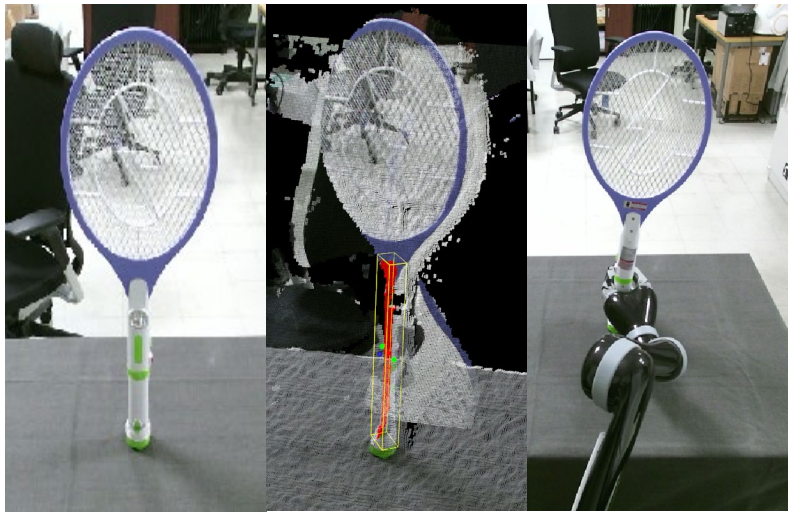
Human-like Object Manipulation Based on Object Affordance Detection and 3D Shape Analysis for Social Robot

Introduction

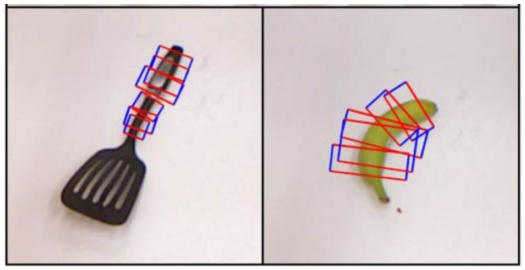


- **Left** is origin RGB image. **Middle** is the result of the proposed architecture. **Right** is the result of applying the architecture to the Jaco robot arm.
- This approach can not only grasp objects but also deal with the object after grasping it.

- The goal of conventional grasp detection is how the robot can grasp objects without thinking of use of the objects.

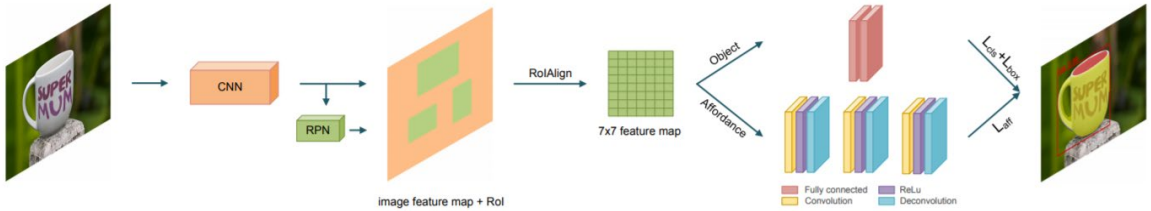
Related works

Grasp detection



- They detect grasp detection in real-time with CNN

Affordance detection

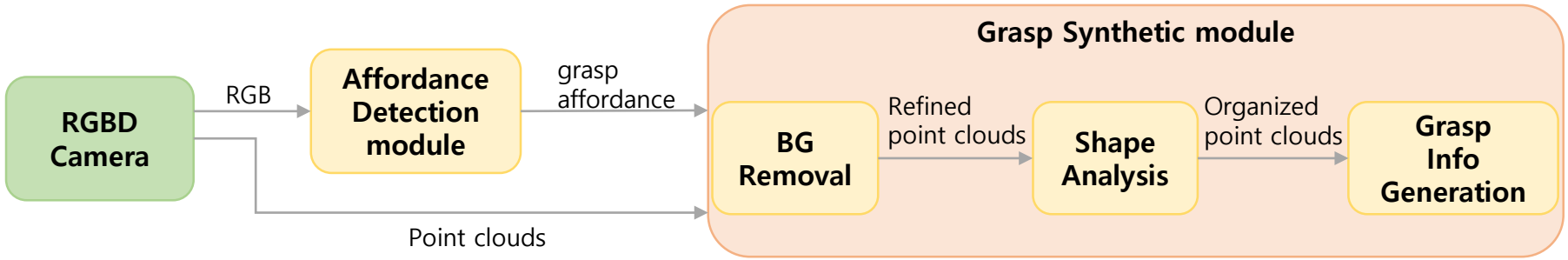


- AffordanceNet detect affordances of objects
- We adopted AffordanceNet for affordance detection

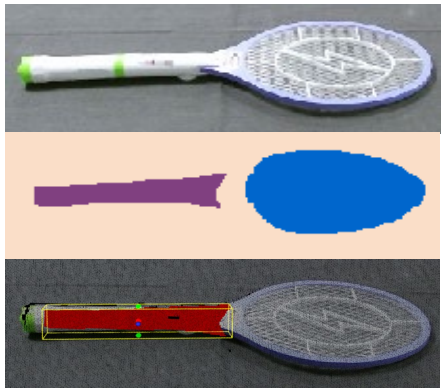
Grasp detection using affordances

- In [16], grasp affordance is used for grasp detection

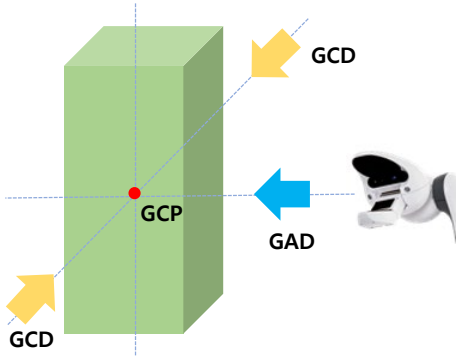
The Architecture of the Grasp Synthesis



- The proposed Architecture of the Grasp Synthesis. The Grasp Synthesis generates Grasp Information of objects for a Robot who is about to grasp the objects.

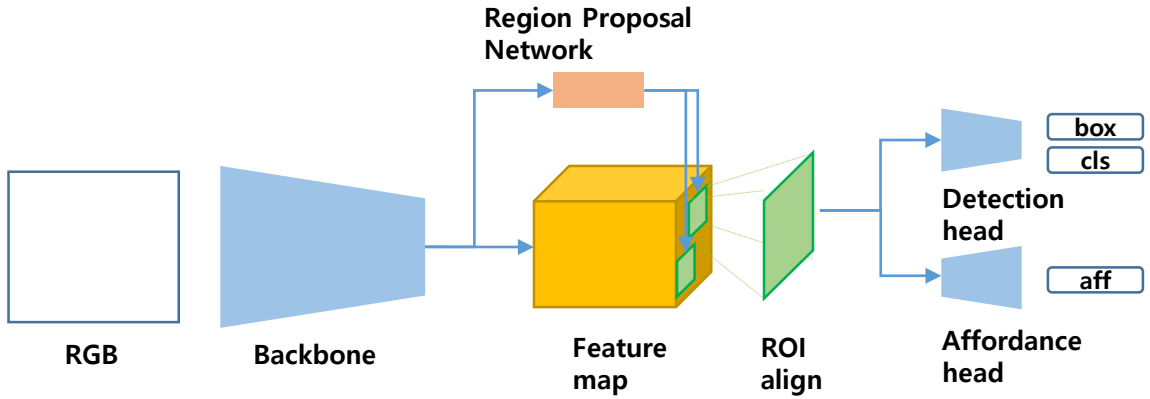


- **Top** is origin RGB image. **Middle** is the results of Affordance Detection module. Each color refers to some pre-defined affordances. **Bottom** is the results of Grasp Synthetic module. The red, green and blue dots denote GCP, GCD and GAD, respectively.



- The illustration of the Grasp Information.
  GCP : The expected Grasp Center Point of a shape
  GCDs : The expected Grasp Close Directions of robot's fingers
  GADs : The expected Grasp Approach Directions of robot's hand

Affordance Detection module



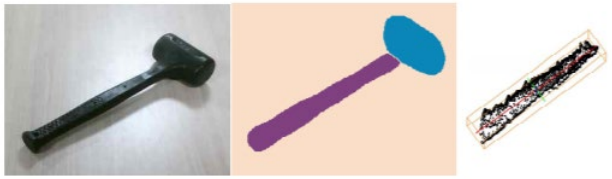
- This is the Architecture of the Affordance Detection module. We applied AffordanceNet which adopted Mask RCNN. It can detect affordance masks and bounding boxes of pre-defined objects.

Objects	Bowl	Monitor	Hammer	Pan	Knife
	Cup	Drill	Racket	Spatula	Bottle
Affordances	Contain	Hit	Cut	Display	Engine
	Support	Pound	w-Grasp	Grasp	

- The pre-defined 10 object classes and 9 affordances.

▪ Grasp Synthetic module

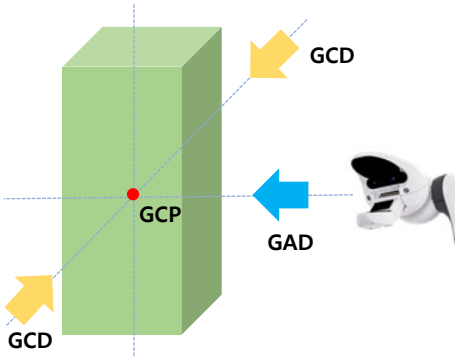
➢ Background Removal & Shape Analysis



(a) origin image (b) affordance detection (c) Point clouds fitted with cuboid

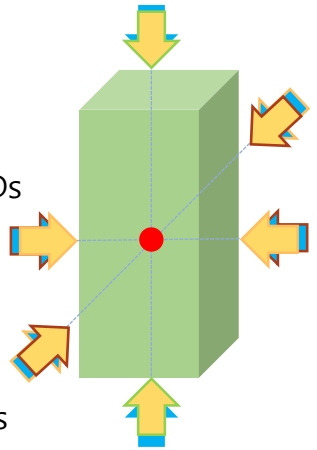
- Grasp affordance mask can be detected through the affordance detection module. To obtain 3D information of the grasp affordance, Grasp Synthetic module receive not only grasp affordance mask but also point clouds as input. But the point clouds may contain background due to the gap between the 2D and 3D sensors. To remove background, clustering is used and the largest cluster is remained.
- The remained point clouds are 3D information of grasp affordances, but not organized.
- We adopted the notion of primitive shapes (pre-defined shapes) such as sphere, cube and cylinder.
- The primitive shape that best matches the refined point clouds are found.

➢ Grasp Information Generation



- GCP : Grasp Center Point
- GAD : Gripper Approach Direction
- GCD : Gripper Close Direction

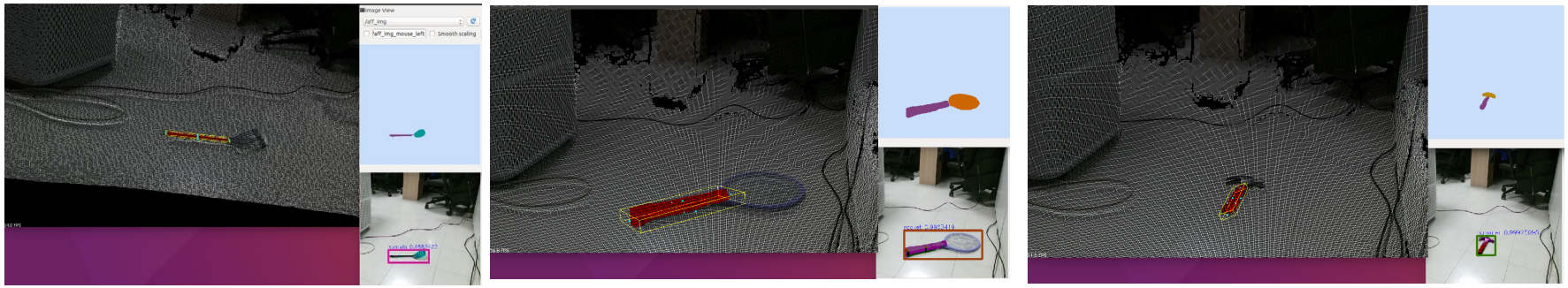
- When the cube shape fits into the grasp affordance, the GADs are decided as the directions from each side of the cube to GCP
- When GADs are decided, the GCDs are orthogonal to each GAD
- The facing GCD vectors are a set
- Each grasp affordance has 1 GCP, 6 GADs and 6 sets of GCDs



Experiments

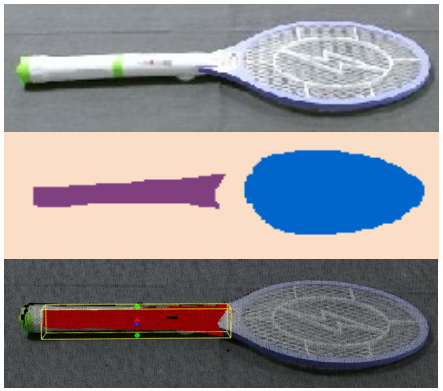
The results of our Architectures

- The results of the Affordance Detection module.

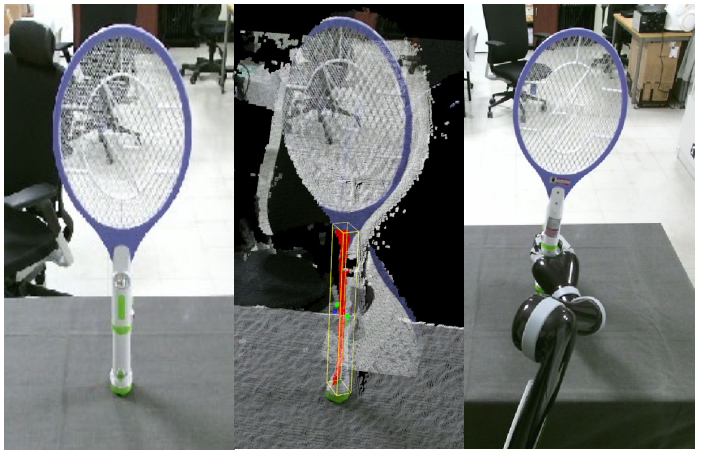


- The results of the Grasp Synthesis. Each figures shows the results of affordance detection and grasp info generation.
- On the right, the Top is the affordances mask and the Bottom is the bounding box of an object.
- On the left, it shows point clouds with RGB.
  The red point clouds are the grasp affordance.
  Yellow bounding boxes are the results of the Shape Analysis.
  The cyan color points are the GADs(Grasp Approach Directions) points.

- **Experiments**
  - The results of our Architectures and applying it to Jaco robot arm



- **Top** is origin RGB image.
- **Middle** is the results of Affordance Detection module. Each color refers to some pre-defined affordances.
- **Bottom** is the results of Grasp Synthetic module. The red, green and blue dots denote GCP, GCD and GAD, respectively. It shows only the one set of GCDs and GADs for avoid duplication.



- The result of the proposed architecture and applying it to the Jaco robot arm

Conclusion

- It provides various grasp information without focusing only on the grasp. It helps the robot to do other tasks after grasping the object

- Affordance detection module achieves 88% accuracy on custom dataset

- We verified that our grasp synthesis using affordance is useful and Jaco robot arm grasped an object using generated grasp information