Generative Perception for Robotic Grasping

Douglas Morrison, Peter Corke, Jürgen Leitner

AUSTRALIAN CENTRE FOR ROBOTIC VISION

The System

Generates a direct mapping from a scene to a representation of grasp affordances.

Anti-podal grasps are encoded as a likelihood of success (Grasp Score) and Grasp Angle for each input pixel.

Transformed into 6-DoF grasp poses using depth and surface normals.

Preliminary results on Cornell Grasping Dataset shown below (grasps extracted by locating local maxima in the Grasp Score).

Advantages

No sampling of grasp candidates – grasps are represented as a continuous space.

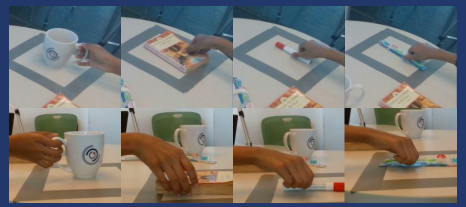
Better represents a robot's environment in terms of the distribution of grasp affordances compared to sampling/discretising the grasping space or generating a single grasp pose per scene.

Future work: integration to existing robotics techniques such as visual servoing or active perception – no more blind grasping.

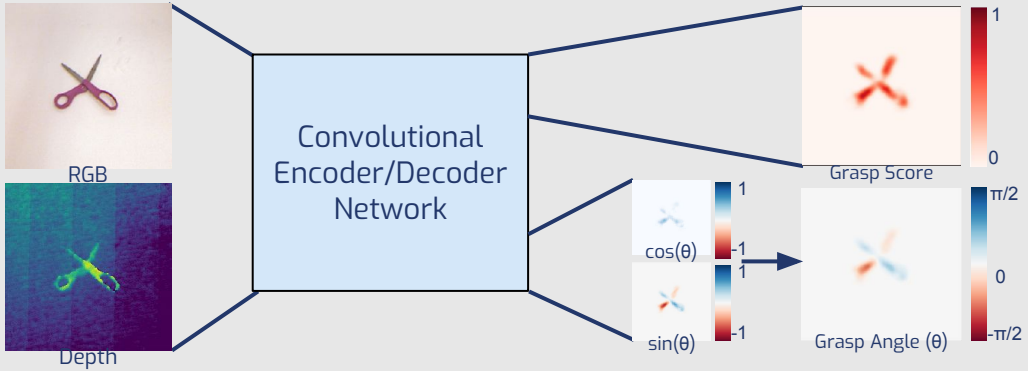
Extending to multi-fingered grasping

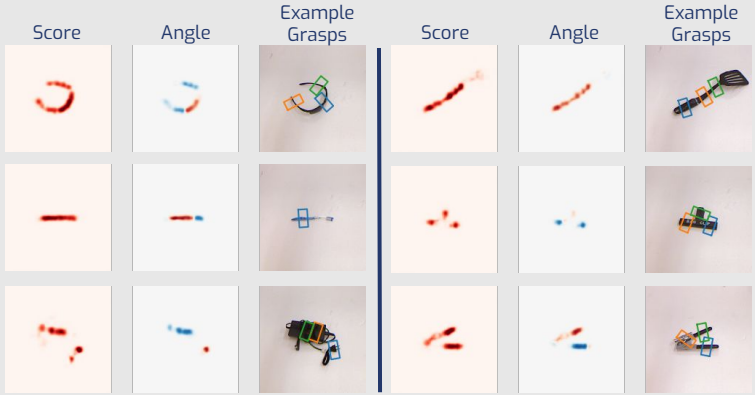
Future work: Learn to generate grasps by encoding finger-wise placement models – not possible with commonly used rectangle representation.

We have collected a video dataset of human grasping for this purpose.



One-to-one generative mapping from image-space to grasping-space.





Contact: douglas.morrison@hdr.qut.edu.au