

USTC FlyingCar Dataset: Multisensor Data for Stronger Autonomous Aerial Work Robot

Journal Title
XX(X):1–10
©The Author(s) 2022
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/ToBeAssigned
www.sagepub.com/

SAGE

Ziming Wang, Yujiang Liu, Erbao Dong* and Yanyong Zhang

Abstract

In this paper we present *USTC FlyingCar Dataset*, dedicated to the development of simultaneous localization and mapping methods and precise 3D working space reconstruction for autonomous aerial work robots. In recent years, many public datasets have play an important roll on the progress of autonomous car and UAVs. However, for the platform of aerial work robots, UAVs are not strong enough and cars can not fly. Thus, to fill in this gap, we create “Giraffe” mapping robot based on bucket truck which equipped with a variety of well calibrated and synchronized sensors: four 3D lidars, two stereo cameras, two monocular cameras, multiple Inertial Measurement Units (IMUs) and GNSS/INS system. Laser tracker is used to record the millimeter accurate position ground truth. We also made its ground twin — “Okapi” mapping robot to gather data as comparison. The main characteristic of the dataset is the application of a typical autonomous driving sensing suite to aerial scenes. Therefore, we name the dataset “FlyingCar”. The vision of our work is to contribute to development of flying cars, especially the take-off and landing of VTOL (Vertical Takeoff And Landing) flying cars in real environments. The dataset is available for download at: <https://ustc-robot.github.io/>.

Keywords

Dataset, aerial work, aerial robotics, mobile robotics, SLAM, computer vision, cameras, lidar, flying car, bucket truck, aerial lift device

1 Introduction

Aerial work plays an important role in our daily life and industrial or agricultural production. Figure 1 illustrates a range of typical aerial work scenes. However, aerial work is characterized by low efficiency and high risk, workers are exposed to the danger of falling from high places, electrocution from overhead power lines, being trapped or squeezed, etc. Thousands of workers are seriously injured or killed each year due to aerial work. If robots can replace the workers to enter dangerous aerial working environments, the efficiency and safety of aerial work will greatly improve and life will be saved.

1.1 Challenges

The complete automation of aerial work faces challenges in several stages. The first is how to lift a heavy robot into the air. In order to be competent for aerial work, robots should be equipped with flexible and powerful robotic arms as an execution system, complex sensors as a perception system and sufficient computing resources for data processing and decision-making, which makes it difficult for the robot to achieve lightweighting at the current technical level. Therefore, the UAVs are limited by the load and cannot become a loading platform. To solve this problem, we find a suitable platform — bucket truck, which is a high-capacity construction vehicle equipped with an extendable, hydraulic boom carrying a large bucket for raising workers to elevated, inaccessible areas. The bucket trucks have both strength and flexibility, its load can usually reach more than 200kg, and any target position in its three-dimensional work space can be reached through the extension of the arms and the rotation

of the joints. A combined prototype of the working robot and the bucket truck is shown in Figure 2 (left).



Figure 1. Typical aerial work scenes in our daily life:
(a) repair and maintenance of electrical power facilities, (b) machinery manufacturing, (c) ship maintenance, (d) building construction, (e) tree trimming, and (f) aerial fire fighting and rescue

University of Science and Technology of China, 96 Jinzhai Road, Hefei, 230026, Anhui, China.

Corresponding author:

Erbao Dong, CAS Key Laboratory of Mechanical Behavior and Design of Materials, Department of Precision Machinery and Precision Instrumentation, University of Science and Technology of China, 96 Jinzhai Road, Hefei, Anhui Province, 230026, China.
Email: {zimingwang, yujiangliu}@mail.ustc.edu.cn, {ebdong, yanyongz}@ustc.edu.cn



Figure 2. Left: Right:

After completing the selection of the aerial platform, the challenge comes from how the robot interacts itself with the complex aerial work environment. Aerial robot or vehicle requires accurate and real-time localization, all-round accurate 3D perception and densification reconstruction of the environment via its visual and inertial system, which is a prerequisite for the next steps to perform object recognition, trajectory planning and control, scene understanding, etc.

Visual and multisensor fusion localization and mapping in aerial work environments presents a set of specific challenges. Many aerial objects are difficult to detection and reconstruction because of small and lack texture such as power lines, trusses, tree branches. Unlike other scenes such as storehouse or traffic, aerial work environment are usually cluttered and unstructured scenes, therefore it is difficult to use some general structured features to improve the effect of the algorithms. And some aerial environments are emptier than ground environments, which makes the matching and loopback detection of visual or lidar odometry more difficult. Compared with ground robots and cars, aerial robots have more degrees of freedom and sudden changes in motion. The rotation of the robot and the motion in three-dimensional space challenge the accuracy and robustness of the algorithm for complex motion and attitude. Aerial work robot is loaded on the huge hydraulic arm of the bucket truck, due to the mechanical characteristics it will shake at low frequency and vibrate at high frequency while working. At the same time, the constantly moving hydraulic arm will become a dynamic disturbance in the environment. These two reasons will affect the accuracy of localization and mapping. Aerial work robots also face some similar difficulties as other robots working outdoors. One of them is a high variety of lighting conditions in different weather, for example, vision sensors will fail caused by direct sunlight or the darkness at night. Enabling aerial work robots to meet these challenges and advance their practical applications is the motivation for us to make this dataset.

1.2 Related Works

Public datasets play essential roles when it comes to a new scene to the application of robots and other automated systems. Public datasets provide unified and fair benchmarks for evaluating various algorithms. Researchers can quickly investigate and verify existing algorithms and develop new algorithms for a new scene without the need to build expensive hardware equipment and tedious calibration and data collection. We have investigated and referenced some outstanding datasets that related to our aerial work dataset. The datasets are divided into two categories: ground and aerial. Table 1 gives a summary for details.

For ground scenes, the most representative one is autonomous driving. Autonomous driving has developed and progressed rapidly and made a big success in the past decade thanks to diverse public datasets. For example one of the most famous among them, the KITTI[1] dataset. A feature of this type of dataset is the abundance of sensors. Because of the load-carrying capacity and sufficient space of ground vehicles, various sensors such as stereo or surround view cameras, 3D Lidars, IMUs, INS/GNSS, etc. can be incorporated into a same spatiotemporal system for data collection. For localization and mapping tasks, ground scenes are usually based on RTK-GPS or Lidar SLAM to generate ground truth values with centimeter-level accuracy.

Aerial autonomous systems have also made progress in recent years. EuRoC dataset equips UAV (Unmanned Aerial Vehiclw) with synchronously triggered high frame rate stereo cameras and IMU for the first time to apply tightly coupled visual-inertial system into localization and attitude estimation of aerial robots. For aerial datasets, micro or small drones have been the aerial platforms used to collect the data. Therefore, due to the payload constraint, aerial platforms usually equipped with only a few cameras and inertial sensors. In indoor or small scenes, motion capture systems are used to record motion trajectories and poses to generate 6DoF millimeter-level ground truth. In outdoor or larger scenes, laser trackers are used to record motion trajectories to generate millimeter-level ground truth.

According to the aforementioned description, there is a gap between ground datasets and aerial dataset. In contrast, one is a large scene with more sensors for flat 2D motion, and the other is a small scene with fewer sensors for aggressive 3D motion. To fill this gap, there are two routes to choose from, one is to choose a drone platform with a larger payload to add sensors such as lidar, and the other is to let the car “fly” into the air for 3D movement. NTU VIRAL dataset[2] also has noticed the gap and they choose the first way to equip a larger UAV with two Ouster 16 lines 3D lidar in addition to a pair of cameras and an IMU. The second way sounds like a fantasy but that’s exactly what we do. Although the current flying car is in its infancy, it is feasible to make the autonomous driving multi-sensor perception platform perform flexible 3D motion within a certain range in the air with the help of the bucket truck. As an aerial dataset, USTC FlyingCar dataset uses a typical suite of sensors for autonomous driving, including steven cameras, four 3D lidars and three IMUs. When designing and making the sensor suite, we try to keep the similarities with the existing autonomous driving datasets as much as possible. For example, our Bumblebee stereo camera is the same as Oxford RobotCar and EU Long-term datasets. One of our horizontal 3D lidars — Velodyne HDL-32E is the same as nuScenes, Oxford Radar RobotCar, EU Long-term and NCLT[3] datasets. Therefore, it would be more convincing to compare algorithms using same hardwares between these datasets and USTC FlyingCar. The USTC FlyingCar dataset also has similarities with the existing aerial datasets. Because it is geared towards more delicate tasks, we obtain the millimeter-level outdoor ground truth based on the laser tracker.

Looking back at Table 1, it can be clearly seen that USTC FlyingCar is the aerial dataset with the most sensors. To our

Table 1. The choice of options.

Dataset	IMU	Sensors		Gruond Truth	Environment	Platform
		Camera	Lidar			
MIT DARPA[4]	N/A	4 point grey (gray): 1392×512x4	Velodyne HDL-64E	RTK GPS/INS	Outdoor (Urban)	Car
Ford Campus (2011 IJRR) [5]	6 axis Xsens MTi-G	4 point grey (gray): 1392×512x4	Velodyne HDL-64E	RTK GPS/INS	Outdoor	Car
KITTI (2013 IJRR) [1]	6 axis OXTS RT3003	4 point grey (gray): 1392×512x4	Velodyne HDL-64E	RTK GPS/INS	Outdoor (Urban)	Car
NCLT (2015 IJRR) [3]	N/A	4 point grey (gray): 1392×512x4	Velodyne HDL-64E	SLAM	Outdoor	ground robot
Oxford RobotCar (2016 IJRR) [6]	6 axis NovAtel SPAN-CPT ALIGN	Point Grey Bumblebee XB3: 1280×960x3 3 Point Grey Grasshopper2: 1024×1024	2 SICK LMS-151 2D LIDAR SICK LD-MRS 3D LIDAR	RTK GPS/INS	Outdoor	Car
Oxford Radar RobotCar [7] (2019 ICRA)	6 axis NovAtel SPAN-CPT ALIGN	Point Grey Bumblebee XB3: 1280×960x3 3 Point Grey Grasshopper2: 1024×1024x3	2 Velodyne HDL-32E	RTK GPS/INS	Outdoor	Car
Rosario (2019 IJRR) [8]	6 axis LSM6DS0	ZED stereo: 672×376x2	N/A	RTK GPS/INS	Outdoor (Agriculture)	ground robot
KAIST Urban (2019 IJRR) [9]	9 axis Xsens MTi-G-300	FLIR FL3-U3-20E4C-C: 1280×560x2	2 Velodyne VLP-16C	SLAM	Outdoor	Car
EU Long- term [9] (2020 IROS)	9 axis Xsens MTi-28A53G25	FLIR FL3-U3-20E4C-C: 1280×560x2	2 Velodyne VLP-16C	SLAM	Outdoor	Car
nuScenes (2020 CVPR)	9 axis Advanced Navigation Spatial	6 Basler acA1600-60gc: 1600x1200x6	Velodyne HDL-32E	RTK GPS/INS	Outdoor	Car
EuRoC (2016 IJRR)	6 axis ADIS16448	2 MT9V034:752×480x2	N/A	6DOF MoCap 3D Laser Tracker	Indoor	UAV
Zurich Urban	9 axis VectorNav VN100	2 uEye 1221 LE: 752×480x2	N/A	3D Laser Tracker	Outdoor	UAV
UZH-FPV	9 axis VectorNav VN100	2 uEye 1221 LE: 752×480x2	N/A	3D Laser Tracker	Indoor Outdoor	UAV
NTU VIRAL (2021 IJRR)	9 axis VectorNav VN100	2 uEye 1221 LE: 752×480x2	2 Ouster OS1-16	3D Laser Tracker	Outdoor	UAV
USTC FlyingCar	9 axis Xsens MTi-G-710	Point Grey Bumblebee XB3: 1280×960x3 Point Grey Bumblebee XB2: 1024x768x2 Hikvision MV-CB016-10GC 1440 × 1080 Hikvision MV-CE060-10UC 3072×2048	Velodyne HDL-32E Velodyne VLP-32C Ouster OS0-128 LiVOX Avia	6DOF MoCap 3D Laser Tracker	Outdoor	Bucket Truck/ ground robot

knowledge, it is also the first dataset in the world to apply the typical autonomous driving sensing suite to aerial scenes. This work is unique in its scope. We believe that our dataset is a significant contribution to provide benchmarks of testing existing algorithms for autonomous system and develop new ones that are more suited to the particularities of aerial work scenes.

Going a step further, we named the dataset “FlyingCar” because our vision is to advance the development of autonomous flying cars. S. S. Ahmed[10] has summarized challenges and strategies toward flying car future adoption. He argues that the ideal flying cars should have a high degree of autonomy and vertical take-off and landing (VTOL) capabilities. Advanced robotics and sensor fusion technology is an important factor in promoting ongoing development of flying car. Besides, the most essential part of the safe operation of a flying car will be the ground/air transition (takeoff/landing). As shown in Figure 2 (right), we draw an imaginary diagram of the flying car taking off and

landing. How to correctly perceive, identify and safely avoid obstacles in the space adjacent to the ground during operation, such as power lines, trees, and buildings, is an important challenge to flying car. Therefore, aside from other parts of a flying car, let us focuses on its multi-sensor fusion perception, localization and mapping during its movement. USTC FlyingCar dataset has the motion characteristics of VTOL like flying cars, i.e. ascent, descent and hover in the z direction, smooth rather than aggressive pitch and rotation in the air. If researchers want to test the performance of existing ground vehicle autonomous driving solutions in the air or develop and validate new algorithms, our dataset will provide real-world data in reference for flying cars during take-off and landing “parking” phases. The USTC FlyingCar dataset will become a bridge for the application of autonomous driving fusion perception algorithms to the field of flying cars.

We organize the rest of the paper as follows: Section 2 describes in detail the various components of the data

acquisition systems and the parameters, characteristics and function of each sensor.

2 Acquisition Systems and Sensor Setup

The data was collected using the “Giraffe” acquisition system and “Okapi” acquisition system. As shown in Figure 3, the “Giraffe” system is an aerial system consisting of (a) multisensor data collection platform, (b) laser tracker ground truth system and (c) bucket truck; the “Okapi” system is a ground system similar to an autonomous vehicle, which equips the same sensors (a) and ground truth record system (b) on a ground robot (d) for data acquisition as a ground motion comparison with the data recorded by aerial system.

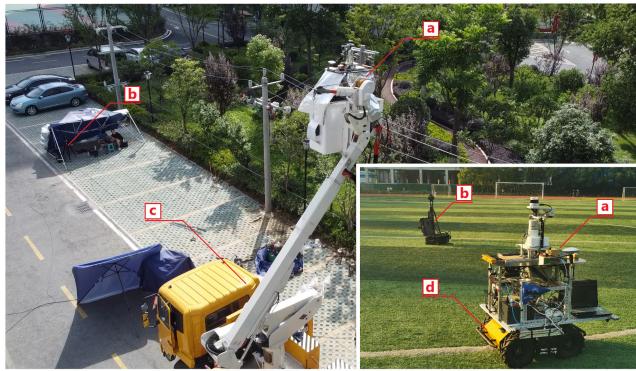


Figure 3. “Giraffe” and “Okapi” acquisition systems:
“Giraffe” aerial system: (a), (b) and (c).
“Okapi” ground system: (a), (b) and (d)
(a) multisensor data collection platform (Fig. 7), (b) laser
tracker ground truth system, (c) bucket truck, (d) ground robot.

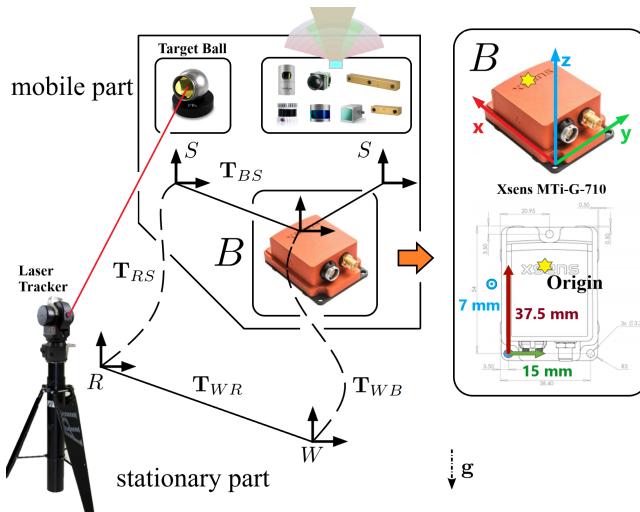


Figure 4. system trans

Both Systems are equipped with the following sensors:

2.1 Inertial Measurement Unit

The main IMU of the acquisition system is an [Xsens MTi-G-710¹](#) INS/GNSS module installed in the center of the system. The body frame is defined to be aligned with the Xsens sensor frame.

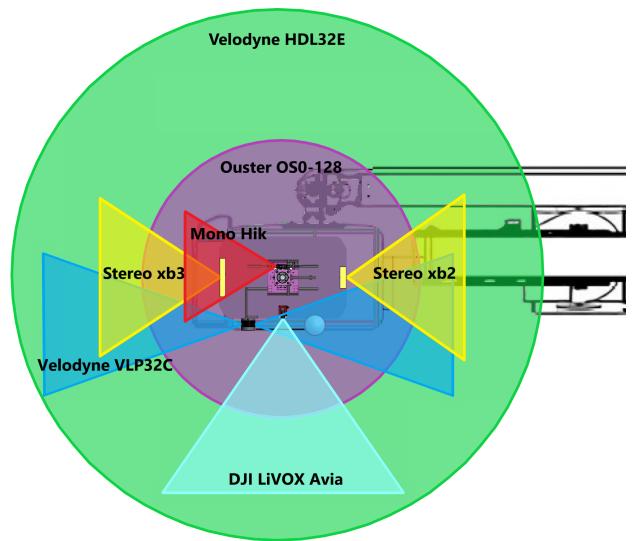


Figure 5. The visual scope of the acquisition system sensors,
top view.

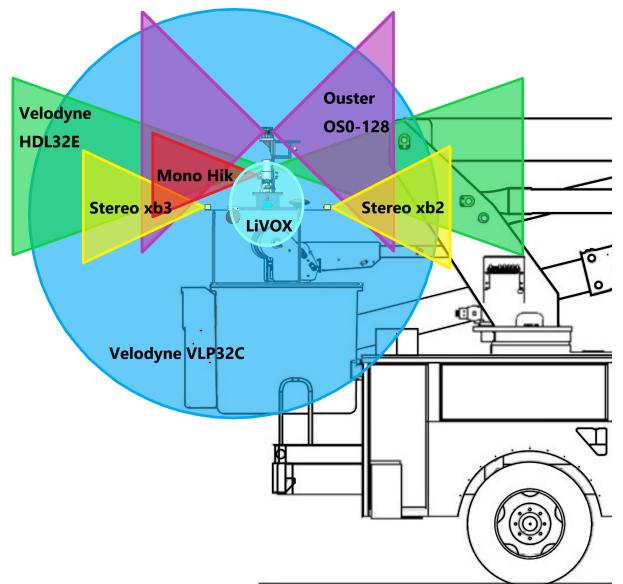


Figure 6. The visual scope of the acquisition system sensors,
side view.

- 1 × Xsens MTi-G-710 INS/GNSS, 9 axis, 400 Hz, accuracy: 0.2° in roll/pitch, 0.8° in heading.

Xsens outputs the three-axis acceleration and three-axis angular velocity in its own coordinate system, and the quaternion attitude in the north-east-down (NED) coordinate system. Xsens is hardware synchronized to the same external GPS clock source with the cameras and lidars in the system, making up visual-inertial and lidar-inertial sensor units together. Two extra 6 axis IMUs are in OS0-128 and LiVOX Avia lidars, as part of lidar-inertial sensor unit.

2.2 3D Lidars

3D lidars are important to 3D scenes accurate perception and understanding. This work includes three different kinds of mainstream lidars — Digital Lidar, Mechanical Lidar and MEMS Lidar to provide aerial autonomous systems with no-blind-spot perception covering a 720° range.

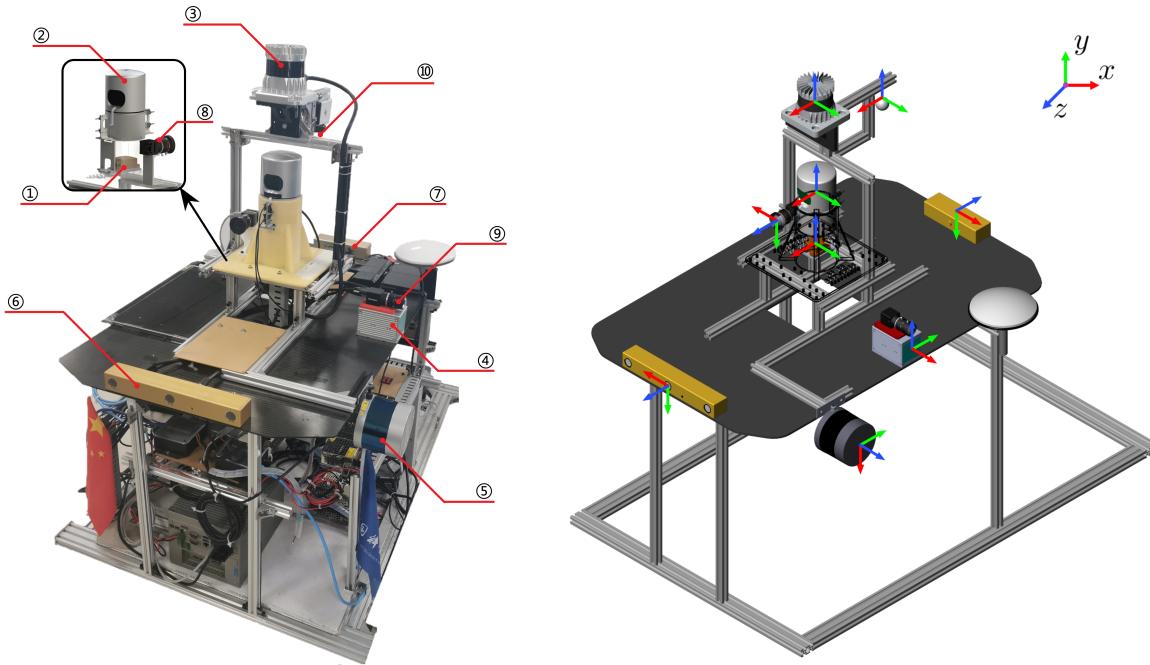


Figure 7. yes

- Digital Lidar: 1 × Ouster OS0-128, 10 Hz, 128 beams, 0.7° angular resolution, ± 1.5 to ± 5 cm distance accuracy, collecting 2.62 million points/second, field of view: 360° HFoV, 90° VFoV ($\pm 45^\circ$), range: 50 m

- Mechanical Lidar: 1 × Velodyne HDL-32E, 5/10 Hz, 32 beams, 1.33° angular resolution, ± 2 cm distance accuracy, collecting 1.39 million points/second, field of view: 360° HFoV, 41.3° VFoV ($+10.67^\circ$ to -30.67°), range: 100 m

- Mechanical Lidar: 1 × Velodyne VLP-32C, 10 Hz, 32 beams, 0.33° angular resolution (non-linear distribution), ± 3 cm distance accuracy, collecting 1.20 million points/second, field of view: 360° HFoV, 40° VFoV (-25° to $+15^\circ$), range: 200 m

- MEMS Lidar: 1 × DJI LiVOX Avia, 10 Hz, 2 cm distance accuracy, collecting 0.24 million points/second, field of view: 70.4° HFoV, 77.2° VFoV (Non-repetitive Scanning), range: 450 m

The visual scope and installation position of each lidar are shown in Figure 5 and 6. Ouster OS0-128 and Velodyne HDL32E lidars are installed in the center of the system as the main source of 3D environmental data for the system. In the actual work process, the point cloud accuracy of Velodyne HDL32E is the highest. And the Ouster OS0-128 lidar is a digital lidar. Digital lidar is based on custom system-on-a-chip (“SoC”) with single photon avalanche diode (“SPAD”) detectors. Therefore, it can not only output pointclouds, but also output depth images and signal-intensity images of lidar and visible light spectrum (Figure 11).

Another 360° 32 beams lidar Velodyne VLP32C is installed vertically on the side of the acquisition system to supplement the blind area of vision. The aerial robot moves in three dimensions, and objects such as wires and branches may suddenly appear in all directions of the robot as obstacles. Therefore, the robot needs 720 degrees of

perception in the horizontal and vertical directions. Besides, since the objects above the robot are few and small, there may only a few points in the point cloud within 180 degrees above the lidar.

A MEMS Lidar LiVOX Avia is mounted horizontally on the sensor platform. A MEMS Lidar LiVOX Avia is mounted horizontally on the sensor platform. The main characteristics of LiVOX Avia is that it has a view like camera and the mode of non-repetitive scanning. Pointclouds from LiVOX lidar scans are uniformly accumulated on the map over time.

2.3 Monocular and Stereo Cameras

- 1 x Point Grey Bumblebee XB3 (BBX3-13S2C-38) trinocular stereo camera, $1280 \times 960 \times 3$, 10Hz, Sony ICX445 CCD, 1/3, 3.75m, global shutter, 3.8mm lens, 66° HFoV, 12/24cm baseline

- 1 x Point Grey Bumblebee XB2 (BBX2-08S2C-38) binocular stereo camera, $1024 \times 768 \times 2$, 10-15Hz, Sony ICX204 CCD, 1/3, 4.65 μ m, global shutter, 3.8mm lens, 70° HFoV, 12cm baseline

- 1 x Hikvision MV-CB016-10GC-C industrial monocular camera, 1440×1080 , 20Hz, Sony IMX296 CCD, 1/2.9, 3.45m, global shutter, 6mm lens (MVL-HF0628M-6MPE), 63.11° HFoV

2.4 Laser Tracker

The laser tracker is the only sensor that is independent of the overall sensor system, as shown in Figure 3 (b).

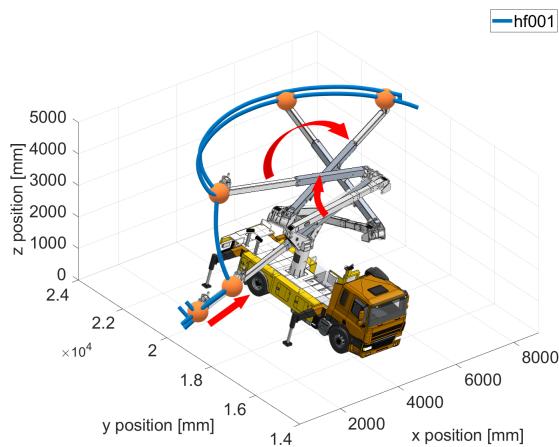
- API T3 Laser Tracker, azimuth: $\pm 320^\circ$ (640° end to end), angular resolution: ± 0.018 arc-seconds, angular accuracy: 3.5m/meter, system resolution: 0.1m, maximum lateral target speed: 4 meters/sec, maximum acceleration: 2 g, internal level accuracy: ± 2 arc-second, linear range: 80 m.

Table 2. The choice of options.

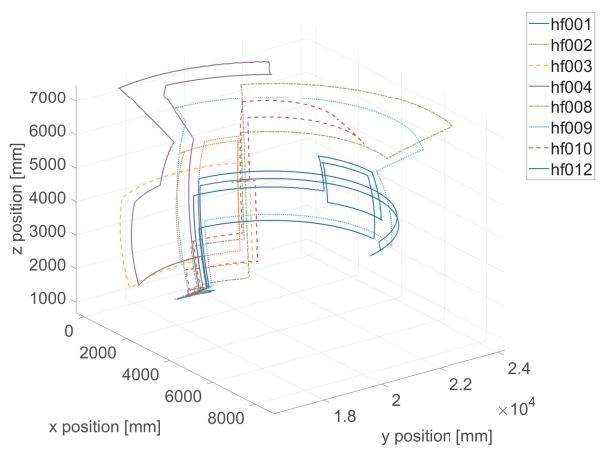
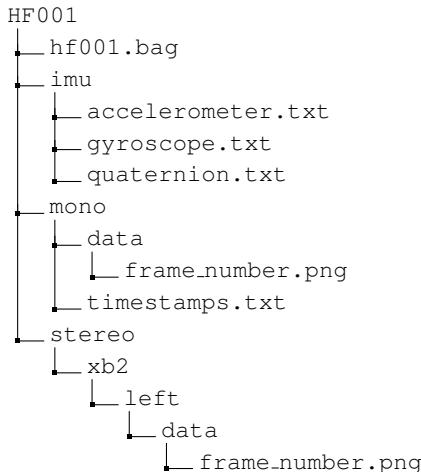
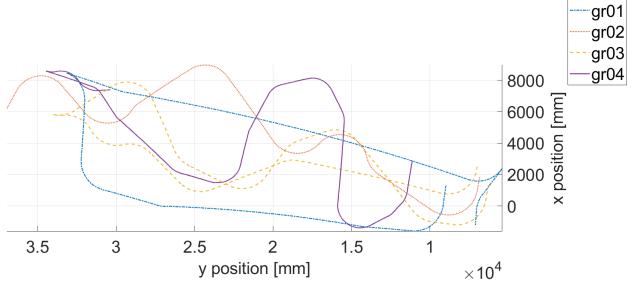
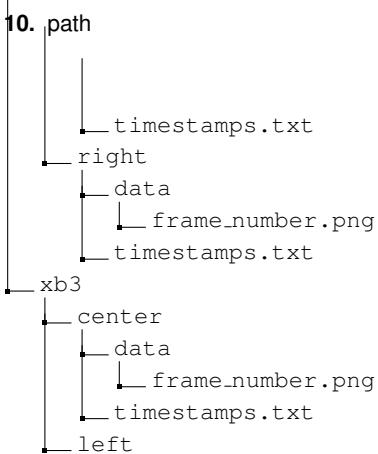
No	Sensor	Model	ROS Topic	Message type	Rate
1	IMU/INS	Xsens MTi-G-710	/imu/data	sensor_msgs/Imu	400Hz
2	Horizontal Lidar 1	Velodyne HDL32E	/velodyne_points_HDL32	sensor_msgs/PointCloud2	5/10Hz (rotate at 10Hz)
3	Horizontal Lidar 2	Ouster OS0-128	/os0.cloud.node imu /os0.cloud.node points	sensor_msgs/PointCloud2	10Hz
4	Horizontal Lidar 3	LiVOX Avia	/livox/lidar /livox/imu	livox_ros_driver/CustomMsg sensor_msgs/Imu	10Hz 200Hz
5	Vertical Lidar 1	Velodyne VLP32C	/velodyne_points_VLP32	sensor_msgs/PointCloud2	10Hz
6	Stereo Camera front	PointGrey Bumblebee xb3	/camera/left/image_raw /camera/center/image_raw /camera/right/image_raw	sensor_msgs/Image	10Hz
7	Stereo Camera back	PointGrey Bumblebee xb2	/cam_xb2/left/image_raw /cam_xb2/right/image_raw	sensor_msgs/Image	10-15Hz
8	Mono Camera 1	Hikvision MV-CB016-10GC-C	/hik_camera/iamge_raw	sensor_msgs/Image	20Hz
9	Mono Camera 2	Hikvision MV-CE060-10UC		sensor_msgs/Image	20Hz

3 Dataset

Representative scenes of aerial work were selected to collect data. The surrounding objects include power lines, trees, buildings, roads, etc. At the same time, we collected data from morning to night and under different weather conditions, in order to ensure that the aerial work robot can work around the clock.

**Figure 8.** path

KITTI format

**Figure 9.** path**Figure 10.** path

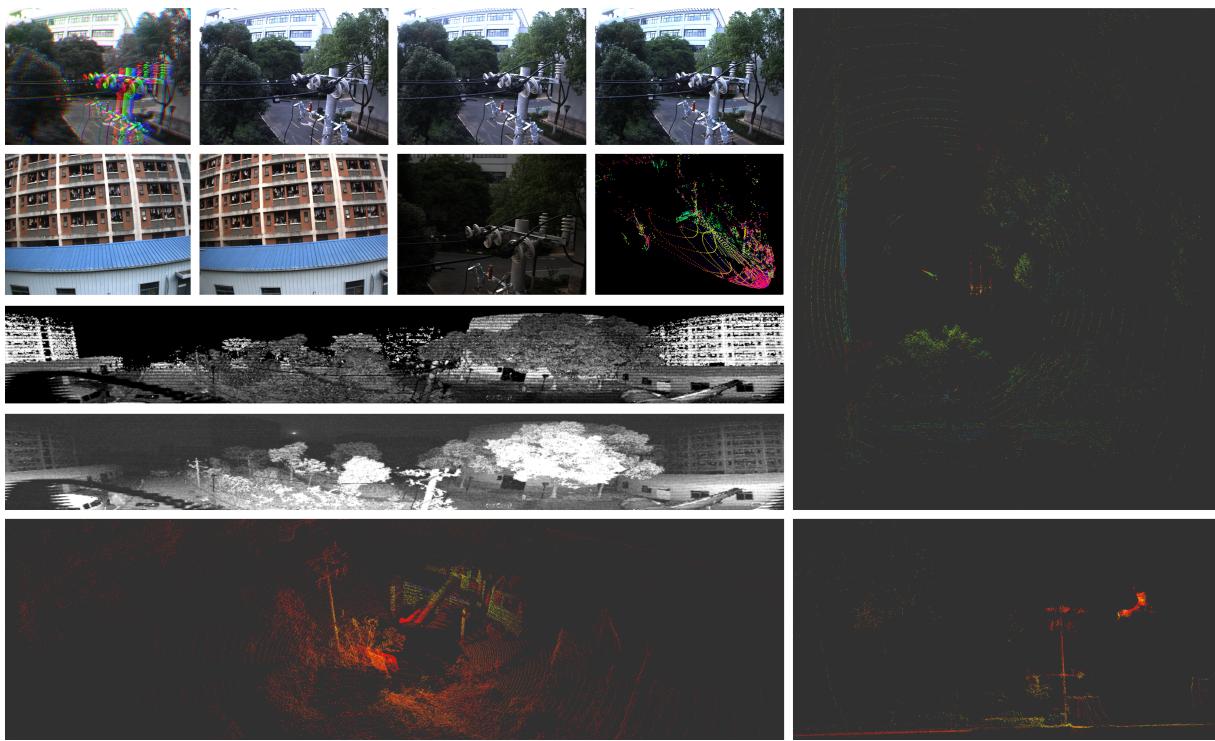
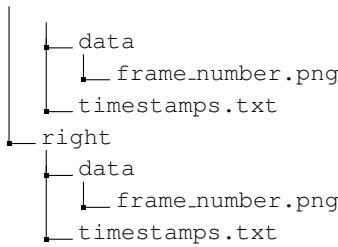


Figure 11. data



4 Sensor Synchronization and Calibration

Sensor spatio-temporal fusion requires accurate time synchronization and spatial calibration of multiple sensors. All sensors in the system are rigidly held in place by aluminium profile brackets, 3D prints and carbon fibre sheets. An FPGA-based hard trigger circuit and NTP synchronization network are used for time synchronization of multiple sensors and data acquisition computers.

4.1 Time Synchronization

The time synchronization module performs the time synchronization of the camera, lidar, IMU and the main control computing module. The first level of the time synchronization module is the GNSS receiving module, which obtains the UTC true time data with nanosecond precision through the satellite. The logic parallel circuit processes the timing information, converts the GNSS signal into PPS signal and NEMA signal, and the lidar is connected to the two signals for time synchronization. At the same time, the PPS signal is connected to the IMU module and the frequency divider module. The PPS signal makes the inertia data timestamp of the IMU consistent with the true value of UTC time; the PPS signal connected to the frequency divider module is based on the The camera needs to be triggered at the frame rate. The camera trigger signal is aligned with the PPS signal at the edge of the

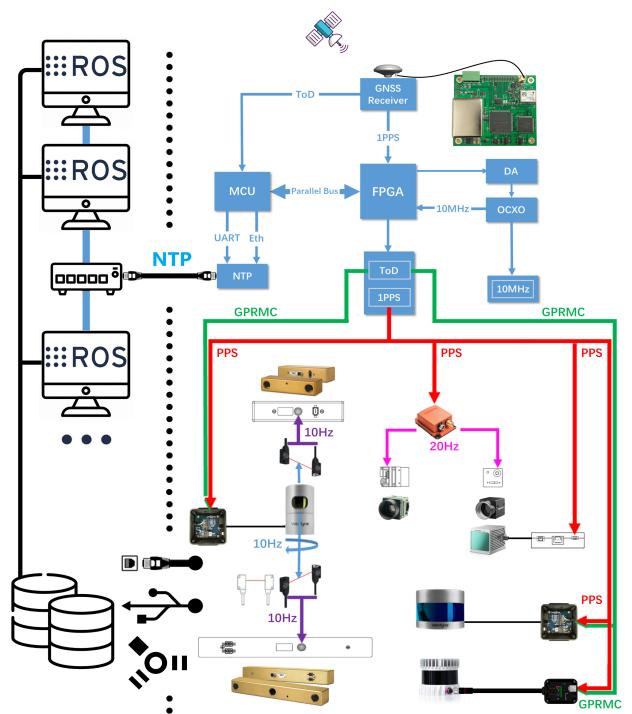


Figure 12. time

whole second, and the delay between the two signals is within a few tens of ns. Therefore, the camera exposure image time is synchronized with the IMU data acquisition time; The camera time is time-synchronized during device manufacture, so the trigger signal triggers all cameras on the serial trigger line simultaneously. The computing master accepts the NTP network data packets converted and sent by

the FPGA, and performs time synchronization through the NTP protocol.

For the horizontal lidar and the stereo camera on the axis, a photoelectric trigger sensor is designed. When the lidar rotates to coincide with the camera's field of view, the excitation switch is turned on and the camera is exposed to collect images to ensure the spatiotemporal synchronization of point clouds and images.

4.2 Mono and Stereo Cameras Calibration

In order to make full use of the metric information of 2D images for 3D tasks, we calibrate the internal parameters of each camera and the external parameters between stereo cameras. The calibration approach we use is proposed by Zhang[11]. A known prior size checkerboard is placed at different distances and attitudes relative to the cameras, the cameras collect images at a fixed frame rate as calibration data and the cameras in same stereo pair are triggered synchronously. Camera parameters are provided in OpenCV format, which are stored in the `camera_name.yaml` calibration file.

The camera parameters are notated as:

- *image_size* $\in \mathbb{N}^2$
- *camera_matrix* $\in \mathbb{R}^{3 \times 3}$
- *distortion_coefficients* $\in \mathbb{R}^5$
- *rectification_matrix* $\in \mathbb{R}^{3 \times 3}$
- *projection_matrix* $\in \mathbb{R}^{3 \times 4}$

Here, the *distortion_coefficients* vector is used to rectify the tangential and radial distortion of images, using pinhole camera distortion model. The *rectification_matrix* is only applicable to stereo cameras, which is used to align the epipolar lines between two stereo images for 3D stereo vision geometry calculation. It is identity matrix for monocular cameras.

The camera projection matrix is used to project objects in the 3D world to the camera 2D image pixels:

$$P_{proj} = \begin{bmatrix} f_x & 0 & c_x & T_x \\ 0 & f_y & c_y & T_y \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (1)$$

The left 3×3 portion is the intrinsic *camera_matrix* for the rectified image. The fourth column $[T_x \ T_y \ 0]^T$ is to translate the optical center of the second camera to the position in the frame of the first camera. For monocular cameras, $T_x = T_y = 0$. The average calibration error of monocular cameras is about 0.08 pixel and average calibration error of stereo cameras is about 0.1 pixel.

4.3 Visual Inertial Calibration

The fusion of visual and inertial sensors will greatly improve the robustness of the visual based SLAM system. The camera provides high resolution external measurements of the environment, while the IMU measures the internal ego-motion of the sensor platform.

The first task is to calibrate the internal parameters of the IMU. The IMU sensor will drift over time, therefore it is necessary to add an error term into the motion model to correct the IMU raw data based on IMU noise model. We fixed the IMU still on the anti-shake optical table for 4

hours and recorded the data. The toolbox `imu_utils` is used for calibration.

IMU internal parameters in the corresponding yaml file are as followed:

- σ_g — gyroscope white noise
- σ_a — accelerometer white noise
- σ_{bg} — gyroscope bias instability
- σ_{ba} — accelerometer bias instability
- rotation matrix: $R_{cam}^{imu} \in \mathbb{R}^{3 \times 3}$
- translation vector: $t_{cam}^{imu} \in \mathbb{R}^{1 \times 3}$

$$T_{cam}^{imu} = \begin{bmatrix} R_{cam}^{imu} & t_{cam}^{imu} \\ 0_{1 \times 3} & 1 \end{bmatrix} \quad (2)$$

4.4 LiDAR Inertial Calibration

- rotation matrix: $R_{lidar}^{imu} \in \mathbb{R}^{3 \times 3}$
- translation vector: $t_{lidar}^{imu} \in \mathbb{R}^{1 \times 3}$

$$T_{lidar}^{imu} = \begin{bmatrix} R_{lidar}^{imu} & t_{lidar}^{imu} \\ 0_{1 \times 3} & 1 \end{bmatrix} \quad (3)$$

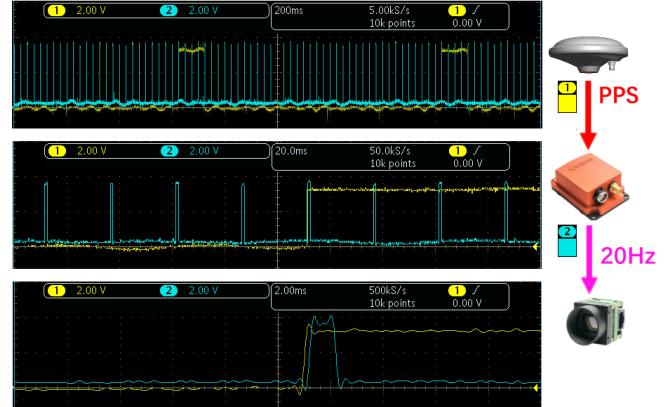


Figure 13. Time synchronization signals in visual-inertial system. The time resolution of the abscissa from top to bottom is 200ms, 20ms and 2ms, from a Tektronix MDO3024 oscilloscope. Yellow: 1Hz PPS signal; Blue: 20Hz camera trigger signal.

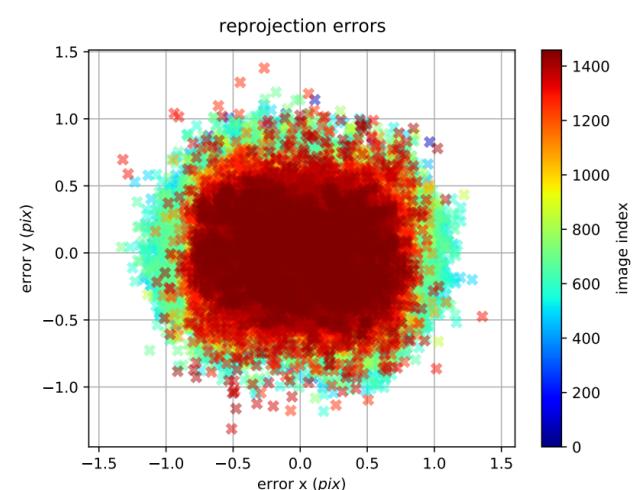


Figure 14. kalibr error

4.5 LiDAR Camera Calibration

Velo2cam Velodyne Lidar to cameras:

- rotation matrix: $R_{velo}^{cam} \in \mathbb{R}^{3 \times 3}$
- translation vector: $t_{velo}^{cam} \in \mathbb{R}^{1 \times 3}$

$$T_{velo}^{cam} = \begin{bmatrix} R_{velo}^{cam} & t_{velo}^{cam} \\ 0_{1 \times 3} & 1 \end{bmatrix} \quad (4)$$

Step1: World/Lidar coordinate system to camera coordinate system:

$$\begin{bmatrix} X_{cam} \\ Y_{cam} \\ Z_{cam} \\ 1 \end{bmatrix} = \begin{bmatrix} R_{velo}^{cam} & t_{velo}^{cam} \\ 0_{1 \times 3} & 1 \end{bmatrix} \begin{bmatrix} X_{velo} \\ Y_{velo} \\ Z_{velo} \\ 1 \end{bmatrix} \quad (5)$$

Step2: Camera coordinate system to image coordinate system

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \frac{1}{Z_{cam}} \begin{bmatrix} f_x & 0 & c_x & T_x \\ 0 & f_y & c_y & T_y \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_{cam} \\ Y_{cam} \\ Z_{cam} \\ 1 \end{bmatrix} \quad (6)$$



Figure 15. yes

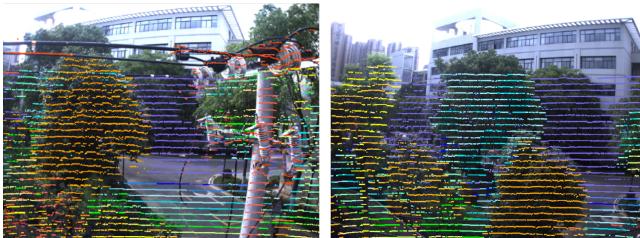


Figure 16. yes

4.6 Ground Truth Alignment

5 Evaluation

Please be aware that the use of this L^AT_EX 2_E class file is governed by the following conditions.

5.1 Copyright

Copyright © 2022 SAGE Publications Ltd, 1 Oliver's Yard, 55 City Road, London, EC1Y 1SP, UK. All rights reserved.

5.2 Rules of use

This class file is made available for use by authors who wish to prepare an article for publication in a SAGE Publications journal. The user may not exploit any part of the class file commercially.

This class file is provided on an *as is* basis, without warranties of any kind, either express or implied, including but not limited to warranties of title, or implied warranties of merchantability or fitness for a particular purpose. There will be no duty on the author[s] of the software or SAGE Publications Ltd to correct any errors or defects in the software. Any statutory rights you may have remain unaffected by your acceptance of these rules of use.

6 Summary and future work

Acknowledgements

This class file was developed by Sunrise Setting Ltd, Brixham, Devon, UK.

Website: <http://www.sunrise-setting.co.uk>

Notes

<https://versalift.com/>.

1.https://ustc-robot.github.io/pdf/MTi_usermanual.pdf.

2.https://ustc-robot.github.io/pdf/MANUAL_USERS_HDL32E.pdf.

3.https://ustc-robot.github.io/pdf/MANUAL_USERS_HDL32E.pdf.

https://github.com/gaowenliang imu_utils.

References

- [1] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, “Vision meets robotics: The kitti dataset,” *International Journal of Robotics Research (IJRR)*, 2013.
- [2] T.-M. Nguyen, S. Yuan, M. Cao, Y. Lyu, T. H. Nguyen, and L. Xie, “Ntu viral: A visual-inertial-ranging-lidar dataset, from an aerial vehicle viewpoint,” *International Journal of Robotics Research*, 2021.
- [3] N. Carlevaris-Bianco, A. K. Ushani, and R. M. Eustice, “University of michigan north campus long-term vision and lidar dataset,” *The International Journal of Robotics Research*, vol. 35, no. 9, pp. 1023–1035, 2016.
- [4] A. S. Huang, M. Antone, E. Olson, L. Fletcher, D. Moore, S. Teller, and J. Leonard, “A high-rate, heterogeneous data set from the darpa urban challenge,” *The International Journal of Robotics Research*, vol. 29, no. 13, pp. 1595–1601, 2010.
- [5] G. Pandey, J. R. McBride, and R. M. Eustice, “Ford campus vision and lidar data set,” *The International Journal of Robotics Research*, vol. 30, no. 13, pp. 1543–1552, 2011.
- [6] 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, 2017.
- [7] D. Barnes, M. Gadd, P. Murcutt, P. Newman, and I. Posner, “The oxford radar robotcar dataset: A radar extension to the oxford robotcar dataset,” *arXiv preprint arXiv: 1909.01300*, 2019. [Online]. Available: <https://arxiv.org/pdf/1909.01300>
- [8] T. Pire, M. Mujica, J. Civiera, 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.
- [9] J. Jeong, Y. Cho, Y.-S. Shin, H. Roh, and A. Kim, “Complex urban dataset with multi-level sensors from highly diverse urban environments,” *The International Journal of Robotics Research*, vol. 38, no. 6, pp. 642–657, 2019.

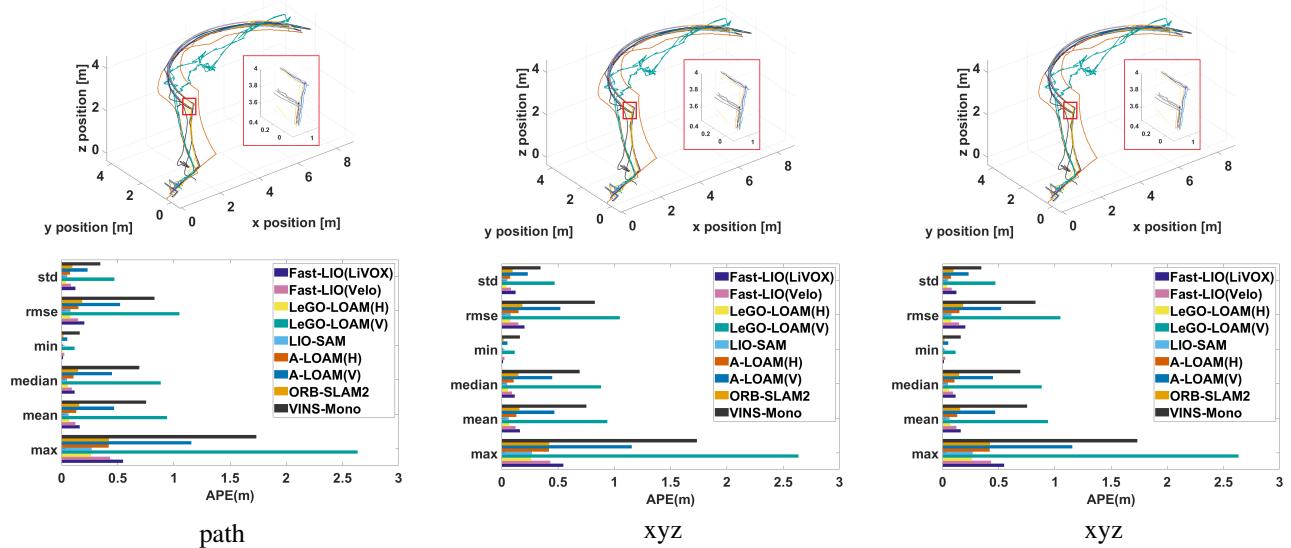


Figure 17. Visual comparisons of original models (SCRN)

Table 3. The choice of options.

Sequence	hf001	hf002	hf003	hf004	hf005	hf006	hf007	hf008	gr001
VINS-Mono	0.472	0.245	0.703	0.813	0.936				
ORB-SLAM2	0.182	0.166	0.136	0.184	0.213				
A-LOAM(H)	0.149	0.152	0.157	0.188	0.168				
A-LOAM(V)	0.520	—	0.422	0.507					
LeGO-LOAM(H)	0.077	0.082	0.165	0.113	0.083				
LeGO-LOAM(V)	1.048	—	0.767	0.479	1.216				
LIO-SAM	0.079	0.080	0.154	0.128	0.173				
Fast-LIO (Velo)	0.147	0.153	0.154	0.193	0.167				
Fast-LIO (LiVOX)		0.201			0.128	0.130			

- [10] S. S. Ahmed, K. F. Hulme, G. Fountas, U. Eker, I. V. Benedyk, S. E. Still, and P. C. Anastasopoulos, “The flying car—challenges and strategies toward future adoption,” *Frontiers in Built Environment*, vol. 6, p. 106, 2020.
- [11] Z. Zhang, “A flexible new technique for camera calibration,” *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.
- [12] M. Grupp, “evo: Python package for the evaluation of odometry and slam.” <https://github.com/MichaelGrupp/evo>, 2017.
- [13] T. Shan, B. Englot, D. Meyers, W. Wang, C. Ratti, and R. Daniela, “Lio-sam: Tightly-coupled lidar inertial odometry via smoothing and mapping,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 5135–5142.
- [14] T. Shan and B. Englot, “Lego-loam: Lightweight and ground-optimized lidar odometry and mapping on variable terrain,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 4758–4765.
- [15] J. Rehder, J. Nikolic, T. Schneider, T. Hinzmann, and R. Siegwart, “Extending kalibr: Calibrating the extrinsics of multiple imus and of individual axes,” in *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2016, pp. 4304–4311.
- [16] J. Beltrán, C. Guindel, A. de la Escalera, and F. García, “Automatic extrinsic calibration method for lidar and camera sensor setups,” *IEEE Transactions on Intelligent Transportation Systems*, 2022.