

A Computational Framework for Visual Perception of Inertial Affordances

Walter A. Talbott and Javier Movellan

UC San Diego

Goal

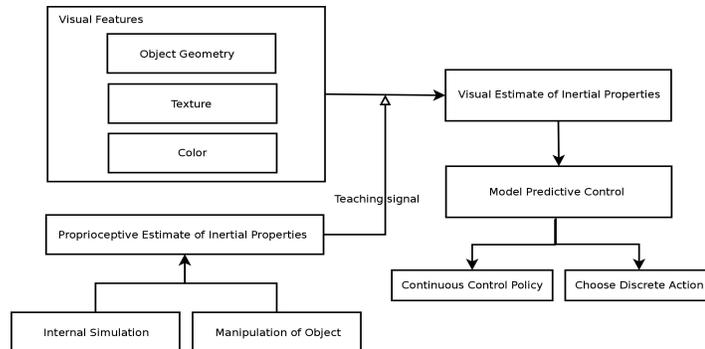
Computational framework that captures the human ability to quickly estimate how to use novel objects.



Which of these objects would you use to mash potatoes?
How would you grasp that object?

Proposed Approach

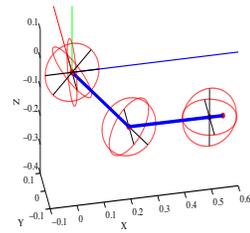
1. **Visual Features** to estimate inertial parameters
2. **Proprioceptive Experience** to teach module 1
3. **Model-Predictive Control** to run internal models and generate real-world behavior



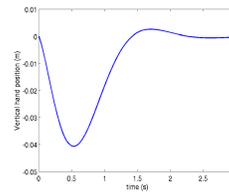
Implementation Overview



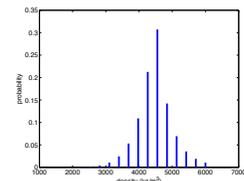
Use vision to estimate object geometry and material. Density and geometry give inertial estimates.



Use MPC to hold object at arm's length



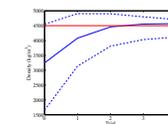
Proprioceptive experience: torque and acceleration



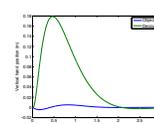
Update belief about material density to best match proprioceptive experience

Experiments 1 and 2: Learning

- Robot given 4 objects
- Robot tries to hold each object level at arm's length
- Vertical displacement of object endpoint is measured
- Hollow decoy object presented last

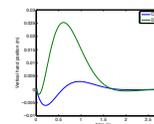
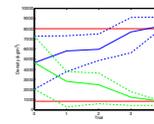
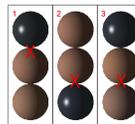


Density estimate improves



Exaggerated response to decoy

- 3 objects, all with same mass
- Procedure similar to Experiment 1
- Decoy object has appearance of first object, but center of mass of second object.

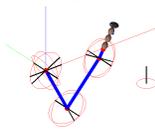


Experiment 3: Generalizing to novel objects and novel tasks

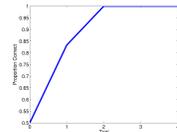
Hold training object level at arm's length



Choose grasp (top or bottom) on novel object that requires least energy to hit target with a given force. Never hold the novel object



Repeat with different random seeds, look at percent correct grasps chosen as a function of number of training objects held.



Conclusions

- Framework learns to manipulate new objects for new tasks
- Framework explains an existing experiment, suggests a new experiment, and predicts results
- Can be implemented on a robot