16833: KinectFusion: Paper Summary

Andrew ID: hdhruva

April 2021

1 Summary

The KinectFusion system takes a real time stream of noisy depth maps from Kinect, and performs dense SLAM, incrementally producing a consistent 3D scene model, while simultaneously tracking the motion of the sensor using all the depth data in each frame. The first stage of surface measurement generates a dense vertex map and normal map pyramid from the raw measurements obtained, as a pre-processing stage. The surface measurement is then integrated into a volumetric truncated global TSDF, as the Surface update reconstruction step. Ray casting of the signed distance function is used to compute dense surface prediction to which the live depth map is aligned. The sensor pose estimation is achieved through multi scale ICP alignment between predicted surface and measured surface, and GPU enables the system to use all the available data at frame-rate. The system is currently not able to perform automatic re localization when the tracking fails for large models. [1]

References

[1] Richard A Newcombe et al. "Kinectfusion: Real-time dense surface mapping and tracking". In: *2011 10th IEEE international symposium on mixed and augmented reality*. IEEE. 2011, pp. 127–136.