RBE MS Thesis Presentation

Ameya Wagh

A Deep 3D Object Pose Estimation Framework for Robots with RGB-D Sensors

Abstract: The task of object detection and pose estimation has widely been done using template matching techniques. However, these algorithms are sensitive to outliers and occlusions, and have high latency due to their iterative nature. Recent research in computer vision and deep learning has shown great improvements in the robustness of these algorithms. However, one of the major drawbacks of these algorithms is that they are specific to the objects. Moreover, the estimation of pose depends significantly on their RGB image features. As these algorithms are trained on meticulously labeled large datasets for object's ground truth pose, it is difficult to re-train these for real-world applications.

To overcome this problem, we propose a two-stage pipeline of convolutional neural networks which uses RGB images to localize objects in 2D space and depth images to estimate a 6DoF pose. Thus the pose estimation network learns only the geometric features of the object and is not biased by its color features. We evaluate the performance of this framework on LINEMOD dataset which is widely used to benchmark object pose estimation frameworks. We found the results to be comparable with the state of the art algorithms using RGB-D images. Secondly, to show the transferability of the proposed pipeline, we implement this on ATLAS robot for a pick and place experiment. As the distribution of images in LINEMOD dataset and the images captured by the MultiSense sensor on ATLAS are different, we generate a synthetic dataset out of very few real-world images captured from the MultiSense sensor. We use this dataset to train just the object detection networks used in the ATLAS Robot experiment.

Thesis Advisor: Prof. Michael Gennert
Thesis Committee: Prof. Emmanuel Agu, Prof. Berk Calli

Monday, April 22, 2019
4:00 p.m. - 5:00 p.m.
85 Prescott St., Suite 201, RBE 209 (Conf. Room)