[16-833] Paper Summaries 2

Bharath Somayajula

March 29, 2023

Contents

1   Kinect Fusion

1. The paper describes a method to perform real time volumetric reconstruction using Kinect sensor of large indoor scenes

2. In the first step, A vertex map and surface normal map are constructed at multiple scales after performing bilateral filter on raw depth obtained from sensor to improve quality of depth

3. In the second step, Pose is then computed between the map and measurement by linearizing the non-linear projective transformation under slow-motion assumption

4. In the third step, with the pose information available, the depth measurement is integrated into the map by using Truncated Signed Distance Function(TSDF)

5. In the fourth step, ray casting is performed on signed distance function to predict the surface which is given by places where SDF is 0

6. The system has been shown to preform both real time reconstruction of an object and indoor environment for AR applications

2   NERF

1. The paper presents a new technique to perform novel view synthesis for a scene using deep learning where the model accepts viewing angle and location as inputs and produces volume density and color as output

2. To preserve high frequency features in outputs, the inputs are first transformed using high frequency sine and cosine functions to map them to a higher dimensional input (20 for position and 8 for viewing angle)

3. Since much of the scene consists of empty space, two networks- coarse and fine- are used to compute color and volume density along a ray

4. The network architecture involves a series a simple fully connected layers followed by ReLU non-linearities with skip connections

5. The loss function used to optimize the networks involve L2 loss between predicted and ground truth color information obtained by COLMAP software on collected data

6. The model achieves SOTA results on DeepVoxels dataset for novel view synthesis task