CS365 Project Proposal

Group 6

Ankit Agarwal, Ripu Singla

Vision Based Grasping

- We propose to make NAO able to grasp object of certain classes.
- Presently we will be working on Coffee Mugs.
- We plan to do the task of grasping based on vision, assuming that we do not have 3-d model available for the object.
- Our work will be based on the paper by Ashutosh Saxena.

Robotic Grasping of Novel Objects using Vision

- Considers the problem of grasping novel objects, specifically ones that are being seen for the first time through vision.
- Most of the work in robot manipulation before it assumes availability of complete 2-d or 3-d model of the object.
- The task of identifying where to grasp an object involves solving a difficult perception problem.
- Attempt to infer grasps directly from 2-d images, even ones containing clutter.

Learning the Grasp Point

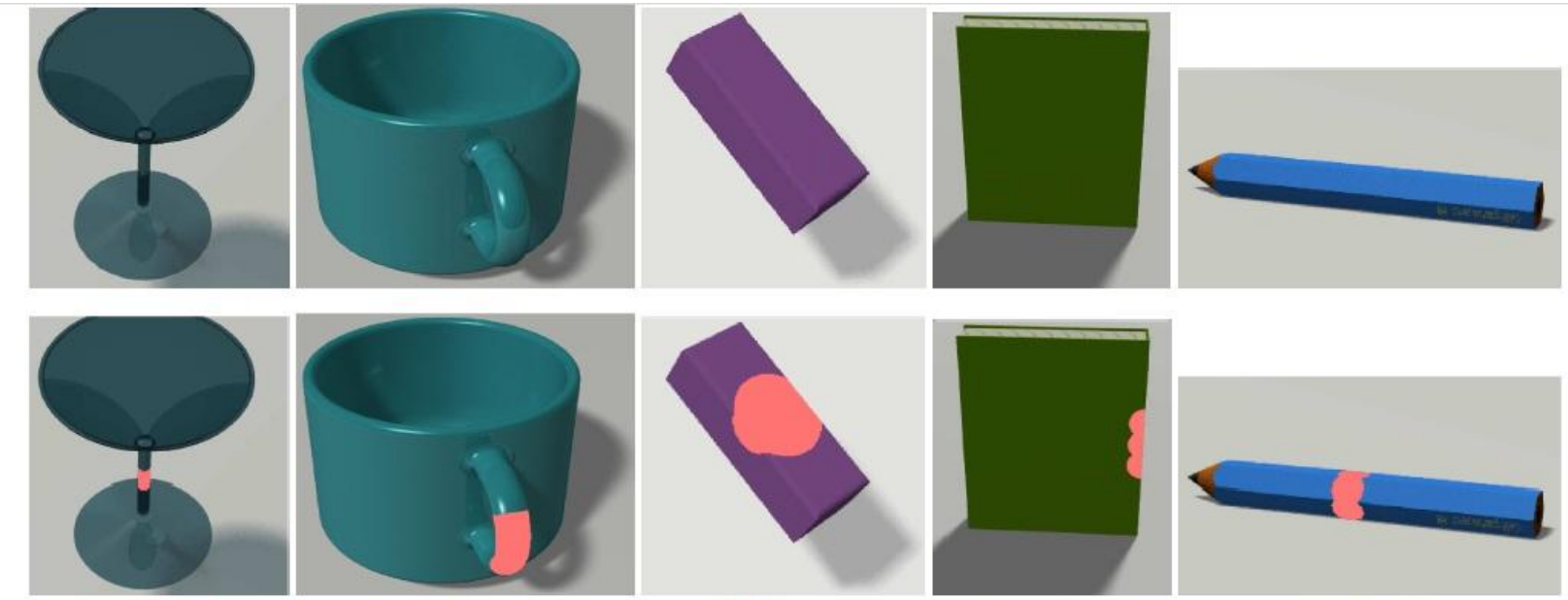
- Even very different objects can have similar subparts, there are certain visual features that indicate good grasps.
- Try to identify the projection of a good grasping point onto the image plane.
- Given two images, predict the 3-d position of the grasping point.
- Synthetic data for training.
- Computes 3 types of local cues: edges, textures and color.

Features

- Divide the image into small rectangular patches.
- Transform the image into YCbCr color space.
- Convolve the intensity channel with 6 oriented edge filters to get features representing edges.
- Apply 9 Laws' masks to intensity channel to compute texture energy.
- Apply first Laws mask to 2 color channels.
- Dimension of feature vector x:

$$1*17*3 + 24*17 = 459$$

Training objects



(a) Martini glass (b) Mug (c) Eraser (d) Book (e) Pencil

Probabilistic Model

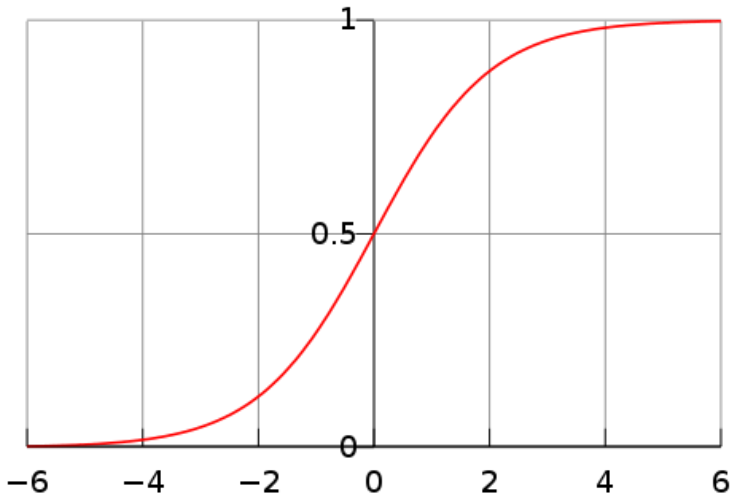
$$P(z(u,v)=1|C) = P(\hat{z}(\hat{u},\hat{v})=1|\hat{C})$$
$$= \int_{\epsilon_u}\int_{\epsilon_v} P(\epsilon_u,\epsilon_v)P(\hat{z}(u+\epsilon_u,v+\epsilon_v)=1|\hat{C})d\epsilon_u d\epsilon_v$$

$$P(\hat{z}(u+\epsilon_u,v+\epsilon_v)=1|\hat{C}) = P(\hat{z}(u+\epsilon_u,v+\epsilon_v)=1|x;\theta)$$
$$= 1/(1+e^{-x^T\theta})$$

$$\theta^* = \arg\max_\theta \prod_i P(z_i|x_i;\theta)$$

Logistic Regression

- Used for prediction of the probability of occurrence of an event by fitting data to a logistic curve.
- $f(z) = 1 / 1+e^{-z}$
- $z = b_0 + b_1x_1 + \ldots + b_nx_n$



Probabilistic Model

$$P(z_i(u,v)=0|C_i) = P(y_{r_1}=0,...,y_{r_K}=0|C_i)$$
$$= \prod_{j=1}^{K} P(y_{r_j}=0|C_i)$$

$$P(y_{r_j}=1|C_i) = 1-(1-P(z_i(u,v)=1|C_i))^{1/K}$$

$$P(y_j=1|C_1,...,C_N) = \frac{P(y_j=1)P(C_1,...,C_N|y_j=1)}{P(C_1,...,C_N)}$$
$$= \frac{P(y_j=1)}{P(C_1,...,C_N)} \prod_{i=1}^{N} P(C_i|y_j=1)$$
$$= \frac{P(y_j=1)}{P(C_1,...,C_N)} \prod_{i=1}^{N} \frac{P(y_j=1|C_i)P(C_i)}{P(y_j=1)}$$
$$\propto \prod_{i=1}^{N} P(y_j=1|C_i)$$

Results

Objects Similar to Ones Trained On

Tested on	Mean Absolute Error (cm)	Grasp-Success rate
Mugs	2.4	75%
Pens	0.9	100%
Wine Glass	1.2	100%
Books	2.9	75%
Eraser/ Cellphone	1.6	100%
Overall	1.80	90.0%

Novel Objects

Tested on	Mean Absolute Error (cm)	Grasp-Success rate
Stapler	1.9	90%
Duct Tape	1.8	100%
Keys	1.0	100%
Markers/Screwdriver	1.1	100%
Toothbrush/Cutter	1.1	100%
Jug	1.7	75%
Translucent Box	3.1	75%
Powerhorn	3.6	50%
Coiled Wire	1.4	100%
Overall	1.86	87.8%

NAO

- Each arm has 5 DOFs (2 at shoulder, 2 at elbow, 1 at wrist)
- All 3 fingers closes simultaneously.
- Can carry upto 300gm using both hands.

Dataset

- We will use subset of the following dataset : http://ai.stanford.edu/~asaxena/learninggrasp/data/mug.tar.gz
- Contains 2001 labeled examples of coffee mug in different configuration.
- The depth_ image gives the depth at each pixel much like a standard gray scale depth image.
- The graspPriorityWidth_ image is used to give a ground truth label for each pixel to indicate whether it is a grasp or not.
- Other files gives information such as camera location.

References

- Robotic Grasping of Novel Objects using Vision - Ashutosh Saxena, Justin Driemeyer, Andrew Y. Ng, Computer Science Department, Stanford University, 2008
- http://ai.stanford.edu/~asaxena/learninggrasp/data.html
- Cooperative Human Robot Interaction with the Nao Humanoid: Technical Description Paper for the Radical Dudes
- Mechatronic Design of NAO Humanoid - David Gouaillier, Vincent Hugel, Pierre Blazevic