ICRA2026 Late Breaking Results Poster presented at
2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

BIM-Informed Visual SLAM for Construction Monitoring

Asier Bikandi-Noya[1], Miguel Fernandez-Cortizas[1], Muhammad Shaheer[1],
Ali Tourani[1], Holger Voos[1], and Jose Luis Sanchez-Lopez[1]

***Abstract*— Monitoring construction sites requires comparing the *as-planned* design with the *as-built* state in real time. Visual SLAM offers a lightweight solution but is prone to trajectory drift in construction environments due to repetitive layouts, textureless surfaces, and occlusions. We augment an existing visual SLAM system with structural priors from the Building Information Model (BIM), associating detected walls with their BIM counterparts and including these correspondences as geometric constraints in the back-end optimization. The system operates in real time and is validated on multiple real construction sites, achieving 25.23% average trajectory error reduction and 7.14% map accuracy improvement over state-of-the-art baselines, with demonstrated resilience to incomplete BIM data and *as-planned/as-built* discrepancies.**

I. Introduction

Construction site monitoring requires estimating the as-built state and comparing it with the as-planned BIM design [1]. Visual SLAM offers a scalable, cost-effective solution using standard RGB-D sensors. However, construction environments challenge visual SLAM severely: repetitive layouts, large textureless walls, and occlusions introduce ambiguous observations that accumulate as trajectory drift, producing geometrically inconsistent maps that are unreliable for monitoring [2].

BIM encodes the complete structural layout from the earliest design phase, a natural source of architectural priors to constrain the SLAM and bound drift. While such integration has been explored in LiDAR-based systems, visual SLAM method with embedded BIM priors to maintain a drift-bounded *as-built* map for construction monitoring remains a challenge. The main contributions of this work are:

- A novel integration of architectural BIM priors into a visual SLAM framework, reducing trajectory drift by enforcing structural consistency between the *as-built* map and the *as-planned* BIM.
- A wall-based initialization and association strategy that establishes correspondences between *as-planned* and *as-built* walls using only two walls as prior information, enabling monitoring from the earliest stages of operation without prior environment exploration.

[1]Authors are with the Automation and Robotics Research Group, SnT, University of Luxembourg. {asier.bikandi, muhammad.shaheer, miguel.fernandez, ali.tourani, holger.voos, joseluis.sanchezlopez}@uni.lu
This work was partially funded by the FNR Luxembourg under project BRIDGES/2025-1/IS/19685965/BARCODE.

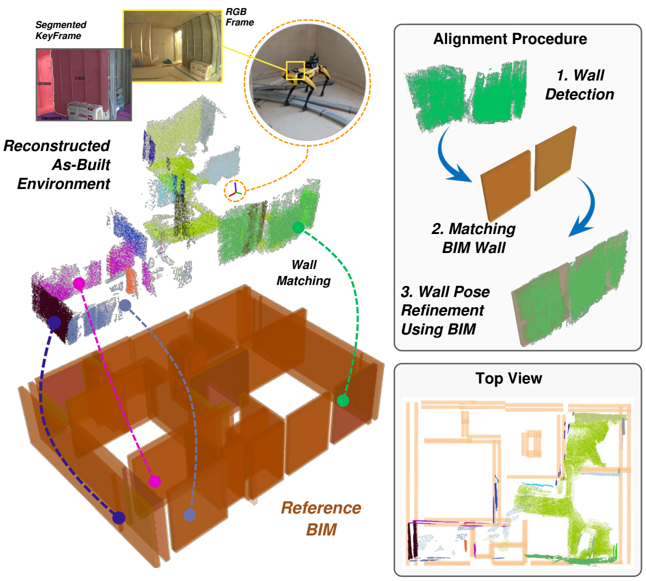

Fig. 1: Proposed BIM-informed RGB-D SLAM system. The figure shows the *as-planned* BIM and the reconstructed *as-built* map, with detected walls matched to their BIM counterparts. A 2D projection illustrates the correct alignment between the SLAM map and the BIM layout.

II. Methodology

A. System Overview

ivS-Graphs extends vS-Graphs [3], an RGB-D SLAM backbone with a hierarchical factor graph back-end, with three BIM-informed modules: (1) initial BIM-SLAM alignment, (2) continuous wall association, and (3) BIM-constrained back-end integration. Both detected and BIM-extracted walls share a unified representation encoding plane parameters (Hessian normal form), centroid, length, and thickness, enabling direct geometric comparison across coordinate frames.

B. Initial Alignment

Before BIM constraints can be applied, the two coordinate frames must be aligned. The only prior required is the specification of two BIM wall IDs expected to be observed first, id_{A_k} and id_{A_l}. The system waits until two detected walls (w_{S_i}, w_{S_j}) with nearly perpendicular supporting planes are found, matches them to w_{A_k} and w_{A_l}, and estimates the transformation ${}^{S}\hat{T}_A \in \mathrm{SE}(3)$ by minimizing plane residuals via least squares. All BIM walls $\mathbf{W_A}$ are then transformed to $\mathcal{S}$ and stored in the Atlas Manager for subsequent continuous association. This design enables immediate operation without requiring extensive prior exploration.

C. Continuous Wall Association

After initialization, each newly detected wall is matched to its best BIM counterpart using a combined score of two complementary metrics: (i) **plane-parameter distance (PPD)**, which measures geometric alignment between infinite planes; and (ii) **lateral centroid distance (LCD)**, which measures the tangential displacement between finite wall centroids, preventing cross-room mismatches caused by accumulated drift. A normalized weighted score combines both metrics, and only candidate pairs within configurable thresholds are accepted.

D. BIM-Constrained Back-End

Matched wall pairs are added as fixed–optimizable edges in the factor graph: BIM walls are fixed nodes representing the trusted *as-planned* reference, while detected walls remain optimizable. The wall-to-wall cost is weighted by the inverse of the match confidence, implementing an **uncertainty-aware formulation** where well-aligned walls exert stronger corrections and ambiguous matches are automatically down-weighted. A Huber kernel further suppresses outlier associations from partial structures or occlusions.

III. Experimental Evaluation

A. Dataset

Since no public dataset combines RGB-D data with the corresponding BIM, we collected our own sequences. RGB-D data was captured with an Intel RealSense D435i and ground-truth trajectories were generated offline from Ouster OS0-64 LiDAR scans processed with S-Graphs [4]. The dataset spans 11 sequences across 2 office buildings and 3 construction sites at different completion stages, totalling 642 m of trajectory length and approximately 62 min of duration, with 5 runs per sequence.

B. Baselines

We evaluate against ORB-SLAM3 [5], BAD-SLAM [6], and DROID-SLAM [7]. Comparison with vS-Graphs [3] directly isolates the effect of adding BIM constraints, as our system builds on it as backbone.

C. Trajectory Estimation

ivS-Graphs reduces mean ATE by **25.23%** over vS-Graphs, with improvements in 10 out of 11 sequences. Gains correlate with trajectory length: sequences exceeding 50 m show an average improvement of 30.54%, confirming that BIM constraints primarily combat long-range drift. The single degraded sequence (office1-4, -33.33%) is the shortest trajectory (37.69 m) where drift is already minimal. BAD-SLAM failed in 6 out of 11 sequences due to sensitivity to textureless surfaces; DROID-SLAM, while more robust, still exhibits significant drift.

D. Map Accuracy

ivS-Graphs achieves the lowest point cloud RMSE in 9 out of 11 sequences, with improvements over vS-Graphs ranging from +0.99% to +17.88%, averaging **7.14%**. Gains are largest in bigger environments where drift-induced map distortion is more pronounced.

E. Robustness Analysis

With **30% of BIM walls missing**, the system incurs only a 5.3% mean ATE increase, as remaining walls continue providing effective drift correction. Under **moderate geometric deviations** ($\Delta d = 0.2$ m, $\Delta\theta = 5°$), the match rate remains 88.0% and ATE increases by 25.8%. Under severe deviations ($\Delta d = 0.8$ m, $\Delta\theta = 15°$), association thresholds trigger rejection and the system gracefully falls back to the visual baseline, degrading predictably rather than catastrophically.

F. Runtime

ivS-Graphs operates at an average of **23.3 FPS**, above the 20 FPS real-time threshold. The initial alignment completes in 0.241 ms. The BIM back-end integration adds minimal overhead compared to the vS-Graphs baseline (24.2 FPS).

IV. Conclusions

We presented *ivS-Graphs*, a visual SLAM system augmented with BIM architectural priors for construction monitoring. By embedding BIM wall correspondences as uncertainty-aware geometric constraints in the factor graph back-end, the system achieves 25.23% ATE reduction and 7.14% map RMSE improvement over the visual SLAM backbone, with robustness to incomplete and imprecise BIM data. Future work will extend the approach to additional BIM elements and investigate automatic initialization strategies.

References

[1] M. Shaheer, J. A. Millan-Romera, H. Bavle, J. L. Sanchez-Lopez, J. Civera, and H. Voos, "Graph-based global robot localization informing situational graphs with architectural graphs," in 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2023, pp. 9155–9162.

[2] T. Sun, Y. Hao, S. Huang, S. Savarese, K. Schindler, M. Pollefeys, and I. Armeni, "Nothing stands still: A spatiotemporal benchmark on 3d point cloud registration under large geometric and temporal change," ISPRS Journal of Photogrammetry and Remote Sensing, vol. 220, pp. 799–823, 2025.

[3] A. Tourani, S. Ejaz, H. Bavle, D. Morilla-Cabello, J. L. Sanchez-Lopez, and H. Voos, "vs-graphs: Integrating visual slam and situational graphs through multi-level scene understanding," arXiv preprint arXiv:2503.01783, 2025.

[4] H. Bavle, J. L. Sanchez-Lopez, M. Shaheer, J. Civera, and H. Voos, "S-graphs+: Real-time localization and mapping leveraging hierarchical representations," IEEE Robotics and Automation Letters, vol. 8, no. 8, pp. 4927–4934, 2023.

[5] C. Campos, R. Elvira, J. J. G. Rodríguez, J. M. M. Montiel, and J. D. Tardós, "Orb-slam3: An accurate open-source library for visual, visual–inertial, and multimap slam," IEEE Transactions on Robotics, vol. 37, no. 6, pp. 1874–1890, 2021.

[6] T. Schöps, T. Sattler, and M. Pollefeys, "Bad slam: Bundle adjusted direct rgb-d slam," in 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 134–144.

[7] Z. Teed and J. Deng, "Droid-slam: Deep visual slam for monocular, stereo, and rgb-d cameras," Advances in neural information processing systems, vol. 34, pp. 16 558–16 569, 2021.