CS 674:Project-2 Report

Aditya Chukka
Net Id - ahc77
aditya.chukka@rutgers.edu

Chaitanya Mitash
Net Id - cm1074
chaitanya.mitash@rutgers.edu

1. GENERATING INITIAL CONFIGURATIONS

The following is the algorithm to generate initial configurations for each scene.

- Get the bounding box co-ordinates for the target-object in each scene.
- Convert the pixel points inside the bounding box into the camera frame using the camera intrinsic parameters.
- Then convert the camera co-ordinates into the world co-ordinates.
- Identify the center of the target-object in world-frame.
- Generate points that lie on the hemisphere of radius 25cm with the center of the target-object as the center of the hemisphere.
- Calculate the vector between the center and these points, which defines the approach axis of the gripper.
- discretize the angles on the plane perpendicular to the approach axis to get multiple orientations.
- These points are used as initial configurations for the end-effector. To visualize these we are stopping the simulator at the point of collision which speeds up the data collection. Find the video of the process here : https://youtu.be/v8MnznV6n5o

2. SCORING

- For each scene we get the target pose at the point of collision or final position from the simulators.
- At this point we also get the pose of the cuboid geometry which is an indicator of the area within the gripper.
- The segmented out image for the target object is obtained, and converted to point cloud in world frame.
- These points are then converted to the frame of the cuboid for each target pose.
- The number of such points, within the dimension of the cuboid is being used to score the grasp. This is an indicator of how much of the object is the gripper actually grasping.

3. TRAINING FEATURES

- Next step was to associate each of the initial configuration to a set of training features. There are two phases of feature collection.
- In the first phase, the points of the target object are converted to the frame of the cuboid geometry. Then we maintain counts along the discretized space of the cuboid. This indicates the shape of the object that could be grasped.
- In the second phase, we get the point cloud of the whole scene in the frame of the gripper and maintain count of points in the approach direction. This indicates the distance of any object from the gripper in initial position. Best is when the points are as far off as possible. This takes care of the obstacles in space.
- We combine these features to get a set of 25 features, for each configuration in the training and test images.

4. LEARNING

- Using the generated features and scores we have trained a bunch of regressors like KNeighbors, SVR.
- Using these we predict scores for each configuration in the test images and get the configuration with the best score.

5. RESULTS

- We extracted the 10 best configurations for each of the test scenes using multiple regressors and compared the results.
- For validation we trained different classifiers with scene1 and scene2 and predicted grasps for scene3. Following are the results :-

  Success with K-Nearest Neighbor :
  k = 5 : 5/10
  k = 10 : 4/10
  k = 15 : 7/10
  k = 20 : 8/10
  k = 25, weights = distance : 5/10
  video(k=20) : https://youtu.be/ZSSAkg5mLHQ

  Success with SVR :
  default(Gaussian kernel) : 4/10
  Linear SVR : 3/10 (These seem good though)
  NuSVR : 4/10 (Good ones, just missing in others)

  Ridge :
  default : 0
  Bayesian : 0
  (It does grasp a few times, but bad quality)

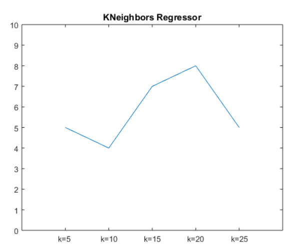

Figure 1: Knn results by tuning k

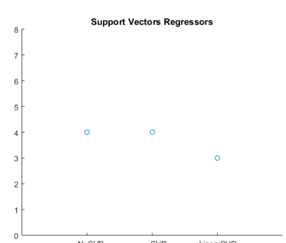

Figure 2: Support Vectors Regression

6. TO IMPROVE

- We started working with the surface normals at each point on the object as the approach direction, however it seemed hard to debug, so went ahead with hemi-spherical sampling and vector to the center.
- The scoring function was kept simple because it is easy to debug. We thought of other heuristics as taking the component of the vectors connecting grasped points to the center of the cuboid

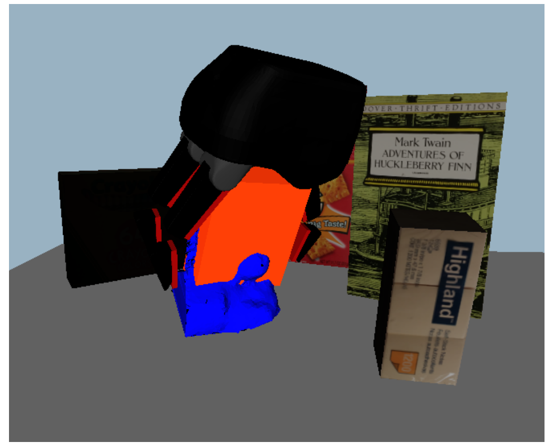

Figure 3: A good grasp, indicated by the collision of the cuboid geometry and the target object

Figure 4: A bad grasp, indicated by the collision of the collision of the fingers with the target object

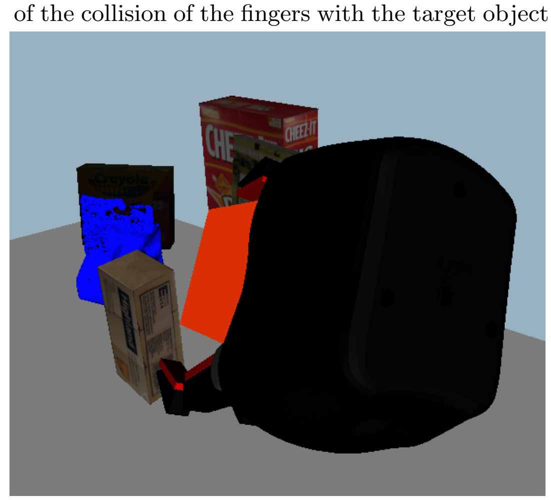

Figure 5: The training feature finds the obstacles in this view closer compared to others