

Performance Analysis of Visual-Inertial Odometry under Degraded Underwater Visual Conditions

Gavin Moroney¹ and Dr. He (David) Zhang²

¹ Department of Electrical Engineering,
² Department of Computer Science,

Grove City College, Grove City, PA 16127

Spring 2026



Abstract: Visual-inertial odometry (VIO) has achieved remarkable progress in recent years and has been extensively deployed in unmanned aerial and ground vehicles for real-time state estimation and navigation. State-of-the-art (SOTA) VIO methods demonstrate strong robustness and accuracy in structured terrestrial environments. However, their performance has not been comprehensively studied in underwater settings, where haze, blurring, and low contrast present significant challenges for reliable visual feature extraction and tracking. In this study, we investigate the performance of several state-of-the-art VIO methods using datasets collected by underwater vehicles equipped with stereo cameras, inertial measurement units (IMUs), barometric pressure sensors, and sonar systems. The findings of this study aim to provide systematic insights into the limitations and adaptability of current VIO approaches in underwater environments and to lay the groundwork for developing robust navigation solutions for autonomous underwater vehicles.

Background



Figure 1. A: UZH-FPV Drone [1]
 B: Indoor dataset. C: Outdoor dataset.

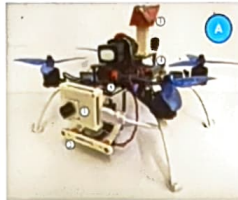


Figure 2. A: Aqua2 underwater robot [2]
 B: The BUS dataset. C: The CAVE dataset.

| Component | OpenVINS (MSCKF) | VINS-Mono | OKVIS |
|-----------|--|---|---|
| Tracking | KLT (Pyramidal Lucas-Kanade optical flow) | KLT (Pyramidal Lucas-Kanade optical flow) | Patch/descriptor matching (Patch-based or descriptor-based search) |
| Backend | Filter (MSCKF) (Multi-State Constraint Kalman Filter) | Optimization (Non-linear optimization with sliding window) | Optimization (Non-linear optimization with sliding window) |

Table 1. SOTA VIO methods: MSCKF [3], [4], VINS-Mono [5], OKVIS [6]

| | MSCKF | VINS-Mono | OKVIS |
|------|--------|-----------|-------------|
| BUS | 0.65 | failed | 0.24 |
| CAVE | failed | 0.37 | 0.27 |

Table 2. RMSE of the pose estimation results of the SOTA VIO with BUS and CAVE datasets

SOTA VIO methods Performance Analysis

MSCKF and VINS-Mono both rely on KLT optical flow for feature tracking, while OKVIS uses patch/descriptor-based matching, which is generally more robust to blur, haze, and low-contrast imagery. MSCKF adopts a filter-based backend that is efficient but can be less tolerant to persistent tracking errors, whereas VINS-Mono and OKVIS use optimization-based sliding-window estimation for improved consistency.

In the underwater datasets, degraded images weaken visual tracking and the slow, smooth vehicle motion provides limited IMU excitation. Under these conditions, VINS-Mono's pose estimation diverged on the BUS dataset, and MSCKF diverged on the CAVE dataset. In contrast, OKVIS remained stable on both datasets and achieved the lowest RMSE values of 0.24 on BUS and 0.27 on CAVE, demonstrating stronger robustness in challenging underwater environments.

Experimental Results: Trajectory and Position Error Visualization

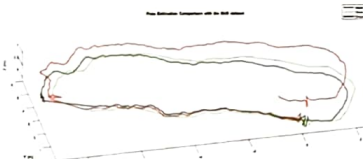


Figure 3. Trajectory Comparison of MSCKF (red) and OKVIS (green) with the BUS dataset

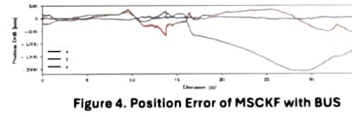


Figure 4. Position Error of MSCKF with BUS

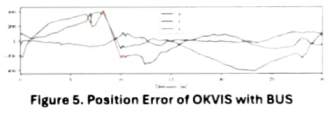


Figure 5. Position Error of OKVIS with BUS



Figure 6. Position Error of VINS-Mono with BUS

Pose estimation diverges when the robot travels around 14 meters



Figure 7. Trajectory Comparison of VINS-Mono (red) and OKVIS (blue) with the CAVE dataset

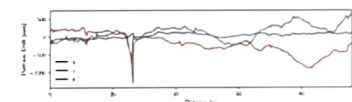


Figure 8. Position Error of VINS-Mono with CAVE

Pose estimation diverges in the beginning of the sequence

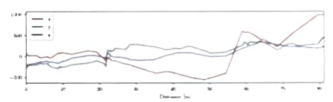


Figure 9. Position Error of OKVIS with CAVE



Figure 10. Position Error of MSCKF with CAVE

References

- [1] Delmerico, Jeffrey, et al. "Are we ready for autonomous drone racing? the UZH-FPV drone racing dataset." In *Proc. of the International Conference on Robotics and Automation (ICRA)*, 2019.
- [2] G. Dudek, M. Jenkin, C. Prahacs, A. Hogue, J. Sattar, P. Giguere, A. German, H. Liu, S. Saunderson, A. Ripsman, S. Simhon, L. A. Torres-Mendez, E. Milios, P. Zhang, and I. Rekleitis, "A visually guided swimming robot." In *Proc. IROS*, Edmonton AB, Canada, Aug. 2005, pp. 1749-1754
- [3] Geneva, Patrick, et al. "Openvins: A research platform for visual-inertial estimation." In *Proc. of the International Conference on Robotics and Automation (ICRA)*, 2020.
- [4] Sun, Ke, et al. "Robust stereo visual inertial odometry for fast autonomous flight." In *IEEE Robotics and Automation Letters* 3.2 (2018): 965-972.
- [5] Qin, Tong, Peiliang Li, and Shaojie Shen. "Vins-mono: A robust and versatile monocular visual-inertial state estimator." In *IEEE transactions on robotics* 34.4 (2018): 1004-1020.
- [6] Stefan Leutenegger, Simon Lynen, Michael Bosse, Roland Siegwart and Paul Timothy Furgale. "Keyframe-based visual-inertial odometry using nonlinear optimization." In *the International Journal of Robotics Research*, 2015.