Learning 3-D Object Orientation from Images

Ashutosh Saxena, Justin Driemeyer and Andrew Y. Ng
Stanford University

We consider the problem of estimating the 3-D orientation of objects from a single image, even in the presence of symmetries of the object. We apply our algorithm both to estimating orientation of new objects from a known object class, and to robotic manipulation, where the task is to grasp (pick up) an object from a previously unknown object class.

Estimating orientation is a fundamental problem in computer vision, but is difficult because (i) The space of orientations is intrinsically non-Euclidean and non-linear, (ii) The presence of symmetries means that orientation is ambiguous, in that there are multiple "correct" orientations.[1] These two properties make orientations difficult to learn using standard learning algorithms.

In this paper, we present a novel representation for orientation that is invariant to symmetries in the object. Our representation applies even to objects that exhibit arbitrary (rotational, reflective, axial, etc.) symmetries in 2-D or 3-D. Using this representation, we further develop novel learning and inference algorithms for estimating orientations of symmetric and asymmetric objects.

We first describe previous methods to learn 3-D orientation, and describe their deficiencies. Specifically, they fail when the range of angles considered does not lie within a small range.[2] In contrast, our method does not assume any restriction on the range of angles considered, and is accurate even in the fully general case.

We extensively evaluate previous methods, and our algorithm, on two tasks: (i) Estimating the 3-D orientation of a new object (drawn from a known object class), and (ii) Choosing the orientation of a robot arm/hand so as to enable it to grasp a new object (where here the previously-unseen object is drawn from a previously-unknown object class). We show that in all cases, our algorithm results in significantly lower error. We have also successfully applied these ideas to enabling our robot to grasp a variety of objects.

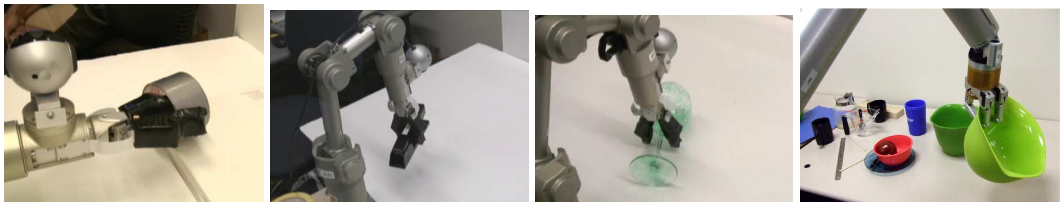

Figure 1: Figure shows the robots grasping a roll of duct tape, a stapler, a wine-glass and a cereal bowl.

References

[1] N. Fisher. *Statistical Analysis of Circular Data*. Cambridge University Press, 1993. 1

[2] P. Mittrapiyanuruk, G. N. DeSouza, and A. C. Kak. Calculating the 3d-pose of rigid objects using active appearance models. In *ICRA*, 2004. 1