Imitation Learning of Whole-Body Grasps

Kaijen Hsiao, Tomás Lozano-Pérez

Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology

Pictures from www.inmagine.com

The Vision: Through demonstration, we can show our robots how to grasp and manipulate all sorts of objects in a human environment. While there are many algorithms to find stable fingertip grasps, sometimes a whole-body grasp is more useful for added stability, or for freeing up hands to grasp other objects (such as an underarm grasp). In addition, many objects/tasks require specific grasps. For instance, tools often require specific grasps in order to use them. Eventually, we would like to get to the point where we could show a robot how to use a power drill once and then have it adapt that grasp to all manner of power drills, jigsaws, and other similar-shaped objects. Or show it a hook grasp of a suitcase and have it recognize handles on other objects and grasp them the same way, or show it how to empty a dishwasher and have it be able to do the same on a new dishwasher configuration.

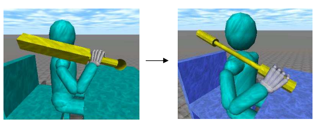
Our work: Below we describe our current working framework, which concentrates on adapting demonstrated whole-body grasps of simple objects modeled by up to 3 primitives. We are working on extensions to arbitrary mesh objects, which would enable us to focus on tool use and tasks such as emptying dishwashers.

ABSTRACT

Humans often learn to manipulate objects by observing other people. In much the same way, robots can use imitation learning to pick up useful skills. A system is detailed here for using imitation learning to teach a robot to grasp objects using both hand and whole-body grasps, which use the arms and torso as well as hands. Demonstration grasp trajectories are created by teleoperating a simulated robot to pick up simulated objects. When presented with a new object, the system compares it against the objects in a stored database to pick a demonstrated grasp used on a similar object. Both objects are modeled as a combination of primitives—boxes, cylinders, and spheres—and by considering the new object to be a transformed version of the demonstration object, contact points are mapped from one object to the other. The best kinematically feasible grasp candidate is chosen with the aid of a grasp quality metric. To test the success of the chosen grasp, a full, collision-free grasp trajectory is found and an attempt is made to execute in the simulation. The implemented system successfully picks up 92 out of 100 randomly generated test objects in simulation.

Problem Statement

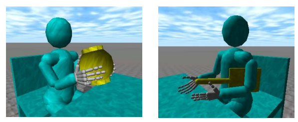
- We wish to enable a simulated robot to learn whole-body grasps by imitation
- A human demonstrates picking up a simulated object
- The robot performs the same grasp on new objects that are different than the training objects



What Are Whole-Body Grasps?

Grasps that can use surfaces besides just fingertips

- Enveloping grasps
- Two-hand grasps
- Under-arm, over-shoulder grasps



Why is this hard?

- Typical methods of finding grasps:
  - Finding individual contact locations:
    - Any contact can go anywhere on object surface
    - General grasp construction/optimization problem must search # of grasps exponential in number of contacts
  - Using taxonomy-based, heuristic methods:
    - Unclear how to generalize to complex new objects
- We want to do complex whole-body grasps:
  - A single grasp can have up to 38 contacts
  - Must find grasp sequences (multiple, linked grasps)
- Thus: learning—adapting demonstrated grasps

Our Approach To Adapting Demonstrated Grasps

Plan Grasp
- Reduce demonstration contacts to representative set
- Generate pre-grasp locations by assuming target object is template after undergoing transformations on primitives
- Pick best kinematically feasible pre-grasp location using grasp quality measure

Test in head
- Test plan in physics-based simulation, wrapping hands around object with low-level controllers

Test on robot
- Test plan with real robot (but we have no robot, so we skip this step)

Modeling Objects with Primitives

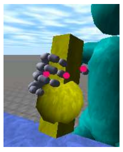
Simplifies search drastically
- Symmetries provide rotational alignments
- Individual primitives provide 'chunking'

Objects currently limited to those modeled by 3 primitives in a line (axis-aligned)

Models generated by hand

Representative Contacts

- Finger contacts are not independent; thus, can use representative contacts
- Akin to concept of virtual fingers
- Track 3 points:
  - Tip of middle or pointer finger
  - Tip of thumb
  - Palm contact / location nearest middle knuckle
- Given representative contacts, use optimization to find pre-grasp location



Outline of Grasp Adaptation Process

1) Demonstrate template grasp sequences

2) Choose template grasp sequence with NN

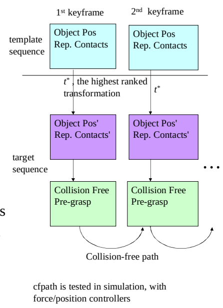
3) Map grasp contacts from template to target using transformations

4) Choose best grasp sequence with a quality measure

6) Generate col. free keyframes

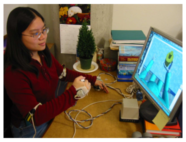
7) Generate col. free trajectory

8) Execute grasp in physics-based simulation



Demonstrating Grasps

- Nest of Birds
- Keyframe recorded when a new contact with the object is made or broken:
  - object position
  - arm positions
  - locations of contacts on object
  - locations of contacts on body parts/table
- Open Dynamics Engine (ODE) for physics of simulated world

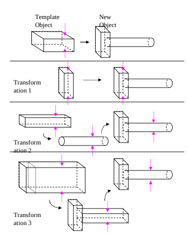


Picking a Template Grasp

- Pick a similar object from the database
- Grasp the new object in the "same way"
- Nearest-neighbor classification
  - Object dimensions
  - Object mass
  - Inertia in each dimension
  - Object z-axis (alignment of primitives)

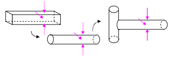
Object Transformations

- But: there are many ways of transforming one object to another!
- How do we map contacts through a transformation?
- And how can we tell which resulting grasp is "best"?

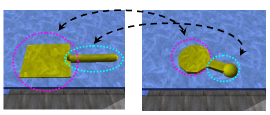


How To Adapt Grasp Contacts From Template To New Object?

- New object should be "similar" to chosen template
- Imagine that new object is just the template object after having undergone a series of transformations
  - Expanding/shrinking
  - Morphing between primitives
  - Adding on pieces
  - Removing pieces
  - Splitting into multiple pieces
  - Joining multiple pieces

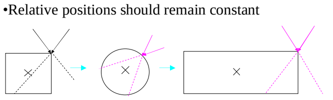


- Now imagine grasping the template object while it undergoes these transformations—if the transformations are not too extreme, grasp should still work (small adaptation of old grasp)
- Equivalent to saying that 'chunks' of old object are grasped in the same way as 'chunks' of new object



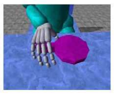
Mapping Contacts Through Transformations

- One transformation is a mapping of 'chunks' on template to 'chunks' on target
- Contacts should be on the appropriate 'chunks' on the new object
- Relative positions should remain constant



Choosing the Best Grasp

- Do optimization to find joint angles that position the hand to best make each proposed set of contacts (IK for hand location)
- Estimate kinematically feasible contact positions (close fingers)
- Eliminate obviously infeasible grasp candidates (due to major collisions or being out of reach)
- Choose the best remaining grasp candidate according to a grasp quality measure

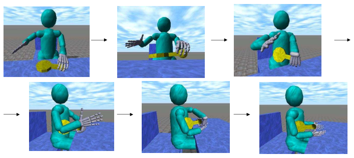


Quality = 0.179109

Quality = 0.000000

Finding Collision Free Keyframes/Trajectory

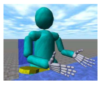
- Once the proposed grasps are ranked and the best chosen, pre-grasp locations (keyframes) can be adjusted to eliminate minor collisions
- Then a collision-free trajectory to traverse the keyframes can be found using a probabilistic roadmap



Executing Trajectories

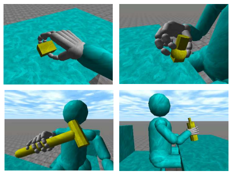
- Once a desired trajectory is found, it must be executed by the robot in the simulated world
- Hybrid position/force control to move arms
- Low-level hand controllers use force control to wrap around object
- If object is dropped (or is never successfully picked up), grasp has failed

Having a collision-free trajectory with a high quality grasp does not guarantee success in a dynamic world!



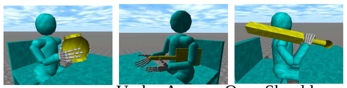
Template Grasps

Our current system has seven demonstrated grasps with which to pick up new objects:



Precision Grasps From Top and Side

Palm Grasps From Top and Side

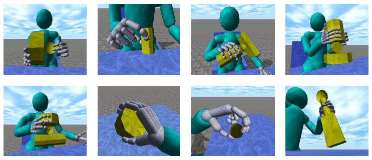


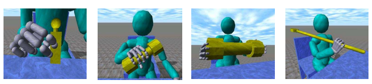
2-Hand Grasp | Under-Arm Grasp | Over-Shoulder Grasp

Results of Current System

Using the seven template grasps, our system successfully grasped 92 of 100 randomly generated objects

Examples of grasps of some randomly generated objects





Examples of grasps of some hand-generated objects