Open source, open hardware ground truth for Visual Odometry and SLAM applications

Janusz Bedkowski, Grzegorz Kisala, Michal Wlasiuk, and Piotr Pokorski,

Abstract—Ground-truth data is essential for VO (Visual Odometry) and SLAM (Simultaneous Localization and Mapping) quantitative evaluation using e.g. ATE (Absolute Trajectory Error) and RPE (Relative Pose Error). Many open-access data sets provide raw and ground-truth data for benchmark purposes. The issue appears when one would like to validate Visual Odometry and/or SLAM approaches on data captured using the device for which the algorithm is targeted for example mobile phone and disseminate data for other researchers. For this reason, we propose an open source, open hardware ground-truth system that provides an accurate and precise trajectory with a 3D point cloud. It is based on LiDAR Livox Mid-360 with a non-repetitive scanning pattern, on-board Raspberry Pi 4B computer, battery and software for off-line calculations (camera to LiDAR calibration, LiDAR odometry, SLAM, georeferencing). We show how this system can be used for the evaluation of various the state of the art algorithms (Stella SLAM, ORB SLAM3, DSO) in typical indoor monocular VO/SLAM.

Keywords—Visual Odometry, LiDAR odometry, SLAM, georeferencing

I. Introduction

Rapid development of hand-held mobile mapping systems [47] incorporating recent advances in SLAM (Simultaneous Localization and Mapping) [53][57] allows the creation of both accurate and precise 3D point cloud and the trajectory. This trajectory can be considered as ground truth looking from Visual Odometry and SLAM applications [55][35]. An alternative approach could be a ground truth reference system using robust visual encoded targets [29]. The limitation of existing approaches is related either to the high complexity of the mobile mapping system [52] or to the overall cost of the solution. For these reasons, we proposed a novel open-hardware mobile mapping system that is:

- hand-held,
- weight less than 1kg,
- LiDAR range up to 40m,
- affordable (open source, open hardware MIT license),
- indoor and outdoor,
- large scale, long-term data acquisition up to 5 hours of continuous scanning with suggested velocity up to 8km/h,
- with the handler for a smartphone.

The contribution of this paper is also related to open-source software that provides an alternative approach to e.g. g^2o[22], gtsam[21], manif[11], ceres [2]. Our aim was to minimize the effort needed for software installation and interoperability between different operating systems. The contribution is:

- camera to LiDAR calibration,
- LiDAR odometry,
- single session alignment,
- multiple session alignment with georeferencing.

Benchmarking [16] is very important for providing advantages and disadvantages of the current state of the art. A great impact was introducing ATE (Absolute Trajectory Error) and RPE (Relative Pose Error) errors by [43]. Recent Hilti-Oxford Dataset [58] provides an interesting approach for LiDAR and Visual SLAM benchmarking. It has been collected on construction sites as well as at the famous Sheldonian Theatre in Oxford, providing a large range of difficult problems for SLAM.

Ground truth data sources are crucial when looking from quantitative and qualitative benchmarking points of view. The dominant one in literature is Visual Odometry-SLAM Evaluation 2012 KITTI [18] that directs the mainstream of many types of research. The second great example is the Hilti-Oxford dataset [58]: a millimeter-accurate benchmark for SLAM. Most of the recent papers incorporate KITTI or/and Hilti datasets for evaluation purposes. It is evident a rapid growth of more and more sophisticated benchmarks for SLAM applications [40]. Such an approach unfortunately introduces an important gap between benchmark and real-live applications. This observation is supported by this paper showing that it is not straight forward to take state of the art Visual SLAM approach and to perform mobile mapping in e.g. typical office environment. Chosen algorithms: StellaSLAM[44], ORBSLAM3[10] and DSO[14][51] have great performance on widely used and accepted by academia benchmarks [16], [9]. Most of these datasets contain data derived from professional cameras such as high fps global shutter and IMU. Such data is typically tested in the mobile robotics domain. Unfortunately, research in the area of smartphones requires a more flexible approach since new devices appear each year. The SLAM technology transition is rather challenging since most of the approaches are not prepared for it. Especially, large resolution cameras with rolling shutters are still an open research topic [37][38][24].

A rationale behind our research is the fact that we are demonstrating the bias of potential use of existing benchmarks. In our case, state-of-the-art algorithms do not solve our real-world task. We have chosen most easiest cases - mapping of small indoor scenes. It means that the proposed methodology can have a positive impact on the research community since the required effort for ground truth data collection can be drastically decreased. The contribution of this paper as is follows:

- First, we analyzed existing ground truth systems.
- Second, we introduced open hardware and methodology

for 3D data collection.
- Third, we show ground truth accuracy assessment.
- Fourth, we show visual SLAM accuracy assessment using state-of-the-art StellaSLAM, ORBSLAM3, and DSO.
- Finally, we conclude the paper with suggestions and discuss future direction.

II. Ground truth systems

Rapid growth of indoor ground truth data interest is evident [56], since it provides useful data for qualitative and quantitative measures for SLAM applications. Cost-effective camera-based ground truth for indoor localization is very helpful [4] for performing preliminary tests. For outdoor a typical approaches are a global positioning system [42] and total stations [48]. Obtaining ground truth is a rather sophisticated procedure than easy technology to use. Moreover, deployment of ground truth technology is not always possible, especially in extreme environments [1], e.g. an analysis of SLAM-based LiDAR data quality metrics for geotechnical underground monitoring [17] is very important from safety point of view. It is related with the work on affordable low-cost handheld LiDAR-based SLAM systems [46]. Further important limitation of the ground truth data source is its accuracy and precision [54]. In most of SLAM applications [13] it seams that we do not need a millimeter accuracy to conduct qualitative and quantitative evaluation. Hence, a centimeter level accuracy is sufficient and it can be obtained with many existing technologies such as Terrestrial Laser Scanner [35],[12]. Such accuracy is also provided by GPS, GNSS receivers with Real Time Kinematics (RTK) [7] which is the most popular technique to collect ground truth data in open-sky outdoor environments [8]. We distinguish following satellite positioning systems, 1: Global Navigation Satellite Systems (GNSS), GPS (Global Positioning System, United States), GLONASS (GLObal Navigation Satellite System, Russian Federation), Galileo (European Global Navigation Satellite Systems Agency (GSA)), BeiDou (approximately translated to “Northern Dipper”, People’s Republic of China), IRNSS (Indian Regional Navigation Satellite System, India) and QZSS (Quasi-Zenith Satellite System, Japan). It is also possible to process such data with PPP (Precise Point Positioning) [45]. PPP relies on carrier-phase measurements as the primary observable to model or estimate effects for centimetre-level resolution. We can incorporate SLAM techniques to improve this data even assuming continent-scale [5], thus optimizing multiple trajectories decreases overall data uncertainty and increases the accuracy.

Mobile mapping systems incorporates LiDAR technology for 3D measurements [26]. Instead of LiDAR multi beam technology non-repetitive scanning patterns [32] are recently investigated since this functionality provides 99% coverage of the surrounding environment even without motion [28]. Affordable non-repetitive scanning pattern mobile mapping systems are currently state-of-the-art for building low-cost ground truth systems. It might be possible that solid-state LiDARs [27] will provide even more affordable applications. There are plenty of commercial ground truth systems in the market since benchmarking is crustal both from academic and industrial points of view. Table I shows some of the ground truth systems that are extensively used by many researchers [9], [39], [33], [35], [34]. It can be seen that the cost is multiple times larger than our approach, but those systems provide much better accuracy and precision.

TABLE I: Commercial ground truth systems

Method	Hardware	Approximate Cost	Reference
Spatio-Temporal Alignment	Leica MS50 laser tracker and scanner, Vicon 6D motion capture system	19,500 USD https://califfsurveying.com/product/leica-nova-ms50	[9]
Optical Motion Capture	16 infrared OptiTrack Flex13 cameras	16 x 1,099 USD = 17,584 USD https://optitrack.com/cameras/flex-13/buy.html	[39]
Optical Motion Capture	18 cameras OptiTrack Prime 41	18 x 6,499 USD = 116,982 USD https://optitrack.com/cameras/prime-41/	[33]
Iterative Closest Point on Point cloud	Laser scanner Leica BLK360	25,900 USD https://shop.leica-geosystems.com/leica-blk/blk360/buy	[35]
Time of Flight / Multiple Frequency Phase-shift	TOPCON GT1205	37,000 USD https://surveyingsupplies.com/products/gt-series-total-station-kits	[34]
GPS	South Galaxy G9 Rover Set	€4,630 https://globalgpssystems.com/gps-receivers/	
GPS	South Galaxy G9	€2,995 https://globalgpssystems.com/gps-receivers/	
GPS	EFIX F8 Rover set	€6,495 https://globalgpssystems.com/gps-receivers/	
ToF	laser 3D scanner	65,000 USD https://www.artec3d.com/portable-3d-scanners/	
ToF	Our	1,000 USD url: anonymized due to double blind review	

III. Open-hardware for 3D data collection

Figure 1 shows the open-hardware hand-held mobile mapping system. Technical details are available at (url: anonymized due to double blind review). The open source project for off-line 3D data processing is available at (url: anonymized due to double blind review). It is compatible with ROS thanks to (url: anonymized due to double blind). The system is composed of LiDAR Livox Mid-360. It is capable of collecting data on USB flash memory. It is equipped with an onboard RaspberryPi4B computer. It can work for more than 5 hours and the weight is around 1kg.

A. Methodology

Ground truth data processing is composed of four modules:
- Camera to LiDAR calibration,
- LiDAR odometry,
- single session refinement,
- multiple sessions refinement with georeferencing.

Camera to LiDAR calibration is implemented based on fundamental paradigm in computer vision [19] - re-projection error. LiDAR 3D point - image pixel pairs form optimization problem that calculates extrinsic parameters ($[R,t]_{LiDAR \leftarrow camera}$) of the system. LiDAR odometry is composed of highly coupled multiple view normal distributions transform with pose graph SLAM incorporated for preserving motion model derived from IMU processed with Madgwick Orientation Filter [30], [23]. Each consecutive batch of 20 poses is processed within the assumption of the sliding window. This approach differs from classic pair-wise matching with pose graph SLAM [49]. Thus, in our approach relative pose constraints (20 consecutive poses) are highly coupled with multi-view normal distributions transformed as a single optimization routine. Once the initial trajectory is calculated it is possible to perform consistency procedure that makes the trajectory smooth.

B. Camera LiDAR synchronization

Both devices used in the setup have different clocks which requires synchronization between them. There are at least two ways in which we can address synchronization problem:

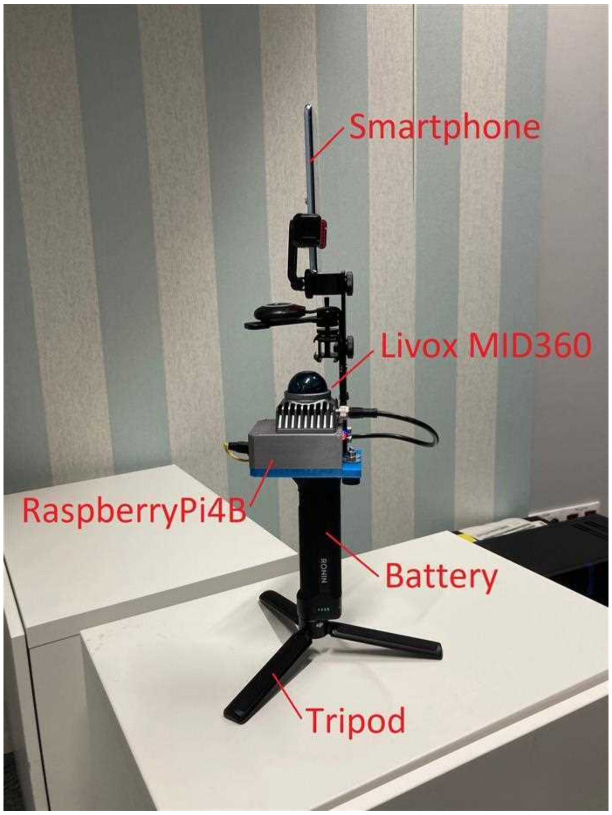


Fig. 1: Open-hardware for 3D data collection composed of LiDAR: Livox Mid-360, on board computer: RaspberryPi4B, battery and tripod.

1: Synchronize system clocks of both devices and use those system clocks as basis for time stamping sensor data. This is the simplest solution which can use existing clock synchronization methods and protocols, such as NTP (network time protocol) and PTP (precision time protocol). The former is available out-of-the-box in many platforms, such as Android itself (in fact Android uses NTP to synchronize system clock over internet). On Unix-like operating systems NTP daemons are also easily obtainable and configurable. The precision of NTP synchronization varies and is dependant on network architecture. Generally, one can expect synchronization accuracy below 100 ms (less in small LAN networks) (https://www.ion.org/publications/abstract.cfm?articleID=14186, https://www.ntp.org/ntpfaq/NTP-s-algo/#5131-how-accurate-will-my-clock-be). PTP is a much more accurate solution, easily guaranteeing microsecond accuracy, however it requires more effort in the setup and puts more requirements on network adapters (such as software or hardware timestamping). As this protocol is not available out-of-the-box on Android, we didn't consider it further.

2: Synchronize data timestamps based on IMU data. Most API used for fetching data (such as Android Camera2 or Motion sensors API) do not provide timestamps associated with system clock, but rather some internal monotonic clock (e.g. time since booting up the system). When one uses raw timestamps provided by the device it is pointless to use time synchronization protocols, as those affect system clock. In such cases direct data synchronization is required. In our experiments we found that best results are achieved through IMU data synchronization (especially via acceleration). This method can achieve decent results due to high frequency and low latency of IMU data. It is however hard to automate and requires good IMU data such that synchronization based on movement patterns is possible.

In our setup we opted to use NTP-based synchronization, which allowed us to achieve synchronization accuracy below 50 ms. With average speeds not exceeding 0.6 m/s this means that inaccuracy due to synchronization is below 3 cm, which is enough for datasets presented here. NTP-based synchronization was chosen as it was the simplest method to setup and automate.

C. Camera to LiDAR calibration

We use built-in Camera2 API to get intrinsic parameters for the camera. It is a deliberate choice over manual calibration. Our goal is to test the algorithms on out-of-the-box devices, such as in commercial applications, where user is not supposed to calibrate the device. Relative camera to LiDAR pose estimation is addressed by solving two subproblems: First a feature matching problem that seeks to establish putative 2D-3D correspondences, and then a Perspective-n-Point problem that minimizes, w.r.t. the camera pose, the sum of so-called Reprojection Errors (RE). Feature matching problem is solved by manual procedure, where end-user finds at least 5 pairs 2D (camera image) - 3D (3D point cloud) correspondences. These pairs form the Perspective-n-Point problem [50] that is minimized with Gauss-Newton optimization routine.

D. Normal Distributions Transform

Normal Distributions Transform [31] is an alternative technique to Iterative Closest Point [6],[20] for point cloud data registration and it is available in a well-known Point Cloud Library [36] open source project. It is limited to the pair-wise matching of two-point clouds, thus a contribution of the proposed research is a novel approach to NDT enabling fusing it with pose graph SLAM. The key element of the NDT is the representation of the data as a set of normal distributions organized in the regular grid over 3D space. These distributions describe the probability of finding a 3D point at a certain position. The advantage of the method is that it gives a smooth representation of the point cloud, with continuous first and second-order derivatives. Thus, standard optimization techniques described in this paper can be applied. Another advantage of NDT over ICP is its much less computational complexity since the consumptive nearest neighbourhood search procedure is not needed. Authors of [3] also elaborate on this advantage. The 3D space decomposition into the regular grid introduces some minor artefacts, but in a presented experiment it is a negligibly small disadvantage. For each bucket from a regular grid containing a sufficient number of measured points, NDT calculates the mean given by the equation (1) and the covariance given by the equation

(2).

$$\boldsymbol{\mu} = \frac{1}{m}\sum_{k=1}^{m} \boldsymbol{P}_k^g \tag{1}$$

$$\boldsymbol{\Sigma} = \frac{1}{m-1}\sum_{k=1}^{m} (\boldsymbol{P}_k^g - \boldsymbol{\mu})(\boldsymbol{P}_k^g - \boldsymbol{\mu})^\intercal \tag{2}$$

The likelihood of having measured point $\boldsymbol{P}_m^g$ is given by equation (3).

$$p(\boldsymbol{P}_m^g) = \frac{1}{(2\Pi)^{\frac{1}{2}}\sqrt{|\boldsymbol{\Sigma}|}} \exp\left(-\frac{(\boldsymbol{P}_m^g - \boldsymbol{\mu})^\intercal \boldsymbol{\Sigma}^{-1} (\boldsymbol{P}_m^g - \boldsymbol{\mu})}{2}\right) \tag{3}$$

Each $p(\boldsymbol{P}_m^g)$ can be seen as an approximation of the local surface within the range of the bucket. It describes the position $\boldsymbol{\mu}$ of the surface as well as its orientation and smoothness given by $\boldsymbol{\Sigma}$. Let $\boldsymbol{\Psi}([\boldsymbol{R}, \boldsymbol{t}]^{3\times4}_{W\leftarrow LiDAR}, \boldsymbol{P}_m^l)$ will be a transformation function of the local measurement point $[\boldsymbol{P}_m^l, 1]^\intercal$ via pose $[\boldsymbol{R}, \boldsymbol{t}]^{3\times4}_{W\leftarrow LiDAR}$ expressed as (4).

$$\boldsymbol{\Psi}([\boldsymbol{R}, \boldsymbol{t}]^{3\times4}_{W\leftarrow LiDAR}, \boldsymbol{P}_m^l) = \boldsymbol{P}_m^g = [\boldsymbol{R}, \boldsymbol{t}]^{3\times4}_{W\leftarrow LiDAR} \begin{bmatrix} \boldsymbol{P}_m^l \\ 1 \end{bmatrix} \tag{4}$$

Thus, the NDT optimization problem is defined as the maximization of the likelihood function given in equation (5).

$$[\boldsymbol{R}, \boldsymbol{t}]^{3\times4,*}_{W\leftarrow LiDAR} = \max_{[\boldsymbol{R}, \boldsymbol{t}]_{W\leftarrow LiDAR}} \prod_{k=1}^{N} p(\boldsymbol{\Psi}([\boldsymbol{R}, \boldsymbol{t}]^{3\times4}_{W\leftarrow LiDAR}, \boldsymbol{P}_m^l)) \tag{5}$$

Furthermore, the optimization problem is equivalent to the minimization of the negative log-likelihood given in equation (6).

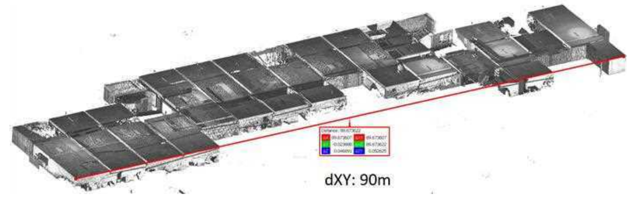
$$[\boldsymbol{R}, \boldsymbol{t}]^{3\times4,*}_{W\leftarrow LiDAR} = \min_{[\boldsymbol{R}, \boldsymbol{t}]^{3\times4}_{W\leftarrow LiDAR}} -\sum_{k=1}^{N} \log\left(p(\boldsymbol{\Psi}([\boldsymbol{R}, \boldsymbol{t}]^{3\times4}_{W\leftarrow LiDAR}, \boldsymbol{P}_m^l))\right) \tag{6}$$

NDT implementation similar to ICP uses using point-to-point observation equation. The only difference is that an information matrix $\boldsymbol{\Omega}$ is calculated as an inverse of the covariance matrix from equation (2).

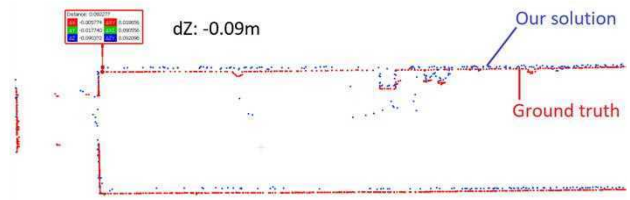
The disadvantage of multi-view NDT is the fact that narrow obstacles such as walls observed from neighbouring rooms can converge to a single entity (the width of the wall should not converge to 0). To discriminate obstacles we remove such observations that correspond to different viewpoints. It means that the flat surface of one room is not converging to this flat surface observed from the neighbouring room. It is implemented as a normal vector geometric check.

E. Single session refinement

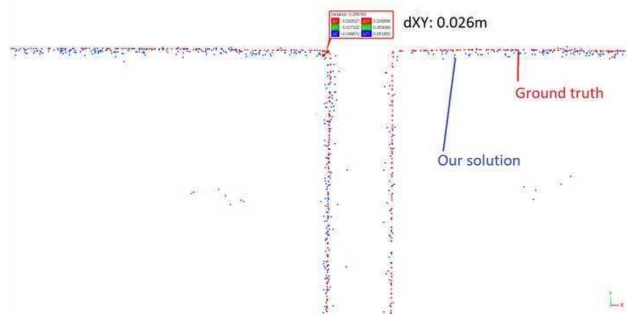
Classic pose graph SLAM[22] is incorporated to optimize manually chosen pairwise matches. Thus, this semi automatic process uses loop closure edges chosen by end user to optimize graph composed of trajectory edges, loop closure edges and motion model. The result is reduced consecutive error of the LiDAR odometry.



(a) Ground truth data set (TLS Z+F IMAGER 5010),



(b) Maximum vertical deviation.



(c) Maximum horizontal deviation.

Fig. 2: Quantitative comparison of our mobile mapping system to ground truth obtained with terrestrial laser scanner survey.

F. Multiple session refinement with georeferencing

Multiple trajectories can be organized into the project. Only one trajectory can be treated as ground truth (obtained e.g. with geodetic survey). Other trajectories will be aligned together and to ground truth based on loop closure edges.

IV. Ground truth accuracy assessment

We performed a quantitative comparison to ground truth data - underground INDOOR scenario 20×90 [m]. It is shown in figure 2. Ground truth data were collected with terrestrial laser scanner survey TLS Z+F IMAGER 5010 that provides point cloud with the milliliter accuracy and precision [41]. We observed that the maximum vertical deviation is less than 10cm and the maximum horizontal deviation is 3cm. This is a satisfactory result that is sufficient, moreover, further investigation on global accuracy is not our main focus since our LiDAR provides 2cm accuracy on a distance 20m (documentation is available here https://www.livoxtech.com/mid-360).

V. Visual SLAM accuracy assessment

For the demonstration of the functionality of the proposed affordable ground truth data system, we performed data collection for several INDOOR scenarios. We evaluated DSO https://github.com/JakobEngel/dso [14][15], OpenVSLAM https://github.com/stella-cv/stella_vslam [44] that is inspired by ORB SLAM https://github.com/UZ-SLAMLab/ORB_SLAM3 [10].

TABLE II: Absolute Trajectory Error

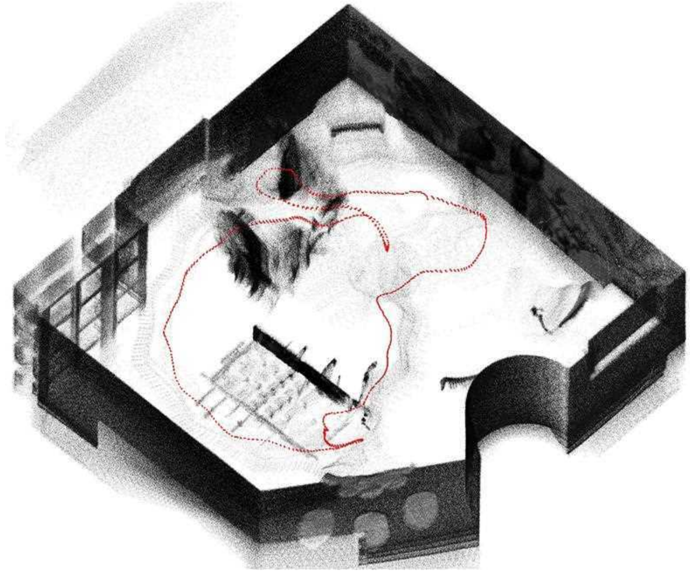
Experiment	LiDAR	DSO	STELLA-VSLAM	ORB-SLAM3
1	reference	0.164169	0.0923437	0.184776
2	reference	0.109849	0.268346	0.290233
3	reference	0.0634539	0.0371597	0.0489425

We selected these three open-source visual SLAM implementations since they can be implemented in smartphones with minimal effort.

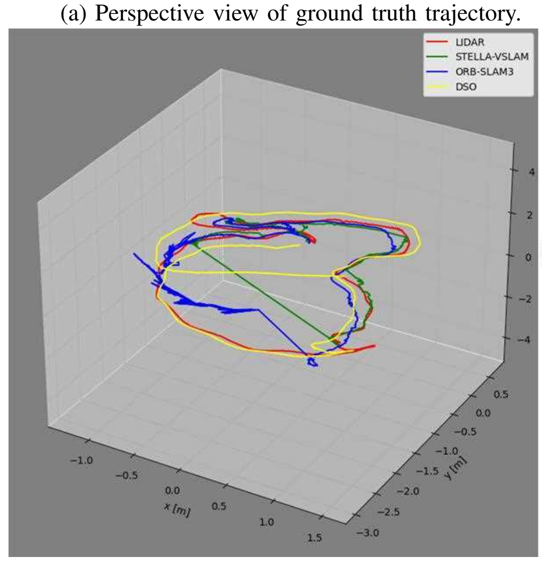
Real-world datasets were recorded using SAMSUNG GALAXY S23 SM-S911B device with ISOCELL GN3 (S5KGN3) electronic rolling shutter sensor locked at 30 FPS as a sequence of corresponding YUV-420-888 images (separate image per YUV component stream) further combined into 8 bit BGR images which after undistortion create a final dataset. Camera characteristics like sensor intrinsic parameters and lens distortion coefficients were retrieved from Android camera API and downscaled from native camera resolution to resolution at which the sequence was recorded and processed using state-of-the-art SLAM frameworks. Before dataset recording started camera was configured so chromatic aberration correction, distortion correction, auto white balance, auto-focus, and color effects were disabled.

Table II collects all ATE (Absolute Trajectory Error) results according to methodology from [43]. This quantitative measure compares trajectory to ground truth. It can be seen that DSO performs best, but there is an issue with this statement since figure 5 shows the failure of all methods. An interesting observation is that OpenVSLAM (current name StallaVSLAM) performs better than its successor ORB SLAM3. So, it is beneficial to refactor existing implementations (Authors of [44] mainly refactored work of [10]). This very simple experiment provides significant observations:

- state-of-the-art visual SLAM algorithms are not out-of-the-box solution for smartphone cameras without IMU,
- ground truth system provides sufficient data for quantitative and qualitative benchmark,
- simple methodology incorporating ATE is not sufficient for correct research statements (it was also observed in [25]).



(a) Perspective view of ground truth trajectory.

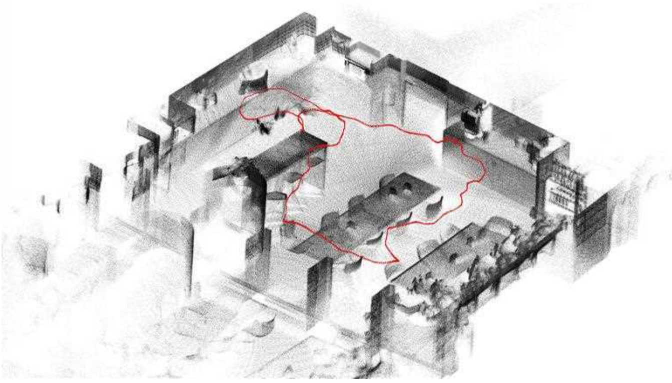


(b) Ground truth, StellaSLAM, ORBSLAM3 and DSO trajectories.

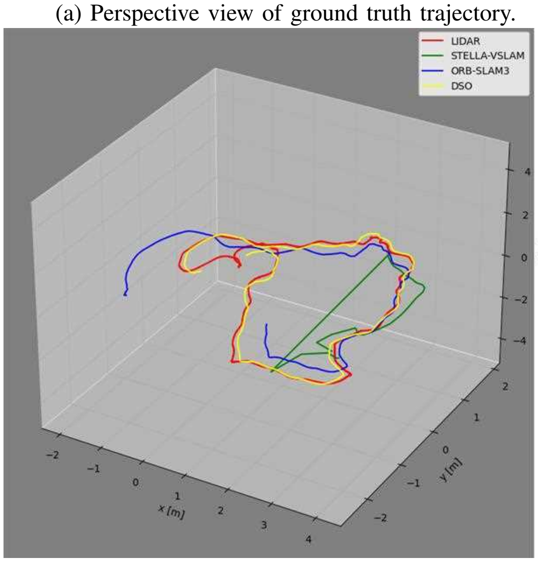
Fig. 3: Experiment 1.

VI. Conclusion

This paper shows how to build a system for ground truth data collection for machine vision, robotics and other mobile mapping applications. It can be used for qualitative and quantitative SLAM evaluation. This research drastically reduces the cost of benchmark data generation, thus many researchers instead of focusing mainly on available datasets will generate new ones. Such an approach removes the existing limitation related to existing benchmarks typically focused on high-end sensors. For this reason, we proposed a novel and affordable ground-truth system that provides an accurate and precise trajectory with a point cloud. It is based on LiDAR Livox Mid-360 with a non-repetitive scanning pattern with affordable IMU, on-board Raspberry Pi 4B computer, battery and software for off-line calculations (LiDAR odometry, SLAM). The software is based on an alternative approach e.g. g2o, GTSAM, manif or Ceres since this lightweight implementation does not require any installation on Linux or Windows. This software is dedicated also to non-programmers. We have shown how this system can be used for the evaluation of various the state of the art algorithms (Stella SLAM, ORB SLAM3, DSO). An open-hardware measurement device specification is available at (url: anonymized due to double blind review) We hope this research will boost machine vision experiments since the proposed solution provides ground truth almost for all scenarios. The project has a rapidly growing community and it addresses the most significant issues with ground truth: cost-effectiveness, scale, ergonomic design, simplicity and interoperability. The accuracy in typical indoor scenarios does not exceed 5cm and the precision 3cm. It is sufficient for qualitative and quantitative SLAM evaluation which was demonstrated in this paper.

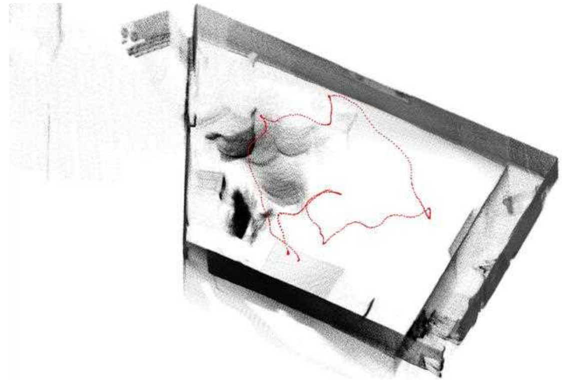


(a) Perspective view of ground truth trajectory.

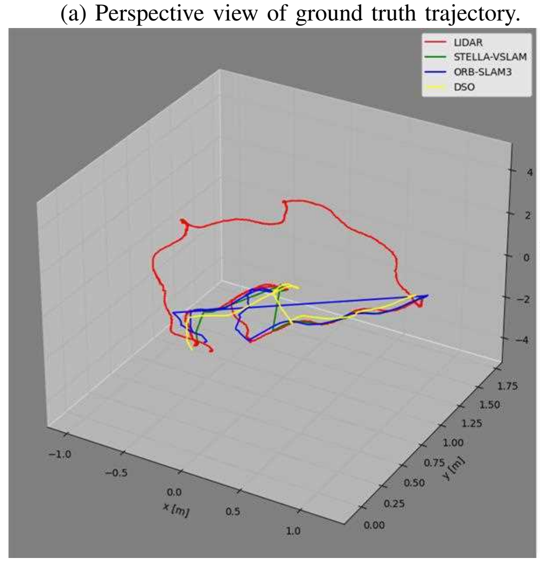


(b) Ground truth, StellaSLAM, ORBSLAM3 and DSO trajectories.

Fig. 4: Experiment 2.



(a) Perspective view of ground truth trajectory.



(b) Ground truth, StellaSLAM, ORBSLAM3 and DSO trajectories.

Fig. 5: Experiment 3.

REFERENCES

[1] Adamek, A., Bedkowski, J., Kaminski, P., Pasek, R., Pelka, M., Zawislak, J.: Underground mining shaft survey – multiple lidar and imu sensors data set. Preprints (October 2023). https://doi.org/10.20944/preprints202310.1388.v1

[2] Agarwal, S., Mierle, K., Team, T.C.S.: Ceres Solver (10 2023)

[3] Bai, C., Xiao, T., Chen, Y., Wang, H., Zhang, F., Gao, X.: Faster-lio: Lightweight tightly coupled lidar-inertial odometry using parallel sparse incremental voxels. IEEE Robotics and Automation Letters **7**(2), 4861–4868 (2022). https://doi.org/10.1109/LRA.2022.3152830

[4] Becker, D., Thiele, F., Sawade, O., Radusch, I.: Cost-effective camera based ground truth for indoor localization. In: 2015 IEEE International Conference on Advanced Intelligent Mechatronics (AIM). pp. 885–890 (2015). https://doi.org/10.1109/AIM.2015.7222650

[5] Bedkowski, J., Nowak, H., Kubiak, B., Studzinski, W., Janeczek, M., Karas, S., Kopaczewski, A., Makosiej, P., Koszuk, J., Pec, M., Miksa, K.: A novel approach to global positioning system accuracy assessment, verified on lidar alignment of one million kilometers at a continent scale, as a foundation for autonomous driving safety analysis. Sensors **21**(17) (2021). https://doi.org/10.3390/s21175691

[6] Besl, P., McKay, N.D.: A method for registration of 3-d shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence **14**(2), 239–256 (1992). https://doi.org/10.1109/34.121791

[7] Blanco, J.L., Moreno, F.A., González, J.: A collection of outdoor robotic datasets with centimeter-accuracy ground truth. Autonomous Robots **27**(4), 327–351 (November 2009). https://doi.org/10.1007/s10514-009-9138-7

[8] Broekman, A., Gräbe, P.J.: A low-cost, mobile real-time kinematic geolocation service for engineering and research applications. HardwareX **10**, e00203 (2021). https://doi.org/https://doi.org/10.1016/j.ohx.2021.e00203

[9] Burri, M., Nikolic, J., Gohl, P., Schneider, T., Rehder, J., Omari, S., Achtelik, M.W., Siegwart, R.: The euroc micro aerial vehicle datasets. The International Journal of Robotics Research (2016). https://doi.org/10.1177/0278364915620033

[10] Campos, C., Elvira, R., Rodríguez, J.J.G., Montiel, J.M.M., Tardós, J.D.: ORB-SLAM3: an accurate open-source library for visual, visual-inertial and multi-map SLAM. CoRR **abs/2007.11898** (2020)

[11] Deray, J., Solà, J.: Manif: A micro lie theory library for state estimation in robotics applications. Journal of Open Source Software **5**(46), 1371 (2020). https://doi.org/10.21105/joss.01371, https://doi.org/10.21105/joss.01371

[12] Dong, Z., Yang, B., Liang, F., Huang, R., Scherer, S.: Hierarchical registration of unordered tls point clouds based on binary shape context descriptor. ISPRS Journal of Photogrammetry and Remote Sensing **144**, 61–79 (2018). https://doi.org/https://doi.org/10.1016/j.isprsjprs.2018.06.018

[13] Dongjae, L., Minwoo, J., Wooseong, Y., Ayoung, K.: Lidar odometry survey: recent advancements and remaining challenges. Intel Serv Robotics **17**, 95—-118 (2024)

[14] Engel, J., Koltun, V., Cremers, D.: Direct sparse odometry. In: arXiv:1607.02565 (July 2016)

[15] Engel, J., Koltun, V., Cremers, D.: Direct sparse odometry. IEEE Transactions on Pattern Analysis and Machine Intelligence (mar 2018)

[16] Engel, J., Usenko, V., Cremers, D.: A photometrically calibrated benchmark for monocular visual odometry. In: arXiv:1607.02555 (July 2016)

[17] Fahle, L., Holley, E.A., Walton, G., Petruska, A.J., Brune, J.F.: Analysis of slam-based lidar data quality metrics for geotechnical underground monitoring. Mining, Metallurgy & Exploration **39**, 1939 – 1960 (2022)

[18] Geiger, A., Lenz, P., Urtasun, R.: Are we ready for autonomous driving? the kitti vision benchmark suite. In: Conference on Computer Vision and Pattern Recognition (CVPR) (2012)

[19] Hartley, R.I., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press, ISBN: 0521540518, second edn. (2004)

[20] He, Y., Liang, B., Yang, J., Li, S., He, J.: An iterative closest points algorithm for registration of 3d laser scanner point clouds with geometric features. Sensors **17**(8) (2017). https://doi.org/10.3390/s17081862

[21] Kaess, M.: Gtsam library (July 2015)

[22] Kummerle, R., Grisetti, G., Strasdat, H., Konolige, K., Burgard, W.: G2o: A general framework for graph optimization. In: ICRA. pp. 3607–3613. IEEE (2011)

[23] Kuo, C.T., Lin, J.J., Jen, K.K., Hsu, W.L., Wang, F.C., Tsao, T.C., Yen, J.Y.: Human posture transition-time detection based upon inertial measurement unit and long short-term memory neural networks. Biomimetics **8**(6) (2023). https://doi.org/10.3390/biomimetics8060471

[24] Lang, X., Lv, J., Huang, J., Ma, Y., Liu, Y., Zuo, X.: Ctrl-vio: Continuous-time visual-inertial odometry for rolling shutter cameras (2022)

[25] Lee, S.H., Civera, J.: What's wrong with the absolute trajectory error? (2023)

[26] Li, A., Liu, X., Sun, J., Lu, Z.: Risley-prism-based multi-beam scanning lidar for high-resolution three-dimensional imaging. Optics and Lasers in Engineering **150**, 106836 (2022). https://doi.org/https://doi.org/10.1016/j.optlaseng.2021.106836

[27] Li, N., Ho, C.P., Xue, J., Lim, L.W., Chen, G., Fu, Y.H., Lee, L.Y.T.: A progress review on solid-state lidar and nanophotonics-based lidar sensors. Laser & Photonics Reviews **16**(11), 2100511 (2022). https://doi.org/https://doi.org/10.1002/lpor.202100511

[28] Li, Q., Yu, X., Queralta, J.P., Westerlund, T.: Multi-modal lidar dataset for benchmarking general-purpose localization and mapping algorithms (2022)

[29] Liao, X., Chen, R., Li, M., Guo, B., Niu, X., Zhang, W.: Design of a smartphone indoor positioning dynamic ground truth reference system using robust visual encoded targets. Sensors **19**(5) (2019). https://doi.org/10.3390/s19051261

[30] Madgwick, S., Vaidyanathan, R., Harrison, A.: An efficient orientation filter for inertial measurement units (imus) and magnetic angular rate and gravity (marg) sensor arrays. Tech. rep., Department of Mechanical Engineering (April 2010)

[31] Magnusson, M., Lilienthal, A.J., Duckett, T.: Scan registration for autonomous mining vehicles using 3d-ndt. J. Field Robotics **24**(10), 803–827 (2007)

[32] Miao, Z., He, B., Xie, W., Zhao, W., Huang, X., Bai, J., Hong, X.: Coarse-to-fine hybrid 3d mapping system with co-calibrated omnidirectional camera and non-repetitive lidar (2023)

[33] Ogiso, S., Mizutani, K., Wakatsuki, N., Ebihara, T.: Robust indoor localization in a reverberant environment using microphone pairs and asynchronous acoustic beacons. IEEE Access **7**, 123116–123127 (2019). https://doi.org/10.1109/ACCESS.2019.2937792

[34] Peng, H., Chen, P., Liu, R., Chen, L.: Spatiotemporal information conversion machine for time-series forecasting. Fundamental Research (Dec 2022). https://doi.org/10.1016/j.fmre.2022.12.009

[35] Ramezani, M., Wang, Y., Camurri, M., Wisth, D., Mattamala, M., Fallon, M.: The newer college dataset: Handheld lidar, inertial and vision with ground truth. In: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE (Oct 2020). https://doi.org/10.1109/iros45743.2020.9340849

[36] Rusu, R., Cousins, S.: 3d is here: Point cloud library (pcl). In: 2011 IEEE International Conference on Robotics and Automation (ICRA). pp. 1–4 (May 2011). https://doi.org/10.1109/ICRA.2011.5980567

[37] Schubert, D., Demmel, N., Usenko, V., Stueckler, J., Cremers, D.: Direct sparse odometry with rolling shutter. In: European Conference on Computer Vision (ECCV) (September 2018)

[38] Schubert, D., Demmel, N., Stumberg, L.v., Usenko, V., Cremers, D.: Rolling-shutter modelling for direct visual-inertial odometry. In: 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE (Nov 2019). https://doi.org/10.1109/iros40897.2019.8968539

[39] Schubert, D., Goll, T., Demmel, N., Usenko, V., Stuckler, J., Cremers, D.: The tum vi benchmark for evaluating visual-inertial odometry. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE (Oct 2018). https://doi.org/10.1109/iros.2018.8593419

[40] Sier, H., Qingqing, L., Xianjia, Y., Queralta, J.P., Zou, Z., Westerlund, T.: A benchmark for multi-modal lidar slam with ground truth in gnss-denied environments (2022)

[41] Stenz, U., Hartmann, J., Paffenholz, J.A., Neumann, I.: High-precision 3d object capturing with static and kinematic terrestrial laser scanning in industrial applications—approaches of quality assessment. Remote Sensing **12**(2) (2020). https://doi.org/10.3390/rs12020290

[42] Stopher, P.R., Shen, L., Liu, W., Ahmed, A.: The challenge of obtaining ground truth for gps processing. Transportation Research Procedia **11**, 206–217 (2015). https://doi.org/https://doi.org/10.1016/j.trpro.2015.12.018, transport Survey Methods: Embracing Behavioural and Technological Changes Selected contributions from the 10th International Conference on Transport Survey Methods 16-21 November 2014, Leura, Australia

[43] Sturm, J., Engelhard, N., Endres, F., Burgard, W., Cremers, D.: A benchmark for the evaluation of rgb-d slam systems. In: 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. pp. 573–580 (2012). https://doi.org/10.1109/IROS.2012.6385773

[44] Sumikura, S., Shibuya, M., Sakurada, K.: OpenVSLAM: A Versatile Visual SLAM Framework. In: Proceedings of the 27th ACM International Conference on Multimedia. pp. 2292–2295. MM '19, ACM, New York, NY, USA (2019). https://doi.org/10.1145/3343031.3350539

[45] Teunissen, P.J., Montenbruck, O.: Springer handbook of global navigation satellite systems, vol. 10. Springer (2017)

[46] Trybała, P., Kasza, D., Wajs, J., Remondino, F.: Comparison of low-cost handheld lidar-based slam systems for mapping underground tunnels. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **XLVIII-1/W1-2023**, 517–524 (2023). https://doi.org/10.5194/isprs-archives-XLVIII-1-W1-2023-517-2023

[47] Trybala, P., Kujawa, P., Romanczukiewicz, K., Szrek, A., Remondino, F.: Designing and evaluating a portable lidar-based slam system. In: The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-1/W3-2023 2nd GEOBENCH Workshop on Evaluation and BENCHmarking of Sensors, Systems and GEOspatial Data in Photogrammetry and Remote Sensing, 23–24 October 2023, Krakow, Poland. pp. 191–198 (2023)

[48] Vaidis, M., Giguère, P., Pomerleau, F., Kubelka, V.: Accurate outdoor ground truth based on total stations. In: 2021 18th Conference on Robots and Vision (CRV). pp. 1–8 (2021). https://doi.org/10.1109/CRV52889.2021.00012

[49] Walcott-Bryant, A., Kaess, M., Johannsson, H., Leonard, J.: Dynamic pose graph SLAM: Long-term mapping in low dynamic environments. In: Conf. on Intelligent Robots and Proc. IEEE/RSJ Intl. Systems, IROS. pp. 1871–1878. Vilamoura, Portugal (Oct 2012)

[50] Wang, P., Xu, G., Cheng, Y., Yu, Q.: A simple, robust and fast method for the perspective-n-point problem. Pattern Recognition Letters **108**, 31–37 (2018). https://doi.org/https://doi.org/10.1016/j.patrec.2018.02.028

[51] Wang, R., Schworer, M., Cremers, D.: Stereo dso: Large-scale direct sparse visual odometry with stereo cameras. In: International Conference on Computer Vision (ICCV). Venice, Italy (October 2017)

[52] Wang, Y., Lou, Y., Song, W., Tu, Z.: A tightly-coupled framework for large-scale map construction with multiple non-repetitive scanning lidars. vol. 22, pp. 3626–3636 (2022). https://doi.org/10.1109/JSEN.2022.3142041

[53] Xu, W., Cai, Y., He, D., Lin, J., Zhang, F.: Fast-lio2: Fast direct lidar-inertial odometry (2021)

[54] Yan, D., Wang, C., Feng, X., Dong, B.: Validation and Ground Truths, pp. 239–260. Springer International Publishing, Cham (2018)

[55] Yoshida, T., Kaji, K., Ogiso, S., Ichikari, R., Uchiyama, H., Kurata, T., Kawaguchi, N.: A survey of ground truth measurement systems for indoor positioning. Journal of Information Processing **31**, 15–20 (2023)

[56] Yoshida, T., Kaji, K., Ogiso, S., Ichikari, R., Uchiyama, H., Kurata, T., Kawaguchi, N.: A survey of ground truth measurement systems for indoor positioning. J. Inf. Process. **31**, 15–20 (2023). https://doi.org/10.2197/IPSJJIP.31.15

[57] You, B., Zhong, G., Chen, C., Li, J., Ma, E.: A simultaneous localization and mapping system using the iterative error state kalman filter judgment algorithm for global navigation satellite system. Sensors **23**(13) (2023). https://doi.org/10.3390/s23136000

[58] Zhang, L., Helmberger, M., Fu, L.F.T., Wisth, D., Camurri, M., Scaramuzza, D., Fallon, M.: Hilti-oxford dataset: A millimeter-accurate benchmark for simultaneous localization and mapping. IEEE Robotics and Automation Letters **8**(1), 408–415 (2023). https://doi.org/10.1109/LRA.2022.3226077