VR-Robo: A Real-to-Sim-to-Real Framework for Visual Robot Navigation and Locomotion

Shaoting Zhu[12*], Linzhan Mou[23*], Derun Li[24], Baijun Ye[13], Runhan Huang[12], Hang Zhao†[123]
*Equal contribution †Corresponding author VR-Robo.github.io

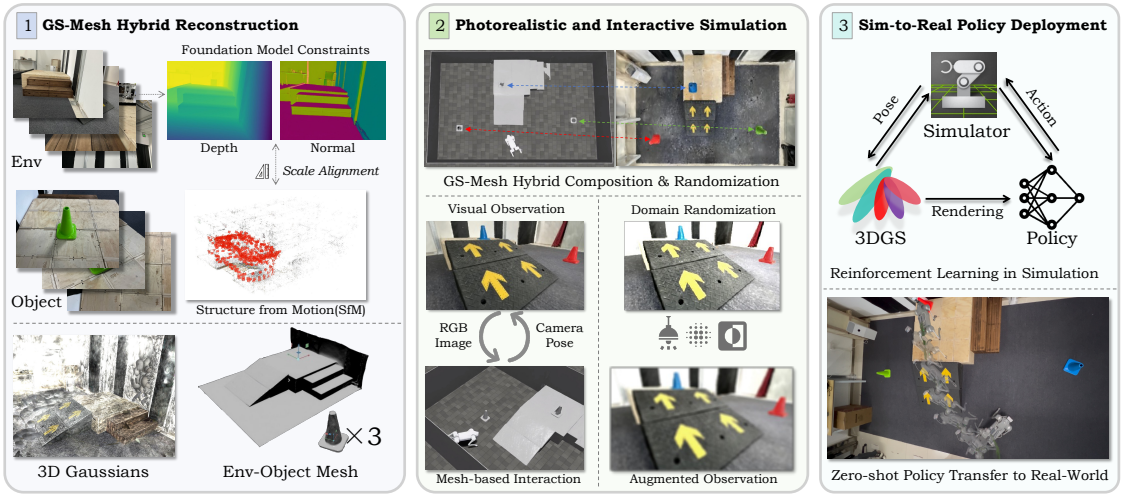

Figure 1. **VR-Robo *real-to-sim-to-real* framework.** We build a realistic and interactive simulation environment with GS-mesh hybrid representation and occlusion-aware composition & randomization for policy training. Finally, we zero-shot transfer the RL policy trained in simulation into the real robot for ego-centric navigation and visual locomotion.

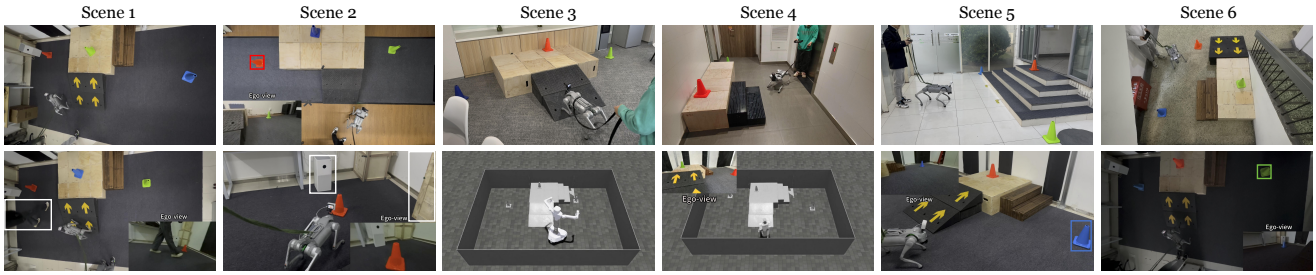

Figure 2. Diverse experiment settings including scenes, conditions, and robot types (better viewed when zoomed in).

Recent success in legged robot locomotion is attributed to the integration of reinforcement learning and physical simulators. However, these policies often encounter challenges when deployed in real-world environments due to sim-to-real gaps, as simulators typically fail to replicate visual realism and complex real-world geometry. Moreover, the lack of realistic visual rendering limits the ability of these policies to support high-level tasks requiring RGB-based perception like ego-centric navigation. This paper presents a *Real-to-Sim-to-Real* framework that generates photorealistic and physically interactive "digital twin" simulation environments for visual navigation and locomotion learning. Our approach leverages 3D Gaussian Splatting (3DGS) based scene reconstruction from multi-view images and integrates these environments into simulations that support ego-centric visual perception and mesh-based physical interactions. To demonstrate its effectiveness, we train a reinforcement learning policy within the simulator to perform a visual goal-tracking task. Extensive experiments show that our framework achieves RGB-only sim-to-real policy transfer. Additionally, our framework facilitates the rapid adaptation of robot policies with effective exploration capability in complex new environments, highlighting its potential for applications in households and factories.