

Leveraging Big Data for Grasp Planning

Reminders

- “ ϵ -metric quantifies the quality of force-closure grasp in terms of the smallest external wrench that breaks a grasp when unit force is applied at the contact points” - unit force maintains the grasp

Overview

- Multiple useful contributions:
 - Generating a large annotated grasp database
 - Verifying both the popular ϵ -metric and the proposed physics metric using crowdsourcing as the ground truth
 - Using the generated database to train a DNN classifier to predict grasp success from the local grasp configuration
 - Comparing the DNN to a more frequent (in this domain) linear logistic classifier
- Using the local shape (called *template*) of the object for evaluating grasp quality
- Similar work on learning grasp mapping from vision already out there
- Annotations are usually provided from ϵ -metric which is deemed unreliable
 - ϵ -metric is a proxy for determining force closure \rightarrow static equilibrium, but good grasps require dynamic stability which can only be analyzed through a dynamic system

Key ingredients

- Multiple objects arranged into four categories: toy bottles, small, medium and large objects
- In total over 500k labeled examples
- Multiple grasps are performed for every object for different standoffs (distance from object collision), approach rays and rolls, but the same preshape is used on all grasps
- Local shape representation for learning
 - Basically a grid projected on the tangent plane constructed in the intersection point between the object and the approach ray
 - **If I understood correctly:** From different viewpoints the pointcloud is generated and the points are projected to this grid as the height value. If this value changes on some viewpoints, this counts as occlusion.
 - Not sure if input is all of the viewpoints, or are multiple viewpoints used to construct a single grid (probably the latter, will double check)

Comments

- Not much to comment on as everything is covered in the paper
- A very nice contribution, only some parts could be explained better, while some other parts are maybe too emphasized - e.g. the logistic regression part
- Still unclear what the local object representation is. TODO: Figure this out