Understanding and applying human grasping to artificial manipulators
(Thesis Proposal)

Yuzuko Nakamura
Computer Science Department
Carnegie Mellon University

May 2, 2017

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 goals of improving the success or naturalness 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 virtual hand controller capable of levering up objects into the hand. We also propose using the results of the bowl transport experiment to create a character whose transporting behavior looks natural and believable.

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.

Contents

Chapter 1

Introduction

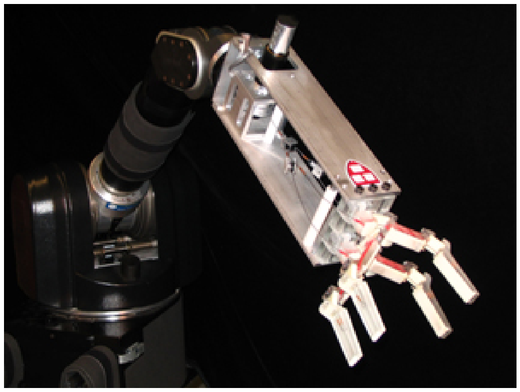
A major goal in the field of robotics is the creation of robots capable of using their hands as humans use their hands. The ability to do so would allow robots to be useful in environments too dangerous for humans and to assist people in the home who have difficulty with activities of daily living.

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 [DH10] (Fig. 1.1a) and the coffee-grounds-filled universal gripper [ABR+12] (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 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.

In this work, we seek to partially fill this gap in our understanding of human manipulation



(a)

(b)

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

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 propose applying the results of those investigations to the control of artificial hands.

1.1 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 [Cut89] or asking subjects to grasp a wide variety of objects [KMI+80]. 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 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. [GHD12, RCRS10, CN00]). 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: (proposed work) 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 behavior of artificial graspers – robotic hands and animated characters. We focus on two potential applications in particular: First, observations of the grasping process reveal that people frequently lever up objects before grasping in order to expose surfaces that are needed in a final grasp. We aim to apply this knowledge of grasping to create physically-plausible animated hands capable of performing these kinds of levering up motions. Second, experiments on the factors affecting grasp strategy reveal that people use different strategies (such as handing off between hands, or using only one hand) based on where objects are and where they need to be moved with respect to the person. We plan to apply this insight to create virtual actors that move naturally and believably. The results will be evaluated by users comparing the movements of different virtual characters.

1.2 Completed work

Completed/in-progress and proposed work are summarized in Fig. 1. A brief summary of each of the four completed projects follows.

1.2.1 Completed project I: Grasps in action

For this project, we sought to answer the question of how many grasps people use in the activities needed for daily living. Two investigators recorded the ways they used their hands over

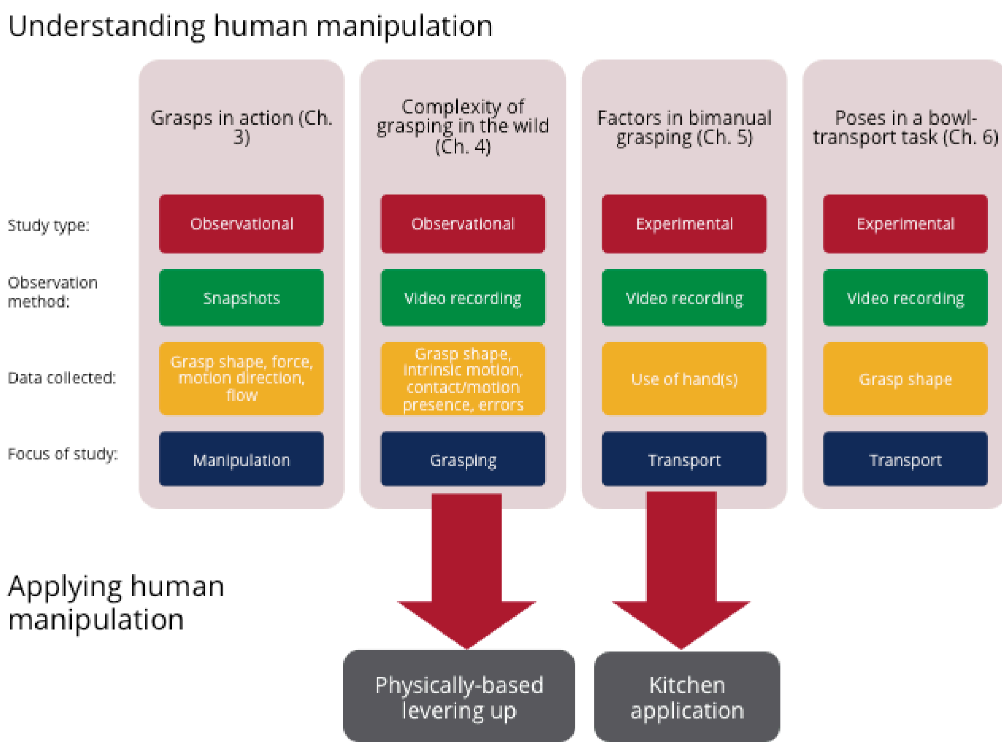


Figure 1.2: Overview of completed and proposed work

the course of the day and augmented this list with activities involved with specific occupations. Examining this data, they realized the same grasps can be used for a wide variety of actions. To differentiate actions, we created an annotation system that specifies (1) force type, (2) direction of force and motion, and (3) flow of force and motion (a quality that describes how restricted or constrained the force or motion is). These annotations were collected into a database that provides a more comprehensive summary of how hands are used in activities of daily living.

1.2.2 In-progress project II: Complexity of grasping in the wild

For this project, we observed the full process of grasping in a natural convenience-store setting, facilitated by the use of a handheld, 120 frames-per-second (FPS) camera. The video was analyzed in detail under the lenses of various taxonomies of manipulation. Our analysis points out some of the complexities of grasping that are not captured by existing taxonomies. These complexities include shifts from one grasp to another that happen very quickly; the way contacts with other surfaces are used to grasp objects; and grasp poses not covered in grasp taxonomies.

1.2.3 In-progress project III: Factors in bimanual grasping

The use of both hands simultaneously when manipulating objects is fairly commonplace, but it is not known what factors encourage people to use two hands as opposed to one during simple tasks such as transport. In this study, we investigated the effect of object size, weight, and position as well as the presence of a balance requirement on whether people used two hands in a bowl-moving task. We found that position and balance had a strong effect on two-handed vs. one-handed strategy, and weight had a weaker effect. We conclude that the choice of strategy is most likely determined by stability requirements and the desire to minimize body rotation and reaching into contralateral space.

1.2.4 In-progress project IV: Poses in a bowl transport task

In the experiment described above, we also recorded grasp pose information, classifying poses into five unimanual poses and three bimanual poses. We present the frequency of these grasps, connect them with existing grasp taxonomies, and report how they were affected by the factors of the experiment (size, weight, balance, and position). We hope the knowledge of how humans choose grasp poses can yield insight on how robots should choose grasp poses.

1.3 Proposed work

1.3.1 Project I: Physically-based levering up

Our observations of grasping in a convenience store setting revealed that people often regrasp items during the process of pulling the object into the hand, because parts of the object they would like to use for a stable final grasp are initially blocked. One common, simple strategy observed was levering up: lifting one edge or corner of an object to expose the underside, and then form a grasp using contacts on that exposed underside. We believe the ability to lever up in this way would be useful for dexterous robots. We propose evaluating a variety of motion controllers to find one that can achieve these motions in a physics simulation.

1.3.2 Project II: Grasping in a virtual kitchen

Previous work on bimanual grasping indicated that the start and goal position of an object has a strong effect on whether people choose to use pure unimanual, pure bimanual, or hand-off strategies. We propose to use this knowledge to animate the motions of a virtual character in a kitchen setting, as they move items to manipulate them, and evaluate whether this knowledge results in more natural character behavior. We focus on position because it was a strong effect and applies to many objects rather than a specific subset (as is the case for the balance factor).

Chapter 2

Review of literature

We review two categories of literature: observations of human manipulation and control of artificial hands.

2.1 Observations of human manipulation

Within this category, there are two main types of research we review here: grasp and manipulation taxonomies and experimental work. 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 grasping and manipulation taxonomies. Another robust line of research investigates how object and task properties affect grasping choices. In this line of research, researchers vary experimental factors and collect information like hand pose or contact points to better understand grasping.

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 [Sch19] and Napier [NT93], 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 [KH70], Edwards et al. [SJEMP02], and Kamakura et al. [KMI+80]. 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 [Cut89], 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 [ZDLRD11] 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. [FPS+09] 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

A number of taxonomies have been developed to express manipulation actions as well. Chang and Pollard [CP09] classify manipulations prior to grasping, with a focus on how the object is adjusted, considering both rigid transformation and non-rigid reconfigurations. Wörgotter and colleagues [WAK+13] discuss how manipulation actions are structured in space and time. Focusing on actions of bringing together and breaking apart, they identify 30 fundamental manipulations that allow sequences of activities to be encoded. Elliott and Connolly [EC84] classify coordinated motions of the hand that are used to manipulate objects, identifying three classes of intrinsic movements: simple synergies such as squeeze, reciprocal synergies such as roll, and sequential patterns such as a rotary stepping motion of the fingers to change contact positions on the object. Bullock et al. [BMD13] encode manipulation instances at a more abstract level, focusing on motion of the hand and relative motion of the hand and object at contact, with the goal of creating a classification scheme that does not assume a specific hand design.

In terms of bimanual manipulation, Guiard laid 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) [Gui87]. Various taxonomies of bimanual manipulation ([GBZ+08, SYN+10]) also start from this framework.

Mechanisms other than grasping and manipulation taxonomies exist for classifying movement, such as Laban Movement Analysis [AL14] and the Facial Action Coding System (FACS) [EF78, CE05]. Observations of great apes are also of interest [Tor85]. For example, Byrne et al. [BB+01] observe over 200 primitive actions, such as pick-out, pull-apart, and rotate-adjust, as necessary to describe feeding behaviors of mountain gorillas.

2.1.3 Experimental studies of manipulation

In this section, we focus on studies of bimanual manipulation in particular, but similar studies exist for unimanual grasping as well.

Studies on infant grasping

There is extensive investigation on when bimanual manipulation skills emerge in infants. Studies such as that of Kimmerle et al. [KFKM10] observe the amount of time very young children spent on different kinds of manipulation (unimanual, symmetric bimanual, and differentiated bimanual). Greaves and colleagues [GIKS+12] 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. In this work, we are interested in the use of bimanual manipulation skills in adults, not the development, although some of the same properties of objects that elicit bimanual manipulation in infants, such as size, weight, stability, and number, may also function the same way for adults.

Effect of object/task properties on grasping strategy

Several experimental studies investigate the effect of object or task properties on large-scale grasping strategy that occurs outside the hand. For example, an experiment by Rosenbaum [Ros08] investigates how the choice to walk to the left or right of a table is influenced by goal position and object position. Participants were asked to grab a bucket from a table before continuing on to one of seven goal locations. Depending on which side of the table the bucket is closer to and how far left or right the final goal position is, the participant can be forced to trade off between reaching across the table for the bucket or taking a less efficient route to the goal. Rosenbaum and colleagues [RCRS10] investigate how goal location affects the usage of left, right, or both hands when grasping or placing for a Tupperware-stacking task.

A study by Cesari and Newell [CN00] investigates how object size and weight influence the number of fingers or hands used. By having participants grasp cubes of different sizes and densities, they fit an equation to describe the transition between four- and five-finger grasps and one- and two-handed grasps. Choi and Mark [CM04] extend Cesari and Newell's cube study to investigate how object distance and weight affect reaching modes (using the arm and shoulder only, using the torso to reach, and standing to reach).

2.2 Control of artificial hands

A major application of manipulation research is using it to design artificial hands. In this section, we review work that has gone into the control of robotic hands as well as the hands of virtual animated characters.

2.2.1 Control of robotic hands

One line of research in robot hand control is implementing reusable manipulation primitives. A tiny sample includes examples ranging from pushing [Mas86], toppling [Lyn99], pivoting [YPL+10], opening doors [KSN10], and moving objects out of the way [SK05] to making pancakes [TKPB11], making cookies [BR], and folding towels [MSCTLA10].

In the area related to levering up when grasping, Trinkle and Paul [TP90] show how to use an object's geometry to analytically find contact points such that applying inward force at these points (squeezing) will cause one of the corners of the object to break contact with the surface it's resting on. The result is the ability to lever an object up into a wrap grasp. However, this work does not regrasp the object into a desired configuration after levering up.

2.2.2 Grasping in virtual characters

Getting virtual characters to interact with their environment in a believable way is an ongoing challenge.

Mordatch et al. [MPT12] use an optimization that includes continuous variables representing the state of contact in order to create hand motion sequences given initial and goal states for the hand and an object. From the motion of the wrist and object, Ye and Liu [YL12] search a space of feasible contact points in order to synthesize a variety of detailed hand motions. Using a Kinect depth camera and a database of ten prerecorded grasping motions, the physics-based controller of Zhao et al. [ZZMC13] generates grasping motions in realtime.

In addition to prehensile grasping, work in the animation of hands has also gone into generating finger gaiting behavior for in-hand manipulation of objects [AK13] and generating non-prehensile hand gestures [JHS12].

Chapter 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 [FPS+09]. 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 [ND03], which has been long developed to describe motion and action, but does not focus on grasping. They also may tie into existing impedance [Hog85] and operational space controllers [Kha87] used in robotics.

This chapter describes our complete process, our annotation scheme, highlights from the full database (viewable online [LFNP14a]), 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 [KKF+97, CWDH88, LHW+94, PMCS06]. 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. [CDC+09] 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 [KKF+97]. 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. [FPS+09, FPS+]. 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://www.cs.cmu.edu/~jialiu1/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. The object related property is another factor that influences the hand shape and motion, but these relationships are not made explicit in our database. In contrast to traditional taxonomy research, which (aside from [WAK+13]) 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 [FPS+] and realized that a wide variety of tasks belonged to one grasp type but involved very different motion.

3.2.1 Hand shape

Our classification of hand shape comes directly out of [FPS+], combined with ideas out of [Nap56]. In [FPS+], 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 [Nap56].

For example, the type can be power grip or precision grip, or intermediate which is in

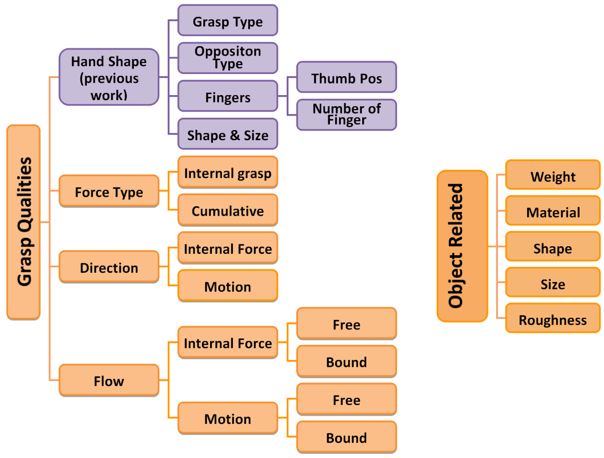


Figure 3.1: Simple Classification of the Database

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).

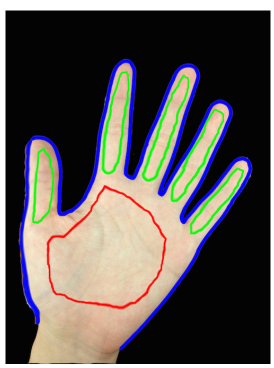


Figure 3.2: Palm, pad, side

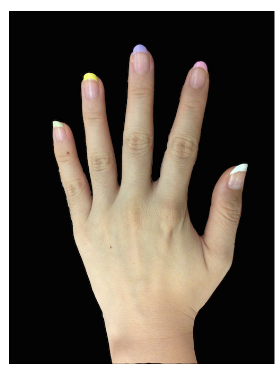


Figure 3.3: Back

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.

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 [Nap56].

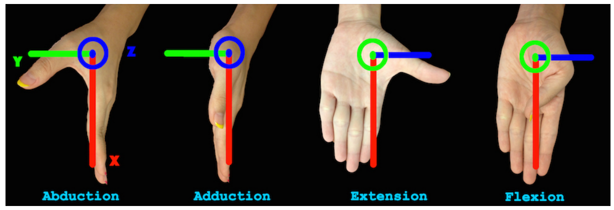


Figure 3.4: Local coordinates of all the types (left hand)

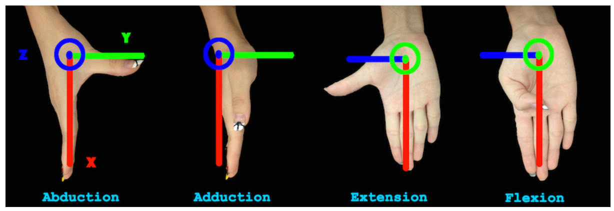


Figure 3.5: Local coordinates of all the types (right hand)

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).

Table 3.1: Force Type Definitions

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 3.2: Internal Force Examples

Example		
Force Type	Squeeze	Hold
Annotation	Squeeze toothpaste	Hold a pan

Table 3.3: Cumulative Force Examples

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

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

Table 3.4: Grab vs. Hold

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

Table 3.5: Direction Examples

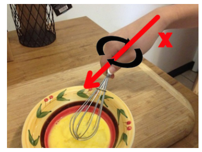
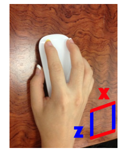
Property	**Possible Values**	**Example**
Direction Subspace	along x/y/z axis	Table 3.6 I
	rotate around x/y/z axis	Table 3.6 II
	plane xy/xz/yz	Table 3.6 III
Coordinate Frame	hand	Table 3.7 I
	global	Table 3.7 II
	object	Table 3.7 III

is halfway between the x- and y-axes (e.g., Table 3.12, second row). Directions that are less constrained or more 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

Table 3.6: Axes Examples

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

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 local 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.

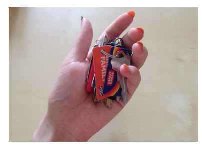
Object coordinates: Finally, occasionally the local coordinates of the object must be used since, in some motions, the object shape decides the direction of motion. If the object is a long stick or string type, we define the direction along the stick to be the x-axis. If the object is rectangular in 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.

Table 3.7: Coordinate Frame Examples

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 3.8: Flow Factor Examples

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

3.2.4 Flow

The effort factor we use here is flow. Flow comes from the Laban Effort / Shape notation [SBGK13]. 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).

Table 3.9: Weight of Object Examples

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

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).

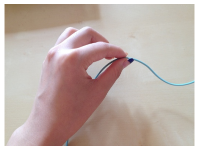
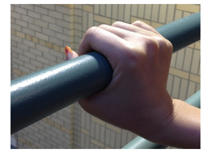
The material of the object also strongly affects grasping strategy. For example, grabbing highly deformable material requires continuous adjustment of grasp shape as the object changes shape. Another example of the effect of material is that people will grab raw meat differently than paper.

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 6.

Table 3.10: Shape & Size & Roughness of Object Examples

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

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. [FPS$^+$], 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 describe 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 [LFNP14a]. 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,

Table 3.11: Large Diameter and Lateral Grasp

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

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

Table 3.12: Large Diameter Cylinder Grasp Examples

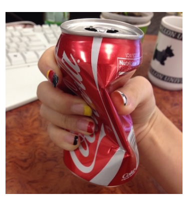
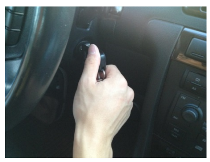
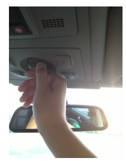
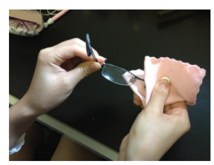
Example		
Force Type	Hold	Hold
Motion Dir	around y axis (hand)	-
Force Dir	-	-z (global)
Flow	Free Motion/ Bound Force	Bound Force
Annotation	Drink water	Hold a bottle
Example		
Force Type	Hold	Grab&Press
Motion Dir	x(45)y (hand)	-
Force Dir	-	z (global)
Flow	Free Motion/ Half Bound Force	Bound Force
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 3.13: Lateral 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

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), 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

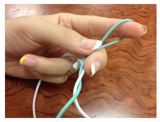
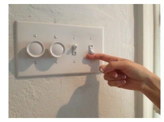
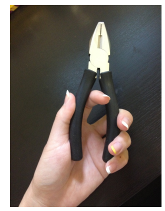
From our observations, the existing taxonomy that served as our starting point [FPS+] has covered many types of grasps. However, there exist some actions which are not represented by their taxonomy, for which we have created new categories in the database. Some of the new entries involve deformable objects. 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 [LFNP14a]. Some, but not all of the new grasp types can be found in other taxonomies, such as those of Kapandji [KH70] and Buckland et al. [SJEMP02].

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 [KMI+80], 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.

Table 3.14: New Type Examples

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

It is interesting to note the similarities between our annotations and the parameters required for impedance control [Hog85] or operational space control [Kha87], 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) [ND03] 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.

Chapter 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. 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 an ad hoc analysis that attempts to note every significant event in the recording.

There is prior and ongoing work that observes human manipulation using better instrumentation, so the natural question is to ask whether such light instrumentation has any advantages. One advantage is that we can observe humans "in the wild". It is easy to observe behaviors, including unanticipated behaviors. It is easy to study a broader range of behaviors. The naked eye is sometimes better than a microscope.

The disadvantage is that the data 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. In fact one motive of this work is to provide the ground truth to support development of automated analytical tools.

The primary long range goal of this work is to develop a database of manipulation behaviors. We should articulate this goal as carefully as possible, to guide our analysis of human manipulation behavior. The database should support both *descriptive* and *prescriptive* processes. By *descriptive*, we mean that an observed behavior can be mapped to description: a composition of behavioral elements already present in the database, along with contextual information and parameters. By *prescriptive*, we mean that such a description can be mapped to the available robotic hardware so that the originally observed behavior is reproduced, at whatever level of detail is desired and supported with available hardware. It should be the aim of our database, and perhaps of the entire field of robotics, that our database is detailed enough and expressive enough to close the loop from behavior, through description and prescription, to behavior.

We are using the term “database” but it implies a notational system and a classification system as well.

The short range goals are two: to analyze and learn from a particular dataset; and to apply lessons learned to develop a notation system to support the systematic development of the manipulation database.

This particular dataset is of a single human subject in a lunch store. The greatest surprise 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 obtaining a desired grasp pose in the presence of clutter is complex, varied, and time consuming.
- Contact-guided motion is common.
- The subject often adjusted the grasp just prior to lifting the object, either from precision to power grasp, or a more subtle change in disposition of fingers.
- There are numerous collisions, between effector and clutter, and between the target item and clutter. 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 a third party 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, 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 hogie; a pasta salad and bowl; a plastic salsa cup with lid; and steel tongs.

A compressed version of the dataset is available online at `http://www.cs.cmu.edu/~dtroniak/nsh_shop_120.webm`.

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 dataset was labeled using several existing taxonomies, as well as through other lenses where a taxonomy does not exist. The annotations were focused around the following:

- Static grasp pose taxonomies created by Cutkosky [Cut89] and Kamakura et al. [KMI+80]. These taxonomies separate different hand shapes used during grasping based on form, function, and the pattern of contacts on the hand.
- Intrinsic (within-hand) hand motion categories observed by Elliott and Connolly [EC84], which describe the small motions a hand uses to manipulate an object already in the hand.
- Bullock et al.'s manipulation taxonomy [BMD13], which creates broad categories of ma-

nipulation based on the presence or absence of contact, prehension, motion, intrinsic hand motion, and motion at contact points. The taxonomy is high-level and doesn't assume any particular hand morphology.

- The lens of errors and recovery from errors.
- General observations of recurring manipulation primitives, similar to gorilla studies.

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 stage in grasping do certain manipulation categories correspond to; what was the hand attempting to do when the error occurred; and what are the recurring strategies involved in grasping?

From all the actions in the video, we selected seven to examine in more detail. These actions contain grasp poses and manipulation actions from the taxonomies we used, but also things that fall outside of the taxonomies. We are in the process of plotting these old and new annotations on a timeline.

4.3 Preliminary results

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.

In practice, the process of forming a grasp is complex. While it's simple to pinch small items between two of 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. 4.1 is an example of how and why manipulation occurs after the hand makes contact with the object but before a final grasp is attained: the object is lifted by the ulnar fingers using the rim of the salad box, exposing the bottom surface (frame 1). Then there is a complicated sequential pattern of finger lifting and recontacting (frames 2-5) that 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 hand in anticipation of grasping

Figure 4.1: Grasping a salad box. Before the final grasp is formed, there is often a complicated manipulation sequence to regrasp the object.

Figure 4.2: An example of grasping. The hand makes contact with the top of the milk bottle and levers it out while simultaneously closing the fingers around the cap to form a final grasp.

2. Contact: compliantly making contact with some part of the object
3. Dealing with clutter: maneuvering fingers into spaces, singulate object, or push 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

For small, light, or unobstructed objects, some of these phases may not be necessary. Also, another thing to note is that sometimes these phases overlap. For example, in Fig. 4.2, tilting out the milk bottle simultaneously deals with clutter (separates the bottle from neighboring ones) and transfers a portion of the bottle's weight into the hand.

In addition, this movement is an example of how the environment is used to manipulate the object before (in preparation for) making a stable wrapping grasp. For example, the flat hand used to rock the bottle out (Fig. 4.2, frame 3) is not a stable grasp found in taxonomies, but works with the environment to nevertheless manipulate the bottle and bring it into the hand.

4.4 Proposed work

This work is not yet complete. We plan to take seven of the motions in the video and create a timeline of annotations made using the taxonomies we picked out. These visual images will illustrate the process of grasping and placing seen through the lenses of these taxonomies, as well as illustrate how many annotations fit within the taxonomy or are new categories.

We plan to collect numerical summaries of the annotations as well, presenting the most popular grasp poses, in-hand manipulation tasks, or manipulation action.

Finally, we plan to discuss the frequency of errors and how they were recovered from.

4.5 Conclusion

Observing video of a motion – especially at the increased time resolution of 120 fps – reveals the complexity of the grasping process. In this work, we sought to observe the whole grasping process in a natural setting and to analyze how well existing manipulation taxonomies cover and describe the elements of grasping. As levering up (and down) were commonly used in the video, we hope that implementing such a skill would be helpful for artificial graspers (Ch. 7).

Chapter 5

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, [KMB03] 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 [Gui87] 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 ([GBZ+08, SYN+10]) based on Guiard's analysis [Gui87] 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 bimanual actions as each hand having its own role.

When grasping and moving objects, people have the option to use one or two hands. In particular, there are two bimanual strategies that people may make use of: handing off, which is when the object is grasped with one hand, but then transferred to the other hand to be placed, and symmetric bimanual grasping, when both hands are used to form a single grasp of the object. The existence of multiple ways of accomplishing the same task raises the questions of how often various strategies are used and how people pick between them. In order to answer these questions, we performed an experiment using different sizes and weights of bowls, difficulty of balancing, and starting positions for the subject to test how these factors affect the frequency of unimanual and bimanual grasping strategies. Our results give an indication of how often bimanual strategies are used, why they are chosen, and how much object and task factors affect the frequency of bimanual strategies in adults.

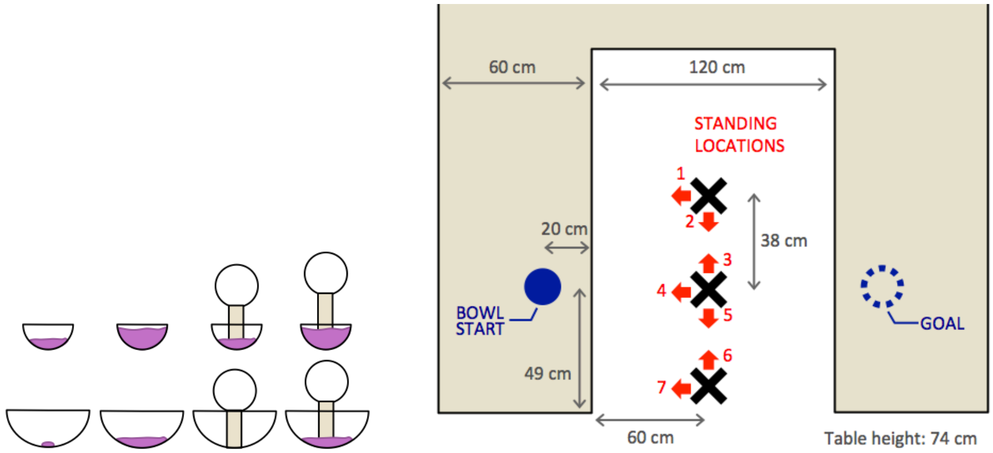
5.1 Experiment 1

5.1.1 Measures and Hypotheses

The 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 balance was important) and position (the position of the bowl relative to the subject). We collected which hand(s) subjects used to pick and place the bowl.

Our hypotheses were as follows:

1. Size: Larger object size would encourage the use of two-handed strategies. Two hands/arms can function similarly to a large manipulator ([BMD13]), so we expect more bimanual usage as the object gets larger.
2. Weight: Heavier object weight would encourage the use of two-handed strategies. Using both arms can spread the load to a more comfortable level at each arm and can also reduce the torque/moment by supporting the object at two places.
3. Balance: The presence of a balance requirement would encourage the use of two-handed strategies. Like using two hands on the handlebars of a bike, it may be easier to make small adjustments by controlling the difference between two forces than to modulate the magnitude of a single force.
4. Position: Start and goal position would affect use of both hands in hand-offs. Handing off may be a way to make use of the reach of both arms without needing to move.



(a) (b)

Figure 5.1: Experimental setup. (a) Bowls can be small or medium; light or heavy; and with or without a balance tube. (b) Participants stand at each of the seven starting positions/facing directions indicated.

5.1.2 Method

Participants

We ran an experiment with 17 participants (5F, 12M; 1 left-handed, 14 right-handed, 2 mixed-handed; mean age = 27.5 (SD = 6.7)). The method was approved by the Disney Research Institutional Review Board.

Procedure

Each participant performed 56 tasks that consisted of moving a bowl from one table to another, where the subject's starting location and facing direction varied. There were 8 different types of bowls and 7 different places/directions that the subject could stand (Fig. 5.1).

The bowls moved were 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 smaller bowl was 12.2 cm x 6.1 cm (diameter, height), while the larger bowl was 20.2 cm x 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 x 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, and a packet of stones added to maintain the above-listed weights.

There was one trial per condition, resulting in 56 trials overall (2 size $\times$ 2 weight $\times$ 2 balance $\times$ 7 directions). Tasks sharing a facing position/direction were presented consecutively in a block to avoid making the subject move around after each trial. Otherwise, the presentation of the blocks and tasks within each block were randomized.

After the trials, hand measurements were collected from the subjects: hand length and width, finger length, middle finger-thumb grasp diameter, and grip strength. Grip strength was measured using a Camry electronic hand dynamometer. One trial was recorded with participants asked to squeeze as hard as they could.

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. The nine possibilities are:

- 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 $\rightarrow$ r) – Hand-off from left hand to right (pick up with left, place with right)
- RL Hand-off (r $\rightarrow$ l) – Hand-off from right to left
- LB Left $\rightarrow$ bi – Pick with left hand, add right to place bimanually
- RB Right $\rightarrow$ bi – Pick with right hand, add left to place bimanually
- BI Bimanual – Pick up, transport, and place with both hands
- BL Bi $\rightarrow$ left – Grab bimanually, place with left hand only
- BR Bi $\rightarrow$ 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[1] and the end was the second when the bowl made contact with the goal table. Duration was calculated as the time in between.

To calculate rotation, first, facing directions of the hip at transport start and end were recorded, rounded to the nearest 45°. For clockwise turns, the end angle was annotated as negative while counter-clockwise turns were annotated as positive. Rotation was defined as the positive difference between the angle of the initial facing direction and the angle of the start of bowl transport, plus the positive difference between transport start and end angles.

Note that there was some ambiguity in categorizing grasp strategy: the second hand sometimes floated near the object, or briefly touched it before dropping away. Because contact was non-existent or brief, these cases were annotated as unimanual motions.

Data Analysis

The analysis in this chapter is in the process of being redone. Instead of using repeated-measures ANOVA, a generalized linear mixed model with a logistic link function was fit to the data using R's lme4 package and glmer function. The response variables analyzed were usage of bimanual, hand-off, and unimanual strategy (three separate analyses with binary outcomes). Size, weight, balance, and position were used as fixed effects in the model. Variation between participants was modeled as a random intercept (without random slopes). A stepwise procedure comparing likelihood ratios was used to eliminate non-significant variables until no more could be removed.

In addition, in order to understand the reason behind people's preference of certain strategies over others, the effect of strategy (bimanual, hand-off, and unimanual) on transport duration and body rotation were investigated using linear models with either duration or rotation as the response, and strategy as the fixed effect (with the four experimental variables added to the model as well to control for them). Participants were modeled using random intercepts. P-values were calculated using the likelihood ratio between the model including strategy and the model without it.

The pattern of the left-handed participant's data differed from that of the other participants; as such, this participant's data was discarded when performing analysis. Their data is

[1] 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.

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

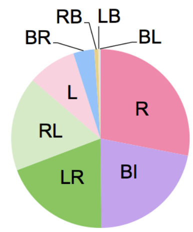


Figure 5.2: Frequencies of each strategy.

not represented in the following results.

5.2 Experiment 1 Results

5.2.1 Strategy frequency overview

Frequencies of grasp strategies are summarized in Fig. 5.2. All strategies were observed at least once, but strategies that involved switching from one to both hands or vice versa were rare.

5.2.2 Effect of experimental variables on grasp strategy

For all three strategies – bimanual, hand-off, and unimanual – balance and position remained in the model. In addition, the balance $\times$ position interaction effect remained in the unimanual model (χ^2(6)=48.78; p < .0001).

Balance as a main effect was significant in bimanual (χ^2(1)=234.5, p < .0001) and hand-off (χ^2(1)=127.4, p < .0001) strategies, but not the unimanual (χ^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. At positions where the unimanual strategy was popular (P1, P4, and P7 – positions involving moving the bowl from front to back), the balance requirement cut down unimanual usage. At the other four

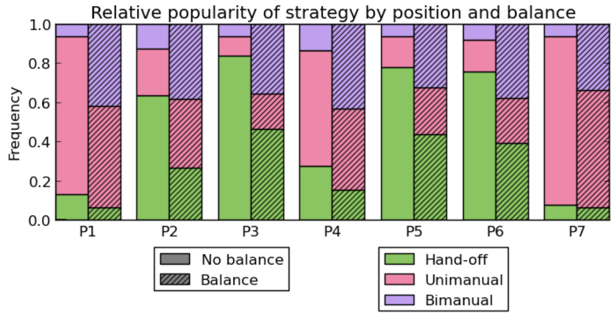


Figure 5.3: Popularity of the three main strategies by position and balance.

positions (positions involving moving the bowl left to right or vice versa), however, unimanual usage increased in the balance case.

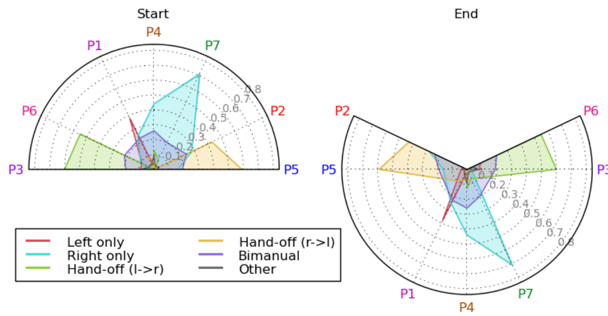
The three strategies were also affected by position (bimanual: χ^2(6)=18.0, p = .006; hand-off: χ^2(6)=370.7, p < .0001; unimanual: χ^2(6)=257.4, p < .0001). Hand-offs were the strategy people used most often at P2, P3, P5, and P6, which involved moving the bowl from left to right or vice versa. The unimanual strategy was used most at P1, P4, and P7, which are the three positions where the bowl is moved from front to back. Fig. 5.3 summarizes these balance and position effects.

Fig. 5.4 also shows balance and position effects, but separates the unimanual strategies into L and R, and the hand-off strategies into LR and RL. It also illustrates the angle at which the bowl starts and where it ends, relative to the participant. For example, at P3, the bowl starts out directly to the left of the participant and is moved to the participant's right. At this position, the hand-off left-to-right (LR) strategy is the most common strategy (used about 60% of the time) followed by the bimanual strategy.

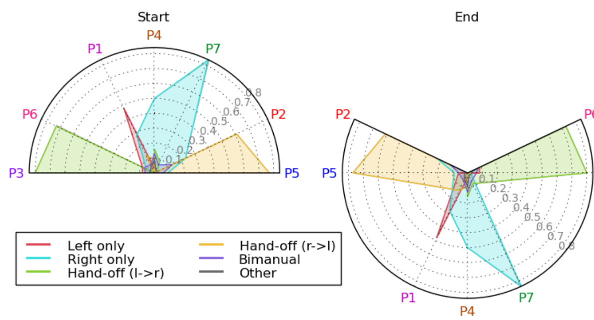
Neither size nor weight were significant in any of the models.

5.2.3 Reason for strategy choice

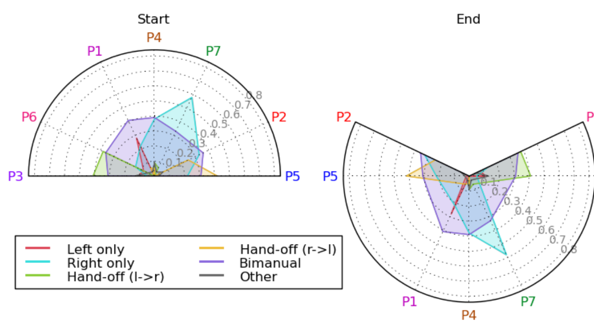
Although some participants use the bimanual strategy as a dominant strategy, its overall usage even in the balance case, only reaches 33%. One possible explanation for this is that the other strategies have some other benefit and so can still be appealing even when balance is a factor.



(a)



(b)



(c)

Figure 5.4: 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.

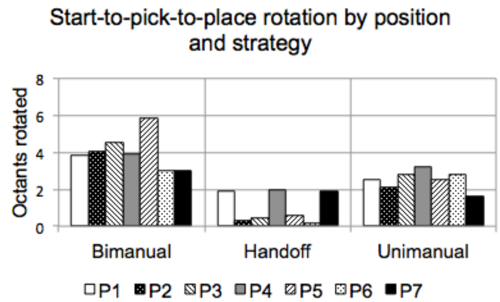


Figure 5.5: Total rotation by position and strategy

The first hypothesis we tested was that hand-off and unimanual strategies were quicker. However, an analysis of whether different strategies have an effect on duration controlling for the experimental factors showed that the effect of strategy on transport duration was only barely significant (χ^2(2)=6.5; p=.039).

The second hypothesis we tested was that hand-off and unimanual strategies require less rotation. Here we found the effect of strategy on rotation was very strong (χ^2(2)=473.5; p < .0001). Bimanual strategies require much more rotation than unimanual strategies, which require much more rotation than hand-off strategies.

Moreover, as one might expect, position strongly affects the rotation needed at each strategy by different amounts. Fig. 5.5 breaks down the average amount rotated at each position by strategy. Although the amount of rotation at each position maintains the ordering of bimanual > unimanual > hand-off (with the exception of P7), the spikes at P1, P4, and P7 for hand-offs could be related to the low popularity of hand-offs at those positions.

5.3 Preliminary observations

Neither size nor weight showed the effect on bimanual usage we expected. However, balance had a strong effect in line with our hypotheses. When balance was needed, hand-offs were used less while the bimanual strategy became more popular. Balance had a mixed effect on unimanual strategy, evening out its usage at positions where it was normally used very often or not often. Together, these effects suggest that the unimanual strategy is between the other two strategies in terms of stability. During balance cases, it “steals” some uses from the hand-off strategy, but loses some to the bimanual strategy.

Position had a strong effect on hand-off and unimanual usage, but not bimanual usage. The unimanual strategy was the most popular of the three strategies when moving the bowl from front to back, while hand-offs were most common when passing from left to right or vice versa. By contrast, usage of the bimanual strategy did not seem to vary as much with position, which is similar to results found by Rosenbaum et al. [RCRS10].

The desire to minimize body rotation seems to be a part of strategy choice. Hand-offs are most popular at the four positions where minimal rotation is required, and unimanual is popular at the other three positions. The bimanual strategy requires the most rotation.

In this experiment, we found a strong effect of position and balance on grasp strategy. However, we didn't find the effect we expected of greater size and weight on bimanual strategy. One possibility is that the bowls were not sufficiently large or heavy enough to encourage bimanual usage. We investigated this possibility in a second experiment.

5.3.1 Proposed work

The analysis of the effect of strategy on rotation and duration needs to be re-run. In addition, plots of all results with standard error bars need to be generated.

5.4 Experiment 2

5.4.1 Measures and Hypotheses

The focus of this experiment was to test if weights and sizes larger than the ones previously investigated can elicit a size/weight effect on bimanual usage. Four bowl sizes and three weights were used.

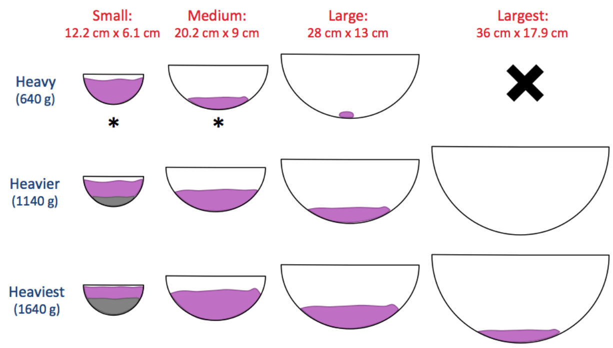


Figure 5.6: The eleven different size/weight combinations for the bowls. The heavier small bowls were padded with lead to increase density. The two bowls marked with an asterisk are repeated from Experiment 1.

5.4.2 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.

Procedure

Each participant performed one trial for each of the 66 conditions tested in the experiment. Each trial consisted of moving a bowl from one table to another. There were 11 size/weight combinations for the bowls (Fig. 5.6) and 3 different places/directions that the subject could stand (Fig. 5.7). There were 2 balance/no balance conditions as in Experiment 1. The entire procedure took around 30 minutes.

Two more IKEA® BLANDA BLANK bowls were added: a large bowl (28 cm x 13 cm (diameter, height), 600 g), and largest bowl (36 cm x 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 bowl were achieved with sealed bags of lead at the bottom.

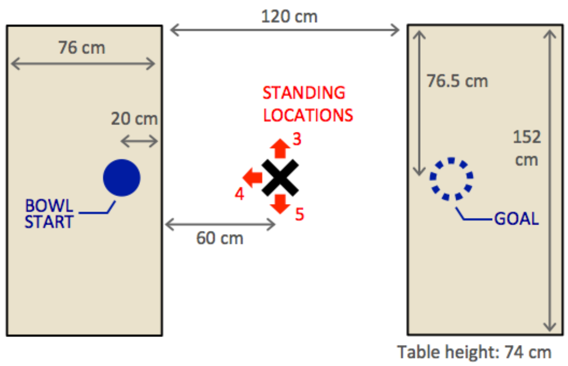


Figure 5.7: Experiment 2 setup with only one starting location (with three facing directions).

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. 5.8), 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 same starting positions clustered together.

The procedure was otherwise identical to the first experiment.

Data Processing

Motion capture data was 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 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 if the facing angle at a snapshot was positive (counter-clockwise from the bowl) or negative (clockwise from the bowl). Hip orientation at the start and end were then used to calculate rotation as in Experiment 1.

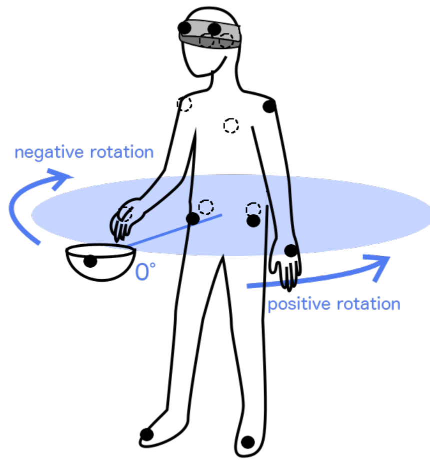


Figure 5.8: Motion capture setup for Experiment 2. The black and dotted circles represent the placement of 16 reflective markers on the front and back of the participant. As in Experiment 1, the direction of the bowl is defined as zero degrees (the direction of the goal is $\pm 180°$) and counter-clockwise rotations are positive angles.

Data Analysis

Analysis was similar to Experiment 1. Grasp strategy and individual pose frequencies were analyzed using a generalized linear model.

5.5 Experiment 2 Results

5.5.1 Basic strategy frequencies and comparison to Experiment 1

Strategy frequencies are summarized in Fig. 5.9. Unlike in Experiment 1, the bimanual strategy (BI) was the most popular strategy.

Compared to Experiment 1 (Fig. 5.2), bimanual usage in Experiment 2 was much higher while hand-off usage was cut down. We can limit the examination to only trials featured in both experiments. These are all three positions 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. 5.10), despite the task being the same.

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

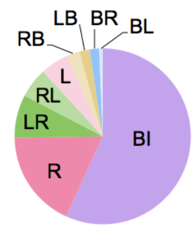


Figure 5.9: Frequencies of each strategy.

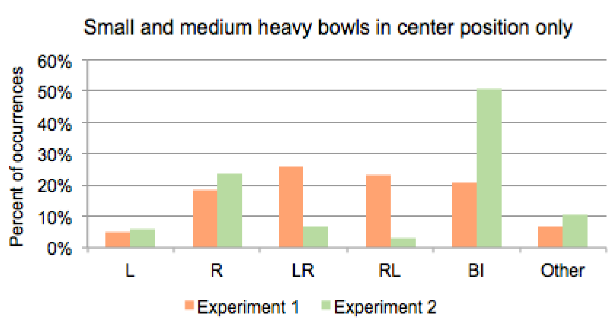


Figure 5.10: Comparison between Experiment 1 and 2 on identical trials.

5.5.2 Effect of experimental variables on grasp strategy

These analyses are in the process of being re-run using generalized linear models.

5.6 Preliminary observations

Our results showed that strategies were strongly influenced by the position of the object/goal and the presence of a balance requirement and less strongly influenced by object weight.

Although bimanual strategy was slightly affected by object start/goal position, as seen by the significant effect found in Experiment 2, this effect of position was not as dramatic as it was for the unimanual and hand-off strategies. This relative position-invariance of bimanual strategy is in line with findings from an earlier study of grasp strategy on Tupperware containers ([RCRS10]).

Our data contained left/right asymmetry in unimanual reaching (e.g. the difference between “left only” and “right only” spikes in Fig. 5.4a), indicating a preference to use the right hand, even if it requires reaching into contralateral space or rotating. This bias toward right-handed reaching matches that found in handedness studies ([GFS14]).

Avoiding reaching across the body and minimizing rotation could be behind the different popularity of strategies by position. Thus, hand-offs become popular when the start and goal position are in different left-right hemispheres as this strategy minimizes contralateral reaching and rotation.

The smallest bowl was often unusual. Part of this may be due to size – in particular, the smallest bowl is smaller than the human hand, possibly opening up a wider variety of feasible grasp poses, which might also influence strategy. However, small bowl’s effect on strategy be due to the fullness of the bowl, a factor not considered in this study.

Finally, 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). This could possibly suggest that the wider set of tasks is capable of influencing people’s default strategy, or it is possible that the changes to the experimental method are responsible. Even within the first possibility, it’s not clear whether greater weights, sizes, fullness, or uncertainty is what is behind the greater bimanual usage. Further investigation is needed.

5.7 Proposed work

For Experiment 2, all analyses must be re-run using linear models, and graphs with error bars must be generated. This includes an analysis of the effect of strategy choice on rotation and duration.

Chapter 6

Poses in a bowl transport task

In the experiments detailed in the previous chapter, we also collected 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).

6.1 Pose classification

Grasp poses were sorted into eight different possibilities based on contact areas and fingers used (Table 6.1).

Poses seemed to be a somewhat continuous space making differentiating between certain pairs of grasps (A vs. E and sometimes B vs. C) ambiguous. If there were at least two finger pads used in the grasp, the trial was annotated as E rather than A. If it appeared the finger pads were being used more than the pad of the palm, the grasp was annotated as B rather than C.

Table 6.1: Codes and descriptions for bowl grasp poses. Comparisons are from Feix et al. [FPS+09] and Liu et al. [LFNP14b] / Ch. 3

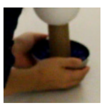
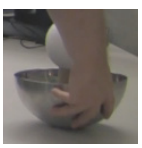
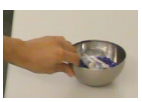
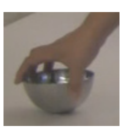
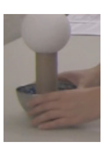
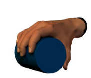
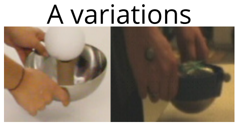
Grasp	Description	Comparison
 A	Lateral pinch – Side of middle finger is the main thing holding the bowl. Thumb hooked on inside of bowl. Index finger extended away from other fingers. Alternatively, the bowl rim can be pinched between thumb and side of index finger. Can also be used bimanually.	 Lateral
 B	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 grasp E.	 Parallel ext.
 C	Cupped bimanual – Palm is the main thing holding the bowl. Fingers cupped and horizontal (cf. B'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
 E	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 bimanual grasp B.	 Extension
 F	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
 G	Overhand (precision disk) – Grab bowl from the top. Uses pads of thumb and all fingers.	 Precision disk
 H	Ring bimanual – 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
 I	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.	??

Table 6.2: Variations on grasp poses.

A variations 	Grasp A can also be bimanually (both images) or by pinching the bowl between thumb and index finger (right).
B variations 	Grasp B encompasses a wide range of poses. The thumb in grasp B can be lightly hooked, not in contact (left), on the outside of the bowl (center), or fully inside the bowl (right).
F variation 	Any grasp involving pinching with the index finger is counted as grasp F including this grasp between index and middle finger.
H variation 	Bimanual grasps can be used unimanually. This is what grasp H looks like when used with one hand.

6.2 Preliminary results

The frequency of the grasp poses for Experiment 1 and 2 are summarized in Figs. 6.1 and 6.2. The extension grasp and open bimanual grasp are the most frequently used unimanual and bimanual grasps.

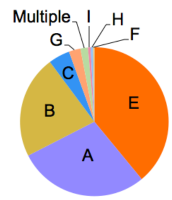
A surprising result is the amount of variety shown in the poses used to grasp the bowls. Rather than there being a single one-handed grasp and a single two-handed grasp, five unimanual and three bimanual poses made an appearance in the experiments.

We are planning on doing an analysis of the effect of the experimental factors (size, weight, balance, and position) on choice of grasp pose. From initial analyses, it seems likely the physical properties of the bowl (size, weight, and the presence of the balance tube) likely affect the choice of grasp while position doesn't.

Initial results also indicate that the open bimanual, cupped bimanual, and extension grasps tend to be used at all sizes, while the other grasps tend to be used on the smallest size bowl in particular.

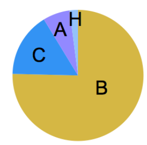
All grasps

	Pose	Frequency	
A	Lateral pinch	416	28%
B	Open bimanual	327	22%
C	Cupped bimanual	68	5%
E	Extension	569	39%
F	Index hook	4	<1%
G	Overhand	37	3%
H	Ring bimanual	8	<1%
I	Closed lateral pinch	8	<1%
	Changing grasps	24	2%



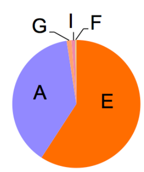
Bimanual grasps

	Pose	Frequency	
A	Lateral pinch	26	7%
B	Open bimanual	294	75%
C	Cupped bimanual	62	16%
H	Ring bimanual	8	2%



Hand-off grasps

	Pose	Frequency	
A	Lateral pinch	251	38%
E	Extension	387	59%
F	Index hook	2	<1%
G	Overhand	8	1%
I	Closed lateral pinch	6	<1%



Unimanual grasps

	Pose	Frequency	
A	Lateral pinch	128	39%
E	Extension	170	51%
F	Index hook	2	<1%
G	Overhand	29	9%
I	Closed lateral pinch	2	<1%

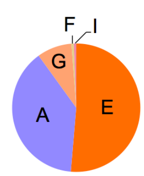
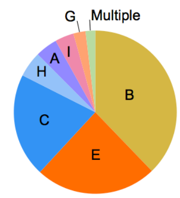


Figure 6.1: Frequencies of grasp poses in Experiment 1.

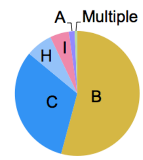
All grasps

	Pose	Frequency	
A	Lateral pinch	82	4%
B	Open bimanual	706	38%
C	Cupped bimanual	384	21%
E	Extension	450	24%
G	Overhand	45	2%
H	Ring bimanual	95	5%
I	Closed lateral pinch	70	4%
	Changing grasps	36	2%



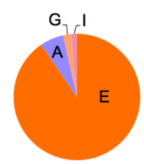
Bimanual grasps

	Pose	Frequency	
A	Lateral pinch	18	2%
B	Open bimanual	650	54%
C	Cupped bimanual	382	32%
H	Ring bimanual	83	7%
I	Closed lateral pinch	59	5%
	Changing grasps	8	<1%



Hand-off grasps

	Pose	Frequency	
A	Lateral pinch	17	6%
E	Extension	257	90%
G	Overhand	6	2%
I	Closed lateral pinch	4	1%



Unimanual grasps

	Pose	Frequency	
A	Lateral pinch	24	10%
E	Extension	166	68%
G	Overhand	39	16%
H	Ring unimanual	8	3%
I	Closed lateral pinch	6	2%
	Changing grasps	1	<1%

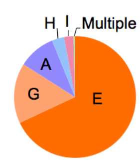


Figure 6.2: Frequencies of grasp poses in Experiment 2

6.3 Proposed work

Analyses of the effect of experimental factors (size, weight, balance, and position) on grasp pose, similar to the one in the previous chapter, have to be rerun, and important results presented in graph form with error bars.

6.4 Discussion

We performed an experiment featuring simple, lipless bowls of four different sizes and four different weights. We recorded the poses that were used, categorized them, and noted variations and poses that fall in between different categories.

Our goal is to also analyze how the experimental factors influenced the choice in grasp pose. We hope that the connection between these factors and choice of grasp poses will help to create artificial graspers that can use object properties to select appropriate poses.

Chapter 7

Proposed work: Physically-based levering up

When picking up objects, the final, secure grasp people intend to use is often not actually feasible in the object's initial pose. 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. 7.1).

In order to deal with this challenge, people will often lever up the phone to expose this surface and regrasp. This action is done very quickly without people being aware they are doing it. The ability to grasp objects by levering up would be a useful addition to robots' toolkits. In this project, we aim to create a motion controller that allows a virtual hand to perform these kinds of levering up motions in simulation.

There are many ways to design a motion controller. What design choices will achieve the goal of creating feasible levering up motions? In this work, we seek to answer some of the following research questions:

- Which is more robust for actions like levering up: force control or position control? Humans grasp compliantly, and compliant force control tends to be more appropriate and safe when a robot hand interacts heavily with an environment, so we hypothesize force control will work best for levering up. Alternatively, position control or a scheme using both position and force control may yield better results for levering up.
- For accomplishing these tasks, is it important to know the surface normals of points in

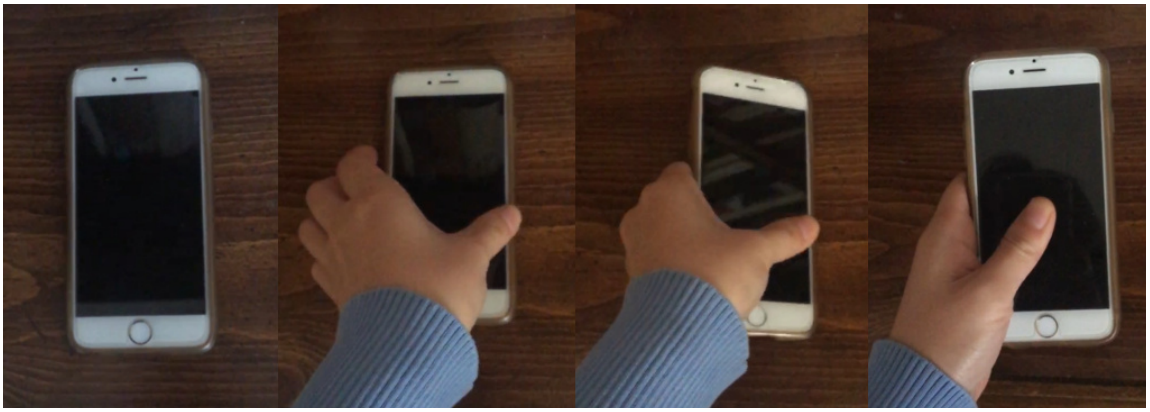


Figure 7.1: When a phone is lying flat on a table, its bottom surface is blocked (far left). However, contacts with this bottom surface are needed in the desired grasp (far right). To achieve this grasp, intermediate levering up is performed.

the workspace? Knowing the normals of contact points might be necessary to successfully lever up. Alternatively, this action might be possible even with uncertainty about the environment.

- Is it better to optimize the entire motion all at once, or is it better to break down the motion into predefined phases and optimize each individually? Breaking down a levering up motion into smaller, simpler phases – for example, an initial phase where the thumb secures one edge of the object (either by pressing in or down on one side), a second phase where opposing fingers lift up the other end of the object, a third phase where the object is regrasped by sequentially establishing contacts on the newly exposed surface, and a fourth phase where the object is lifted up – may make it easier to design the controller needed to accomplish levering up. We hypothesize a multi-phase controller will be effective in levering up. Alternatively, levering up may be accomplished with fewer phases, or even just one phase.
- Observations of humans indicate they take turn flexing different fingers while levering up, in case one fails to make contact with the object. Is this kind of error recovery strategy key to levering up successfully? We hypothesize incorporating this error anticipation strategy into the virtual hand's control algorithm will help us achieve our levering up goals.
- What contact model is needed to accomplish these motions? Complex contact models are more physically realistic, but complicate optimization. We hypothesize these motions can be accomplished using a simple rigid-body contact model with friction cone approximation. Alternatively, more complicated friction model involving soft contacts

may work better.

7.1 Method

We will start with a simple system: one test case (the phone picking up example above), a motion capture example of a human performing the action, and several different control schemes that control the joint torques of a virtual hand so as to accomplish the levering up.

For the example of human levering up, we already have a levering up motion captured using marker-based technology. This motion will guide an initial levering up motion in a physics simulation.

We will implement a control scheme with the following features:

1. Compliant force control
2. Four predefined phases of motion: securing with thumb, levering up with finger, form grasp with other fingers, and final grasp adjustment after lifting. The motion at each phase will be based on mocap example.
3. Rigid body contact model with friction cone approximation

The control scheme will be formulated as an optimization problem, and will be implemented using a pre-existing physics simulation that has a Shadow Hand model as the hand, PD-control, and uses CMA optimization. The control scheme will be altered if it fails to be able to lever up the phone, and the success and reliability of the alternative control schemes will be reported and compared.

Chapter 8

Proposed work: Grasping in a virtual kitchen

The work on bimanual grasping in the previous chapters indicated that the start and goal position of an object has a strong effect on whether people choose to use pure unimanual, pure bimanual, or hand-off strategies. We propose to use this knowledge to animate the motions of a virtual character in a kitchen setting, as they move items to manipulate them, and evaluate whether this knowledge results in more natural character behavior. We focus on position because it was a strong effect and applies to many objects rather than a specific subset (as is the case for the balance factor).

We hypothesize that an animated character that chooses one- or two-handed strategies based on object start/goal position will be perceived as more convincing and human-like than a character that always uses one strategy or picks from the strategies at random.

8.1 Method

8.1.1 Motion capture clips

First, we will collect motion capture clips of unimanual, bimanual, and hand-off strategies using three different start and goal directions. For each of the three strategies, twelve motions will be captured:

1. Using that strategy to move an item from front to left
2. Move item from front to right
3. Move item from front to back
4. Move item from left to front
5. Move item from left to right
6. Move item from left to back
7. Move item from right to front
8. Move item from right to left
9. Move item from right to back
10. Move item from back to front
11. Move item from back to left
12. Move item from back to right

In total, 36 motions will be captured.

8.1.2 Character animation and control

We then animate a character using these motions in a kitchen environment. One control animated character selects strategies (motion clips corresponding to those strategies) at random, while another control character only uses the unimanual strategy. The test character chooses a motion clip based on the goals of minimizing body rotation and having to reach across the body into contralateral space.

8.1.3 Evaluation

Finally, we will conduct a user study where participants watch clips of each type and compare the realism/believability of the character motions.

Chapter 9

Summary of proposed work

9.1 Expected contributions

Contributions from completed / in-progress work:

- An expanded grasp taxonomy created from observing people using their hand over a typical day. In total, we observed 179 actions that might be performed in daily life, and sorted the poses used into 73 types, 40 of which are not covered in the cumulative taxonomy of Feix and colleagues [FPS+09]. Our survey of how people use their hands is more comprehensive than previous work that focuses on gravity-independent wrap grasps.
- An understanding of the range of actions that might be performed once each grasp has been established. Our work shows that a single grasp pose – for example, a large diameter grasp – can be used in many ways – for example, to hold, press, squeeze, or twist objects. Our descriptions of the different forces and motions that a grasp pose is used for indicates what set of skills robots would need to be able to perform *after* grasping to manipulate as well as humans.
- An annotation system that allows people to describe how stable grasp poses are used in actions. In our work, we tried out various ways of describing force and motion such that someone can learn the annotation system with little training. We eventually found that using action verbs to describe force types, three types of coordinate systems to describe motion, and flow were helpful and easy to use. Such a system could be useful for instructing robots how to move.

- High-framerate video and analysis of picking and placing motions that shed light on the entire grasping process, especially in situations where the surfaces needed to form a stable wrap grasp are blocked. With high-framerate video becoming cheaper and more ubiquitous, we used it to capture a large amount of slow-motion examples of picking and placing in naturalistic settings, which include clutter – 90 actions in total. Such a large collection of slow-motion videos allows us to find patterns in the way people grasp objects, at a very fine level of detail. We used this work to identify multiple strategies that have been little studied, including levering up; simultaneously moving objects away from clutter and into the hand; and creating gaps between objects and clutter to allow a finger to be inserted.
- Analysis of factors that influence the use of one or both hands when grasping objects. In particular, we studied the use of hand-offs, which has not gotten much attention in grasping studies. Our results show the advantages and disadvantages of each strategy that influences their usage: the pure bimanual strategy helps to balance an object but requires more body rotation. By contrast, hand-offs are undesirable when balance is important, but greatly reduce body rotation when the object is being transported left-to-right or vice versa. The pure unimanual strategy is between the other two in terms of both stability and amount of rotation required.
- Knowledge about how humans grasp a simple lipless bowl. Rather than finding that a single pose dominates, we found there are many ways to grasp a bowl. The popular poses may be used as a reference to analyze quality of grasps for robots.

Expected contributions from proposed work:

- An evaluation of current taxonomies' ability to describe and prescribe motion. Such an evaluation would help identify gaps in our knowledge of grasping, and inspire new systems of motion description that can be used for robotic grasping.
- A database of whole-body motion capture motions (of transporting objects) and detailed hand motion capture (of levering up objects). Making this kind of data publicly available can help other people doing research in grasping animation or robotics.
- A controller capable of levering objects up and regrasping them in simulation. To our knowledge, there is no work that addresses this specific grasping skill.
- A virtual character in a kitchen setting that represents a step toward naturalistic interactions with the environment by animated characters.
- An understanding of whether the factors that affect choice of grasping strategy also affect the perception of naturalness of grasping strategy to outside observers.

9.2 Timeline

Task	Chapter	Deadline	Venue
Factors in bimanual grasping	5	June 1	Acta Psychologica
Complexities of grasping in the wild	4	July 15	HUMANOIDS 2017
Literature search	7/8	July	
Kitchen mocap session	8	July	
User study IRB application	8	July	
Implementing levering optimization	7	August	
Exploring optimization alternatives	7	November	
Physically-based levering up	7	January	SIGGRAPH 2018
Kitchen animation clips	8	February	
Conduct user study	8	March	
Analysis of study results	8	April	
Grasping in a virtual kitchen	8	May 22	SAP 2018

Bibliography

[ABR+12] John R Amend, Eric Brown, Nicholas Rodenberg, Heinrich M Jaeger, and Hod Lipson. A positive pressure universal gripper based on the jamming of granular material. *IEEE Transactions on Robotics*, 28(2):341–350, 2012.

[AK13] Sheldon Andrews and Paul G Kry. Goal directed multi-finger manipulation: Control policies and analysis. *Computers & Graphics*, 37(7):830–839, 2013.

[AL14] Naoko Abe and Jean-Paul Laumond. Dance notations and robot motion. In *Proceedings of the 1st Workshop of the Anthropomorphic Motion Factory at LAAS-CNRS '14*, Toulouse, France, 2014.

[BB+01] Richard W Byrne, Jennifer M Byrne, et al. Manual dexterity in the gorilla: bimanual and digit role differentiation in a natural task. *Animal Cognition*, 4(3-4):347–361, 2001.

[BMD13] Ian M Bullock, Raymond R Ma, and Aaron M Dollar. A hand-centric classification of human and robot dexterous manipulation. *Haptics, IEEE Transactions on*, 6(2):129–144, 2013.

[BR] M. Bollini and D. Rus. Cookies, anyone? http://web.mit.edu/newsoffice/2011/cookies-anyone.html.

[CDC+09] Young Sang Choi, Travis Deyle, Tiffany Chen, Jonathan D Glass, and Charles C Kemp. 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, 2009.

[CE05] J. Cohn and P. Ekman. Measuring facial action by manual coding, facial EMG, and automatic facial image analysis. *Handbook of Nonverbal Behavior Research Methods in the Affective Sciences*, 2005.

[CM04] Hyeg Joo Choi and Leonard S Mark. Scaling affordances for human reach actions. *Human Movement Science*, 23(6):785–806, 2004.

[CN00] Paola Cesari and Karl M Newell. Body-scaled transitions in human grip configurations. *Journal of Experimental Psychology: Human Perception and Performance*, 26(5):1657, 2000.

[CP09] Lillian Y Chang and Nancy S Pollard. Video survey of pre-grasp interactions in natural hand activities. June 2009.

[Cut89] Mark R Cutkosky. On grasp choice, grasp models, and the design of hands for manufacturing tasks. *Robotics and Automation, IEEE Transactions on*, 5(3):269–279, 1989.

[CWDH88] C Collin, DT Wade, S Davies, and V Horne. The Barthel ADL Index: a reliability study. *Disability & Rehabilitation*, 10(2):61–63, 1988.

[DH10] Aaron M Dollar and Robert D Howe. The highly adaptive SDM hand: Design and performance evaluation. *The International Journal of Robotics Research*, 29(5):585–597, 2010.

[EC84] John M Elliott and KJ Connolly. A classification of manipulative hand movements. *Developmental Medicine & Child Neurology*, 26(3):283–296, 1984.

[EF78] P. Ekman and W. Friesen. *Facial action coding system: A technique for the measurement of facial movement.* Consulting Psychologists Press, Palo Alto, 1978.

[FPS$^+$] T. Feix, R. Pawlik, H. Schmiedmayer, J. Romero, and D. Kragic. Human grasping database. `http://grasp.xief.net`. Accessed: 2017-04-05.

[FPS$^+$09] T. Feix, R. Pawlik, H. Schmiedmayer, J. Romero, and D. Kragic. A comprehensive grasp taxonomy. In *Robotics, Science and Systems: Workshop on Understanding the Human Hand for Advancing Robotic Manipulation*, 2009.

[GBZ$^+$08] Gerhard Grunwald, Christoph Borst, J Zöllner, et al. 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*, 2008.

[GFS14] Claudia LR Gonzalez, Jason W Flindall, and Kayla D Stone. Hand preference across the lifespan: effects of end-goal, task nature, and object location. *Frontiers in Psychology*, 5, 2014.

[GHD12] René Gilster, Constanze Hesse, and Heiner Deubel. Contact points during multi-digit grasping of geometric objects. *Experimental Brain Research*, 217(1):137–151, 2012.

[GIKS+12] Susan Greaves, Christine Imms, Lena Krumlinde-Sundholm, Karen Dodd, and Ann-Christin Eliasson. 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, 2012.

[Gui87] Yves Guiard. Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. *Journal of Motor Behavior*, 19(4):486–517, 1987.

[Hog85] Neville Hogan. Impedance control: An approach to manipulation: Part ii,Äîimplementation. *Journal of Dynamic Systems, Measurement, and Control*, 107(1):8–16, 1985.

[JHS12] Sophie Jörg, Jessica Hodgins, and Alla Safonova. Data-driven finger motion synthesis for gesturing characters. *ACM Transactions on Graphics (TOG)*, 31(6):189, 2012.

[KFKM10] Marliese Kimmerle, Claudio L Ferre, Kathleen A Kotwica, and George F Michel. Development of role-differentiated bimanual manipulation during the infant's first year. *Developmental Psychobiology*, 52(2):168–180, 2010.

[KH70] Ibrahim Adalbert Kapandji and Louis Henri Honoré. *The physiology of the joints: annotated diagrams of the mechanics of the human joints*, volume 1. E. & S. Livingstone London, 1970.

[Kha87] Oussama Khatib. 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, 1987.

[KKF+97] Bruno Kopp, Annett Kunkel, Herta Flor, Thomas Platz, Ulrike Rose, Karl-Heinz Mauritz, Klaus Gresser, Karen L McCulloch, and Edward Taub. 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, 1997.

[KMB03] Marliese Kimmerle, Lynda Mainwaring, and Marlene Borenstein. The functional repertoire of the hand and its application to assessment. *American Journal of Occupational Therapy*, 57(5):489–498, 2003.

[KMI+80] Noriko Kamakura, Michiko Matsuo, Harumi Ishii, Fumiko Mitsuboshi, and Yoriko Miura. Patterns of static prehension in normal hands. *American Journal of Occupational Therapy*, 34(7):437–445, 1980.

[KSN10] E. Klingbeil, A. Saxena, and A. Y. Ng. Learning to open new doors. In *IEEE/RSJ Conference on Intelligent Robots and Systems (IROS)*, 2010.

[LFNP14a] Jia Liu, Fangxiaoyu Feng, Yuzuko Nakamura, and Nancy Pollard. A taxonomy of everyday grasps in action. `http://www.cs.cmu.edu/~jialiu1/database.html`, 2014. Accessed: 2017-04-05.

[LFNP14b] Jia Liu, Fangxiaoyu Feng, Yuzuko C Nakamura, and Nancy S Pollard. A taxonomy of everyday grasps in action. In *Humanoid Robots (Humanoids), 2014 14th IEEE-RAS International Conference on*, pages 573–580. IEEE, 2014.

[LHW+94] John Michael Linacre, Allen W Heinemann, Benjamin D Wright, Carl V Granger, and Byron B Hamilton. The structure and stability of the functional independence measure. *Archives of Physical Medicine and Rehabilitation*, 75(2):127–132, 1994.

[Lyn99] Kevin M Lynch. Toppling manipulation. In *IEEE International Conference on Robotics and Automation (ICRA), 1999*, pages 2551–2557, 1999.

[Mas86] Matthew T Mason. Mechanics and planning of manipulator pushing operations. *The International Journal of Robotics Research*, 5(3):53–71, 1986.

[MPT12] Igor Mordatch, Zoran Popović, and Emanuel Todorov. Contact-invariant optimization for hand manipulation. In *Proceedings of the ACM SIGGRAPH/Eurographics symposium on computer animation*, pages 137–144. Eurographics Association, 2012.

[MSCTLA10] J. Maitin-Shepard, M. Cusumano-Towner, J. Lei, and P. Abbeel. Cloth grasp point detection based on multiple-view geometric cues with application to robotic towel folding. In *IEEE International Conference on Robotics and Automation*, 2010.

[Nap56] John R Napier. The prehensile movements of the human hand. *Journal of Bone and Joint Surgery*, 38(4):902–913, 1956.

[ND03] Jean Newlove and John Dalby. *Laban for all*. Routledge, 2003.

[NT93] John Russell Napier and Russell Tuttle. *Hands*. Princeton University Press, 1993.

[PMCS06] N Pollock, MA McColl, A Carswell, and T Sumsion. The Canadian occupational performance measure. *Client-centred Practice in Occupational Therapy: A Guide to Implementation*, pages 103–114, 2006.

[RCRS10] David A Rosenbaum, Chase J Coelho, Jewels D Rhode, and Joseph P Santamaria. Psychologically distinct classes of motor behavior inferred from individual differences: Evidence from a sequential stacking task. *Journal of Motor Behavior*, 42(3):187–194, 2010.

[Ros08] David A Rosenbaum. Reaching while walking: Reaching distance costs more than walking distance. *Psychonomic Bulletin & Review*, 15(6):1100–1104, 2008.

[SBGK13] Ali-Akbar Samadani, Sarahjane Burton, Rob Gorbet, and Dana Kulic. 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, 2013.

[Sch19] Ing G Schlesinger. Der mechanische aufbau der künstlichen glieder. In *Ersatzglieder und Arbeitshilfen*, pages 321–661. Springer, 1919.

[SJEMP02] D. J. Buckland S. J. Edwards and J. D. McCoy-Powlen. Developmental and functional hand grasps. Oct 2002.

[SK05] M. Stilman and J. Kuffner. Navigation among movable obstacles: Real-time reasoning in complex environments. *International Journal of Humanoid Robotics*, 2(4):479–503, 2005.

[SYN+10] Dragoljub Surdilovic, Yelve Yakut, T-M Nguyen, Xuan Ba Pham, Axel Vick, and R Martin-Martin. 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, 2010.

[TKPB11] M. Tenorth, U. Klank, D. Pangercic, and M. Beetz. Web-enabled robots – robots that use the web as an information resource. *IEEE Robotics and Automation Magazine*, 18, 2011.

[Tor85] Takashi Torigoe. Comparison of object manipulation among 74 species of non-human primates. *Primates*, 26(2):182–194, 1985.

[TP90] Jeffrey C. Trinkle and Richard P. Paul. Planning for dexterous manipulation with sliding contacts. *The International Journal of Robotics Research*, 9(3):24–48, 1990.

[WAK+13] Florentin Worgotter, Eren Erdal Aksoy, Norbert Kruger, Justus Piater, Ales Ude, and Minija Tamosiunaite. A simple ontology of manipulation actions based on hand-object relations. *Autonomous Mental Development, IEEE Transactions on*, 5(2):117–134, 2013.

[YL12] Yuting Ye and C Karen Liu. Synthesis of detailed hand manipulations using contact sampling. *ACM Transactions on Graphics (TOG)*, 31(4):41, 2012.

[YPL+10] Eiichi Yoshida, Mathieu Poirier, Jean-Paul Laumond, Oussama Kanoun, Florent Lamiraux, Rachid Alami, and Kazuhito Yokoi. Pivoting based manipulation by a humanoid robot. *Autonomous Robots*, 28(1):77–88, 2010.

[ZDLRD11] Joshua Z Zheng, Sara De La Rosa, and Aaron M Dollar. 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, 2011.

[ZZMC13] Wenping Zhao, Jianjie Zhang, Jianyuan Min, and Jinxiang Chai. Robust realtime physics-based motion control for human grasping. *ACM Transactions on Graphics (TOG)*, 32(6):207, 2013.