Abstract

In today's rapidly advancing digital days, the boundaries between our physical surroundings and virtual spaces are becoming less distinct. This transition is significantly facilitated by the rise of Extended Reality (XR), revolutionizing the way we interact with both real and virtual environments. To realize an interactive and immersive XR experience, perceiving 3D spatial information and reconstructing real-world details are essential, since these two fundamental processes together digitize our physical world within XR systems. Although 3D vision research is active in current deep learning era, few techniques of 3D perception and reconstruction have been successfully utilized in XR applications given that: (i) acquiring specialized, labeled large-scale data for diverse XR scenarios is difficult; (ii) data-driven visual algorithms remain vulnerable to complex environmental variables (e.g., lighting) in XR development; (iii) computational cost of deep neural networks is hardly affordable to stand-alone XR devices; (iv) 3D sensors inevitably collect incomplete, irregular, or noisy data given occlusion and hardware restrictions; and (v) the synergy between 3D perception and reconstruction is rarely explored for XR utilization.

This project aims to develop an integrated framework for 3D perception and reconstruction in the context of XR. By deploying this framework, XR systems can be provided with high-fidelity digitized real-world context to achieve intelligent interactions and immersive experiences. Particularly, we intend to highlight the efficiency, robustness, and adaptability of our framework across various XR applicable scenarios such as healthcare, education, entertainment, etc. In general, the contributions and innovations of this project are in three main aspects:

- First, the framework will be modularized into functionally distinct components, including the key modules of input, 3D perception, 3D reconstruction, XR integration, optimization, utility, output, etc. By modularizing the framework, each component can be developed, tested, and optimized independently, enabling easy upgrades, scalability, and integration of new techniques and algorithms.
- Second, cutting-edge research on 3D perception and reconstruction related techniques will be performed. Particularly, data-efficient learning algorithms will be our research emphasis to overcome the data issue and enhance the robustness for XR applications. We will also explore interpretable 3D neural networks, understand their decision-making processes, and ensure their adaptability to different XR scenarios.
- Third, we will conduct comprehensive studies to systematically evaluate the framework's performances. By concluding this project, the framework will be delivered as an open-source product, which is not only technically robust but also user-friendly, widely accessible, and adaptable to various XR applications and platforms.

The PI of this project possesses a unique blend of practical know-how and theoretical expertise, given extensive hands-on experience in XR industry and a strong academic research record in 3D vision. Building on the PI's preliminary results of framework development and algorithm research, this project will become an invaluable resource for the XR community.