NYU-MMVC-LAB Weekly Report

CHENG QIAN

May 8, 2017

1 Motivation

This week we've been fully focusing on the ongoing project with Prof. Rizzo, in hope of working out a few papers for the upcoming ACVR workshop. The goals include:

a. Current demo run on TX2 for speed comparison

b. Signal/Traffic light detection with Deep learning method/CV feature extraction

c. Door/Doorknob recognition, joy stick control/orientation

d. Zebra strip detection, intersection walking and orientation ("you are 5 feet away from xxx/off center"), plus angle information

2 Methods

2.1 Jetson TX2

At first we were having some displaying issue as the screen was filled with sparkles. Then after several attempts Liang has found out that there's only one computer in the lab can run the demo on TX2 without any display issue, and he has reinstalled the system with all necessary working environment such as CUDA and OpenCV etc. I think we can show the progress on this Thursday.

2.2 Doorknob detection

We want to try two different methods for training if time permits:

a. Retrain the YOLO network by tuning its parameters and feeding it with manually collected photo data (this will be done by the volunteer with Prof. Rizzo), and we'll write a Python program to store the information of bounding boxes in each photo for further training/testing

b. Build a doorknob classifer and integrate this classifier into a single neural network model for better computation efficiency (inspired by the YOLOv2), although at this moment we're not sure how long this would take and we're a little concerned with the potential sacrifice in accuracy in exchange for speed

2.3 Traffic light and Zebra strips

We believe that all these object detection process can share the similar, if not the same, procedure and we want to start with doorknob to see how it goes. If the experiment results prove to be promising then we'll generalize the model and expand it for all three objects. At this writing the data is being collected, and we find the rest is relatively easy to accomplish once object detection is done.
By "rest" I mean:

a. intersection walking

b. alert the user with angle/signal information

c. simple orientation with audio output