Offline Model-Based Reinforcement Learning for Active SLAM Using Sparse Point Cloud Registration Rewards

M. Briliauskas(1)

(1) Vilnius University, Institute of Data Science and Digital Technologies

In this work, we study active SLAM decision-making using offline model-based deep reinforcement learning on sparse reconstructions generated by SLAM systems such as ORB-SLAM2 [1], ORB-SLAM3 [2], and DPVO [3]. Our contribution is threefold: we propose an offline model-based DRL approach operating on sparse point clouds, present a reward formulation derived from sparse geometric reconstructions, and introduce a partial-overlap registration algorithm for sparse point clouds.

Active SLAM requires selecting robot actions that improve mapping and localization during exploration of unknown environments. This is especially important in safety-critical and GPS-denied settings, where learning must rely on previously collected data rather than extensive online interaction. While reinforcement learning has increasingly been considered for such tasks, many existing approaches rely on dense environment representations, handcrafted heuristics, or direct online interaction with the environment. This motivates the use of offline model-based DRL, where policies are improved from previously collected transitions.

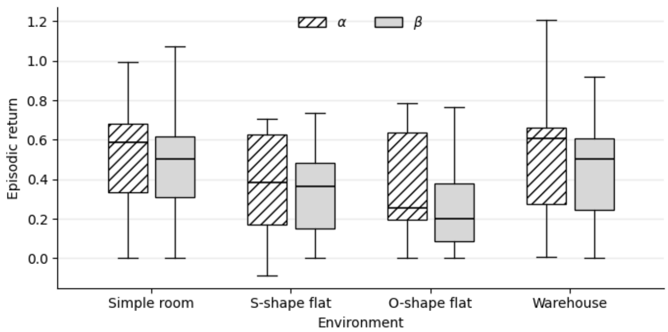

Figure 1: Distribution of episodic returns for the learned policy α and uninformed random-action baseline β across four simulated indoor environments.

In the proposed approach, sparse point cloud reconstructions are used as the basis for decision-making in active SLAM. Although directly available from modern visual SLAM pipelines, this representation is challenging due to sparsity, partial registration overlap, and reconstruction noise. Preliminary experiments in four simulated indoor environments show (see Figure 1) that, even under limited offline data conditions (4468-34662 steps per environment), the learned policy consistently outperforms a random-action baseline. The observed return improvements range from 12.8% to 62.9%, with an average improvement of 22.7% across all tested environments.

[1] R. Mur-Artal et al., IEEE Trans.Robot., 33 (5), 1255–1262, 2017.

[2] C. Campos et al., IEEE Trans.Robot., 37 (6), 1874–1890, 2021.

[3] Z. Teed et al., Adv.Neural Inf.Process.Syst., 36, 39033–39051, 2023.