HILTI 2026 SLAM Challenge - Submission Report

Simon Boche, Jaehyung Jung, Sebastián Barbas Laina, Stefan Leutenegger

ABOUT THE TEAM

Team Name: Mobile Robotics Lab (MRL)
Affiliation: TU Munich / ETH Zurich

I. DESCRIPTION OF THE APPROACH

OKVIS2-X [1] is a multi-sensor (including cameras, IMU, LiDAR, and GNSS) SLAM system tightly-coupled with dense volumetric occupancy mapping within a factor graph optimization framework. For the challenge specifically, OKVIS2-X was configured with a stereo-inertial setup without the occupancy mapping due to the lack of enough overlap between the two cameras. The following summarizes the configuration and modifications used in this challenge.

- **Static Mask**
  We define static image masks to remove keypoints detected from self-occluded areas. As shown in Fig. 1, we adapted the mask provided by the organizer and use the same masks for the all evaluation.

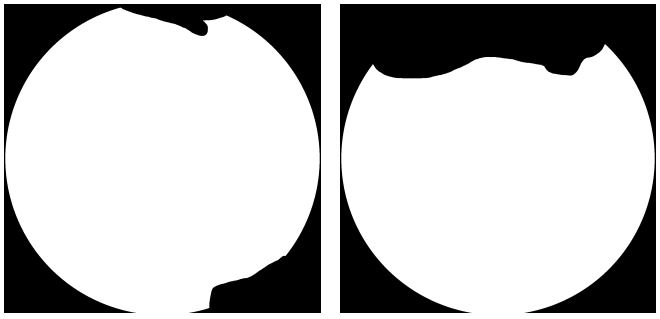

Fig. 1. Left and right image masks for removing self-occluded areas.

- **EUCM Camera Model**
  To account for wide-angle fish eye cameras, we employ the enhanced unified camera model (EUCM) [2] with analytic Jacobians.
- **Online calibration of IMU-to-camera extrinsics**
  We model time-invariant transformation between IMU to cameras in reprojection and relative pose residuals.
- **Refinement through final bundle adjustment**
  After consuming all sensor data in a causal way, we refine the keyframe trajectory and observed landmarks in a final bundle adjustment given the initial estimate from the causal estimation. In this way, we estimate the trajectory given all sensor measurements.
- **Image median filter**
  To mitigate image noise, particularly in dark scenes, we applied OpenCV's `cv::medianBlur` with a kernel size of 3 throughout the evaluation.

The authors are with the Mobile Robotics Lab (MRL), Technical University of Munich (TUM) and ETH Zurich.

II. TIMINGS

OKVIS2-X is a real-time SLAM system. We have included the timings for the main components of the SLAM system in Table I when processing the sequence `floor1_2025_12_02_run_1`, the longest in the dataset (duration: $\approx 280s$). Note that for this challenge evaluation, we processed each sequence in a synchronous manner, meaning processing all available data without dropping any frames. It is important to note that OKVIS2-X is a multi-threaded system, therefore the total processing time is not the sum of the individual components. The total processing time, including the final bundle adjustment, is 409.91 seconds (circa 389 seconds without the final BA refinement).

TABLE I
AVERAGE RUNTIME FOR THE MAIN COMPONENTS OF OKVIS2-X IN THE SEQUENCE FLOOR1_2025_12_02_RUN_1

Task	Average per frame [ms]	Total time [s]
Keypoint detection and description	12.6	104.32
Landmark matching	17.5	144.98
Optimisation	19.3	159.64
Marginalization	6.4	53.39

III. THE FULL RESULTS

All our results were obtained with a unique configuration file, meaning, all sequences were run with the same set of parameters and not fine-tuned to specific sequences. Due to the randomness (e.g., RANSAC) in OKVIS2-X, we ran six independent runs and report the best score within the six runs. For clarity we report all results in the Table II.

REFERENCES

[1] S. Boche, J. Jung, S. B. Laina, and S. Leutenegger, "OKVIS2-X: Open keyframe-based visual-inertial SLAM configurable with dense depth or lidar, and gnss," *IEEE Transactions on Robotics*, 2025.
[2] B. Khomutenko, G. Garcia, and P. Martinet, "An enhanced unified camera model," *IEEE Robotics and Automation Letters*, vol. 1, no. 1, pp. 137–144, 2015.

TABLE II
LOCALIZATION SCORES OF 6 RUNS.

	Run 0	Run 1	Run 2	Run 3	Run 4	Run 5	Best
floor_1_2025-05-05_run_1	90.81	**95.12**	94.27	89.52	93.25	93.96	-*
floor_1_2025-07-07_run_1	**94.13**	91.11	90.39	83.73	92.06	91.94	94.13
floor_1_2025-12-02_run_1	94.95	95.59	96.18	**96.21**	94.83	95.13	96.21
floor_2_2025-05-05_run_1	82.27	88.77	**92.47**	90.58	91.87	90.94	-*
floor_2_2025-10-28_run_1	95.25	94.96	95.21	94.88	95.07	**96.52**	-*
floor_2_2025-10-28_run_2	91.82	91.62	90.90	92.38	91.99	**92.64**	-*
floor_2_2025-12-02_run_1	79.14	80.82	73.43	67.49	**82.88**	76.17	82.88
floor_2_2025-12-03_run_1	93.81	90.58	94.95	91.98	**95.34**	94.19	95.34
floor_3_2025-05-19_run_1	94.67	94.76	94.72	94.59	95.42	**95.72**	95.72
floor_3_2025-12-02_run_1	80.26	82.48	91.31	88.97	86.90	**91.83**	91.83
floor_4_2025-05-19_run_1	94.55	90.86	89.66	96.16	93.59	**96.44**	96.44
floor_4_2025-12-02_run_1	96.62	94.32	95.09	93.24	**97.29**	96.29	97.29
floor_5_2025-12-02_run_1	86.79	**92.65**	92.61	83.93	85.68	83.57	92.65
floor_6_2025-06-18_run_1	95.80	94.31	92.40	94.83	**96.29**	94.11	96.29
floor_6_2025-07-07_run_1	94.27	**97.27**	95.73	96.00	95.02	95.68	97.27
floor_6_2025-12-02_run_1	92.97	**95.13**	92.27	90.60	89.62	95.06	95.13
floor_6_2025-12-02_run_2	94.33	90.24	**96.24**	94.09	93.18	93.13	96.24
floor_7_2025-12-02_run_1	95.00	95.51	**96.61**	95.75	94.12	95.84	96.61
floor_7_2025-12-02_run_2	91.37	91.66	91.33	91.18	**92.45**	90.24	92.45
floor_7_2025-12-03_run_1	90.58	88.05	89.45	89.10	**92.98**	91.58	92.98
floor_EG_2025-10-16_run_1	95.48	94.25	95.18	94.08	**96.76**	94.30	96.76
floor_EG_2025-12-02_run_1	**96.37**	94.37	91.61	93.91	93.58	94.19	96.37
floor_EG_2025-12-02_run_2	94.20	**95.43**	94.33	94.94	94.75	94.92	95.43
floor_UG1_2025-05-19_run_1	81.53	74.17	75.70	77.01	77.26	**83.86**	83.86
floor_UG1_2025-06-18_run_1	82.24	79.76	69.46	67.61	83.23	**83.48**	83.48
floor_UG1_2025-10-16_run_1	88.88	92.52	91.61	**94.00**	87.23	91.19	-*
floor_UG1_2025-12-02_run_1	88.44	76.80	**91.30**	85.76	88.48	90.04	91.30
floor_UG1_2025-12-02_run_2	84.54	85.28	81.83	84.17	84.81	**87.71**	87.71
floor_UG1_2025-12-03_run_1	**96.78**	95.53	94.27	95.30	94.70	96.70	96.78
floor_UG2_2025-12-02_run_1	**95.35**	89.53	93.11	92.93	91.62	90.51	95.35
Total	2284.17	2250.45	2259.51	2233.55	2282.85	**2292.61**	2336.50

Best scores per sequence are in bold.
Best column refers to best result per sequence over all 6 runs.
*Sequences that are not considered in the challenge are grayed out and not considered in the score.