



## DISSERTATION DEFENSE

### Nilesh Kulkarni



#### Reconstructing Full 3D Geometry and Interaction Dynamics from Real World data

Friday, August 30, 2024

4:30pm – 6:30pm

Virtual – [Zoom](#)

Passcode: 083024

**ABSTRACT:** 3D scene understanding is crucial for inferring spatial representations grounded in the real world. These representations are leveraged for various applications in robotics, virtual and augmented reality, and image/video manipulation. Fundamental problems such as 3D reconstruction, object affordances, and scene understanding from perceptual inputs broadly enable several applications to leverage 3D representations. One of the bottlenecks to training such models is the need for high-quality paired perceptual data with corresponding supervision, e.g., accurate 3D geometry for reconstruction. It is challenging to train deep networks with limited 3D data. In this thesis, we answer the question, “How do we train these models with inexpensive, noisy real data or use pseudo supervisory signals for large-scale learning?” specifically, we train and propose models for 3D reconstruction and human-object interaction. We propose an Articulation-aware canonical surface map, ACSM, that can learn 3D reconstruction of objects through geometrical self-supervision using web-scale data. This enables the reconstruction of object categories, such as birds and quadrupeds, where acquiring 3D supervision data is expensive. For reconstructing 3D scenes, our introduced Directed Ray Distance Function (DRDF) enables 3D reconstruction from a single or few images. Our representation, DRDF, effectively handles actual 3D scan data, typically a non-watertight mesh or posed RGBD image collection. DRDF is flexible with data constraints and applicable to various real datasets. Our models can also infer human-object interaction trajectories when trained with limited human-object interaction poses.

**CHAIR:** Prof. David Fouhey and Prof. Justin Johnson