# Refining Object Recognition, Localization, and Grasping with Tactile Feedback

Jonathan Binney, Kaijen Hsiao, and Matei Ciocarlie
Willow Garage Inc., Menlo Park, CA, USA Email: {binney, hsiao, matei}@willowgarage.com

***Abstract*—Recognizing, localizing, and grasping objects in unstructured environments is often difficult because of noisy sensor data, occlusions, and calibration error, which result in uncertainty in both object shape and pose. However, feedback received while attempting to grasp an object can be used to refine object recognition and localization results, which in turn can improve the likelihood of a successful grasp. To that end, we track probabilities for an entire set of object shape and pose hypotheses resulting from visual object recognition, select grasps that maximize the probability of grasp success based on those hypotheses, and refine the results of object recognition and localization based on tactile and proprioceptive feedback obtained while attempting to grasp the object.**

## I. Approach and Example

In the approach presented here, we assume that we have a database of objects to be recognized and grasped with both shape (mesh model) and pre-computed grasps. The flow chart and pictures on the left show a typical sequence of grasps: a can in front of the robot is seen by the robot's head Kinect. Visual object recognition algorithms are used to generate an initial shape and pose distribution for the object (shown in blue, offset from the object for visualization purposes). Based on the generated distribution, a set of grasps is sorted by estimated probability of success, and the first kinematically-feasible, collision-free grasp is executed. In this example, the object is shorter than expected, and so the robot grasps the air just above it. The robot's proprioception is used to eliminate particles with tall object shapes that should have been hit. A new grasp is planned based on the updated object distribution, and this time contact is made with both fingertips. The location of the fingertip contacts is used to update the distribution again, and now the robot knows the object's shape and pose, enabling it to re-plan a successful grasp for the final pickup.

## II. Object Detection

In our implementation, two object detection algorithms are used to generate object shape/pose hypotheses: the first is the Tabletop Object Detector, which uses ICP to recognize well-separated, rotationally-symmetric objects on a table. The second is TOD (Textured Object Detector), which uses a bag-of-features approach to recognize textured objects [1]. Both methods provide quality scores for their detection results, which can be used to prune hypotheses and generate the initial probability distribution for likely object shape/pose hypotheses using Naive Bayes (left side of the Bayes Net diagrams at center-right). Additional object shape/pose particles are generated by sampling poses around each detection result, to account for calibration error and sensor noise.

## III. Bayesian Grasp Planner

The Bayesian grasp planning module [2] selects grasps that maximize the probability of grasp success over the entire set of object shape/pose hypotheses. It assumes the existence of grasp generators that can produce a set of grasps to consider, which in this case is simply the union of the pre-generated grasps for all the object shape/pose hypotheses. It also assumes the existence of a grasp evaluator that can evaluate the quality of an arbitrary grasp given a particular object shape/pose hypothesis, which in this case is a regression-based grasp evaluator based on the stored grasps for each object. The probability of success for each proposed grasp is estimated by marginalizing over the shape/pose hypotheses (right side of Bayes Net diagrams in center-right). The figure shows grasps planned for a wine glass, based on several shape/pose hypotheses; the grasps shown at the bottom right are those most likely to work given all the hypotheses and their probabilities.

## IV. Belief Update and Sensor Models

Each grasp attempt results in an update to the robot's belief state, represented by the weighted set of particles containing object shape/pose hypotheses. Tactile sensors on the gripper provide both positive (fingertip sensor contact locations are expected to be at the surface of the object) and negative (contact sensor should not penetrate an object without feeling contact) measurements. The likelihood of positive (contact) measurements is represented using a Gaussian on the distance to the nearest surface of a particular shape/pose hypothesis. The likelihood of negative (no contact) measurements that would be inside the object is a similar Gaussian; negative measurements outside the surface are assigned constant likelihoods. No resampling step is used for the particles.

## V. Conclusion

If the goal is merely to grasp the object, this approach provides a framework for doing so in the face of uncertainty in object shape and pose. However, we may separately be concerned with the object's identity and the exact pose of the object in the hand, both to ensure that we have the correct object, and to improve subsequent object placement. Using the presented framework, we increase our likelihood of successfully grasping, recognizing, and localizing the object all at once, based on both visual and tactile sensing. See http://bit.ly/R0G7Y8 for a video of the presented approach.

## REFERENCES


[1] Object recognition kitchen. http://wg-perception.github.io/object_recognition_core/.

[2] Kaijen Hsiao, Matei Ciocarlie, and Peter Brook. Bayesian Grasp Planning. In *ICRA 2011 Workshop on Mobile Manipulation: Integrating Perception and Manipulation*, 2011.