Individual lab report #7

2/16/2017

Jin Zhu
Team E

Teammates: Michael Beck, Akshay Bhagat, Matt Lauer, Leo Lu, Jin Zhu

1. Individual Progress

I worked on bin localization and Faster R-CNN since the last progress review.

1.1 Bin localization error estimation

One of my task was to determine the error in alignment between CAD model and Kinect point cloud while using different AprilTag location and camera location. I used a ROS server to transform point cloud from Kinect frame to world frame and publish the point cloud. A ROS client was used for processing the point cloud data and find the error. The error means how different were the point cloud from Kinect to the point cloud of the bin CAD model. The larger error means worse localization accuracy. Figure 1 is a graph explaining the 3 major ROS nodes. The major functions were implemented and tested individually, but haven't been tested as a whole.

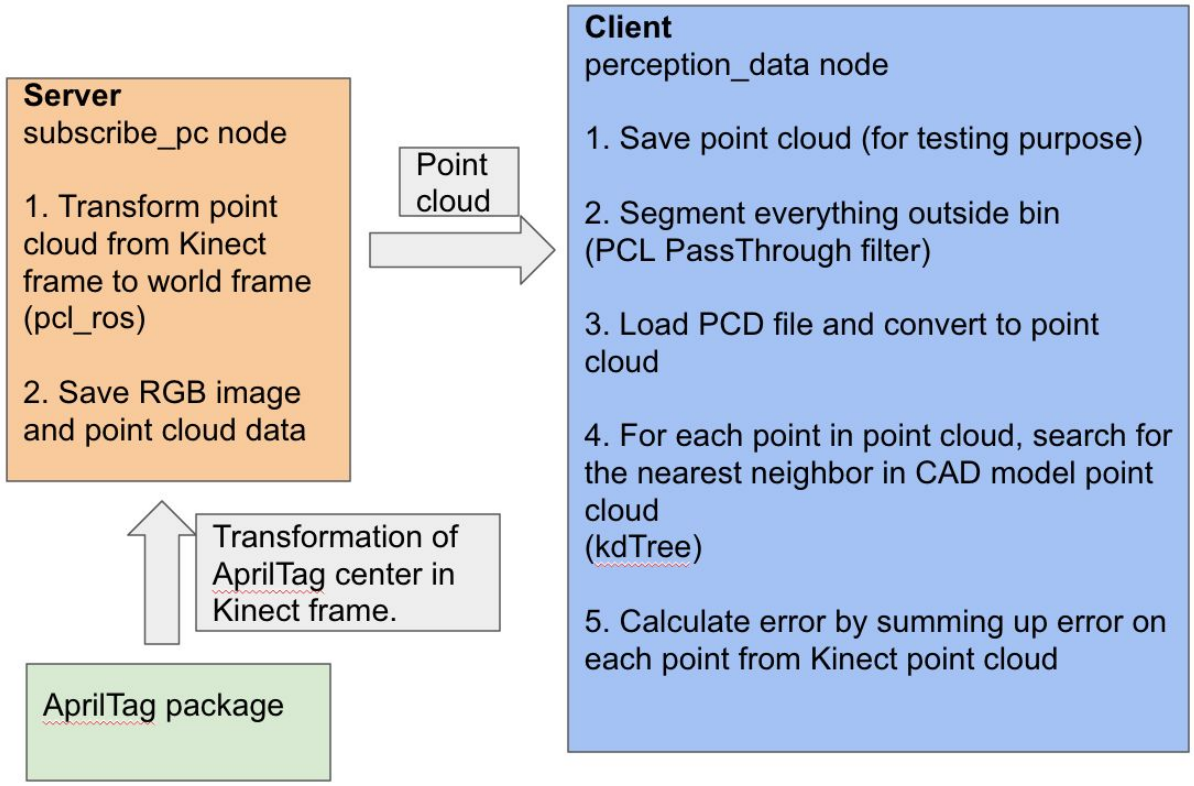

Figure 1. ROS nodes to calculate the error of localization

1.2 Faster R-CNN

The goal for object detection since last progress review was to get a comparison between accuracy of the performance of faster-rcnn and FCN. Our MSCV teammate Sharon was working on evaluating the accuracy of FCN, and I was working on evaluation of faster-rcnn. FCN is based on pixelwise segmentation, and faster-rcnn is based on bounding box.

These two object detection methods would be compared using the following metric, as shown in Figure 2.

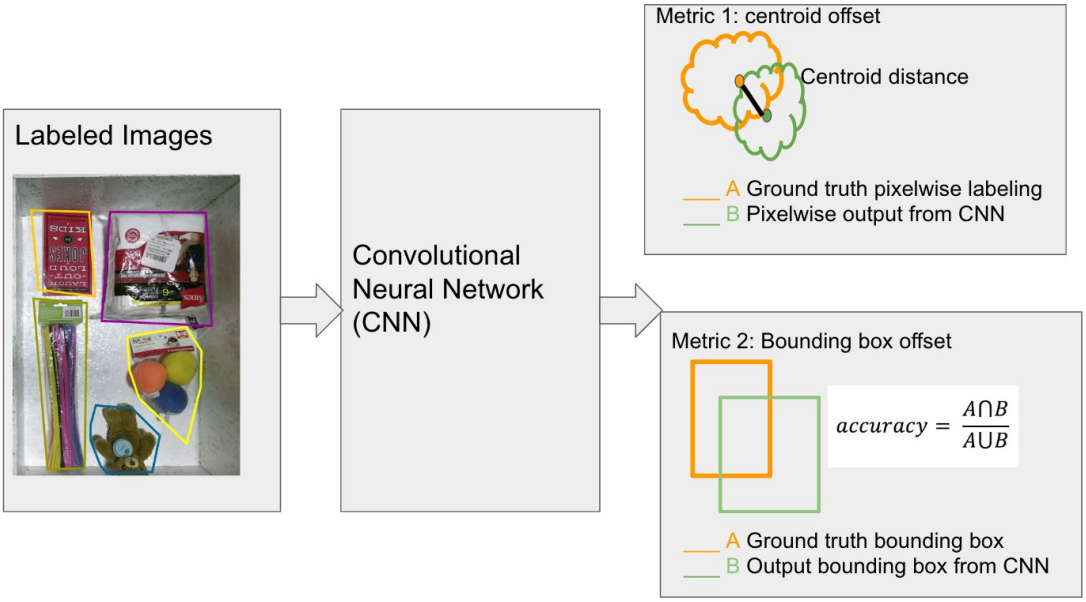

Figure 2. CNN comparison metric

Following are the results of occluded item identification result using Faster R-CNN trained over 27000 iterations. There were 79 images, 59 for training and 20 for testing. The next step is to train on all the 150 we had and use the result to compare with FCN. Figure 3 showed successful identification result on occluded item "laugh_out_loud_joke_book" even when the joke book was occluded by the bear. Figure 4 showed the successful identification of the two parts of joke book when it's occluded. For the situation in Figure 4, the comparison metric will use the centroid and bounding box of each of the two parts of joke book to compare with the ground truth.

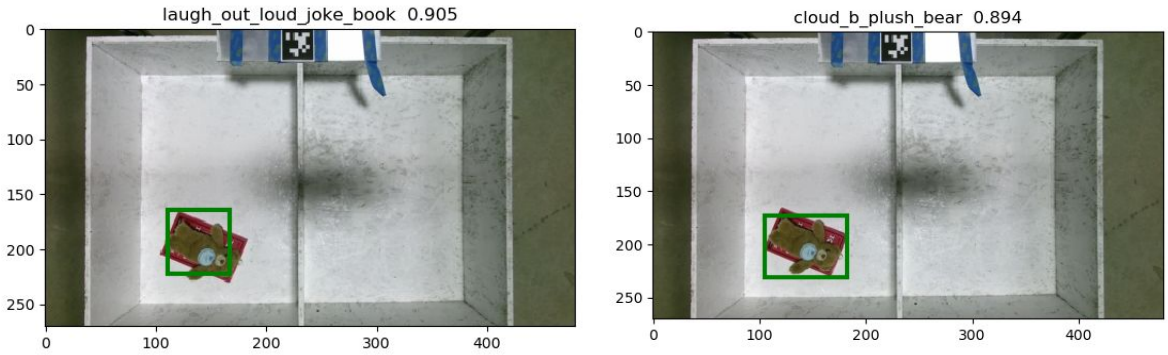

Figure 3. Identification of joke book and bear

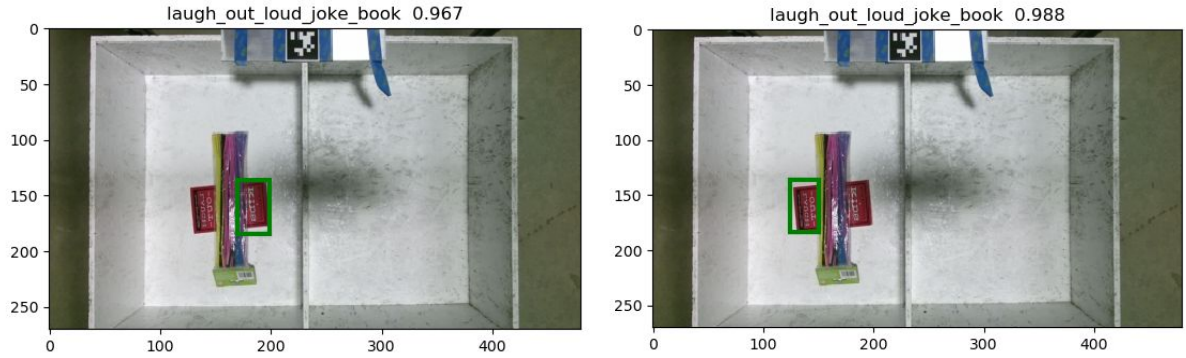

Figure 4. Identification of the exposed two parts of joke book

2. Challenges and problems

2.1 Version control

Insufficient version control impacted the progress of the team and also myself.

My laptop broke this week, so the latest code and the screenshots of visualization was not saved or pushed to github. Since this task is now moved under Akshay and I am moved to Faster R-CNN, I haven't taken time to reproduce the result and provide it here in this report.

Also, one of the team computer had issue and had to do a fresh Ubuntu installed last Friday. This caused the lag in our progress, since we devoted time from Friday afternoon to Sunday to fix the computer. Luckily we were able to retrieve files from the computer before do a clean install of Ubuntu. This computer crashing alarmed us about version control and the risk of computer crushing. Since computer is the central piece for us during the competition, we need to have a fully backed up working version before the competition.

2.2 Learning curve for Faster R-CNN

I was having trouble to finish training the Faster R-CNN on the new dataset. Last week was spent mostly on familiarize myself with Faster R-CNN and redo the ground truth labeling to better fit our need.

2.3 Inefficient time spending for lack of understanding on task

The new 100 images with items ranging from unoccluded to partially occluded was labeled using label me. It was not clear how the occluded items should be labeled, as shown in Figure 5. The occluded item had to be labeled as separate part, as shown in the light green boxes in Figure 5.a, instead of as a whole (Figure x.b).

The groundtruth labeling had to be redo and caused some unnecessary time lost. The lesson was do not jump into working when not clear about the task.

a b

Figure 5. Correct and incorrect labeling for ground truth

3. Teamwork

Lots of collaboration made the MVP possible for progress review. Leo, Matt and Michael worked together to fix the compute issue after one of the computer crash, they also integrated the MVP system together. Matt and Michael also worked on grasping code and linear actuator control. Akshay made the camera mount for Kinect and integrated faster-rcnn with ROS. Michael also laid out the rest of test plans for progress review and SVE.

I received a lot of help from Leo when learning ROS service and client. Leo also helped me with general coding issues when I worked on Faster R-CNN. Sharon and Akshay helped me with learning how to train Faster R-CNN.

4. Plans

The original plan was getting the MVP up and running for this progress review. Individual subsystem works now, but we need to solve the issue of integrate the whole system together. Also, we need to resolve the issue of Kinect signal over USB extension cord and find a solution to use multiple Kinect with stable performance.

Following is the new division of work we will do for the rest of the semester. This changed a little bit compared with the beginning of semester, taking into consideration team members' course selection that would help specializing in specific area of the project.

Akshay: Localization and shelf
Leo: Grasping, software in general
Matt: Planning
Jin: Object identification, project course deliverables schedule
Michael: Program manager, grasping

The division of work is now more clear, and some remaining work since last progress review changed task owner to better match this new division. For example, the remaining job of bin localization accuracy estimation was passed from me to Akshay, because this would be more relevant with his specialty in localization this semester.

The next big milestone for the team is to get deformable grasping code implemented and finish shelf localization.

My next step would be finishing the Faster R-CNN comparison with FCN and looking for methods to identify unknown deformable objects.