ENGG 6100 Machine Vision

Assignment 3

Due: Presentation March 22, 2019 at 10 am
Report March 22, 2019 at 11:30 pm

Problem description

With the explosive growth of online shopping and commerce, there is a significant interest in automation of various operations in a typical online vendor warehouse. These includes filling and packing various orders that may include one or more items. In 2015 Amazon Inc. started the first competition to develop robotics systems that can extend currently available automation systems to include picking up and packaging of various objects available in a typical warehouse. The Amazon Picking Challenge (APC) involved perception, motion planning and grasping of 25 different objects placed in a semi-structured way inside the bins.

Dataset

The Rutgers University team published (http://arxiv.org/abs/1509.01277) a dataset that contains over 10,000 depth and RGB registered images, complete with hand annotated 6-DOF poses for 24 of the APC objects. They also provided 3D CAD mesh models for the 25 APC objects, which may be used for training object recognition algorithms or machine learning models. Please refer to this link http://www.pracsyslab.org/rutgers_apc_rgbd_dataset for additional details about the structure of the dataset.

Requirements

Design and implement a pose estimation algorithm for detecting and estimating the pose of at least five (5) objects in the dataset. Report your results in terms of translational and rotational error, similar to Figure 4 in the Rutgers Dataset paper. You are allowed to implement any existing pose estimation algorithm/model as long as you demonstrate knowledge as to how the techniques work, and the benefits/limitations of using them with respect to this dataset. Submit a report that summarizes your experiments, results, and findings via a discussion.

Getting Started

We suggest using Python or MATLAB to implement this assignment. Before working with the Rutgers dataset, you may wish to follow the OpenCV tutorial documentation to get comfortable with basic techniques for pose estimation of simple 3D objects from 2D images, e.g https://docs.opencv.org/3.1.0/dc/d2c/tutorial_real_time_pose.html and https://opencv-python-tutroals.readthedocs.io/en/latest/py_tutorials/py_calib3d/py_pose/py_pose.html#pose-estimation

You can use http://www.meshlab.net/ to get started with importing and viewing the object meshes.

Tips working with YAML files

The YAML files contain the ground truth labels for the object rotation and translation with respect to the camera ("object_rotation_wrt_camera" and "object_translation_wrt_camera" respectively). Note that you can ignore the labels w.r.t the base.
You shouldn't need to manually modify these files in any way. They can be loaded in Python with OpenCV as follows:

```
import os
import cv2
PATH = "path/to/rutgers_apc_dataset_sample_download"
FILE = "cheezit_big_original-pose-F-1-1-2.yml"
fs = cv2.FileStorage(os.path,join(PATH, FILE), cv2.FILE_STORAGE_READ)
fn = fs.getNode("object_rotation_wrt_camera")
print(fn.mat())
```

Output:

```
[[0.020486953054560247, -0.021702218741474447, -0.9995545481267953],
 [-0.9988029010550874, 0.04397413031596131, -0.021426307077247082],
 [0.04441954652614233, 0.998796954020
```