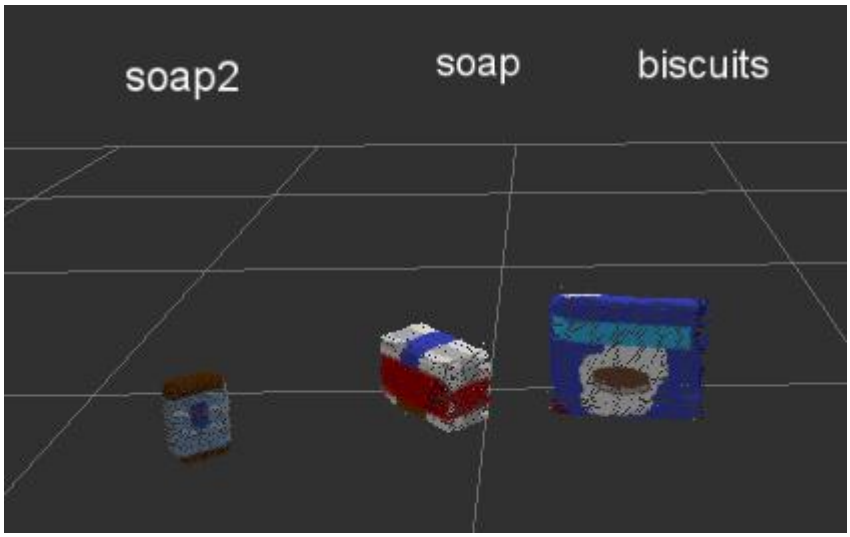
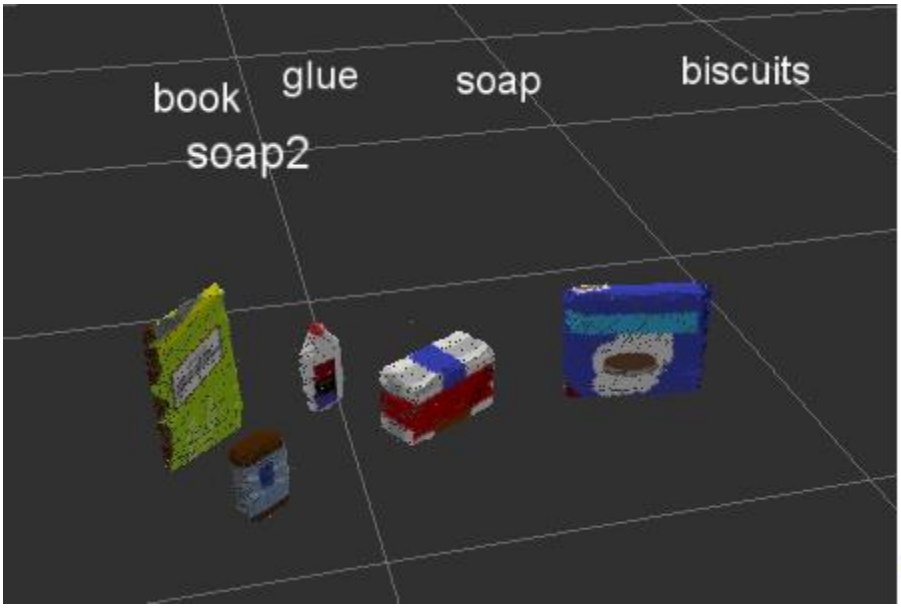
The perception pipline was created by first applying several filters to obtain an accurate point cloud. The filters that were applied were statistical outlier filtering, voxel grid downsampling, passthrough filtering, and RANSAC plane segmenting. RANSAC plane segmenting is used to detect inliers and outliers in the point cloud and separate the objects from the table. After the filters has been adjusted to properly capture the scene, the objects were segmented into separate clusters.





These screenshots show the object after all the filters has been applied. The object recognition was then created by capturing histogram features of the filtered objects. The object detection system got 100% on worlds one and two. The third world got 6/8 correct, but would detect an incorrect amount of objects. Capturing more features during sensor_stick training helped better detect objects.

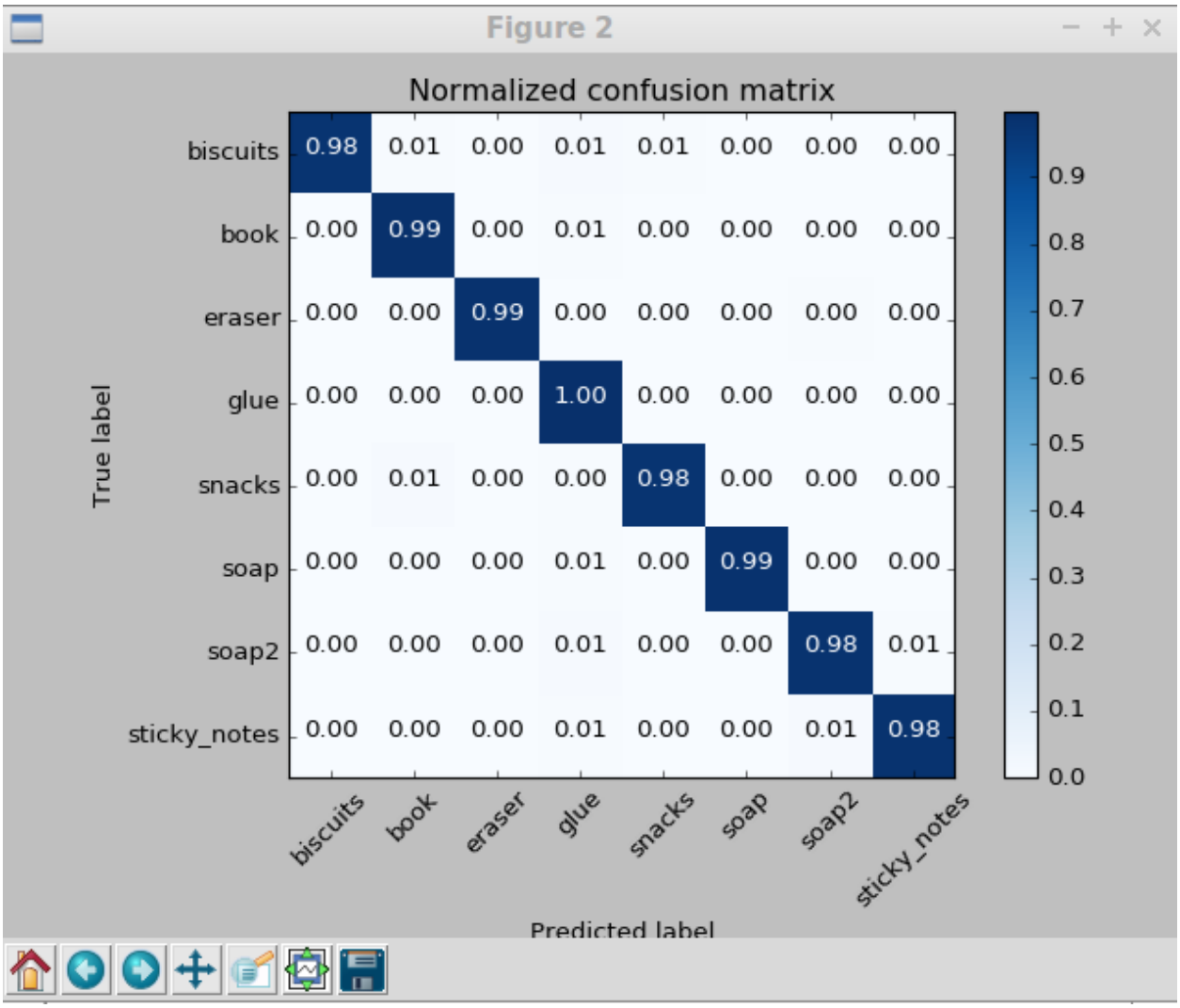
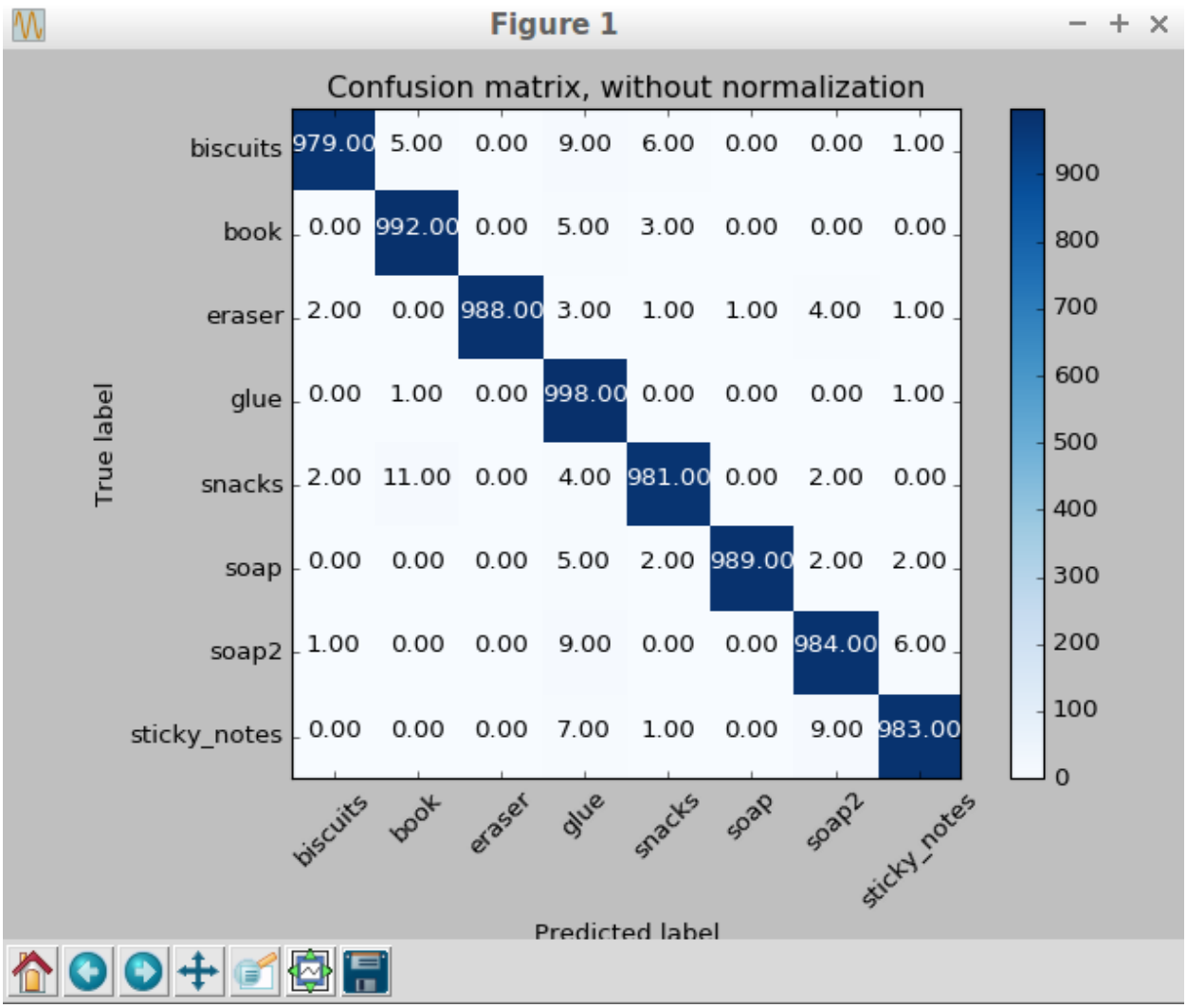


Figure 2
Normalized confusion matrix
biscuits 0.98 0.01 0.00 0.01 0.01 0.00 0.00 0.00
book 0.00 0.99 0.00 0.01 0.00 0.00 0.00 0.00
eraser 0.00 0.00 0.99 0.00 0.00 0.00 0.00 0.00
glue 0.00 0.00 0.00 1.00 0.00 0.00 0.00 0.00
snacks 0.00 0.01 0.00 0.00 0.98 0.00 0.00 0.00
soap 0.00 0.00 0.00 0.01 0.00 0.99 0.00 0.00
soap2 0.00 0.00 0.00 0.01 0.00 0.00 0.98 0.01
sticky_notes 0.00 0.00 0.00 0.01 0.00 0.00 0.01 0.98
True label
biscuits
book
eraser
glue
snacks
soap
soap2
sticky_notes
Predicted label
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0



The training data for world 1 and 2 captured 1000 features for each object. World 3 has a lower detection accuracy, because it was training on 50 features each. The train_svm settings were set to ‘rbf’. Slight improvements to accuracy were achieved by using this classifier method than ‘linear’ setting.