Reading Assignment (NeRF/KinectFusion)

NeRF

- **Contributions**: It introduces a novel dense 3D reconstruction method that leverages depth data from an RGB-D sensor to create highly detailed surface maps in real time. Its primary contribution is a system that tracks the camera pose while incrementally fusing depth data into a volumetric map, allowing for continuous scene reconstruction.
- **Methods**: The system employs voxel-based volumetric mapping using a Truncated Signed Distance Function (TSDF) to represent the 3D surface, enabling efficient data fusion and updating. It also uses ICP alignment to refine camera pose estimation based on the current depth data and the accumulated 3D model. Also, a raycasting method is used for fast surface extraction, allowing to interactively view and inspect the reconstructed map in real-time.

KinectFusion

- **Contributions**: NeRF introduced a novel approach to 3D scene representation by encoding scenes in a neural network that generates photo-realistic images from novel viewpoints, a key advancement in view synthesis. Its contribution lies in transforming 2D images into a 3D radiance field using a neural network, enabling accurate, continuous representations of complex scenes.
- **Methods**: NeRF represents the scene as a volumetric radiance field, using a fully connected neural network to learn the density and color at any 3D coordinate and viewing direction. The network is trained using volumetric rendering techniques, minimizing the difference between predicted and actual pixel colors in multiple training images. It leverages positional encoding to capture high-frequency details, improving accuracy in scene details and lighting variations. During rendering, a process known as ray marching is used to accumulate color and density along rays cast from the camera to synthesize new views. This approach enables NeRF to render complex scenes with realistic occlusions, reflections, and lighting effects.