A Self-Supervised Framework for Online Robot-Camera Calibration NY, USA

Advised by Dr. Shuran Song at Columbia CS 05/2020 - ongoing

PROBLEM Camera extrinsics can be computed from associations between points in the world and pixels in the image. The best procedures today are offline and cumbersome, requiring recalibration every time the camera is moved. Pose tracking can be used in place of markers but needs many real-world labels to be accurate. Calibrating a scene on-the-fly should be possible without anything more than what the camera can see.

ALGORITHM Articulated objects, like robots, can be tracked without collecting real-world labels by leveraging their 3D CAD models instead. Because robots in real-life settings will closely resemble the geometry of their models, I pre-train a network directly in simulation to estimate joint states from a depth image. Next, I fine-tune the network with unlabeled, real-world depth images by first treating them as partial point clouds of the robot, then comparing the partial point cloud to the robot's CAD model, and finally backpropagating a loss through my differentiable forward kinematics.

RESULTS The framework requires no more knowledge than the model of the object in the scene. It is able to both predict joint states and supervise its predictions from the same depth image. In addition, it can estimate the values of underactuated joints whose ground-truth is not available. Real world experiments are currently in progress.

PUBLICATION **Abhi Gupta**, Jingxi Xu, Peter Allen, and Shuran Song. "A Self-Supervised Framework for Online Robot-Camera Calibration." **Working paper.**