3-D RECONSTRUCTION OF SCENES FROM 2-D IMAGES

Abdul Basit (20246646) -- (Supervised by: Dr Hedi Tabia)

ABSTRACT:

This paper presents recent advances in 3-D Reconstruction of scenes using only 2-D images. Traditional approaches rely on geometrical methods. Modern techniques leverage deep learning and involve techniques like NeRF, Gaussian splatting, Sparse Fusion, etc, which learn the 3-D structure directly from 2-D images through optimization or learning. This review serves as a foundation of knowledge for further research.

CONTEXT:

The reconstruction of 3-D models from 2-D images is a long-standing challenge in computer vision. It has applications in AR/VR, autonomous systems, robotics, and the metaverse. The previous methods require special equipment like Multiview geometry or depth sensors, which makes them complex and hardware dependent. But with the emergence of volume rendering techniques, the reconstruction of 3D novel models is now done through images.

METHODS:

NeRF (Neural Radiance Fields)

Introduced a method to represent 3D scenes as a continuous function using an MLP that outputs color and density. Used volume rendering with ray sampling to synthesize novel views from multiple 2D images.

INSTANT NERF

Replaced expensive positional encoding with multi-resolution hash grid encoding and a smaller MLP for faster training. Achieved real-time rendering and training in seconds using optimized CUDA implementations.

PIXEL NERF

Conditioned the NeRF model on pixel-aligned CNN features, enabling 3D reconstruction from a single or few images. Learned a generalizable scene-independence prior to removing the need for per-scene optimization.

NEURAL ANGELO

Introduced a method for extracting high-fidelity 3D surfaces from videos by combining neural rendering and surface tracking. Captures fine geometric details using structured neural representations.

GAUSSIAN SPLATTING

Used 3D Gaussians instead of volume rendering to represent and render scenes efficiently. Enabled real-time, photorealistic novel view synthesis without neural networks or ray marching.

SPARSE FUSION

Fused sparse view NeRF reconstruction with diffusion models to complete missing geometry and texture. Leveraged generative priority for high-quality 3D reconstruction from minimal input views.

COMPARATIVE ANALYSIS:

Model	Strengths	Weakness
NeRF	High-quality rendering, photorealism	Slow training, inference, and high memory
Pixel-NeRF	Works with a few images & fast inference.	Lower fidelity than NeRF with few views
Instant-NeRF	Extremely fast training & rendering	Limited to static scenes, resolution depends on hash levels
Neural Angelo	Produces detailed geometry from video	Computationally heavy pipeline
Gaussian Splatting	Real-time performance with high quality	Struggles with occlusions and dynamic scenes
Sparse Fusion	Real-time & Robust to sparse depth and efficient fusion	Depends on the quality of sparse input

RESULTS:

(a) (b) (c)

Figure 1: This figure represents the 3-D scenes generated (a) by NeRF, (b) by Sparse Fusion (c) by Pixel Nerf.