MultimodalStudio: A Heterogeneous Sensor Dataset and Framework for Neural Rendering across Multiple Imaging Modalities

Authors: Federico Lincetto[1], Gianluca Agresti[2], Mattia Rossi[2], Pietro Zanuttigh[1]
Affiliation: 1 - University of Padova, 2 – Sony Europe Limited

Abstract
Neural Radiance Fields (NeRF) have shown impressive performances in the rendering of 3D scenes from arbitrary viewpoints. While RGB images are widely preferred for training volume rendering models, the interest in other radiance modalities is also growing. However, the capability of the underlying implicit neural models to learn and transfer information across heterogeneous imaging modalities has seldom been explored, mostly due to the limited training data availability. For this purpose, we present MultimodalStudio (MMS): it encompasses MMS-DATA and MMS-FW. MMS-DATA is a multimodal multi-view dataset containing 32 scenes acquired with 5 different imaging modalities: RGB, monochrome, near-infrared, polarization and multispectral. MMS-FW is a novel modular multimodal NeRF framework designed to handle multimodal raw data and able to support an arbitrary number of multi-channel devices. Through extensive experiments, we demonstrate that MMS-FW trained on MMS-DATA can transfer information between different imaging modalities and produce higher quality renderings than using single modalities alone. We publicly release the dataset and the framework, to promote the research on multimodal volume rendering and beyond.