Structured Semantic 3D Reconstruction (S23DR) Challenge 2025 Write-up

Yuliia Moliashcha
Ukrainian Catholic University
Lviv, Ukraine
moliashcha.pn@ucu.edu.ua

Abstract

This report details our 3rd place solution for the Structured Semantic 3D Reconstruction (S23DR) Challenge 2025. We improve upon the official baseline by expanding the 2D semantic feature detection to include more architectural elements, and by introducing a geometric validation step that filters implausible 3D edges against the sparse COLMAP point cloud. This approach produces more geometrically accurate wireframes, achieving a 45% relative improvement on the mean Hybrid Structure Score on the private leaderboard.

1. Introduction

The reconstruction of structured, semantic 3D models is a key challenge in computer vision. The S23DR Challenge 2025 [1] addresses this by tasking participants with generating 3D wireframes from posed images.

Our approach, which secured 3rd place, enhances the provided baseline solution. We focused on improving two key areas: first, by extending the 2D feature detector to recognize a richer set of architectural semantics, and second, by building a validation module to ensure all 3D edges are geometrically consistent with the sparse SfM point cloud. This report details our methodology and presents an analysis of our results.

2. Methodology

Our method enhances the handcrafted baseline pipeline. The core idea is to process each image view to generate 2D wireframe proposals, lift them into 3D space, and then merge and validate them into a final model. Our contributions focus on improving the initial feature detection and enforcing geometric consistency on the final 3D graph.

2.1. Expanded Semantic Feature Detection

The baseline constructs a 2D wireframe from `gestalt` maps by identifying vertices from `apex` and `eave_end_point` classes. To create a more detailed initial graph, we expanded the vocabulary of detected semantic features. We added `flashing_end_point` to the set of detectable vertices. The list of line-like classes used for edge fitting was also augmented to include `flashing`, `hip`, `step_flashing`, and `transition_line`. This allows the model to capture finer architectural details from the segmentation data.

2.2. Geometric Edge Validation

One of the limitations of the baseline is the potential for ”phantom” edges – edges that appear valid in a 2D projection but are geometrically implausible in 3D. To address this, we introduced a geometric validation module that filters edge proposals by verifying their support against the high-fidelity COLMAP point cloud.

The process begins by constructing a KD-Tree from the COLMAP points for efficient nearest-neighbor searches. For each candidate edge, we sample N equidistant points along its vector and query the KD-Tree to find the distance from each sample point to its closest neighbor in the point cloud. An edge is accepted into the final wireframe only if a minimum of k sample points fall within a distance threshold τ of the cloud; otherwise, it is pruned as geometrically unsupported. In our final submission, we used parameters $N = 3$, $k = 3$, and $\tau = 0.5$m.

2.3. Robust Merging

To aggregate the 3D vertices generated from all views, we replaced the baseline's merging logic with a more robust Disjoint Set Union (DSU) algorithm. This correctly handles transitive relationships (i.e., if A is close to B, and B is close to C, all three are merged), ensuring each vertex cluster is coherent and preventing duplicates.

2.4. Parameter Tuning

Systematic experimentation led us to refine several key hyperparameters. To improve the precision of the final 3D model, we tightened the vertex merge radius to 0.4m (from 0.5m) and reduced the final outlier pruning distance to 3.0m (from 4.0m). To make the initial proposal stage more inclusive, we increased the 2D edge threshold to 25.0 (from 10.0), ensuring a wider set of candidate connections for our 3D validation module to filter.

3. Experiments and Results

Our method was evaluated against the official baseline on the private leaderboard. The primary metric for the challenge is the Hybrid Structure Score (HSS), which is a composite of vertex F1-score and edge IoU. A comparison of the performance is presented in Table 1.

Table 1. Performance comparison of our final method against the official baseline on the private leaderboard.

Metric	Baseline	Ours	Rel. Imp. (%)
hss_mean	0.264	**0.3827**	**+45.0%**
hss_q5	0	0	0%
hss_q25	0.1765	**0.2985**	**+69.1%**
hss_q50	0.2732	**0.4044**	**+48.0%**
hss_q75	0.3604	**0.4943**	**+37.2%**
hss_q95	0.4897	**0.5934**	**+21.2%**
corner_f1_mean	0.3844	**0.483**	**+25.7%**
edge_iou_mean	0.2095	**0.3247**	**+55.0%**

Our final method achieves a mean HSS of 0.3827, which is a 45.0% relative improvement over the baseline. This performance increase is primarily driven by a 55.0% gain in `edge_iou_mean`, which we attribute to the geometric validation module's ability to filter incorrect edges. Additionally, the `corner_f1_mean` increased by 25.7%, indicating an improvement in vertex accuracy.

The `hss_q5` score of zero shows that our method, like the baseline, does not produce a scoring output for the most challenging 5% of scenes. This suggests a lack of robustness on certain edge cases and identifies a clear area for future improvement.

4. Conclusion

In this report, we detailed our 3rd place solution for the S23DR Challenge. Our method improved upon the baseline by combining an expanded semantic feature vocabulary with a 3D geometric validation module. This strategy yielded a 45% relative improvement in HSS.

References

[1] Jack Langerman, Dmytro Mishkin, and Huang Yuzhong. S23dr competition at 2nd workshop on urban scene modeling @ cvpr 2025. https://huggingface.co/usm3d, 2025. 1