# Individual lab report #6

2/2/2017

*Jin Zhu*
*Team E*

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

**1. Individual Progress**

I worked on data collection and bin localization during the past 3 weeks.

1.1 Data collection

I collected 50 images for 5 item classes from last year's competition for Faster R-CNN training. The images are collected using a mockup for one bin of the current storage system design. The Kinect was mounted on 80/20 bars, looking top-down at the bin, as shown in Figure 1. This set of data included 1, 3, 4, or 5 objects per bin. Occlusion is not included in this set of data. Data was collected using a ROS node to save the RGB data as .jpg files.

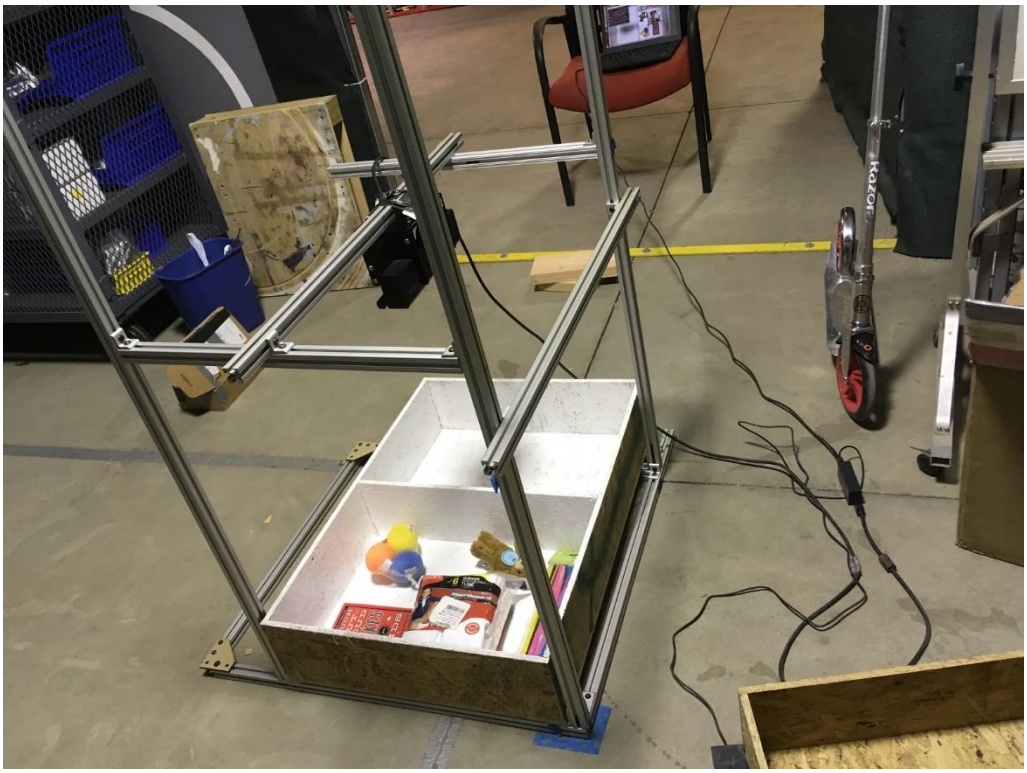

Figure 1. Kinect setup for data collection

Figure 2. RGB image ground truth annotated using LabelMe

Those collected images were annotated using LabelMe. Faster R-CNN needs bounding box, and FCN needs pixel wise labeling. Polygons fitting closely to the item boundary was used, so it can be used by both Faster R-CNN and FCN.

### 1.2 AprilTag size selection for bin localization

AprilTag would be used for localizing the bin, so I tried to see how small can the AprilTag size be, since larger AprilTag at the front face of the bin would sacrifice the accessibility of the bin space in Michael's design.

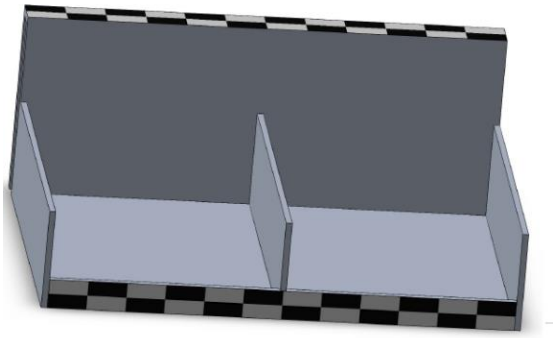

Figure 3. Michael's bin design with AprilTag

The AprilTag package outputs the position of AprilTag in the Kinect camera frame, in the form of x,y,z and quaternions. I tried the smallest identifiable size first, which was 1.5 cm wide. Since too small AprilTag wouldn't have enough pixels for accurate localization. Figure 3 showed an experiment setup which had two AprilTags separated 30 cm apart, and the x coordinate of the two AprilTags was around 3 cm more than the expected 30 cm. The AprilTag size was increased to 5.7 cm for current design for better accuracy.

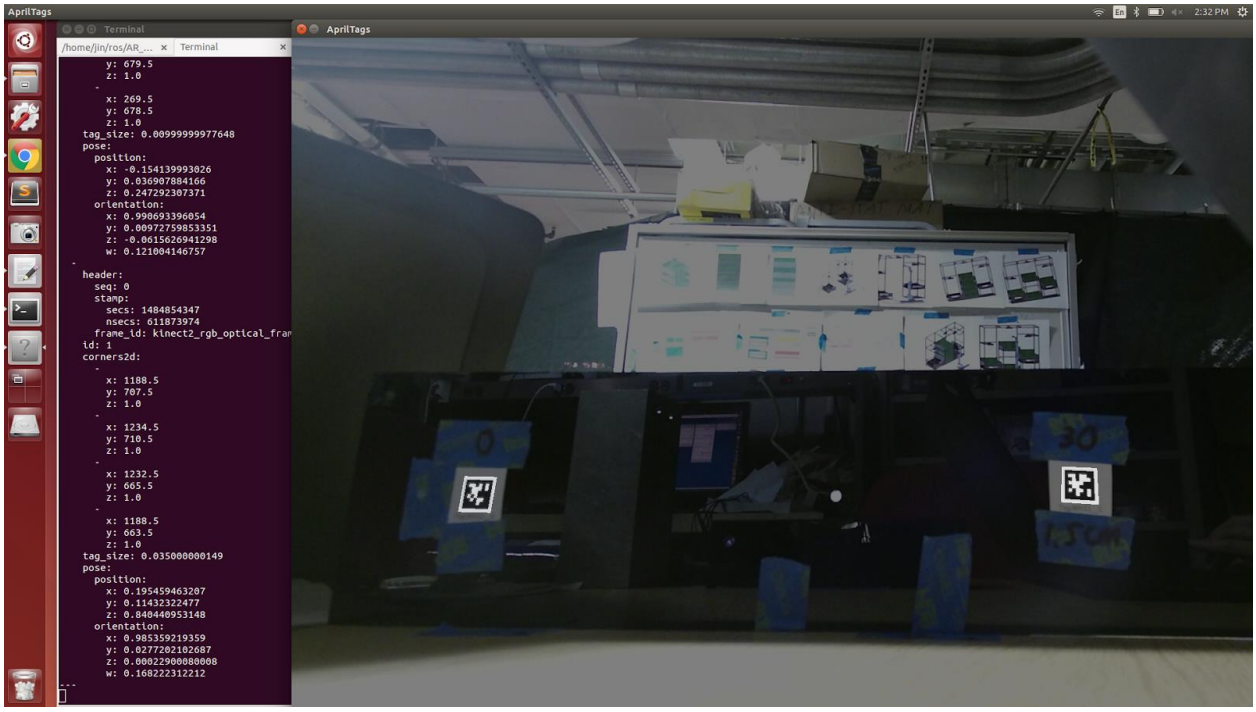


Figure 3. AprilTag identification using 1.5 cm tags

To deal with the issue of losing space for large AprilTag size, a new AprilTag position would be tried for later design, as shown in Figure 4. Since one bin wouldn't be right below another bin after the shelf was unfolded, an AprilTag hanging blow the edge wouldn't sacrifice accessibility of the bin.

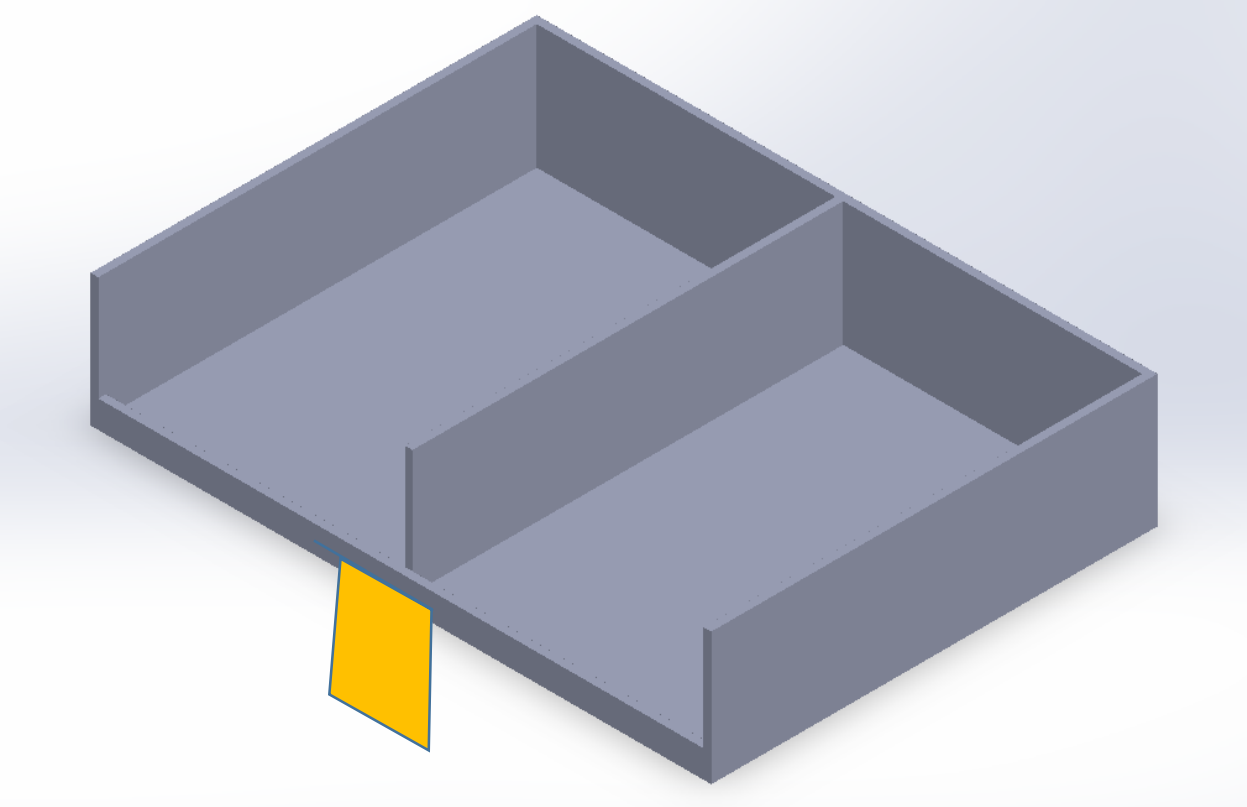

Figure 4. Possible new AprilTag placement. The yellow tag indicates the position for AprilTag.

1.3 Projection of point cloud and bin CAD model in rviz scene

Projecting point cloud on bin CAD model is helpful for localization and visualize the current state of the system.

Bin origin was set at the position where the AprilTag was positioned, as shown in Figure 5. The center of AprilTag, center of bin and the world coordinate were set to be the same, and the transformation output from the AprilTag package was used to give a position of Kinect camera center in the world frame. The point cloud and bin were separately projected on rviz scene (Figure 6), and overlapped in rviz scene (Figure 7).

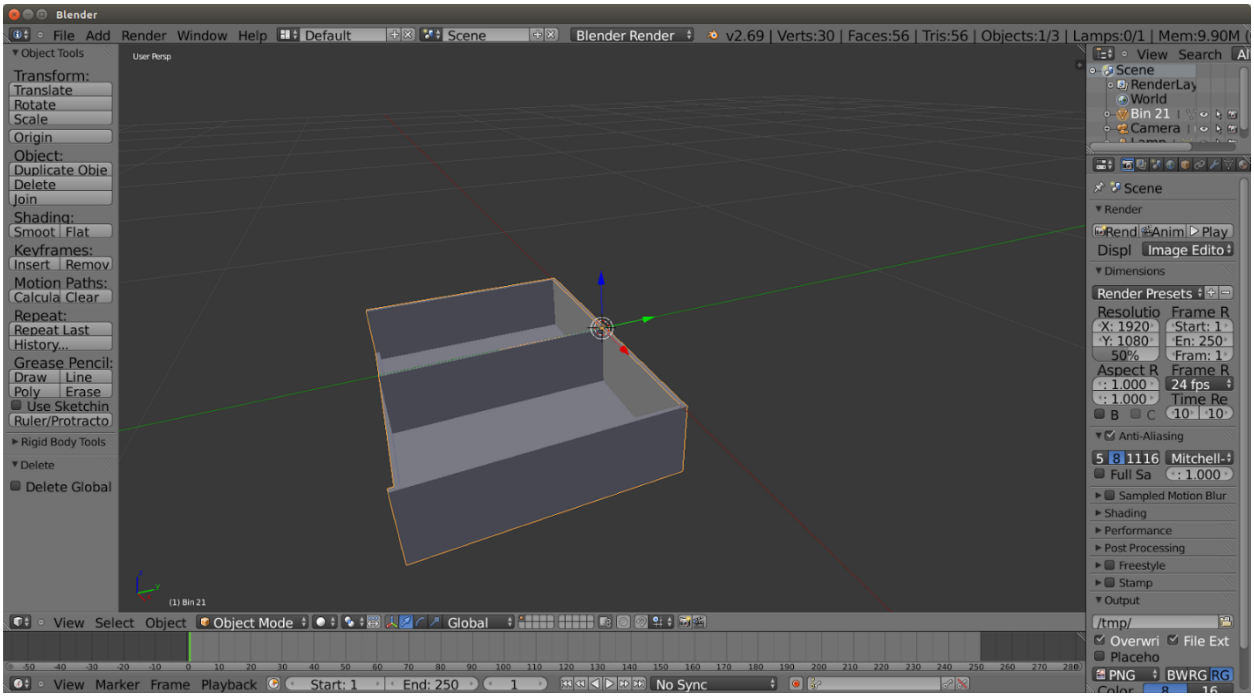

Figure 5. Adjusted origin for CAD model using Blender

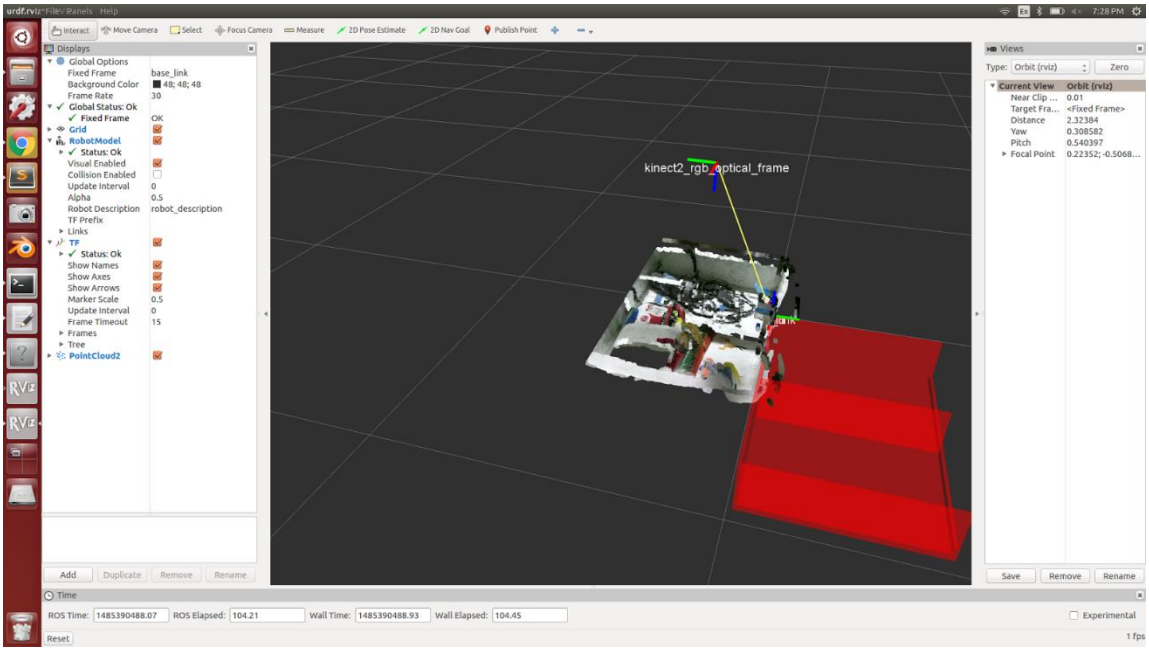

Figure 6. Point cloud and CAD without alignment

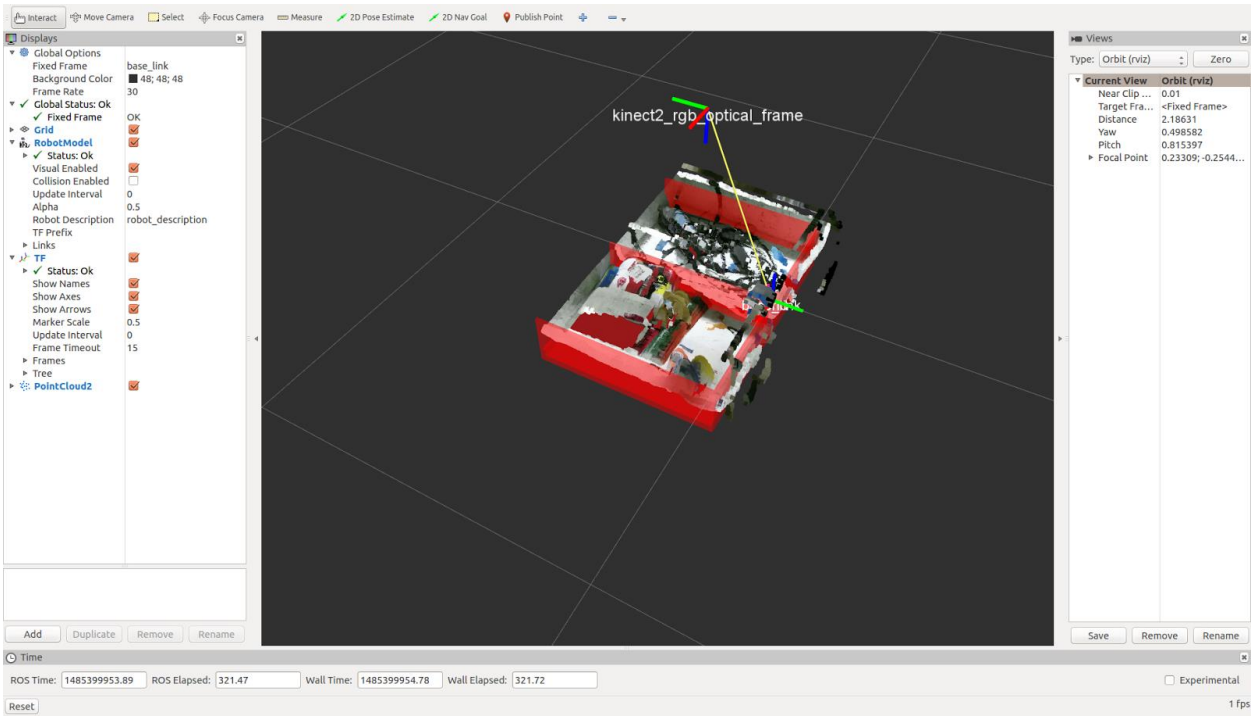

Figure 7. Aligned point cloud and CAD model

## 2. Challenges

2.1 Learning new tools and choosing the suitable tool

One challenge was learning tools such as Blender to reallocate the origin of CAD model, and URDF to project the CAD model into rviz scene. Also, before Alex suggested Blender, I was having difficulties to find a suitable tool for redefining the origin of CAD model.

2.2 ROS tf transformation

Another challenge I had was understanding the transformation between different frames. This was something taught in Manipulation, Mobility and Control, but I didn't learn it solid enough. The process of learning this and gain a more solid understanding used a lot of time, but would be helpful for the later tasks in localization.

## 3. Teamwork

Michael worked on creating CAD for gripper idealization and the shelf system. Michael and Akshay both worked on making the storage system wood mockup and created a bill of materials for the real storage system. The storage system bin positioning was based on Matt's work to find the optimum location to work with the arm. Akshay worked on train Faster-RCNN on the 50 images. He had 100% accuracy for the 10 test images. Matt worked on getting e-graphs planning working for UR10 and verified system viability

through planning scene. Leo worked on run PERCH on last year's items. PERCH would be used for non-deformable object identification and pose estimation.

This semester we started to use Trello to keep track of each other's progress better. Michael created high level subsystem descriptions for the competition and added granular Trello task cards for each person every week. The Trello cards for project course due dates and tasks such as team presentation would be maintained by me.

My teammates helped me a lot when I had troubles with my tasks. Matt pointed me to the correct direction for using URDF to project CAD model in rviz scene. Leo helped me when I have questions with programming and ROS. Michael helped me to get a high level understanding of task, so that if things such as certain packaged didn't work I can find something else as replacement for achieving the same goal. Akshay explained the ROS tf transform to me.

**4. Plans**

The next step would be getting a MVP ready. Which includes using multiple Kinects for item identification, picking some items from the new wood shelf mockup, making UR10 and linear actuator work together and identifying objects using an RGB base CNN.

Next week I will work on evaluating different positioning of AprilTag and overhead Kinect. The position that gives best alignment between CAD model and point cloud would be used for localization later. I will convert the CAD model to a point cloud using the Point Cloud Library. The error in alignment will be represented by the difference between the point cloud from Kinect and the point cloud from the CAD model. I will also work on multi-camera extrinsic calibration, so that the system knows the relative position between end effector, Kinect, and bins.