Self-Supervised Metric-Scaled Monocular 3D Perception Framework for Low-Cost Autonomous Systems

Ali, Ahmed (School: Applied Technology School for Nuclear Energy in El-Dabaa)
Dahman, Mohamed (School: Applied Technology School for Nuclear Energy in El-Dabaa)

In the recent decade, autonomous systems like self-driving cars, drones, and robots have become widely used. They rely heavily on 3D perception of the surrounding environment and mapping technologies to navigate and interact effectively. These systems depend on high-precision range sensors like Lidar for accurate 3D awareness, which is unsuitable for low-cost systems that require lightness, affordability, and portability. Many deep-learning approaches for depth estimation and monocular mapping have been proposed, having limitations in scalability, requiring large amounts of labeled data, constraints in reliability, and providing relative scale-ambiguous depth maps. In this work, we introduce a metric-scaled monocular 3d perception framework that combines optimization-based fusion of local and global position sensors and deep-learning-based monocular visual odometry (VO). The proposed framework is able to train depth estimation models from unlabeled monocular RGB sequences and sensor fusion readings leveraging IMU, barometer, and GPS data. Our framework is able to recover metric depth from dense relative depth maps and correct errors in predicted depth maps. The experimental comparisons show significant improvement in our high resolution depth estimation model trained using the proposed framework and high-accuracy depth recovery at large distances, addressing the reliability and scalability of deep-learning-based 3d perception approaches for low-cost systems.

Awards Won: