibration of the cameras with respect to any IMU by using the Kalibr tool [26].

4) *Laser Range Finder*: All data sets feature a 30Hz data stream of a Gamin Lidar Lite v3, which is a laser-based range measurement sensor. This sensor provides distance to ground measurements at a resolution of 1cm with 2.5cm standard deviation and operates at a distance of up to 40m - according to the manufacturer's description. It was noticed

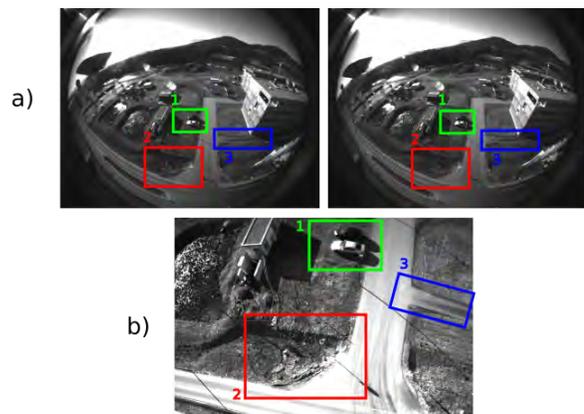


Fig. 5. Example image of the outdoor data set, showing the over-lapping FoV of the forward-down-facing stereo camera (a) and the downward facing navigation camera (b). Labels 1-3 show the same locations in the environment. **Note:** The image of the navigation camera is rotated by -90° and cropped to the region of interest.

throughout the recording of the data sets that this sensor provides sporadic zero measurements beginning at distances of 25-30 meters depending on the properties of the ground surface. The detection and rejection of these measurements is thus straightforward.

This sensor is co-mounted with the high-resolution navigation camera to provide the same setup which is used by the Mars-Helicopter Ingenuity [27]. The LRF is mounted facing down in the same direction as the navigation camera (see Fig. 6) to provide distance to ground measurements for a known pixel cluster within the navigation image.

The data set provides two options for the calibration of the rotation between the camera and the LRF. The calibration routine for the first option provides an image stream with a visible checker/fiducial target board for pose tracking and associated LRF and IMU measurements. The data set contains measurements while the system was excited in all dimensions. A calibration can be performed with the Kalibr tool [26].

The second option is to use the data for the static setup in which the camera and the LRF are facing a wall at a defined distance. With the help of an indicator card, the spot in which the infrared LRF laser hits the wall is made visible and can thus be marked in the image and associated with a pixel location. A calibration that was generated with the second method is provided with the data set.



Fig. 6. This illustration shows the co-mounted camera and laser range finder setup which provides the same sensor setup as for the Mars-Helicopter Ingenuity described by [28].

5) *Real-Time Kinematic (RTK) GNSS*: The sensor suite includes two off-the-shelf UBLOX C94-M8P RTK GNSS

development boards with standard ceramic patch antennas. An RTK GNSS system requires a static base station which communicates corrections to the GNSS modules mounted on the vehicle. This results in highly accurate GNSS measurements with an accuracy of 1 to 16cm standard deviation. For the available data sets, the threshold for the calibration of the RTK base station was set to 0.5m.

The UBLOX C94-M8P module provides a Ultra High Frequency (UHF) module for the communication between the base station and the receivers on the vehicle. This communication channel was replaced by ZigBee modules for an extended range and reliability of the presented setup.

An RTK GNSS module has different modes of operation, which indicate the level of accuracy in position. The three primary modes are non-RTK standard GNSS, RTK float, and RTK fixed mode. Non-RTK represents the standard GNSS information without additional corrections. The RTK-float state signals that no unique solution for the current constellation exists. This state still benefits from RTK corrections and provides measurements with 10 to 16cm accuracy. The RTK-fix is the most accurate state and provides an accuracy of below 1cm. The receiver's status depends on a multitude of factors such as the number of satellites, the time each satellite was tracked, and the SNR posed by the environment.

The majority of the recorded data sets provide measurements with RTK-fix measurements. However, certain conditions can cause sporadic RTK-float states. The GNSS receivers also provide an estimate for the measurement covariance for significantly more detail on the measurement accuracy. This information is also provided with each data set. We used the binary information of fixed and float RTK states for the detection of EMI issues which is outlined in Section III-B6.

As illustrated by Figure 2, the two GNSS antennas are placed on aluminum rods with a distance of 1.2m, centered at the midpoint of the vehicle and rotated by -45° with respect to the front of the vehicle represented by the positive x-axis of the main IMU. The reason for the large baseline between the GNSS antennas is that the GNSS measurements are used for the position and rotational ground truth. Given the setup shown by Figure 7, an accuracy of $\eta = 1\text{cm}$ for one antenna, and the baseline of $B = 1.2\text{m}$, the rotational error in the worst case is $\epsilon = 0.95^\circ$. The method used for ground truth generation is detailed in Section V.

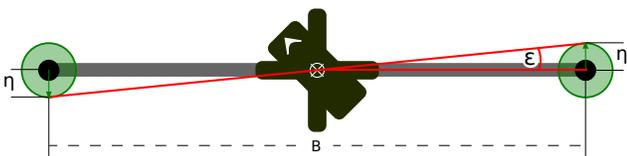


Fig. 7. Rotational dual GNSS error ϵ based on measurement accuracy η and the antenna baseline B .

6) *Electromagnetic Interference*: The topic of EMI in the context of high-frequency communication channels such as Bluetooth and especially USB3 has been discussed and analyzed in the literature [29]. Because the GNSS ceramic patch antennas for this system are standard consumer-grade

products, the antennas do not make use of additional active interference rejection and provide no dedicated shielding against EMI. Initial GNSS quality tests showed impaired signal quality mainly due to the low SNR of individual satellite signals. This is caused by the onboard electronics such as the non-shielded computation boards, sensors, and high-frequency data lines hosted by the platform. Thus the antennas required significant additional shielding (see Fig. 2) to reach an SNR, which allows the operation with high accuracy.

Different shielding designs, made from copper sheets and constructed by hand, as shown by Figure 8 have been tested. In all tests, shielding variation a) showed the most improvement for the SNR. Option c) mainly failed due to possible ground reflections of vehicle-centered EMI sources. Option b) showed good static results, but tests showed that long-term tracking of satellite signals was interrupted during rotational movement. The RTK receiver requires observations of satellite signals over more extended periods of time. Shielding option b), however, causes a barrier of signals to one side of the hemisphere, which changes based on the rotation of the vehicle. Thus, shielding variation a) was ultimately used for all data recordings. In addition to the apparent antenna shielding, the final setup required individual shielding of USB3 connection points and additional USB3 inline low-pass filter³ and shielding, which are partially enclosing the computational platforms and the GNSS receivers.

As mentioned in the previous section, the quality of the GNSS also depends on the environment. The transition data was recorded in a semi-urban environment prone to multipath effects and a higher noise floor. On the contrary, the Mars analog data sets recorded in the Negev desert have a very low noise floor. Such circumstances can be observed in the corresponding covariances of the GNSS measurements. However, we made sure that the data in the data sets always have high-quality GNSS measurements and thus highly accurate ground truth.

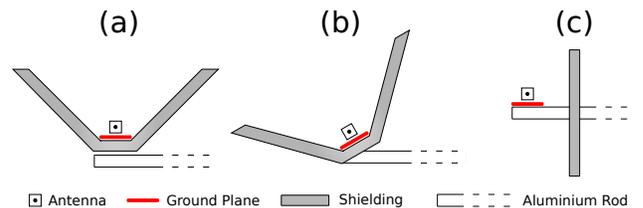


Fig. 8. Tested EMI preventive variations for dedicated RTK Antenna shielding to archive sub-centimeter position accuracy and subsequent orientation accuracy for the ground truth method described in Section V-A. a) upward-opening with 45° side and base shielding, b) similar to (a) but antenna and shielding tilted 40° in the vertical plane, c) fully vertical shielding, no ground encapsulation besides antenna ground plane.

7) *Fiducial Marker*: The transition and indoor data sets provide fiducial marker landmarks. The primary purpose of these markers is to provide ground truth pose information in the transition area. When moving from outdoors to indoors, the transition area poses the challenge in which highly accurate GNSS position measurements degrade in terms of accuracy as the vehicle moves closer towards the building. At the

³Wuert Elektronik 829993STICK

same time, the motion capture system does not detect the vehicle yet. Thus, a field of 108 fiducial markers on eight rigid platforms is used to generate ground truth information for this trajectory segment. Each marker is a uniquely identifiable ArUco marker⁴ generated with the library introduced in [30]. The markers become visible while the GNSS accuracy is still sufficient. During a handover phase, the GNSS and marker pose information is aligned to provide a common ground truth reference frame. After this handover phase, only the marker poses are used to generate ground truth information until the vehicle reaches the inside of the building and the area which is covered by the motion capture system. This represents the second handover phase in which a transition from marker ground truth to the motion capturing system is performed.

The system uses two board types, five narrow field and three wide-field marker boards. The difference between these boards is the marker size. The wide field boards carry a 50cm marker which becomes visible at the height of $< 10\text{m}$. These boards are used for the outdoor area. The narrow field boards feature smaller markers to cover lower-altitude detections during the indoor transition phase.

The fiducial marker field is calibrated based on a dedicated recording with the navigation camera which covers all markers across boards. Marker poses $T_{\text{cam_marker}}$ are expressed with respect to the navigation camera frame. First, all transformations between visible markers in each image $T_{ij} = T_{\text{cam}_i}^T T_{\text{cam}_j}$ with $j < i \in \mathbb{N}$ are extracted. This condition is applied to reduce complexity. Once all images have been analyzed, the individual transformations T_{ij} are filtered for outliers not laying within 1σ standard deviation and the mean is generated.

A main marker is defined, which builds the reference frame for all other markers. All individual calibrations, which may not be observed together with the main marker, are built by forming a graph search with the shortest path regarding the number of needed translations between the main marker and individual markers. Since this can include erroneous translations, the pose of each marker with respect to the main marker is improved by building multiple randomized paths between the main and individual markers, which are then averaged using the geometric median. Finally, each marker is referenced to the main marker. The board which is carrying the main marker is associated to a motion capturing object. Thus, all markers can be expressed with respect to the motion capture system. This static calibration (visualized by Fig. 9) is used to generate poses for the vehicle in the transition scenario case.

The dedicated layout for this location is shown by Figure 13 and the location of the markers after calibration is shown by Figure 9. A second purpose for these boards is the reference of UWB modules located outside the building and mounted to the marker boards. The UWB modules are placed on three boards and known locations for UWB anchor calibration.

Markers can also be used as a visual aid and are present for the indoor recordings for this purpose as well. Prepared camera relative poses are provided with the data sets as well

as the marker calibrations. All tools to export this data from raw data streams are open-sourced with the data set.

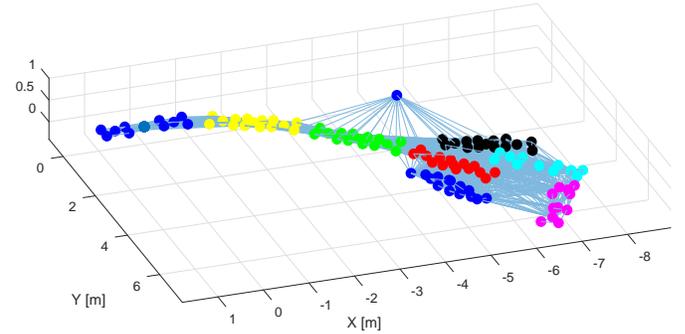


Fig. 9. The graph shows the placement of individual fiducial markers and their respective carrier boards in the outdoor/indoor transition area after the calibration routine. Colored dots represent individual markers detected, a dedicated color is used for each carrier board, and the lines connecting the dots represent observed relative translations between markers during the calibration procedure. The designed layout is shown in Figure 13.

8) *Ultra-Wideband Modules:* The indoor and transition data sets also feature UWB measurements. In case of the transition data set, the UWB modules are placed close to the building entrance and can be used for position estimates in the area in which the GNSS signals are degrading. For this purpose, the vehicle is carrying a main anchor, and three additional UWB anchors are placed on the ground. The UWB positions are shown in Figure 13 and 14. Each UWB measurement is associated with a tag.

The position of the ground UWB anchors is determined by known positions on the fiducial marker boards, which can be related to the indoor motion capturing system (see Sec. III-B7). The range in which the UWB modules provide measurements depends on environmental occlusions. While the vehicle is in the air, first measurements are received from a distance of 30m. A reliable stream of measurements is available at a range of $< 15\text{m}$ between UWB anchors. This includes partial areas of the indoor area.

9) *Computation Module Time Synchronization:* Inter-module time synchronization is essential to ensure coherent data association of isolated modules. As shown in Figure 1, the system makes use of three main modules for the transition experiment case. The two embedded platforms carried by the vehicle, and a dedicated platform for processing the data of the motion capture system. All platforms run an NTP server through Chrony⁵, which is set up to synchronize the system clocks at a higher rate compared to default settings. The NTP time reference is set up to be on module 1 of the vehicle companion boards. Module 2, and the motion capture system run an NTP client, which is synchronizing to the reference time source. Dedicated tests, utilizing the boards General Purpose Input/Outputs (GPIOs) and an oscilloscope, have shown that this setup allows a time synchronization accurate to $100\mu\text{s}$ between all modules.

⁴ArUco Dict. 7x7 250

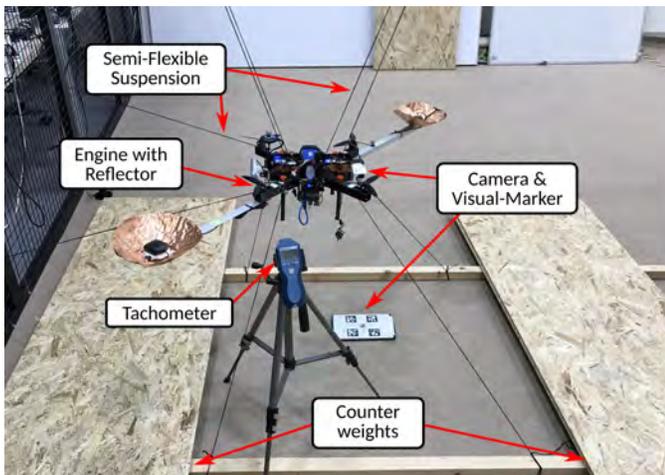


Fig. 10. The image shows the vibration test bench to which the vehicle is semi-rigidly attached. Variations of static RPM and RPM sweeps of the engines provide a baseline for the vibration characteristics of the vehicle for subsequent analysis of the three IMU signals in nominal condition.

C. Vibration Test Bench

The data set also provides additional data for a vehicle-specific vibration analysis. This data is especially interesting for advanced IMU pre-filtering. The high-frequent vibration characteristics of the vehicle showed under sampling effects for the low-rate IMUs (RealSense and PX4 @200Hz). The RealSense is also mounted off-center, which increases this effect. Further investigation also showed that the IMU of the PX4 autopilot was hardware dampened (see Fig. 3). To capture all vibration dynamics of the vehicle without hardware dampening and at a high enough sampling rate, a dedicated rigidly attached high-rate 900Hz IMU was added to the sensor suite. Figure 11 shows a signal comparison between all IMUs. Please note that the Figure shows the norm of the linear acceleration per IMU, the change in magnitude for the LSM9DS1 IMU can be explained by the lever arm. The change in magnitude for the RealSense IMU can be caused by the lever arm as well as hardware dampening and settings from the manufacturer which could not be changed to RAW.

The resulting data set is suitable for an in-depth analysis of conditional sensor signal characteristics, and enables the development of adaptive noise canceling and other IMU pre-filtering techniques for improved trajectory tracking.

Data for the vibration analysis was recorded with the test bench shown in Figure 10. The flight platform was restrained by ten ropes, eight ropes to restrain the vehicle in the vertical axis, and two ropes to further restrain rotations in the horizontal plane. The goal was to provide a semi-rigid constraint that prevents large movement of the platform but allows the nominal vibration characteristics of the vehicle that it exhibits during unconstrained, real-world flight.

The recorded data for this scenario includes the three IMU data streams, magnetometer, pressure, downward-facing LRF, and nominal motor speeds. The PX4 firmware was adapted to stream the nominal RPM values reported by the internal

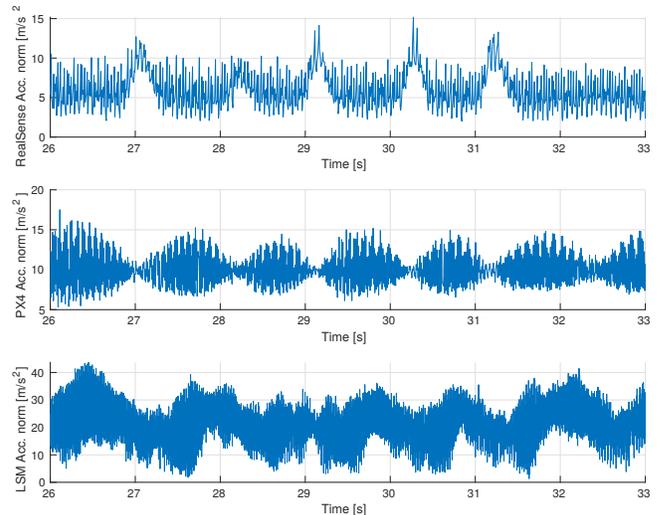


Fig. 11. This plot shows a comparison for the norm of the linear acceleration signal from the BMI055 RealSense IMU (top), the ICM20689 PX4 IMU (middle), and the LSM9DS1 IMU (bottom). The vehicle was suspended in the test bench shown by Figure 10, and the motor speeds were set to 15500 RPM. The RealSense IMU was mounted off-center and thus with a lever arm, and the PX4 IMU was mounted in the center of the vehicle. Both IMUs have a measurement rate of 200Hz and show under sampling effects with corresponding correlations. The LSM9DS1 IMU was also mounted off-center but has a measurement rate of 900Hz which results in a more accurate sampling of the vibration signals.

mixer module. A tachometer, which provides the direct RPM of a single-engine, is used to scale the nominal motor speeds reported by the autopilot. In addition, the video streams of both cameras are recorded as well. The downward-facing navigation camera records a small five-element fiducial marker board, and the forward-tilted stereo camera observes another fiducial marker board. The pose information of the marker throughout the vibration data sequences can be used to refactor the vehicle pose data and exclude low-frequent movement - if desired. It further serves the purpose of verifying the credibility of the camera sensor mounting design.

The vibration data is recorded in nine segments with motor speeds ranging from 10% to 100% of the full nominal motor speeds. The vehicle is static before each sequence, and the motors are at a full stop before and after each segment is recorded. All four engines receive the same thrust signal through software commands which ensure fixed and equal RPMs for all motors during the recorded sequence. A more detailed analysis of the vibration data is presented in section IV-D.

D. Environment

The data set provides experiment data for three different environments. This includes the environmental transition scenario at the campus at the University of Klagenfurt, a solely outdoor setup at a model airfield, and an outdoor Mars analog setup at the Ramon crater in the Negev desert of Israel. All experiments are performed with the same vehicle setup, however, external sensor sources vary depending on the environment.

⁵<https://chrony.tuxfamily.org/>

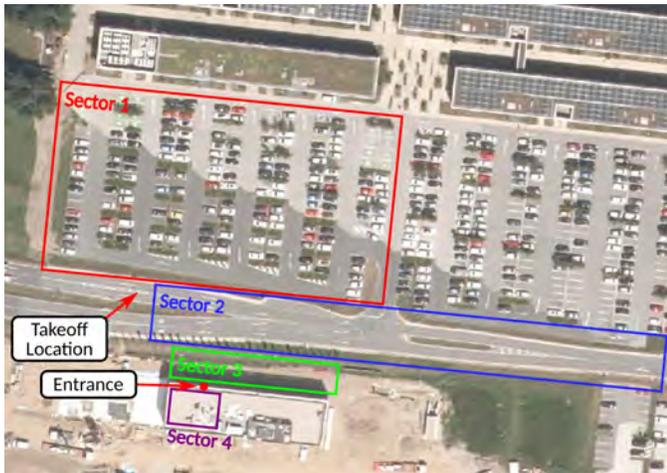


Fig. 12. Location map and flight sectors for the Klagenfurt Dronehall indoor - outdoor transition experiments ⁷

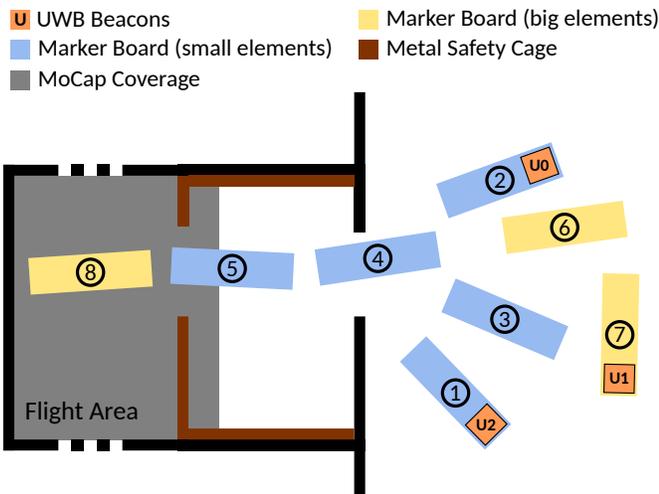


Fig. 13. ArUco fiducial marker field design for the outdoor-indoor transition area at the Dronehall Klagenfurt. Small markers on the edge of each board provide higher accuracy for board-to-board calibration. Indoor markers are small to medium-sized due to the close proximity of the vehicle during the experiment. The outdoor marker boards carry marker that allow reliable detection up to 10m altitude. This Setup ensures sufficient overlay in which high RTK GNSS quality is given, and the navigation camera can detect the marker. The image further shows the motion capture coverage and the placement of the UWB modules.

1) *University of Klagenfurt*: The location at the University of Klagenfurt (environmental info in Tab. II) is used for two types of experiments. First, the previously mentioned indoor data sets in a controlled environment for the initial tests of algorithms with the flight platform. This scenario makes use of the "Dronehall", a motion capture environment with an area of 150m² and 10m in height. This area features 37 cameras and provides millimeter and sub-degree accuracy. Experiments in this environment also provide fiducial markers on the ground, visible by the navigation camera, and three UWB modules placed at different heights. Light conditions for these experiments are constant.

The second type of experiment is the outdoor to indoor



Fig. 14. The entrance of the Dronehall which the aerial vehicle is flying through for the recording of the transition phase data sets. The image illustrates the setup of the marker boards and the distribution of UWB modules.

TABLE II
GNSS COORDINATES AND MAGNETIC VARIATION AT THE UNIVERSITY OF
KLAGENFURT TEST SITE - AUSTRIA

Location:	Klagenfurt am Wörthersee - Austria
Latitude:	46° 36' 48.8124" N
Longitude:	14° 15' 44.5962" W
Magnetic Declination:	+4° 9'
Declination:	Positive (East)
Inclination:	63° 8'
Magnetic field strength:	48300.8nT

transition scenario. These experiments make use of the surrounding outdoor area and the indoor area of the Dronehall as illustrated in Figure 12. Sectors 1 and 2 are the main outdoor areas of the science park. Sector 3 is the transition area in which the GNSS quality starts to degrade based on the altitude of the vehicle. The lower the position of the vehicle, the higher the GNSS signal occlusion due to the building. Figure 15 further illustrates the GNSS signal degradation as the vehicle moves closer to the building at ground level. Sector 4 represents the indoor location of the Dronehall.

The process of gaining and losing the RTK fix of the GNSS modules shows a hysteresis effect. Once the observation of satellites is lost after a longer period, regaining the same accuracy is not instantaneous and requires significant time. Another environmental factor is the influence of the building structure on the magnetic field. Figure 18 shows the change of the magnetic field strength when transitioning from outdoor to indoor. Such real-world effects are highly relevant for real-world applications and the presented work will enable the development and deployment of robust solutions to work with such effects.

This location made use of the fiducial marker setup as described in Section III-B7.

2) *Model Airfield Klagenfurt*: The location at the model airfield in Klagenfurt (environmental info in Tab. III) provides a purely outdoor data set with partially visible fiducial marker and UWB modules. The UWB modules are also mounted onto the marker boards and are calibrated with respect to the GNSS ENU reference frame.

3) *Mars Analog Desert*: The desert location at the Ramon crater in Israel was used for the Mars analog simulation AMADEE20⁸ and was the experiment area for the Mars ana-

⁷Map source: Land Kaernten - KAGIS

⁸OeWF - AMADEE20

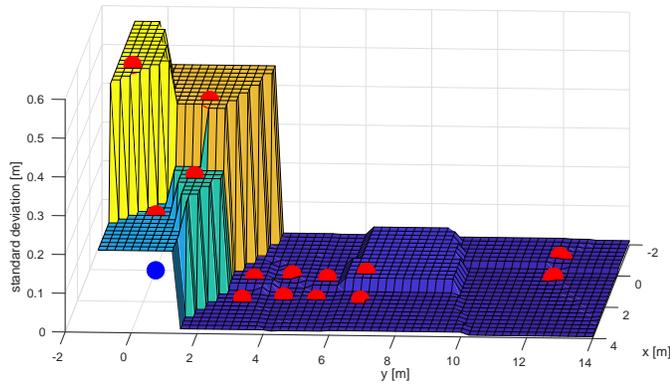


Fig. 15. This map illustrates the RTK GNSS accuracy in close proximity to the Dronehall building. Red dots show the XY position at which the measurements are taken, and the z-axis shows the sample standard deviation of different static measurement sequences with more than 2000 position measurements each. The surface shows the *nearest interpolation* of the data points, and the blue dot indicates the location of the entrance. This shows that the position accuracy towards the building degrades as expected. It also shows that the accuracy is sufficient between four and eight meters, which is the area where the fiducial markers are placed for ground truth during the transition phase.

TABLE III
GNSS COORDINATES AND MAGNETIC VARIATION AT THE KLAGENFURT
MODEL AIR FIELD TEST SITE - AUSTRIA

Location:	Klagenfurt am Wörthersee - Austria
Latitude:	46° 36' 24.7212" N
Longitude:	14° 16' 44.835" W
Magnetic Declination:	+4° 9'
Declination:	Positive (East)
Inclination:	63° 8'
Magnetic field strength:	48300.8nT

log Helicopter experiments. The experiments were performed in a wider area, but information on a general environmental reference point is shown by Table IV.

This environment does not feature external synthetic visual cues or sensors such as UWB. The only external component was the RTK GNSS base station. The focus for this environment was on Mars analog structures, visual elements, terrain topographies, and specific maneuvers. Surface structures vary from flat and sand only, to flat and stone-paved patches (see Fig. 16). Special attention was given to terrain structures with elements of high elevation, such as small canyons, craters, and crater walls.

TABLE IV
GNSS COORDINATES AND MAGNETIC VARIATION AT THE RAMON
CRATER, NEGEV DESERT TEST SITE - ISRAEL

Location:	Negev Desert, Ramon Crater - Israel
Latitude:	30° 35' 59.7438" N
Longitude:	34° 52' 2.3088" W
Magnetic Declination:	-10° 929'
Declination:	Positive (East)
Inclination:	43° 57'
Magnetic field strength:	39808.4nT

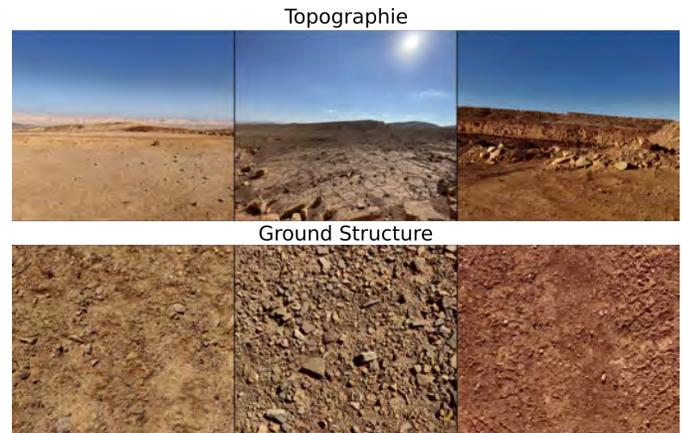


Fig. 16. Externally recorded images as a reference for terrain structures. The top row shows the topography, and the bottom row shows a close-up of the ground structure. **Left:** Sand patches and no terrain, **Middle:** Stone paved medium terrain, **Right:** High terrain elevation and sandy ground structure.

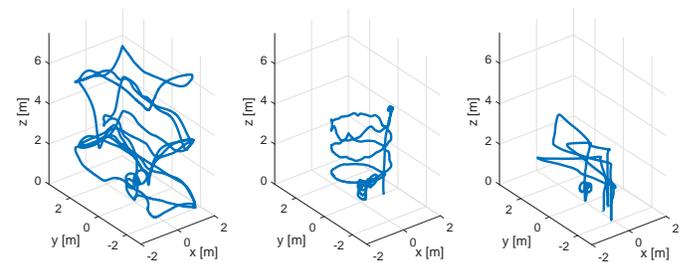


Fig. 17. Indoor motion capture flight patterns. **Left:** Ascending square trajectory, **Middle:** Upwards spiral, **Right:** Translations with pick and place elements.

IV. EXPERIMENT - DATA

This section describes the patterns and scenarios which were performed for each location. In general, the majority of data sets have an initialization phase. This initialization phase consists of a smooth impulse in the direction of the IMU's z-axis and consecutive smooth excitations in 6 DoF. This can be used for VIO algorithms requiring initial IMU excitement and for initial convergence. However, since this feature is not desired in all cases, the copter is placed down to the ground after initialization, followed by a short steady pause before the actual flight maneuver. Thus, the initial phase can be cropped. A list of the majority of recorded flights is shown in Table V.

A. University of Klagenfurt

1) *Motion Capture - Indoor:* The indoor data sets cover three patterns at moderate speed: An upward spiral movement, a square trajectory, and a pick and place scenario in which the vehicle moves to multiple locations and performs a touch-down and take-off maneuver. The three patterns are visualized by Figure 17. The pick and place scenarios have impulses in linear acceleration during the touch-down segments. The pick and place scenario also features UWB measurements for triangulation.

2) *Outdoor to Indoor Transition:* The transition data sets provide a variety of flight patterns. All data sets start at

TABLE V
TABLE OF MAIN DATASETS. THE LOCAL LOCATIONS FOR THE EXPERIMENTS ARE SHOWN IN FIGURES 12 AND 21.

No.	Test-Site	Pattern	Dist.[m]	max. Vel. [m/s^2]	max. Height [m]	Aspects
Dronehall Indoor						
1		Ascending square pattern	60	1	7	
2		Upward spiral	25	0.3	5	
3		Pick and place	42	1	3	with UWB
Outdoor to indoor transition						
1		Const. vel. along sector 2, to sector 4	80	2	20	with UWB
2		Square in sector 1, sector 2, to sector 4	120	2	25	with UWB
3		Square in sector 1, to sector 4	180	2.5	32	with UWB
Model Airfield						
1		Square pattern	260	6	25	with UWB
Mars Analog Desert						
1	Zone 1	Circular pattern	65	2.3	6	
2	Zone 1	Circular pattern	98	3.5	19	
3	Zone 1	Circular pattern	190	5	20	
4	Zone 1	Traverse forth and back	145	4.5	25	
5	Zone 1	Traverse forth and back	27	2	9	
6	Zone 1	Traverse forth and back	48	2	13	
7	Zone 1	Traverse forth and back	48	3	13	
8	Zone 1	Cliff dive	80	4	10	
9	Zone 2	Traverse forth and back	70	4	15	
10	Zone 2	Traverse forth and back	79	4.5	22	
11	Zone 2	Traverse forth and back	158	6	40	
12	Zone 2	Cliff dive	127	5.5	23	
13	Zone 2	Cliff dive	60	2	12	
14	Zone 2	Traverse landing at new location	226	5.3	22	
15	Zone 3	Science close-up	90	6	2.5	
16	Zone 3	Traverse landing at new location	50	4.5	15	
17	Zone 3	Traverse forth and back	162	5	15	
18	Zone 3	Multiple takeoff and landing	5	1.5	3.5	
19	Zone 3	Multiple takeoff and landing	113	1.6	3	
Calibration data						
-		Magnetometer intrinsic and extrinsic		-	-	
-		LRF + IR indicator (static)		-	-	
-		LRF + IMU (moving)		-	-	
-		All IMU 4h static		-	-	
-		Vehicle vibration data		-	-	
-		Navigation camera intrinsic and extrinsic - Klagenfurt experiments		-	-	
-		Navigation camera intrinsic and extrinsic - Mars analog experiments		-	-	
-		Stereo camera intrinsic and extrinsic - Klagenfurt experiments		-	-	
-		Stereo camera intrinsic and extrinsic - Mars analog experiments		-	-	

approximately the same location. Figure 12 shows the flight sectors and Figure 19 shows the trajectory and map overlays.

Data sets were recorded throughout the day, which resulted in varying light conditions between recordings. The lighting conditions for the transition data sets are especially challenging. Sectors 1 and 2 provide natural light and possible reflections from the ground varying based on the height and rotation of the vehicle. Sector 3 has shadow casting because of the nearby building, and sector 4 provides artificial light after transitioning to the indoor location. As described in Section III-B3, the exposure time of the navigation camera was fixed and as low as possible to prevent motion blur. The gain was set to "auto" with a reference value that provides the best result for the changing light conditions. The majority of trajectories perform a square over Sectors 1 and 2 before transitioning to the indoor location, starting at a high altitude of 20m. One dedicated data set also provides a constant velocity segment along sector 2, to expose possible unobservability for VIO algorithms.

A subsection of data sets also shows lens flare effects

for the stereo camera when facing the sun at high altitudes. Another real-world effect is the change of the magnetic field strength when transitioning from indoor to outdoor, shown by Figure 18. Such effects are desired since the presented work focuses on aspects of real-world sensor effects.

B. Model Airport Klagenfurt

The model airport data set focuses on a simplified scenario for visual-inertial odometry and UWB triangulation. The environment mainly provides a planar grass field with a partial agricultural area. Similar to the outdoor-indoor transition setup, the three UWB modules are mounted on fiducial marker boards to calculate a global reference. However, instead of using the motion capture system, the GNSS-based ground truth is used to express the location of the board and the UWB module position consecutively.

C. Mars Analog Desert

The Mars analog flight data focused on two subjects: The acquisition of realistic vision data and the performance of sci-

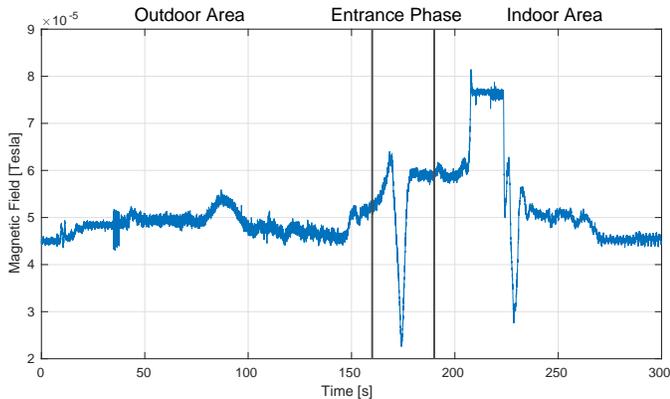


Fig. 18. This graph shows the change of the magnetic field during the transition phase from outdoor to indoor. Such real-world effects can either be treated as outliers or as additional information to adapt to changing magnetic field as described by [25].

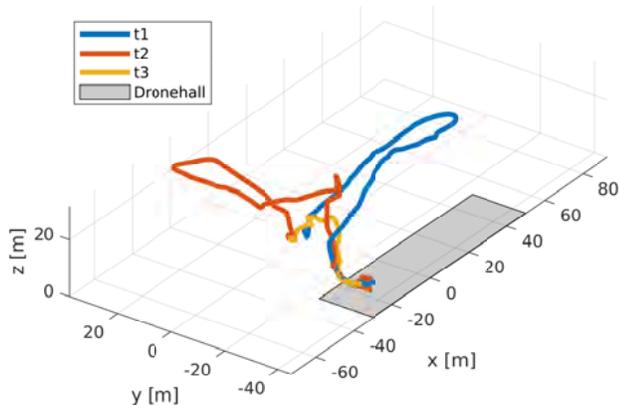


Fig. 19. Trajectory and overlays for the outdoor to indoor transition data sets. The shaded area marks the indoor area.

entific flight tasks. Thus, the recorded data covers challenging localization elements such as higher speed traverse at high altitude, constant velocity segments, and crater wall fly-overs. For the scientific flight tasks, the data set provides science close-up flights with transitions from intermediate altitude towards an object, a semi-circular scan of the object, and raise to the initial altitude. Another scientific element is the scanning of crater walls (see Fig. 20), thus, data sets 9, 13 and 14 provide longer transitions with the 3D camera facing the crater wall. Figure 21 shows the majority of flight patterns performed in the Ramon Crater.



Fig. 20. Representative image of a crater wall in the Negev desert. This location, among others was used to record data for a possible crater wall mapping task. The image was taken with an external device.

D. Vehicle Vibration Analysis

It is well known that parasitic effects in measured inertial data can lead to drastic decreases in performance of classical state estimation methods if not accounted for. Apart from noise and biases of the inertial sensor itself, these also include vibrations of the vehicle that couple into the IMU at various frequencies. One major source of vibrations on UAVs are the motors. While it may be difficult to completely account for vibrations in classical state estimation approaches, it has been hypothesized that latent features in IMU data such as RPM or velocity dependent vibrations are beneficial for learning based methods [31], [21]. In this section, we show that the motors introduce a resonant frequency in the IMU spectrum that is characteristic of their speed. This resonance is particularly pronounced in the inertial data obtained from the high-rate IMU.

During flight, the PX4 provides the rate values for each of the four motors at 100Hz. The PX4 rate values range from 0 to 2000 on a scale with arbitrary units. To calibrate these rate values to actual RPM values, the data obtained with the vibration measurement setup described in Sec. III-C for seven different motor speeds is used to derive an analytic relationship between PX4 rate values and measured RPM. The corresponding data is shown in Tab. VI, and plotted on the left hand side of Fig. 22 together with a quadratic polynomial with coefficients $a_0 = 168.5541$, $a_1 = 12.1870$, and $a_2 = -0.0023$. To analyze the frequencies of the res-

TABLE VI
RPM CALIBRATION

Relative RPM	PX4 Rate Values	Measured RPM
10%	297	3560
20%	486	5605
40%	864	9000
60%	1242	11800
80%	1620	14000
100%	1999	15500

onances for the different motor speeds, we computed the averaged power spectral densities of the high-rate inertial data for each vibration test run using Matlab's `spectrogram` method, and extracted the frequency of the main peak using Matlab's `findpeaks` method. Plotting the true motor RPMs against the peak frequencies shows a linear relationship. The corresponding line fit with coefficients $a_0 = 10.6666$ and $a_1 = 0.0161$ is shown on the right side of Fig. 22 along with the data. Although this data was recorded in a very controlled experimental setup, a clear relationship between the resonance in the spectrogram and the PX4 motor rate values exists also for real-flight data. Figure 23 shows the spectrogram of the high-rate IMU for one of the flights in gray. Overlaid in color are the expected resonances computed from the reported PX4 rate values for each motor using the same linear fit parameters. The inset shows a magnified portion of the whole spectrogram. As can be seen, the expected resonant frequencies correlate very well with the observed resonances. We attribute the remaining differences to experimental effects

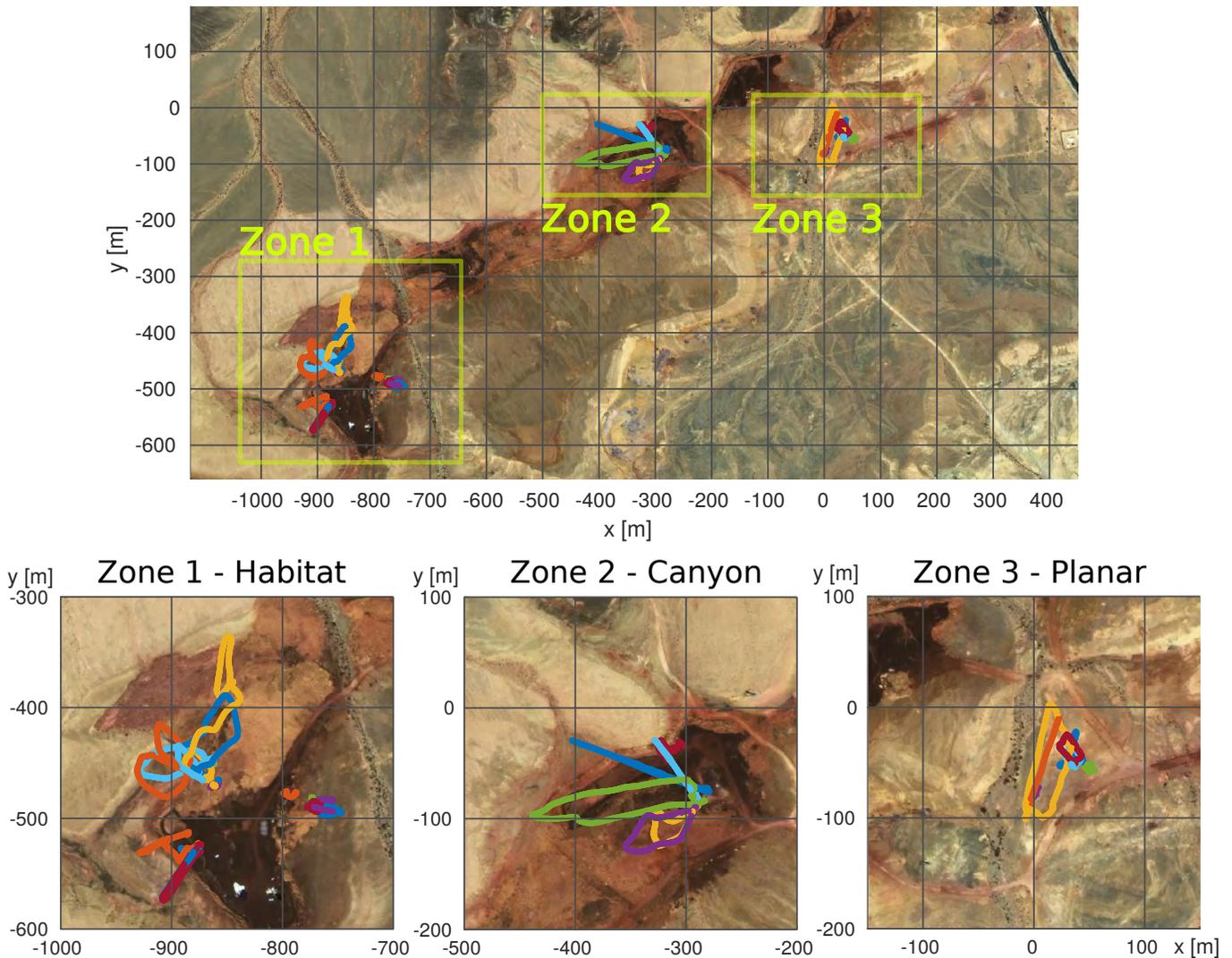


Fig. 21. Trajectory and map overlays for the Mars analog data sets. Zone 1 shows the habitat in a small valley where a part of the cliff mapping data was recorded. Zone 2 shows a steep and lengthy canyon in which cliff mapping data and data with high elevation changes were acquired. Zone 3 was a planar only region that was also used for the science close-up maneuvers. - Map ©OpenStreetMap contributors

in the real-flight scenario. In particular, the actual RPMs may differ from the PX4 rate values. Nonetheless, these results indicate the presence of motor speed dependent resonances in the recorded inertial data that can be of interest in particular to the learning community.

V. GROUND TRUTH

This section describes the generation of 6 DoF ground truth data for the outdoor data sets, sensor time calibration, as well as the alignment of the pose information for the outdoor to indoor transition data sets. The tools which were used to generate the ground truth will be open-sourced together with this data set.

A. Vehicle 6 DoF Pose

The notation of transformations used by this work is as follows. \mathbf{R}_{A_B} is the rotation of frame B expressed in frame A . ${}_A\mathbf{p}_{AB} \equiv \mathbf{p}_{AB}$ is the translation from frames A to B expressed in frame A . As an example ${}_C\mathbf{p}_{AB} = \mathbf{R}_{CA} {}_A\mathbf{p}_{AB}$.

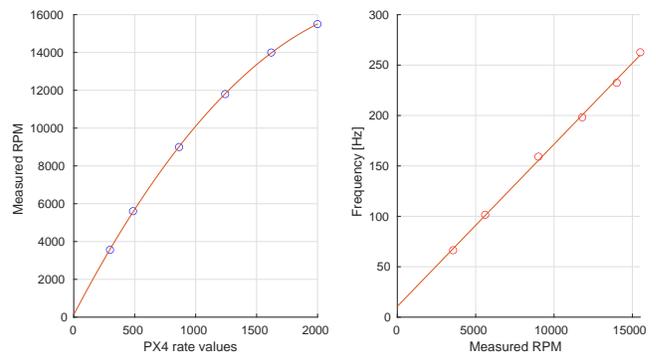


Fig. 22. Calibration of PX4 rate values (left), and resonant frequencies as a function of motor RPM (right).

As mentioned before, we chose to use individual raw measurements for the generation of the ground truth data because of the high measurement accuracy. Using a recursive

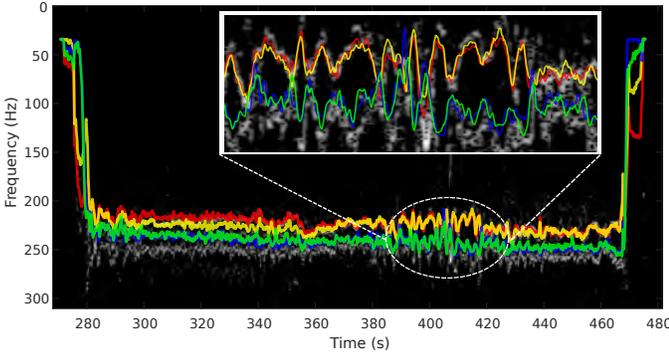


Fig. 23. Prediction of spectral resonances based on PX4 rate values (color) and measured resonances (grey) for real-flight data.

algorithm or a graph-based optimization for ground truth data generation can cause biased results for comparisons against this ground truth data. Original raw sensor data is provided to allow the use of different methods for the generation of ground truth.

We aim to generate 6 DoF ground truth information of an IMU-frame with respect to the GNSS world-frame (ENU). The provided ground truth for the vehicle is referenced to the PX4 IMU. This is done based on two GNSS position measurements and a known calibration of the GNSS antenna positions with respect to the IMU-frame. The method relies on the generation of a virtual GNSS sensor frame which provides 6 DoF. The calculation of the pose for the IMU frame is then achieved by applying transformations of known calibration transformations. Using the two position measurements and generating a virtual directional vector, this setup only provides rotational information in 2 DoF. Thus, we are using the measurements of the directional magnetic vector as support and to fix the ground truth frame for a 3 DoF rotation. As a result, this method will lead to a full-frame definition in 3D with 3 DoF orientation and 3 DoF position. The reference frames and sensor information for the ground truth method are shown in Figure 24. For the remainder of this document, the world-

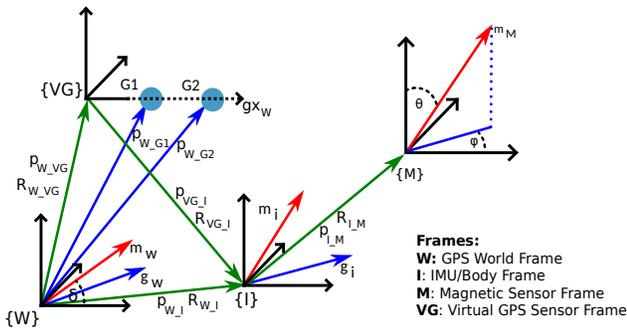


Fig. 24. Ground truth sensor frame setup. RTK GNSS and Magnetometer sensor-frames with respective transformations for the described 6 DoF ground truth calculation.

frame $\{W\}$ is equal to the GNSS world-frame, following the East-North-Up (ENU) convention $ENU \rightarrow XYZ$. The virtual GNSS frame $\{VG\}$, IMU/Body frame $\{I\}$, and the frame of the magnetic-sensor $\{M\}$ are placed on a rigid body.

The GNSS position measurements \mathbf{p}_{W_G1} and \mathbf{p}_{W_G2} are expressed with respect to the world frame.

The magnetic field \mathbf{m}_W is assumed to be locally static with respect to the world-frame and the local magnetic variation, i.e. the inclination and declination, is known. In the IMU frame, the magnetic variation depends on the rotation of the vehicle. This effect needs to be taken into account. Thus, a system that describes the rotation of the vehicle needs to use the following set of information: The magnetic vector expressed in the IMU and the world-frame \mathbf{m}_w and \mathbf{m}_i , respectively, as well as the virtual GNSS vectors in world and IMU frame \mathbf{g}_w and \mathbf{g}_i , respectively. Since only direction is relevant, all vectors are normalized for this analysis. In addition, we can make use of the orthogonal vectors $\mathbf{c}_w = \mathbf{g}_w \times \mathbf{m}_w$ and $\mathbf{c}_i = \mathbf{g}_i \times \mathbf{m}_i$ which renders a possible linear least squares problem full-rank. We present three different methods to analyze the resulting data. The naive approach is to solve the resulting linear least squares problem directly:

$$\mathbf{y} = \mathbf{A} \mathbf{x} \quad (1)$$

$$[\mathbf{g}_w; \mathbf{m}_w; \mathbf{c}_w] = \mathbf{R}_{wi} [\mathbf{g}_i; \mathbf{m}_i; \mathbf{c}_i] \quad (2)$$

However, an additional adaptation is needed because the result is not guaranteed to be a rotation matrix. Another approach is given by the following system, which allows a non-linear optimization on the tangent space and ensures that the result is a rotation matrix:

$$\begin{bmatrix} \mathbf{g}_w \\ \mathbf{m}_w \\ \mathbf{c}_w \end{bmatrix} = \begin{bmatrix} e^{[\omega]} & 0 & 0 \\ 0 & e^{[\omega]} & 0 \\ 0 & 0 & e^{[\omega]} \end{bmatrix} \begin{bmatrix} \mathbf{g}_i \\ \mathbf{m}_i \\ \mathbf{c}_i \end{bmatrix} \quad (3)$$

With Jacobian:

$$\frac{\partial}{\partial \omega} h(\omega) = [-e^{[\omega]} [\mathbf{g}_i]; -e^{[\omega]} [\mathbf{m}_i]; -e^{[\omega]} [\mathbf{c}_i]] \quad (4)$$

Finally, our preferred method because of its simplicity, forms a least squares problem according to Wahba's problem [32]. This approach ensures a rotation matrix and allows to introduce weights for individual inputs. For the present application, it is of interest to associate a higher weight to the GNSS-based information because it is less likely to be disturbed by the environment in contrast to the magnetometer. The magnetometer measurements only serve the purpose of supporting the third degree of freedom for the orientation information. The weight α for the GNSS vector is chosen heuristically and is set to 50:1.

$$\mathbf{A} = \alpha \mathbf{g}_w \mathbf{g}_i^T + \mathbf{m}_w \mathbf{m}_i^T + \mathbf{c}_w \mathbf{c}_i^T \quad (5)$$

$$\mathbf{A} = \mathbf{USV}^T \quad (6)$$

$$\mathbf{R} = \mathbf{UMV}^T \quad (7)$$

$$\text{with } \mathbf{M} = \text{diag}([1; 1; \det(\mathbf{U})\det(\mathbf{V})]) \quad (8)$$

The position of the IMU expressed in the world-frame, given that GNSS point measurement $G2$ is the origin of the virtual GNSS-frame, is then calculated as $\mathbf{p}_{WI} = \mathbf{p}_{W_G2} + \mathbf{R}_{W_VG} \mathbf{R}_{VG_I} (-\mathbf{p}_{I_G2})$. The three methods for rotation estimation described above are provided with the data set tools and can be chosen accordingly.

In this context, it is also important to note that the rotation can be incorrect if the magnetometer provides incorrect information. This can be the case where the vehicle is picked up at the beginning of an experiment and the magnetic field is disturbed by the person of the field crew. The increased weight on the GNSS vector when calculating the 3 DoF rotation reduces this effect during this time, and such disturbance does not occur during the main phase of an experiment.

B. Sensor Time Calibration

The synchronization for the computation platforms of the vehicle are described in Section III-B9, however, internal system delays can still lead to minor time shifts of the sensor data. Thus, all sensors that are required for the generation of ground truth data are additionally time-synchronized in post-processing. This includes the two RTK GNSS signals, the PX4 IMU and magnetometer, as well as the motion capture system.

To synchronize the two GNSS signals, we first calculate their velocity and use their convolution to find the time offset which provides the highest overlap of the two signals \dot{p}_{g1} and \dot{p}_{g2} . Using the synchronized GNSS information, the virtual orientation vector for the two GNSS measurements expressed in the world frame is built (see the calculation of VG in Section V-A). In order to synchronize the magnetometer data to the virtual GNSS vector, the time derivative of the horizontal projection for VG and the yaw rotation component measured by the magnetometer are used.

After this step, the time derivative of the rotational ground truth for the outdoor pose information and the motion capture pose is used for the synchronization towards the angular velocity of the PX4 IMU. This provides all sensor time-offsets to generate accurate ground truth information for the vehicle.

C. Transition Segment Alignment

As described in the previous sections, the indoor ground truth based on motion capture data and the fiducial marker-based position of the vehicle are expressed in the reference frame of the motion capture system because the main marker board is a known motion capture object. Thus, in order to gain continuous ground truth for the transition data sets, only the outdoor pose information based on the RTK GNSS and magnetometer solution and the fiducial marker-based vehicle pose need to be aligned. Since accurate RTK GNSS solutions become spars towards the entrance of the building, GNSS data from multiple entry approaches are used to increase the accuracy of the alignment. The final alignment between the outdoor GNSS information and the fiducial marker-based trajectory is done by using a least-squares solution to Wahba's problem, shown by e.g. [33]. The resulting ground truth after the alignment is shown by Figure 25. Should this solution be too inaccurate for a specific purpose, the local ground truth for each segment can be used.

VI. PROVIDING THE DATA

The data is provided in human-readable plain text format as dedicated CSV files with headers and uncompressed PNG

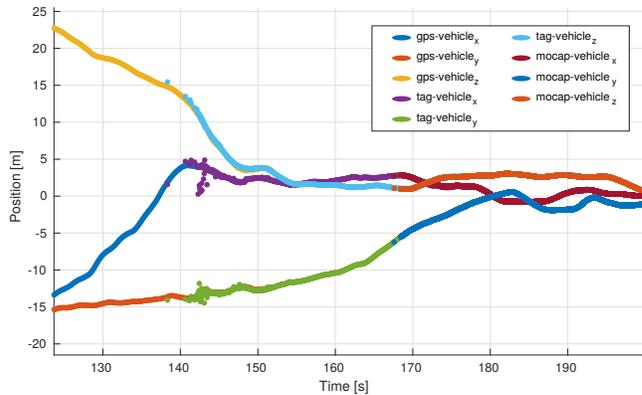


Fig. 25. Alignment of the ground truth segments for the outdoor area, the marker based transition section and the indoor motion capture pose. The shown trajectory builds the continuous ground truth for the outdoor-indoor datasets.

images with attached timestamps in a CSV file. Individual sensor calibration files are provided in YAML format. The structure of the files is outlined in Figure 28 and 29. A script for the conversion of this data into ROS bagfiles is provided with the toolbox. The generation of other data formats is possible by using the provided script as a template. The calibration of all sensors is provided with respect to the PX4 IMU and allows to construct any relative calibration that might be necessary.

Each outdoor data set provides two ground truth files. This ground truth is calculated using the RTK GNSS (8Hz) and the Magnetometer data (80Hz). Thus, two files are provided, `ground_truth_8hz.csv` provides data calculated based on RAW GNSS measurements and time-matched magnetometer measurements, while `ground_truth_80hz.csv` provides data based on interpolated RTK GNSS measurements and the raw magnetometer data. The usage may depend on the scenario and the choice is left to the user. Because the ground truth for the indoor data sets is generated using the motion capture system, `ground_truth.csv` provides this data for clear association.

The ground truth data for the transition dataset contains the segments for outdoor, transition, and indoor poses. The `ground_truth.csv` file for these scenarios contains the aligned pose information. Since data from multiple sources with multiple rates are used, the ground truth file also has segments with different frame rates.

Further, all data for individual sensor calibration is provided. This includes static IMU data for the generation of the Allan variance, magnetometer intrinsic calibration data, and camera to IMU calibration sequences.

Measurements are not hardware synced, which is the case for most real-world applications and is a strength of the provided data set.

VII. USABILITY OF THE DATA SET

This section provides two bench-marking scenarios that illustrate the data sets usability and how the provided ground truth relates to the estimated results. The first scenario

uses a state-of-the-art EKF that utilizes multiple sensors and includes a loosely coupled VIO component for an outdoor to indoor transition dataset. The second scenario specifically addresses the VIO navigation aspects by deploying a state-of-the-art VIO framework that solely uses the navigation camera stream and the main IMU.

A. Usage and Result with a state-of-the-art multi-sensor EKF

This test uses MaRS, a modular sensor fusion framework introduced by [34], to process outdoor to indoor transition data set one (see Tab. V). This represents a possible application in which a continuous state estimate is generated while the sensor data streams are selected/switched based on their availability.

The transition data sets consist of three elements as detailed in Section IV. The first section begins outdoors, and the filter uses the PX4 IMU, RTK GNSS measurements (position and velocity), vision-based pose as well as one magnetometer. As the vehicle approaches the building (referring to the map in Fig. 12), sector three is entered, and the quality of the GNSS measurements degrade, which triggers a χ^2 rejection test. After the number of rejections reaches a certain threshold, the sensor is considered unreliable and temporarily removed as an input sensor. The estimate continues with VIO pose information and magnetometer. After passing through the transition phase, the vehicle enters sector four by navigating through the entry of the building. This area is prone to magnetic disturbances, and consecutive magnetometer measurement rejections, which recover after the indoor area of the building is entered. At this point, the filter initializes a reference frame for the motion capture measurements, as soon as they become available, to reference the measurements in the navigation world frame.

The result for this experiment is shown by Figure 26. As can be seen, the estimate matches the ground-truth data well, despite the transition phase in which sensor switching and consecutive reference frame adaptations occurred. The difference in the indoor section results from a vision drift accumulated during the vision-only phase with challenging vision inputs. The data set aims to provide the means to improve corresponding algorithms to improve estimation performance for the presented scenarios in the future.

B. Usage and Evaluation of Example Data with State-of-the-Art VIO

As a validation of the usability of the data sets for VIO algorithm evaluation, raw IMU data as well as camera images from the Mars analog data set 18 were used to run OpenVins [35], an open-source, state-of-the-art Multi-State Constrained Kalman Filter (MSCKF) based VIO framework. Figure 27 shows the estimated trajectory compared with RTK-GNSS based ground truth. The estimated trajectory and ground truth are aligned along the four unobservable degrees of freedom of VIO, with the Umeyama alignment algorithm [36] computed over the full trajectory. The error of the estimate is calculated after the trajectory alignment and based on the absolute trajectory error (ATE) introduced by [37]. The absolute trajectory error of the position is 7.5m and

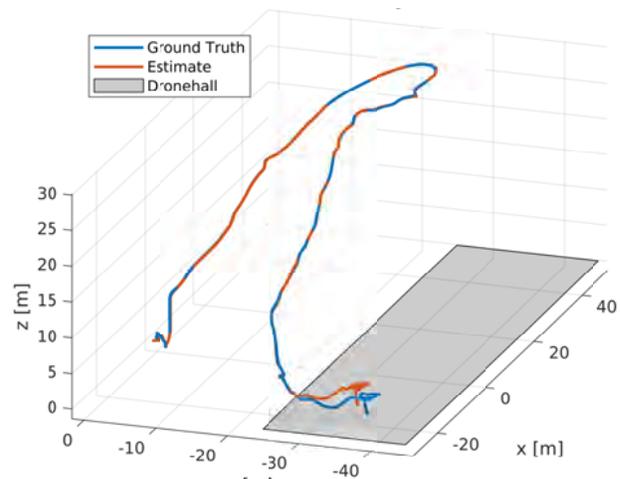


Fig. 26. State-of-the-art EKF filtering results for transition data set one, overlaid with the ground truth provided by the data set. The shaded area represents the indoor segment.

the rotation error is 7° . It is important to note that this data set poses particular challenges to vision-based algorithms. During the takeoff phase, the shadow of the platform with spinning propellers is in view of the navigation camera, resulting in moving features being tracked even though no motion is experienced by the platform itself. Moreover, during the flight, an almost featureless ground is seen from the camera resulting in bad feature tracking (see the top left segment of Fig. 27). Novel approaches as introduced by [38] are currently tackling this issue.

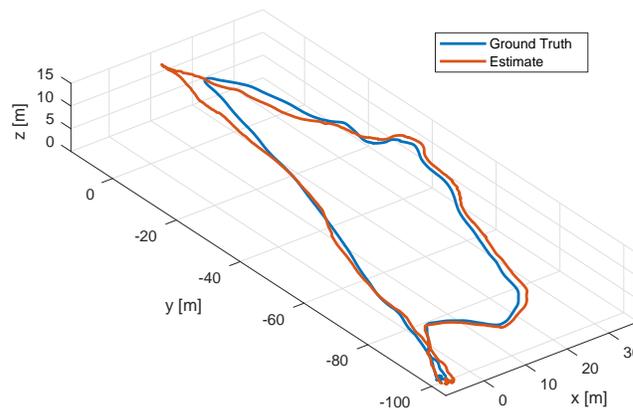


Fig. 27. Illustrative visual inertial localization example using a planar Mars analog data set with state-of-the-art methods. The estimated trajectory in red in comparison to the provided ground truth of the data set (blue). The estimated relative trajectory was globally aligned with the ground truth.

VIII. LESSONS LEARNED

This project posed challenges in terms of system and sensor setup as well as environmental difficulties. A few aspects have already been outlined in the System Setup Section (III). However, this section concludes the main difficulties that had to be overcome.

The generation of ground truth data posed multiple difficulties. First, the acquisition of highly accurate GNSS measurements was not straightforward because the vehicle hosts

several computation boards as well as high-frequency data lines, which cause EMI. Thus, the electronic components and the GNSS antennas required customized shielding (see Fig. 2 and 8) to reduce interfering signals to a level at which the GNSS provides position accuracy of 1cm. Section III-B6 discusses this issue in detail. Another difficulty was the generation of continuous ground truth information for the outdoor to indoor transition area. This is due to RTK GNSS signals not being available close to the building area, and the motion capture system not covering the outside of the building entrance. Thus, a field of fiducial markers was used to provide ground truth for this transition area. This pose data is aligned with valid RTK GNSS data for the outdoor area and the motion capture reference frame via marker-object association. This method allows for global and local ground truth for each of the three segments of the outdoor-indoor transition trajectories.

Finally, the recording of high amounts of data, as done for this project, was a challenge. Special attention was given to the two onboard embedded platforms to reach balanced computational loads and, most importantly, the full use of available data interfaces. As shown by Figure 1, each board makes use of two storage entities with different data throughputs. USB3 interfaces also needed to be shared between storage devices and sensors. Thus, the data communication and recording needed to be designed to make use of the maximum data capacity that each interface and storage unit can provide. Overall, the system needed to write to an SD and SSD medium on each of the two embedded boards, thus requiring a total of four storage locations to record the data with the given measurement rate.

IX. CONCLUSION AND FUTURE WORK

This work introduced the INSANE data set collection, which aims to provide multi-disciplinary in-flight data with a versatile sensor suite that is subject to real-world sensor effects. The flight scenarios address various research domains for vehicle localization. This work discussed individual aspects of the sensors and their integration in detail. The raw data for customized sensor calibrations and the analysis of specific sensor properties such as intrinsic, extrinsic and behavior for vibration behavior in a dedicated test bench setup are provided.

The quality of a data set is directly correlated with the accuracy and uninterrupted availability of the accompanying ground truth information. Thus we presented a carefully designed setup to directly measure highly accurate ground truth as raw data. In contrast to other work, we do not need to filter this data to achieve the high accuracy which prevents any filter-induced artifacts such as biases and inconsistencies.

Finally, we presented two show-cases which demonstrate the usability of the data sets with comparisons to the provided ground truth.

The lessons learned throughout this work are a crucial stepping stone for the development and extension of novel flight platforms. As for future aspects, it is planned to extend the data set with new scenarios and sensor setups over time. Thus, the open-sourced data set is not a static entity and will grow over time, following the same format for compatibility. Given

the versatility of the presented setup, we believe that this data will enable researchers to develop and test future algorithms addressing real-world challenges in various disciplines.

ACKNOWLEDGMENTS

A portion of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004).

The authors would like to thank the Austrian Space Forum (OeWF) for the possibility to perform tests and record datasets during the Analog Mars Mission AMADEE-20⁹.

The authors would also like to express a special thanks to the technical personnel, Fred Arneitz, and Patrik Grausberg for their support of the presented work.

FUNDING

Research was sponsored by the Army Research Office and was accomplished under Cooperative Agreement Number W911NF-21-2-0245. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

This work has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement 871260.

REFERENCES

- [1] T. A. Rodrigues, J. Patrikar, A. Choudhry, J. Feldgoise, V. Arcot, A. Gahlaut, S. Lau, B. Moon, B. Wagner, H. S. Matthews, S. Scherer, and C. Samaras, "Data collected with package delivery quadcopter drone," 2021. [Online]. Available: https://kithub.cmu.edu/articles/dataset/Data_Collected_with_Package_Delivery_Quadcopter_Drone/126834531
- [2] B. Congram and T. D. Barfoot, "Relatively lazy: Indoor-outdoor navigation using vision and GNSS," in *2021 18th Conference on Robots and Vision (CRV)*. IEEE, May 2021, pp. 1–8. [Online]. Available: <https://doi.org/10.1109/crv52889.2021.00015>
- [3] C. Brommer, D. Malyuta, D. Hentzen, and R. Brockers, "Long-duration autonomy for small rotorcraft UAS including recharging," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, Oct. 2018, pp. 7252–7258. [Online]. Available: <https://ieeexplore.ieee.org/document/8594111>
- [4] A. Fornasier, M. Scheiber, A. Hardt-Stremayr, R. Jung, and S. Weiss, "VINSEval: Evaluation Framework for Unified Testing of Consistency and Robustness of Visual-Inertial Navigation System Algorithms," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 2021, pp. 13 754–13 760. [Online]. Available: <https://doi.org/10.1109/icra48506.2021.9561382>
- [5] W. Wang, D. Zhu, X. Wang, Y. Hu, Y. Qiu, C. Wang, Y. Hu, A. Kapoor, and S. Scherer, "TartanAir: A dataset to push the limits of visual SLAM," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, Oct. 2020, pp. 4909–4916. [Online]. Available: <https://doi.org/10.1109/iros45743.2020.9341801>
- [6] A. Antonini, W. Guerra, V. Murali, T. Sayre-McCord, and S. Karaman, "The blackbird UAV dataset," *The International Journal of Robotics Research*, vol. 39, no. 10-11, pp. 1346–1364, Mar. 2020. [Online]. Available: <https://doi.org/10.1177/0278364920908331>

⁹OeWF - AMADEE20

- [7] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The KITTI dataset," *The International Journal of Robotics Research*, vol. 32, no. 11, pp. 1231–1237, Aug. 2013. [Online]. Available: <https://doi.org/10.1177/0278364913491297>
- [8] W. Maddern, G. Pascoe, C. Linegar, and P. Newman, "1 year, 1000 km: The oxford RobotCar dataset," *The International Journal of Robotics Research*, vol. 36, no. 1, pp. 3–15, Nov. 2016. [Online]. Available: <https://doi.org/10.1177/0278364916679498>
- [9] W. Maddern, G. Pascoe, M. Gadd, D. Barnes, B. Yeomans, and P. Newman, "Real-time kinematic ground truth for the oxford robotcar dataset," 2020. [Online]. Available: <https://arxiv.org/abs/2002.10152>
- [10] T. Pire, M. Mujica, J. Civera, and E. Kofman, "The rosario dataset: Multisensor data for localization and mapping in agricultural environments," *The International Journal of Robotics Research*, vol. 38, no. 6, pp. 633–641, 2019. [Online]. Available: <https://doi.org/10.1177/0278364919841437>
- [11] F. Walch, C. Hazirbas, L. Leal-Taixe, T. Sattler, S. Hilsenbeck, and D. Cremers, "Image-based localization using lstms for structured feature correlation," in *2017 IEEE International Conference on Computer Vision (ICCV)*, October 2017, pp. 627–637. [Online]. Available: <https://doi.org/10.1109/iccv.2017.75>
- [12] Y. Liu, W. Gao, and Z. Hu, "A large-scale dataset for indoor visual localization with high-precision ground truth," *The International Journal of Robotics Research*, vol. 41, no. 2, pp. 129–135, Oct. 2021. [Online]. Available: <https://doi.org/10.1177/02783649211052064>
- [13] D. Lee, S. Ryu, S. Yeon, Y. Lee, D. Kim, C. Han, Y. Cabon, P. Weinzapfel, N. Guérin, G. Csurka, and M. Humenberger, "Large-scale Localization Datasets in Crowded Indoor Spaces," 2021. [Online]. Available: <https://arxiv.org/abs/2105.08941>
- [14] B. Pfrommer, N. Sanket, K. Daniilidis, and J. Cleveland, "PennCOSYVIO: A challenging Visual Inertial Odometry benchmark," *Proceedings - IEEE International Conference on Robotics and Automation*, pp. 3847–3854, 2017. [Online]. Available: <https://doi.org/10.1109/icra.2017.7989443>
- [15] D. Schubert, T. Goll, N. Demmel, V. Usenko, J. Stuckler, and D. Cremers, "The TUM VI Benchmark for Evaluating Visual-Inertial Odometry," *IEEE International Conference on Intelligent Robots and Systems*, pp. 1680–1687, 2018. [Online]. Available: <https://doi.org/10.1109/iros.2018.8593419>
- [16] M. Burri, J. Nikolic, P. Gohl, T. Schneider, J. Rehder, S. Omari, M. W. Achtelik, and R. Siegwart, "The EuRoC micro aerial vehicle datasets," *The International Journal of Robotics Research*, vol. 35, no. 10, pp. 1157–1163, Jan. 2016. [Online]. Available: <https://doi.org/10.1177/0278364915620033>
- [17] J. Delmerico, T. Cieslewski, H. Rebecq, M. Faessler, and D. Scaramuzza, "Are we ready for autonomous drone racing? the uzr-fpv drone racing dataset," in *2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 6713–6719. [Online]. Available: <https://doi.org/10.1109/icra.2019.8793887>
- [18] L. Meyer, M. Smísek, A. Fontan Villacampa, L. Oliva Maza, D. Medina, M. J. Schuster, F. Steidle, M. Vayugundla, M. G. Müller, B. Rebele, A. Wedler, and R. Triebel, "The MADMAX data set for visual-inertial rover navigation on Mars," *Journal of Field Robotics*, vol. 38, no. 6, pp. 833–853, 2021. [Online]. Available: <https://doi.org/10.1002/rob.22016>
- [19] M. Scheiber, A. Fornasier, R. Jung, C. Böhm, R. Dhakate, C. Stewart, J. Steinbrener, S. Weiss, and C. Brommer, "Flight stack for reproducible and customizable autonomy applications in research and industry," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 11 283–11 290, 2022. [Online]. Available: <https://ieeexplore.ieee.org/document/9849131>
- [20] C. Stewart, "Skiffos: Minimal cross-compiled linux for embedded containers," 2021. [Online]. Available: <https://arxiv.org/abs/2104.00048>
- [21] J. Steinbrener, C. Brommer, T. Jantos, A. Fornasier, and S. Weiss, "Improved State Propagation through AI-based Pre-processing and Down-sampling of High-Speed Inertial Data," *2022 International Conference on Robotics and Automation (ICRA)*, pp. 1–8, 2022.
- [22] Freescale, "Allan Variance: Noise Analysis for Gyroscopes," *Freescale white paper*, p. 9, 2015. [Online]. Available: http://cache.freescale.com/files/sensors/doc/app_note/AN5087.pdf
- [23] J. F. Vasconcelos, G. Elkaim, C. Silvestre, P. Oliveira, and B. Carneira, "Geometric approach to strapdown magnetometer calibration in sensor frame," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 2, pp. 1293–1306, 2011. [Online]. Available: <https://doi.org/10.1109/taes.2011.5751259>
- [24] K. Papafotis and P. P. Sotiriadis, "A Fast and Accurate Accelerometer and Magnetometer Alignment Algorithm," *IEEE Sensors Journal*, vol. 20, no. 24, pp. 15 061–15 067, Dec. 2020. [Online]. Available: <https://doi.org/10.1109/jsen.2020.3011579>
- [25] C. Brommer, C. Böhm, J. Steinbrener, R. Brockers, and S. Weiss, "Improved state estimation in distorted magnetic fields," in *2020 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, Sep. 2020, pp. 1007–1013. [Online]. Available: <https://doi.org/10.1109/icuas48674.2020.9213913>
- [26] J. Rehder, J. Nikolic, T. Schneider, T. Hinzmann, and R. Siegwart, "Extending kalibr: Calibrating the extrinsics of multiple IMUs and of individual axes," *Proceedings - IEEE International Conference on Robotics and Automation*, vol. 2016-June, pp. 4304–4311, 2016. [Online]. Available: <https://doi.org/10.1109/icra.2016.7487628>
- [27] D. S. Bayard, D. T. Conway, R. Brockers, J. H. Delaune, L. H. Matthies, H. F. Grip, G. B. Merewether, T. L. Brown, and A. M. S. Martin, "Vision-based navigation for the NASA mars helicopter," in *AIAA Scitech 2019 Forum*. American Institute of Aeronautics and Astronautics, Jan. 2019, p. 1411. [Online]. Available: <https://doi.org/10.2514/6.2019-1411>
- [28] H. F. Grip, J. Lam, D. Bayard, D. T. Conway, G. Singh, R. Brockers, J. H. Delaune, L. H. Matthies, C. Malpica, T. L. Brown, A. Jain, A. M. S. Martin, and G. B. Merewether, "Flight control system for NASA's mars helicopter," in *AIAA Scitech 2019 Forum*. American Institute of Aeronautics and Astronautics, Jan. 2019, pp. 1–20. [Online]. Available: <https://doi.org/10.2514/6.2019-1289>
- [29] H.-n. Lin, C.-c. Lu, H.-Y. Tsai, and T.-W. Kung, "The Analysis of EMI Noise Coupling Mechanism for GPS Reception Performance Degradation from SSD / USB Module," *IEICW*, pp. 350–353, 2014.
- [30] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, and R. Medina-Carnicer, "Generation of fiducial marker dictionaries using Mixed Integer Linear Programming," *Pattern Recognition*, vol. 51, pp. 481–491, 2016. [Online]. Available: <https://doi.org/10.1016/j.patcog.2015.09.023>
- [31] C. Chen, X. Lu, A. Markham, and N. Trigoni, "Ionet: Learning to cure the curse of drift in inertial odometry," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, Apr. 2018. [Online]. Available: <https://doi.org/10.1609/aaai.v32i1.12102>
- [32] G. Wahba, "A least squares estimate of satellite attitude," *SIAM Review*, vol. 7, no. 3, pp. 409–409, Jul. 1965. [Online]. Available: <https://doi.org/10.1137/1007077>
- [33] O. Sorkine-Hornung and M. Rabinovich, "Least-squares rigid motion using svd," *Computing*, vol. 1, no. 1, pp. 1–5, 2017.
- [34] C. Brommer, R. Jung, J. Steinbrener, and S. Weiss, "MaRS: A modular and robust sensor-fusion framework," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 359–366, Apr. 2021. [Online]. Available: <https://ieeexplore.ieee.org/document/9286578>
- [35] P. Geneva, K. Eickenhoff, W. Lee, Y. Yang, and G. Huang, "OpenVINS: A research platform for visual-inertial estimation," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, Paris, France, 2020, pp. 4666–4672. [Online]. Available: https://github.com/rpng/open_vins
- [36] S. Umeyama, "Least-squares estimation of transformation parameters between two point patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 4, pp. 376–380, Apr. 1991. [Online]. Available: <https://doi.org/10.1109/34.88573>
- [37] Z. Zhang and D. Scaramuzza, "A tutorial on quantitative trajectory evaluation for visual-(inertial) odometry," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 7244–7251.
- [38] A. Hardt-Stremayr and S. Weiss, "Monocular visual-inertial odometry in low-textured environments with smooth gradients: A fully dense direct filtering approach," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, May 2020, pp. 7837–7843. [Online]. Available: <https://doi.org/10.1109/icra40945.2020.9196881>

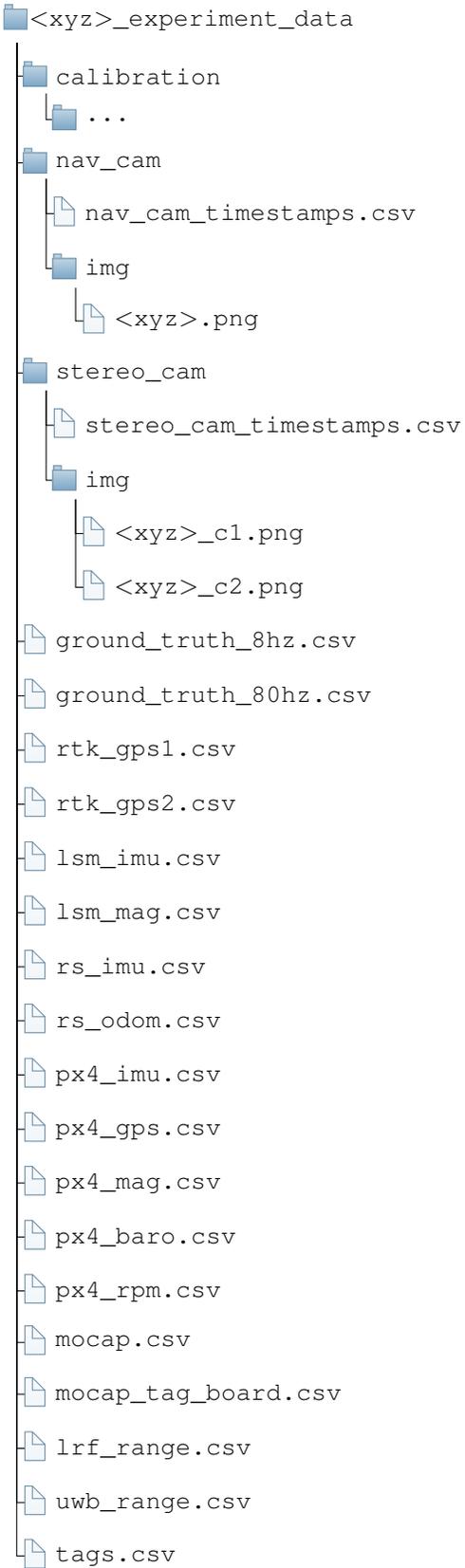


Fig. 28. Main data directory

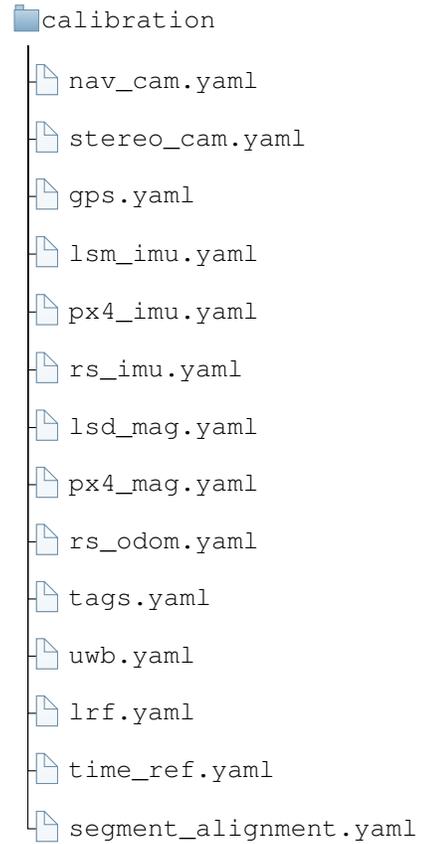


Fig. 29. Calibration directory