Understanding and applying human grasping to artificial manipulators

Yuzuko Nakamura

CMU-CS-18-116
October 2018

School of Computer Science
Department of Computer Science
Carnegie Mellon University
Pittsburgh, PA

Thesis Committee:

Nancy S. Pollard, Chair
Jessica K. Hodgins
Roberta L. Klatzky
Carol A. O'Sullivan, Trinity College Dublin

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy.

This research was sponsored by the National Science Foundation under grant numbers IIS-1218182, IIS-1344361, and IIS-1637853; the Boeing Company under grant number CMUBAGTA1AA017415; and Disney Research. The views and conclusions contained in this document are those of the author and should not be interpreted as representing the official policies, either expressed or implied, of any sponsoring institution, the U.S. government or any other entity.

Keywords
grasping, manipulation, grasp taxonomies, dexterous manipulation, bimanual manipulation, motion capture, high framerate video, soft robotics

Abstract

Though a small part of the body, the human hand is complex and remarkably versatile and multipurpose. Much work has gone into understanding the hand, such as understanding the physical capabilities of the human hand, how humans develop manipulation skills throughout the lifespan, how people translate task requirements into grasping strategy, and so on. Despite that, human manipulation is still not well understood. For example, how many grasps or manipulation actions do people use in daily life? How often and under what circumstances do people use both hands simultaneously instead of one? A better understanding of how humans grasp can improve our ability to control robotic hands, which are still far behind human hands in dexterity.

In our work we have used a variety of methods to observe how humans grasp and manipulate in natural, everyday settings. We have used photos taken throughout a normal day; high-framerate video in a specific setting (that of a convenience store); and cameras and motion capture systems in the context of a controlled experiment involving transporting a bowl from one location to another. In these studies we found that a single grasp pose can be used for a variety of actions, were able to observe the grasping process in detail, and found that minimizing body rotation plays a large role in the use of one hand vs. two in transport tasks.

We propose applications of some of the main findings of these studies to the goal of improving the success of grasping performed by robotic hands and virtual characters. In particular, we propose using the detailed grasping behavior found in the high-framerate video to create a simple soft hand capable of executing several ways of turning precision grasps into power grasps.

This work thus presents the results and insights from investigations of human manipulation and lays out ways in which those insights can be used to improve the capabilities of artificial manipulators.

Acknowledgments

Thank you to all the people who have advised me, inspired me, supported me, and advocated for me: to Nancy Pollard, Cinda Heeren, Margaret Fleck, Carol O'Sullivan, Jessica Hodgins, and Roberta Klatzky.

Thank you to the people who helped teach me early on how to do research: Junggon Kim, Jia Liu, Xiaoyu Feng, and Daniel Troniak. Thank you to Justin Macey for all the mocap help over all the years, and to Jonathan King for the fabrication help.

Thanks to the friends I made along the way, who have all helped me out of pinches: Dana, Yair, Sarah, John Wright, Danny, Deby, Kristy, Stefan, and João. Thanks to Se-Joon for being a part of this journey from start to finish. Best of luck with what comes after the PhD!

Thank you to the entire Graphics Lab for honing my presentation skills. Thanks to Yumika Amemiya for typesetting help and additional figures for this thesis.

Finally, thank you to Liam Bright, who has over the years provided me with an unbelievable amount of emotional support.

Contents

Listing of figures

Listing of tables

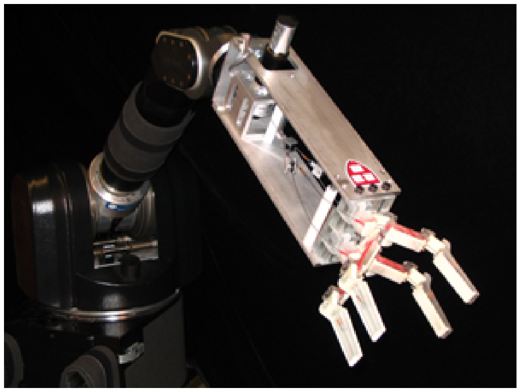
1
Introduction

A major goal in the field of robotics is the creation of robots capable of using their hands as humans use their hands. The ability to do so would allow robots to be useful in environments too dangerous for humans and to assist people in the home who have difficulty with activities of daily living.

People have pursued various lines of research to get closer to achieving this goal. For example, improvements in the design of robotic hands/grippers, computers' visual understanding of images (computer vision), and control schemes robust to uncertainty have yielded improvements to the ability of robotic hands to grab and manipulate objects.

Although we use the human hand as a model and as the goal for robotic manipulation, anthropomorphic robot hands that resemble the human hand are complicated and hard to control. In practice, the most successful, reliable hands for grasping have been simpler and non-anthropomorphic ones, such as the SDM hand [Dollar and Howe, 2010] (Fig. 1.1a) and the coffee-grounds-filled universal gripper [Amend et al., 2012] (Fig. 1.1b).

Part of the problem is the complexity of control and the lack of sensors on robotic hands, in contrast to human hands which have a dense network of touch receptors in the skin of the hands. However, part of the problem is that we do not have a detailed enough understanding of how humans use and control their hands to know how much robotic manipulation is limited by these shortcoming.



(a)

(b)

Figure 1.1: Non-anthropomorphic hands. (a) SDM hand. (b) Universal gripper.

In this work, we seek to partially fill this gap in our understanding of human manipulation and apply the knowledge gained to the task of creating artificial manipulators. We observe grasping in everyday situations in order to understand how people use their hands to manipulate their environment, and we conduct controlled experiments to understand how environmental cues change people's choice of grasping strategy. We then apply the results of those investigations to the control of artificial hands.

1.1 List of definitions

Dexterity In general, refers to the ability of a hand to successfully manipulate its environment, which can include both *grasping* and *manipulation*. Sometimes, dexterity refers specifically to minute motions of the fingers needed to perform difficult manipulation tasks.

End effector The point of interaction with the environment for a robot manipulator. This word can refer to a robot hand or finger, which may not necessarily resemble a human hand or finger.

Grasping Generally refers to when a hand is able to hold an object securely in a gravity-independent way. However, the notion of grasping can also be extended to cases when a hand is able to secure an object with help from the environment or gravity.

Grasp pose The shape of a hand when grasping an object.

Intrinsic manipulation / within-hand manipulation *Manipulation* that occurs entirely within the hand without help from the environment or arm.

Manipulation Any way of using one's hands to effect changes in the environment, such as changing an object's 6-D position and orientation, applying forces to objects, or using objects as tools.

Pick and place A type of *manipulation* that involves grabbing and moving objects. Mainly involves forming a stationary grasp pose that allows an object to be securely attached to the arm via the hand and then moved with body or arm movement.

Prehension Synonymous with grasping. Literally laying hands on something and acquiring it into the hand.

1.2 Research questions

Research question #1: How many grasp poses do people use in daily life? An answer to this question helps us understand how complicated the space of human grasping is, and also what may be needed to create manipulators that are as dexterous as humans. Various taxonomies have been created by observing engineers as they work [Cutkosky, 1989] or asking subjects to grasp a wide variety of objects [Kamakura et al., 1980]. We add to this line of research by observing the grasps used during the daily life of two subjects and supplementing this with more occupation-specific grasps.

Research question #2: How do people acquire objects into their hands? This question is concerned with understanding the whole process of grasping – given a final hand shape (final grasp), what intermediate steps are taken in order to lift the object into the final grasp? Here, we take advantage of the greater ubiquity of slow-motion (120 frames per second) cameras to better answer this question. We record video that details the grasping process of a subject in a natural setting as they grab and replace items in a convenience store.

Research question #3: What factors influence grasp strategy? The first two questions deal with the "how" of grasping – here we deal with the "why". What cues in the environment do people pick up on that affect their choice of grasping strategy? Possible cues include properties of the object (its size, weight, shape, material, etc.) as well as aspects of the environment or task (object location, presence of obstacles or constraints, difficulty of desired task, etc.). A rich body of research exists that investigates the connection between cues and grasping strategy (e.g. [Gilster et al., 2012; Rosenbaum et al., 2010; Cesari and Newell, 2000]). We add to this line of research by focusing on factors that affect use of one or both hands when grabbing and moving an object. The factors tested were object size and weight, object starting position, and the importance of maintaining balance in whether people choose to use one or two hands in a bowl-transporting task. We also consider how choice of hand shape is affected by these factors.

Research question #4: Can we use the insights from observing human grasping to make better graspers? We can use our observations of human manipulation in order to inform the design and control of robotic hands. Observations of the grasping process reveal that people make use of strategies capable of turning a weak or environment-aided grasp into a final power grasp. We created a robot hand capable of performing these actions.

1.3 Contributions

This thesis seeks to better understand the considerations underlying grasping and manipulation, and to use those insights to advance the development of dexterous robotic hands. The contributions of this work are as follows:

- Creation of a database containing 179 actions from daily life, useful for benchmarking artificial manipulators (Chapter 3).

- Insights from the above database about grasping that there are more hand shapes than previous taxonomies would suggest, and that the same grasp can be used in multiple ways (Chapter 3).

- A method for annotating action that can distinguish the different ways in which a single grasp may be used in practice by adding information on the type of force being applied, motion and force direction, and whether there are constraints on the motion or force (Chapter 3).

- Using handheld slow-motion cameras to create a video dataset containing 91 slow-motion picking and placing tasks in clutter. From that dataset, we also created a set of annotations for eight of the actions through the lenses of three existing manipulation taxonomies and two other lenses focused on errors and the role of contact. We also compiled a list of ways within-hand movements were used by the actor captured in the dataset. All three are useful for benchmarking either artificial manipulators or vision systems (Chapter 4).

- Insights from the above database about grasping, such as that intrinsic motions are utilized in simple pick-and-place tasks; that contact with the environment is exploited to guide or aid manipulation; and errors occur frequently and are recovered from quickly (Chapter 4).

- Evaluation of the ability of current taxonomies to describe manipulation. Taxonomies for grasp poses, within-hand movements, and describing general manipulation are all useful for describing grasping strategy, but have limitations in their ability to capture common variations on grasp poses, environment-aided manipulation, and multitasking (Chapter 4).

- Investigation into the reasons why people use one or two hands when transporting objects, useful for creating realistic animated characters and robot manipulators suitable for interacting with humans. The insight from this investigation is that different transport strategies involve different amounts of rotation or amounts of stability, which drives selection of transport strategy (Chapter 6).

- Investigation into the unimanual and bimanual grasps people use on a single object (a bowl), and why some grasps are preferred to others, useful for creating realistic behavior in artificial graspers. The main insight from this investigation is that there are many ways to grab a simple bowl, some of which become more appealing depending on the size, weight, and presence of an obstacle in the center of the bowl (Chapter 7).

- A process that goes from observed manipulations to a physical hand capable of executing those manipulation actions. As a proof of concept, we create a soft foam hand capable of executing simple precision-to-power regrasping actions (Chapter 8).

2

Review of literature

In this chapter, we review four areas of grasping research that we build on in this thesis: grasp and manipulation taxonomies (Section 2.1), experimental investigations of grasping (Section 2.2), and design and control of artificial hands (Section 2.3). The first two sections touch on ways people observe and learn from human manipulation. The third section briefly reviews the field involving the creation of robotic manipulators and the control of artificial hands.

2.1 Grasp and manipulation taxonomies

A large amount of work has gone into cataloguing and categorizing the grasp poses people use – this categorization work has led to the development of various taxonomies that seek to map the space of grasping and manipulation. We first start with earlier literature that seeks to categorize and understand the space of human grasp poses, and then move on to work that broadens the scope of classification to manipulation, movement, and non-human hands.

2.1.1 Grasp taxonomies

Grasp taxonomies based on shape and function have existed for many years. The earliest well-known such taxonomies are those of Schlesinger [1919] and Napier and Tuttle [1993], which led the way in discriminating major hand shapes and grasp functions. Grasp taxonomies have been developed for tasks of everyday living, including those of Kapandji and Honoré [1970], Edwards et al. [2002], and Kamakura et al. [1980]. Kamakura and colleagues, for example, classified static prehensile patterns of normal hands into 14 patterns under 4 categories (power grip, intermediate grip, precision grip and grip involving no thumb). They illustrated detailed contact areas on the hand for each grasp and analyzed for which objects the grasp may be used.

Perhaps the most widely cited taxonomy in robotics is that of Cutkosky [1989], which includes 16 grasp types observed in skilled machining tasks. The Cutkosky taxonomy consists of a hierarchical tree of grasps, with categories classified under power and precision. Moving from left to right in the tree, the grasps become less powerful and the grasped objects become smaller. Zheng and his colleagues [2011] used this taxonomy to capture the daily activities of a skilled machinist and a house maid, giving for the first time a count of how frequently different grasps are used.

Feix et al. [2009] recently developed a comprehensive taxonomy of grasps that brings together previous research with their own observations. They propose a definition of a grasp as follows: “A grasp is every static hand posture with which an object can be held securely with one hand,” and identify 33 grasp types that are distinct from one another and fit this definition. We use this work as a starting place in Chapter 3.

2.1.2 Manipulation taxonomies

Going beyond static grasping, a number of taxonomies have been developed to express manipulation actions as well. A classic work in studying manipulation is Elliott and Connolly’s [1984] observations and categorization of ways people were able to manipulate objects entirely within the hand. This is called within-hand, intrinsic, or dexterous manipulation. They identify three classes of intrinsic movements: simple synergies such as squeeze where all participating fingers perform the same motion (flex, extend, abduct, or adduct); reciprocal synergies such as roll where participating fingers manipulate the object with complementary movements; and sequential patterns that rely on a time sequence of movements, such as a rotary stepping motion of the fingers to change contact positions on the object. Ma and Dollar [2011] seek to classify within-hand manipulation in a way that is not tied to the morphology of the human hand, separating manipulation into six broad strategies capable of changing an object's orientation: using another grasper or environment to ungrasp and regrasp; exploiting the hand's kinematic redundancy to reorient the object without contact changes; sequential patterns of contact changes; pivoting around two contact points; manipulation featuring rolling; and manipulation featuring sliding.

Wörgotter and colleagues [2013] take an object-centric approach. Focusing on actions of bringing together and breaking apart, they identify 30 fundamental manipulations that allow sequences of activities to be encoded. Bullock et al. [2013] take a hand-centric approach. They encode manipulation at a more abstract level, focusing on motion of the hand and relative motion of the hand and object at contact. Their classification scheme does not assume a specific hand design and is applicable to all stages of the grasping process including pre-grasping.

Other taxonomies focus on a particular subset of manipulation scenarios. Chang and Pollard [2009] classify manipulations prior to grasping, with a focus on how the object is adjusted, considering both rigid transformation and non-rigid reconfigurations. Heinemann et al. [2015] also focus on elements that may be present or absent in the grasping process, such as contacting the support surface with fingers, sliding or rotating the object along the support surface, edge grasps (moving the object to the edge of the surface to expose its underside), and flips that raise one edge of a thin object. Dafle et al. [2014] look at the ways in which a hand in combination with gravity or contacts with the environment can manipulate an object, calling this extrinsic dexterity.

Guiard [1987] focused on bimanual manipulation, laying out three different types of bimanual manipulation: independent (hands performing tasks not requiring coordination), symmetric (hands coordinating on one task with similar roles), and differentiated (hands coordinating but given different roles). Various taxonomies of bimanual manipulation [Grunwald et al., 2008; Surdilovic et al., 2010] also start from this framework.

Mechanisms other than grasping and manipulation taxonomies exist for classifying movement, such as Laban Movement Analysis [Abe and Laumond, 2014] and the Facial Action Coding System (FACS) [Ekman and Friesen, 1978; Cohn and Ekman, 2005]. Observations of great apes are also of interest [Torigoe, 1985]. For example, Byrne et al. [2001] observe over 200 primitive actions, such as pick-out, pull-apart, and rotate-adjust, as necessary to describe feeding behaviors of mountain gorillas. We draw upon many of these more general taxonomies to describe important aspects of motion in Chapters 3 and 4, and contribute manipulation taxonomies of our own in those chapters.

2.2 Experimental studies of manipulation

Taxonomization of manipulation is one way to understand how manipulation works. Another robust line of research seeks to understand human manipulation by varying object and task properties in a controlled way and observing how those variations affect grasping choices such as hand pose or contact points. We first review studies that study grasp poses, and then focus on work in the area of bimanual manipulation.

2.2.1 Effect of object/task properties on grasp pose

There is a very large literature investigating how properties affect the pose people use to grasp the object. The properties investigated include object shape [Touvet et al., 2014; Gilster et al., 2012; Sartori et al., 2011], object size [Park et al., 2014; Cesari and Newell, 2000], presence of obstacles [Voudouris et al., 2012], shape and size of affordances such as handles [Fuller and Trombly, 1997], start/goal location [Touvet et al., 2014; Voudouris et al., 2012], and task to be performed after grasping [Crajé et al., 2011; Sartori et al., 2011]. The measure of grasp can be either full hand pose described by joint angles [Park et al., 2014; Touvet et al., 2014], choice of contact points for each finger [Crajé et al., 2011; Gilster et al., 2012; Sartori et al., 2011; Voudouris et al., 2012], or number of fingers used [Cesari and Newell, 2000; Fuller and Trombly, 1997; Gilster et al., 2012].

In our work in Chapter 7, we study the effect of object size and weight, start and goal location, and presence of a balance tube inside the bowl on grasp pose. Unlike the above work, we classify grasp pose using broad categories of hand shape – similar to the work of grasp taxonomies – and not precise quantitative data like joint angles. Our work is also unusual in its use of bowls as the manipulated object. Most of the studies above use simple objects like cylinders [Park et al., 2014; Gilster et al., 2012], spheres [Voudouris et al., 2012], or cubes [Cesari and Newell, 2000], or everyday objects like bottles [Crajé et al., 2011; Sartori et al., 2011] or mugs [Fuller and Trombly, 1997].

2.2.2 Effect of object/task properties on bimanual usage

Cesari and Newell [2000] investigate how object size and weight influence the number of hands used to grasp the object. By having participants grasp cubes of different sizes and densities, they fit an equation to describe the transition between unimanual and bimanual grasps. Rosenbaum and colleagues [2010] investigate how goal location affects the usage of left, right, or both hands when grasping or placing for a Tupperware-stacking task.

Although we are interested in the use of bimanual manipulation skills in adults in this thesis, some studies on infants explore bimanual hand usage in more detail than adult studies, and the result of these studies on infants may apply to adults as well. Studies such as that of Kimmerle et al. [2010] observe the amount of time very young children spent on different kinds of manipulation (unimanual, symmetric bimanual, and differentiated bimanual). Greaves and colleagues [2012] review a body of literature investigating what sorts of toy properties can encourage various kinds of bimanual manipulation (bimanual reaching, holding, handing off, turning, symmetric, and asymmetric) in developing children.

We contribute to existing work on what factors lead to the selection of bimanual strategies in Chapter 6. We focus on two understudied aspects of bimanual grasping: (1) the role of a balance requirement on grasp strategy, which has been studied in the context of pre-grasp strategy [Chang et al., 2009] but not bimanual strategy, and (2) when and how the hand-off strategy is used, which has not been considered outside of infant studies.

2.3 Design and control of artificial hands

A major application of manipulation research is using it to design artificial hands.

The creation of modern robotic hands goes back to at least the early 1980s with the creation of the three-fingered Stanford/JPL hand [Salisbury and Craig, 1982] and the anthropomorphic Utah/MIT hand [Jacobsen et al., 1986]. Recent anthropomorphic hands like the Shadow Hand [Shadow Robot Company, 2018] and the DLR/HIT Hand [Liu et al., 2008] match the human hand in terms of the degrees of freedom in the joints, and there has been some success in getting these hands to manipulate (rotate) objects within the hand [Kumar et al., 2016; Rajeswaran et al., 2017]. However, the rigid material of these hands makes control somewhat fragile and sensitive to perturbations. Adding compliance to the hand promises to add stability to hand motions. Soft robotic hands have been getting more attention for their ability to conform to unknown geometry, which not only might result in better performance, but also simplify the design of the hand and reduce the costs of making it. The RBO Hand 2 [Deimel and Brock, 2016] and the soft gripper by Zhou et al. [2017] are two recent examples of compliant hands that make use of pneumatic actuation, while King et al. [2018] create inexpensive tendon-driven hands made of foam. Work has also been done to design soft robots with a particular manipulation goal in mind. For example, Deimel et al. [2017] co-design the morphology and control of a pneumatic hand to improve grasping performance, and Bern et al. [2017] provide a way to create plush tendon-driven robots given posing goals.

In our work, we aim to design soft hands for the purpose of accomplishing dexterous tasks. We make use of the foam hands in King et al.'s work [2018] and Bern et al.'s Soft IK framework [2017] to create a dexterous robot hand in Chapter 8.

3

Everyday grasps in action

Grasping has been well studied in the robotics and human subjects literature, and numerous taxonomies have been developed to capture the range of grasps employed in work settings or everyday life. But how completely do these taxonomies capture grasping actions that we see every day? Our goal with this work was to build a taxonomy / database that captured at least 90 percent of everyday grasping and manipulation actions. Towards this goal, two subjects recorded all actions accomplished during a typical day, with a focus on critical humanoid robot capabilities such as home care and manipulation in unstructured environments such as a home or workplace. For each observed grasp or manipulation action, our subjects attempted to classify it using the Comprehensive Grasp Taxonomy of Feix and colleagues [2009]. In all, 179 distinct grasping actions were captured and classified.

As a result of this study we found that many grasping actions could be classified in the existing taxonomies. However, we also found that a single grasp could be employed in very different actions. Existing taxonomies did not consider the differences between these actions. To capture those differences, we propose an extended set of annotations related to features of the grasps in action: force (§3.2.2; §3.2.3; Table 3.1), motion (§3.2.3), and flow (§3.2.4). Our goal for this annotation scheme was to communicate motion, force, and flow information as precisely as possible while still allowing individuals with light training to understand and classify grasps or communicate differences to a robot. In addition, we found 40 grasp types which could not be well captured by existing taxonomies, including actions of pushing, grasping while pressing a button or lever, and grasping with extension (inside-out) forces. We believe our database is an improvement on our prior work, because we characterize human grasps by taking into account forces and motion exerted after a grasp is achieved. These added properties have intriguing similarities to aspects of dance notation such as Laban Movement Analysis [Newlove and Dalby, 2003], which has been long developed to describe motion and action, but does not focus on grasping. They also may tie into existing impedance [Hogan, 1985] and operational space controllers [Khatib, 1987] used in robotics.

This chapter describes our complete process, our annotation scheme, highlights from the full database (viewable online [Liu et al., 2014a]), and connections to Laban notation and robotic control schemes that may allow this work to bridge the gap between describing human manipulation and prescribing robotic manipulation.

3.1 Methods

To begin with, we studied previous literature that measured self-care and mobility skills for patient rehabilitation [Kopp et al., 1997; Collin et al., 1988; Linacre et al., 1994; Pollock et al., 2006]. The measured skills listed in these papers such as dressing, eating, and grooming, were useful to our study because they covered most of the typical and important tasks humans need to do even for those who are disabled. Our initial list of actions was a union of the tasks mentioned in those papers; however, we realized that many patients only needed a robot to do a certain amount of ancillary work and that not all everyday motions were captured in these studies. In fact, in work like Choi et al. [2009] where tasks are ranked by importance, tasks like buttoning, putting on socks, and personal hygiene are discarded because they received a low ranking and are difficult for a robot to accomplish. These less important tasks are not only part of daily life but also require the use of hands, and so are especially important to our study.

We next observed several people's life from when they woke up in the morning until when they went to bed at night. These people included teenagers, adults, and elderly people. We captured all the hand gestures that the person would use and all the motions into hundreds of tasks. However, we felt this wasn't enough since there are many experienced hand gestures people are capable of doing but may do less commonly than everyday life, and that the task collection so far was biased toward the office settings of the subjects. Therefore, we expanded our task list to include specific tasks that people from different careers would accomplish in their workplace.

After that, we further separated the compound tasks into small task components and movement pieces, like what Kopp et al. did [Kopp et al., 1997]. For example, wearing a T-shirt was broken down into three basic tasks: (1) arms in T-shirt sleeves, (2) grab the neck hole and move head through neck hole, and (3) pull down and straighten shirt. We collapsed similar gestures together and classified these movements into the existing 33-grasp database of Feix et al. [Feix et al., 2009; Feix et al.]. When we encountered daily-use hand gestures that were not in the basic database, including grasping, pressing, squeezing and lifting, we added them to the database.

Our final database contains 73 database categories, of which 50 are grasp types, 4 are press types, 10 are grasp and press types, 2 are extend types, and 7 are other hand types. We also illustrate where each movement may be used in daily life with corresponding pictures. The database can be accessed at: http://graphics.cs.cmu.edu/projects/graspsinaction/database.html.

3.2 Overview of annotation system

Fig. 3.1 shows the classification we've developed in order to distinguish the different manipulation actions we've encountered in our observations. The focus of previous literature has generally been on hand shape (highlighted in purple). With our observations, we divided all the different tasks by four general features: (1) hand shape, (2) force type, (3) direction, and (4) quality. Object properties are also factors that influence the hand shape and motion, but these relationships are not made explicit in our database. In contrast to traditional taxonomy research, which (aside from Wörgotter et al. [2013]) focuses mostly on static hand shape, our research focuses on motion related grasp tasks: in our database, both force and motion properties affect the action of a simple task. The rationale behind this focus on motion came about when we separated all the small tasks into the existing grasp taxonomy of Feix et al. and realized that a wide variety of tasks belonged to one grasp type but involved very different motion.

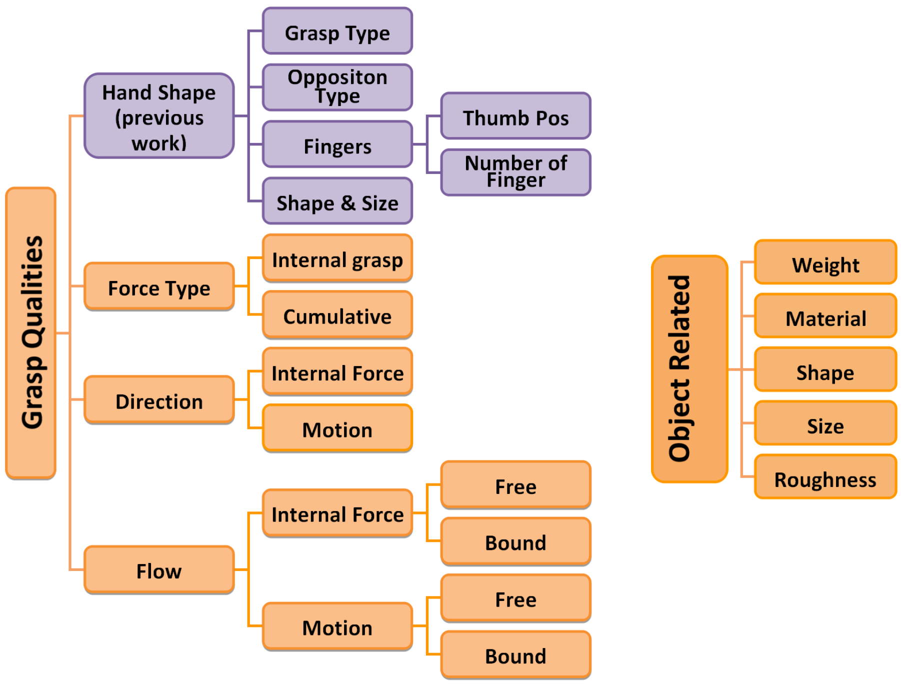


Figure 3.1: Simple Classification of the Database

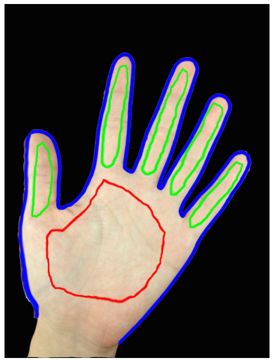


Figure 3.2: Palm, pad, side

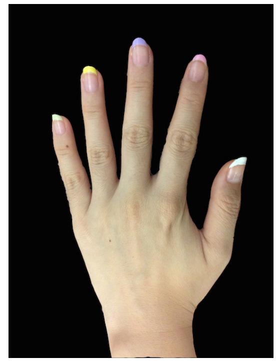


Figure 3.3: Back

3.2.1 Hand shape

Our classification of hand shape comes directly out of Feix et al., combined with ideas out of Napier [1956]. In Feix et al., they separated all different hand shapes by certain characteristics: type of the grasp, opposition type, thumb position, and which fingers are involved. The shape and size of the hand during each grasp is from Napier [1956].

For example, the type can be power grip or precision grip, or intermediate which is in between. A power grip is usually applied by partly flexed fingers and the palm with countering pressure, while a precision grip is more of a pinching of the object between fingers.

Opposition type means which part of the hand is used mostly. It includes palm (red in Fig. 3.2), pad (green), side (blue), and back (Fig. 3.3).

For the fingers, the thumb position is classified as ABD, ADD, EXT, or FLX (Fig. 3.4). It is also important to indicate which fingers (2: index finger, 3: middle finger, 4: fourth finger, 5: little finger) are used in each gesture.

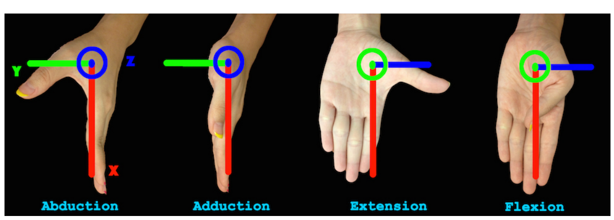


Figure 3.4: Local coordinates of all the types (left hand)

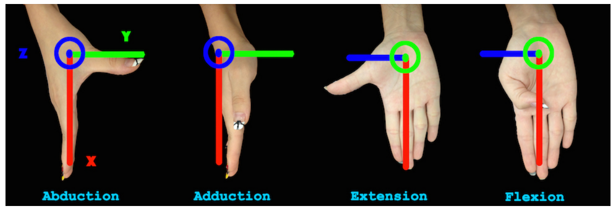


Figure 3.5: Local coordinates of all the types (right hand)

From another point of view, we can separate the gesture by the shape and size of the object we hold, like ball, large/medium/small diameter cylinder, disk and so on [Napier, 1956].

3.2.2 Force Type

There are many different ways in which forces can be distinguished or described: axis direction, the magnitude of the force, where force is being exerted, and so on. However, we found that describing forces using verbs from the English language made it clear what force is being used. We use 20 verbs to describe these forces (Table 3.1).

Although we don't make a distinction in our database, it's interesting to note that these force words imply an internal grasp force (exerted by the hand), or a cumulative / external force (exerted by the wrist or whole arm), or both. Table 3.2 shows two examples of internal forces (squeezing a tube of toothpaste and grabbing the handle of a pan). Table 3.3 shows two examples of cumulative forces: shooting a basketball and pushing down on a door handle. Both tasks involve the internal force of grabbing while the cumulative force is shoot or press.

In our database, both force and motion are important. For this reason, "grab" and "hold" are not the same, even though they feature the same motion (i.e. no motion). We define grab as touching or securing an object that is resting on a surface. We define hold with a gravity factor, where the hand/arm is applying an upward force to counteract gravity (Table 3.4).

Force Type	Definition
Break off	Remove a part of an object
Extend	The hand applies outward forces from the inside of the object
Grab	Holding or securing an object without opposing gravity
Hold	Grasp object in a way that resists gravity
Lever	Lift or move (something) with a lever
Lift	Apply upward force greater than necessary to resist gravity
Place	Put something in a specified position
Press	Push object in a single direction (different presses exist: gentle, forceful, light, quick, etc.)
Pull	Pulling something in a single direction
Punch	Press or push (something) with a short, quick movement
Put in	Insert
Roll	Cause to move in a circular manner
Rub	Move something back and forth along the surface of (something) while pressing
Scratch	Rub a surface or object with something sharp or rough (with the hand directly or a tool)
Squeeze	Apply compressive force around object greater than needed to just hold object
Take out	Remove one object from another
Throw	Cause something to move out of your hand and through the air by quickly moving your arm forward
Turn	Flipping or rifling through pages
Twist	Use torsional force to rotate an object around a central point
Swing	Move with a smooth, curving motion like waving hand or swinging arm

Table 3.1: Force Type Definitions

Example		
Force Type	Squeeze	Hold
Annotation	Squeeze toothpaste	Hold a pan

Table 3.2: Internal Force Examples

Example		
Force Type	Throw	Grab&Press
Annotation	Shoot a basketball	Press down a door handle

Table 3.3: Cumulative Force Examples

3.2.3 Direction

In order to specify the direction of a task, we need to specify the direction subspace and the coordinate frame as shown in Table 3.5.

In order to specify the direction of a force or motion, we need to specify the direction subspace and the coordinate frame as shown in Table 3.5. The direction subspace describes a subset of the six-dimensional space within which the motion is occurring. Examples of direction subspaces that we use include: (1) along a linear axis, (2) rotation around an axis, (3) movement within a plane, or (4) inwards/outwards (towards or away from the center of an object). We note that the motion direction can be very different from the force direction. For example, when we zip a zipper, the internal force direction of the hand is *inwards* for the zipper (i.e. grab the zipper tightly), but the direction of motion is *along* the zipper. Similarly, the internal force direction is *inwards* to hold the egg beater but the direction of motion is around the x-axis (Table 3.6). We use the notation x(45)y to describe movements along an axis that is halfway between the x- and y-axes (e.g., Table 3.12, second row). Directions that are less constrained or more

Example		
Force Type	Grab	Hold
Annotation	Grab the ladder	Hold a laundry detergent

Table 3.4: Grab vs. Hold

Property	Possible Values	Example
Direction Subspace	along x/y/z axis	Table 3.6 I
	rotate around x/y/z axis	Table 3.6 II
	plane xy/xz/yz	Table 3.6 III
Coordinate Frame	hand	Table 3.7 I
	global	Table 3.7 II
	object	Table 3.7 III

Table 3.5: Direction Examples

difficult to describe are captured in freeform text (e.g., “a cone about the x-axis” or “various”).

Most of the time, we use the local coordinates of the hand to describe the direction of movement. However, we also sometimes use global coordinates of the world or local coordinates of the object, depending on which is most useful for each motion.

Hand coordinates: The local coordinates of the hand are defined as follows: The direction of the four fingers is defined as the x-axis. The y-axis is defined as coming out of the palm in the ventral/palmar direction. The z-axis is defined as the thumb pointing away from the little finger for both hands (Figures 3.4 and 3.5). This results in using either the left hand rule for left hand or right hand rule for right hand to compute the z-axis. This unorthodox use of coordinate frames results in symmetrical descriptions of movements and grasps using the two hands. Local coordinates of the hand are mostly used when the motion is along one of the hand coordinate axes. For example, Table 3.7, first column, shows rubbing the hands along the lo-

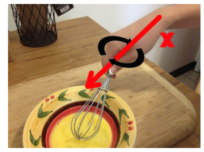
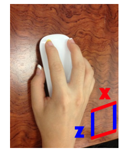
Example			
Axes	motion along x/-x(object)	motion around z axis	motion along xz plane
	force toward zipper	force toward egg beater	force against the mouse surface
Annotation	Zip a zipper	Beat eggs with egg beater	Move a mouse

Table 3.6: Axes Examples

cal x-axis.

Global coordinates: Global coordinates of the world are used when the motion is along the direction of gravity or within a coordinate system that could be fixed to our local environment. For example, when we dribble a basketball, we maneuver the ball within a coordinate frame fixed to the world, not the hand or the ball (Table 3.7, second column). The direction of gravity is defined as the global z-axis.

Object coordinates: Finally, occasionally the local coordinates of the object must be used since, in some motions, the object shape decides the direction of motion. If the object is a long stick or string type, we define the direction along the stick to be the x-axis. If the object is rectangular in shape, we define the direction along the long side to be the x-axis and the direction along the short side as the z-axis. For example, when we pull out measuring tape, the motion direction is along the tape's long dimension: the x-axis (Table 3.7, third column).

Many motions or forces can be described naturally in multiple coordinate frames. For example, plugging in a charger could be expressed in the coordinate frame of the charger, the wall, or the hand. We asked our subjects to make the annotations that were most intuitive for them. The important point is that all three coordinate frames are useful, as different actions may focus on different frames of reference.

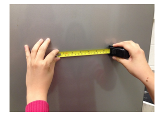
Example			
Coord Frame	Hand	Global	Object
Axes	motion along y/-y	motion along z/-z	motion along x/-x
Annotation	Rub hands	Dribble basketball	Measure with a tape measure

Table 3.7: Coordinate Frame Examples

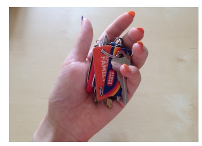
Example		
Flow	Bound	Free
Annotation	Stick key into key hole	Hold keys

Table 3.8: Flow Factor Examples

3.2.4 Flow

The effort factor we use here is flow. Flow comes from the Laban Effort / Shape notation [Samadani et al., 2013]. It refers to "attitude toward bodily tension and control" and can be *free*, *bound* and *half-bound*. Free refers to the moving direction of the gesture being very casual, while bound refers to the action being very stiff or tightly controlled. The half bound annotation is used when the action is bound along one or more axes and free along the rest. For example, in Table 3.13, the flow of motion in dragging toilet paper is half-bound because in the plane that is perpendicular to the axis of the toilet paper, the motion is still free. Our informal observation is that most of the time we specify an action as being free or bound depending on whether the action includes a goal location. For example, if we try to plug in a charger into a wall or stick a key into a lock, the motion is bound, but if we just throw the key for fun, the action is entirely free (Table 3.8).

Example		
Object weight	Light	Heavy
Annotation	Grab an empty box	Hold a heavy box

Table 3.9: Weight of Object Examples

3.2.5 Object related factors

Most grasps depend on the object our hands manipulate, thus object related factors are also important features for describing hand gestures. From our observations, weight is a big factor since it affects both internal and cumulative force applied on the object. A simple example is when we hold an empty box or a full box. If the box is empty, we tend to grab the top piece of the box, but if the box is heavy, we would starting from the bottom by lifting it up (Table 3.9).

The material of the object also strongly affects grasping strategy. For example, grabbing highly deformable material requires continuous adjustment of grasp shape as the object changes shape. Another example of the effect of material is that people will grab raw meat differently than paper. The shape and size of the object affects hand shape. We usually pinch a thin wire but grab a thick string, see Table 3.10. Finally, the friction coefficient of an object determines how hard we grab the object. The thick string in Table 3.10 is rougher then the exercise bar, which will affect the force needed to prevent slipping in both cases. We explore object-related factors in more detail in Chapter 7.

3.3 Results

Our main result is an annotated database of grasping actions observed in our study. The database contains 73 grasp types, including the 33 types enumerated in Feix et al., along with 40 additional types. Each of these 73 types includes one or more annotated examples. Examples are annotated with force type, motion direction, force direction, and flow to more fully de-

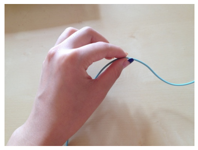
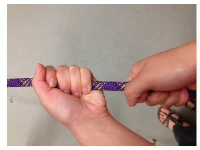
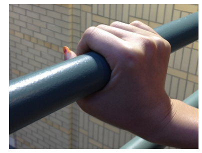
Example			
Size	Thin	Thick	Thick
Roughness	Slippery	Rough	Slippery
Annotation	Grab a wire	Grab a rope	Grab exercise bar

Table 3.10: Shape & Size & Roughness of Object Examples

Name	Large Diameter	Lateral
Picture		
Type	Power	Intermediate
Opp.Type	Palm	Side
Thumb Pos	Abd	Add
VF2	2-5	2
Shape	Cylinder/Cuboid	Card piece
Size	Large Diameter	Thin

Table 3.11: Large Diameter and Lateral Grasp

scribe the grasp in action. Each of the 179 total examples differs from the others by at least one annotation.

One additional result listed here is a count of force types, which can be found in Table 3.1 (frequency column). In this table, we can see, for example, that *hold* (41), *grab* (32), *press* (31) and *pull* (18) make up the majority of tasks that we observed in our study.

The full database can be found on our website [Liu et al., 2014a]. In this chapter, we describe two of the 73 grasp type entries (§3.3.1 and §3.3.2) as well as listing some of the new grasp types (Section 3.3.3).

3.3.1 Large diameter cylinder

The first grasp type we examine is the large diameter cylinder grasp. In a large-diameter grasp (Table 3.11, Left), the hand shape is appropriate for a larger-diameter cylinder-shaped object, and all five fingers are used. The opposition type is palm. The thumb is abducted.

Our entire database entry for this grasp is shown in Table 3.12, and we see that this single entry in the grasp taxonomy contains a variety of different examples. Force types are varied, including *hold*, *grab*, *squeeze*, *press*, and *twist*. Even with the same force type, other annotations can differ. For example, as shown in Table 3.12 (top), the action of *drink water* involves motion around the y-axis, while holding a bottle does not involve any motion. The flow can vary even within the same task. As shown in Table 3.12 (bottom), the motion of squeezing a towel is free, but the force is bound.

3.3.2 Lateral

The second grasp type we review is the lateral grasp. As shown in Table 3.11, Right, in the lateral grasp, the hand shape is more suitable for a thin card-shaped object, which is pinched between the thumb and index finger. The opposition type is side, and the pad of the thumb is used. The thumb is adducted.

For some very similar tasks, the direction and flow can be different. As shown in Table 3.13 first row, the flow of motion in putting on gloves and dragging toilet paper are different. Putting on gloves is bound since the direction of motion is set along the arm. But dragging toilet paper is half-bound.

The two tasks in Table 3.13 second row appear almost identical, but the direction of motion is different in terms of hand coordinates. Twisting the key happens around y-axis of the hand (the axis out of the palm), and twisting the knob happens around the x-axis of the hand (the direction aligning with the forearm).

Some motions are in the same direction but with different force types and flow as shown in Table 3.13 third row. In this case, the force based interactions are both in the xy-plane of the hand (or equivalently the object),

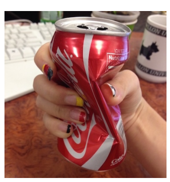
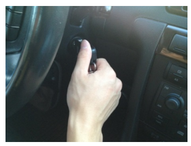
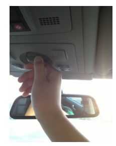
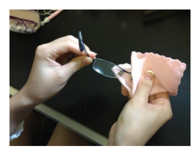
Example		
Force Type	Hold	Hold
Motion Dir	around y axis (hand)	-
Force Dir	-	-z (global)
Flow	Free Motion/ Bound Force	Bound Force
Annotation	Drink water	Hold a bottle
Example		
Force Type	Hold	Grab&Press
Motion Dir	x(45)y (hand)	-
Force Dir	-	z (global)
Flow	Free Motion/ Half Bound Force	Bound Force
Annotation	Throw paper	Grab cabbage
Example		
Force Type	Squeeze	Twist
Motion Dir	-	around z axis (hand)
Force Dir	inwards (hand)	inwards (hand)
Flow	Bound Force	Free Motion/ Bound Force
Annotation	Squeeze an empty soda can	Squeeze towel to dry

Table 3.12: Large Diameter Cylinder Grasp Examples

Example		
Force Type	Pull	Pull
Motion Dir	-x (hand)	xz plane (hand)
Force Dir	-	-
Flow	Bound Motion/ Bound Force	Half Bound Motion/ Bound Force
Annotation	Put on gloves (along the arm)	Drag toilet paper

Example		
Force Type	Twist	Twist
Motion Dir	around y axis (hand)	around x axis (hand)
Force Dir	-	-
Flow	Bound Motion	Bound Motion
Annotation	Twist the key to start up the car	Twist the knob in car

Example		
Force Type	Hold	Rub/Stroke
Motion Dir	xy plane (hand)	xy plane (hand)
Force Dir	-	inwards (hand)
Flow	Free Motion/ Half Bound Force	Half Bound Motion/ Bound Force
Annotation	Give card to someone	Wipe glasses

Example		
Force Type	Hold	Hold
Motion Dir	z (global)/ -z (global)/ around x axis (hand)	around x axis (hand)
Force Dir	-	-
Flow	Free Motion/ Bound Force	Half Bound Motion/ Bound Force
Annotation	Eat with spoon	Pour washing powder

Table 3.13: Lateral Grasp Examples

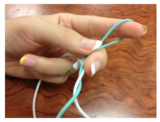
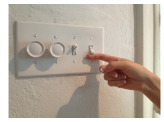
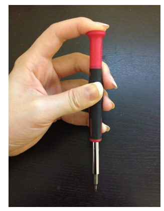
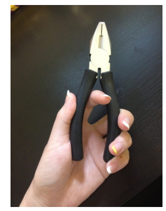
Example				
Annotation	Tie	Shuffle cards	Lift up the switch	Scratch
Example				
Annotation	Press perfume bottle	Open soda bottle	Use screwdriver	Use pliers

Table 3.14: New Type Examples

but one example has free motion while gently holding the grasped object and the other has motion relative to the object that is constrained to maintain forceful contact for cleaning. These differences are reflected in the differing annotations.

3.3.3 New Types

From our observations, the existing taxonomy that served as our starting point [Feix et al.] has covered many types of grasps. However, there exist some actions which are not represented by their taxonomy, for which we have created new categories in the database. Some of the new entries involve deformable objects. Some are very specific gestures such as opening a soda can and tying shoes. Overall, we have added 40 new categories. We illustrate 8 of them in Table 3.14. All classifications and annotations can be found in our database [Liu et al., 2014a]. Some, but not all of the new grasp types can be found in other taxonomies, such as those of Kapandji and Honoré [1970] and Edwards et al. [2002].

3.3.4 Discussion

Effective grasp taxonomies capture not only hand shape, but also the nature of contact between the hand and object. The best in this regard is perhaps the Kamakura taxonomy [Kamakura et al., 1980], which illustrates in great detail regions on the hand that come in contact with the object. The patterns and extent of these regions reveals much, especially when considering grasp control and robot hand design.

However, we find annotating only shape and contact to be insufficient to convey important differences between everyday actions; in part because this set of actions is more broad than grasping, but also because many grasps that may look similar from a snapshot involve very different intentions – different uses of the hand to accomplish a task. We find that to communicate these differences, we need to express the type of force, directional information, and stiffness information for the action.

It is interesting to note the similarities between our annotations and the parameters required for impedance control [Hogan, 1985] or operational space control [Khatib, 1987], where one expresses a task in terms of the desired impedance or motion/force/stiffness properties of the manipulator. Annotations such as those we propose here could form the starting point for a learning-from-demonstration or coaching system where the user indicates to the robot coordinate frames and directions best suited for position control and force control, along with indications of the level of force or stiffness required for the task. In particular, we found the use of English language verbs very promising for conveying the type of force desired in a way that was intuitive for our subjects, and the use of multiple coordinate frames (hand, object, and world) make it easier to specify axes along which motion and force should be emphasized or constrained. It is of great interest to us to explore mechanisms for translating such annotations into robot controllers and allowing users to provide feedback to adjust those controllers in a language that is natural to them.

The similarities between our classification scheme and Laban Movement Analysis (LMA) [Newlove and Dalby, 2003] are also intriguing and invite further exploration. Perhaps we may consider the static grasps of the conventional taxonomies as Shape Forms – static shapes that the hand may take while grasping an object. Annotation mechanisms within the category of Space may capture our intent when annotating motion and force directions, where we consider natural coordinate frames and landmarks that serve to orient the action. Annotation mechanisms within the category of Effort were motivating to us when considering how to discriminate between grasps. Although we did not make direct use of the Action Effort verbs (Float, Punch, Glide, Slash, Dab, Wring, Flick, and Press), many of them are represented in our force list of Table 3.1. In addition, we attempted to directly adopt the Effort category of Flow to allow users to discriminate between stiff and tightly controlled vs. free or flowing intent. We are interested to explore further how theory and practical experience from LMA may allow us to create more precise and comprehensive annotations.

Although there are similarities between our annotation scheme and LMA categories, there are also differences. For example, although our verb list is similar to the Action Effort verbs, there are verbs in our list that may fit one or more Action Effort verbs depending on how the action is performed. For example, in our database subjects used “Press” for forcefully supporting a cabbage for cutting and also for lightly pressing a small button, which may correspond to different Action Effort verbs such as “Press” and “Dab.” In addition, there are items in our verb list that do not correspond well to the Action Effort verbs, such as “Put In” and “Take Out.” The largest conceptual difference seems to be that our subjects considered verbs in our list to express *what* the hand was doing, as opposed to *how* the action was performed. Given this conceptual difference, it is interesting to see the level of similarity we do see in the two sets of verbs.

We also found that we needed to give our lightly trained users a great variety of verbs as options to specify force intent. We have listed 20 such verbs in Table 3.1 and have no doubt that a more extensive survey of everyday actions will require adding others. Intent of an action as it affects function and appearance of grasping appears to be challenging to capture and communicate in a manner that can discriminate between actions that are evidently different to both the performer and the observer.

One limitation of this database is that we need a more accurate system for describing the direction of motion and force that accommodates directions that do not perfectly align with an easily identifiable single axis. However, interestingly, this situation appears to be uncommon.

We can also ask whether all entries in our database are relevant for humanoid robots. We believe that as robots become more pervasive, especially in home, health care, and rehabilitation scenarios, a large majority of the grasps depicted here will become of interest. However, we did not attempt to make this distinction.

It may be possible to organize this database from a different point of view, such as making the force types or motion types the central classification rather than grasp type. We chose grasp type as the first level of organization in order to be consistent with existing taxonomies. However, it is interesting to consider whether a different organization may lead to a simpler or more intuitive way of describing these results.

4

Complexities of grasping in the wild

Whereas the work in the previous chapter looked at still snapshots of grasp poses, in this work, we examine the whole time sequence of grasping. The growing ubiquity of high-framerate video cameras in phones gives us the opportunity to observe human grasping at a fairly high level of temporal detail. In addition, the availability of these handheld cameras means that it is now feasible to capture a large number of grasping actions "in the wild" i.e. in everyday settings such as cluttered workspaces. The large number of actions and the everyday setting allows behaviors such as mistakes to be captured, and the high framerate reveals detailed finger movement and the making and breaking of contact.

We observed one human subject taking items from store shelves, counter, and bins, and replacing them. The subject was recorded using a single hand-held camera at 120 frames per second. We then analyzed the video using several classification systems, as well as ad hoc analyses that attempt to note high-level events in the recording not captured in the other taxonomies.

Unfortunately, RGB video is not amenable to automated analysis. Humans have to watch the video and record their observations. It is our hope that eventually this process can be partially automated using video analytics and behavior recognition, and that our annotations can function as ground truth data for future automated analytical tools.

The primary long range goal of this work is to develop an annotation system capable of describing manipulation behavior performed by one actor in such a way that the manipulation can be copied by another actor. Such an annotation system would need to be detailed and expressive enough to note all elements critical for duplicating the motion, but also flexible/abstract enough to be applicable across different robot hardware and hand morphologies.

The contributions of this chapter are as follows: (1) a summary and analysis of a slow-motion video dataset featuring interaction with a wide variety of objects, and (2) application of those findings toward the development of an annotation system able to capture important elements of grasping.

The greatest surprise in our analysis of this dataset was the variety and complexity of behaviors we saw, even though the task domain is mostly picking and placing. Other lessons learned from this study include:

- The process of grasping in the presence of clutter can be complex, sometimes involving adjustment of a grasp or exploiting the environment, yet occurs quickly.
- Contact-guided placing is common.
- Collisions between effector and clutter or between object and clutter are commonplace. Error recovery is quick when it is necessary at all.
- Expected patterns of behavior based on grasp taxonomies and other prior work were observed but less frequently than we expected.

4.1 Dataset

The dataset analyzed consists of a collection of RGB videos of a single subject manipulating objects in a convenience store. The videos were captured by one of the researchers using the iSight camera on an iPhone 5S (120 frames per second, 1280x720 resolution).

Continuous video capture of the entire visit was infeasible due to limitations in disk space and battery; thus videos were captured discontinuously and subsequently trimmed and pieced together to form a single video. In total, 91 interactions between the subject and 60 convenience store objects were observed and analyzed. These interactions collectively took place over a period of 3 minutes and 9 seconds of discontinuous video.

The subject was given instruction on which items to manipulate as she moved about the store. On occasion, the subject was encouraged to increase the variety of manipulation actions when possible, such as to twirl a turnstile or regrasp an apple. When finished, the subject attempted to replace the items back in their original locations. The subject has identified herself as being right-handed.

Objects manipulated by the subject include beverage bottles, cans, cups and Tetra Paks; salad dressing, tea, salt and cream packets; dry condiment shakers; a refrigerator door; various packaged foods, such as ice cream, potato chips and candy bars; plastic knives, forks, and spoons; napkins; a plastic sign; a plastic bag; a turnstile; an apple; a pizza box; a wrapped hoagie; plastic salad boxes; a plastic sauce cup with lid; and steel tongs.

A compressed version of the dataset and annotations are available online at http://graphics.cs.cmu.edu/projects/graspinginthewild/nsh_shop_120.webm and http://graphics.cs.cmu.edu/projects/graspinginthewild/annotations.zip.

4.2 Methodology

The captured video was viewed and analyzed with the aim of noting any significant events or processes that would be helpful for instructing a robotic actor to be able to replicate the manipulation. The researchers manually labeled the dataset using several existing taxonomies, as well as through other lenses where a taxonomy does not exist:

- Static grasp pose taxonomy created by Feix et al. [2009]. This taxonomy collects poses from previous taxonomies and separates hand shapes based on function (power, precision, or intermediate), thumb position (abducted or adducted) and which surfaces of the hand are used to secure the object (palm, finger pads, or sides of fingers).
- Intrinsic (within-hand) hand motion categories observed by Elliott and Connolly [1984], which describe motions a hand uses to manipulate an object already in the hand.
- Bullock et al.'s manipulation taxonomy [2013], which creates broad categories of manipulation based on the presence or absence of contact (C), prehension (P), motion (M), intrinsic hand motion (W), and motion at contact points (A). The taxonomy is high-level and doesn't assume any particular hand morphology.

- The lens of errors and recovery from errors.
- The lens of contacts and when they are important to execution of a motion, either aiding or constraining manipulation.

Annotating the video through these lenses often involved noting intention as well – why that choice of grasp; what is the purpose and end effect of a particular intrinsic hand motion; what was the hand attempting to do when the error occurred?

In a first pass, the entire video was annotated through each of the above lenses. We then focused on a small number of actions that contained examples of interesting recurring phenomena (e.g. levering up, regrasps, errors/error recovery), and ranged from the simplest actions (milk bottle place #1) to the most complicated actions (cutlery pick #2) observed. We cleaned up the annotations for these motions to make them consistent and to use more fine-resolution frame numbers instead of seconds, and then plotted their annotations in the form of a timeline. The eight motions selected were:

- Zone bar pick (0:11-0:22 – Fig. 4.3)
- Zone bar place (0:22-0:29 – Fig. 4.3)
- Mountain Dew pick (1:01-1:08 – Fig. 4.4)
- Milk bottle place #1 (3:41-3:46 – Fig. 4.4)
- Pepsi cup pick (4:09-4:15 – Fig. 4.4)
- Cutlery pick #2 (5:56-6:34 – Fig. 4.5)
- Lay's chips pick (8:25-8:36 – Fig. 4.6)
- Pizza box pick (9:35-9:46 – Fig. 4.6)

In the timelines, we used color to distinguish between annotations that fell within a taxonomy (grey blocks) and new ones not found in that taxonomy (green blocks). For annotations using the Bullock, Ma, and Dollar (BMD) taxonomy, the "new" annotations correspond to moments when multiple actions are being performed by different parts of the hand – for example, two fingers holding an object in a stable grasp while the rest form a grasp of a second object. While these moments could be annotated as a single BMD category (usually C [P,NP] M W A), we decided to annotate the actions of the different units of the hand separately to be more descriptive of what is happening. The downside is that this way of annotation is more complicated.

Due to the general and comprehensive nature of the BMD taxonomy, an annotation was possible at every point in time during grasping except when the hand is off-screen or occluded. Gaps in the BMD timeline correspond to these situations.

When analyses had no annotations associated with them, their empty timelines were excluded from the figure. For example, there were no miscellaneous annotations in the Mountain Dew pick action and no intrinsic manipulation annotations in the milk bottle place action (see Fig. 4.4).

4.3 Results

Annotations for the selected clips are shown in Figs. 4.3-4.6. The accompanying video shows these motions. This section outlines insights obtained from these and other annotations.

4.3.1 The process of forming a grasp is complex.

The high framerate video reveals detailed grasping strategies that are hard to see in normal 30 fps video. The examples shown in the video indicate that the process of forming a grasp is as complex and worthy of notice as the final achieved grasp pose itself. While it is simple to pinch small items between two or more fingers and instantly form a grasp that way, many of the grasps observed featured some kind of hand pose adjustment between the time of making contact and forming the final grasp. Fig. 4.1a is an example of how and why adjustments occur between contact and final grasp: first, ulnar fingers use the rim of the box to lift one side, exposing the bottom surface (frame 1). Then a complicated sequential pattern of finger lifting and recontacting (frames 2-5) results in the final grasp (last frame). This final grasp involving the bottom surface of the box is much more secure, but not possible until the bottom surface has been lifted up enough for fingers to be placed underneath.

In general, we find that the process of forming a grasp has multiple phases:

1. Approach and preshaping: changing the pose of the arm or hand in anticipation of grasping
2. Contact: compliantly making contact with some part of the object
3. Dealing with clutter: maneuvering fingers into spaces, singulating an

(a)

(b)

(c)

(d)

Figure 4.1: Examples of (a) regrasping into a stronger grasp, (b) simultaneous levering out and grasp formation, (c) contact-guided placing, and (d) error correction (pinky is withdrawn from bin while approaching).

object, or pushing its surfaces away from nearby surfaces.

4. Taking weight: bracing or adjusting pose to take full weight of object
5. Lift: able to move object with full arm now that stable grasp has been formed
6. Grasp adjustment: to more comfortable grasp

For small, light, or unobstructed objects, some of these phases may not be necessary. Sometimes singulating the object and pulling it further into the hand to form a grasp happen simultaneously (see Fig. 4.1b).

Ungrasping involves similar phases but in reverse (for example, touching down and letting go of weight instead of taking weight and lifting). Similar to grasping, many ungrasping motions are not just opening the hand to break contact; they instead involve some kind of in-hand motion or grasp change before contact is broken. Approximately 25 of the 53 grasping examples (47%) feature post-contact grasp adjustments before a final grasp, and 16 of the 48 placing examples (33%) feature pre-release grasp adjustment (see Fig. 4.1c).*

*A single action in the video could contain multiple grasping and placing examples, so the total number of grasps and places is greater than the number of actions captured.

4.3.2 Environment-aided grasping

Before prehension is achieved, the human hand is nevertheless able to manipulate an object (for example, lifting a corner or edge up, tilting an object out, singulating an object by pressing down, etc.). The way it does this is by using the environment as a “finger” of sorts, which provides an opposing surface that a hand can use to “grasp” an object securely enough to manipulate it. Being able to exploit these environmental contacts appears to be important for grasping objects when a normal pinch grasp is not feasible.

We also found that gaps in the environment are also exploited in order to aid grasping. Fingers can be inserted into gaps and extended in order to create more space, as in the case of the soymilk pick (1:29). The pizza box pick (Fig. 4.6) is an example of both exploiting a gap to contact the side of the box and then using that contact to form an environment-aided grasp.

4.3.3 Insights from the grasp taxonomy analysis

Fig. 4.2 shows new and in-between grasps found in the video. (a) The placement of the index finger is flexible and can be abducted away from other fingers, resulting in variations on existing grasps. (b) There exists a family of lateral grasps involving the side of fingers other than the index finger, possibly in conjunction with the index finger to strengthen the grasp. (c) Storage grasps involving the ulnar fingers or the crease between the thumb and index finger are specialized grasps that allow manipulation or a second grasp to be performed by unused fingers. (d) Deformable objects like potato chip bags resulted in unusual grasps that use a mix of side and pad opposition. (e) Some in-between grasps were found like an apple grasp in between the precision sphere and precision disk grasps, and a milk bottle grasp similar to a tripod grasp but stronger and more stable.

We also observed objects initially grasped with a weak/precision grasp being regrasped into a power grasp. Figs. 4.1a and 4.5 are examples of this.

Although we focused on stable grasp poses e.g. times when there is no motion occurring within the hand, the cutlery pick action (Fig. 4.5) was an exception. During this action, small motions within the hand (such as

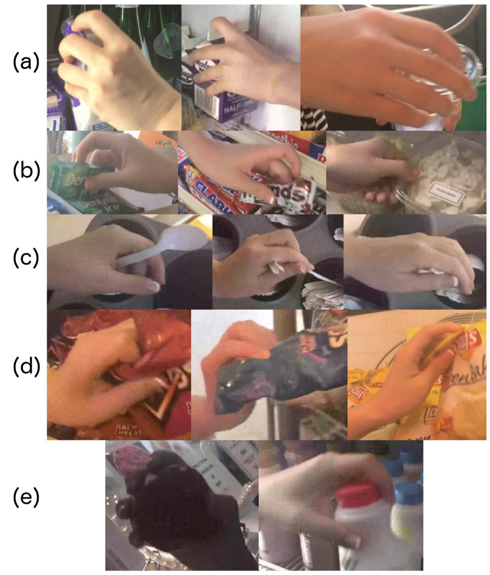


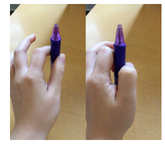
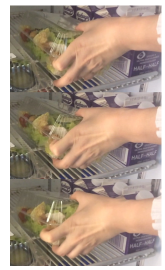
Figure 4.2: New and in-between grasps observed in the video. (a) Variations with index finger extended or placed in a different area than other fingers. (b) Variations on lateral pinch grasp using middle and ring fingers. (c) Storage grasps allow manipulation and multiple grasps. (d) Deformable objects create a variety of opposition types (side and pad). (e) In-between grasps.

lifting the middle finger) can instantly change one grasp into another (e.g. from prismatic 2 finger to inferior pincer).

4.3.4 Insights from intrinsic manipulation analysis

We observed intrinsic movements like squeezing a bottle into the palm; interdigital steps to reorient stick-like objects in the hand; and rocking objects back or forth to help remove them from clutter. In particular, we noticed from the cutlery pick action (Fig. 4.5) that the interdigital step is a broad, high-level category that contains various smaller motions that people use to reorient objects in the hand.

The Elliott and Connolly taxonomy is designed for motions to manipulate objects already grasped. However, we noticed intrinsic hand motion happening before prehension. We collected all intrinsic motions found in the dataset, both non-prehensile and prehensile motions and summarized them in Tables 4.1-4.3. We connect these motions with those found in previous literature ([Elliott and Connolly, 1984] for prehensile manipulation, [Heine-

Motion	Equivalents	Examples	Images
Levering to lift edge	Full roll (EC); In-grasp manipulation (MD); **Flip** (H+)	Zone bar pick 0:14; Dentyne Ice pick 0:30; Tea bag pick #1 4:38; Ice cream pick #1 11:49	
Squeeze to bring object into hand *	Squeeze (EC); In-grasp manipulation (MD)	Mtn Dew pick 1:03; Napkin pick 6:54	
Sequential regrasp to bring object further into hand *	New sequential pattern (EC); Finger gaiting (MD)	Salad box pick 2:28; Plastic bag pick 7:14; Hoagie pick 10:13; Ice cream pick #2 12:17	

* New motions
EC = [Elliott and Connolly, 1984]; H+ = [Heinemann et al., 2015]; MD = [Ma and Dollar, 2011]

Table 4.1: Motions that are able to be used prehensilely and non-prehensilely.

mann et al., 2015] for non-prehensile manipulation, and [Ma and Dollar, 2011] for both) where relevant.

4.3.5 Insights from Bullock, Ma, and Dollar (BMD) analysis

Throughout the picking/placing process, the hand is very rarely still, with either the whole arm, individual fingers, or both moving for the entire time in most examples. This analysis reveals that the human hand is very efficient when grasping, parallelizing work. For example, Fig. 4.1d shows approach to an object (whole-arm motion) occurring at the same time as error correction (within-hand motion to pull the pinky finger out of the way).

One limitation of the BMD taxonomy is that there is no way to annotate the common scenario when motion is occurring both outside the hand and within the hand simultaneously (i.e. a motion-within-hand (W) plus a motion-not-within-hand (NW) annotation), or when some contacts are changing while others are static (motion-at-contact (A) plus motion-not-at-contact (NA) annotation). In other words, within-hand and at-contact mo-

Motion	Equivalents	Examples	Images
Use ulnar fingers to push environment away *†	None	Soymilk pick 1:32; Cracker Jack place 9:15 & 9:17	
Flat hand squeeze to manipulate object and bring into hand *†	Squeeze (EC)	Milk bottle pick #1 3:37; Tongs pick 11:28	

* New motions † Related to dealing with clutter

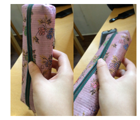
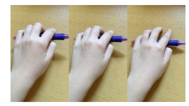
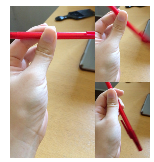
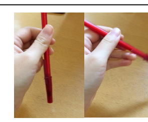
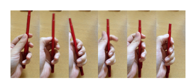
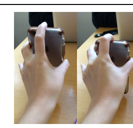
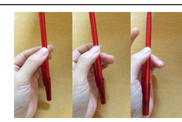
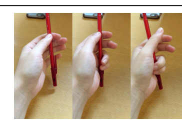
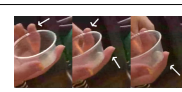
Table 4.2: Motions that are only used non-prehensilely.

tion "mask" external motion or still contacts. The ability to indicate both are occurring simultaneously complicates the process of annotating motion, but may be important for the goal of being able to instruct robots to copy human grasping actions.

As the Pepsi cup pick (Fig. 4.4) indicates, full arm motion with a stable grasp pose can denote very different kinds of forces and motions. It can denote a smooth motion (pulling an object out of a space), or the shaking used during part of this motion. It is not able to distinguish between these two types of motion, which makes sense as the BMD taxonomy was designed to be augmented with other manipulation taxonomies. In particular, the taxonomy from the previous chapter (Ch. 3) may be a good choice to use here.

4.3.6 Errors and error recovery

The subject was instructed not to take any particular care when grasping. As a result, errors are observed from time to time, appearing in 13 of the 91 captured actions. Errors were corrected very quickly and the intended motion eventually succeeded with only one exception (tea packet push (4:53-5:04)). Fig. 4.1d shows an example of a quickly-corrected error, where a finger slips into a bin and is lifted without interrupting the grasping motion. Other errors we noticed included an edge of the object hitting other objects, pinches missing/failing to secure an object, and actions failing to insert an object into the intended location.

Motion	Equivalents	Examples	Images
Rock to rotate object	**Rock** (EC); In-grasp manipulation (MD)	Zone bar pick 0:18	
Inverse squeeze to drop object gently	**Squeeze** (EC); In-grasp manipulation (MD)	Mtn Dew place 1:09; Salad box place 2:36; Cutlery place #2 (knife) 6:34 & 6:37; Doritos place #1 8:07; Doritos place #2 8:21	
Drop / flex to rotate a stick (a part of interdigital step)	**Interdigital step** (EC); Finger-pivoting/tracking / sliding (MD)	Cutlery pick #1 (spoon) (gravity) 5:22; Cutlery pick #1 (knife) (flexion) 5:33; Cutlery pick #2 (spoon) (flexion) 6:05; Cutlery pick #2 (knife) (both) 6:24	
Extend finger to push stick outward (a part of interdigital step)	**Interdigital step** (EC)	Cutlery place #1 (fork) 5:43; Cutlery pick #2 (spoon) 6:01	
Use thumb to move object down into hand (a part of linear step)	**Linear step** (EC); Sliding (MD)	Cutlery pick #2 (fork) 6:18	
Linear step to inch up a stick	**Linear step** (EC); Finger gaiting (MD)	Cutlery pick #2 (knife) 6:27	
Squeeze variation using first three fingers in a tripod *†	Squeeze (EC); Sliding (MD)	Soymilk place 1:39	
Swap index finger for middle finger *‡	Regrasping (MD)	Cutlery pick #1 (spoon) 5:20; Cutlery pick #2 (fork) 6:14; Cutlery pick #2 (knife) 6:25	
Regrasp into ulnar grasp *‡	Regrasping (MD)	Cutlery pick #1 (spoon) 5:24; Cutlery pick #2 (spoon) 5:59; Cutlery pick #2 (spoon) 6:06	
Adjust contact points to make room for other grasper *‡	Regrasping (MD)	Sauce cup pick 10:44	

* New motions † Related to dealing with clutter ‡ Regrasp

Table 4.3: Motions that are only used prehensilely.

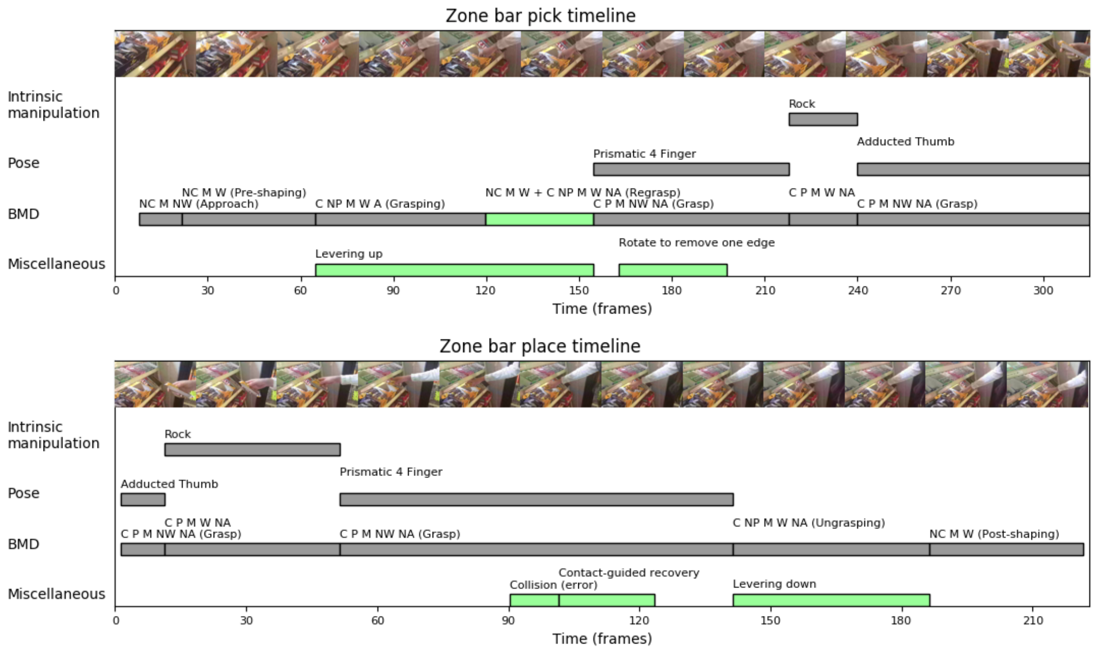


Figure 4.3: Timelines for Zone bar pick and place.

4.3.7 Insights from contact analysis

One analysis of the video focused on contacts and noted whenever contact was important to the motion. These motions fell into two categories: (1) contact was established purposefully in order to aid the motion (contact guidance), and (2) haptic feedback rather than visual feedback was possibly driving the action.

In the first case, contact was helpful for completing a motion. In 13 of 48 placing actions, an initial contact between a corner of the object was first established, and then the constraints created by that contact were used to guide the object into place. Fig. 4.1c is an example of such a movement where a contact is established.

The second case contained most examples of error correction as well as motions that were incidentally contact-heavy. For example, the Pepsi cup grab (Fig. 4.4) involved contacts that needed to be broken; this task was accomplished by shaking the cup.

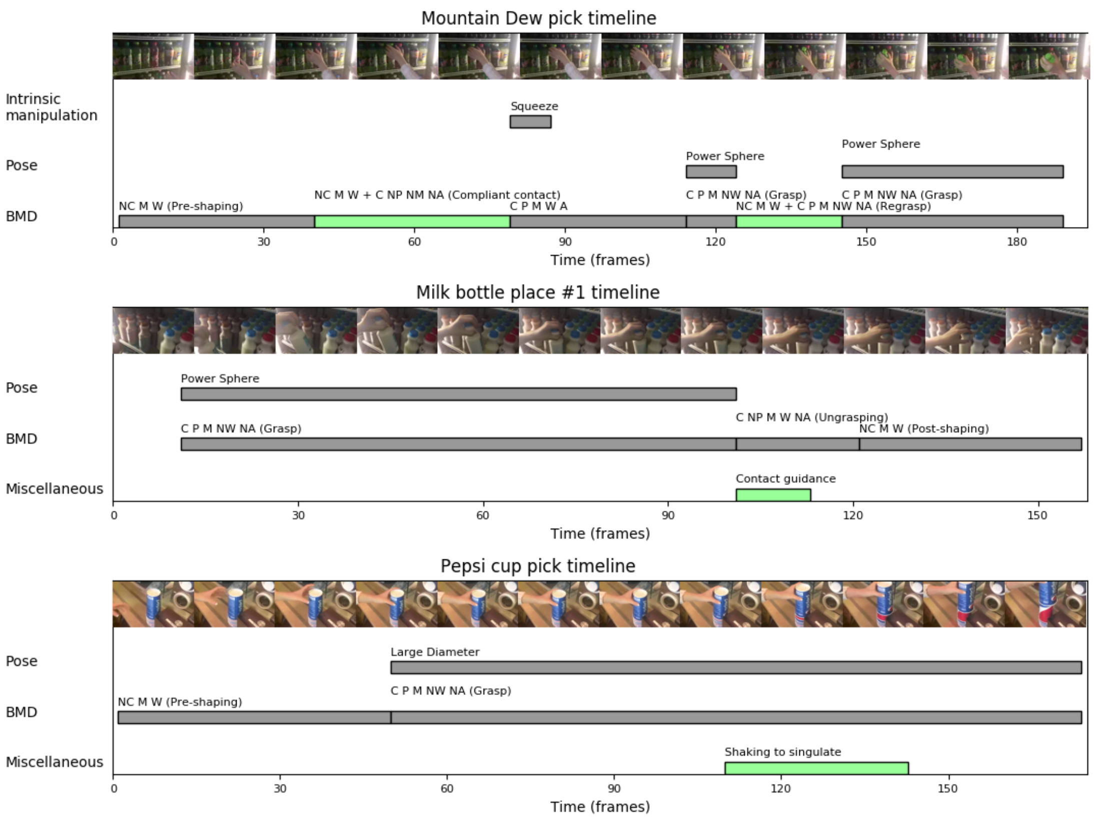


Figure 4.4: Timelines for Mountain Dew pick, milk place #1, and Pepsi cup pick.

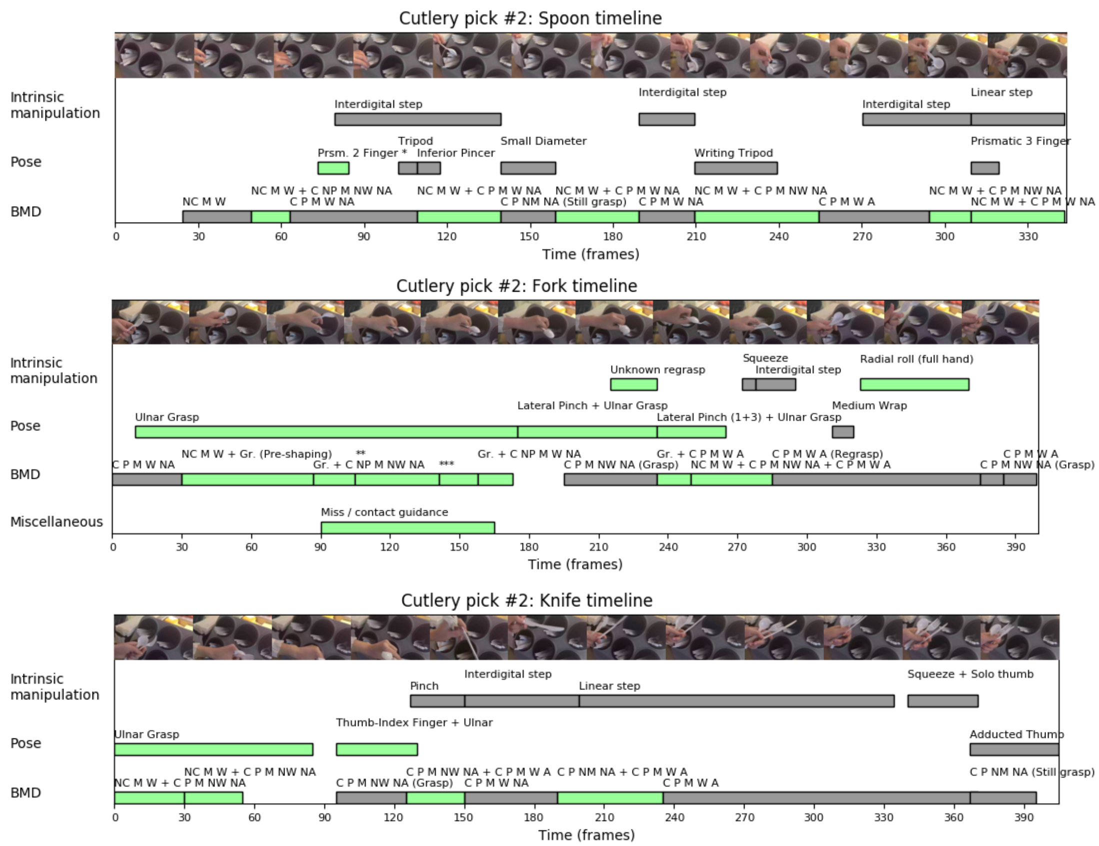


Figure 4.5: Timelines for cutlery pick #2. *This grasp is capable of sliding the spoon out, but part of the spoon is supported by the environment, so this grasp is an environment-aided prehensile grasp. **NC M W + C P NM NA + C NP M W A (different fingers holding, preshaping, and manipulating). ***NC M W + C P M NW NA + C NP M NW A (same as previous but with full-arm motion)

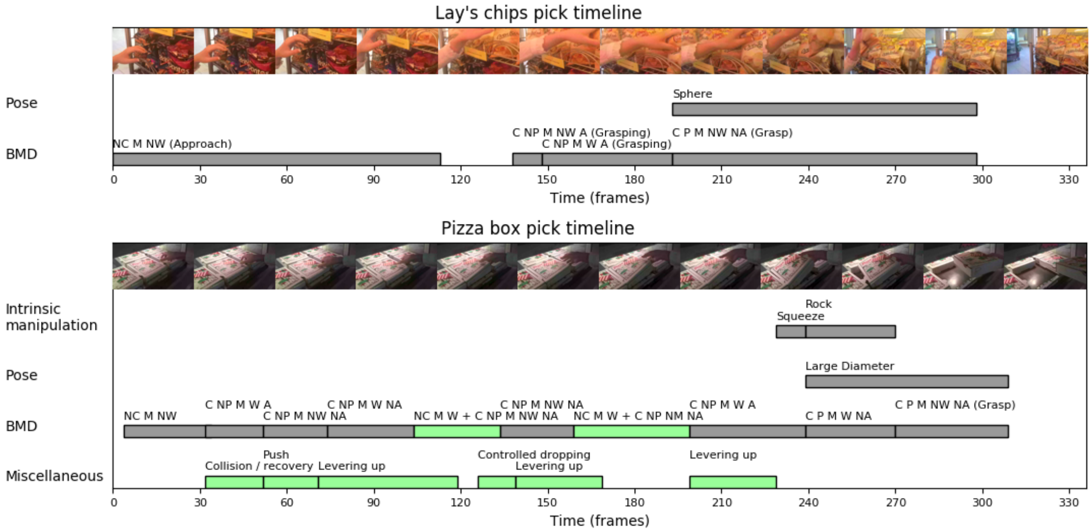


Figure 4.6: Timelines for Lay's chips pick and pizza box pick

4.4 Discussion

Our video analysis reveals that the grasping process is surprisingly complex but fast. It takes advantage of environmental contacts and touch feedback. An initial non-prehensile "grasp" is often used to manipulate the object to make a final power grasp possible.

To create an annotation system sufficient to describe and prescribe manipulation, it is helpful to use taxonomies for grasp poses, intrinsic manipulation, and generic manipulation. Grasp poses often reflect the end goal of the action e.g. a stable grasp of an object that is suitable for transporting and placing. However, pose taxonomies need to be extended to describe the flexible aspects of the grasp (for example, to instruct a robot that the index finger can be separated from other fingers and be used to tip the object out) and to include storage grasps. The intrinsic manipulation taxonomy is useful for describing manipulation of already-grasped objects and could be extended to include non-prehensile manipulation. The generic manipulation taxonomy lives up to its goals of being general enough to describe all manipulation without being tied to any one hand morphology, and is useful for segmenting motions into phases. However, its main limitation is the difficulty in describing multitasking/concurrent manipulation, which occurs regularly in human grasping.

In addition to those elements, a prescriptive annotation system needs to also describe the role of the environment in grasping and be able to convey intent. Force types and object motion may help fill this intention gap. Different force types can achieve different objectives, like smooth movement for transport vs. shaking to break contacts/suction force. Object motion hints at purpose as well. For example, some non-prehensile manipulation is done for the purpose of lifting a corner of the object up (changing the object's configuration). How this is achieved (e.g. by motion within the hand) is as important as knowing the purpose (e.g. to change the object configuration to expose the object's bottom side).

Our work has implications for robotic grasping. For example, compliant contact and contact guidance found in human grasping suggests that ongoing work on compliant control is important. In addition, dexterous in-hand manipulation seems important even for simple pick-and-place tasks. If human hands are any indication, work on the more difficult task of dexterous manipulation, including non-prehensile manipulation, will aid work on the forming of stable cage grasps. Finally, the intrinsic motions in our dataset (Tables 4.1-4.3) can serve as a benchmark for the manipulation capabilities of anthropomorphic robot hands, alongside other taxonomies.

The dataset has several limitations. First, the motions are primarily picking and placing motions that are performed by a single subject who is aware of being recorded. Because only one subject was recorded, some aspects of grasping may be idiosyncratic to her. Second, the objects are usually grasped without any intention of being used or placed in a different location.

4.5 Conclusion

In this work we captured a dataset of slow-motion actions in a convenience store setting. We analyzed this video through the lenses of different manipulation taxonomies – a grasp pose taxonomy, an intrinsic manipulation action taxonomy, and a generic manipulation taxonomy – as well as through lenses focused on errors and contacts. We found that the process of grasping is complex and deserves more focus, particularly in situations with clutter or environmental constraints. Grasping is not only complex but also *quick* – with multiple goals being worked toward at the same time, such as one motion both singulating an object and drawing it into the hand

– and *heavily reliant on touch* for corrections. The process of annotating elements of manipulation is time-consuming and at times reliant on high-level understanding of the video such as being able to infer the intention of a motion or series of motions. As such, there are many challenges to using human motion examples to inform robotic grasping. However, awareness of the complexity and strategy involved in grasping may help us design more robust and effective grasping processes. In particular, we seek to apply the insights related to pre-grasping manipulation from this work to artificial graspers (Chapter 8).

5

Recommendations for a manipulation annotation system

The previous chapters investigated how people grasp in various settings, including natural, cluttered environments. In this chapter, we use the insights from those studies in order to make a recommendation for an annotation system capable of noting the important aspects of grasping for transferring manipulation skills to robots.

5.1 Anthropomorphic vs. non-anthropomorphic annotation systems

When considering the problem of transferring manipulation skills from humans to robots, one major issue is the translation from the morphology of the human hand to any arbitrary robot hand morphology. While some robot hands purposely mimic the shape and function of the human hand, many differ from the hand in terms of number of fingers, number of degrees of freedom, the shape and rigidness of the palm, the mechanism by which the thumb abducts to oppose the fingers, and so on.

There are several possible ways to deal with differences in hand morphologies. One way is to take an object-centric view of manipulation. Object-centric descriptions can be especially useful if the object is changed in some way: translated, rotated, broken apart, etc. Wörgotter et al. [2013] and Leidner et al. [2015] produced two taxonomies that took an object-centric view of manipulation. In their work, Wörgotter et al. [2013] looked at different ways objects can be interacted with: rearranged (hit, turned, stirred, pushed/pulled, rubbed, levered), destroyed (cut, scratched, squeezed), taken apart, or combined. Leidner et al. [2015] created a taxonomy useful for describing tool usage. Tasks are classified based on how a hand alone or a combined hand-tool system manipulates an object or the environment. The importance of object-centric features such as contact, friction, deformation (of either the environment or the tool), and penetration is noted, as well as the presence of intrinsic hand motion (a hand-centric feature).

Object-centric views can be helpful for transferring manipulation skills from humans to robot hands because they ignore details of the hand. For example, Gupta et al. [2016] use object trajectories in order to transfer manipulation skills from humans to robots. The robot learns to produce the same change in the object without necessarily mimicking the human-specific strategy that was used to accomplish that manipulation. However, the drawback of object-centric views of manipulation is that the amount of helpful information they give about how a manipulation might be accomplished is limited, especially if the object is stationary – for example, in a stable grasp.

Another way of dealing with differences in hand morphologies is to take a hand-centric view of manipulation but one that abstracts out details that are specific to the human hand. For example, you can contrast how Elliott and Connolly [1984] describe and classify intrinsic manipulation with how Ma and Dollar [2011] do. For the former, motions are collected from observations of people and descriptions are provided of how the human hand accomplishes these motions. Note that this work also features some morphology-independent abstraction as well in its classification of types of synergies involved between the fingers in order to accomplish the motion. Our list of intrinsic manipulations in Chapter 4 (Tables 4.1-4.3) is also anthropomorphic. By contrast, Ma and Dollar [2011] outline general strategies of manipulating an object – ungrasping and regrasping it, sequential contact changes, pivoting around two contact points, and so on – that could potentially be used by a hand of any morphology. This taxonomy potentially gives more guidance as to how manipulations might be accomplished without making assumptions about the function of the hand performing them.

Figure 5.1: Adducted thumb featuring both palm and side opposition.

5.1.1 Hands at different scales

Another non-anthropomorphic way of describing manipulation is the taxonomy by Bullock et al. [2013] that was used in Chapter 4. This generic taxonomy of manipulation is useful for segmenting manipulation into phases and can be applied to both human and non-anthropomorphic hands. However, as we found in Chapter 4, when a hand is multitasking and using different parts of itself to perform different types of manipulation, it is difficult to pick an appropriate way of describing this manipulation. It is as if the human hand contains several "sub-hands." This insight is not new. Bullock et al. discussed the difficulty in defining a hand when they presented their taxonomy. In earlier work, Iberall [1987] described different types of grasps called opposition types, capable of securing and/or manipulating an object. These opposition types can be assigned to different parts of the hand, and multiple oppositions can be in play in a single grasp pose. For example, the adducted thumb (Fig. 5.1) features a side opposition between thumb and index finger as well as a palm opposition between the palm and the last three fingers. The ulnar grasp we found in Chapter 4 is a palm opposition grasp using the fourth and fifth fingers, leaving the first three fingers free to manipulate and form a second grasp independently. At different scales, different grasps can seem analogous to each other. For example, a tripod grasp forms multiple contacts around the edge of a small object in the same way a precision disk forms contacts around the edge of a large round object (Fig. 5.2).

Bullock et al.'s warning about the difficulty of defining a hand extends in the other direction to when the whole hand functions as one part of a larger "super-hand" system, such as when two hands – or even a hand and

Figure 5.2: Similar grasps at different scales.

other parts of the body such as the abdomen, leg, head, etc. – work together to function like a large hand. The bimanual grasps featured in Chapter 7 are examples of how two hands collaborate to form a larger hand system. Some bimanual grasps involve forming a one-handed grasp at two locations of the object, which can be seen as two independent unimanual grasps or a large hand-arm “super-hand” with “sticky” end effectors making contact at two places on the object. Other bimanual grasps like the cupped bimanual grasp involve each hand making a non-prehensile grasp of the object, such that the two hands together are capable of securely grasping the object.

An annotation system needs to be able to recognize when subparts of the hand are functioning separately and independently from each other, or when the hand is functioning in conjunction with other parts of the body to manipulate an object. Such a system would need to be able to describe manipulation happening at different scales and switch between the different scales.

5.1.2 The environment as a hand

Another useful thing that falls out of abstracting the morphology of the hand is that the environment also begins to look like a manipulator. For example, a table supporting an object looks analogous to a flat hand grasp (Fig. 5.3). We saw in Chapter 4 that environment-aided manipulation is common. Contacts are established purposely during placing in order to help constrain and guide the motion of the object. Obstacles that cage the object are exploited in order to manipulate the object before grasping with the hand. Dafle et al. [2014] show how the environment, including grav-

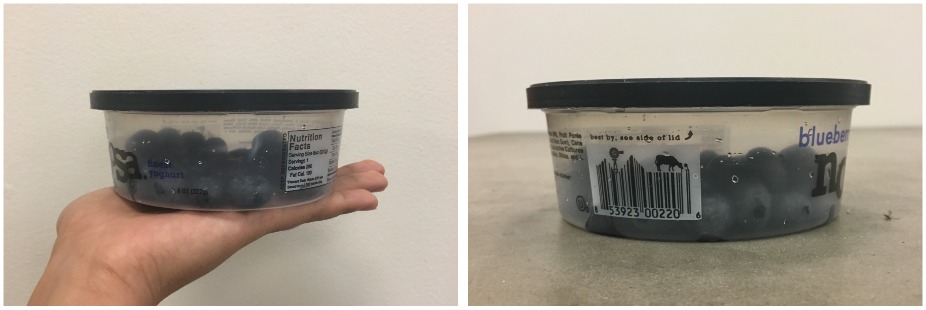


Figure 5.3: Flat hand grasps by hand and table supporting an object.

ity forces, can be used to aid manipulation. For example, surfaces can be used as contact points that allow an object to be reoriented in the hand without intrinsic hand motion by changing the position/orientation of the hand with respect to the environment.

In order to capture what people are doing when they manipulate and explain what purpose that action has, it is often important to understand how the environment is being used. A good annotation system should be able to understand how contacts between the environment and an object can be part of a larger "grasp" so that a robot manipulator can also exploit the environment in a similar way.

5.2 Recommendations for describing grasp poses and manipulation

To describe static poses, we draw on the opposition types of Iberall [1987] and her observation of how different sub-grasps can be combined to form grasps poses. We then use these grasp types to understand what kinds of manipulation are possible, at points drawing upon Ma and Dollar [2011].

- Wrap grasps (equivalent to palm opposition and power grasps): These grasps involve curling a long finger around an object. These grasps are particularly stable because they maximize contact surface area and also physically block the object from moving in various directions. One or more fingers can perform this grasp, and the strength of the grasp increases with more fingers. Examples of wrap grasps: power grasps like the various cylinder grasps, hook grasps and hooking fingers through loops, the ring grasp, and ulnar grasps (Fig. 5.4a). Robot hands that perform wrap grasps include the Hirose Soft Gripper [Hirose and Umetani, 1978] and other tentacle-like grippers.

- Precision pinch grasps (includes any two-point side opposition or pad opposition grasps): These grasps involve holding an object between two surfaces (fingers or finger groups). The two contacts are the minimum needed to hold the object, but allow the object to be manipulated in a variety of ways: rotated around the axis connecting the two fingers by an outside force, or slid in a direction perpendicular to that axis, or to be manipulated by moving the position of the fingers relative to each other. Examples of precision pinch grasps: palmar pinch, lateral, tip pinch, etc. (Fig. 5.4b).
- Power pinch grasps (equivalent to side opposition). In these grasps, an object is held between two surfaces and a larger contact surface area is used. It is useful for securely holding flat and light objects. However, because of the minimal number of contact areas, it isn't entirely resistant to rotations around the axis between the two contact surfaces. Examples of power pinch grasps: lateral, extension, parallel extension, and adduction grasps (Fig. 5.4c). Parallel grippers make use of this type of grasp.
- Multipoint precision grasp (equivalent to pad opposition grasps with 3+ contacts): Adding more fingers/contacts restricts the movement of the object beyond what is achieved in a pinch grasp. The presence of multiple contact points eliminates some forms of manipulation; however, manipulation can still occur by moving the position of the fingers relative to each other, including sequential regrasps (finger gaiting). Examples of multipoint precision grasps: tripod, precision disk, prismatic, and sphere grasps (Fig. 5.4d).
- Non-prehensile point contact (external force): A single contact not capable of grasping the object but capable of imparting forces on it. Examples of no-nprehensile point contacts: flat hand grasp, thumb and index finger in various variations.

Sometimes a single grasp pose can include multiple grasps, such as the adducted thumb and stick grasps that feature a pinch (side opposition) with thumb and index and wrap (palm opposition) with the other three fingers. At other times, multiple objects can be handled by assigning different grasps to different parts of the hand, such as an ulnar grasp (wrap) that uses the ring and pinky fingers, leaving the top three fingers free to form a tripod grasp on another object. Also, note that there is grey area between different types of grasps. For example, small randomness in the placement of

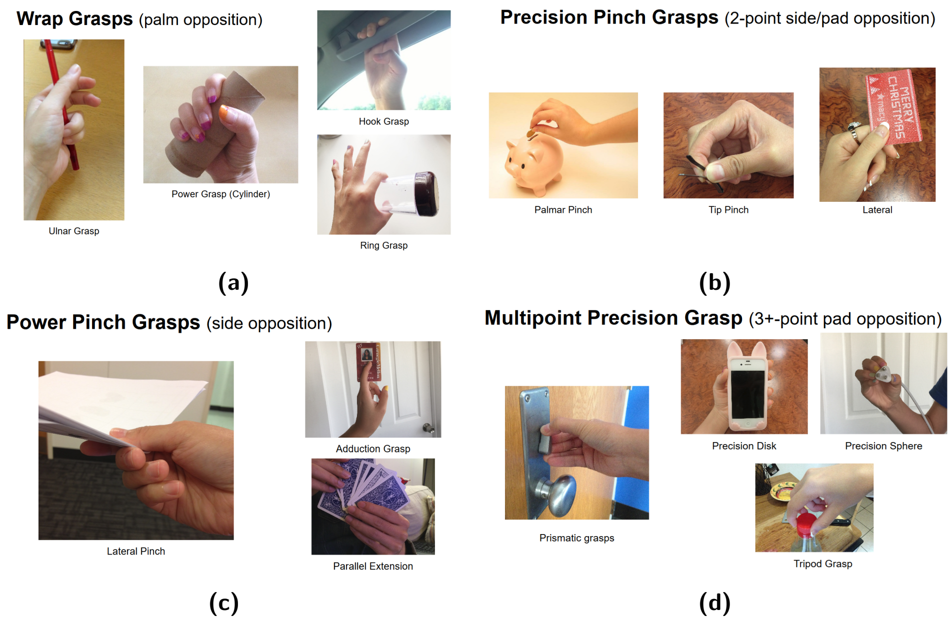


Figure 5.4: Examples of the four grasp types

the fingers can turn a tripod grasp (multipoint precision) into a two-finger extension grasp (precision pinch with the first three fingers) – see Fig. 5.5. The human hand's ability to transition seamlessly between different types of grasps is key to its dexterity. Most robotic hands have the ability to employ different types of grasps. For example, the three-fingered Barrett Hand [Townsend, 2000] is capable of wrapping grasps as well as precision and power pinching. Anthropomorphic robot hands can accomplish all types.

Next we discuss the types of manipulation possible while in each grasp. We draw on the non-anthropomorphic types of manipulation put forward

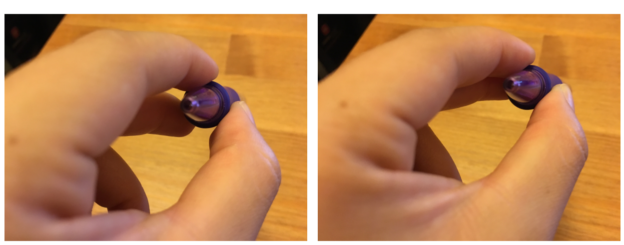


Figure 5.5: Small differences in finger placement can make a grasp more like a tripod grasp (left) or two-finger extension grasp (right).

by Ma and Dollar [2011]: regrasping, in-grasp manipulation, finger gaiting, finger pivoting/tracking, rolling, and sliding.

- Wrap grasps: As pointed out by Napier [1956], there is a trade-off between precision manipulation and stability. Wrap grasps have high stability at the cost of minimal ability to manipulate an object. Types of manipulation possible from a wrap grasp are simply to ungrasp (remove contacts from the object) or to regrasp (add a second grasp and then release this one - see Fig. 5.6). These options are available to any grasp.
- Precision pinch grasps: At the other end of the scale, precision pinch grasps maximize the manipulative capability of the hand at the cost of stability.
  - Rotation via external force (equivalent to finger pivoting): Between a two-point grasp, objects of any type can be rotated around the axis between the two points by an external force (applied by gravity, a finger, or the environment). The interdigital step uses this type of rotation. In addition, for long objects, a torque can be applied away from the grasp location to rotate the object in other directions. This type of manipulation is what is happening when objects are levered up from a support surface. See Fig. 5.7 for examples of each.
  - Translation via external force (equivalent to sliding): By lightening up the contact forces, a precision pinch grasp can allow the object to slip or be pushed across the fingers by an external force. An example of this is the squeeze variation we observed (see Fig. 5.8).
  - Manipulation via finger movement (equivalent to in-grasp manipulation): Extra degrees of freedom allow the fingers involved in a precision pinch to move without moving the hand. This type of manipulation can accomplish translation by moving the fingers but maintaining a similar relationship between them, as in the squeeze motion, or can accomplish rotation by changing the position of the fingers relative to each other (see Fig. 5.9).
- Power pinch grasps: Similar to wrap grasps, these types of grasps are meant for stability. The options of ungrasping and regrasping are available for these grasps. While it is possible to rotate an object even while in a power pinch, if such rotations are desired, it is more energy-efficient to switch to a nearby precision pinch grasp.
- Multipoint precision grasp: These grasps are in between the wrap grasp and precision pinch grasp. In addition to the ungrasping and

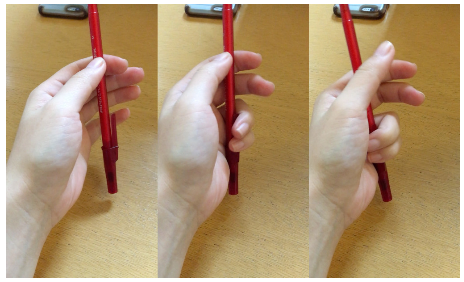


Figure 5.6: Regrasp example

Figure 5.7: Two types of rotation using an external force. In the left, the axis of rotation is between the two points of the grasp. In the right, the two points of the "grasp" are between the thumb and table, and a long lever arm allows rotation in any direction.

regrasping options available to all grasps, they have some ability to do in-grasp manipulation (like a rock – see Fig. 5.9) and sequential regrasps (finger-gaiting – see Fig. 5.10). By removing fingers, it is easy to turn a multipoint precision grasp into a precision pinch.

We summarize this way of classifying grasps and manipulation in Fig. 5.11. The human hand can switch rapidly between different types of grasps in order to take advantage of the manipulation or stability capabilities of each. Some grasp transitions are as simple as adding or removing fingers, but others themselves require some manipulation in order to transition to a different grasp. Depending on the morphology of the hand and the environment, certain grasps, grasp combinations, and intrinsic manipulations become possible. Ideally, a manipulation system should be able to recognize these grasp and manipulation capabilities, including ones that take advantage of the environment.

Our system focuses on the capabilities of a grasp such as the ability to impart forces only, to apprehend, or to perform various types of intrinsic

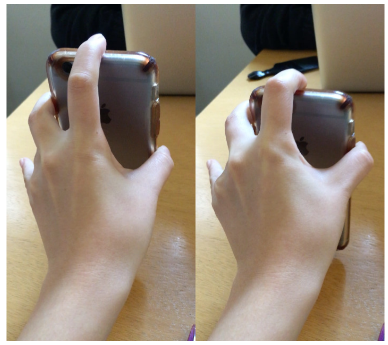


Figure 5.8: Translation example: The precision pinch also allows controlled sliding.

Figure 5.9: Three types of manipulation using finger movement. In the top example, the motion of the two fingers are similar and so have the effect of translating the object without much change to its orientation. In the middle, the two points of the "grasp" are between the flat hand and the shelf, and by moving the location of the hand only, the object is rotated. The last example is another example of finger movement accomplishing object rotation but with a multipoint precision grasp.

Figure 5.10: Sequential regrasp example

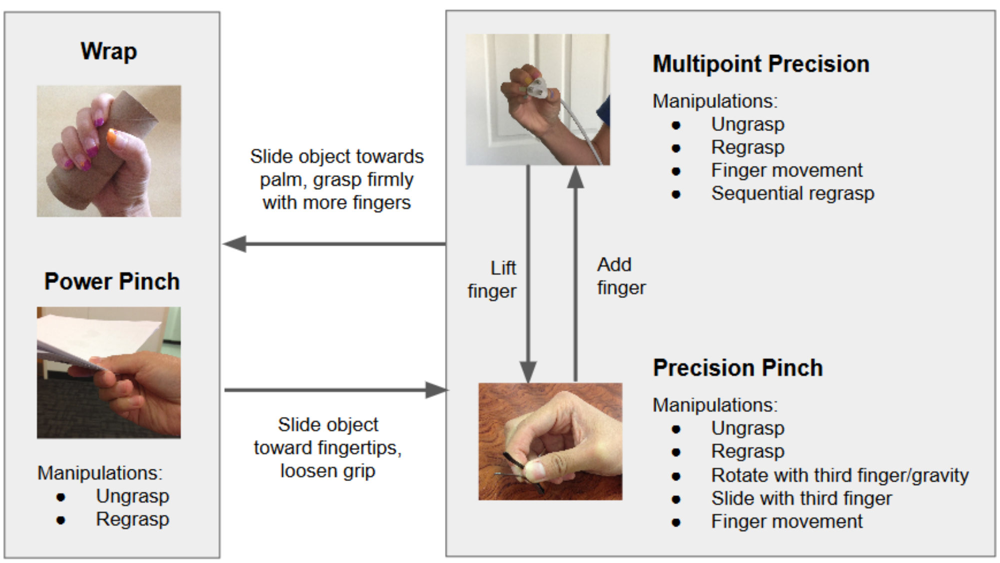


Figure 5.11: Summary of grasp types, types of manipulation possible at each, and general patterns of transitioning between them. Not shown: the non-prehensile point contact.

manipulation. As a result, our system distinguishes grasps based on the following two qualities: (1) the number of **contact groups** (contact directions) and (2) the size of contact areas. A contact group is defined as a set of contacts that are applied in a similar location on the object with a similar normal direction. As more or larger contact points are added, the *location* of contact becomes less coherent, pushing one contact group to begin to look like several contact groups. Likewise, as a finger curls around and passively conforms to an object, the *direction* of contact continuously changes, meaning that one contact group becomes many contact groups.

Fig. 5.12 summarizes how we divide grasp types, and shows grasps that fall in between categories. The difference between one or two contact groups and "many" depends on the number of fingers and amount of curl, as described in the previous section. Some grasps can be done with the ends of fingers to result in a light precision grasp (bottom half of the graph), while other grasps can use more contact surface area to transform a precision grasp into a power grasp (top half of the graph). The lateral is an example of grasp that can seamlessly transition between precision and power. The boxes in Fig. 5.12 are for the purpose of distinguishing grasps clearly in one category vs. grasps that are in-between categories. However, the axes in the graph should be thought of as generally continuous.

We chose the features of number and size of contact groups because they alter the basic capabilities of the grasp. For example, with one contact group, prehension is not possible; imparting forces is the only action that can be taken. Two contact groups allow prehension as well as many types of intrinsic manipulation. For certain object geometries, two contact groups is also sufficient to create a power grasp that largely restricts manipulation, achievable by increasing the size of contact areas. Increasing the number of contact groups but keeping the contact area small results in a precision grasp that is secure but still allows some types of manipulation. By contrast, a wrap grasp simultaneously maximizes the number of contact directions and the area of the contacts, resulting in a very secure grasp. Additionally, both dimensions roughly indicate how much of the object's force/torque space is restricted. As you move toward the top and/or right of the graph, the object's degrees of freedom are increasingly blocked.

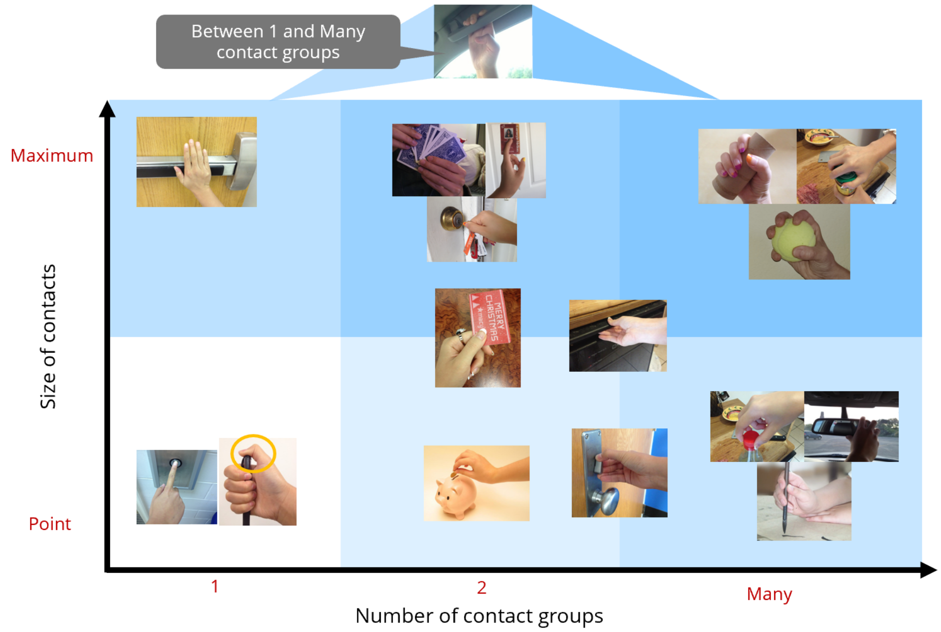


Figure 5.12: Axes underlying our pose classification: number of contact groups and size of contacts. Some grasps are in-between types.

5.2.1 Describing and understanding variations of basic pose types

The system we proposed focuses on elements that change the capabilities of a grasp: the number of contact groups and the size of contact areas. Under the proposed system, the number of basic grasp types is small, and the number of ways of achieving a particular grasp type or of combining multiple grasp types is very large. Variations on a basic grasp pose differ in the following ways:

- Which fingers and how many are involved in the grasp, including **redundant fingers** (extra fingers that don't change the basic grasp type)
- What parts of the fingers are being utilized in the pose
- How far those parts are from the base of the fingers

By changing these properties of a grasp, many variations of the basic grasp types can be generated. Grasp variations involving a different number of fingers and different location of contact can look different, but it is sometimes possible to smoothly transition between them. An example is

Figure 5.13: The lateral and extension type grasps demonstrate how the location of contact and number of fingers involved in a grasp can seamlessly vary.

the lateral and extension type grasps, where changing the finger(s) opposing the thumb and contacting farther away from the base of the fingers can smoothly turn a lateral pinch into an extension grasp (Fig. 5.13).

Various task characteristics can determine which variation of a basic grasp is chosen. Some task characteristics that influence more specific grasp pose include:

- Grasp force: Some grasps are capable of exerting larger forces than others and with less effort. If the object is heavy, then a grasp capable of withstanding large forces is necessary. Adding redundant fingers to a pose can strengthen it, so may be desirable.
- Force direction: Being required to exert a force in a particular direction may require certain fingers or otherwise constrain the space of feasible grasps.
- Multitasking: If it's important to accomplish multiple grasps at the same time, this affects the number of fingers that can be dedicated to a task.
- Graspable areas: Some objects have surfaces that should or should not be grasped. For example, the broad sides of a CD or the length of a key that is to be inserted into a lock are surfaces that shouldn't be part of a grasp, while the trigger of a tool is a surface that should be part of a grasp (usually a non-prehensile component of a composite grasp).
- Size and shape of the intended grasping location of an object: The size of the object determines e.g. how many fingers can be used in a grasp. As mentioned earlier, the shape of the object determines the number of contact groups needed to form a fairly stable power grasp.
- Manipulation ability needed: If fine manipulation is important, contacting the object in the extremities of the fingers allows for more

movement.

All of the above factors place constraints on the more specific grasp pose that can be formed. There are also various properties of the hand design itself that affect the set of possible grasp poses and choice of grasp pose:

- Number of fingers: This affects the ability of the hand to form composite grasps.
- Placement of fingers: This determines what fingers are able to oppose each other, in what direction forces may be exerted, etc.
- Finger length: Longer fingers can grip larger objects, are better able to wrap, have more ability to move an object at the fingertip, and have longer reach.
- Surface/contact properties of parts of the finger: Different parts of the human hand have different properties. The finger pads are sticky. The inner (ventral) side of fingers in general is soft/deformable, while the backs (dorsal side) of the fingers are bony and more rigid, and the nail is even more so. The contact properties of robot hands can be freely customized and do not have to resemble those of the human hand. In either case, areas of the hand with different contact properties may be more or less desirable in forming a grasp.

To summarize, the system we detail here for describing grasp pose and in-hand manipulation focuses on two aspects of the grasp pose: number of contact groups and size of contact area. These properties are strongly tied to the manipulation capabilities of the hand, in particular the ability or inability to apprehend objects, and the trade-off between stability and manipulation capability (i.e. power vs. precision trade-off). However, there are many aspects of grasping that are not described by this system that result in variations on these basic grasp types. These variations can be conceived of in a hand-centric way (first list) or a task-centric way (second list). The whole set of possible grasp variations is itself influenced by the hand design (possible design considerations in the third list).

5.3 Annotating other parts of manipulation

Aside from grasp pose and intrinsic manipulation, what other aspects of manipulation should be noted? As mentioned in Chapter 4, the Bullock et al. [2013] taxonomy is useful for segmenting manipulation into phases, and can be augmented with additional taxonomies. For example, once a stable grasp has been formed, the object can be manipulated in various ways. The force types, motion and force directions, and constraints on motion and force featured in Chapter 3 can be useful here.

In addition, it is important to note errors and error correction strategies. The fact that humans frequently make mistakes and experience failures when manipulating is possibly reassuring to people who work on dexterous robots. However, it suggests that being able to detect errors – which includes understanding the intent behind a motion or manipulation – and recover from them is potentially an important skill. Errors can occur at various stages of grasping and placing. They can occur while approaching the object in the form of unintended collisions with parts of the environment. They can occur while closing the hand around the object in the form of a miss (the fingers fail to contact the object) or an ejection (the object flies out of the grasp, changing its position). Errors can also occur during transport – dropping the object. Finally, errors can occur during placing in the form of an unsatisfactory place (the object doesn't end up in the desired configuration) or a post-place drop (the object was unstable and fell). Collisions and misses provide helpful information about the environment and object, assuming that the hand has tactile sensing capabilities. The other types of errors require the manipulation to be redone.

5.4 Annotation example

If we were to redo the annotations in Chapter 4 with the insights and knowledge we have now, the procedure might look like this:

1. Use Bullock et al. [2013] taxonomy to segment motion into phases and to identify when different parts of the hand are acting as independent units.
2. Note any errors or unintended object motion.
3. For stable grasps, identify the grasp type (using the above system) and the type of action the stable grasp is being used for (transport, tool usage, etc.).
4. When the hand or object is moving outside of a stable grasp, identify the purpose of this movement. Also identify the way the hand (or hand-environment system) accomplishes it using the above system.

Because the motions in Chapter 4 were largely pick-and-place actions, the type of action a stable grasp was used for was not important outside of a few motions. However, if it is important to describe the action, the system in Chapter 3 or object-centric taxonomies like that of Wörgotter et al. [2013] and Leidner et al. [2015] might be used.

5.5 Discussion

In this chapter, we used the insights from our observations of grasping in Chapters 3 and 4 to highlight some difficulties in describing grasping. Most of these difficulties center around defining a hand or a grasp due to the way the hand can multitask and multiple hands or parts of the environment can collaborate to form a grasp. We recommended bringing the environment into understandings of grasping by using existing non-anthropomorphic ways of describing grasps and manipulations.

We drew upon Iberall's opposition types [1987] as a starting place, and altered them slightly to be less anthropomorphic and to include non-prehensile "grasps." The main benefit of using these basic grasp types is that they are able to describe a large array of grasp poses and manipulation, including environment-aided ones. By breaking down grasps into smaller components, this system is able to describe many of the grasps found in Chapters 3 and 4 that are missing from previous taxonomies, especially composite grasps and grasps involving specialized fingers. By not tying grasp description to the morphology of the human hand, a wide range of manipulation is also able to be described, including environment-aided manipulation like that found in Chapter 4 (see Tables 4.1 and 4.2).

Another main feature of our proposed system is that it focuses on how different types of grasps change the prehensile and manipulative capabilities of the hand, resulting in a set of grasp types that differ based on number of contact groups and size of contact area. Interestingly, previous work [De Souza et al., 2012] also supports the idea that Iberall's opposition types are connected to the precision vs. power dichotomy and are thus useful for understanding the purpose of a grasp. By studying grasping in more detail, we reveal the limitations of current taxonomies and use those insights to create a more promising system.

6

Factors in bimanual grasping

While it is difficult to assess how much of daily human manipulation involves both hands simultaneously (bimanual manipulation), various people surmise that bimanual manipulation is the predominant form of manipulation. For example, Kimmerle et al. [2003] claim “The majority of activities of daily living are typically executed bimanually, for example, getting dressed, cooking, eating, and the majority of tool uses”, while Guiard [1987] reviews various handedness inventories and finds that slightly more than half of the tasks listed are bimanual.

In addition, bimanual manipulation can appear in many different forms. The bimanual activities referenced above are mainly pure bimanual actions – ones where each hand is given a different role and each hand is necessary in order to effectively accomplish the task. However, bimanual actions also come in other flavors, like simple one-handed tasks for each hand that happen to overlap in time (bimanual multi-tasking) or ones where an extra hand helps in the handling of larger or heavier objects but is not strictly necessary. Various taxonomies of bimanual manipulation [Grunwald et al., 2008; Surdilovic et al., 2010] based on Guiard’s analysis [1987] differentiate between non-coordinated bimanual actions, where the two hands are each performing their own one-handed task independent of the other hand, and coordinated bimanual actions, where the two hands have to coordinate in space and/or time. Within the coordinated type, they define symmetric/anti-symmetric bimanual actions as both hands performing the same task simultaneously, and asymmetric or differentiated biman-

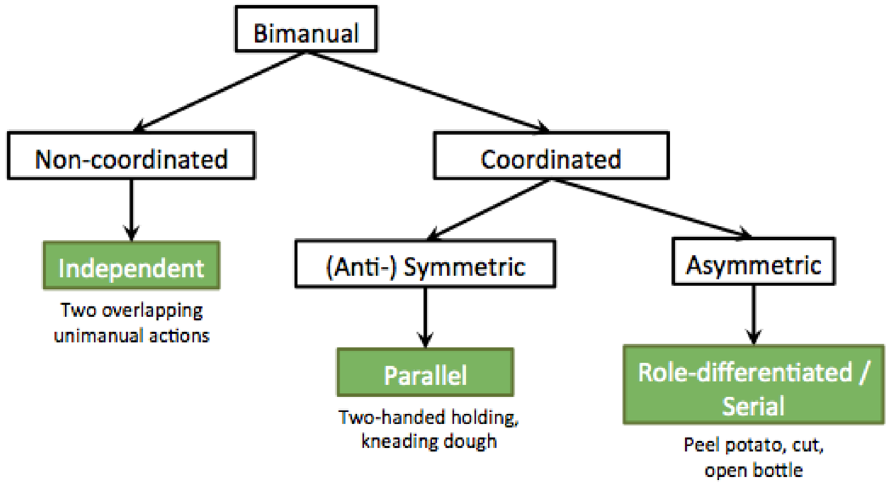


Figure 6.1: Types of bimanual manipulation as laid out by Guiard [1987].

ual actions as each hand having its own role (Fig. 6.1).

In the case of transporting objects, multiple strategies capable of accomplishing the task are available: a person can use one hand exclusively to transport the object, can grab the object with two hands, or pick an object up with one hand and transfer it to the other hand. A key question in studying motor behavior is understanding what criteria determine how one motor plan is selected over alternative movement strategies. It is possible that these complicated choices are determined by an underlying minimum principle such as time or energy minimization [Engelbrecht, 2001], which allows people to select a single motion plan that is responsive to arbitrary starting and ending positions of the transport task; size, weight, and shape of object to be transported, etc.

Researchers have previously shown that size and weight of an object affects whether people grasp it with one hand or two hands [Cesari and Newell, 2000]. In addition, they found that hand length can be used to fairly accurately predict the transition point when an object starts to be handled with two hands. The weight of the subject's hand also has some ability to predict the weight at which that subject will transition from one to two hands, but there is a greater amount of unexplained variation in the weight case than in the size case. Researchers have also shown that the end goal affects the usage of the left and right hand for grasping in a Tupperware-stacking task [Rosenbaum et al., 2010]. When the end goal is to the right, people walk from left to right stacking containers along the way, using mainly their left hand to grab and place. The opposite is true when the end goal is toward the left. Alternatively, some people use both hands simultaneously (symmetric bimanual strategy) to grab and place, although this strategy was used less frequently and was less responsive to end goal location.

None of these studies, however, fully consider the task of object transport. The task of transporting an object opens up new strategies including one where the object is grabbed with one hand but then handed off to the other hand, which places it (hand-off strategy). Cesari and Newell [2000] only consider the act of grasping (apprehending) an object, briefly lifting it, and replacing it. As such, handing off between hands is not a strategy under consideration. While Rosenbaum et al. [2010] consider the whole process of transporting objects to a final destination, they consider grasping actions independently from placing actions, meaning that it is unfeasible to identify instances where an object might have been handed off between hands. In order to fully understand the choice of using one vs. two hands in the task of object transport, handing-off actions must be explicitly considered. An open question not answered by these studies is how object and task properties affect entire transport strategy, including not just usage of the pure unimanual and symmetric bimanual strategies but usage of the hand-off strategy as well.

The study in this chapter investigates two questions: the first is what effect object and task factors have on the use of unimanual, bimanual, and hand-off transport strategies, and the second is what is the underlying reason those strategies are chosen. We expect that the same effects of object size and weight that affected use of one- and two-handed grasping would manifest in transport as well. Two hands can function as a large manipulator [Bullock et al., 2013] and using two hands can spread the weight of an object to a more comfortable load at each hand. While there is no theoretical or empirical work on the effect of size or weight on hand-offs in adults, an infant study researching the development of manipulation skills over time [Palmer, 1989] recorded when infants handed objects off hand-to- hand (switching), finding that heavier objects were handed off less frequently, although no explanation was offered for why this might be.

We also expected that transporting an object that requires its balance to be carefully maintained would push people to use the symmetric bimanual strategy. It has been found that manipulating an object that must be carefully balanced has the effect of increasing task difficulty, which in turn influences selection of (pre-)grasp strategy [Chang et al., 2009]. In particular, this study found that the difficulty of the object-balancing task had the effect of increasing the amount of pregrasp rotation people performed. The pregrasp rotation of the object put the hand configuration into a region that was shown to have greater lifting capabilities. This range of angles may be related to the "comfortable" mid-range of movement where more precision of hand motion can be applied (see Rosenbaum et al. [1996] and the middle-is-faster effect). It is possible that the use of two hands simultaneously might have a similar effect of increasing precision of control. An infant study by Palmer [1989] indicates that when surfaces are stable (hard rather than foam), infants spend more time holding an object in a single hand, which may be because unstable surfaces make the unimanual strategy more difficult. However, there has not been work specifically investigating whether the need to carefully maintain an object's balance has an effect on people's choice of one vs. two hands.

Hand-offs have not been studied much in previous literature. Studies have shown that object location influences the choice of left and right hands when grasping. In particular, studies on handedness find people prefer to not cross the midline when reaching. For example, Gonzalez et al. [2014] found that for right-handed participants, over 95% of objects located to the right of the participant's midline were reached for with the right hand, while 65-90% of objects located to the left of the participant's midline were reached for with the left hand. Hand-offs from one hand to the other may be used as a way to avoid crossing the midline when grabbing and placing, so would be used when the start and goal location are on different sides of the body.

The second question we sought to answer was what explains the choice between the unimanual, bimanual, and hand-off strategies. In particular, we wanted to investigate whether minimum principles are a plausible explanation for choice of transport strategy. The time it takes to execute an action and the metabolic energy consumption in executing it are common minimum principles used in biology to explain behavior, and may be useful for understanding motor behavior as well [Engelbrecht, 2001]. In this study, we considered the explanatory ability of two possible costs: the quickness with which the movement could be executed, and the amount of rotation each strategy requires in order to execute. In these experiments, we wanted to test if either of these measures – movement duration and body rotation – had the ability to explain people's transport strategy choices.

The following set of experiments seeks to answer these questions related to the use of one and two hands in transport. The first experiment focuses on a larger set of start and goal positions, while the second experiment focuses on a larger set of object sizes and weights.

6.1 Experiment 1

6.1.1 Measures and Hypotheses

The first goal of this experiment was to determine how various object and task properties affect whether people use one or two hands to transport a bowl. The object properties varied were bowl size and weight. The task properties varied were balance (whether the bowl's balance was important) and configuration (the start and goal position of the bowl relative to the subject). We collected which hand(s) subjects used to pick and place the bowl. Our expectations were as follows: Larger object size, heavier object weight, and the presence of a balance requirement would encourage the use of the symmetric bimanual strategy. Start and goal position would affect the use of hand-offs, as people would use their left hand to pick/place when the bowl/goal was in the left hemispace and use their right hand to pick/place when the bowl/goal was in the right hemispace.

The second goal was to investigate the reason underlying strategy selection. In order to answer this question, we collected movement time and amount of hip rotation. We then compared how the choice of strategy and experimental conditions affected the movement time and rotation. We expected that strategies that people favor and use frequently would be quicker or involve less body rotation.

6.1.2 Method

Participants

We ran an experiment with 16 participants (4F, 12M; 14 right-handed, 2 mixed-handed (self-reported handedness, with a prompt "The dominant hand is the one typically used for writing, brushing teeth, throwing, using a spoon, opening a box (the one on the lid, etc.)"); mean age = 27.8 (SD D 6.8)). In addition, a left-handed participant was recruited and data collected. However, the pattern of this participant's data differed noticeably from that of the other participants, for example, right-handed participants used their right hand unimanually more often than they did their left hand, and this was reversed for the left-handed individual. As such, this participant's data were discarded and are not represented in the following results. The method was approved by the Disney Research Institutional Review Board, and the informed consent of all participants was obtained in accordance with the Declaration of Helsinki.

Procedure

Each trial consisted of the participant moving a bowl from one table to another. The experiment varied bowl size (two conditions) and weight (two conditions), the presence within the bowl of a tube with a ball balanced on top (two conditions), and the subject's starting location and facing direction. There were seven different standing place-facing direction combinations (hereafter called "configurations") in which the subject could stand (Fig. 6.2). There was one trial per condition, resulting in 56 trials overall per participant (2 size $\times$ 2 weight $\times$ 2 balance $\times$ 7 configurations).

The bowls moved were metal IKEA® BLANDA BLANK bowls of two different sizes. The BLANDA bowls were chosen due to their simple, symmetric geometry – in particular, their lack of a lip that could be used for grasping – and their similar shape across sizes. The "small" bowl was 12.2 cm $\times$ 6.1 cm (diameter, height), while the "medium" bowl was 20.2 cm $\times$ 9 cm. The "light" bowls were filled with aquarium stones to the total weight of 290 g, while bowls in the "heavy" condition were filled to 640 g total.

In the "balance" condition, a toilet paper roll (4.1 cm diameter $\times$ 10.5 cm height) with a 4" (10 cm diameter) styrofoam ball balanced on top was used to add the difficulty of balancing to the moving task. The roll was inserted into and stabilized by the aquarium stones inside the bowl. For bowls without enough stones to stabilize the roll, the roll was attached to adhesive putty at the bottom of the bowl. The roll and ball were removed in the "no balance" condition.

There were seven possible configurations (Figure 6.2b). The experiment consisted of seven blocks of eight trials. Within a block, all trials shared

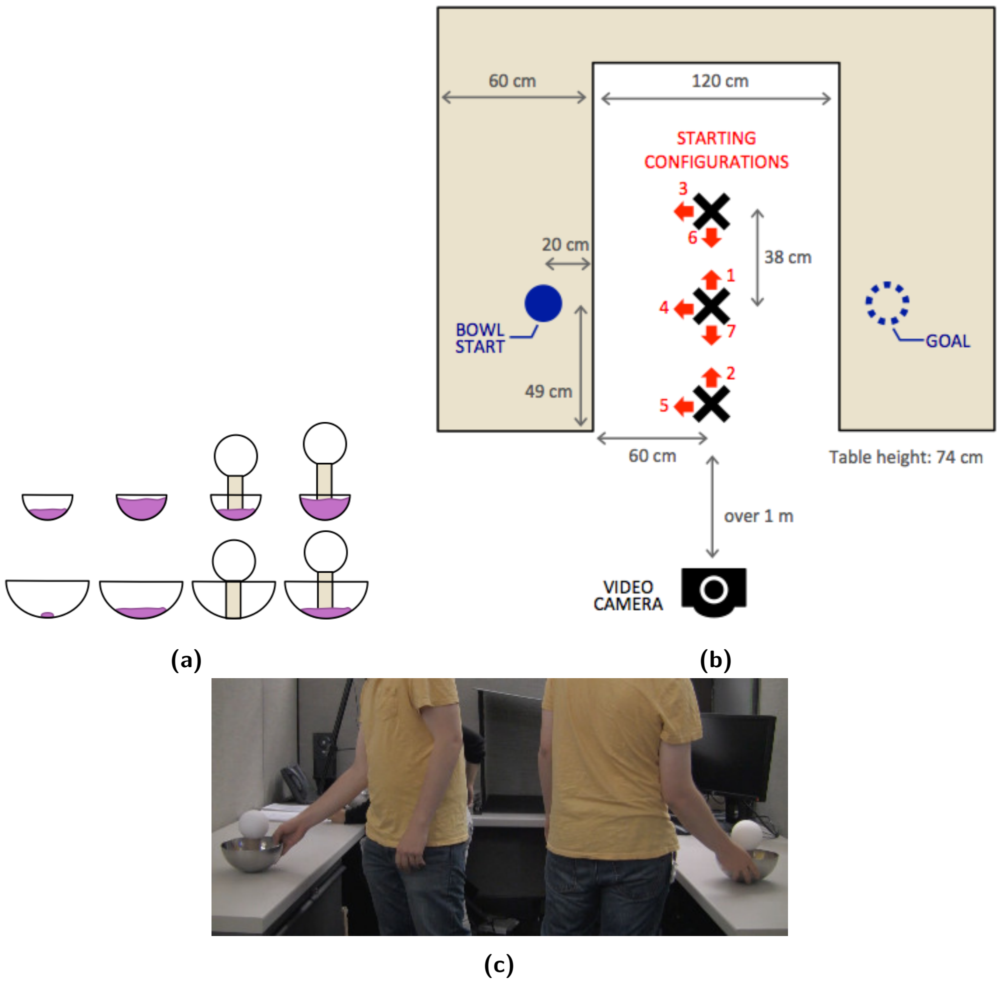


Figure 6.2: Experimental setup. (a) Bowls can be small or medium; light or heavy; and with or without a balance tube. (b) The seven arrows in this diagram indicate the seven possible starting configurations of the participant, which consist of a standing location and a facing direction. There are three standing locations with either two or three facing directions, yielding a total of seven possible start configurations. (c) Screenshots of video collected as part of the experiment. These screenshots feature the start and end of object transport within a trial.

the same starting configuration. This clustering of trials by starting configuration was done to avoid making the subject move around after each trial. The presentation of these blocks was randomized, and the presentation of trials within each block was randomized.

At the start of the experiment, the participant was instructed to not knock over the styrofoam ball used in the "balance" condition. If the ball fell from the tube, the trial was repeated. The error was recorded but the trials with errors were not included in the analysis – only successful trials were analyzed. Participants were only required to start the trial at a particular spot and facing a particular way; once the trial started, they were allowed to walk around the experimental space freely while transporting the bowl.

The trials were videotaped with an ordinary video camera that included the participant, start location, and goal location in the frame. The entire procedure including instruction and obtaining consent took under 30 minutes.

Data Processing

Videos were reviewed by the researcher, and the following annotations were made: (1) grasp strategy, (2) approximate transport duration, and (3) approximate hip rotation. Strategies were differentiated by which hand(s) were used for grasping and placing (left, right, or both hands). Using this way of distinguishing transport strategies, there are nine possible strategies:

- L Left only – One-handed pick up, transport, and place with left hand
- R Right only – One-handed transport with right hand
- LR Hand-off (l → r) – Hand-off from left hand to right (pick up with left, place with right)
- RL Hand-off (r → l) – Hand-off from right to left
- LB Left → bi – Pick with left hand, add right to place bimanually
- RB Right → bi – Pick with right hand, add left to place bimanually
- BI Bimanual – Pick up, transport, and place with both hands
- BL Bi → left – Grab bimanually, place with left hand only
- BR Bi → right – Grab bimanually, place with right hand only

For duration, the start of transport was considered to be the second when a stable grasp was formed* and the end was the second when the bowl made contact with the goal table. Duration was calculated as the number of seconds in between.

To calculate rotation, first, facing directions of the hip at transport start and end were recorded, rounded to the nearest 45°.For example, in Fig. 6.2c, the participant's hips at the start of transport faced forward-left while at the end, the participant was facing a direction between straight backward and backward-right. This was determined to be closer to the backward direction. The facing direction of the hip (as opposed to the shoulder or chest) was chosen because its orientation was easiest to estimate visually. The angles of the hip's facing direction at trial start, transport start, and transport end were recorded. Rotation was then defined as the octants rotated between trial start and transport start, plus the octants rotated between transport start and transport end. For the trial depicted in Fig. 6.2c, this participant started the trial facing the bowl, rotated roughly one octant at the time of grasping, and then rotated counterclockwise roughly five octants to place.

Data Analysis

First, we analyzed the effect of the experimental factors (size, weight, balance, and configuration) on the response of choice of transport strategy using a mixed-effects generalized linear model with a logistic link function (a generalized linear mixed model or GLMM). This model was fit to the data using the glmer function of R's lme4 package [Bates et al., 2015]. This analysis method was chosen because it was capable of handling both binary response data and the repeated measures experimental design. The response variables analyzed were usage of bimanual, hand-off, and unimanual strategy (three separate analyses with binary outcomes). Size, weight, balance, configuration, and their interactions were used as fixed effects in the model. Variation between participants was modeled as a random intercept. Because models had difficulty converging when random slopes were added, random slopes were not included in the model. A stepwise procedure comparing likelihood ratios (using ANOVA) was used to eliminate non-significant variables until no more could be removed (a significance level

*When grasping, subjects would first move and adjust their fingers on the bowl; then their fingers would stop moving for a moment as the participant braced to take on the load of the bowl. This solidifying of the grasp pose right before lifting was considered the moment a stable grasp is formed.

of .01 was used to determine which factors to keep). For effects remaining in the model, plots showing the mean probability of a strategy being used under each condition and an estimation of the standard error of that mean were generated using the effect function of R's effects package [Fox, 2003].

In order to understand the reason behind people's preference of certain strategies over others, a second analysis investigated the effect of strategy (bimanual, hand-off, and unimanual) on transport duration and body rotation using linear models with duration or rotation as the response; strategy and the four experimental variables as fixed effects; and participant as a random intercept. A final model was selected by removing nonsignificant effects using likelihood ratios.

6.2 Experiment 1 Results

6.2.1 Strategy frequency overview

Frequencies of grasp strategies are summarized in Fig. 6.3. All nine possible strategies were observed at least once. However, the strategies we were mainly interested in—the symmetric bimanual strategy (BI), the two hand-off strategies (LR, RL), and the two unimanual strategies (L, R)—were much more common than the four "mixed" strategies (LB, RB, BL, BR) that involved changing the number of hands grasping the bowl during transport. These four mixed strategies were used in less than 5% of trials. We therefore focus on the bimanual, hand-off, and unimanual strategies in our analysis.

6.2.2 Effect of experimental variables on grasp strategy

For all three strategies – bimanual, hand-off, and unimanual – balance and configuration remained in the model. In addition, the balance × configuration interaction effect remained in the unimanual model ($\chi^2(6) = 48.8$; $p < .0001$).

Balance as a main effect was significant in bimanual ($\chi^2(1) = 235$, $p < .0001$) and hand-off ($\chi^2(1) = 127$, $p < .0001$) strategies, but not the uniman-

	Strategy	Frequency	
L	Left only	79	9%
R	Right only	252	28%
LR	Hand-off (l → r)	174	19%
RL	Hand-off (r → l)	153	17%
LB	Left → bi	3	<1%
RB	Right → bi	5	<1%
BI	Bimanual	195	22%
BL	Bi → left	1	<1%
BR	Bi → right	34	4%

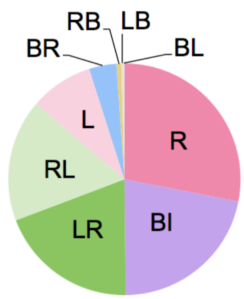


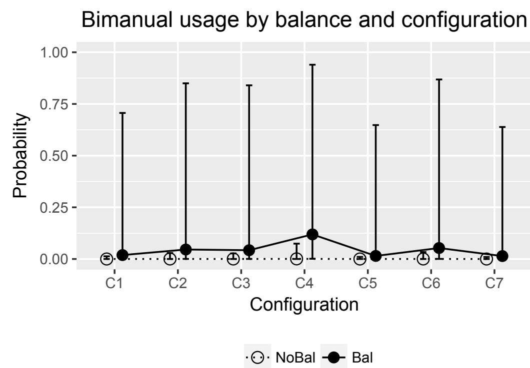
Figure 6.3: Frequencies of each strategy.

ual ($\chi^2(1) = 3.1$, p = .080). When the balance requirement was in play, the bimanual strategy was more likely, the hand-off strategy less likely, and had a more complicated effect on the unimanual strategy. In configurations where the unimanual strategy was frequently used (C3, C4, and C5 configurations involving moving the bowl from front to back; see Figure 3), the balance requirement cut down unimanual usage. In the other four configurations (ones involving moving the bowl left hemispace to right hemispace or vice versa), however, unimanual usage increased in the balance case.

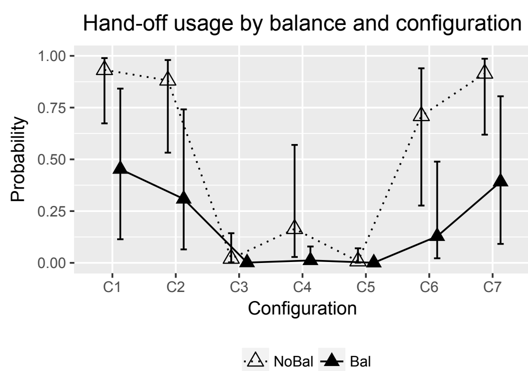
The three strategies were also affected by configuration (bimanual: $\chi^2(6) = 18.0$, p = .006; hand-off: $\chi^2(6) = 371$, p < .0001; unimanual: $\chi^2(6) = 257$, p < .0001). Hand-offs were the strategy people used most often at C1, C2, C6, and C7, which involved moving the bowl from left to right or vice versa. The unimanual strategy was used most at C3, C4, and C5, which are the three configurations where the bowl is moved from front to back. Fig. 6.4 summarizes these balance and configuration effects.

Fig. 6.5 provides a useful way of visualizing configuration and balance effects. It arranges the raw strategy usage data† at each configuration to be at the angles where the bowl starts and ends relative to the participant. For example, at C1, the bowl starts out directly to the left of the participant and is moved to the participant's right. In this configuration, the hand-off left-to-right (LR) strategy is the most common strategy (used about 60% of

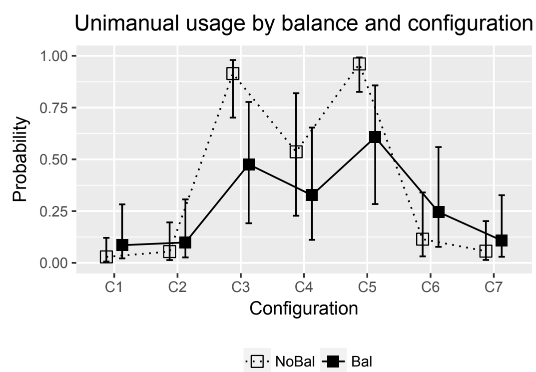
†These data separate out the two unimanual strategies (L and R) and the two hand-off strategies (LR and RL) and also show raw frequency of each strategy averaged over participants, rather than the predicted probabilities of Fig. 6.4 that account for random variation between participants.



(a)

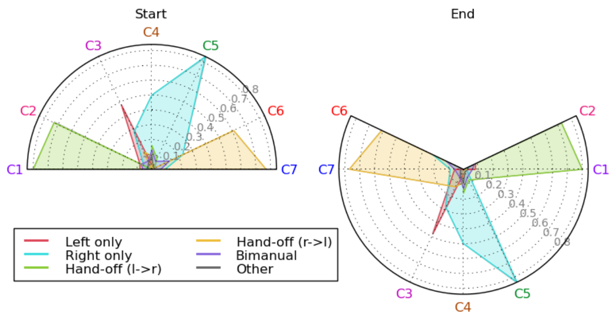


(b)

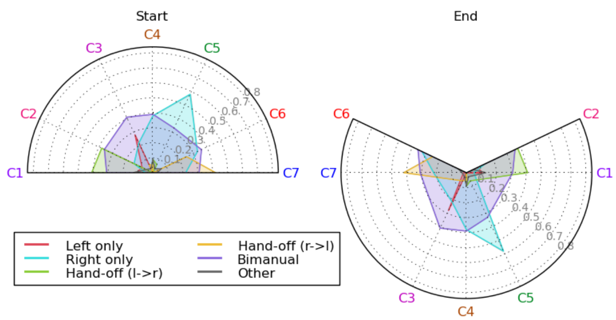


(c)

Figure 6.4: Effects of balance and configuration on the usage of the three strategies: the main effects for the bimanual and hand-off strategies, and the significant interaction effect for the unimanual strategy. The bars signify estimated standard error of the mean in log-odds space.



(a)



(b)

Figure 6.5: Angle plots showing popularity of strategies at each position. Data is plotted at starting and ending angles, from the perspective of someone facing up. (Down indicates the goal is behind the subject; left and right indicate toward the left and right hands.) Strategy popularity is shown for (a) all cases, (b) no balance cases, and (c) balance cases only.

the time) followed by the bimanual strategy.

Neither size nor weight was significant in any of the models. Size and all related interaction effects were able to removed from the full model (unimanual: $\chi^2(28) = 18.3$, p = .92), or from a partial model after the removal of weight (bimanual: $\chi^2(14) = 13.3$, p = .51; hand-off: $\chi^2(14) = 20.9$, p = .10). Weight and interaction effects involving weight were able to be removed from the full model (bimanual: $\chi^2(28) = 24.2$, p = .67; hand-off: $\chi^2(28) = 24.2$, p = .67) or from a partial model after the removal of size (unimanual: $\chi^2(14) = 13.8$, p = .47).

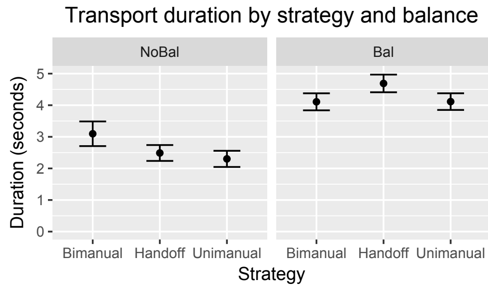


Figure 6.6: Transport duration by strategy and balance.

6.2.3 Reason for strategy choice

First, we investigated the possibility that minimizing movement time might be underlying people's strategy choices. Using the generalized linear model that had duration as a response variable, both strategy ($\chi^2(2) = 18.9$, $p < .0001$) and the strategy $\times$ balance interaction ($\chi^2(2) = 40.2$, $p < .0001$) were significant. Examining the significant strategy $\times$ balance interaction effect (Fig. 6.6) reveals that the bimanual strategy is slower than the other two strategies in the no-balance case, while the hand-off strategy is slower in the balance case, which potentially explains the lower hand-off selection and higher bimanual selection in the balance case found in the first analysis. However, duration does not explain why people often decline to use the bimanual strategy in the balance case, or why the unimanual and hand-off strategies are so dominant in certain configurations.

The second possibility we investigated was that the desire to minimize rotation might be underlying strategy choice. Strategy ($\chi^2(2) = 475$, $p < .0001$) and the strategy $\times$ configuration interaction ($\chi^2(2) = 196$, $p < .0001$) remained in the rotation model. As Fig. 6.7 illustrates, (1) bimanual strategies require more rotation than unimanual strategies, which generally (except at C7) require more rotation than hand-off strategies; and (2) configuration affects the rotation needed at each strategy by different amounts. In particular, the hand-off strategy needs more rotation at C3, C4, and C5, which could be responsible for the low popularity of hand-offs in those configurations.

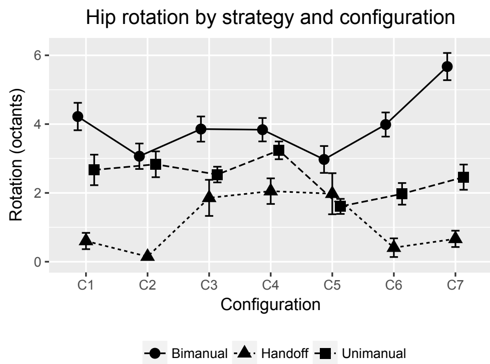


Figure 6.7: Total rotation by strategy and configuration. The y-axis is the average number of octants rotated for each strategy at each configuration.

6.3 Experiment 2

The previous experiment did not yield an effect of size or weight, as expected from previous work. It is possible that the bowl weights and sizes used did not span a sufficiently broad range to include the transition point where individuals switch from one-handed to two-handed grasping, as found in Cesari and Newell [2000]. The focus of this experiment was to test if weights and sizes larger than the ones previously investigated could elicit a size/weight effect on bimanual usage. Four bowl sizes and three weights were used. In addition, we replaced the method of collecting movement time and rotation through visual inspection of video with a more accurate motion capture system. Finally, we collected information on step counts and head and chest rotation for analysis and comparison with the hip rotation measure used in Experiment 1.

6.3.1 Measures and Hypotheses

We hypothesized a greater range of sizes and weights would elicit a switch from unimanual strategy to bimanual strategy as the dominant transport strategy as observed in previous work. In addition, we hypothesized the balance and configuration effects on strategy and the strategy effects on movement time and rotation found in the first study to appear in this study as well.

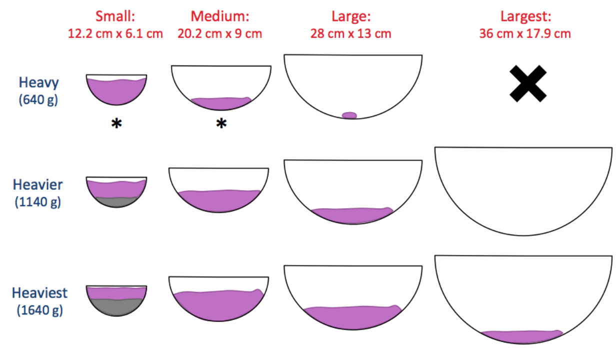


Figure 6.8: The eleven different size/weight combinations for the bowls. The heavier small bowls were padded with lead to increase density. The two bowls marked with an asterisk are repeated from Experiment 1.

6.3.2 Method

Participants

We ran an experiment with 16 participants (6F, 10M; 15 right-handed, 1 mixed-handed; mean age = 26.2 (SD = 6.1)). The method was approved by the Carnegie Mellon University Institutional Review Board, and the informed consent of all participants was obtained in accordance with the Declaration of Helsinki.

Procedure

Each trial consisted of moving a bowl from one table to another. There were 11 size/weight combinations for the bowls (Fig. 6.8) and three possible starting configurations (Fig. 6.9). There were two balance/no balance conditions as in Experiment 1. There was one trial per condition, resulting in 66 trials overall per participant (11 bowls × 3 configurations × 2 balance).

Two more IKEA® BLANDA BLANK bowls were added: a large bowl (28 cm × 13 cm (diameter, height), 600 g), and largest bowl (36 cm × 17.9 cm, 1110 g). Three weight levels were used: the “heavy” condition of Experiment 1 (640 g), as well as a “heavier” condition (1140 g) and a “heaviest” condition (1640 g). There was no heavy condition for the largest bowl because it weighed more than 640 g when empty. Greater weights for the smallest

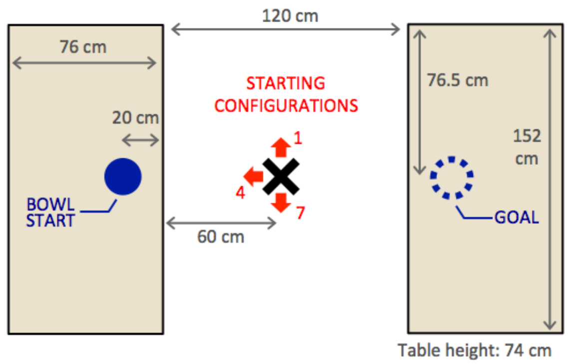


Figure 6.9: Experiment 2 setup with only one starting location (with three facing directions).

bowl were achieved with sealed bags of lead at the bottom.

The main difference between this experiment and the previous one is the use of motion capture technology (Vicon system, 120 fps resolution) to more accurately determine transport times and facing angles. Reflective markers were placed on various parts of the participants (Fig. 6.10), including the middle of the back of their hand, and on each bowl. The bowl was oriented with the marker at the "12 o'clock" position from the participants' point of view to minimize interference during grasping.

Unlike the previous experiment, all 66 trials were fully randomized, with facing direction allowed to change from trial to trial rather than clustering trials with the same starting configurations together. The procedure was otherwise identical to the first experiment. The entire procedure including instruction, obtaining consent, and using motion capture markers took 30–35 minutes.

Data Processing

Motion capture data were used as an alternate way to calculate transport duration and rotation. For determining both of these, transport start and end were determined by when the velocity of the marker on the bowl fell below a 0.1 m/s threshold in each direction starting from the peak velocity timestep. Duration was defined as the time between these two timesteps.

The orientation of the hip at transport start and end was calculated as the vector from the midpoint of the back hip markers to the midpoint of

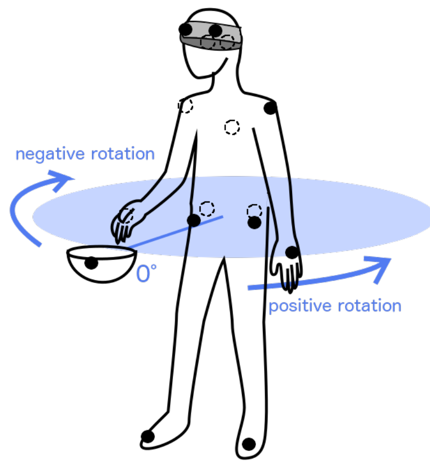


Figure 6.10: Motion capture setup for Experiment 2. The black and dotted circles represent the placement of 16 reflective markers on the front and back of the participant. The direction of the bowl is defined as zero degrees (the direction of the goal is ±180°) and counterclockwise rotations are positive angles.

the front hip markers. The direction to the bowl was defined as zero degrees and samples taken between transport start and end were used to determine which direction the participant rotated between the two time points. Hip orientation at the start and end were then used to calculate rotation as in Experiment 1.

For head orientation, a similar procedure was used to calculate the head facing direction from four markers. Chest orientation was calculated by finding the direction normal to the line connecting the shoulder markers and choosing the facing direction to be the one further (greater than 90°) from the back marker. For the head, torso, and hip, transport rotation was defined as the rotation from the moment of picking the bowl up to the moment of placing it; total rotation was defined as transport rotation plus the amount of rotation from the starting configuration to bowl picking.

Data Analysis

Analysis was identical to Experiment 1. Grasp strategy usage was analyzed using three generalized linear mixed models (GLMM). The effect of strategy on duration, rotation, and step count was analyzed using a linear mixed model that included the experimental factors, grasp strategy, and their interactions. Inclusion of a factor in a final model was determined using likelihood ratios between the model with the factor included and one without it. One benefit of using generalized linear mixed models is that they are capable of handling the unbalanced experimental design caused by the lack of a Heavy Largest bowl. To calculate group means for effects involving both size and weight, the lsmeans package of R was used [Lenth, 2016].

To explore the effect of hysteresis, we tested whether the number of bimanual uses in directly preceding trials affected whether participants used the bimanual strategy again. All trials under conditions shared between Experiments 1 and 2 (small or medium bowls of the heavy weight, in configurations C1, C4, and C7) were analyzed regardless of whether their preceding trials were also shared. Participants who either never or always used bimanual strategy affected the results of this analysis and so were removed. The remaining trials were analyzed using a GLMM to compare if using the bimanual strategy in none or all of the trials in the preceding set had a significant effect on the outcome. This analysis was done for one, two, and three previous trials.

6.4 Experiment 2 Results

6.4.1 Basic strategy frequencies and comparison to Experiment 1

Strategy frequencies are summarized in Fig. 6.11. Similar to Experiment 1, the four mixed strategies (LB, RB, BL, and BR) were used in a small proportion of the trials (3.4%). Unlike in Experiment 1, the bimanual strategy (BI) was the most popular strategy. We can limit the examination to only trials featured in both experiments. These are all three configurations of Experiment 2, the small and medium sizes at the "Heavy" weight only, and with both no-balance and balance cases included. Even so, the pattern of strategies is drastically different (Fig. 6.12), despite the task being the same.

	Strategy	Frequency	
L	Left only	55	5%
R	Right only	189	18%
LR	Hand-off (l $\rightarrow$ r)	84	8%
RL	Hand-off (r $\rightarrow$ l)	58	5%
LB	Left $\rightarrow$ bi	20	2%
RB	Right $\rightarrow$ bi	25	2%
BI	Bimanual	600	57%
BL	Bi $\rightarrow$ left	7	<1%
BR	Bi $\rightarrow$ right	18	2%

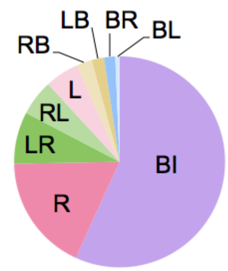


Figure 6.11: Frequencies of each strategy.

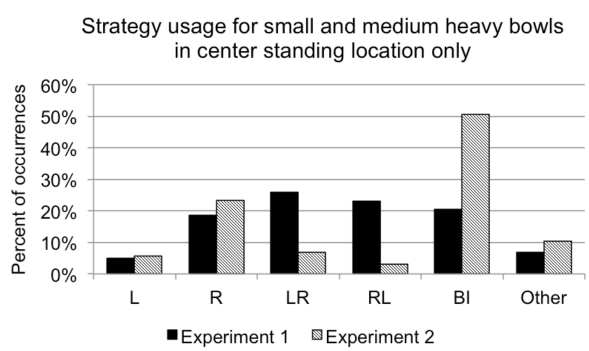


Figure 6.12: Comparison between Experiment 1 and 2 on identical trials.

6.4.2 Effect of experimental variables on grasp strategy

The models for all three strategies included significant effects of size and balance, as well as the size × balance interaction for the bimanual and unimanual strategies. In addition, the model for bimanual strategy also had a main effect of weight. These effects are summarized visually in Fig. 6.13.

In the model for bimanual usage, the effects remaining were size ($\chi^2(3) = 14.1$, $p = .003$), weight ($\chi^2(2) = 21.4$, $p < .0001$), balance ($\chi^2(1) = 251$, $p < .0001$), and the size × balance interaction effect ($\chi^2(3) = 14.6$, $p = .002$). Fig. 6.13 indicates that heavier weights increase bimanual usage slightly. It also indicates that bimanual usage is nearly maxed out in the balance condition, while, in the no-balance condition, small bowls are markedly likely to be handled with two hands, more so than larger bowls. However, beyond that point, increasing bowl size pushes people to use the bimanual strategy more often.

For the hand-off strategy, the three effects remaining in the model were size ($\chi^2(3) = 49.5$, $p < .0001$), balance ($\chi^2(1) = 104$, $p < .0001$), and configuration ($\chi^2(2) = 84.9$, $p < .0001$; Fig. 6.14) main effects. The hand-off strategy is less often used at the smallest bowl size (Fig. 6.13). The balance and configuration effects are similar to those found in Experiment 1: balance cuts down hand-off usage, and hand-offs are used more frequently to transport left-to-right or vice versa than front-to-back.

For unimanual usage, the effects that remained in the model were the main effects of balance ($\chi^2(1) = 44.5$, $p < .0001$) and configuration ($\chi^2(2) = 29.2$, $p < .0001$) as well as the size × balance interaction ($\chi^2(3) = 17.9$, $p = .0005$). The main effect of size was not significant ($\chi^2(3) = 4.39$, $p = .22$). Unlike in Experiment 1 where the effect of balance depended on the starting configuration, in Experiment 2 the balance condition cut down unimanual usage in all configurations. The configuration effect (Fig. 6.14) was similar to Experiment 1, with most unimanual usage when moving the bowl front to back (C4). The size × balance interaction (Fig. 6.13) shows that unimanual usage declines as bowl size increases for the no-balance case only.

The bimanual strategy was the only strategy that had a weight effect. Weight and its interaction effects were removed from the full hand-off ($\chi^2(42) = 41.4$, $p = .50$) and unimanual ($\chi^2(42) = 42.3$, $p = .46$) models. Un-

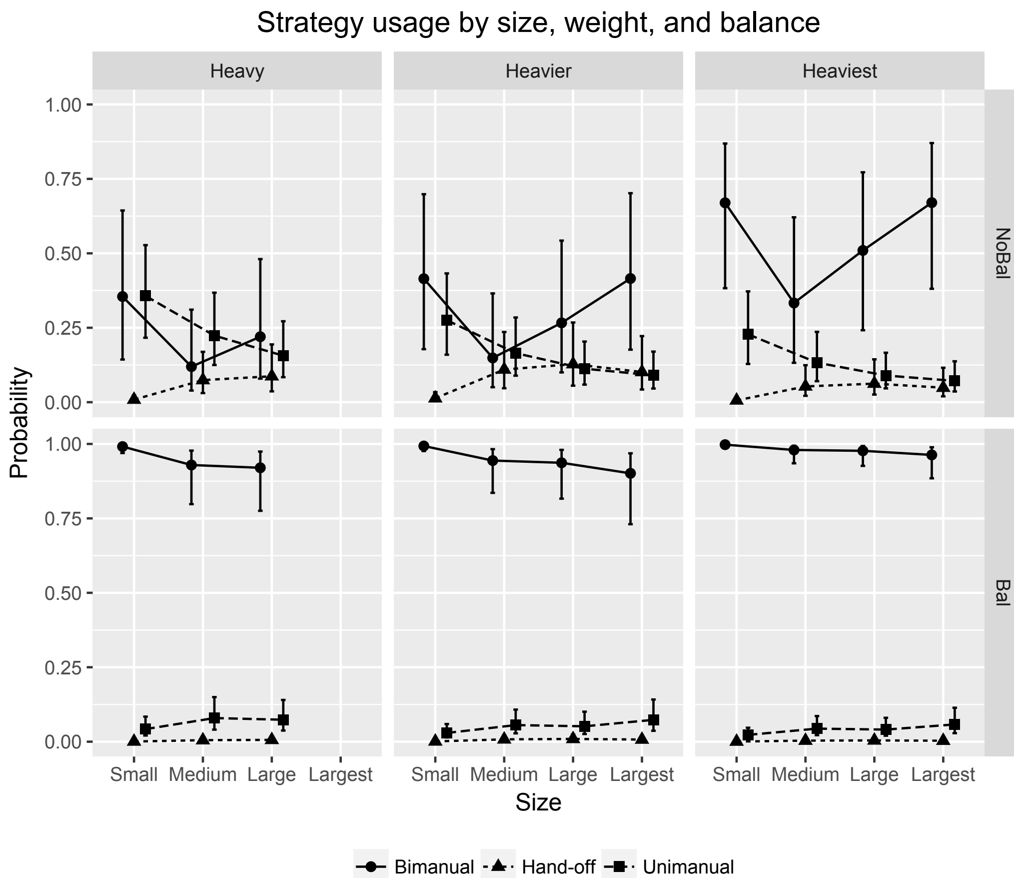


Figure 6.13: Effects of size, weight, and balance on the usage of the three strategies. Size, weight, balance, and the size × balance interaction effect are significant for bimanual usage; size and balance main effects are significant for hand-off usage, and the balance main effect and size × balance interaction effect are significant for unimanual usage.

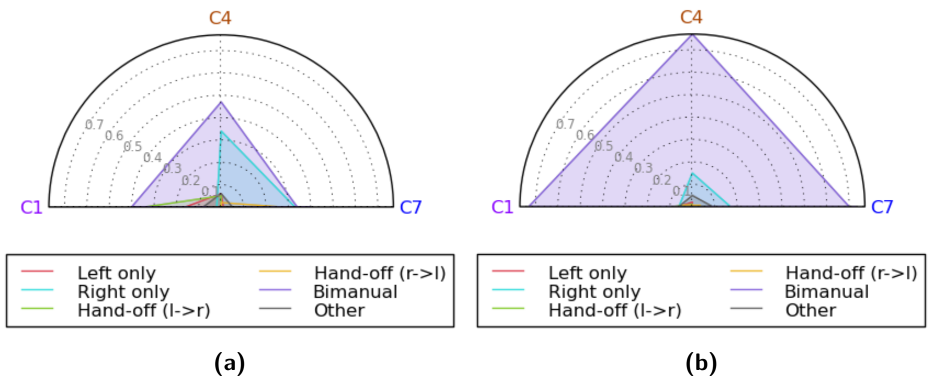


Figure 6.14: Angle plot showing strategy usage at all configurations for (a) no-balance cases and (b) balance cases.

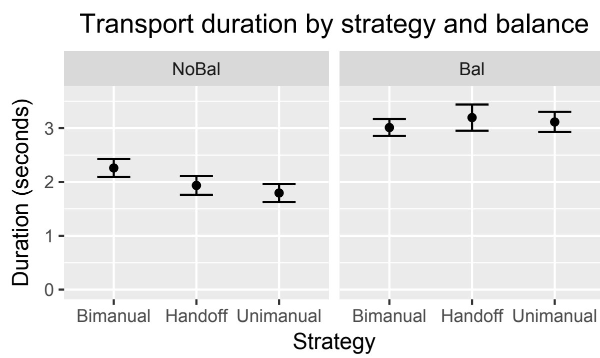


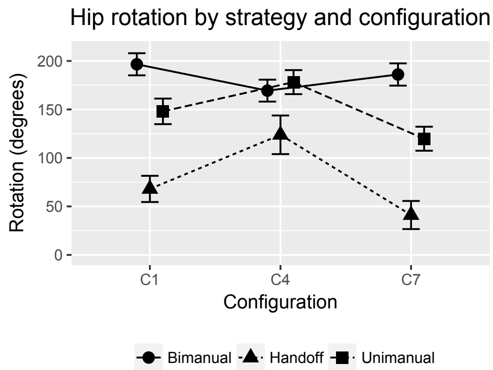
Figure 6.15: Transport duration by strategy and balance.

like the other two strategies, configuration was able to removed from the bimanual model ($\chi^2(44) = 59.5$, p = .059).

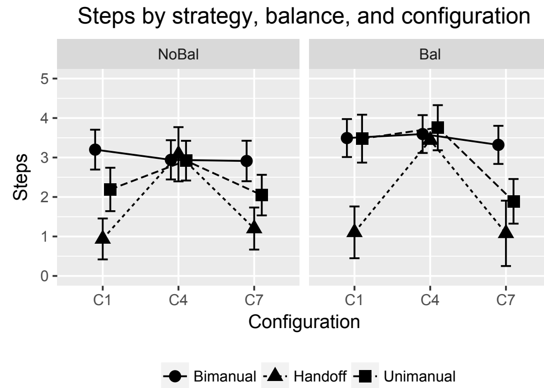
6.4.3 Effect of strategy on duration, rotation, and step count

For duration, the results in Experiment 2 match the first experiment closely. Both strategy ($\chi^2(2) = 24.1$, $p < .0001$) and strategy $\times$ balance ($\chi^2(2) = 66.0$, $p < .0001$) were significant, with the bimanual strategy taking longer in the no-balance case but competitive in the balance case (Fig. 6.15).

For rotation, similar to Experiment 1, both strategy ($\chi^2(2) = 474$, $p < .0001$), and the strategy $\times$ configuration interaction ($\chi^2(2) = 66.0$, $p < .0001$; Fig. 6.16a) were significant. In addition, the strategy $\times$ size ($\chi^2(6) = 21.5$, p =



(a)



(b)

Figure 6.16: Total hip rotation by strategy and configuration compared to step count by strategy, balance, and configuration.

.002) and strategy × balance ($\chi^2(2) = 16.0$, $p < .001$) interactions were also significant. Although the mean rotations for unimanual and hand-off strategies were slightly higher using the motion capture in Experiment 2, the interaction effect is similar to the first experiment (compare with C1, C4, and C7 in Figure 5). The main exception is that at C4, the unimanual strategy requires more rotation than the bimanual strategy.

For step count, the strategy ($\chi^2(2) = 222$, $p < .001$), strategy × size ($\chi^2(6) = 21.1$, $p = .002$), and strategy × balance × configuration interaction ($\chi^2(4) = 18.3$, $p = .001$; Fig. 6.16b) were significant. The configuration pattern is similar to the rotation results (Fig. 6.16a), except for unimanual at C1.

6.4.4 Correlations between measures

Fig. 6.17 contains information on the correlation between hip rotation and other measures – duration, step count, and other rotation measures. Because duration and the rotation measures were continuous, they were compared using the Pearson's R correlation coefficient. Because steps were discrete, we use a boxplot to compare rotation and steps. Hip rotation and duration have low correlation ($R^2 = .280$). By contrast, there is a moderately strong relationship between hip rotation and step count (Fig. 6.17, right). The correlation between hip rotation and other rotation measures is high indicating hip rotation is acceptable to use as a proxy for other kinds of

Correlation with total hip rotation	R^2
Duration	.280
Transport hip rotation	.932
Total torso rotation	.910
Transport torso rotation	.891
Total head rotation	.567
Transport head rotation	.741

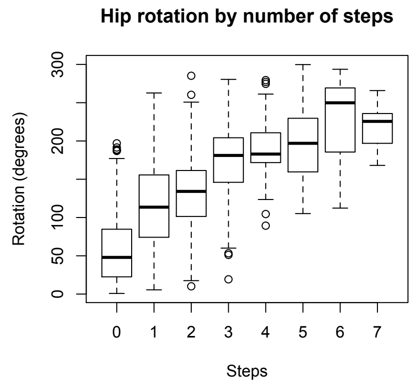


Figure 6.17: The relationship between (left) hip rotation and duration, and hip rotation and other rotation measures (Pearson's R) and (right) hip rotation and step count.

rotation.

6.4.5 Hysteresis analysis

The results of the analysis of the previous trials was that whether the k previous trials had zero or k bimanual uses (for k = 1, 2, 3) did not have a significant effect on whether the current trial would be bimanual in Experiment 1 (k = 1: $\chi^2(1) = 2.46$, p = .117; k = 2: $\chi^2(1) = 1.86$, p = .172; k = 3: $\chi^2(1) = 3.48$, p = .062), but was significant in Experiment 2 for three previous trials (k = 1: $\chi^2(1) = 0.594$, p = .441; k = 2: $\chi^2(1) = 4.45$, p = .035; k = 3: $\chi^2(1) = 9.00$, p = .0027). This difference is mostly likely due to the lower bimanual usage in Experiment 1, which makes the dataset analyzed smaller, as Experiments 1 and 2 have similar trends (Fig. 6.18).

6.5 Discussion

The factors that affect the use of one or two hands in object transport, especially the strategy of handing off between hands, are not well understood. In these experiments, we wanted to investigate the effect of object and task properties on the selection of bimanual, hand-off, and unimanual strategy, and to identify principles that might be underlying this selection.

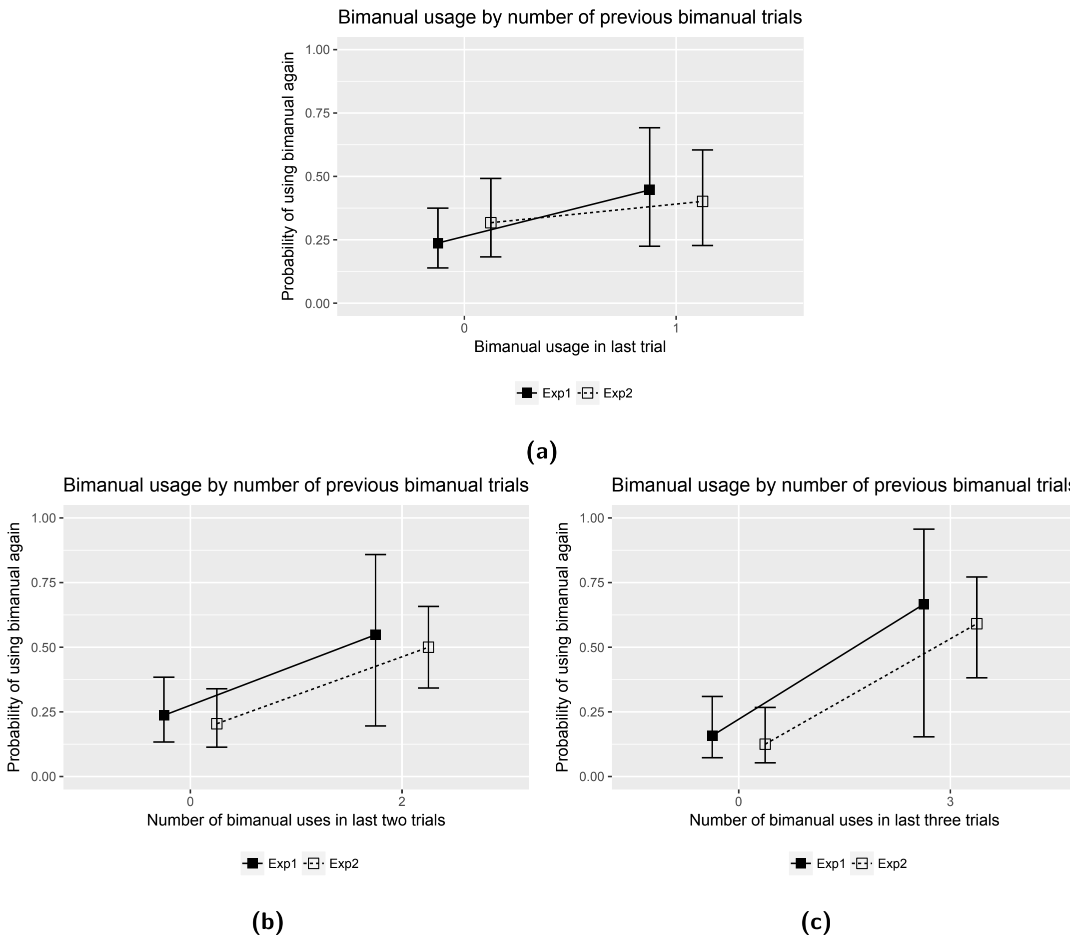


Figure 6.18: Effect of previous bimanual usage on current trial's bimanual usage for Experiments 1 and 2.

Previous work examining grasping [Cesari and Newell, 2000] has found that increasing the size of an object will cause people to transition from one-handed to two-handed grasping. In our work examining transport, we found similar effects of size. However, these effects were weaker than we had expected based on the Cesari and Newell results. One possible explanation for a weak size effect is that the typical grasping location of a bowl does not increase as the bowl gets bigger. Specifically, subjects typically pinched the rim of the bowl (see next chapter for more details), which is similarly thin across bowl sizes. By contrast, in previous studies, the objects manipulated were cubes or toys, where all graspable dimensions increase in width simultaneously. Previous work [Feix et al., 2014] has shown that grasps are overwhelmingly formed around an object's thinnest dimension, so the consistently thin bowl lip may explain the relatively weak effect of size on bimanual usage. Further investigation could clarify the effect of an object's thinnest dimension on choice of transport strategy.

Another unexpected effect of size on strategy is that bimanual usage for the small bowl was particularly high. Part of the strategy choice may be due to the fullness of the bowl, a factor not considered in this study. The fullness of the bowl appeared to make unimanual grasps more difficult: most, but not all, unimanual grasps were formed by pinching the rim of the bowl. These grasps required placing the thumb inside the bowl, which is more difficult to do when the bowl is full. By contrast, most bimanual grasps involved forming multiple contacts around the outside surface of the bowl and thus were not affected by the bowl's fullness. In the future, using fillers of higher density could be used to test whether bowl fullness was influencing choice of strategy.

The Cesari and Newell [2000] study also found a similar effect of weight on causing people to transition from onehanded to two-handed grasping. The effect of weight on transport strategy in our experiments was also weaker than we expected. The absence of a weight effect in Experiment 1 may be because the heavy bowl was not sufficiently heavy to affect people's strategy choices. However, Experiment 2 also contained a weak effect of weight found only in the bimanual strategy. This weak effect may be due to high usage of the bimanual strategy in general (discussed below).

Palmer [1989] found an effect of weight on hand-offs. However, we did not find an effect of greater weight on discouraging hand-offs. It is possible this effect only applies to handing-off as an idle action (as opposed to a transport strategy) or to infant development. Using a finer step size in weights may help to clarify how weight causes people to transition between bimanual, hand-off, and unimanual strategies.

In these experiments, we expected that the use of two hands is a strategy that decreases the difficulty of a balancing task, similar to the effect of pregrasp rotation found by Chang et al. [2009]. Our data strongly support this possibility in two ways. First, use of the bimanual strategy increases when balance is necessary (Figs. 6.4 and 6.13). Second, although the bimanual strategy is slowest when balance is not required, it becomes faster than hand-offs and as fast as the unimanual strategy in the balance case (Fig. 6.6 and 6.15), making it the strategy with the smallest increase in movement time going from no-balance to balance cases.

Previous work [Rosenbaum et al., 2010] has investigated the effect of start and goal location on the use of left, right, and both hands. Similar to their work and other handedness studies (e.g., Gonzalez et al. [2014]), we observed that the left hand is more often used to pick and place on the left side of the body and the right hand on the right side (see Fig. 6.5). Matching Rosenbaum and colleagues' work, we also found that two-handed picking and placing were much less responsive to object start/goal position than one-handed picking and placing.

We also wished to extend Rosenbaum et al.'s work to distinguish between pure unimanual transport and hand-offs. Our findings indicate that hand-offs function as an alternative to the unimanual strategy. Hand-offs and unimanual strategies each dominate at disjoint sets of configurations. We expected that hand-offs would be used when the start and goal are located in different left/right hemispaces. Our findings support this guess, with configurations with this property being dominated by hand-off usage, while hand-off usage is dramatically cut down when this property does not hold (Fig. 6.5). Although we found that hand-offs function as an alternative to the unimanual strategy, we also found that hand-offs seem to be less stable than the unimanual strategy. This is indicated by relatively longer movement times in balance cases (Figs. 6.6 and 6.15) and being disfavored compared to the unimanual strategy in balance cases (Fig. 6.4).

The second major question we investigated was the underlying reason behind choices of transport strategy. Minimization has been a guiding principle when trying to explain motion choices Engelbrecht [2001], and our results support that minimal principles may be useful for explaining selection of transport strategy. Specifically, our results indicate that the desire to minimize body rotation is likely underlying people's choices of transport strategy. First, the large amount of body rotation necessary for bimanual transport could explain why the seemingly less stable hand-off and unimanual strategies were widely used even in the balance cases of Experiment 1 (see Fig. 6.5b). Second, the usage of hand-offs corresponds closely to configurations where less rotation is performed. Our results also indicate that other measures of rotation and step count are strongly tied to hip rotation.

Although we did not investigate it in this study, it is also worth asking why these different strategies entail different amounts of body rotation. One possible reason why people rotate different amounts is that the reach of the arms changes with reaching angle, so rotation may be used to change reaching length, including equalizing the reaching length of both hands for a bimanual grasp. Factors such as different comfort and lifting ability at different points within a joint's range of motion [Chang et al., 2009] may also be important. More work is needed to determine in detail the biomechanical considerations underlying the different amount rotated for each strategy in each configuration.

Comparing the two experiments, we see that the less precise methodology of the first experiment nevertheless yielded similar results to the motion capture technology used in the second. We also found that the second of our experiments had a significantly larger amount of bimanual strategy usage than the first experiment, even when comparing identical trials (unchanged bowl size and weight). Our hysteresis analysis indicates it is possible that previous trials affect the strategy choice in the next trial, meaning that the different bowl sizes and weights used in Experiment 2 could have affected strategy usage on the shared bowl sizes and weights. Another possibility is that the act of wearing motion capture markers could make people more self-conscious about their motions and affect their strategy choices. A third possibility is that changing the starting configuration frequently as in Experiment 2 and not in Experiment 1 may have encouraged people to use a single transport strategy (the bimanual strategy) by default rather than adapting their strategy to the starting configuration. Further investigation is needed.

This is a preliminary study with a small number of participants. Therefore, the results should be interpreted conservatively. However, overall, our work indicates that the choice of hands in transport is highly responsive to task demands. In this work, we focused on the hand-off strategy, finding that it is similar to the unimanual strategy, but is less stable. It is mainly used in configurations that involve transporting an object between left and right hemispaces, where it reduces the amount of body rotation needed to complete the transport task. For bimanual transport, we found that using two hands is a strategy that can be employed to reduce the difficulty of maintaining an object's balance, similar to pregrasp rotation. However, it requires more body rotation and effort. The selection of bimanual, hand-off, and unimanual transport strategy appears to balance these considerations of stability and effort.

7

Poses in a bowl transport task

In the experiments detailed in the previous chapter, we also collected grasp pose information. There are many possible grasp poses that people can use to grab bowls. In this chapter, we propose a classification scheme for these poses, connect the poses to existing grasp taxonomies, and report how the pose usage was affected by the factors of the experiment (size, weight, balance, and position).

7.1 Pose classification

Grasp poses were sorted into eight different possibilities based on contact areas and fingers used: five unimanual grasps (Table 7.1) and three bimanual grasps (Table 7.2). Poses seemed to be a somewhat continuous space making differentiating between certain pairs of grasps (lateral vs. extension and sometimes open vs. cupped bimanual) ambiguous. If there were at least two finger pads used in the grasp, the trial was annotated as extension rather than lateral pinch. If it appeared the finger pads were being used more than the pad of the palm, the grasp was annotated as open bimanual rather than cupped bimanual.

Table 7.3 shows variations of the eight grasps that were encountered.

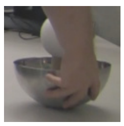
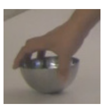
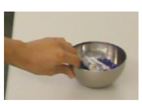
Grasp	Description	Comparison
	**Extension** – Pads of 2+ fingers are the main thing holding the bowl. Thumb hooked on inside of bowl. Fingers curve around bowl, creating a fan shape. Related to the open bimanual grasp.	 Extension
	**Lateral pinch** – Side of middle finger is the main thing holding the bowl. Thumb hooked on inside of bowl. Index finger extended away from other fingers. Alternatively, the bowl rim can be pinched between thumb and side of index finger. Can also be used bimanually.	 Lateral
	**Overhand (precision disk)** – Grab bowl from the top. Uses pads of thumb and all fingers.	 Precision disk
	**Closed lateral pinch** – Thumb resting on top of bowl, not hooked on the inside. Support comes from thumb and pad/side of index and middle fingers. All fingers are curled.	Unknown
	**Index hook** – Collection of several distinct but similar poses where the index finger is hooked on the inside of bowl. Bowl held by pinching between index finger and other fingers.	 Palmar pinch, Middle-over-index grasp

Table 7.1: Codes and descriptions for unimanual bowl grasp poses. Comparisons are from Feix et al. [2009] and Liu et al. [2014b] / Chapter 3.

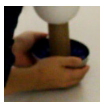
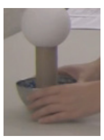
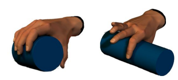
Grasp	Description	Comparison
	**Open bimanual** – Finger pads are the main things holding the bowl. Thumb may be lightly hooked on the bowl rim, not in contact, on the outside of the bowl, or fully inside the bowl. (It is not essential to the grasp in any case.) Hooking the thumb and releasing with the other hand turns this into the extension grasp.	 Parallel ext.
	**Cupped** – Palm is the main thing holding the bowl. Fingers cupped and horizontal (cf. open bimanual pose's vertical direction). Thumbs do not oppose fingers and may be resting on bowl rim or floating. They are not essential to the grasp. Can also be used unimanually.	 Flat hand cupping
	**Ring** – The two thumbs and fingers form a ring around the bowl. Other fingers free-floating or used as optional extra support. Can also be used unimanually.	 Large diameter, Ring

Table 7.2: Codes and descriptions for bimanual bowl grasp poses. Comparisons are from Feix et al. [2009] and Liu et al. [2014b] / Chapter 3.

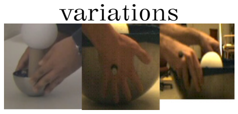
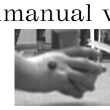
Lateral variations 	The lateral pinch grasp can also be bimanually (both images) or by pinching the bowl between thumb and index finger (right).
Open bimanual variations 	The open bimanual grasp encompasses a wide range of poses. The thumb can be lightly hooked, not in contact (left), on the outside of the bowl (center), or fully inside the bowl (right).
Index hook variation 	Any grasp involving pinching with the index finger is counted as an index hook, including this grasp between index and middle finger (middle-over-index garsp).
Ring bimanual variation 	This grasp can also be used unimanually. This is what the ring looks like when used with one hand.

Table 7.3: Variations on grasp poses.

7.2 Data analysis

Data analysis was done in the same way as in Chapter 6. We used a generalized linear mixed model (GLMM) with logistic link function using R's ***lme4*** package [Bates et al., 2015]. Separate analyses were done for each grasp – whether the grasp was used or not was a binary outcome – restricted to a single strategy at a time. For example, an analysis of the effects of the experiment variables on the usage of the open bimanual pose was done analyzing bimanual trials only, while two analyses for the extension grasp were performed: one for hand-offs only and one for unimanual trials only. Analysis was done only for grasps that had at least 20 uses within a particular strategy. A stepwise procedure comparing likelihood ratios was used to eliminate non-significant variables until no more could be removed (using a significance level of .001). For some grasps, scaling issues in the data made fitting some models difficult. In those cases, smaller models were used, and forward selection was used to see if adding back factors was significant. For effects remaining in the model, plots showing the mean probability of a strategy being used under each condition and an estimation of the standard error of that mean were generated using the effect function of R's ***effects*** package [Fox, 2003]. Note that if there is an infinite or missing error bar, that means that the grasp was never observed in that condition.

7.3 Results

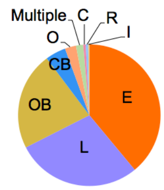
7.3.1 Overall pose frequencies

The frequency of the grasp poses for Experiment 1 and 2 are summarized in Figs. 7.1 and 7.2. "Changing grasps"/"multiple" indicates that one hand used multiple grasp poses during the transport, most often used for mixed strategies that add or remove a hand during transport, changing a bimanual grasp into a unimanual one, or vice versa. Grasp changes can also occur occasionally in other transport strategies by regrasping the bowl.

The extension grasp and open bimanual grasp are the most frequently used unimanual and bimanual grasps. The most rare grasp is the index hook grasps, which only appeared four times in the two experiments.

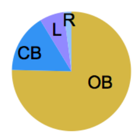
All grasps

	Pose	Frequency	
E	Extension	569	39%
L	Lateral	416	28%
OB	Open bimanual	327	22%
CB	Cupped	68	5%
O	Overhand	37	3%
	Changing grasps	24	2%
C	Closed lateral	8	<1%
R	Ring	8	<1%
I	Index hook	4	<1%



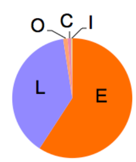
Bimanual grasps

	Pose	Frequency	
OB	Open bimanual	294	75%
CB	Cupped	62	16%
L	Lateral	26	7%
R	Ring	8	2%



Hand-off grasps

	Pose	Frequency	
E	Extension	387	59%
L	Lateral	251	38%
O	Overhand	8	1%
C	Closed lateral	6	<1%
I	Index hook	2	<1%



Unimanual grasps

	Pose	Frequency	
E	Extension	170	51%
L	Lateral	128	39%
O	Overhand	29	9%
I	Index hook	2	<1%
C	Closed lateral	2	<1%

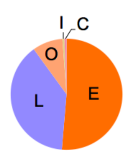
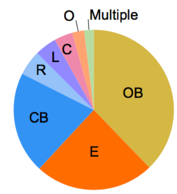


Figure 7.1: Frequencies of grasp poses in Experiment 1.

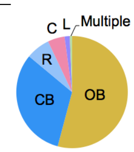
All grasps

	Pose	Frequency	
OB	Open bimanual	706	38%
E	Extension	450	24%
CB	Cupped	384	21%
R	Ring	95	5%
L	Lateral	82	4%
C	Closed lateral	70	4%
O	Overhand	45	2%
	Changing grasps	36	2%



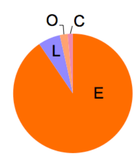
Bimanual grasps

	Pose	Frequency	
OB	Open bimanual	650	54%
CB	Cupped	382	32%
R	Ring	83	7%
C	Closed lateral	59	5%
L	Lateral	18	2%
	Changing grasps	8	<1%



Hand-off grasps

	Pose	Frequency	
E	Extension	257	90%
L	Lateral	17	6%
O	Overhand	6	2%
C	Closed lateral	4	1%



Unimanual grasps

	Pose	Frequency	
E	Extension	166	68%
O	Overhand	39	16%
L	Lateral	24	10%
R	Ring (unimanual)	8	3%
C	Closed lateral	6	2%
	Changing grasps	1	<1%

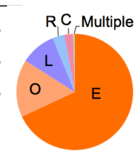


Figure 7.2: Frequencies of grasp poses in Experiment 2

7.3.2 Effect of experimental variables on grasp pose

For Experiment 1's bimanual trials, we analyzed the open, cupped, and lateral bimanual grasps. For the hand-off trials, we analyzed the extension and lateral grasps. For the unimanual trials, we analyzed the extension, lateral, and overhand grasps. For the bimanual cupped and overhand unimanual grasps, some combinations of variables had no covariance and so a model could not be fit. In these cases, we used smaller models and forward selection to determine whether or not variables could be removed from the model.

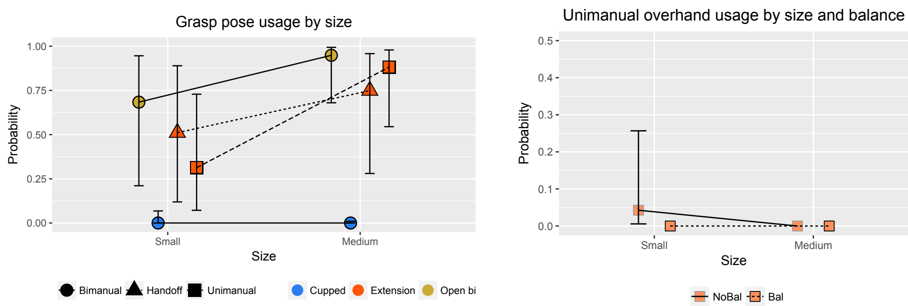
For the open and cupped bimanual grasps and the extension grasp, only size remained in the model (open bi: $\chi^2(1) = 30.6$; $p < .0001$; cupped: $\chi^2(1) = 14.0$; $p = .0002$; extension (hand-off): $\chi^2(1) = 12.6$; $p = .0004$; extension (unimanual): $\chi^2(1) = 49.4$; $p < .0001$). Fig. 7.3a shows these effects. The open bimanual and extension grasp are used more frequently for the medium bowl than for the small bowl, while the usage for the cupped grasp goes down for the medium bowl.

For the unimanual overhand grasp, both size ($\chi^2(1) = 47.7$; $p < .0001$) and balance ($\chi^2(1) = 30.8$; $p < .0001$) were significant. As Fig. 7.3b shows, the overhand grasp is used only in the small no-balance case (missing or infinite error bars indicate that the grasp was never observed in that condition)

For the lateral grasp, in both the hand-off and unimanual cases, only weight remained in the model (hand-off: $\chi^2(1) = 12.1$; $p = .0005$; unimanual: $\chi^2(1) = 12.0$; $p = .0005$). Lateral grasp usage went down in the heavy condition (Fig. 7.4).

For Experiment 2, we fit a model to four bimanual poses (open bimanual, cupped, ring, and closed lateral), one hand-off poses (extension), and three unimanual poses (extension, overhand, and lateral). However, the unimanual lateral grasp had no factors that remained in the model.

Fig. 7.5 shows the significant effects for bimanual grasps. For the open bimanual grasp, size ($\chi^2(3) = 177$; $p < .0001$), balance ($\chi^2(1) = 40.4$; $p < .0001$), and size $\times$ balance ($\chi^2(3) = 19.1$; $p = .0003$) remained in the model. For the cupped grasp, only balance remained in the model ($\chi^2(1) = 18.8$; $p < .0001$). For the ring grasp, both size ($\chi^2(3) = 140$; $p < .0001$) and balance ($\chi^2(1) = 29.4$; $p < .0001$) were significant. In particular, the ring grasp was



(a) **(b)**

Figure 7.3: Significant size and balance effects.

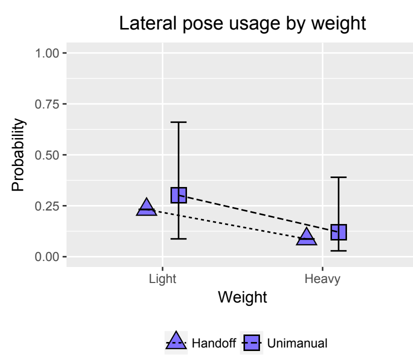


Figure 7.4: Significant weight effects for the lateral grasp.

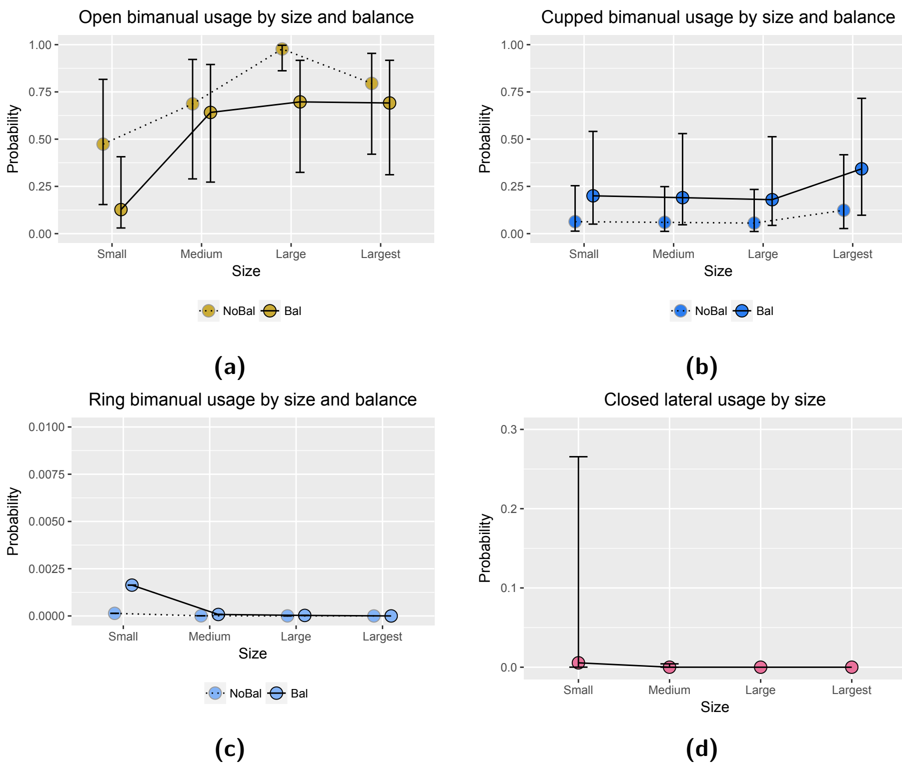
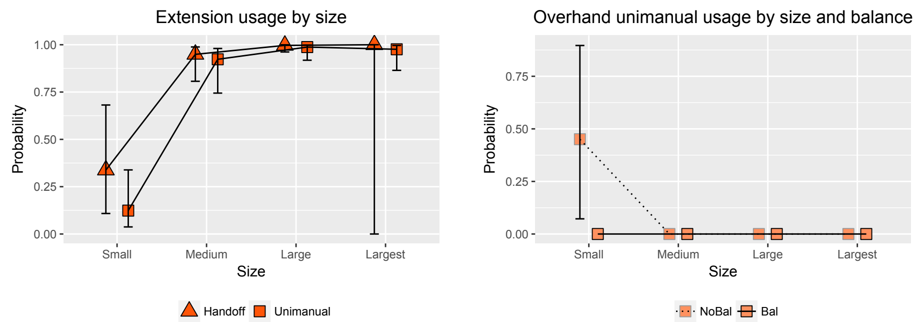


Figure 7.5: Significant effects for bimanual poses: (a) open bimanual grasp (size $\times$ balance), (b) cupped grasp (balance), (c) ring (size + balance), (d) closed lateral (size).

mainly used on small balance bowls. Finally, size was significant for the closed lateral grasp ($\chi^2(3) = 174$; $p < .0001$), with usage being highest in the small bowl case and dropping off.

Fig. 7.6 shows the significant effects for one-handed grasps used in hand-offs and unimanual transport. For the extension grasp, only size remained in the model for both hand-off cases ($\chi^2(3) = 73.8$; $p < .0001$) and unimanual cases ($\chi^2(3) = 156$; $p < .0001$). For the overhand grasp, in the unimanual case, both size ($\chi^2(3) = 109$; $p < .0001$) and balance ($\chi^2(1) = 36.8$; $p < .0001$) remained.

Throughout both Experiment 1 and 2, the overhand grasp was only used on small no-balance bowls. Although the index hook grasp was used too rarely to analyze, it only appeared for light bowls in Experiment 1 and did not appear in Experiment 2.



(a) **(b)**

Figure 7.6: Significant effects for hand-off/unimanual poses: (a) extension grasp (size), (b) overhand grasp (unimanual: size + balance).

7.4 Discussion

We performed an experiment featuring simple, lipless bowls of four different sizes and four different weights. In some trials, the bowl contained a balance tube, and the starting and ending goal varied. We recorded the poses that were used, categorized them, and noted variations and poses that fall in between different categories.

A surprising result is the amount of variety shown in the poses used to grasp the bowls. Rather than there being a single one-handed grasp and a single two-handed grasp, at least seven grasp poses were used in the experiment. Six could be used unimanually (extension, lateral, overhand, closed lateral, index hook, and ring) and five could be used bimanually (open bimanual (extension), cupped, ring, lateral, and closed lateral). Variations of these poses could also be considered to be additional grasps.

Grasp poses are most affected by the size of the bowl. First, the extension and open bimanual grasps are used more often for larger bowls than for the smallest bowl in both Experiments 1 and 2. On the opposite end, there are some poses that seem to only or mostly be used on the smallest bowl size: the overhand, ring, and closed lateral grasps. With the overhand grasp, only the smallest bowl will fit in the hand in this grasp. For the ring grasp, the greater the curvature of the bowl rim, the more comfortable it is to have the thumb opposing the fingers in this grasp. It is less clear why the closed lateral grasp is used more for small bowls, but it could possibly be that the fullness of the small bowls in Experiment 2 that caused participants to avoid hooking the thumb in the bowl, which yields this kind of grasp. Finally, there are poses that seem to be used somewhat consistently across bowl sizes, like the cupped and lateral grasps.

Balance also plays a strong role in determining grasp choice. This can be for two reasons: one is that the grasp can be good or bad at stabilizing the bowl, and the other is that the balance tube itself is a physical obstacle that encourages alternate strategies. In the case of the overhand grasp, this grasp was only ever observed on small no-balance bowls in both experiments because it is only possible to make this grasp on a small bowl if there's nothing sticking out of the top. By contrast, the cupped and ring grasps are *more* frequently observed for balance bowls. Because the cupped grasp is more frequently used in balance cases on all sizes of bowls including the larger bowls where the balance tube does not interfere much with forming a grasp around the edge of the bowl, it is possible that the cupped grasp might be more stable compared to the open bimanual grasp. For the ring grasp, however, it's unclear whether greater stability or the physical presence of the balance tube causes this grasp to be preferred.

Weight only played a role in the lateral grasp and possibly the index hook grasps. In Experiment 1, lateral usage went down in the heavy condition. Consistent with this, Experiment 2, which featured the three heaviest weight conditions only, featured lower lateral grasp usage than Experiment 1. The index hook grasps were only used in the light condition in Experiment 1. They were used too rarely for a proper analysis, but in other work, I have observed people using the awkward and fairly weak middle-over-index grasp to grab very light objects (cloths) that were far away and required stretching in order to reach. Though weak, this grasp was capable of extending the reach of the arm because it can be done with both the hand and fingers fully extended. In any case, the lateral grasp and index hook grasps are pinches rather than wrap grasps and so may have less ability to lift heavy objects than other grasps.

Finally, configuration was able to be removed from the models of all grasps in both experiments. This suggests that object location does not play a strong role in selecting type of grasp.

These analyses are done on fairly small slices of a small dataset, and so the results reported here should be interpreted conservatively. However, the results of this analysis suggest that at different sizes, different parts of the bowl become salient for forming a grasp. That is, different parts of the bowl's geometry register as affordances. At the smallest size, the rim of the bowl offers the same affordance for a precision disk grasp that a CD might – but only if there is nothing sticking out of the bowl. Alternatively, the rim of the bowl can resemble a large cylinder, allowing for a large cylinder or ring grasp. The thinness of the bowl at all sizes allows it to be pinched like a card, although a pinch grasp may be fatiguing when the bowl is heavy. The most common grasp used, however, is one that uses finger pad contacts on the outside of the bowl either bimanually or unimanually by hooking the thumb inside the bowl.

In Chapter 3, we assumed that one task generally corresponded to one hand pose, and only briefly acknowledged the role that object-related factors play in grasping (Section 3.2.5). It is possible that bowls of different sizes are especially flexible in the ways they can be grasped. However, the work in this chapter gives some indication of the diversity of grasp poses that can be used for a simple task, what affordances they exploit, and the properties (e.g. strength) of different grasps that might make them more or less preferred.

8

Dexterous grasping with soft robot hands

Chapter 4 demonstrated that the process of grasping is complex. A complex grasping process is necessary when the desired final grasp is infeasible in the object's initial pose, thus requiring a grasp adjustment to be made after an initial grasp is established. For example, for a phone lying on a table, the secure grasp people intend to use (and do end up using) involves making contacts on the back of the phone, which is inaccessible when it is lying on the table (Fig. 8.1).

These kinds of within-hand grasp adjustments are common, both to form a stable grasp in the first place (non-prehensile manipulation) and to manipulate the object in hand after forming a stable grasp (prehensile manipulation). Examining the non-prehensile movements featured in Tables 4.1 and 4.2, we find that many of them involve manipulating an object so that it can be brought into the palm. By doing so, a fairly precarious precision grasp can be strengthened into a power grasp. We focus on these motions capable of turning precision grasps into power grasps. Many of these motions require at least three fingers: two to grasp the object stably enough to move or manipulate it, and one more to take advantage of the object's new configuration to create a new grasp capable of putting the object further into the hand/palm.

Of these motions, three of them can be accomplished by three fingers

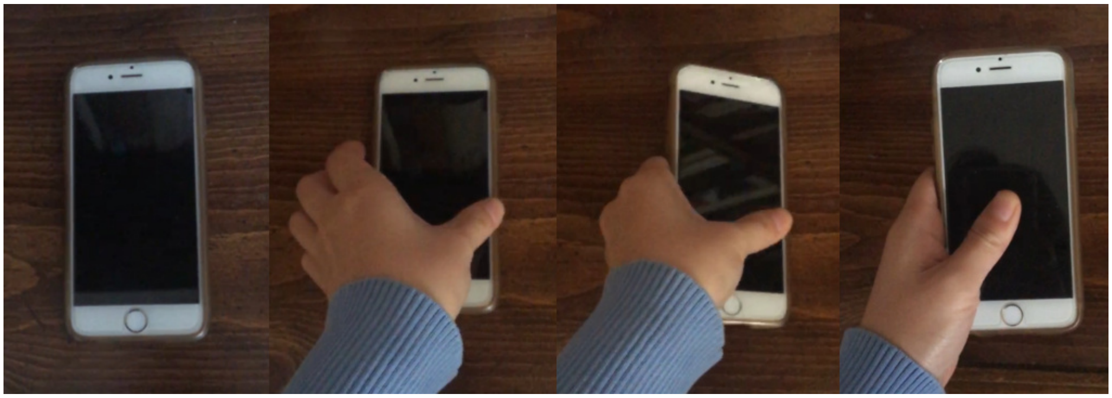


Figure 8.1: When a phone is lying flat on a table, its bottom surface is blocked (far left). However, contacts with this bottom surface are needed in the desired grasp (far right). To achieve this grasp, intermediate levering up is performed.

and in a kind of extruded 2-D space, or 2.5-D space (Fig. 8.2). The first is a full roll that levers up an object. A third finger located behind the left finger (into the page) would be able to take advantage of the exposed underside to form a contact. The second motion is a squeeze that pulls the object further into the palm. Again, a third finger located behind either finger would be able to recontact the object in its new configuration. The third motion tilts the object by flexing a finger. This motion is mostly used in the presence of clutter to singulate an object from nearby objects and create enough space for a third finger (located behind the flexing finger) to contact the side of the object.

In this work, we sought to create a robot hand capable of performing these precision-to-power transitions. In order to do so, we used an existing system capable of designing soft, tendon-driven robots [Bern et al., 2017] to design a tendon network potentially able to perform all three motions. When manipulating an object with a tendon-driven hand, the tendons are responsible not only for posing the fingers but also applying forces at the fingertips. We therefore use an extension of the Soft IK system that goes beyond simple posing and incorporates the ability to specify applied forces. The poses and forces needed also change over time, so multiple poses for each motion are needed.

Our process of designing a 2.5-D two-finger hand is summarized in Fig. 8.3. First, we divide the three motions into keyframes and calculate forces that need to be applied at the fingertips in each frame. Then we use a Soft IK poser with external forces to make sure that the tendon contractions ex-

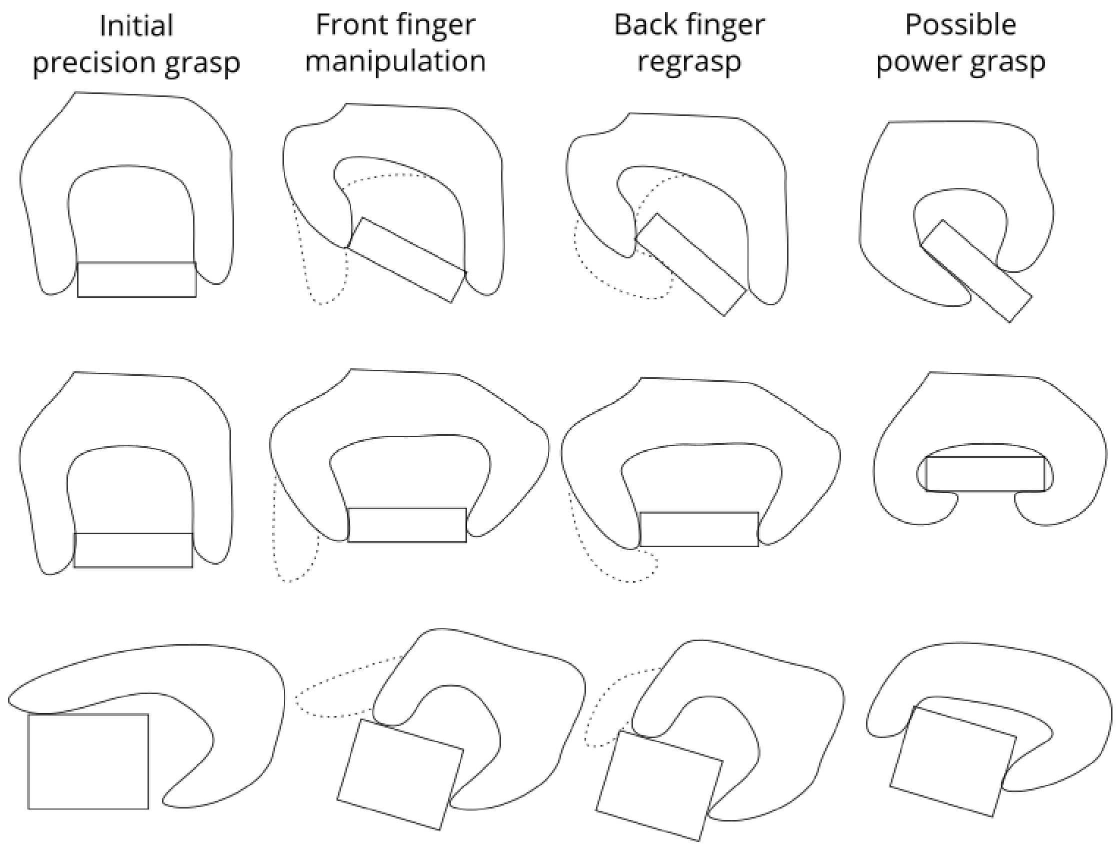


Figure 8.2: Two-finger two-dimensional intrinsic manipulation motions (columns 1-2) observed in natural grasping capable of turning a precision grasp (column 1) into a power grasp (column 4): a full roll (used to lever up objects), a squeeze (used to bring objects closer to the palm), and a squeeze from a flat hand (to tilt object and bring it into the palm). A third finger located behind the manipulating finger is capable of taking advantage of the object's new configuration to form a stronger grasp.

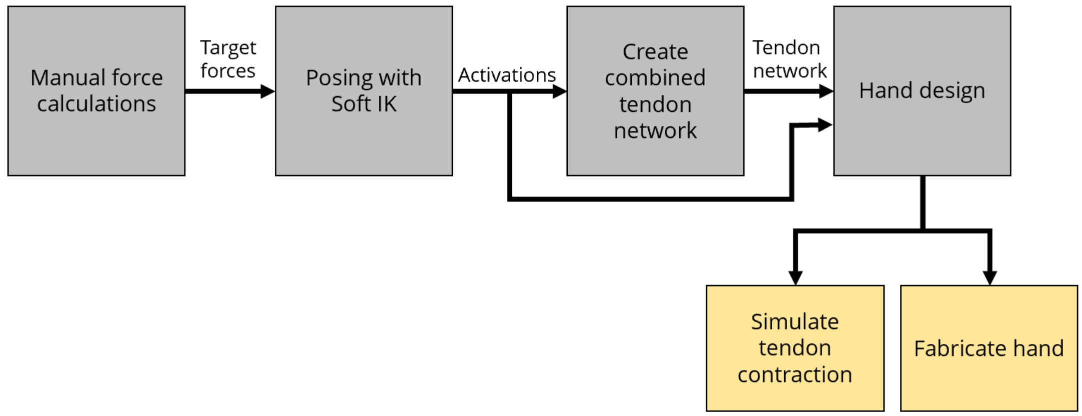


Figure 8.3: Design process

ert a force at the fingertips. We use the tendon contractions that are found for each motion and keyframe in order to create a combined tendon network. This hand design is then tested in simulation to see how well it can perform the desired motions.

8.1 Background

Our work is based on Bern et al.'s Soft IK system [2017]. This system allows users to input a rest pose (undeformed mesh) and an activated pose (target pose $\boldsymbol{x}'$) for a soft planar robot. The system will then solve for a network of tendons within the soft body that, when pulled, will best approximate the activated pose. The system works by finding a solution to an optimization where the objective function being minimized is deviation from the desired pose (Eq. 8.1).

$$O = \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}')^T \boldsymbol{Q}(\boldsymbol{x} - \boldsymbol{x}') + R \tag{8.1}$$

In the above equation, $\boldsymbol{Q}$ is a matrix that filters out unimportant node positions. The variables the system optimizes over are the actuations of contractile elements ($\boldsymbol{\alpha^c}$) within the body of the robot. Finally, R represents the regularization terms:

$$R = w_0 U(\alpha^c) + w_1|\alpha^c - \alpha^0|^2 + w_2|\alpha^c|^2 \tag{8.2}$$

We punish (1) the strain energy U of all tendons (a piecewise C^2 polynomial that is non-negative for all values of α^c), (2) the magnitude of the deformation of the tendon (the deviation of current tendon length from what is expected given the amount of contraction), and (3) the magnitude of contractions themselves.

The Soft IK system makes several assumptions that allows the gradient of O with respect to $\boldsymbol{\alpha^c}$ ($\frac{\delta O}{\delta \alpha^c}$) to be quickly calculated, allowing the optimization to be performed in real time and to be responsive to user input. These assumptions include (1) modeling the strain energy for contractile elements as a smooth (C^2 continuous) function with respect to the amount of deformation, to make calculating derivatives easier; (2) a quasi-static assumption that the forces have come to an equilibrium, which means that forces are not changing with respect to the tension of the contractile elements ($\boldsymbol{\tau}$). This means that the following equation holds:

$$\frac{d\boldsymbol{F}}{d\boldsymbol{\tau}} = \frac{\delta \boldsymbol{F}}{\delta \boldsymbol{\tau}} + \frac{\delta \boldsymbol{F}}{\delta \boldsymbol{x}}\frac{\delta \boldsymbol{x}}{\delta \boldsymbol{\tau}} = 0 \tag{8.3}$$

This equation allows us to calculate $\frac{\delta \boldsymbol{x}}{\delta \boldsymbol{\tau}}$, since we know $\frac{\delta \boldsymbol{F}}{\delta \boldsymbol{\tau}}$ (can be calculated based on the direction of the contractile element) and $\frac{\delta \boldsymbol{F}}{\delta \boldsymbol{x}}$ (can be calculated by finding the Hessian of the deformation energy of the mesh with respect to the pose). This deformation energy includes (1) the energy stored in the mesh modeled as finite elements, due to stretching or shearing, (2) the strain energy of the contractile elements, (3) the energy of pins, which are stiff springs constraining the location of certain points on the mesh, and (4) energy imparted on the object by the application of any external forces. The partial derivative of position $\boldsymbol{x}$ with respect to tensions $\boldsymbol{\tau}$ can then be used to calculate the gradient of the objective function with respect to actuation:

$$\frac{\delta O}{\delta \boldsymbol{\alpha^c}} = \frac{\delta O}{\delta \boldsymbol{x}}\frac{\delta \boldsymbol{x}}{\delta \boldsymbol{\tau}}\frac{\delta \boldsymbol{\tau}}{\delta \boldsymbol{\alpha^c}} \tag{8.4}$$

After solving for a set of actuations that best approximates the desired pose, a tendon network can be selected by choosing high-activation nodes of the mesh and joining them, pruning the nodes selected if needed to create a simpler tendon network.

8.2 Method

We make use of the simulation-based Soft IK system to get an idea of where tendons should be placed on a soft hand in order to exert a desired force from a desired pose. First, for each of the three motions in Fig. 8.2, we assume a specific object shape, and we separate the motion into two keyframes indicating a different configuration of the object (see Fig. 8.4). We then find forces at the fingertips that could accomplish the desired manipulation. We then run the Soft IK system to pose the hand in the presence of external forces, which are the opposite of the target forces we calculated in the previous step. The results of that system are tendon contractions (the pink segments in Fig. 8.5) that we manually combine into a single system. We go into more detail on each of these steps below.

We use a different object for each of the three motions: a flat rectangle in motion 1 (phone-like: .07 m wide × .008 m high; 150 g), a cube in motion 2 (pen-like: .02 m; 50 g), and a tall rectangle in motion 3 (bottle-like: .067 m wide × .15 m high; 420 g). For all motions, we use the same coefficient of friction between the finger and the object ($\mu = 0.8$). We use a safety margin of 50% for heavier objects and 60% for the lightest object, within the range of safety margins people use [Westling and Johansson, 1984; Hiramatsu et al., 2015]. Finally, we assume a location where certain fingers are contacting the object (within .004 m of the corner for the first motion and within .04 m of the corner for the third motion).

We separate each motion into two keyframes featuring different configurations of the object (see Fig. 8.4). When the object's configuration changes, the target pose of the hand and the forces that should be applied also change. For the first motion (lever), the first keyframe features the object lying flat on the table and the second keyframe features the object rotated around its lower-right corner by 15 degrees. For the second motion (squeeze), the first keyframe features the object on the ground and the second keyframe features the object lifted by 0.02 m. In the third motion (tilt out), the first keyframe features the object resting on the table, and the second keyframe features the object rotated around its lower-right corner by 7 degrees.

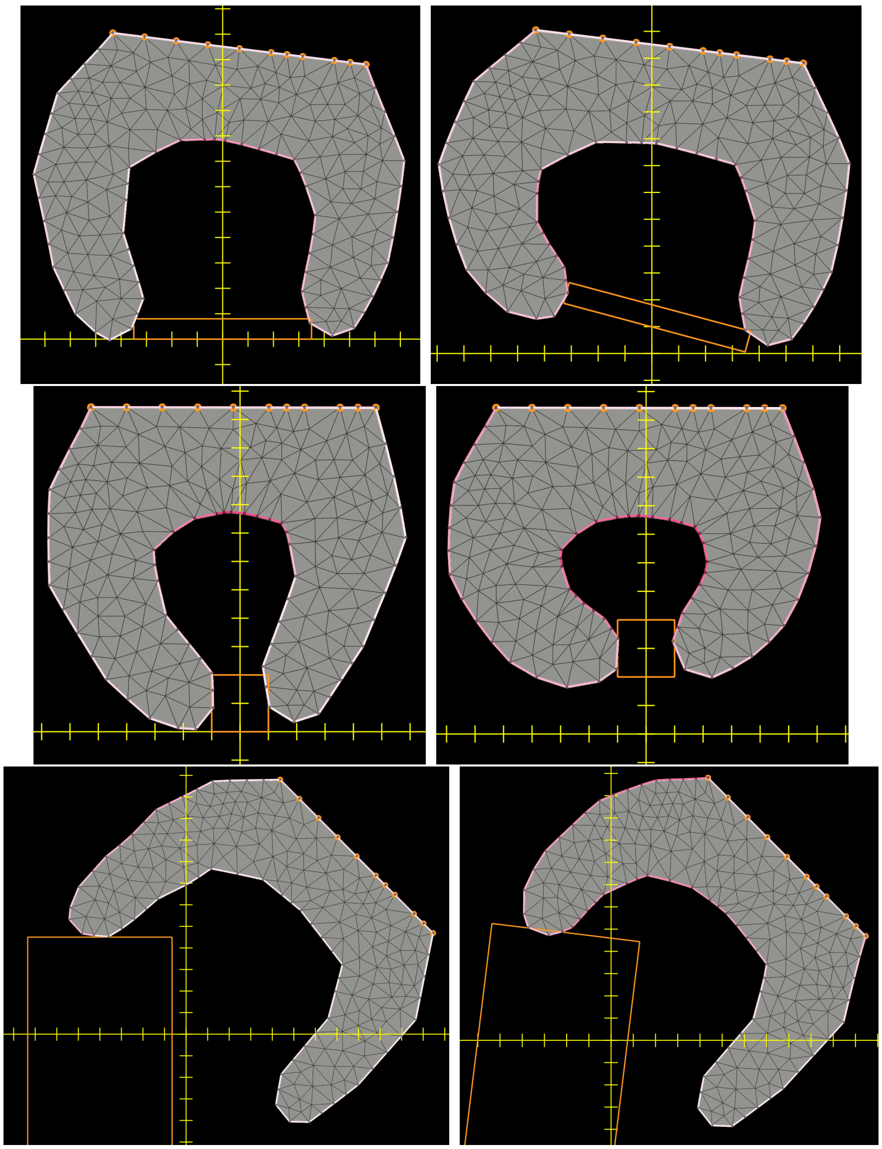


Figure 8.4: Keyframes for all three motions.

Motion	Frame	Finger	Thumb
1: Lever	1	(2.25, 1.2)	(-2.25, -4.22)
	2	(2.48, 0.58)	(-2.48, -2.65)
2: Squeeze	1	(1, 0.5)	(-1, 0.5)
	2	(1, 0.5)	(-1, 0.5)
3: Tilt out	1	(2.1, -4.0)	none
	2	(1.6, -4.23)	none

Table 8.1: Target forces in newtons calculated from the object properties given in the text. The positive x-axis is to the right and the positive y-axis is up.

From these assumptions, it's possible to calculate forces at the fingers capable of manipulating the object in the desired way (pivoting around a pivot point in the first and third motions, lifting an object in the second motion), including a margin of safety above the minimum necessary force. The details of this manual calculation are contained in Appendix A. The calculated forces for both frames of all three motions are summarized in Table 8.1.

These calculated forces are then used as external forces within the Soft IK system. A force of the opposite direction is imposed on a node in the mesh and the solver tries to find activations that move the node to the desired position while resisting the external force. This Soft IK posing with external forces is done for each keyframe of the motion. For some motions, the motion was not feasible without reorienting the hand. For the first motion, we rotated the hand 7° clockwise and for the third motion, we rotated the hand 45° clockwise.

After posing the mesh for each keyframe of each motion, we inspected the active tendons – the parts of the outer surface of the hand that are contracted. We manually combined them into a single tendon system that would be capable of accomplishing all three motions. We then tested the performance of that combined tendon system in a finite element method (FEM) simulation by activating the tendons in the presence of the same external forces we used in the Soft IK system. For each frame, we first started with no external forces and activations similar to Soft IK results that were also created without external forces. From this initial pose – resembling the target pose for this frame but without any force applied – the magnitude of the external force was gradually increased until the full force was being applied. The activations from the initial pose were gradually changed to resemble our Soft IK results in the presence of the full external forces. Small adjustments were then made to the activation levels to resemble the target pose as closely as possible.

After obtaining these activations, we evaluated how close to the goal pose we were able to get. We calculated the difference between the target position used in the Soft IK system and the location of the same nodes after the tendons have been contracted in the simulation.

8.2.1 Hardware evaluation

We used the method of King et al. [2018] to fabricate a pair of two-finger foam robot hands. A mold was 3-D printed, and used to cast two hands made of foam (FlexFoam-iT! X). A custom-knit glove was glued to the hand, and tendons (braided fishing line) were sewn to the glove to match the combined tendon network. The end of these strings were then wrapped around Dynamixel AX-12A motors.

We then posed each hand so as to mirror the manual activation results we obtained from simulation: (1) an initial pose only with no forces (frame 1), (2) the pose resulting from the activations used to achieve the frame 1 pose with external forces, and (3) the pose resulting from the activations used to achieve the frame 2 pose with external forces. By cycling through each pose, interpolating between each pair of poses, the ability of the robot hands to manipulate objects was evaluated.

8.3 Results

Fig. 8.5 shows the results from the Soft IK solver for each of the motions and keyframes. From those results, we created a combined system of five tendons (Fig. 8.6).

Using this combined network, we found activations in order to resemble the desired pose in the presence of external forces. We used the images in Fig. 8.5 in order to initialize values of tendon contractions and adjust them as necessary to approximate the desired pose. The tendon activations were further fine-tuned using an optimization implementing Powell's method. Table 8.2 shows the final activations that were obtained and

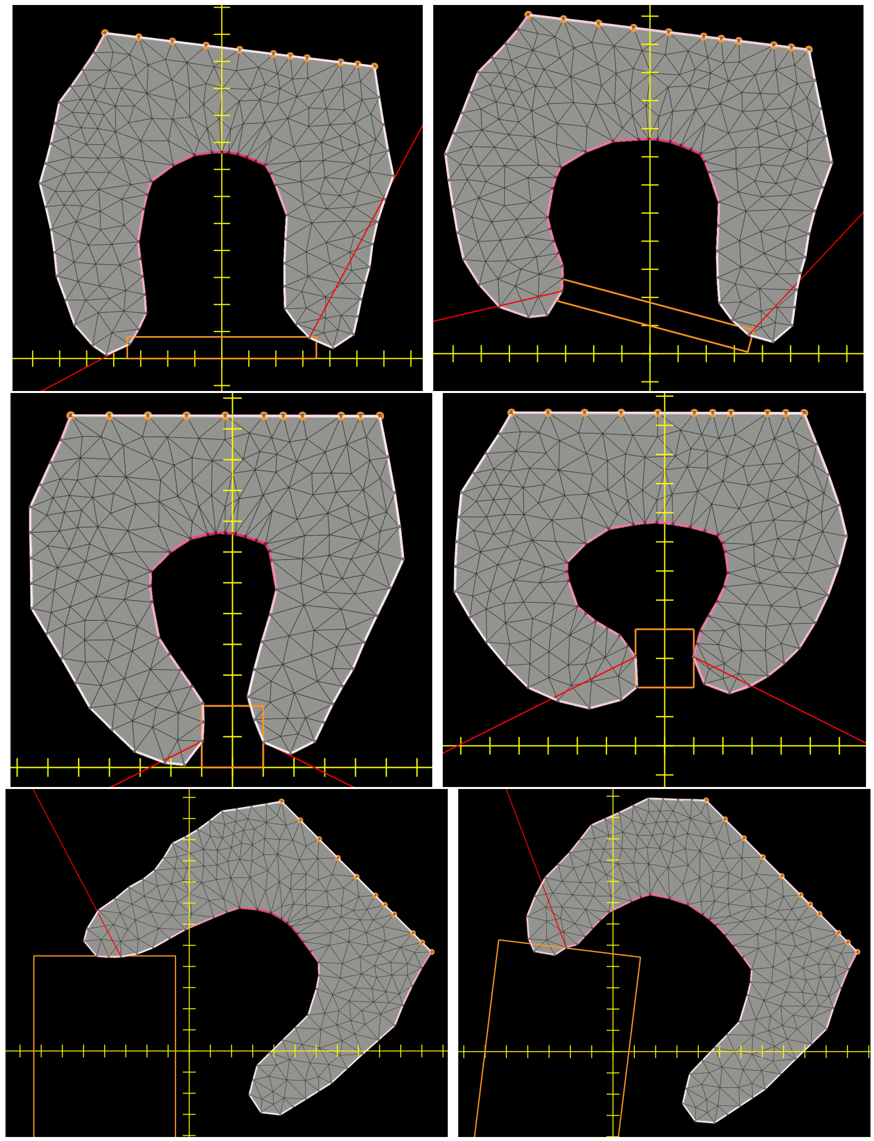


Figure 8.5: Soft IK results from posing the object in the presence of external forces. The darker pink indicates more contraction.

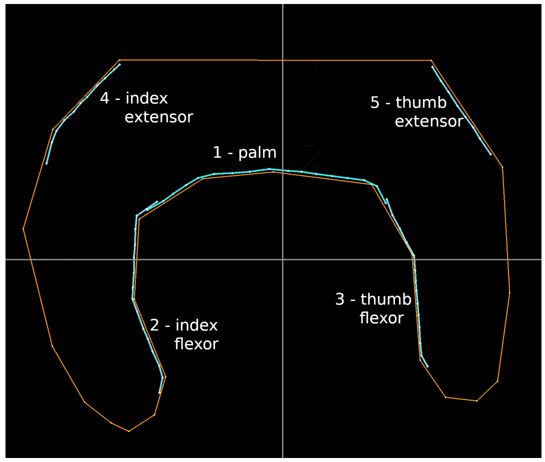


Figure 8.6: Combined tendon network.

Fig. 8.7 shows the final poses in the presence of external forces. Table 8.3 shows how far off the poses using the combined tendon network were from the initial goal poses. The nodes we positioned were often off by several millimeters, and were particularly severe for the positioning of the thumb in motion 2 frame 1.

The errors in Table 8.3 most likely come from two sources. The first possible source of errors is the simplifications of physics that are made by the Soft IK system in order to create an optimization that is quick to perform, such as assumptions about the strain energy function of the material. The second potential source of error is that the Soft IK system is able to actuate its elements in a more fine-grained manner than the simulation of the combined tendon network. Information is lost by combining the tendon contractions for all motions and frames into a single network with only five tendons. In addition, the Soft IK system yields results where individual contractile elements are allowed to contract different amounts – see the way in which neighboring segments in Fig. 8.5 can have different amounts of contraction (shade of pink). By contrast, the combined tendon network forces all points along a tendon to be contracted by the same amount. We chose tendons so as to group segments with similar activation together into the same tendon, but information is lost by having a small number of tendons, each with uniform contraction along its length.

The activation levels we used (Table 8.2) provide a starting place for controlling a robotic hand to perform these motions. The initial activations in

Motion	Frame	Forces	Tendon 1	Tendon 2	Tendon 3	Tendon 4	Tendon 5
1	1	No force	0.13	0	0.1	0.02	0.01
		Final	0.44	0.06	0.1	0.06	0
	2	Final	0.43	0.28	0.1	0.36	0
2	1	No force	0.48	0.08	0.12	0.23	0
		Final	0.54	0.14	0.18	0.16	0
	2	Final	0.26	0.36	0.44	0.02	0
3	1	No force	0	0	0.04	0.24	0.03
		Final	0.44	0	.01	0.03	0.08
	2	Final	0.26	0.21	0.03	0	0.2

Table 8.2: Activations used in combined system of tendons. Tendon 1: palm. Tendon 2: index flexor. Tendon 3: thumb flexor. Tendon 4: index extensor. Tendon 5: thumb extensor.

Motion	Frame	Error (mm)		
		Finger	Thumb	Average
1	1	1.44	2.66	2.05
	2	4.11	1.92	3.02
2	1	1.27	6.72	4.00
	2	0.29	2.62	1.46
3	1	4.55	-	-
	2	1.76	-	-

Table 8.3: Posing error (distance from the target position of each finger in mm).

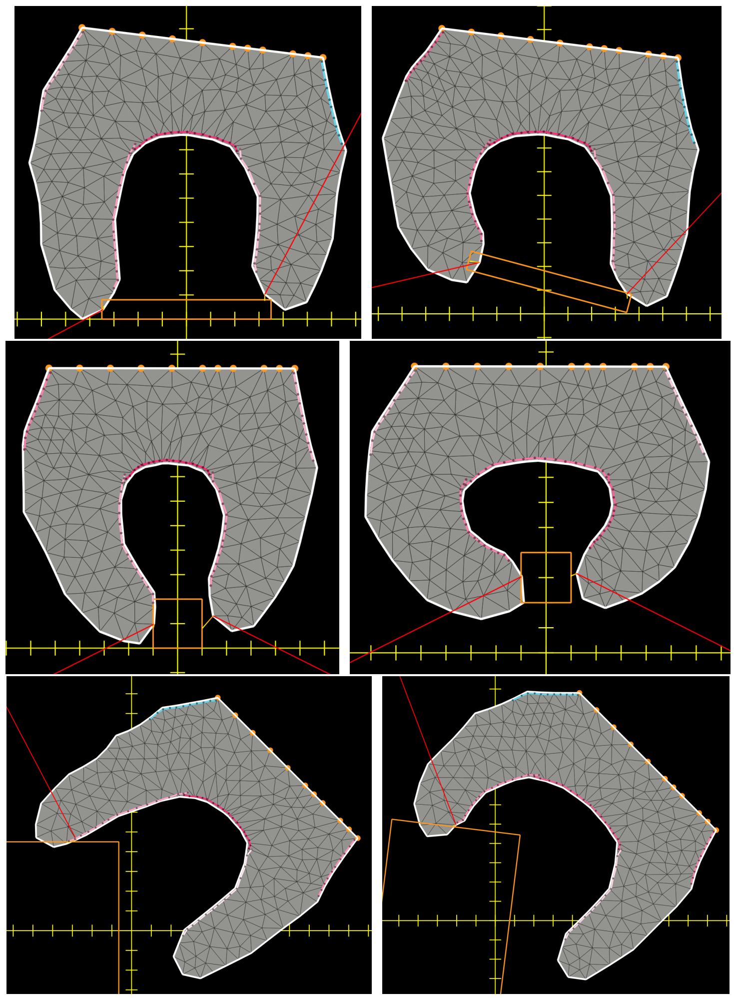


Figure 8.7: Results from manually contracting the combined tendon network.

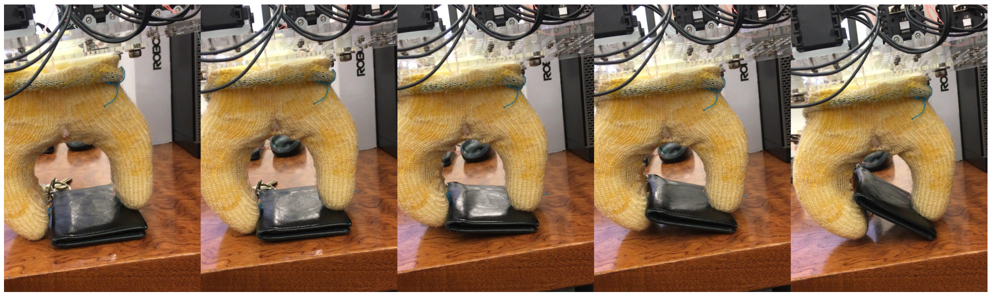


Figure 8.8: Hardware results for motion 1. Screenshot 1: initial position. Screenshots 2-3: motion with front fingers. Screenshots 4-5: motion with back fingers.

frame 1 provide a way to achieve the pose without any particular force application. By interpolating between the initial and final activation sets, the hand should move closer toward applying the desired force. Then the activations can be changed from the ones used in frame 1 to the ones used in frame 2 to attempt to transition between the poses and forces applied. We use the activations and poses as a starting place in our hardware evaluation.

8.3.1 Hardware evaluation

A pair of foam robot hands was fabricated in order to evaluate the performance of our hand design in real life. Because of the low usage of tendon 5, it was dropped from the physical robot and only tendons 1-4 were used. The sequence of poses for each hand was initialized as described in Section 8.2.1. The poses were then adjusted so that each hand was capable of performing the desired manipulation on an object (levering up, squeezing, or tilting out). The motors then alternated between running through the pose sequences for each hand: first, the front hand would go from the initial rest pose to the frame 1 poses and to the frame 2 pose. It would then hold the frame 2 pose while the back hand would run through rest pose, frame 1 poses, and frame 2 pose. For motions 2 and 3, it would then hold on the frame 2 pose while the first hand reset and ran through them again.

The results of these alternating motions on lightweight objects are shown in Figs. 8.8-8.10. The results demonstrate how repeating the same pattern of activation multiple times is capable of manipulating an object into a more secure power grasp, as we hypothesized.

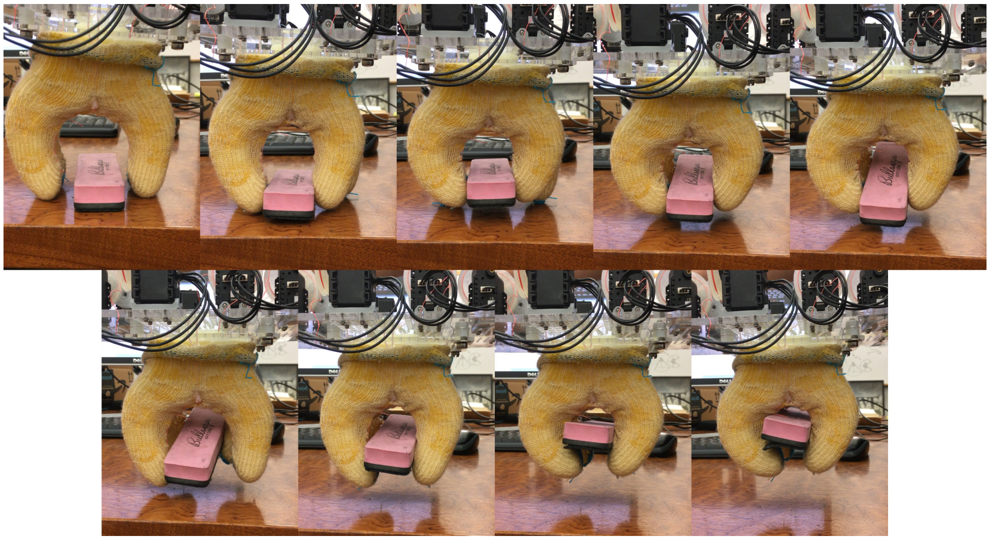


Figure 8.9: Hardware results for motion 2. Screenshot 1: initial position. Screenshots 2-3: motions with front fingers. Screenshots 4-5: motions with back fingers. Screenshots 6-8: resetting front fingers and repeating motion. Screenshot 9: final power grasp.

Figure 8.10: Hardware results for motion 3. Screenshots 1-3: initial position and motion with front fingers. Screenshots 4-5: motion with back fingers. Screenshots 6-7: front finger reset and first frame of motion.

Several issues arose during the execution of these motions. The first major issue is that the hands were not able to lever up heavy objects such as a cell phone. We were unable to find a set of activations capable of lifting one edge of the phone. Instead, the fingers (both finger and thumb) frequently slipped and slid across the surface of the phone, indicating an issue with applying forces within the friction cone. Another issue was that one of the hands experienced significant out-of-plane motion when contracting its palm tendon. The palm tendon ran along the side of the hand, so contracting it was capable of also curling the body of the hand. As a result, in motions that made significant use of the palm tendon such as motion 2, the fingers of the hands were distant from each other (in the direction of into the page). This behavior limited what objects could be picked up because they needed to be sufficiently deep in order to allow the fingertips of both hands to contact them.

8.4 Discussion

In this work, we sought to design a hand able to perform motions we observed in Chapter 4 that have not been well-studied. We analyzed types of intrinsic motions and selected a set of three motions that could be performed by a 2.5-D pair of two-finger hands. We demonstrated a pipeline for designing a hand potentially capable of performing the three motions, and evaluated the performance of this hand design in simulation and on a real robot. Our results constitute an interesting proof of concept about the usefulness of pre-grasping motions, which can turn a precarious initial grasp into a stable power grasp. In particular, our results show that repeating two frames of a manipulation, alternating between hands, is capable of gradually manipulating a light object into the hand. This result is promising because it shows that pre-grasping manipulation can possibly increase the ability of a hand to form a stable grasp, especially in the presence of clutter.

However, the major limitation of our hardware experiments is that the hands were unable to manipulate heavier objects – more specifically, they had difficulty exerting forces within the friction cone. One possibility is that the fingers and objects had a lower coefficient of friction than expected. Experimenting with a different covering for the fingertips to increase friction may help. Another possibility is that the finger may have been exerting forces too quickly and suddenly and not allowing time for the fingertips to make contact with the object and apply a squeezing force. Changing the code that interpolates between poses to transition between them at a slower pace may result in better performance.

In the future, our method could be extended to different sets of motions and different hand shapes. The Soft IK framework we made use of in this paper has been extended to handle 3-D meshes, so future work could go beyond the almost planar nature of the motions we chose. Rather than keeping the shape of the hand fixed, as we did in this work, altering the process to explore many hand shapes would better achieve the goal of creating a hand tailored toward performing certain kinds of manipulations.

9 Conclusion

In this work, we employed a variety of ways to study how humans grasp and use their hands in order to help bridge the gap between robot and human manipulation. Chapters 3 and 4 both observe the hand in natural, everyday settings. These investigations had two aims. The first was to create a method for describing important aspects of manipulation in a way that would aid a robotic manipulator in copying the motion. The second aim was to map out the space of manipulation. Chapter 6 and 7 used controlled experiments to investigate how people's grasp strategies respond to object and task properties. Finally, we applied the insights from studying manipulation to designing and fabricating a soft hand (Chapter 8).

9.1 Summary of insights

One major insight of the work in this thesis is that the problem of grasping and manipulation is more complex than what current work might indicate. In Chapter 3, we found many grasp shapes beyond what is included in the Cutkosky [1989] or Feix et al. [2009] taxonomies. Some of these are not included in previous taxonomies because they are bimanual grasps, non-prehensile grasps used to press or lever objects, or unusual grasps, like one that holds an elastic object by stretching it from within. Such grasps might be included in or excluded from a taxonomy depending on how that taxonomy defines a grasp. Other poses were variations on exist-

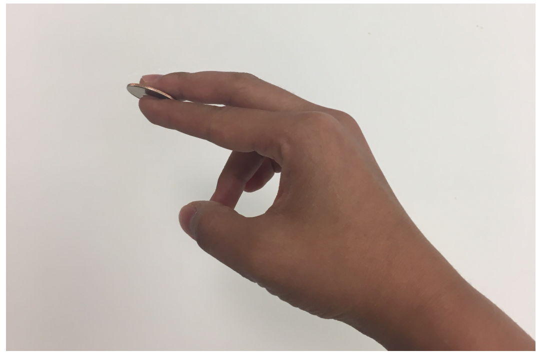


Figure 9.1: Middle-over-index finger grasp in holding a coin.

ing grasps where the index finger or thumb was constrained to be placed at a particular location. Grasp variations like these were also found in the observations of Chapter 4 (see Fig. 4.2a. Finally, some grasps are useful but rare, like the middle-over-index finger grasp (Fig. 9.1), which seems awkward but was used spontaneously by the participants in the experiments of Chapter 7. In the work featured in Chapters 4 and 7, we also encountered poses that appear to be in between two canonical grasp poses (see Fig. 4.2e), which seems to suggest that the space of grasping is somewhat continuous rather than made up of discrete grasps. In Chapter 7, we also saw that one object can be grasped in many ways even in the context of the simple task of transporting a bowl.

The process of forming a grasp around an object is also surprisingly complex. In the simplest case, grasping an object can consist of approaching it, opening the hand (preshaping), and then closing the hand around the object. However, in Chapter 4, we saw that the hand often makes small adjustments to the grasp after first contact in order to deal with clutter or to strengthen an initially weak grasp.

In the face of this complexity, how do we simplify the problem of grasping and manipulation to be more tractable? Despite the large number of poses found in Chapter 7, our findings indicate that these poses are highly responsive to parts of the object's geometry in ways that cohere with existing taxonomies. The weight effects we found also suggest the possibility that different grasps have different properties like grip strength that affect their selection. In Chapter 6, we also found that the choice of one- or two-handed strategy was consistently and strongly influenced by two factors: configuration (start and goal location) and whether balance was important. In other words, our work suggests that physical properties might be underlying the variation we see in grasp strategy, and understanding those properties might allow the problem of grasping to be simplified.

We also found that the complexity of grasping and manipulation applies not just to static poses but to the whole process of grasping. In Chapter 4, we found that the grasping process often involves quick contact adjustments, errors and error corrections, intrinsic hand motions, and different types of manipulation happening in parallel. More investigation into all of these areas is needed; however, our preliminary observations suggest that small intrinsic hand motions are often for the purpose of dealing with clutter and of forming stable grasps more robustly. These two purposes have the potential to be organizing principles that make sense of the complexity of the grasping process and that improve grasping performance in real-world settings.

In Chapter 5 we looked at the ways in which current taxonomies fail to capture multitasking, some pose variations, and environment-aided manipulation – some of the complicated features of grasping in the wild. In that chapter, we proposed relying on non-anthropomorphic descriptions of grasping to decompose complicated poses and actions into simpler parts. We propose describing grasps in terms of contact strategies that enable different kinds of manipulation.

We also applied the insights from Chapter 4 on how intrinsic motion is used to the area of artificial hands. In the work in Chapter 8, we designed a robot gripper for the purpose of performing non-prehensile manipulation helpful for forming stable cage grasps.

In addition to the insights gained about grasping, the work we did during our observations in this thesis resulted in datasets that have been useful for computer vision and robotics research. For example, the grasp poses featured in the database of Chapter 3 were used by Rogez et al. [2015] to perform grasp recognition using a broader set of “grasps” (including non-prehensile grasps) than found in previous grasp pose taxonomies. Similar to other grasp pose taxonomies, our taxonomy can be used as categories in computer vision tasks. Unlike other taxonomies, however, our work spans a broader set of manipulation behavior than just prehensile manipulation. Additionally, the list of intrinsic manipulation actions in Chapter 4 (Tables 4.1-4.3) were used to inspire example motions for robotic hands to demonstrate their manipulation capabilities [Hazard, 2018; King et al., 2018].

In general, the data we collected is potentially useful for benchmarking manipulation. The purpose of a benchmark for manipulation is to provide a standard way of evaluating the performance of artificial hands, so that the results coming out of different research groups can be compared. Quispe et al. [2018] outline three types of skills evaluated by manipulation benchmarks: (1) theoretical manipulation capabilities expressed as the workspace of the hand and arm and only limited by hardware; (2) simple grasping and manipulation tasks that make use of vision and sensing but require little planning; and (3) whole tasks including unimanual and bimanual tasks that potentially require multiple steps and planning. Grasp pose taxonomies are an example of the first type of benchmark, and the expanded set of grasp poses in Chapter 3 would work well for that, as would the more basice grasp types discussed in Chapter 5. However, more interesting is that the larger database of 179 actions featured in Chapter 3 could be used as a benchmark of the second type – it outlines a set of actions performed by hands during tasks of daily living including detailed information about the action taken after grasping (force type, motion direction, etc.). One way to evaluate the performance of robotic hands would be to see how many of the 179 actions the hand is capable of accomplishing.

9.2 Limitations and future work

9.2.1 Further study on grasping and manipulation in slow motion

Our work in Chapter 4 demonstrates the insights on grasping and manipulation that slow-motion video can provide. However, it is only an initial investigation. Studying multiple subjects would help to generalize the observations, and expanding the domain beyond simple pick-and-place tasks might reveal new grasping strategies.

One issue with this kind of work, however, is that it is time-consuming to thoroughly review and annotate the collected videos. Any future work on observing human grasping through simple RGB video would benefit significantly from a computer vision system that is able to detect, for example, the presence or absence of intrinsic hand motion, stable grasp poses, object motion, and errors or noteworthy events. Automated annotation can help streamline the task of annotation. However, the role of a human annotator is still an important and indispensable one. Because we use these videos in order to understand grasping, there are many aspects of grasping and manipulation that we don't know to look out for yet, and for that reason, annotating in an open-ended fashion is key.

A wider variety of videos captured and more efficient ways to summarize them have the potential to yield insights into human grasping. Current intrinsic manipulation and pre-grasp manipulation taxonomies seem far from complete, based on our findings in Chapter 4. It is our hope that further investigation will be able to give us a better picture of the full space of the intrinsic manipulation capabilities of the human hand.

9.2.2 Automatic grasp pose selection in artificial manipulators

One of the long-term goals of studying human grasping behavior is to allow a robot or virtual character to choose appropriate grasp poses for a grasping task. Empirical work such as that of Feix et al. [2014] are a step toward making connections between grasping task and what pose humans select. Ideally, it would be possible to use the results of studies on human grasping to enable an artificial manipulator to select a grasp strategy that has a high chance of success based on object and task properties. I envision a system that would have physics-based objectives and constraints that form an optimization problem that could be optimized to choose a grasp pose.

The work in Chapter 6 suggests that a trade-off between effectiveness at a task and energy-efficiency might make for a good objective. Effectiveness could include the ability to impart the necessary forces, the ability to maintain stability, the likelihood of a task being successfully completed, and so on, while efficiency includes the total amount of energy exerted to accomplish the task. In Chapter 5, Section 5.2.1 suggests some possible factors that affect the effectiveness or efficiency of the grasp strategy, and sketches some ways in which the task requirements and the design of the hand constrain the space of possible grasps. Using these insights, a possible optimization for the purpose of grasp selection of final power grasps could look something like the following:

System inputs: The system would require the user to specify the geometry of the grasped object, including non-graspable and must-grasp areas; the geometry and a mechanical model of the hand; and the task objectives, such as forces that must be exerted by the hand or by an object connected to the hand.

System outputs / Optimization variables: The variables of the optimization would include the grasp type each finger is involved in; what areas of the finger are involved in the grasp; and the position and orientation of the hand with respect to the object. The optimal value of these variables would be the output of the system.

Optimization objective: The objective of the optimization would be to minimize some measure of energy or effort, such as minimizing the joint torques needed to hold an object in some pose while exerting the desired forces. An important part of calculating this cost would be knowing in detail how the fingers contact the object when, for example, wrapped compliantly around the object as in a wrap grasp. It's possible this cost can be calculated analytically based on the mechanics of the hand, but it might better to calculate this cost using a physics simulation that is better able to model the deformable aspects of the hand.

Optimization constraints: There are many possibilities for optimization constraints and how they might be implemented. One constraint is the requirement that a particular force be exerted in a particular direction. This requirement constrains which finger(s) can be used and how the hand has to be oriented. Another constraint is the avoidance/usage of non-graspable/must-grasp areas. These constraints could also be implemented as part of the optimization objective (for example, punishing deviation from the goal force direction or punishing amount of contact in non-graspable regions) and also possibly calculated using a physics simulation.

The above system could also be extended beyond stable power grasps to look at sequences of grasps. Being able to generate an entire grasp sequence using optimization would be useful for showing how to manipulate an object within the hand or how to grasp an object in the presence of clutter.

9.2.3 Applications for unimanual and bimanual strategies

In the future, we hope that our work on unimanual and bimanual transport in Chapter 6 might find a useful application to virtual characters and robots navigating home environments like the kitchen. Our observations of how people behave in a transport task could be applied to the automatic creation of natural-looking arm and hand behavior in virtual characters and robots.

Our work in Chapter 6 left open the possibility that there are lower-level biomechanical considerations that drive how much rotation is required for bimanual, hand-off, and unimanual strategies for a particular start and goal location. Extending our work to investigate these underlying considerations could result in a way of generating reaching and transport motions that would be applicable beyond the particular situations we studied – for example, generalizable to any start/goal location pair. The result would be that, given any start and goal location, it would be possible to determine how much people rotate on average for each strategy: bimanual, hand-off (considering left-to-right hand-offs separately from right-to-left hand-offs), and unimanual (considering left only transports separately from right only transports).

It would then be possible to use those results to generate natural-looking motion for artificial manipulators. One possible system for generating transport motions is the following:

The system would take as input object/task properties (size, weight, and balance requirement) and the start and goal location for an object. Using the results of the study proposed above, the system would be able to use the start and goal location to automatically calculate how much rotation would be required in order to perform each strategy. We could then use the results of our studies in Chapter 6 to select the strategy people are most likely to use given the object and task properties and the cost of (rotation required by) each strategy. For example, a function could be fit to the data of our studies to predict how people trade off between rotation and the benefits such as greater strength or stability that come with using more energy-intensive strategies. The strategy most likely to be used would be selected and an animation clip would be generated.

Note that, in order to look realistic, the system might also need to incorporate realistic full-body motion as well, such as stepping behavior and torso bending behavior. Our studies contain full-body motion capture data that could help with the task of generating realistic full-body motion. However, further study on these aspects of manipulation may also be needed.

The behavior of this virtual character could then be evaluated through a study asking users to rate and compare clips of motions selected based on the results of these studies vs. motions selected by randomization or by relying on a single strategy, or motions generated using previous approaches to human motion generation such as that of Bai et al. [2012].

9.2.4 Additional experiments on grasping and manipulation

The experiments in Chapters 6 and 7 are very narrow in scope and focused, involving a single object (bowls), only a single, simple task (transport), and two specific aspects of manipulation (use of one or two hands and grasp pose). Do our findings on the diversity of grasp poses generalize to other objects aside from bowls? What can we learn about aspects of manipulation aside from choice of hands and grasp pose by studying behavior in detail? We have also seen people use two hands in other tasks in a kitchen setting, such as opening and closing cabinets and doors, using a faucet, and so on. Are there simple rules determining how the two hands coordinate on these tasks, similar to what was found in transport?

9.2.5 Extending the intrinsic dexterity of a manipulator

The three motions we focus on in Chapter 8 were chosen because of their simplicity – they could be performed with three fingers, with the thumb set to be opposing the other two fingers, and feature motion restricted to a single plane. The full range of human manipulation is much more complicated than that, featuring five fingers, including a thumb whose opposition to the fingers can change, and motion in full 6-D space. Some motions also involve a complex sequence of contact changes capable of adjusting the grasp while maintaining a stable grasp on the object the entire time. Being able to accomplish these motions is no less important than the subset of motions we chose to focus on. How could one design a robot hand capable of accomplishing all the motions in Tables 4.1-4.3 or in the taxonomy by Elliott and Connolly [1984]?

We used five tendons to accomplish the three motions of Chapter 8. How many tendons would be needed for a fully dexterous hand? What is the simplest way to construct such a hand? These questions are difficult but important to answer if we wish to have robotic hands on par with that of humans.

9.3 Final thoughts

In this thesis, we aim as much as possible to understand the gap between robotic hands as they are today and human manipulators, which function as a kind of "gold standard" for manipulation skills. We do this through multiple studies of human grasping, and attempt to apply the results to design interesting new robotic hands. It is our hope that continuing to observe and analyze human grasping will get us closer and closer to the goal of robotic manipulators capable of performing the activities of daily living that require skillful use of the hands.

A

Manual force calculations

This chapter contains the full details of the manual force calculations for each of the three motions featured in Chapter 8.

A.1 Motion 1: Lever

In the first frame of the motion, the object is lying flat on the surface (see Fig. A.1). Each finger exerts a force on the object, which must be within the friction cone to maintain a stable contact point. The first finger exerts an upward force to lift the left edge of the object, while the second finger exerts a downward or down-leftward force to prevent the object from moving. The first force must press inward (to the right) to prevent the finger from slipping, and the second force must cancel out this force, either by exerting a leftward force with the finger or by creating a frictional force resisting motion between the bottom-right corner of the object and the support surface. In these calculations we choose to exert a force with the finger itself.

Finally, the torque created by the fingers must be enough to cancel out the torque created by gravity. Both the tangent and normal components of the first force f_1 contribute to a clockwise torque capable of lifting the left edge of the object. By contrast, the gravity force and the inward force of the second force f_2 contribute to a counterclockwise torque. The down-

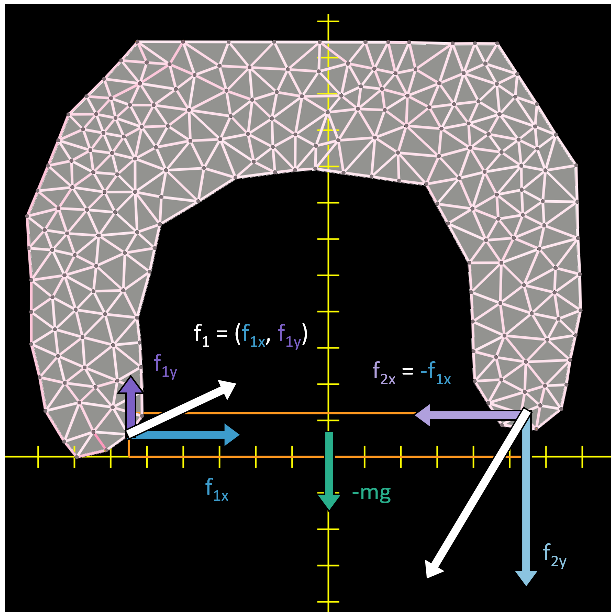


Figure A.1: Motion 1 (lever), frame 1.

ward force does not exert any torque if the contact is located exactly above the corner of the object. However, we assume that the contact will not be made exactly on the corner and so the downward force also exerts a counterclockwise torque with lever arm of $\epsilon = .004$ m. We also assume the left contact is located halfway up the side of the object. In order for there to be clockwise torque, the following inequality must be true:

$$wf_{1y} + \frac{h}{2}f_{1x} > \frac{w}{2}mg + hf_{2x} + \epsilon f_{2y} \tag{A.1}$$

In order to be within the friction cone, f_{1y} must be less than μf_{1x}, where $\mu = 0.8$ is the coefficient of friction. Assuming a 50% safety margin, we can set $f_{1x} = 1.5\frac{f_{1y}}{\mu}$. f_{2x} is the same as f_{1x} but in the opposite direction. Finally, in order to be within the friction cone, f_{2x} must be less than μf_{2y}. If $f_{2x} = 1.5\frac{f_{1y}}{\mu}$ and we use a 50% safety margin, we can set $f_{2y} = 1.5^2\frac{f_{1y}}{\mu^2}$. This means that Eq. A.1 can be written in terms of f_{1y} and various object properties:

$$wf_{1y} + \frac{h}{2}\frac{1.5}{\mu}f_{1y} > \frac{w}{2}mg + h\frac{1.5}{\mu}f_{1y} + \epsilon\frac{1.5^2}{\mu^2}f_{1y} \tag{A.2}$$

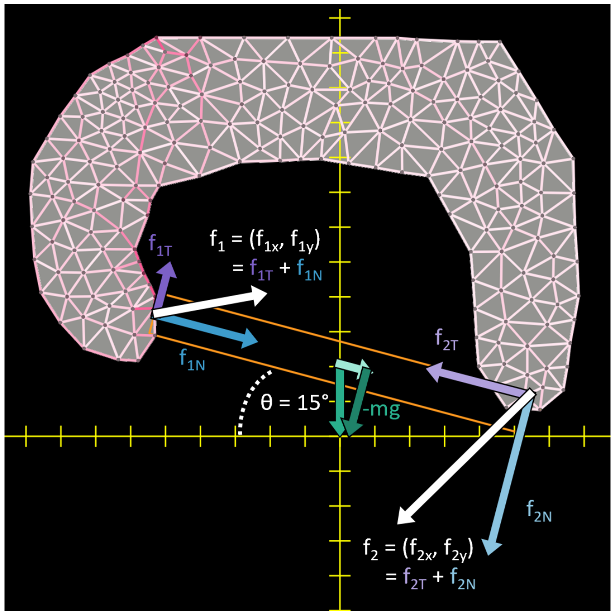


Figure A.2: Motion 1 (lever), frame 2.

By plugging in values $w = 0.07$ m, $h = 0.008$ m, $m = 150$ g, $\mu = 0.8$, and $\epsilon = 0.004$ m, we get that $f_{1y} > 1.06$ N. If we set $f_{1y} = 1.2$ N, we can calculate all the other components of the forces. The final forces are $f_1 = (2.25, 1.2)$ and $f_2 = (-2.25, -4.22)$.

The second frame (Fig. A.2) features the object rotated by 15°. The torque analysis here is similar, with the tangent and normal components of f_1 exerting clockwise torque, and the tangent and normal components of f_2 exerting counterclockwise torque. The main difference is that the gravity force is split into clockwise and counterclockwise components, which makes overcoming its torque easier. Therefore, the same forces from frame 1, rotated by 15°, should work for frame 2.

However, if each force is rotated, the x-component of the second force ($f_{2x} = 3.27$ N) becomes larger than the x-component of the first force ($f_{1x} = 2.48$ N) resulting in a leftward force being exerted overall on the object's center of mass. Instead of rotating f_2, we instead find a tangent and normal component that, when rotated 15° and added together, balances out the x-component of the first force. Assuming a 50% safety margin, we can find

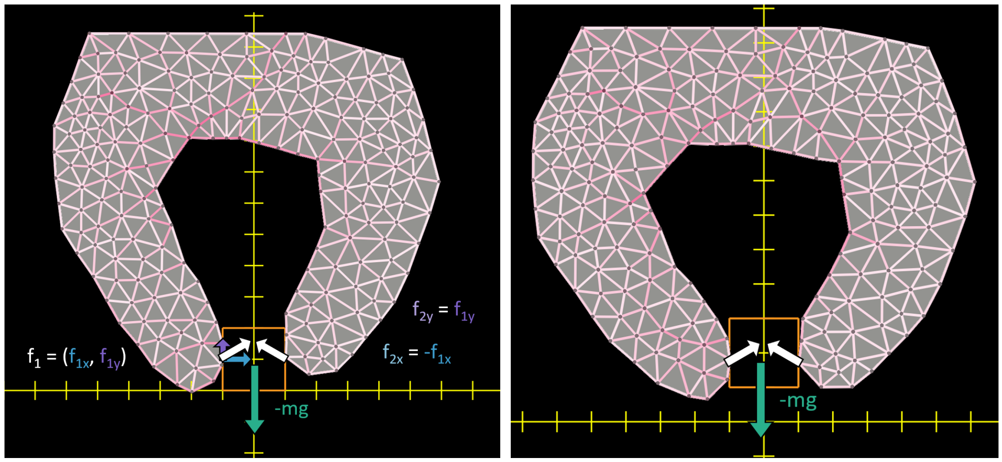


Figure A.3: Motion 2 (squeeze), frames 1 and 2.

the force we want by satisfying the following two equations:

$$\begin{cases} f_{1x} = T_2 \cos\theta + N_2 \sin\theta \\ N_2 = 1.5\frac{T_2}{\mu} \end{cases} \tag{A.3}$$

Solving the above set of equations yields a tangent force magnitude T_2 of 1.71 N and a normal force magnitude N_2 of 3.20 N. These f_2 components, which contribute counterclockwise torque, are smaller than in the first frame of motion, which means that clockwise torque is still being applied to the object overall. After rotating this combination of forces 15° and adding them together, the forces we get are $f_1 = (2.48, .58)$ and $f_2 = (-2.48, -2.65)$.

A.2 Motion 2: Squeeze

This motion requires, at both frames, a lifting force capable of overcoming gravity, within the friction cone (see Fig. A.3). Using $m = 50$ g, we get that an upward force of at least 0.49 N is needed. Using upward/tangential components of 0.5 N for each force is more than sufficient. Assuming a 60% safety margin and $\mu = 0.8$, we get an inward component of 1 N.

The final forces are $f_1 = (1, 0.5)$ and $f2 = (-1, 0.5)$. These target forces

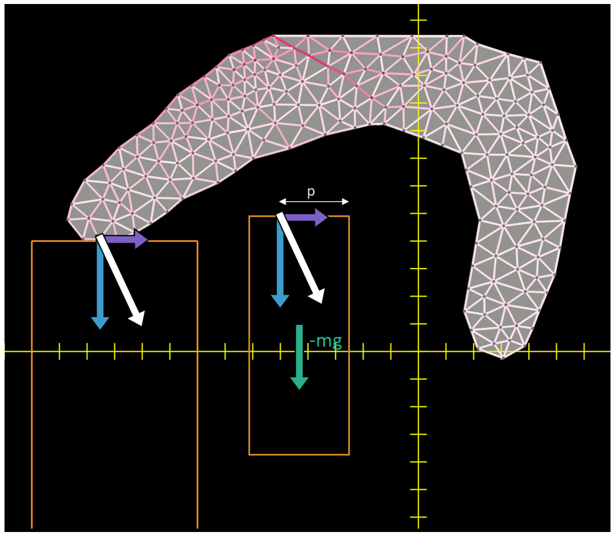


Figure A.4: Motion 3 (tilt out), frame 1.

are used for both frames. The only difference between frames is the object location.

A.3 Motion 3: Tilt out

In the first frame of this motion, the object is standing upright (see Fig. A.4). One finger exerts a down-right force in order to rotate the object clockwise. The force must be within the friction cone to prevent slippage between the hand and object.

The rightward/tangential component of this force creates a clockwise torque; however, gravity and the downward/normal component of the force contribute to a counterclockwise torque. In order for there to be clockwise torque overall, the following inequality must be true:

$$hT > pN + \frac{w}{2}mg \tag{A.4}$$

In the above equation T is the magnitude of the tangential component, N is the magnitude of the normal component of the force, and $p = .04$ m is how far from the right edge of the object the contact is located. $h = 0.15$ m,

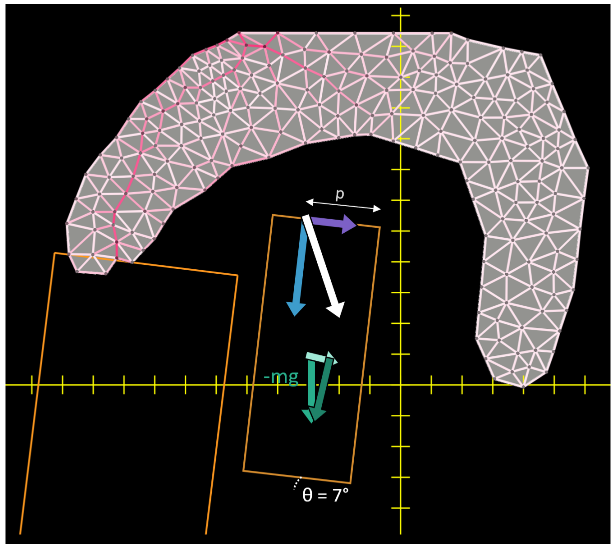


Figure A.5: Motion 3 (tilt out), frame 2.

$w = 0.067$ m, and $m = 420$ g are the height, width, and mass of the object. Using a safety margin of 50%, we get $N = 1.5\frac{T}{\mu}$. We can then express the inequality in terms of T:

$$(h - (1.5\frac{p}{\mu}))T > \frac{w}{2}mg \tag{A.5}$$

Plugging in values for h, p, μ, w, m, and g, we get that T must be greater than 1.84 N. We set $T = 2.1$ N and calculate that $N = 3.93$ N. The final force is thus $f = (2.1, -4)$.

The second frame features the object rotated 7° clockwise (Fig. A.5. Like with the first motion, the only change here is that the gravity force gets distributed between clockwise and counterclockwise torques, meaning that overcoming the counterclockwise torque is easier.

We use the same forces from frame 1 of this motion, but rotated by 7°. The final force is roughly $f = (1.6, -4.23)$.

References

Abe, N. and Laumond, J.-P. (2014). Dance notations and robot motion. In *Proceedings of the 1st Workshop of the Anthropomorphic Motion Factory at LAAS-CNRS '14*, Toulouse, France.

Amend, J. R., Brown, E., Rodenberg, N., Jaeger, H. M., and Lipson, H. (2012). A positive pressure universal gripper based on the jamming of granular material. *IEEE Transactions on Robotics*, 28(2):341–350.

Bai, Y., Siu, K., and Liu, C. K. (2012). Synthesis of concurrent object manipulation tasks. *ACM Transactions on Graphics (TOG)*, 31(6):156.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.

Bern, J. M., Chang, K.-H., and Coros, S. (2017). Interactive design of animated plushies. *ACM Transactions on Graphics (TOG)*, 36(4):80.

Bullock, I. M., Ma, R. R., and Dollar, A. M. (2013). A hand-centric classification of human and robot dexterous manipulation. *Haptics, IEEE Transactions on*, 6(2):129–144.

Byrne, R. W., Byrne, J. M., et al. (2001). Manual dexterity in the gorilla: bimanual and digit role differentiation in a natural task. *Animal Cognition*, 4(3-4):347–361.

Cesari, P. and Newell, K. M. (2000). Body-scaled transitions in human grip configurations. *Journal of Experimental Psychology: Human Perception and Performance*, 26(5):1657.

Chang, L. Y., Klatzky, R. L., and Pollard, N. S. (2009). Selection criteria for preparatory object rotation in manual lifting actions. *Journal of Motor Behavior*, 42(1):11–27.

Chang, L. Y. and Pollard, N. S. (2009). Video survey of pre-grasp interactions in natural hand activities.

Choi, Y. S., Deyle, T., Chen, T., Glass, J. D., and Kemp, C. C. (2009). A list of household objects for robotic retrieval prioritized by people with ALS. In *Rehabilitation Robotics, 2009. ICORR 2009. IEEE International Conference on*, pages 510–517. IEEE.

Cohn, J. and Ekman, P. (2005). Measuring facial action by manual coding, facial EMG, and automatic facial image analysis. *Handbook of Nonverbal Behavior Research Methods in the Affective Sciences*.

Collin, C., Wade, D., Davies, S., and Horne, V. (1988). The Barthel ADL Index: a reliability study. *Disability & Rehabilitation*, 10(2):61–63.

Crajé, C., Lukos, J. R., Ansuini, C., Gordon, A. M., and Santello, M. (2011). The effects of task and content on digit placement on a bottle. *Experimental Brain Research*, 212(1):119–124.

Cutkosky, M. R. (1989). On grasp choice, grasp models, and the design of hands for manufacturing tasks. *Robotics and Automation, IEEE Transactions on*, 5(3):269–279.

Dafle, N. C., Rodriguez, A., Paolini, R., Tang, B., Srinivasa, S. S., Erdmann, M., Mason, M. T., Lundberg, I., Staab, H., and Fuhlbrigge, T. (2014). Extrinsic dexterity: In-hand manipulation with external forces. In *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, pages 1578–1585. IEEE.

De Souza, R., Bernardino, A., Santos-Victor, J., and Billard, A. (2012). On the representation of anthropomorphic robot hands: Shape versus function. In *Humanoid Robots (Humanoids), 2012 12th IEEE-RAS International Conference on*, pages 791–798. IEEE.

Deimel, R. and Brock, O. (2016). A novel type of compliant and underactuated robotic hand for dexterous grasping. *The International Journal of Robotics Research*, 35(1-3):161–185.

Deimel, R., Irmisch, P., Wall, V., and Brock, O. (2017). Automated co-design of soft hand morphology and control strategy for grasping. In *Intelligent Robots and Systems (IROS), 2017 IEEE/RSJ International Conference on*, pages 1213–1218. IEEE.

Dollar, A. M. and Howe, R. D. (2010). The highly adaptive SDM hand: Design and performance evaluation. *The International Journal of Robotics Research*, 29(5):585–597.

Edwards, S. J., Buckland, D. J., and McCoy-Powlen, J. D. (2002). *Developmental and functional hand grasps*. Slack.

Ekman, P. and Friesen, W. (1978). *Facial action coding system: A technique for the measurement of facial movement*. Consulting Psychologists Press, Palo Alto.

Elliott, J. M. and Connolly, K. (1984). A classification of manipulative hand movements. *Developmental Medicine & Child Neurology*, 26(3):283–296.

Engelbrecht, S. E. (2001). Minimum principles in motor control. *Journal of Mathematical Psychology*, 45(3):497–542.

Feix, T., Bullock, I. M., and Dollar, A. M. (2014). Analysis of human grasping behavior: Object characteristics and grasp type. *IEEE transactions on haptics*, 7(3):311–323.

Feix, T., Pawlik, R., Schmiedmayer, H., Romero, J., and Kragic, D. Human grasping database. http://grasp.xief.net. Accessed: 2017-04-05.

Feix, T., Pawlik, R., Schmiedmayer, H., Romero, J., and Kragic, D. (2009). A comprehensive grasp taxonomy. In *Robotics, Science and Systems: Workshop on Understanding the Human Hand for Advancing Robotic Manipulation*.

Fox, J. (2003). Effect displays in R for generalised linear models. *Journal of Statistical Software*, 8(15):1–27.

Fuller, Y. and Trombly, C. A. (1997). Effects of object characteristics on female grasp patterns. *American Journal of Occupational Therapy*, 51(7):481–487.

Gilster, R., Hesse, C., and Deubel, H. (2012). Contact points during multidigit grasping of geometric objects. *Experimental Brain Research*, 217(1):137–151.

Gonzalez, C. L., Flindall, J. W., and Stone, K. D. (2014). Hand preference across the lifespan: effects of end-goal, task nature, and object location. *Frontiers in Psychology*, 5.

Greaves, S., Imms, C., Krumlinde-Sundholm, L., Dodd, K., and Eliasson, A.-C. (2012). Bimanual behaviours in children aged 8–18 months: A literature review to select toys that elicit the use of two hands. *Research in Developmental Disabilities*, 33(1):240–250.

Grunwald, G., Borst, C., Zöllner, J., et al. (2008). Benchmarking dexterous dual-arm/hand robotic manipulation. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, Workshop on Performance Evaluation and Benchmarking, Nice, France*.

Guiard, Y. (1987). Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. *Journal of Motor Behavior*, 19(4):486–517.

Gupta, A., Eppner, C., Levine, S., and Abbeel, P. (2016). Learning dexterous manipulation for a soft robotic hand from human demonstrations. In *Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on*, pages 3786–3793. IEEE.

Hazard, C. (2018). Automated design of manipulators for in-hand tasks. Technical Report CMU-RI-TR-18-36, Robotics Institute, Carnegie Mellon University.

Heinemann, F., Puhlmann, S., Eppner, C., Élvarez-Ruiz, J., Maertens, M., and Brock, O. (2015). A taxonomy of human grasping behavior suitable for transfer to robotic hands. In *Robotics and Automation (ICRA), 2015 IEEE International Conference on*, pages 4286–4291. IEEE.

Hiramatsu, Y., Kimura, D., Kadota, K., Ito, T., and Kinoshita, H. (2015). Control of precision grip force in lifting and holding of low-mass objects. *PloS one*, 10(9):e0138506.

Hirose, S. and Umetani, Y. (1978). The development of soft gripper for the versatile robot hand. *Mechanism and machine theory*, 13(3):351–359.

Hogan, N. (1985). Impedance control: An approach to manipulation: Part ii,Äîimplementation. *Journal of Dynamic Systems, Measurement, and Control*, 107(1):8–16.

Iberall, T. (1987). The nature of human prehension: Three dextrous hands in one. In *Robotics and Automation. Proceedings. 1987 IEEE International Conference on*, volume 4, pages 396–401. IEEE.

Jacobsen, S., Iversen, E., Knutti, D., Johnson, R., and Biggers, K. (1986). Design of the utah/mit dextrous hand. In *Robotics and Automation. Proceedings. 1986 IEEE International Conference on*, volume 3, pages 1520–1532. IEEE.

Kamakura, N., Matsuo, M., Ishii, H., Mitsuboshi, F., and Miura, Y. (1980). Patterns of static prehension in normal hands. *American Journal of Occupational Therapy*, 34(7):437–445.

Kapandji, I. A. and Honoré, L. H. (1970). *The physiology of the joints: annotated diagrams of the mechanics of the human joints*, volume 1. E. & S. Livingstone London.

Khatib, O. (1987). A unified approach for motion and force control of robot manipulators: The operational space formulation. *IEEE International Conference on Robotics and Automation (ICRA), 1987*, 3(1):43–53.

Kimmerle, M., Ferre, C. L., Kotwica, K. A., and Michel, G. F. (2010). Development of role-differentiated bimanual manipulation during the infant's first year. *Developmental Psychobiology*, 52(2):168–180.

Kimmerle, M., Mainwaring, L., and Borenstein, M. (2003). The functional repertoire of the hand and its application to assessment. *American Journal of Occupational Therapy*, 57(5):489–498.

King, J. P., Bauer, D., Schlagenhauf, C., Chang, K.-H., Moro, D., Pollard, N., and Coros, S. (2018). Design, fabrication, and evaluation of tendon-driven foam manipulators. In *2018 IEEE RAS International Conference on Humanoid Robots*. Submitted: Preprint.

Kopp, B., Kunkel, A., Flor, H., Platz, T., Rose, U., Mauritz, K.-H., Gresser, K., McCulloch, K. L., and Taub, E. (1997). The Arm Motor Ability Test: reliability, validity, and sensitivity to change of an instrument for assessing disabilities in activities of daily living. *Archives of Physical Medicine and Rehabilitation*, 78(6):615–620.

Kumar, V., Todorov, E., and Levine, S. (2016). Optimal control with learned local models: Application to dexterous manipulation. In *Robotics and Automation (ICRA), 2016 IEEE International Conference on*, pages 378–383. IEEE.

Leidner, D., Borst, C., Dietrich, A., Beetz, M., and Albu-Schäffer, A. (2015). Classifying compliant manipulation tasks for automated planning in robotics. In *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*, pages 1769–1776. IEEE.

Lenth, R. V. (2016). Least-squares means: The R package lsmeans. *Journal of Statistical Software*, 69(1):1–33.

Linacre, J. M., Heinemann, A. W., Wright, B. D., Granger, C. V., and Hamilton, B. B. (1994). The structure and stability of the functional independence measure. *Archives of Physical Medicine and Rehabilitation*, 75(2):127–132.

Liu, H., Wu, K., Meusel, P., Seitz, N., Hirzinger, G., Jin, M., Liu, Y., Fan, S., Lan, T., and Chen, Z. (2008). Multisensory five-finger dexterous hand: The DLR/HIT Hand II. In *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, pages 3692–3697. IEEE.

Liu, J., Feng, F., Nakamura, Y., and Pollard, N. (2014a). A taxonomy of everyday grasps in action. http://graphics.cs.cmu.edu/projects/graspsinaction/database.html. Accessed: 2018-08-05.

Liu, J., Feng, F., Nakamura, Y. C., and Pollard, N. S. (2014b). A taxonomy of everyday grasps in action. In *Humanoid Robots (Humanoids), 2014 14th IEEE-RAS International Conference on*, pages 573–580. IEEE.

Ma, R. R. and Dollar, A. M. (2011). On dexterity and dexterous manipulation. In *Advanced Robotics (ICAR), 2011 15th International Conference on*, pages 1–7. IEEE.

Napier, J. R. (1956). The prehensile movements of the human hand. *Journal of Bone and Joint Surgery*, 38(4):902–913.

Napier, J. R. and Tuttle, R. (1993). *Hands*. Princeton University Press.

Newlove, J. and Dalby, J. (2003). *Laban for all*. Routledge.

Palmer, C. F. (1989). The discriminating nature of infants' exploratory actions. *Developmental Psychology*, 25(6):885.

Park, J., Seo, N. J., Son, J., Kim, W., and Cheong, J. (2014). Postural variation of hand precision grips by object size. *Journal of Mechanical Science and Technology*, 28(5):1641–1651.

Pollock, N., McColl, M., Carswell, A., and Sumsion, T. (2006). The Canadian occupational performance measure. *Client-centred Practice in Occupational Therapy: A Guide to Implementation*, pages 103–114.

Quispe, A. H., Amor, H. B., and Christensen, H. I. (2018). A taxonomy of benchmark tasks for robot manipulation. In *Robotics Research*, pages 405–421. Springer.

Rajeswaran, A., Kumar, V., Gupta, A., Schulman, J., Todorov, E., and Levine, S. (2017). Learning complex dexterous manipulation with deep reinforcement learning and demonstrations. *arXiv preprint arXiv:1709.10087*.

Rogez, G., Supancic, J. S., and Ramanan, D. (2015). Understanding everyday hands in action from RGB-D images. In *Proceedings of the IEEE international conference on computer vision*, pages 3889–3897.

Rosenbaum, D. A., Coelho, C. J., Rhode, J. D., and Santamaria, J. P. (2010). Psychologically distinct classes of motor behavior inferred from individual differences: Evidence from a sequential stacking task. *Journal of Motor Behavior*, 42(3):187–194.

Rosenbaum, D. A., van Heugten, C. M., and Caldwell, G. E. (1996). From cognition to biomechanics and back: The end-state comfort effect and the middle-is-faster effect. *Acta psychologica*, 94(1):59–85.

Salisbury, J. K. and Craig, J. J. (1982). Articulated hands: Force control and kinematic issues. *The International journal of Robotics research*, 1(1):4–17.

Samadani, A.-A., Burton, S., Gorbet, R., and Kulic, D. (2013). Laban effort and shape analysis of affective hand and arm movements. In *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*, pages 343–348. IEEE.

Sartori, L., Straulino, E., and Castiello, U. (2011). How objects are grasped: the interplay between affordances and end-goals. *PLOS ONE*, 6(9):e25203.

Schlesinger, I. G. (1919). Der mechanische aufbau der künstlichen glieder. In *Ersatzglieder und Arbeitshilfen*, pages 321–661. Springer.

Shadow Robot Company (2018). Dexterous hand. [Online; accessed 18-June-2018].

Surdilovic, D., Yakut, Y., Nguyen, T.-M., Pham, X. B., Vick, A., and Martin-Martin, R. (2010). Compliance control with dual-arm humanoid robots: Design, planning and programming. In *Humanoid Robots (Humanoids), 2010 10th IEEE-RAS International Conference on*, pages 275–281. IEEE.

Torigoe, T. (1985). Comparison of object manipulation among 74 species of non-human primates. *Primates*, 26(2):182–194.

Touvet, F., Roby-Brami, A., Maier, M. A., and Eskiizmirliler, S. (2014). Grasp: combined contribution of object properties and task constraints on hand and finger posture. *Experimental brain research*, 232(10):3055–3067.

Townsend, W. (2000). The BarrettHand grasper–programmably flexible part handling and assembly. *Industrial Robot: an international journal*, 27(3):181–188.

Voudouris, D., Smeets, J., and Brenner, E. (2012). Do obstacles affect the selection of grasping points? *Human Movement Science*, 31(5):1090–1102.

Westling, G. and Johansson, R. (1984). Factors influencing the force control during precision grip. *Experimental brain research*, 53(2):277–284.

Wörgotter, F., Aksoy, E. E., Kruger, N., Piater, J., Ude, A., and Tamosiunaite, M. (2013). A simple ontology of manipulation actions based on hand-object relations. *Autonomous Mental Development, IEEE Transactions on*, 5(2):117–134.

Zheng, J. Z., De La Rosa, S., and Dollar, A. M. (2011). An investigation of grasp type and frequency in daily household and machine shop tasks. In *IEEE International Conference on Robotics and Automation (ICRA), 2011*, pages 4169–4175. IEEE.

Zhou, J., Chen, S., and Wang, Z. (2017). A soft-robotic gripper with enhanced object adaptation and grasping reliability. *IEEE Robotics and Automation Letters*, 2(4):2287–2293.