Course Announcement

EE290T: 3D Reconstruction and Recognition
Instructor: A. Zakhor, avz@berkeley.edu,
Class times: Mondays from 1:00 pm to 4:00 pm, Wang room, 531 Cory Hall.
Pre-requisite: basic linear algebra & probability and statistics

3D sensors and capture devices are becoming more ubiquitous and cheaper thanks to recent popularity of virtual reality, augmented reality applications such as gaming and entertainment, autonomous driving, 3D mapping and a variety of other applications. Introduction of structured light Kinnect sensor in Xbox gaming system from Microsoft seven years ago resulted in great deal of activity in the research community in the general area of 3D scene reconstruction. More recently the use of 3D cameras and sensors in mobile devices such as Tango and iPhone has powered applications such as face recognition. The advent of solid state 3D sensors will enable yet another wave of applications that were inconceivable before.

The theory and algorithms used for processing, analyzing, understanding, and recognizing 3D data is not as mature or well developed as for 2D images. For example, while an image can be thought of as a regular, dense two dimensional array of numbers, for 3D data, there are multiple representations ranging from 3D point cloud , height fields and 3D mesh, all of which are irregular, sparse, and hard to run standard computational algorithms such as convolution on. In addition, adding a third dimension makes dealing with such 3D data sets more compute intensive due to the massive sizes of the data sets.

This course deals with reconstruction and recognition in 3D. The reconstruction portion of the course deals with 3D computer vision and recovery of scenes using stereo cameras, moving cameras, structured light cameras, and touches upon Simultaneous Localization and Mapping (SLAM) as well as bundle adjustment algorithms. The recognition portion of the course covers classical approaches as well as the modern convolutional neural networks with emphasis on 3D.

Topics include: camera and projection models, camera calibrations, single view geometry, epipolar geometry, stereo systems, structure from motion, fitting and matching, detectors and descriptors, image classification and 2D object detection, convolutional neural networks, 3D object detection, segmentation, and scene understanding; 3D recognition using multiple sensor modalities. Last 1/3 of the class is devoted to review of recent literature on 3D recognition methods.