Grasping Familiar Objects Using Shape Context

Jeannette Bohg and Danica Kragic
Centre for Autonomous Systems
Computer Vision and Active Perception Laboratory

KTH, Stockholm, Sweden

Goal

- Re-use grasp experience on objects familiar in shape
- Infer grasp hypotheses for an object
  - Precision Grasp with Barrett Hand
  - 6 DoF to determine (Position and Orientation)
- Input: stereo image of the object

Related Work

Two different movements in grasping of unknown/familiar objects:

1. **Reconstruction of Object in 3D** (e.g. Hübner et al. IROS 2008, Detry et al. ICDL 2009, Richtsfeld et al. ECCV 2008)
   - search stable grasp configurations dependent on 3D structure
   - How to select a hypothesis?
   - How to deal with a high reconstruction error?
2. **Reasoning about 2D Information** (e.g. Saxena et al. NIPS 2006, Stark et al. ICVS 2008)
   - exploit structural cues in a monocular image directly to infer a grasp position
   - avoid difficult 3D reconstruction
   - How to infer a hand orientation?

Approach

1. Detect **where** to grasp from 2D shape.
   - 2D Grasping Point Detection
   - Relative Shape Representation
   - Experiments

2. Infer **how** to grasp from position of grasping point within (minimal) 3D reconstruction.

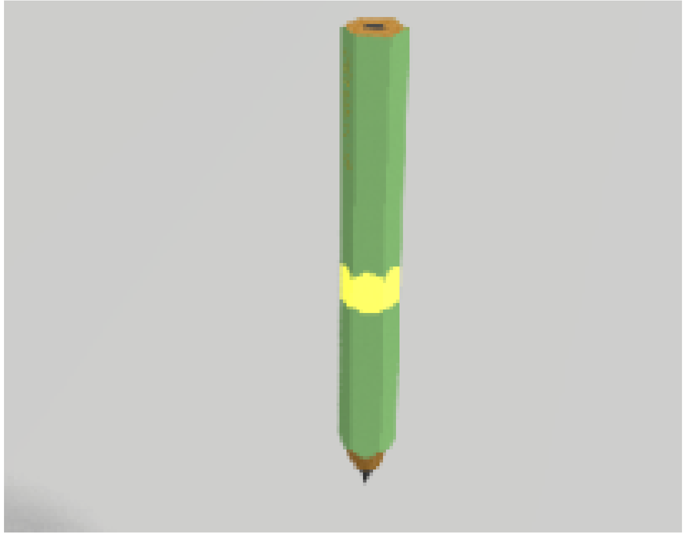
2D Grasping Point Detection

- **Grasping point**: 10×10 px patch in a monocular image, contact point of the palm of the hand
- Supervised training of a grasping point model on labeled database (Saxena et al.)

- 2-class Support Vector Machine: Grasping point/ Non-Grasping point

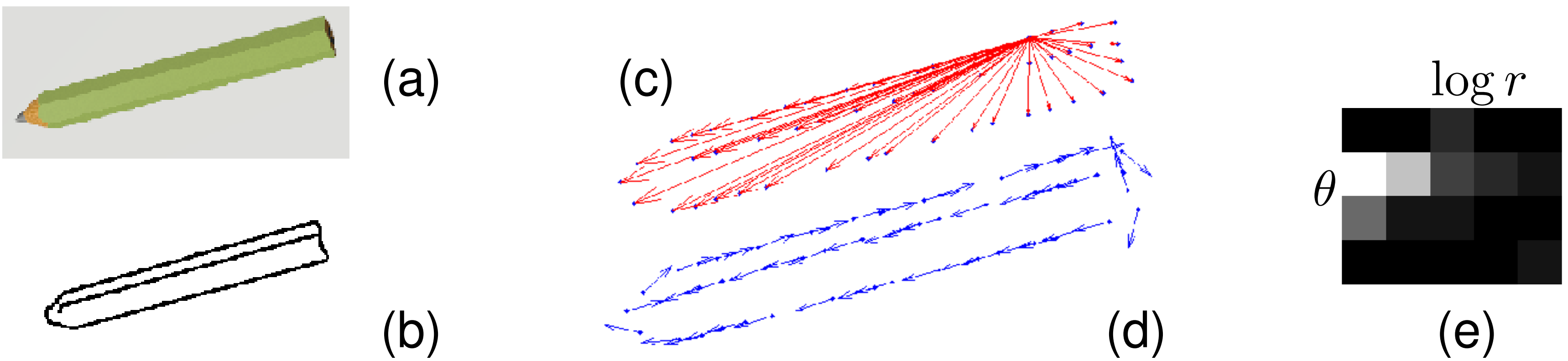
Representation

- Grasping point representation through **relative global shape** as opposed to local appearance



Feature Vector

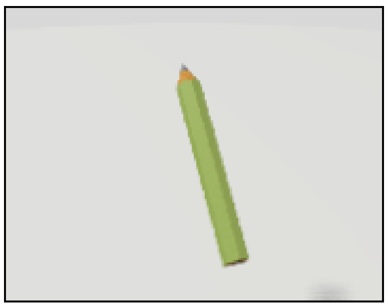
- Shape Context (Belongie et al. NIPS 2000)

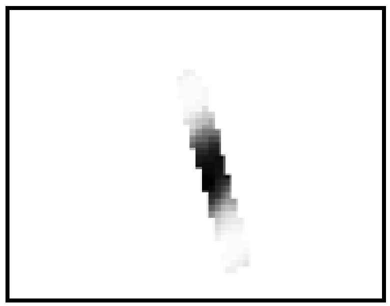


- Feature Vector: merged histograms of all points in a patch and in three spatial scales

Example Result

- Toy example: grasping point model trained on pencils
- *Graspability* distribution over picture of pencil





- Model encodes center of mass
- Local maxima in output are grasping points

Evaluation on Synthetic Images

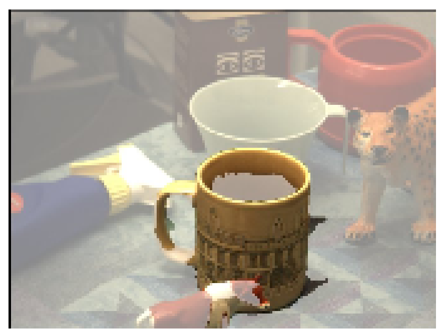
Comparison to appearance based method by Saxena et. al, NIPS 2006

Table 1: Accuracy of the models trained on different objects.

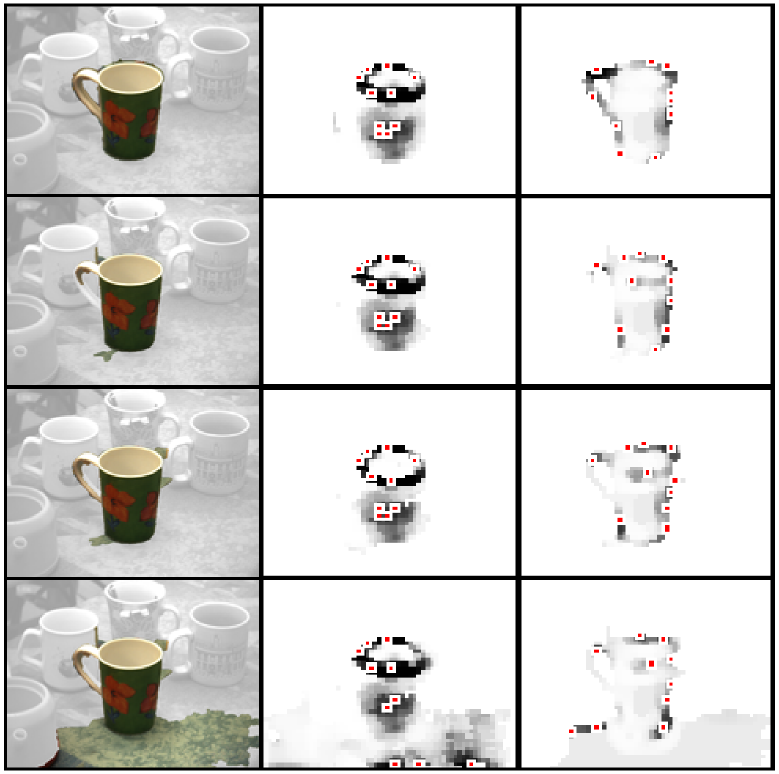
	Shape	Appearance
Pencil	82.55%	77.07%
Cup & Mug	88.01%	83.85%
Pencil, Martini & Eraser	85.65%	74.92%
Pencil & Mug	84.64%	74.32%

Evaluation on Real Images

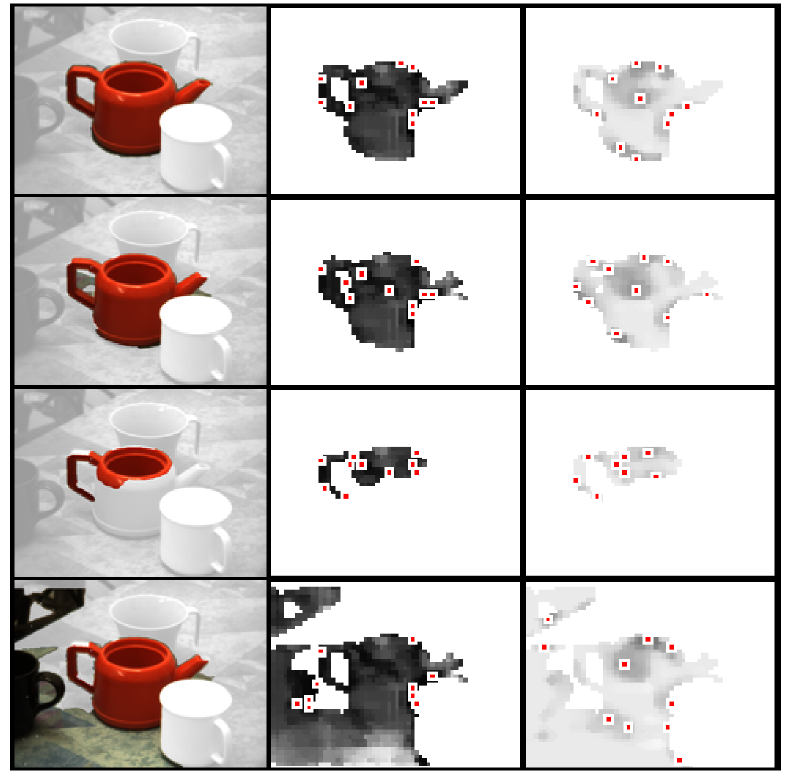
- How does different quality of segmentation affect the selection of grasping points?
- Goal: Grasping points in two images of same object differently segmented should correspond to each other.



Evaluation on Real Images (2)



(a) Textured Cup.



(b) Partly Occluded Teapot

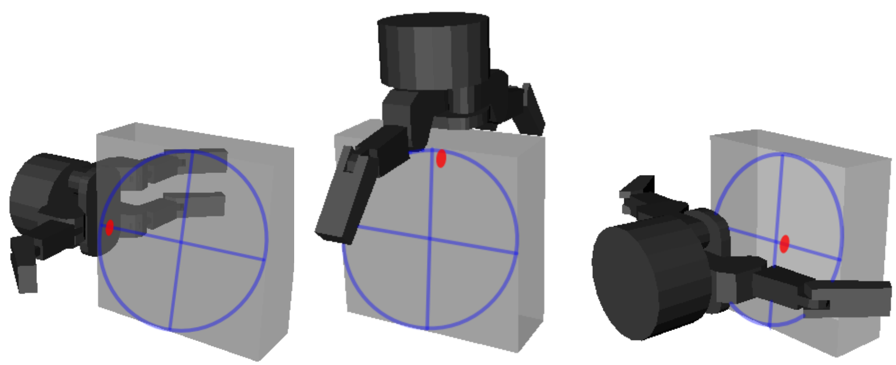
Left: Segmented image, **Middle**: Local Appearance, **Right**: Relative Shape.

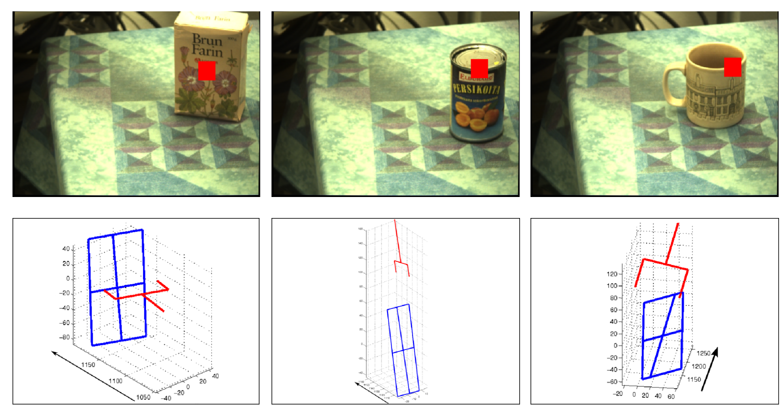
From 2D Grasping Points to 6 DoF Grasp Hypotheses

- Use 2D information as a cue to search for grasp hypothesis in (minimal) 3D reconstruction!

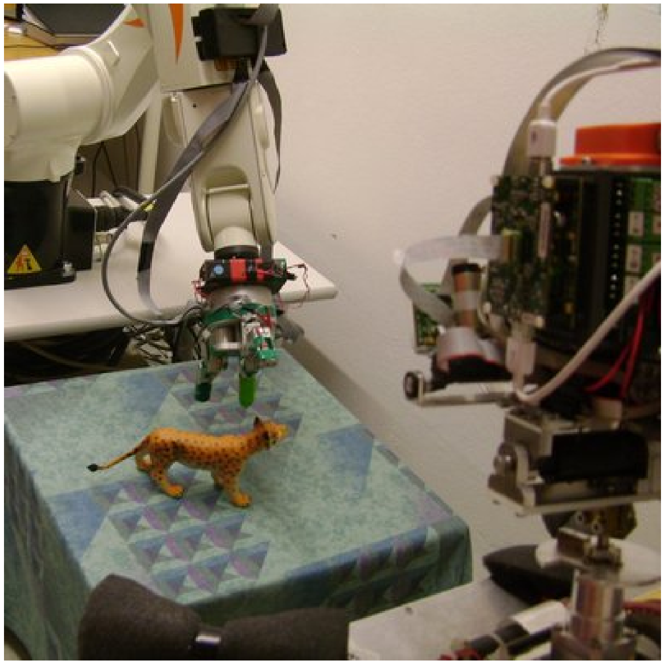
Infer Full Barrett Hand Configuration

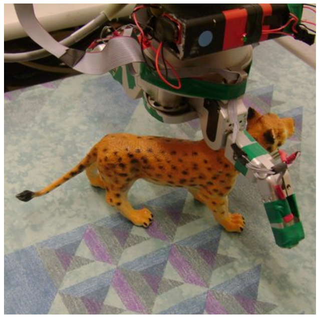
1. 3D Coordinates of Grasping Point through Triangulation
2. Orientation of Barrett Hand
   - object pose approximation through oriented plane
   - hand orientation through grasping position relative to plane

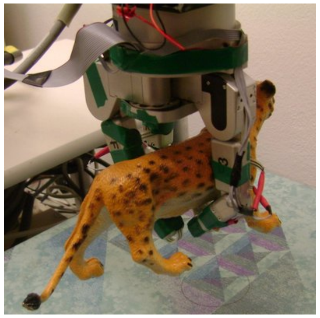


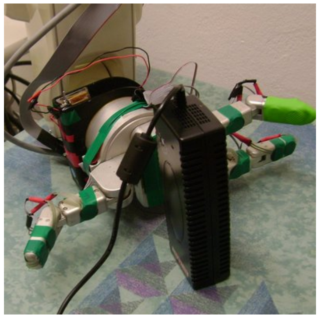


Grasping with the Real System









Conclusions

- 2D object shape provides discriminative features for grasping position
- Based on grasp position within minimal 3D object representation full grasp configuration can be inferred
- Future Work
  - Minimal 3D representation (robust and probabilistic)
  - Integration with tactile grasp controller

Thank you for the Attention!

Questions?

References:

Hübner et al., Selection of Robot Pre-Grasps using Box-Based Shape Approximation, IROS, 2008.
Detry et al., Learning Object-specific Grasp Affordance Densities, ICDL, 2009.
Richtsfeld et al., Grasping of Unknown Objects from a Table Top, ECCV Workshop, 2008.
Saxena et al., Robotic Grasping of Novel Objects, NIPS 2006.
Zillich et al., Functional Object Class Detection Based on Learned Affordance Cues, ICVS 2008.
Belongie et al., Shape Context: Functional Object Class Detection Based on Learned Affordance Cues, NIPS 2000.