Hilti SLAM Challenge 2023: Benchmarking Single + Multi-session SLAM across Sensor Constellations in Construction

Ashish Devadas Nair[1], Julien Kindle[1], Plamen Levchev[1], and Davide Scaramuzza[2]

***Abstract*— Simultaneous Localization and Mapping systems are the key enabler for positioning in both handheld and robotic applications. The Hilti SLAM Challenges organized over the past years have been successful at benchmarking some of the world's best SLAM Systems with high accuracy. However, more capabilities of these systems are yet to be explored, such as platform agnosticism across varying sensor suites and multi-session SLAM. There exists no dataset plus benchmark combination publicly available, which considers these factors combined. The Hilti SLAM Challenge 2023 Dataset and Benchmark addresses this issue. Additionally, we propose a novel system for robot based SLAM benchmarking with lidar-observable fiducials. Results from the challenge show an increase in overall participation, single-session SLAM systems getting increasingly accurate, successfully operating across varying sensor suites, but relatively few participants performing multi-session SLAM.**

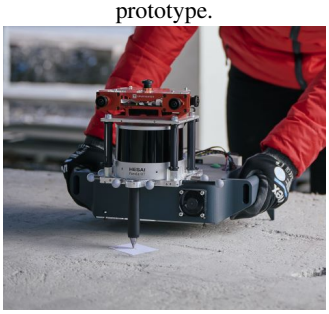
***Index Terms*— SLAM, Localization, Mapping, Sensor Fusion, Dataset, Benchmark, Control Points, Construction.**

I. Introduction

The world of Simultaneous Localization and Mapping techniques has matured steadily as algorithms, sensors, computational capabilities, and simulation systems evolve. Multiple datasets and benchmarks have addressed various aspects of accuracy and robustness. However, there is no publicly accessible live benchmark to evaluate multi-session SLAM.

The Hilti SLAM Challenge 2022 [1] proposed a system where SLAM trajectories were compared with surveyed points in a scene, where during data collection for SLAM, a human operator physically placed a handheld device at survey points. There exists a known rigid body transform between the sensors' reference frame and the spike that made physical contact with the survey point, as seen in Figure 1b. However, this approach of physical contact is not extensible for mobile robot-based SLAM evaluation, since it may not have human-guided fine positioning of a spike. A solution is needed for this. Furthermore, most SLAM benchmarks today use the IMU reference frame to evaluate trajectories, hence the extrinsic transformation of this estimated survey point from the measuring sensor's reference frame to the IMU reference frame needs to be accurate, for accurate Absolute Trajectory Error (ATE) calculation [2], [3]. Allowing SLAM benchmark submissions to include custom extrinsics in ATE evaluation would address effects of any inaccuracies introduced by the calibration data provided with a SLAM dataset.

(a) Trailblazer - drilling robot prototype.



(b) Phasma - Hand-held scanner prototype.

(c) Trimble X7 - Terrestrial Laser Scanner, used for Ground Control Point (GCP) extraction.

Fig. 1: Our SLAM data acquisition devices (a,b); and Ground Truth extraction approach (c).

We also draw into context the field of construction, where SLAM-based systems may need to expand on a previous map, and possibly collaboratively work with another device or robot. Given these needs, we designed the Hilti SLAM Challenge 2023 to deliver the following:

- A Construction Robot platform (Figure 1a) with a sensor suite as described in section III. This is in addition to the handheld Phasma (Figure 1b) prototype used in our previous SLAM challenge [1].
- A lidar-observable fiducial marker and its center position estimation algorithm. A survey point on the ground with such a fiducial is here referred to as a Ground Control Point (GCP), as seen in Figure 1c.
- A multi-device, single and multi-session SLAM Dataset comprising of three locations.
- Evaluation System + Benchmark that includes options for single- and multi-session SLAM evaluation.
- The option for participants to include their own calibration files in their SLAM trajectory submissions.

[1]Corporate Research and Technology, Hilti Group, Liechtenstein. <givenname>.<lastname>@hilti.com

[2]Robotics and Perception Group, University of Zurich, Switzerland. sdavide@ifi.uzh.ch

II. Related Work

Our previous Hilti SLAM Challenge iterations from 2021 [4] and 2022 [1] have set the bar high in pushing the limits of SLAM in construction. Both consisted of a human-operated Lidar, inertial, and visual sensor setup, with evaluations of ATE. The 2022 Challenge made use of a Terrestrial Laser Scanner (here referred to as TLS) to extract sub-centimeter sparse ground-truth (here referred to as GT) - a methodology we chose to adopt.

While addressing SLAM benchmarking across multiple platforms and sessions, the closest fit is the ICCV SLAM Challenge 2023 (October 2023), which encompassed the real-world SubT-MRS [5] and the simulated TartanAir [6] datasets to provide dense ATE and Relative Pose Error RPE evaluations. However, its ground truth trajectory accuracy used for evaluation in real world cases is in the +/- 10cm range [5] despite a TLS based GT map, due to lidar frame scan to map alignment error. It also does not evaluate multi-session SLAM.

In the construction context, it is worth mentioning ConSLAM [7], [8] and ConPR [9]. However, for most construction-related applications consistently high-accuracy positioning across various conditions is the key driver of better quality control and human-equivalent task performance for robots, emphasizing the need for high accuracy GT that the non-Hilti datasets fail to provide.

III. Hardware

In contrast to past SLAM Challenges [1], [4], this year's datasets include trajectories from our Trailblazer robot prototype (Figure 1a), exhibiting unique behaviors such as pronounced vibrations during turns and primarily planar motion. Trailblazer is equipped with a horizontally mounted Robosense Bpearl LiDAR, four Luxonis OAK-D stereo cameras, and an XSens MTi-670 IMU. These components are fixed on a steel plate to ensure stability in sensor positioning, with the LiDAR directly synchronized to the PC and the cameras and IMU via a PTP-to-trigger board for exposure synchronization.

A. Calibration

For IMU calibration, we sampled its static measurements at 200 Hz over 11.5 h, using Allan Variance estimation [10], [11] to assess noise and drift. Stereo cameras underwent intrinsic calibration against a 6x6 april tag grid (60 cm square), with individual datasets processed via kalibr [12]. Extrinsic calibration leveraged movement on a hand pallet track with six 6x6 april tag grids in a hemispherical arrangement, enabling full-motion excitation. A custom variant of MultiCal [13][1], specifically adjusted to fix the extrinsics of stereo camera pairs as determined from intrinsic calibration, produced the default extrinsic calibration.

[1] https://github.com/Hilti-Research/multical

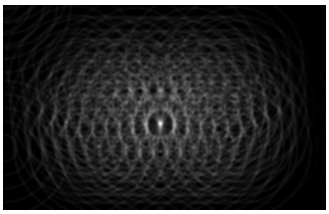
B. Registering of Ground Control Points

Evaluating SLAM trajectory accuracy requires linking the robot's pose to ground truth Terrestrial Laser Scanner (TLS) measurements, necessitating a system to detect Ground Control Points (GCPs). Given the high accuracy of LiDAR SLAM algorithms in previous challenges, a method for LiDAR-based GCPs detection is preferred. One such method is LiDARTag [14] which allows detecting apriltags from LiDAR data. The sparsity of the BPearl LiDAR, however, would require over proportionally large targets. To overcome this, we designed a GCP detector for circular targets, depicted in Figure 2c. This detector projects LiDAR scan regions to a ground plane, identifying significant intensity changes with a 1D Canny Edge detector and recording these in a Hough Space as potential circles matching the GCP. The process culminates in smoothing the Hough Space and pinpointing the most likely 3D position for each GCP.

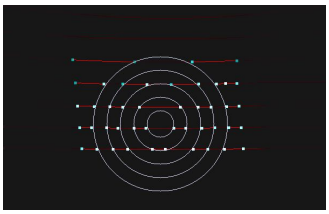
Algorithm 1 LiDAR GCP Detection Algorithm

Input: Set of cropped LiDAR scans $\mathbf{S}_{\text{lidar}}$
Output: Estimated position of GCP center

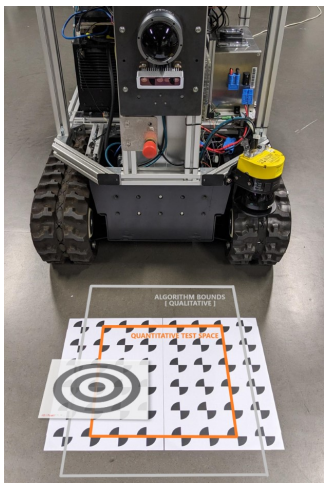
$T_{\text{plane,lidar}} \leftarrow \text{fitPlane}(\mathbf{S}_{\text{lidar}})$ ▷ Fit plane to points
$\mathbf{U}_{\text{plane}} \leftarrow \text{proj}(T_{\text{plane,lidar}}, \mathbf{S}_{\text{lidar}})$ ▷ Project points to plane
$H \leftarrow 0$ ▷ Initialize empty Hough Space
for $U \in \mathbf{U}_{\text{plane}}$ **do** ▷ Iterate over all scans
  $C \leftarrow 0$ ▷ Initialize empty set of corners
  for $r \in U$ **do** ▷ Iterate over all rings
    $C \mathrel{+}= \text{detectEdges}(r)$ ▷ 1D Canny edge
  end for
  $H \mathrel{+}= \text{houghSpace}(C)$ ▷ Add Hough Space votes
end for
$H \leftarrow \text{gaussianBlur}(H)$ ▷ Smoothen Hough Space
$v_{\text{max}} \leftarrow \text{argmax}(H)$ ▷ Find most likely position in plane
return $T^{-1}_{\text{plane,lidar}}\ v_{\text{max}}$ ▷ Transform to 3D Space



(a) Circular Hough Space representation of a GCP detection.



(b) Detected edges of LiDAR intensity, and the resulting most likely fit of the GCP overlayed.



(c) Pre-surveyed Test Grid, annotated with quantitative (orange), qualitative (grey) test bounds, and an overlayed GCP target..

Fig. 2: GCP detector visualizations (a,b) and test setup (c).

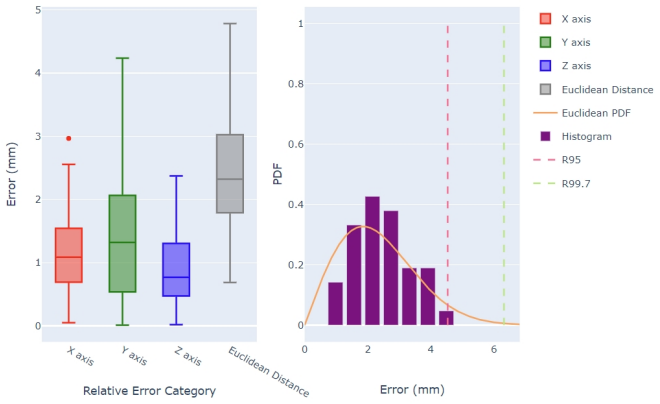


Fig. 3: GCP Estimation Accuracy Evaluation. Left: Relative Error box plots. Right: Rayleigh Distribution of 3DoF Euclidean Distance Error.

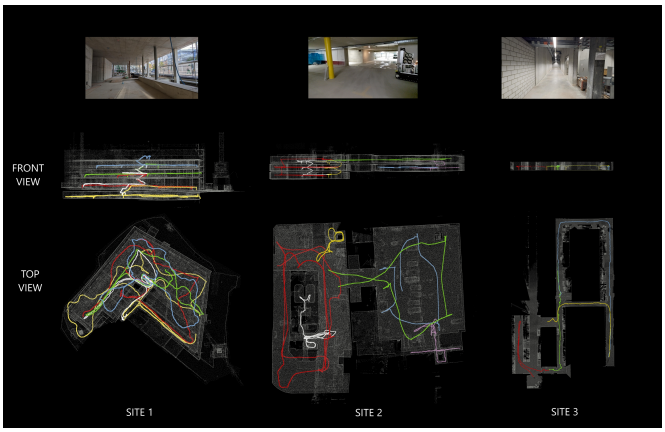


Fig. 4: Locations: Images, Downsampled TLS Scans, and globally aligned FastLio2/AdaLio based Trajectories.

We tested the accuracy of our method by locating the center of ground control points (GCPs) arranged on a 6x6 grid, as illustrated in Figure 2c. Using the Kabsch Algorithm [15], we calculated the transformation between the LiDAR's position and the grid. Our findings, summarized in Figure 3, show a maximum grid positioning error of 4.7mm. By applying a Rayleigh distribution to our data, we established a 95% confidence error threshold of 4.54mm and a 99.7% threshold of 6.32mm, respectively.

IV. Dataset

A. Location and Data Overview

The Hilti SLAM Challenge 2023 dataset features two primary locations (Site 1 and Site 2) and one secondary location (Site 3) as seen in Figure 4. The reason for having a primary-secondary split is discussed in section V-C. Site 1 is a multi-storey new construction with over $4000\,\text{m}^2$ of floor space, explored via our handheld device, Phasma. It contains sequences with bright to dark transitions, narrow corridors, with multiple areas re-visited. Site 2 is a three-storey underground parking lot under renovation with over $7500\,\text{m}^2$ of floor space, challenging manhattan world assumptions with non-parallel walls, ramps, and floors containing gradients for storm water drainage. It has sequences from both Trailblazer and Phasma, which have varying degrees in overlap difficulty. Site 3 is a tunnel corridor network, parts of which are repurposed as a warehouse. It contains sequences collected exclusively from Phasma, with multiple door transitions, repetitive structured rooms, and induced system level issues.

The dataset format has been kept similar to the 2022 Challenge [1], with rosbags containing timestamped lidar, cameras and IMU, and GCP occurence topics. We also provided calibration files and the urdf model for Trailblazer[2] in addition to the already public Phasma urdf. In order to protect privacy we decided to blur faces of passers-by and car license plates from all the relevant camera streams.

[2]https://github.com/Hilti-Research/trailblazer_description

B. Ground Truth Extraction

The ground truth acquisition methodology was maintained the same as the 2022 Challenge's [1], with the primary change being the scanner model - a Trimble X7 TLS (Figure 1c) was used instead of the Z&F Imager 5016. This device change decision is due to the auto registration capabilities of the X7's scanning software, which greatly reduced the processing time by many hours. 92%, 84%, and 90% of registered scans are within 3mm of registration uncertainty or less on a per site basis. Scans containing higher uncertainty were leaf scans containing relatively few connections in the bundle adjustment refinement graph - none of the Challenge's GCPs were extracted from those scans.

V. Benchmarking and Evaluation System

A. Scoring System Update - Increased Accuracy Range

Taking into consideration the increase in accuracy of SLAM systems over the past years, as well as cumulative inaccuracies of multi-session systems, we updated our scoring methodology. For every GCP evaluated on a trajectory, the score s_i was defined as follows:

$$s_i = \begin{cases} 20 & \text{if } e_i < 0.005m \\ 10 & \text{if } 0.005m \leq e_i < 0.01m \\ 6 & \text{if } 0.01m \leq e_i < 0.03m \\ 5 & \text{if } 0.03m \leq e_i < 0.06m \\ 3 & \text{if } 0.06m \leq e_i < 0.1m \\ 1 & \text{if } 0.1m \leq e_i < 0.4m \\ 0 & \text{if } e_i \geq 0.4m \end{cases} \tag{1}$$

where e_i is the absolute distance error at the ith ground control point on that trajectory.

The total score for each dataset S_j is the percentage of the maximum possible score (i.e. if all points had <0.5 cm error and scored 20 points), except for one dataset that was added later, as explained in a section V-C.

$$S_j = \left(\frac{1}{20N} \sum_{i=0}^{N} s_i \right) \times 100 \tag{2}$$

Organization	Algorithm	Sensors			Odometry		SLAM Backend			Const.	Results		
		Lidar	IMU	Cam	Type	Real-Time	Global BA	Causal	LC	Tuning	GCP Coverage	RMSE ATE(m)	Score
Lidar-Based System Results													
1. KAIST URL	Based on [16]–[19]	✓	✓		Filter	✓	✓	✗	✓	✗	100%	0.033	1177.64
2. HKU-MaRS	Based on [20]–[23]	✓	✓		Filter+Opt	✓	✓	✗	✗	✓	81.82%	*0.012	934.40
3. Innopolis Univ.	Strelka (Not published)	✓	✓		Filter	✓	✓	✗	✓	✗	89.09%	*0.016	840.35
4. SNU RPM Lab	Based on [18], [19], [24]	✓	✓		Filter	✗	✓	✗	✓	✗	98.18%	+0.059	789.48
5. Tsinghua Univ.	FT-LVIO [25]	✓	✓	✓(1)	Filter	✓	✗	✓	✗	✗	100%	0.047	727.35
6. ANYbotics	FrankenPharos (Not published)	✓	✓		SW Opt	✓	✗	✓	✗	✓	100%	0.058	657.61
7. B. Kim et. al.	Based on [17], [19], [20], [26]	✓	✓		Filter	✓	✓	✗	✓	✓	100%	0.048	629.30
8. NTU IOT	Based on [20], [22], [27]	✓	✓		Filter	✓	✓	✗	✗	✗	100%	0.067	597.65
Vision-Based System Results													
1. Tencent XR	MAVIS [28]		✓	✓	SW Opt	✓	✓	✗	✓	✓	87.27%	0.051	452.20
2. KAIST URL	Stereo UV SLAM [29]		✓	✓	SW Opt	✓	✓	✗	✓	✓	78.18%	+0.190	200.80
3. ASL ETHZ	Based on [30], [31]		✓	✓	SW Opt	✓	✓	✗	✓	✗	100%	0.189	121.20

*: RMSE ATE is only calculated for data that is submitted. Cases where teams receive lower/better RMSE ATE numbers but poor ranking is as an indicator that sequence trajectories are either incomplete or skipped

+: This team receives a higher score than subsequent leaderboard entries despite lower GCP coverage and higher RMSE ATE, due to its best GCP estimates being in a higher scoring band.

TABLE I: Single-session SLAM leaderboards - Top entries

Organization	Algorithm	Sensors			Fully	Results		
		Lidar	IMU	Cam	Automated	GCP Coverage	RMSE ATE	Score
Lidar-Based System Results								
1. ETHZ ASL	Based on [20], [31]	✓	✓	✓	✓	100%	0.294	36.7
2. Innopolis Univ	Strelka (Not published)	✓	✓		✗	12.72%	0.359	1.7
Vision-Based System Results								
1. Tencent XR	MAVIS [28]		✓	✓	✓	83.63%	0.287	27.0
2. ETHZ ASL	Based on [30], [31]		✓	✓	✓	100%	0.397	15.3

TABLE II: Multi-session SLAM leaderboards

where N is the number of ground truth points evaluated in each dataset. This denominator normalizes the score for a particular run to be between 0 and 100, regardless of the number of GCP's evaluated in each dataset. The final score is then a sum of the scores from each sequence.

B. Participant Calibration Data Inclusion

Since the GCP is observed by Trailblazer in the Lidar frame and participant trajectory outputs are in the IMU frame, we provided participants with the option to submit their own extrinsic calibration between LiDAR and IMU in the MultiCal format [13], for ATE calculation. For teams that did not perform their own calibration, we used our own internally calibrated parameter values.

C. System Vulnerabilities and Mitigation Measures

During the challenge we found that it's possible for teams to experimentally estimate the exact coordinates of the ground control points via a simple euclidean distance formula via the error graph analysis plots that we provide as feedback, and the timestamps at which the system is positioned at a GCP. Manually adjusting the trajectories to achieve better results is not in the spirit of the challenge. As a mitigation measure, we rejected entries with sparse trajectories. To offset effects of any future occurences, we introduced Site 3 as a secondary location, containing only Phasma sequences and excluding error analysis plots. We also decided to double the score per trajectory of this dataset. Formula (2) was modified for this particular dataset:

$$S_3 = \left(\frac{1}{20N} \sum_{i=0}^{N} s_i \right) \times 200 \tag{3}$$

D. Extension Towards Multi-session SLAM

The new stream introduced for the 2023 challenge was the introduction of Multi-session SLAM evaluation. For teams that opted to do so, they submitted all their trajectories in a common reference frame. The multi-session scores are published on separate leaderboards for both LiDAR driven and vision driven systems. The scoring brackets and normalization of total scores remain the same as in single-session. One subtle difference to note is that in multi-session evaluation all trajectories for each particular location are considered as one complete trajectory and normalized as such.

VI. The 2023 Challenge

A. Analysis

The Hilti SLAM Challenge 2023 saw an increase in participation over the previous iterations with 69 unique teams participating, compared to 42 from 2022, and 27 from 2021. This year, we witnessed multiple interesting trends. The Lidar single session category as seen in Table I is led by academia, with KAIST securing the top spot. Hierarchical Bundle adjustment [23] is a new high-performing system backend created by runner-up HKU-MARS Lab. They also decided to skip Loop Closure altogether. There are also high-performing open-loop odometry-only systems in the top teams such as Tsinghua University's FT-LVIO [25] system and ANYBotics' FrankenPharos[3], which secured the 5th and 6th spots on the leaderboard. We continue to see very limited camera-augmented lidar SLAM systems in the top 10 teams, with FT-LVIO [25] being the only entry of that kind. The Vision/IMU Single Session Category is led by big-tech, with Tencent Games' XR Lab with their MAVIS system securing first place. The category also finds all top SLAM frontends being optimizer based. On comparing Lidar-based SLAM vs. Vision-based SLAM, the RMSE ATE accuracy gap appears to be closing when compared to the 2022 Challenge leaderboard. Interestingly, none of the top-performing teams in the displayed leaderboards chose to include their own extrinsic calibration in their final submissions.

[3] https://submit.hilti-challenge.com/submission/1edf8227-fc31-6db8-bbd5-c34ba577a9cc/report

The Multisession SLAM tracks for lidar and vision-based systems (Table II) saw relatively limited participation, with two teams participating in each track, which was anticipated given the limited number of multi-session systems available. Both vision- and lidar-driven multi-session systems received similar scores at similar accuracies in the 30cm range of RMSE ATE. But this changed as of late 2023 when multiple new systems populated the lidar multi-session leaderboard[4]. A more in-depth analysis of the challenge and its results can be found on our full-length pre-print manuscript available on ArXiV[5].

B. Known Issues

Our ground truth system is based on the sparse ground truth extracted from a unified terrestrial laser scan on a per-site basis, without dense trajectory coverage. Regarding calibration, we acknowledge that our system does not have an external validation methodology for accuracy. Hence, we released calibration datasets and provided teams with the option to apply their own extrinsic calibration for evaluation. Lastly, in our robot calibration, the rotation of the roll and pitch axes has been limited to the amount of flex that the manual pallet jack could tolerate.

VII. Conclusions

With the 2023 SLAM Challenge iteration, we are delighted to extend our dataset and benchmarking abilities into multi-device and multi-session SLAM. We witness top-performing single-session lidar SLAM systems attain sub 2cm accuracy in general handheld scenarios. We also see the accuracy gap between vision- and lidar-driven SLAM reducing. Our live leaderboards continue to be available for public use. We hope that the SLAM systems using our dataset and benchmark are deployed to benefit mankind.

Acknowledgments

We would like to thank all the SLAM teams worldwide for their continued participation and engagement, as well as Hilti team members both current and former for their contributions.

References

[1] L. Zhang, M. Helmberger, L. F. T. Fu, D. Wisth, M. Camurri, D. Scaramuzza, and M. Fallon, "Hilti-oxford dataset: A millimeter-accurate benchmark for simultaneous localization and mapping," *IEEE Robotics and Automation Letters*, vol. 8, p. 408–415, Jan. 2023.

[2] M. Grupp, "evo: Python package for the evaluation of odometry and slam.." `https://github.com/MichaelGrupp/evo`, 2017.

[3] 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)*, pp. 7244–7251, 2018.

[4] M. Helmberger, K. Morin, B. Berner, N. Kumar, G. Cioffi, and D. Scaramuzza, "The hilti slam challenge dataset," *IEEE Robotics and Automation Letters*, vol. 7, p. 7518–7525, July 2022.

[5] S. Z. et al., "Subt-mrs: A subterranean, multi-robot, multi-spectral and multi-degraded dataset for robust slam," 2023.

[6] W. W. et al., "Tartanair: A dataset to push the limits of visual slam," 2020.

[7] M. T. et al., "Conslam: Periodically collected real-world construction dataset for slam and progress monitoring," in *Computer Vision–ECCV 2022 Workshops: Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part VII*, pp. 317–331, Springer, 2023.

[8] M. Trzeciak, K. Pluta, Y. Fathy, L. Alcalde, S. Chee, A. Bromley, I. Brilakis, and P. Alliez, "Conslam: Construction data set for slam," *Journal of Computing in Civil Engineering*, vol. 37, no. 3, p. 04023009, 2023.

[9] D. LEE, M. Jung, and A. Kim, "ConPR: Ongoing construction site dataset for place recognition," in *IROS 2023 Workshop on Closing the Loop on Localization: What Are We Localizing For, and How Does That Shape Everything We Should Do?*, 2023.

[10] W. Gao, "imu_utils." `https://github.com/gaowenliang/imu_utils`, 2018.

[11] O. J. Woodman, "An introduction to inertial navigation," Tech. Rep. UCAM-CL-TR-696, University of Cambridge, Computer Laboratory, Aug. 2007.

[12] P. Furgale, J. Rehder, and R. Siegwart, "Unified temporal and spatial calibration for multi-sensor systems," in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1280–1286, 2013.

[13] X. Zhi, J. Hou, Y. Lu, L. Kneip, and S. Schwertfeger, "Multical: Spatiotemporal calibration for multiple imus, cameras and lidars," in *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 2446–2453, 2022.

[14] J.-K. Huang, S. Wang, M. Ghaffari, and J. W. Grizzle, "Lidartag: A real-time fiducial tag system for point clouds," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 4875–4882, 2021.

[15] W. Kabsch, "A solution for the best rotation to relate two sets of vectors," *Acta Crystallographica Section A*, vol. 32, pp. 922–923, Sep 1976.

[16] H. Lim, D. Kim, B. Kim, and H. Myung, "Adalio: Robust adaptive lidar-inertial odometry in degenerate indoor environments," 2023.

[17] H. Lim, S. Yeon, S. Ryu, Y. Lee, Y. Kim, J. Yun, E. Jung, D. Lee, and H. Myung, "A single correspondence is enough: Robust global registration to avoid degeneracy in urban environments," 2022.

[18] A. Segal, D. Haehnel, and S. Thrun, "Generalized-icp," in *Proceedings of Robotics: Science and Systems*, (Seattle, USA), June 2009.

[19] F. D. et al., "borglab/gtsam: 4.2a8," Nov. 2022.

[20] W. Xu, Y. Cai, D. He, J. Lin, and F. Zhang, "Fast-lio2: Fast direct lidar-inertial odometry," 2021.

[21] C. Yuan, W. xu, X. Liu, X. Hong, and F. Zhang, "Efficient and probabilistic adaptive voxel mapping for accurate online lidar odometry," 2022.

[22] Z. Liu, X. Liu, and F. Zhang, "Efficient and consistent bundle adjustment on lidar point clouds," 2022.

[23] X. Liu, Z. Liu, F. Kong, and F. Zhang, "Large-scale lidar consistent mapping using hierarchical lidar bundle adjustment," *IEEE Robotics and Automation Letters*, vol. 8, no. 3, pp. 1523–1530, 2023.

[24] D. He, W. Xu, N. Chen, F. Kong, C. Yuan, and F. Zhang, "Point-lio: Robust high-bandwidth light detection and ranging inertial odometry," *Advanced Intelligent Systems*, vol. 5, no. 7, p. 2200459, 2023.

[25] Z. Zhang, Z. Yao, and M. Lu, "Ft-lvio: Fully tightly coupled lidar-visual-inertial odometry," *IET Radar, Sonar & Navigation*, vol. 17, no. 5, pp. 759–771, 2023.

[26] K. Koide, M. Yokozuka, S. Oishi, and A. Banno, "Voxelized gicp for fast and accurate 3d point cloud registration," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 11054–11059, 2021.

[27] P. Dellenbach, J.-E. Deschaud, B. Jacquet, and F. Goulette, "Ct-icp: Real-time elastic lidar odometry with loop closure," 2021.

[28] Y. Wang, Y. Ng, I. Sa, A. Parra, C. Rodriguez, T. J. Lin, and H. Li, "Mavis: Multi-camera augmented visual-inertial slam using se2(3) based exact imu pre-integration," 2023.

[29] H. Lim, J. Jeon, and H. Myung, "Uv-slam: Unconstrained line-based slam using vanishing points for structural mapping," *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 1518–1525, 2022.

[30] S. Leutenegger, S. Lynen, M. Bosse, R. Siegwart, and P. Furgale, "Keyframe-based visual-inertial odometry using nonlinear optimization," *The International Journal of Robotics Research*, vol. 34, no. 3, pp. 314 – 334, 2015. Published online before print 15 December 2014.

[31] A. C. et al., "maplab 2.0 – A Modular and Multi-Modal Mapping Framework," *IEEE Robotics and Automation Letters*, vol. 8, no. 2, pp. 520–527, 2023.

[4] `https://www.hilti-challenge.com/leader-board-2023.html`

[5] `https://arxiv.org/abs/2404.09765`