Contents

1 Introduction *Malthe Høj-Sunesen*

According to ISO 8373 [ISO, 2012], at least two different types robots exist: Industrial robots and service robots. An industrial robot is defined as aa "automatically controlled, reprogrammable, multipurpose manipulator programmable in three or more axes", while a service robot is defined as a "robot that performs useful tasks for humans or equipment excluding industrial automation applications". The classical application of an industrial robot is to have the robot do a predefined behavior repeatedly, while service robots are still very much under development. Due to hardware and software concerns, robots in the industry have previously not seen adaptive behavior, so elements must be aligned in a specific way. Humans, and indeed most animals, are able to look at objects and grasp accordingly. A lot of research is going into making the robot able to understand what it is "looking" at much like humans can, and how to grasp it. This research into grasping[1] objects using only visual cues is the focus point for this report.

[1]For the purposes of this report, *grasping* is to pick up an object.

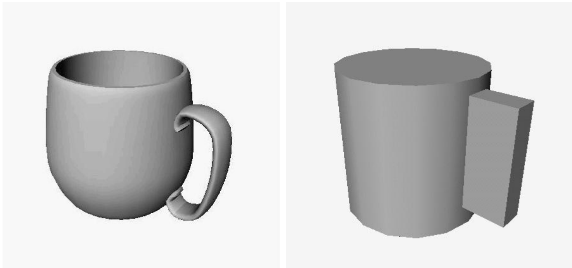


Figure 1: "A mug and its primitive representation"; a cylinder and a box. From [Miller et al., 2003].

2 Grasping

2.1 Motivation *Malthe Høj-Sunesen*

Humans spend years learning how to grasp objects. Babies have a hard time figuring out how to grasp even the most simple objects, and parents solve that problem by giving babies and children plastic cutlery, bouncy, soft toys and always walking around with an eye on each finger. We come to expect of a child to drop toys, knock over glasses, and the like.

A robot is not allowed to fail in the same way. When a robot's hand grasps something we expect it to not let it go — or worse, drop it — before it is supposed to. In a tightly controlled production line that is not a problem. Using embodied AI the parts can be aligned perfectly for the robot and the robot can assemble the parts correctly.

In a not so tightly controlled environment among people it is a bigger problem. If a service robot is supposed to clean up mess left after a human, it is almost guaranteed that the parts are not aligned as a robot could predict. If an industrial robot can figure out the best grasp autonomously for an object it would decrease operator dependency, leading to faster setup and lower costs for the company.

2.2 Simplifying objects to primitive models *Malthe Høj-Sunesen*

Biederman suggested that elements can be broken down to geons, basic elements describing one feature of an object.

In [Miller et al., 2003] the idea behind geons is used to help a robot simulator find good grasps. The robot knows how to grasp each shape primitive (equivalent to geon). Any object is then reduced to its shape primitives where applicable. This allows a simulator to know which points are good to grasp, resulting in simpler calculations. An example of this reduction can be seen in Figure 1 with shape primitive building bricks shown in Figure 2.

Reducing the visual information in this way will give the simulator a

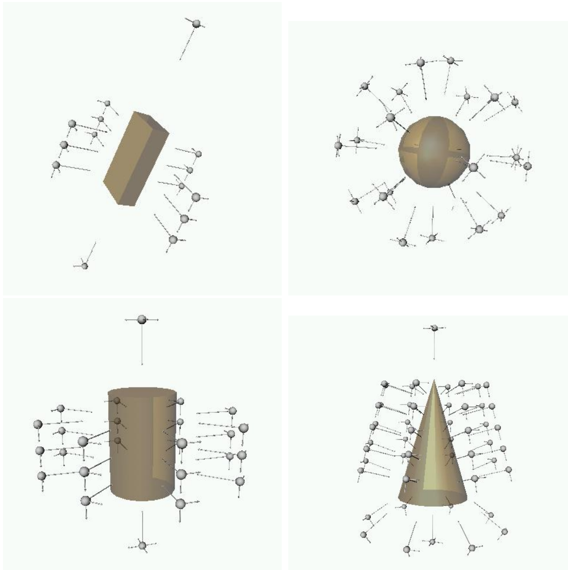


Figure 2: "Examples for grasp generation on single primitives. The balls represent starting positions for the center of the palm. A long arrow shows the grasp approach direction, and a short arrow shows the thumb direction. In most grasp locations, two or more grasp possibilities are shown, each with a different thumb direction." From [Miller et al., 2003].

simpler task, as it does not have to simulate thousands of possible grasps but only the grasps based on the preshape grasps per primitive representation. An example of the found grasps can be seen in Figure 3.

2.3 Learning to grasp through attempts *Malthe Høj-Sunesen*

Much like [Miller et al., 2003] in Section 2.2 tried to emulate how the human vision works according to Biederman, so do [Detry et al., 2011] try to emulate how a child learns to grasp objects. Any parent will tell you that their child did not quite know how to actively grasp[2] toys from the beginning. Where to hold is one of the problems.

The approach in [Detry et al., 2011] is to let a robotic platform learn how to grasp a single object. Using stereo vision, the 3D features of an object can be calculated. The system will then try to calculate where that object can be picked up. An example of where the system calculates a toy pan can be grasped can be seen in Figure 4 on the following page. s

After calculating where the object can be grasped, the robot will start to grasp the object, time and time again. In [Detry et al., 2011], the robot performed more than 2000 grasps. During the grasp trials the robot and

[2]Let alone letting it go again!

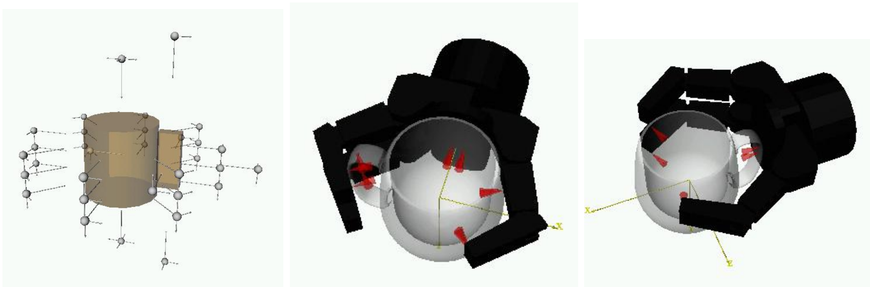


Figure 3: The primitive mug representation and the two best grasps. The red cones indicate point-of-contact. From [Miller et al., 2003].

.

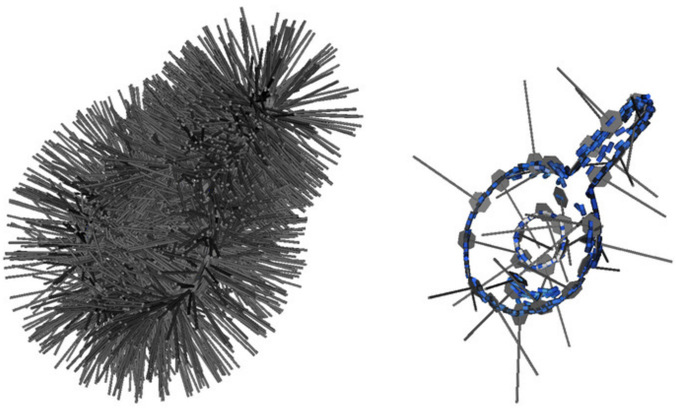
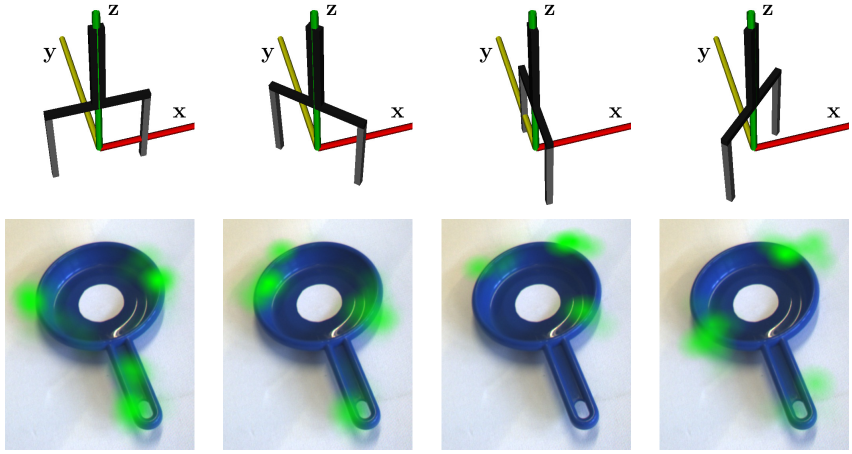


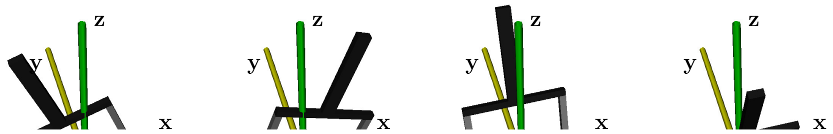
Figure 4: Left: A full graph of the possible grasp positions. Right: Clearer showing of what the sticks mean; the stick is the robotic hand's translation while the paddle at the end is the point where the fingers close. From [Detry et al., 2011]

vision system will see if the grasp is stable, ie. if the object is not moving. Using all the trials the robot system can build a model of the possibility that a grasp will be successful.

Learning Grasp Affordance Densities



(a) Downward grasp | (b) 45° rotation around **z** | (c) 90° rotation around **z** | (d) 135° rotation around **z**



References

[Detry et al., 2011] Detry, R., Kraft, D., Kroemer, O., Bodenhagen, L., Peters, J., Krüger, N., and Piater, J. (2011). Learning grasp affordance densities. *Paladyn, Journal of Behavioral Robotics*, 2:1–17.

[ISO, 2012] ISO (2012). Robots and robotic devices — vocabulary. ISO 8373:2008(en), International Organization for Standardization, Geneva, Switzerland.

[Miller et al., 2003] Miller, A., Knoop, S., Christensen, H., and Allen, P. (2003). Automatic grasp planning using shape primitives. In *Robotics and Automation, 2003. Proceedings. ICRA '03. IEEE International Conference on*, volume 2, pages 1824–1829 vol.2.