Robust Indoor Mapping on a Smart Walking Aid via Visual SLAM

Martin Obert, Subashkumar Rajanayagam and Stefan Twieg
Department of Electrical, Mechanical and Industrial Engineering, Anhalt University of Applied Sciences,
Bernburger Str. 55, Köthen, Germany
martin.obert@hs-anhalt.de, stefan.twieg@hs-anhalt.de, subashkumar.rajanayagam@hs-anhalt.de

Keywords: Visual SLAM, RTAB-Map, Indoor Mapping, Smart Walking Aid, Assistive Technology, Edge Computing.

Abstract: The underlying research project aims to enhance independent mobility for elderly users by enabling indoor navigation assistance in a smart walking aid. A robust mapping system is crucial for such a device to ensure safety, effectively avoid obstacles, and create an optimal navigation path for users. Current approaches often rely on expensive sensors (e.g. LIDAR) or suffer from drift in unstructured environments. This paper proposes an optimized Visual SLAM solution based on the RTAB-Map framework for indoor mapping with a walking aid-mounted RGB-D camera. The method leverages appearance-based loop closure detection and parameter optimization to reliably generate both 2D occupancy grids and 3D point clouds under challenging conditions, like rooms located on different levels and transitions between rooms that are difficult to detect due to their layout. Indoor experiments demonstrate that the optimized system successfully handles typical walking aid motions and indoor scene features, achieving accurate reconstruction of corridors and rooms. Mapping accuracy is validated against the floor plan, showing lower drift than the baseline Visual SLAM setup. The results indicate that RTAB-Map, with proper parameter tuning, provides a solid and cost-effective mapping solution for assistive walking aid platforms, facilitating safe navigation support for the elderly.

1 INTRODUCTION

By the middle of the 2030s, the number of people in Germany at retirement age is expected to increase by approximately 4 million to at least 20 million, as seen on Figure 1. The number of individuals aged 80 or over will remain stable initially, but is expected to rise significantly after the mid-2030s [1], likely leading to a higher demand for long-term care. These demographic changes pose challenges for society that are already visible in today's age structure.

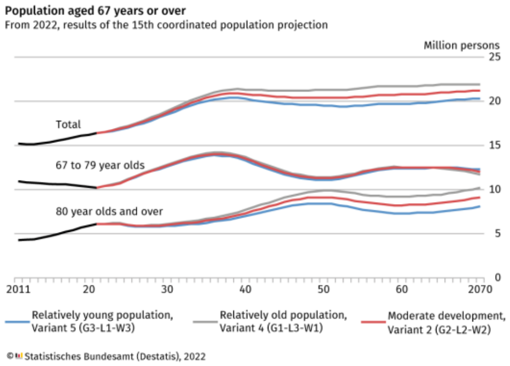

Figure 1: German population projection [1].

To address this trend, the AktiMuW[1] project [2] was initiated to provide an ageing population with a degree of autonomy by integrating technical solutions into a conventional walking aid. In addition to key functions such as sensor fusion-based analysis of body posture [3] and machine learning-based detection of potentially hazardous objects to respond to the high risk of injuries and mortality in road traffic accidents that pedestrians face with increasing age [4], another goal is to assist users in navigating indoor spaces (e.g., living environments, care facilities), which motivates integrating a Visual SLAM (Simultaneous Localization and Mapping using cameras) system on the walking aid. Visual SLAM [5] enables a device to map unknown environments and localize itself without GPS. RGB-D and stereo cameras provide both visual and depth, making them well-suited for robust indoor SLAM tasks. Combining with robotic interaction and reinforcement learning enables to handle objects [6].

This work investigates the use of RTAB-Map (Real-Time Appearance-Based Mapping) [7] on a walking aid-mounted stereo camera and an embedded GPU board. RTAB-Map is an open-source Visual SLAM framework that incrementally builds 3D maps

---

[1] https://www.hs-anhalt.de/aktimuw/uebersicht.html

with appearance-based loop closure. It is ROS2-compatible (Robot Operating System 2) [8] and can run on resource-limited hardware.

Few studies explicitly explore vision-based SLAM on walker-like aids. Notably, Panteleris and Argyros developed the c-Walker – a rollator platform equipped with an RGB-D (Kinect) camera – and implemented a Visual SLAM system that simultaneously tracks moving objects and builds a 3D map of the static environment. They report real-time mapping on the walker with low-end CPUs [9]. Feltner et al. also describe a Kinect-based rollator for the visually impaired, pointing out that a Kinect and Visual SLAM approach is promising for the c-Walker [10]. In summary, the only published SLAM work on rollator-style devices we found involves specialized walkers with RGB-D cameras (e.g. the c-Walker). To our knowledge, no prior work has run modern feature-based SLAM (RTAB-Map, ORB-SLAM3, etc.) on a standard push-style rollator. Thus, applying an up-to-date visual SLAM framework on a commodity walker as a modular system remains novel.The main contribution integrates a ZED Mini stereo camera and Jetson Orin Nano into a ROS2-based SLAM pipeline; correcting TF transforms for consistent spatial frames; optimizing SLAM parameters (e.g., switching from GTFF/ORB (Good Features To Track / Oriented FAST and Rotated BRIEF) [11] to KAZE (jap. wind) [12] feature descriptor); and demonstrating stable 2D/3D mapping results in real indoor environments.

2 COMPARISON OF SLAM FRAMEWORKS

Recent SLAM systems for indoor embedded use were compared to find the most suitable approach. ORB-SLAM3 (Oriented FAST and rotated BRIEF) is highly accurate (monocular/stereo/RGB-D, can fuse IMU) but requires a powerful CPU/GPU to run in real time; it “outperforms others in accuracy” on desktop hardware [13]. VINS-Mono and other VIO systems (Visual-Inertial Odometry) use inertial data to improve robustness, but still strain embedded CPUs [14]. DSO/SVO (Semi-/Direct Sparse Odometry) are direct methods (no feature extraction) that can run on grayscale cameras; they track well on textured scenes but easily drift or fail in low-texture or changing light [15]. In one study on an Nvidia Jetson TX1, DSO had much higher drift (Absolute Trajectory Error RMSE ≈0.46 m) than ORB-SLAM (≈0.17 m) or RTAB-Map [15]. RTAB-Map is a graph-based SLAM that supports RGB-D or stereo and loop closures in ROS2; it is widely used for robotics and can run on CPU-only boards. In the same Jetson test, RTAB-Map achieved one of the lowest errors (≈0.16 m) and compared favorably with ORB-SLAM [15]. RTAB-Map was chosen for our walker prototype because it supports RGB-D input, is ROS2-ready, and has proven performance on embedded platforms [16]

3 SYSTEM COMPONENTS

In this section, we outline the system architecture for our indoor mapping solution, featuring a robust integration of hardware and software components. The Nvidia Jetson Orin Nano and Stereolabs ZED Mini provide the necessary computational and sensory capabilities. The ROS 2-based software stack orchestrates data processing and communication, enabling real-time map generation via Visual SLAM.

3.1 Hardware

Nvidia Jetson Orin Nano: Selected for its balance of compute performance (up to 67 TOPS at 25 W) and compact form factor. The single-board computer’s (SBC) Ampere GPU generalizes to various parallel SLAM tasks, while its ARM CPU cores handle ROS2 middleware and peripheral integration.

Stereolabs ZED Mini: A stereo camera providing synchronized 1080p RGB streams at up to 30 fps and depth maps via hardware disparity calculation. It features a built-in IMU (6-axis gyroscope + accelerometer) at 800 Hz sampling, facilitating visual-inertial odometry.

Walking Aid Platform: A standard four-wheel walking aid instrumented with a rigid mount for the camera and an under-seat housing for the Jetson Orin Nano. Power is supplied via a portable battery capable of a stable 20 V/5 A output (per USB-PD profile).

Smartphone: A regular Android smartphone was selected for wireless and convenient visualization of the SBC’s desktop.

The fully assembled walking aid platform can be seen in Figure 2. The SBC is mounted in a specially designed device carrier under the seat. The stereo camera rests in a holder on the bend of the left handle from the user's perspective.

3.2 Software

The software architecture mainly focuses on ROS 2 (Humble Hawksbill). It provides Data Distribution

Service (DDS)-based publish–subscribe messaging, node lifecycle management, and real-time QoS configuration. For camera the ZED Wrapper is an official driver, packaging the ZED SDK into ROS2 nodes, exposing topics for left/right RGB images, depth images, point clouds, IMU data, and visual-inertial odometry. With RTAB-Map, a ROS2 node encapsulating the RTAB-Map core functions. It subscribes to camera and odometry topics, executes SLAM threads internally, and publishes 2D occupancy grids, 3D point clouds, and pose trajectories. Rviz2 is used for visualization and real-time monitoring of SLAM outputs, including map rendering and trajectory display.

Figure 2: AktiMuW walking aid platform with stereo camera and SBC mounted under the seat.

4 PRELIMINARY WORK

Initial map outputs exhibited warped and misaligned point clouds. Inspection of the ROS2 transformation tree revealed the absence of a transform edge between the base_link and zed_left_camera_optical_frame, preventing correct pose propagation. Additionally, the ZED optical frame conventions differed from ROS's REP-103 [17] conventions, causing rotated axes to be interpreted incorrectly.

4.1 Frame Correction

A zero-offset static transform was published to connect base_link to zed_left_camera_optical_frame:

```
ros2 run tf2_ros
static_transform_publisher 0 0 0 0 0 0
base_link zed_left_camera_optical_frame
```

This closed the TF gap, but the point cloud remained rotated as seen in Figure 3.

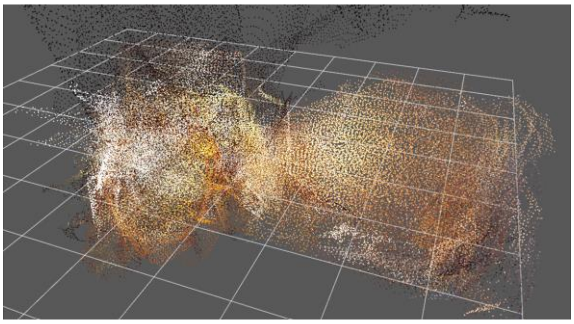

Figure 3: Distorted point cloud of a room after a 360° horizontal camera turn.

4.2 Axis Alignment

After examining the optical frame conventions of ROS2 and the camera, it became apparent that there is a discrepancy in the alignment of the optical axes (see Table 1).

Table 1: Comparison of ZED Wrapper and ROS2 Optical Frame Conventions and required Transformation.

Axis	ROS2 Optical Frame Convention	ZED Wrapper Optical Frame Convention	Required Transformation
X / Roll	Forward	Right	Right (ZED) → -90° → Forward (ROS2)
Y / Pitch	Left	Down	Not required due to other applied transformations
Z / Yaw	Up	Forward	Forward (ZED) → -90° → Up (ROS2)

The camera's optical axes require 90° ($-\pi/2$ radians) rotations about the roll and yaw axes to align with ROS2 conventions.

Additionally, the camera center sits approximately 90 centimetres above the walking aid's base. The final static transform command was:

```
ros2 run tf2_ros
static_transform_publisher 0 0 0.9 -
1.5708 0 -1.5708 base_link
zed_left_camera_optical_frame
```

After this update, the point cloud aligned correctly with the physical environment as seen in Figure 4:

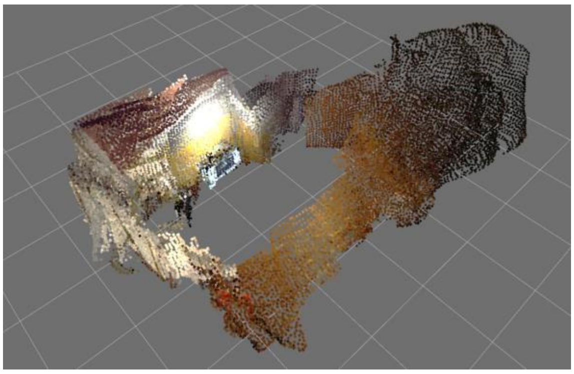

Figure 4: Corrected point cloud of a room after a 360° horizontal camera turn.

4.3 Selection of a Test Environment

The selected test environment is intended to resemble a living environment due to its varied structure and arrangement. A 47.52 m² office space, rich in differentiable characteristics (features), alternates with a 25.2 m² feature-poor entrance hall with complex geometry, which leads into further small bathrooms as seen in Figure 5.

Figure 5: Floor plan of the selected test environment.

Another distinctive characteristic is that the office space is situated 73 cm above the entrance hall, which means that the mapping algorithm is also required to overcome a certain height difference.

5 MAPPING RESULTS

As mentioned beforehand, the entrance hall and the two bathrooms have very few distinctive features, so this area of the test environment was primarily used for the baseline performance test and optimization process.

5.1 Baseline Performance

With the default settings (camera angle 0°, no depth limit, default settings for RTAB map functions), the initial mapping in the entrance hall exhibited several deficiencies as seen in Figure 6.

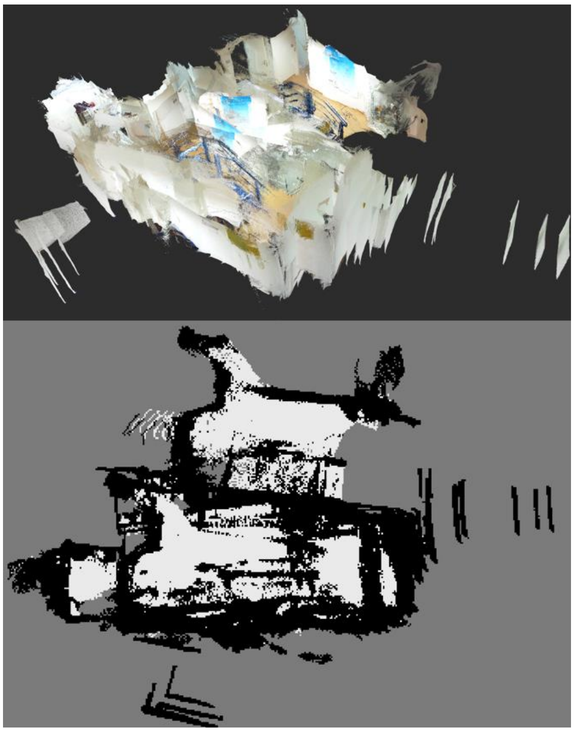

Figure 6: 3D point cloud and 2D occupancy grid with default settings.

Comparing reality vs. reconstructed map the main weakness is about:

1) Repeated Walls: Consecutive walls were often rendered multiple times in slightly shifted positions, creating "ghost" wall artifacts in the 3D point cloud and the 2D occupancy grid.
2) Poor Corridor Geometry: The long corridor segment was barely recognizable in outline, but

walls failed to align properly, leading to kinks and misalignments that did not match the actual room geometry.

3) Loss of Feature Tracking: In particularly difficult to differentiate areas, such as the bathrooms, feature tracking was often discontinued, which required the effort of identifying the last camera view detected.
4) Loop-Closure Gaps: Upon walking a closed loop, no automatic closure of the map occurred.

5.2 Experimental Challenges

During the baseline experiments, several challenges emerged that significantly affected SLAM performance on the walking aid platform:

- Reflective and low-texture surfaces: Glossy floor tiles and plain, bright bathroom walls provided very few reliable features, which led to tracking loss and occasional map discontinuities.
- Repetitive corridor geometry: Long corridors with visually similar wall and floor patterns caused perceptual aliasing, resulting in false loop-closure matches and duplicated wall segments.
- Dynamic disturbances: Moving people occasionally entered the scene, corrupting stereo matches and temporarily degrading odometry estimation.
- Illumination changes: Transitions between rooms with different lighting conditions introduced exposure artifacts, reducing feature matching robustness.

5.3 Optimized Result

After systematic parameter tuning, a set of parameters could be found that enabled a stable mapping result:

- Camera Angle: Tilted downward by 20° to the ground to minimize ceiling and window capture.
- Depth Range Limit: Reduced to a maximum of 5 m to exclude noisy distant measurements and reduce point-cloud clutter.
- Odometry: Switching from purely ZED Mini IMU-based odometry to visual-only odometry considerably improved path tracking performance.
- Feature Detector: Switched from GFTT/ORB to KAZE to enhance feature extraction in unvaried environments like brightly tiled bathrooms.
- Feature Tracking and Loop-Closure Optimization: Lowered the minimum number of inliers for feature matching (Vis/MinInliers) from 20 to 8 and raised the maximum number of features extracted per image (Vis/MaxFeatures) from 500 to 2000, enhancing image recognition and therefore loop-closure detection as seen in Figure 7.

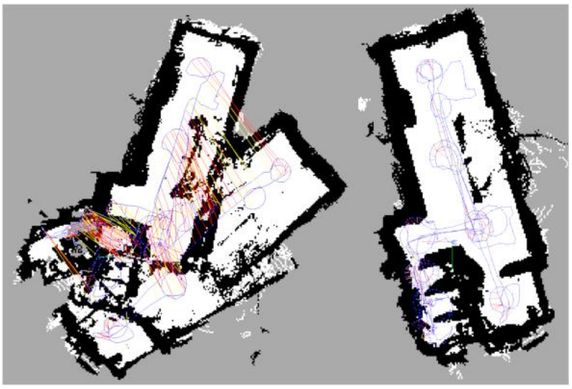

Figure 7: 2D occupancy grid before (left) and after loop closure (right).

5.4 Application to the Room Complex

The optimized mapping parameters were next applied to the full room complex, encompassing both the feature-rich office space and the adjacent entrance hall with the two bathrooms. A single uninterrupted path was executed, starting with the entrance hall and the bathrooms and proceeding to the office via stairway transition.

Despite the low-texture bright and in places tiled walls, the entrance hall and bathrooms were mapped without major gaps. The 20° camera tilt directed more features toward the floor-wall boundary, enabling the SLAM algorithm to maintain tracking in these challenging areas. The height difference between the entrance hall and office level was captured as a smooth vertical offset in the 3D map, with the staircase geometry correctly represented as seen in Figure 8.

As also seen in Figure 8, the office area was reconstructed with high fidelity. Desks, chairs, and shelving units appeared as distinct clusters in the 3D point cloud. The 2D occupancy grid accurately delineated walking aisles between desk clusters as seen in Figure 9.

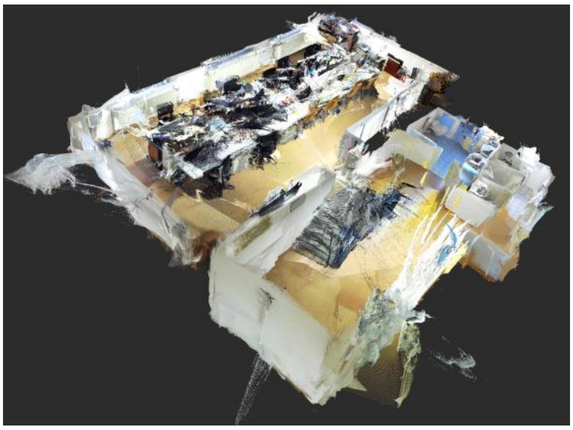

Figure 8: 3D point cloud of the complete test environment with visible offset between room levels.

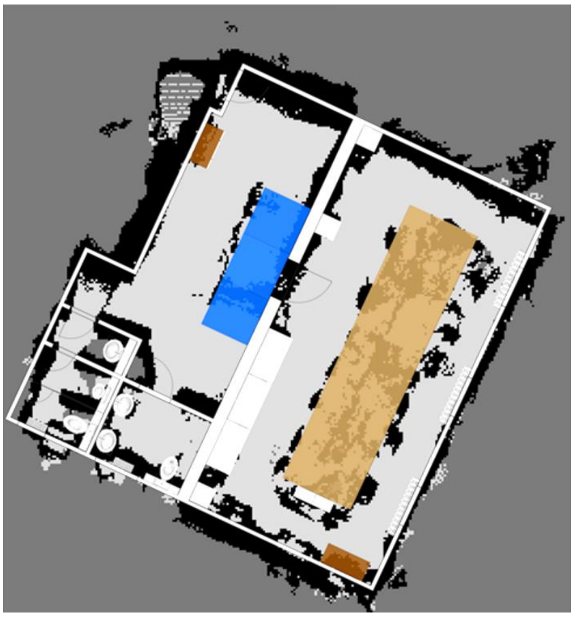

Figure 9: 2D occupancy grid of the complete test environment with validation against the floor plan.

Each rotation at the waypoint returned successfully closed prior loops, resulting in a fully connected 2D grid of the entire complex. No discontinuities were observed at room transitions or while overcoming the staircase.

6 CONCLUSIONS

Reliable indoor mapping is a critical enabler for navigation in smart assistive devices. This work shows that, even on a resource-constrained mobile platform like a smart walking aid, Visual SLAM can deliver accurate, consistent maps, provided that the system is carefully calibrated and adapted to real-world constraints.

By combining a stereo camera, an SBC and the open-source RTAB-Map framework, a full SLAM pipeline was deployed and optimized for real-time use. Key modifications, including depth limiting, camera tilt, axis-aligned TF corrections, and parameter tuning, proved essential for handling texture-poor environments, minimizing mapping artifacts, and preserving loop consistency.

The system did not just perform thoroughly in a controlled test zone; it successfully scaled to a full floor, including stairs, bathrooms, and low-feature areas. The resulting maps were geometrically accurate, loop-closed, and complete, forming a solid foundation for downstream tasks like path planning and user guidance.

These findings demonstrate that high-quality indoor SLAM is within reach even for compact, affordable assistive systems. With careful tuning, visual sensing alone can match the reliability once reserved for more expensive LIDAR-based setups. Future work on this topic will include the implementation of user-interface-guided real-time navigation.

6 FUTURE WORK

Future work will focus on transforming the presented SLAM-based mapping system into a fully integrated indoor navigation solution for assistive walking aids.

One important step will be the implementation of a user-friendly interface that allows elderly users to request navigation assistance and receive intuitive visual or haptic guidance. To further increase safety, the system will be extended with real-time obstacle detection and avoidance features, enabling reliable operation even in highly dynamic environments with moving people or pets.

Energy efficiency is another priority, as optimizing computation pipelines and sensor duty cycles will be essential to ensure extended operating times on battery-powered platforms. In addition to on-device improvements, the integration of Bluetooth Low Energy (BLE) beacons and other infrastructure-based localization technologies will be investigated to provide absolute positioning in large or feature-poor areas, such as open corridors or lobbies, where Visual SLAM performance can be limited. Furthermore, enriching geometric maps with semantic labels – for example, identifying doors, furniture, or hazard zones – offers the potential for context-aware navigation and improved caregiver support.

Finally, scalability will be addressed through the exploration of cloud-based map optimization and multi-floor map sharing, enabling operation in complex care facilities without overloading the embedded platform. These developments aim to evolve the current prototype into a robust, comprehensive navigation solution that seamlessly integrates on-board SLAM with infrastructure support, offering greater autonomy and safety to users.

ACKNOWLEDGEMENTS

The research project underlying this contribution was funded by the Federal Ministry of Education and Research under the grant number FKZ 03WIR3122A. The responsibility for the content of this publication lies with the author.

REFERENCES

[1] Statistisches Bundesamt (Destatis), “4 million more people aged 67 or over will live in Germany in 2035,” Press Release No. 511, 2022. [Online]. Available: https://www.destatis.de/EN/Press/2022/12/PE22_511_124.html. [Accessed: May 2025].

[2] Translationsregion für digitalisierte Gesundheitsversorgung, “Aktiv im Alter durch Multisensorische Umfeldwahrnehmung – pflege- und versorgungswissenschaftlich evaluative Begleitung der co-kreativen Entwicklung.” [Online]. Available: https://inno-tdg.de/projekte/aktimuw.

[3] I. Omeko, S. Twieg, and M. Obert, “Sensor-based gait analysis: A comparative study of ultrasonic and laser sensors for gait monitoring in rollator-assisted walking,” Proc. Int. Conf. Applied Innovation in IT, vol. 13, pp. 213–222, 2025, doi: 10.25673/119236.

[4] T. Fang, F. Xu, and Z. Zou, “Causal factors in elderly pedestrian traffic injuries based on association analysis,” Applied Sciences, vol. 15, no. 3, 2025, doi: 10.3390/app15031170.

[5] I. Abaspur Kazerouni, L. Fitzgerald, G. Dooly, and D. Toal, “A survey of state-of-the-art on visual SLAM,” Expert Systems with Applications, vol. 205, Art. no. 117734, 2022, doi: 10.1016/j.eswa.2022.117734.

[6] V. Kurumbaparambil, S. Rajanayagam, and S. Twieg, “Modular robotic reinforcement learning platform for object manipulation,” Proc. Int. Conf. Applied Innovation in IT, vol. 12, no. 1, pp. 213–224, 2024, doi: 10.25673/115703.

[7] M. Labbé and F. Michaud, “RTAB-Map as an open-source lidar and visual SLAM library for large-scale and long-term online operation,” Journal of Field Robotics, pp. 416–446, 2019, doi: 10.1002/rob.21831.

[8] S. Macenski, T. Foote, B. Gerkey, C. Lalancette, and W. Woodall, “Robot Operating System 2: Design, architecture, and uses in the wild,” Science Robotics, vol. 7, no. 66, 2022, doi: 10.1126/scirobotics.abm6074.

[9] P. Panteleris and A. A. Argyros, “Vision-based SLAM and moving objects tracking for the perceptual support of a smart walker platform,” Lecture Notes in Computer Science, pp. 407–423, 2015, doi: 10.1007/978-3-319-16199-0_29.

[10] C. Feltner, J. Guilbe, S. Zehtabian, S. Khodadeh, L. Boloni, and D. Turgut, “Smart walker for the visually impaired,” in Proc. IEEE Int. Conf. Communications, 2019, doi: 10.1109/ICC.2019.8762081.

[11] S. Tareen and R. Raza, “Potential of SIFT, SURF, KAZE, AKAZE, ORB, BRISK, AGAST, and 7 more algorithms for matching extremely variant image pairs,” in Proc. 4th Int. Conf. Computing, Mathematics and Engineering Technologies, 2023, doi: 10.1109/iCoMET57998.2023.10099250.

[12] P. F. Alcantarilla and A. Bartoli, “KAZE features,” Lecture Notes in Computer Science, vol. 7577, 2012, doi: 10.1007/978-3-642-33783-3_16.

[13] A. Pahlevani, “Which SLAM algorithm should I choose,” Slambotics, 2025. [Online]. Available: https://www.slambotics.org/blog/which-slam-algorithm-should-i-choose. [Accessed: May 2025].

[14] T. Qin, P. Li, and S. Shen, “VINS-Mono: A robust and versatile monocular visual-inertial state estimator,” IEEE Transactions on Robotics, 2018, doi: 10.1109/TRO.2018.2853729.

[15] P. Kaveti, A. Gupta, D. Giaya, Z. Zhang, and H. Singh, “Challenges of indoor SLAM: A multi-modal multi-floor dataset for SLAM evaluation,” in Proc. IEEE Int. Conf. Automation Science and Engineering, 2023, doi: 10.1109/CASE56687.2023.10260618.

[16] M. Filipenko and I. Afanasyev, “Comparison of various SLAM systems for mobile robot in an indoor environment,” 2025, doi: 10.48550/arXiv.2501.09490.

[17] T. Foote and M. Purvis, “REP 103: Standard units of measure and coordinate conventions,” ROS.org, 2010. [Online]. Available: https://www.ros.org/reps/rep-0103.html. [Accessed: May 2025].