ELite++: Efficient Lifelong LiDAR Mapping via Window-Based Registration and Voxel Ratio Maps

Geonhui Lee[1], Hyeonjae Gil[1], Ayoung Kim[1] and Minwoo Jung[1*]

[1]Department of Mechanical Engineering, Seoul National University,
Seoul, 08826, Korea ([geongeon, h.gil, ayoungk, moonshot]@snu.ac.kr)

Abstract: Long-term LiDAR mapping requires clean and consistent static maps across repeated sessions in dynamic environments. Existing lifelong mapping pipelines combine session alignment, dynamic filtering, and map update, but repeated dense point-wise updates and candidate-dependent alignment can limit scalability. To tackle these issues, we present ELite++, an efficient extension of ELite that combines automatic window-based registration, aligned-session dynamic object removal, and compact Voxel Ratio Map integration. ELite++ removes transient objects while preserving static structures through hit-exposure voxel stability, local ground protection, and structure-aware refinement. Experiments on diverse LiDAR datasets show improved dynamic object removal over existing DOR methods, preserved multi-session alignment quality, and runtime reduced to 11% of the original pipeline.

Keywords: Lifelong mapping, dynamic object removal, multi-session registration, static map maintenance.

1. INTRODUCTION

Accurate localization and consistent map maintenance are fundamental for autonomous navigation, and long-term operation further requires maps to adapt to environmental changes over time [1, 2]. In real-world LiDAR mapping, repeated sessions often contain odometry drift, partial overlap, and scene changes from moving or temporarily parked objects; directly accumulating them leaves transient artifacts that degrade localization and spatial analysis. Lifelong mapping addresses this by aligning new sessions to an existing map and maintaining persistent static structures over time [3, 4].

Achieving this goal requires not only aligning sessions, but also deciding which observations should remain in the map. Although multi-session SLAM and map-merging systems improve alignment across sessions [5–8], they do not explicitly maintain a clean static representation under dynamic scene changes. ELite [4] addresses this issue through local and global ephemerality, enabling short- and long-term transiency modeling. However, it still relies on reliable initial session-to-map associations, and repeated dense point-wise ephemerality updates can become costly as sessions accumulate. This motivates an efficient formulation that automatically verifies session alignments and accumulates long-term static evidence as compact voxel-level statistics.

Before such static evidence is committed to the maintained map, transient observations must be reliably filtered. Existing Dyanmic Object Removal (DOR) methods remove many dynamic traces by exploiting temporal consistency and visibility cues [9–14]. However, local visibility or voxel-wise decisions may still cause ground erosion, bottom residues, and fragmented misclassifications, which can accumulate as persistent artifacts in lifelong mapping. Thus, DOR should operate on aligned multi-view observations and support long-term static evidence accumulation across sessions.

To address these challenges, we propose ELite++, an efficient lifelong mapping framework for static map maintenance. ELite++ combines automatic window-based session registration, aligned-session DOR, and compact Voxel Ratio Map integration. The registration module verifies local query submaps, fuses trusted window-level poses, and interpolates low-confidence intervals, while the DOR module estimates hit-exposure voxel stability with ground-aware and point-level refinements. Instead of repeatedly updating dense point-wise ephemerality, Voxel Ratio Maps accumulate voxel-level hit and exposure statistics across sessions, reducing update time while preserving alignment quality and static map cleanliness. The main contributions of this paper are summarized as follows:

- We present ELite++, an efficient extension of ELite that automates window-based registration and maintains compact Voxel Ratio Maps, reducing alignment effort and enabling faster lifelong updates.
- We propose a ground-aware DOR module that reduces ground erosion, bottom residues, and static-structure loss, providing cleaner voxel evidence for robust map integration.
- We validate ELite++ on diverse LiDAR datasets, showing improved DOR performance, preserved alignment quality, and runtime reduced to 11% of the original pipeline for continuous lifelong mapping.

2. RELATED WORKS

2.1 Multi-Session Map Maintenance

Aligning point clouds from multiple robots or temporal sessions is commonly addressed as a multi-session SLAM or map-merging problem. Loop-closure-based methods, such as DiSCo-SLAM [5] and AutoMerge [6], detect inter-session overlaps and optimize the resulting constraints through pose graph optimization or global smoothing. Recent frameworks such as Multi-Mapcher [7] further reduce reliance on explicit loop closures using large-scale map-to-map registration and anchor node-based optimization. While these methods achieve effective inter-session alignment and globally consistent map construction, their main focus remains geometric consistency rather than static map maintenance under dynamic scene changes.

Lifelong mapping methods further consider repeated map updates in changing environments. LT-mapper [3] and ELite [4] address LiDAR-based lifelong mapping through modular map maintenance and ephemerality-based map updating, respectively, while LAMM [8] incorporates DOR into multi-session point-cloud map fusion. These frameworks show the importance of combining alignment and map cleaning, but their multi-stage processing can increase update time in large-scale or repeated-session settings. Moreover, maintaining a clean static map remains challenging when dynamic filtering suffers from visibility ambiguity, ground geometry, or residual artifacts near dynamic objects.

Building on this perspective, our framework treats multi-session registration as the first stage of static map maintenance rather than as an independent alignment objective. The proposed window-based registration aligns local query submaps to the maintained map, and the aligned session is subsequently used for spatio-temporal DOR. This coupling allows transient objects to be filtered before map integration, supporting cleaner lifelong map updates.

2.2 Dynamic Object Removal

DOR is essential for constructing static LiDAR maps, as moving vehicles, pedestrians, and temporarily observed objects can remain as ghost artifacts. Early occupancy- and visibility-based approaches discretize space and infer temporal consistency from sensor observations. OctoMap [9] introduced probabilistic octree-based occupancy mapping, while Peopleremover [10] removes dynamic objects through voxel traversal. Range-image and visibility-based methods, such as Removert [11], further construct static maps by removing dynamic traces and recovering falsely removed static points.

Recent methods improve static mapping with ground reasoning, pseudo-occupancy, instance-level reasoning, and online map consistency. ERASOR [12] identifies dynamic regions using egocentric pseudo-occupancy and ground-aware analysis, while ERASOR2 [13] extends this direction with instance-aware dynamic point rejection. Other approaches combine ground segmentation with occupancy mapping [15] or perform online dynamic removal through

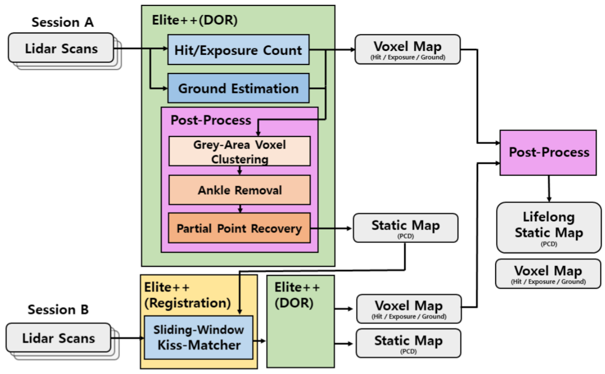

Fig. 1.: The overview pipeline of ELite++

scan-to-map and map-to-map reasoning [14]. Despite these advances, visibility- or voxel-based decisions can still suffer from ground erosion, bottom residues, and isolated misclassifications.

Accordingly, our DOR module addresses these failure cases through spatio-temporal voxel analysis on the aligned session. It combines repeated hit-exposure observations, normal-based ground protection, and cluster-level structural refinement to suppress transient objects while preserving persistent static structures before map integration.

3. METHOD

3.1 System Overview

The proposed lifelong mapping system updates an existing map using newly acquired LiDAR sessions. Following the general formulation of ELite [4], the system incrementally maintains a static map over repeated sessions. We focus on two components that directly affect long-term map quality: (i) window-based multi-session registration and (ii) DOR before map integration. Fig. 1 shows the overview pipeline of ELite++.

Let the t-th session be $\mathcal{S}_t = \{(\mathcal{P}_{t,i}, \mathbf{T}_{t,i})\}_{i=1}^{N_t}$, where $\mathcal{P}_{t,i}$ is the i-th LiDAR scan and $\mathbf{T}_{t,i} \in \mathrm{SE}(3)$ is its local pose. The registration module estimates aligned poses $\widetilde{\mathbf{T}}_{t,i}$ and constructs the aligned session map $\mathcal{M}_{S_t} = \bigcup_i \widetilde{\mathbf{T}}_{t,i}\mathcal{P}_{t,i}$. The DOR module removes dynamic regions from $\mathcal{M}_{S_t}$ and outputs the cleaned static session map $\widehat{\mathcal{M}}_{S_t,\mathrm{st}}$. The cleaned static session map and its voxel statistics form a session-level Voxel Ratio Map $\mathbf{V}_{S_t}$, which is integrated into the maintained voxel-statistic map $\mathbf{V}_t$.

3.2 Window-based Multi-session Registration

The registration module estimates a reliable session-to-map transformation under partial overlap, odometry drift, and dynamic scene changes. Since individual scans provide limited spatial context, we construct overlapping local query submaps and sequentially track their alignments to the existing map. Trusted window-level alignments are fused to obtain scan-wise poses, while low-confidence intervals are recovered through trajectory interpolation.

The previous map is voxel-downsampled to form a compact registration target $\mathcal{C}_{t-1}$. For the query session, overlapping temporal windows are defined as $\mathcal{W}_u = \{u, \ldots, u+L-1\}$, where L is the window length and the starting index u is sampled every r scans. The scans in each window are transformed into the coordinate frame of the first scan and aggregated as

$$\mathcal{Q}_u = \bigcup_{i \in \mathcal{W}_u} \mathbf{T}_{t,u}^{-1} \mathbf{T}_{t,i} \mathcal{P}_{t,i}. \tag{1}$$

This aggregation provides denser geometric constraints than a single scan and reduces boundary effects through window overlap.

Each query submap $\mathcal{Q}_u$ is sequentially matched against $\mathcal{C}_{t-1}$ using KISS-Matcher [16], while the current tracking prior $\mathbf{T}_{\text{init}}$ is updated from trusted windows and carried to the next window. KISS-Matcher estimates a candidate transformation $\mathbf{T}^u_{\mathcal{M} \leftarrow \mathcal{Q}}$ through feature-based global registration and robust outlier pruning. Since $\mathcal{Q}_u$ is represented in the local frame of scan u, the corresponding session-level candidate is $\mathbf{T}^u_{\text{abs}} = \mathbf{T}^u_{\mathcal{M} \leftarrow \mathcal{Q}} \mathbf{T}^{-1}_{t,u}$. The candidate is verified by a fitness score f_u, the ratio of transformed query points within τ_d of the target map. If $f_u \geq \tau_f$, the window is trusted and $\mathbf{T}_{\text{init}}$ is updated by $\mathbf{T}^u_{\text{abs}}$.

For each scan $\mathcal{P}_{t,i}$, we collect trusted window estimates covering the scan:

$$\mathcal{A}_i = \left\{ \mathbf{T}^u_{\text{abs}} \mathbf{T}_{t,i} \mid i \in \mathcal{W}_u,\ f_u \geq \tau_f \right\}. \tag{2}$$

If $\mathcal{A}_i$ is non-empty, the final pose $\widetilde{\mathbf{T}}_{t,i}$ is obtained by fusing the trusted estimates. Otherwise, it is recovered by distance-weighted interpolation from neighboring trusted poses. This scan-level fusion avoids reliance on a single selected window and maintains a continuous aligned trajectory across low-confidence intervals. The resulting aligned session is then passed to the DOR module for spatio-temporal analysis in a common coordinate frame.

3.3 Dynamic Object Removal

After registration, the aligned session may still contain dynamic or temporarily observed objects. We analyze it using a spatio-temporal voxel representation, where candidates are estimated from observation consistency and refined by ground protection, DBSCAN clustering, ankle verification, and intra-voxel restoration.

For each scan index i, neighboring aligned scans within a temporal radius k are aggregated into

$$\mathcal{Z}_{t,i} = \bigcup_{j=i-k}^{i+k} \widetilde{\mathbf{T}}_{t,j} \mathcal{P}_{t,j}. \tag{3}$$

The submap is discretized into voxels. For each voxel v, we accumulate a hit count $H(v)$ and an exposure count $E(v)$. The hit count measures direct point observations, while the exposure count is computed by raycasting along the sensor line-of-sight; voxels traversed before a surface return increase their exposure values. The voxel stability score is

$$R(v) = \frac{H(v)}{H(v) + E(v) + \eta}, \tag{4}$$

where η is a small constant. A high $R(v)$ indicates stable occupancy, whereas a low $R(v)$ indicates an unstable dynamic candidate.

Before dynamic classification, ground regions are protected. Ground anchors are extracted from points whose normals align with the vertical direction, *i.e.*, $|n_z| > \tau_{\text{slope}}$. The representative ground height h_g is estimated from these anchors, and voxels satisfying $|v_z - h_g| < \delta_{\text{ground}}$ are excluded from dynamic removal. This prevents visibility-based filtering from eroding road or terrain surfaces in sloped or sparsely observed regions. After excluding protected ground voxels, voxels with low stability scores are selected as dynamic candidates.

The remaining candidate voxels are grouped into spatial objects using DBSCAN on voxel centroids with search radius ϵ and minimum neighborhood size $N_{\min}$. For each cluster $\mathcal{C}_k$, the average stability score is computed as

$$\bar{R}(\mathcal{C}_k) = \frac{1}{|\mathcal{C}_k|} \sum_{v \in \mathcal{C}_k} R(v). \tag{5}$$

Clusters with $\bar{R}(\mathcal{C}_k) < \tau_{\text{cluster}}$ are labeled as dynamic. This object-level decision suppresses scattered noise and improves transient object removal over voxel-wise classification.

Finally, point-level refinement is applied to handle bottom residues and over-removed static structures. First, an ankle verification rule examines each protected ground voxel and its vertical upper neighbor v_{above}. If v_{above} is dynamic, points inside the ground voxel are re-evaluated by height; points satisfying $p_z > \bar{h}_v + \delta_{\text{tol}}$ are removed as lower-extremity dynamic residues, while the remaining ground points are preserved. If v_{above} is static or empty, all points in the ground voxel are locked as static.

Second, intra-voxel point restoration recovers sparse static structures enclosed by dynamic labels. Isolated dynamic voxels are identified using a 26-neighborhood adjacency search, and their point covariance is analyzed by PCA. The planarity score is defined as $\Pi_v = (\lambda_2 - \lambda_3)/\lambda_1$, where $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$ are the sorted eigenvalues. Each point p in an isolated dynamic voxel is restored with probability

$$P_{\text{rec}}(p) = \min\left(1,\ \frac{R(v)}{2}(1 + \Pi_v)\right). \tag{6}$$

This favors recovery of stable and planar structures such as walls or sparse static surfaces, while retaining

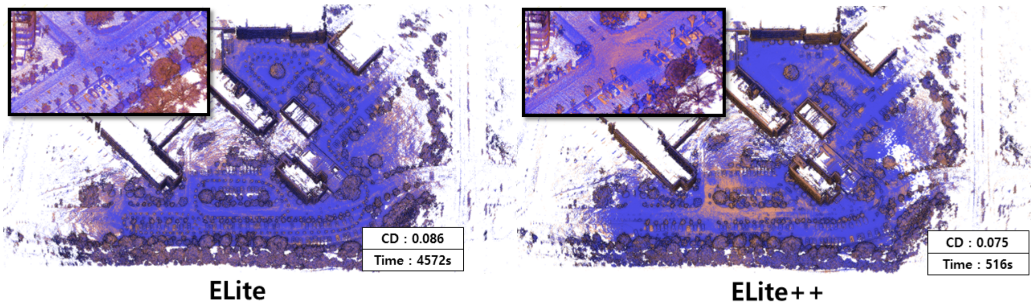

Fig. 2.: Multi-session static map construction on `ParkingLot01-02`. ELite++ maintains Chamfer-distance-based alignment quality comparable to ELite while reducing runtime. Zoomed regions show that ELite retains dynamic-object ghost artifacts, whereas ELite++ produces a cleaner static map.

the voxel-level dynamic label for downstream processing. After refinement, points in verified dynamic regions are removed from the session map:

$$\widehat{\mathcal{M}}_{\mathcal{S}_t,\mathrm{st}} = \mathcal{M}_{\mathcal{S}_t} \setminus \widehat{\mathcal{M}}_{\mathcal{S}_t,\mathrm{dyn}}. \tag{7}$$

3.4 Static Map Integration

The cleaned static representation is integrated using compact Voxel Ratio Maps rather than raw point-cloud fusion. Similar to the long-term evidence accumulation in ELite [4], our update maintains persistent observations across sessions. However, instead of point-wise global ephemerality, we store voxel-level hit and exposure statistics that represent occupancy stability. For each session $\mathcal{S}_t$, the system stores voxel keys, hit counts $H(v)$, exposure counts $E(v)$, and ground flags. Let $\mathbf{V}_{t-1}$ and $\mathbf{V}_{\mathcal{S}_t}$ denote the historical and current session voxel ratio maps, respectively.

For overlapping voxel keys, a density-adaptive decay factor balances historical evidence with the current observation. Given $N_{\mathcal{S}_t}(v) = H_{\mathcal{S}_t}(v) + E_{\mathcal{S}_t}(v)$, the historical statistics are scaled by

$$\alpha_v = \max\left(\alpha_{\min}, \min\left(\alpha_{\max}, \exp\left(-\gamma \frac{N_{\mathcal{S}_t}(v)}{N_{\max}}\right)\right)\right), \tag{8}$$

where $N_{\max}$ is the largest session intensity, γ controls decay sensitivity, and $\alpha_{\min}, \alpha_{\max}$ bound the update range. The voxel statistics are updated by

$$\begin{aligned} H_t(v) &= \alpha_v H_{t-1}(v) + H_{\mathcal{S}_t}(v), \\ E_t(v) &= \alpha_v E_{t-1}(v) + E_{\mathcal{S}_t}(v). \end{aligned} \tag{9}$$

The ground flag is updated from the current session to reflect the latest local terrain observation.

Because integration is performed in voxel-statistic space, the system avoids repeated global nearest-neighbor association or raw point-cloud fusion. The updated map $\mathbf{V}_t$ is then refined using the same voxel-level filtering rules, allowing dynamic regions and residual noise to be removed before generating the final static map.

Table 1.: Registration Results in Multi-Session Environment

Sequence	Method	Metrics		
		RMSE ↓	CD ↓	Time(s) ↓
`LT-ParkingLot(01-02)`	ELite	**0.076**	0.086	4572
	ELite++ (Ours)	0.084	**0.075**	**516**
`KAIST(01-04)`	ELite	0.106	**0.446**	12732
	ELite++ (Ours)	**0.103**	0.541	**1205**

4. EXPERIMENTS

We evaluate the proposed framework on multi-session alignment and DOR using diverse LiDAR datasets: KITTI (`KITTI00`, `KITTI05`) [17], Argoverse 2 (`AV2.0`) [18], LTMapper (`Parkinglot01-02`) [3], MulRan (`KAIST01`) [19], HeLiPR (`KAIST04`) [20], and CU-Multi (`Kittredge Loop`, `Main Campus`) [21]. They cover urban driving, campus-scale mapping, parking-lot environments, and multi-session or multi-robot scenarios.

For multi-session alignment, we report registration accuracy using RMSE and Chamfer Distance (CD). For DOR, we compare against ERASOR [12], Removert [11], OctoMap [9], and OctoMap with ground filtering, using dynamic accuracy (DA), static accuracy (SA), average accuracy (AA), and elapsed time. All experiments are conducted on a desktop with an AMD Ryzen 5 7500F CPU, GeForce RTX 4060 GPU and 32GB RAM.

4.1 Multi-session Experiments

We evaluate multi-session static map maintenance on `ParkingLot01-02` and `KAIST01-04`. As shown in Table. 1, ELite++ achieves RMSE and CD comparable to ELite, while reducing the processing time to about 11% of the original runtime. Considering that ELite requires an initial session-to-map candidate for alignment, these results show that ELite++ preserves alignment quality while automatically verifying local query submaps and fusing trusted window-level poses.

Furthermore, Fig. 2 qualitatively shows the static maps generated on `ParkingLot01-02`. ELite retains dynamic-object ghost artifacts, whereas ELite++ produces a cleaner static map by suppressing transient objects before integration. These results indicate that ELite++ improves automation and update

Table 2.: Quantitative comparison of DOR baselines

Method	KITTI00				KITTI05				AV2.0			
	SA(%) ↑	DA(%) ↑	AA(%) ↑	Time(s) ↓	SA(%) ↑	DA(%) ↑	AA(%) ↑	Time(s) ↓	SA(%) ↑	DA(%) ↑	AA(%) ↑	Time(s) ↓
Removert	**99.44**	41.53	64.26	36.86	**99.42**	22.28	47.06	182.94	**98.74**	34.10	58.02	440.43
ERASOR	66.70	98.54	81.07	**5.57**	69.40	99.06	82.92	**12.42**	77.51	**99.18**	87.68	508.66
OctoMap	68.05	**99.69**	82.37	194.19	66.28	**99.24**	81.10	535.60	69.04	97.50	82.04	5421.68
OctoMap w/ GF	86.09	98.84	92.25	212.38	86.55	97.94	91.88	570.60	77.88	86.18	81.92	5503.41
Elite++ (Ours)	97.86	98.24	**98.05**	66.02	98.39	94.47	**96.41**	114.83	94.76	88.90	**91.78**	**438.19**

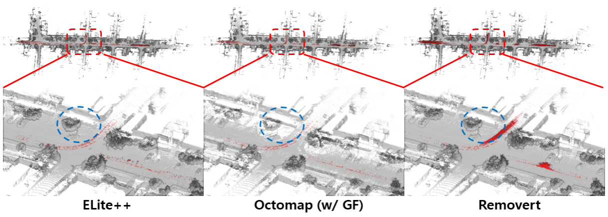

Fig. 3.: Qualitative DOR comparison. Red points indicate remaining dynamic objects, while the blue region highlights over-removed static structures. ELite++ removes dynamic traces while better preserving static regions.

efficiency while maintaining alignment quality. Detailed DOR performance against existing baselines is evaluated in the next section.

4.2 Dynamic Object Removal

As shown in Table. 2, ELite++ achieves the highest AA across all evaluated sequences. Compared with OctoMap w/ GF, which is the strongest existing baseline on the KITTI sequences, our method consistently achieves better overall accuracy. This shows that ELite++ provides a more balanced treatment of static and dynamic points. Removert preserves static objects well but fails to sufficiently remove dynamic objects, whereas ERASOR and OctoMap show the opposite tendency by improving DA while sacrificing SA. In contrast, ELite++ maintains both SA and DA at high levels, resulting in superior AA.

Fig. 3 further illustrates this behavior qualitatively. Removert leaves many unremoved dynamic objects, highlighted by red residual points, indicating insufficient dynamic filtering. OctoMap w/ GF removes dynamic objects more aggressively and achieves a removal level comparable to ELite++, but it also removes static regions that should be preserved, as shown in the blue highlighted area. By contrast, ELite++ suppresses dynamic traces while preserving nearby static structures, leading to a cleaner static map.

4.3 Multi-Robot Experiments

We evaluate ELite++ on the CU-Multi dataset to verify its applicability to multi-robot mapping scenarios. The dataset provides synchronized multi-run LiDAR sequences in outdoor campus environments, including `Main Campus` and `Kittredge Loop`. Fig. 4 shows qualitative mapping results. Despite different trajectories and partial overlaps across runs, ELite++ aligns the sessions into a consistent map. The zoomed regions show locally consistent structures in overlapping areas, indicating that the proposed window-based registration can provide reliable alignment for multi-robot static map maintenance.

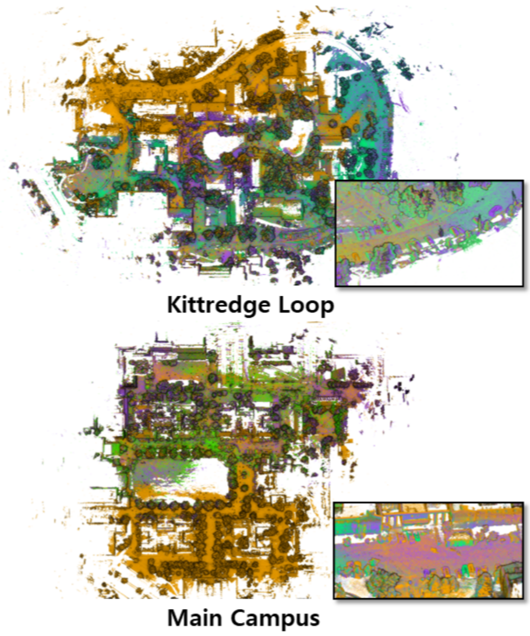

Fig. 4.: Qualitative multi-robot mapping results on the CU-Multi dataset. ELite++ aligns synchronized multi-run LiDAR sessions in `Main Campus` and `Kittredge Loop`. Different colors indicate different runs, and the zoomed regions show locally consistent structures in overlapping areas.

Table 3.: Ablation results for DOR

Recovery	Bottom	Clust.	SA(%) ↑	DA(%) ↑	AA(%) ↑
×	×	×	97.60	80.75	88.78
✓	×	×	**98.70**	80.74	89.27
✓	✓	×	98.67	82.25	90.09
✓	✓	✓	98.39	**94.47**	**96.41**

4.4 Ablation on Dynamic Object Removal

Table. 3 analyzes the role of each refinement in the proposed DOR module. The voxel-wise baseline relies on local observation stability, which preserves many static points but often leaves dynamic objects only partially removed. Static voxel recovery mitigates over-removal of persistent structures, while bottom-residue removal targets low-height dynamic parts that remain near the protected ground surface. Cluster-level classification then resolves fragmented voxel decisions by enforcing spatial consistency over connected regions. Together, these components balance static preservation and dynamic removal, leading to a cleaner static map with fewer residual artifacts.

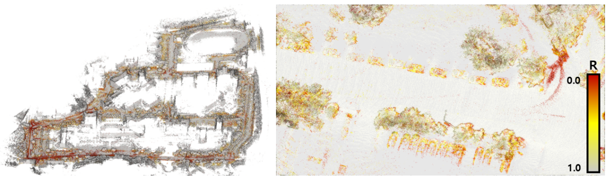

Fig. 5.: Long-term voxel evidence from KAIST01 and KAIST04. Both the full map and zoomed view visualize accumulated voxel stability scores. Low-stability regions correspond to intermittently observed objects, providing cues for dynamic-region filtering before final map generation.

4.5 Long-term Voxel Evidence

We qualitatively analyze the long-term voxel evidence accumulated by the proposed Voxel Ratio Map. Similar to the global ephemerality analysis in ELite [4], this experiment shows whether accumulated long-term evidence can indicate transient regions across multiple sessions. Unlike ELite, which maintains point-wise global ephemerality, our method represents long-term occupancy stability using voxel-level hit and exposure statistics.

Fig. 5 shows the result constructed from KAIST01 and KAIST04. The visualization demonstrates that intermittently observed regions, such as parked or moving vehicles, obtain low stability scores, while persistent static structures remain highly stable. This indicates that Voxel Ratio Maps can accumulate long-term static evidence and provide cues for dynamic-region filtering before final map generation.

5. CONCLUSION

This paper presented ELite++, an efficient lifelong LiDAR mapping framework for static map maintenance. ELite++ combines automatic window-based registration, aligned-session DOR, and compact Voxel Ratio Map integration to reduce update cost while preserving map quality. Experiments on diverse LiDAR datasets show improved DOR performance over existing baselines, maintained alignment quality, and runtime reduced to 11% of the original pipeline. These results demonstrate the practicality of ELite++ for continuous lifelong mapping in dynamic environments.

REFERENCES

[1] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, and J. J. Leonard, "Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age," *IEEE Trans. Robot.*, vol. 32, no. 6, pp. 1309–1332, 2017.

[2] R. B. Sousa, H. M. Sobreira, and A. P. Moreira, "A systematic literature review on long-term localization and mapping for mobile robots," *J. of Field Robot.*, vol. 40, no. 5, pp. 1245–1322, 2023.

[3] G. Kim and A. Kim, "Lt-mapper: A modular framework for lidar-based lifelong mapping," in *Proc. IEEE Intl. Conf. on Robot. and Automat.*, 2022, pp. 7995–8002.

[4] H. Gil, D. Lee, G. Kim, and A. Kim, "Ephemerality meets LiDAR-based lifelong mapping," in *Proc. IEEE Intl. Conf. on Robot. and Automat.*, 2025.

[5] Y. Huang, T. Shan, F. Chen, and B. Englot, "Disco-slam: Distributed scan context-enabled multi-robot lidar slam with two-stage global-local graph optimization," *IEEE Robot. and Automat. Lett.*, vol. 7, no. 2, pp. 1150–1157, 2021.

[6] P. Yin, S. Zhao, H. Lai, R. Ge, J. Zhang, H. Choset, and S. Scherer, "Automerge: A framework for map assembling and smoothing in city-scale environments," *IEEE Trans. Robot.*, vol. 39, no. 5, pp. 3686–3704, 2023.

[7] H. Lim, D. Kim, and H. Myung, "Multi-mapcher: Loop closure detection-free heterogeneous lidar multi-session slam leveraging outlier-robust registration for autonomous vehicles," *IEEE Trans. Intell. Vehicles*, 2025.

[8] H. Wei, R. Li, Y. Cai, C. Yuan, Y. Ren, Z. Zou, H. Wu, C. Zheng, S. Zhou, K. Xue *et al.*, "Large-scale multi-session point-cloud map merging," *IEEE Robot. and Automat. Lett.*, vol. 10, no. 1, pp. 88–95, 2024.

[9] A. Hornung, K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, "Octomap: An efficient probabilistic 3d mapping framework based on octrees," *Autonomous Robots*, vol. 34, no. 3, pp. 189–206, 2013.

[10] J. Schauer and A. Nüchter, "The peopleremover—removing dynamic objects from 3-d point cloud data by traversing a voxel occupancy grid," *IEEE Robot. and Automat. Lett.*, vol. 3, no. 3, pp. 1679–1686, 2018.

[11] G. Kim and A. Kim, "Remove, then revert: Static point cloud map construction using multiresolution range images," in *Proc. IEEE/RSJ Intl. Conf. on Intell. Robots and Sys.*, 2020, pp. 10 758–10 765.

[12] H. Lim, S. Hwang, and H. Myung, "Erasor: Egocentric ratio of pseudo occupancy-based dynamic object removal for static 3d point cloud map building," in *IEEE Robot. and Automat. Lett.*, vol. 6, no. 2, 2021, pp. 2272–2279.

[13] H. Lim, L. Nunes, B. Mersch, X. Chen, J. Behley, H. Myung, and C. Stachniss, "Erasor2: Instance-aware robust 3d mapping of the static world in dynamic scenes." in *Proc. Robot.: Science & Sys. Conf.*, 2023.

[14] T. Fan, B. Shen, H. Chen, W. Zhang, and J. Pan, "Dynamicfilter: an online dynamic objects removal framework for highly dynamic environments," in *Proc. IEEE Intl. Conf. on Robot. and Automat.*, 2022, pp. 7988–7994.

[15] M. Arora, L. Wiesmann, X. Chen, and C. Stachniss, "Mapping the static parts of dynamic scenes from 3d lidar point clouds exploiting ground segmentation," in *Proc. European Conf. on on Mob. Robots*. IEEE, 2021, pp. 1–6.

[16] H. Lim, D. Kim, G. Shin, J. Shi, I. Vizzo, H. Myung, J. Park, and L. Carlone, "Kiss-matcher: Fast and robust point cloud registration revisited," in *Proc. IEEE Intl. Conf. on Robot. and Automat.*, 2025, pp. 11 104–11 111.

[17] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The kitti dataset," *Intl. J. of Robot. Research*, vol. 32, no. 11, pp. 1231–1237, 2013.

[18] B. Wilson, W. Qi, T. Agarwal, J. Lambert, J. Singh, S. Khandelwal, B. Pan, R. Kumar, A. Hartnett, J. K. Pontes, D. Ramanan, P. Carr, and J. Hays, "Argoverse 2: Next generation datasets for self-driving perception and forecasting," in *Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks*, 2021.

[19] G. Kim, Y. S. Park, Y. Cho, J. Jeong, and A. Kim, "Mulran: Multimodal range dataset for urban place recognition," in *Proc. IEEE Intl. Conf. on Robot. and Automat.*, 2020, pp. 6246–6253.

[20] M. Jung, W. Yang, D. Lee, H. Gil, G. Kim, and A. Kim, "Helipr: Heterogeneous lidar dataset for inter-lidar place recognition under spatiotemporal variations," *Intl. J. of Robot. Research*, vol. 43, no. 12, pp. 1867–1883, 2024.

[21] D. Albin, M. Mena, A. Thomas, H. Biggie, X. Sun, D. Woods, S. McGuire, and C. Heckman, "Cu-multi: A dataset for multi-robot data association," *arXiv preprint arXiv:2505.17576*, 2025.