Thesis Proposal

Algorithms, Implementation, and Studies on Eating with a Shared Control Robot Arm

Laura V. Herlant

August 15, 2016

The Robotics Institute
Carnegie Mellon University
Pittsburgh, PA 15213

Thesis Committee:

Siddhartha S. Srinivasa, CMU RI
Jodi Forlizzi, CMU HCII
Christopher G. Atkeson, CMU RI
Leila A. Takayama, UC Santa Cruz

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

Abstract

People with upper extremity disabilities are gaining increased independence through the use of assisted devices such as wheelchair-mounted robotic arms. However, the increased capability and dexterity of these robotic arms also makes them challenging to control through accessible interfaces like joysticks, sip-and-puff, and buttons that are lower-dimensional than the control space of the robot. The potential for robotics autonomy to ease control burden within assistive domains has been recognized for decades. While full autonomy is an option, it removes all control from the user. When this is not desired by the human, the assistive technology in fact has made them *less* able and discards useful input the human might provide, leveraging for example their superior situational awareness, that would add to system robustness.

This thesis takes an in-depth dive into how to add autonomy to an assistive robot arm in the specific application of eating, to make it faster and more enjoyable for people with disabilities to feed themselves. While we are focused on this specific application, the tools and insights we gain can generalize to the fields of deformable object manipulation, selection from behavior libraries, intent prediction, robot teleoperation, and human-robot interaction. The nature of the physical proximity and heavy dependence on the robot arm for doing daily tasks creates a very high-stakes human-robot interaction.

We propose a system that is capable of fully autonomous feeding by (1) predicting bite timing based on social queues, (2) detecting relevant features of the food using RGBD sensor data, and (3) automatically selecting a goal and a food-collection motion primitive to bring a bite from the plate to the operator's mouth. We propose investigating the desired level of autonomy through user studies with an assistive robot where users have varying degrees of control over the bite timing, bite selection, action selection, control mode-switching, and direct teleoperation of the robot to determine the effect on cognitive load, acceptance, trust, and task performance.

Contents

1
Introduction

Assistive machines—like powered wheelchairs, myoelectric prostheses and robotic arms—promote independence and ability in those with severe motor impairments [Hillman et al., 2002, Prior, 1990, Sijs et al., 2007, Huete et al., 2012, Yanco, 1998]. As the state-of-the-art advances, more dexterous and capable machines hold the promise to revolutionize ways in which those with motor impairments can interact within society and with their loved ones, and to care for themselves with independence.

However, as these machines become more capable, they often also become more complex. Which raises the question: how to control this added complexity? A confounding factor is that the more severe a person's motor impairment, the more limited are the control interfaces available to them to operate. The control signals issued by these interfaces are lower in dimensionality and bandwidth. Thus, paradoxically, a greater need for sophisticated assistive devices is paired with a diminishing ability to control their additional complexity.

Traditional interfaces often cover only a portion of the control space of more complex devices like robotic arms [Tsui et al., 2008]. For example, while a 2-axis joystick does fully cover the 2-D control space (heading, speed) of a powered wheelchair, to control the end-effector of a robotic arm is nominally a 6-D control problem. This already is a challenge with a 2-D control interface, which is only exasperated if limited to a 1-D interface like a Sip-N-Puff or switch-based head array [Nuttin et al., 2002, Valbuena et al., 2007, Mandel et al., 2009, Prenzel et al., 2007, Luth et al., 2007, Simpson et al., 2008, Firoozabadi et al., 2008, Vaidyanathan et al., 2006, Galán et al., 2008].

(a) Wheelchair Joystick

(b) Sip-N-Puff

(c) Head Array

Figure 1.1: Traditional control interfaces for electric wheelchairs or other assistive devices.

Challenge 1: Low degree of freedom input to control an increasingly complex and high degree of freedom robot.

A common technique to control a high-dimensional system like an arm with a low-dimensional input like a joystick is through switch-

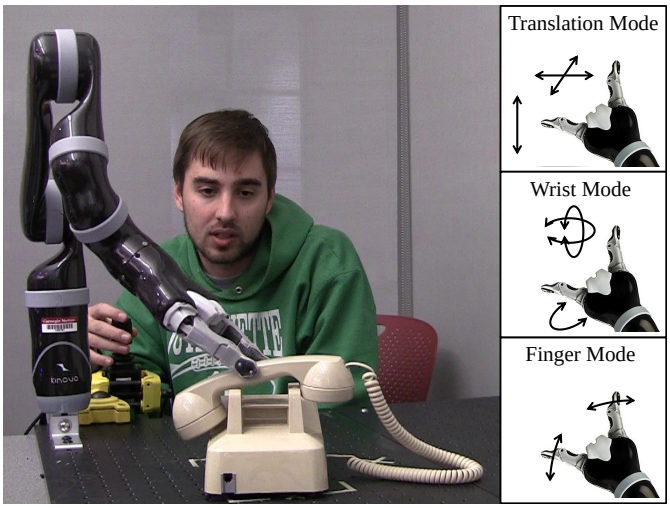


Figure 1.2: An able-bodied user is dialing 911 by directly teleoperating the MICO arm with modal control. The three control modes are translation, wrist, and gripper mode. A 3-axis joystick is used as primary input and buttons atop the joystick are used to change mode.

ing between multiple control modes, such as those shown in fig. 1.2. However, our interviews with daily users of the Kinova JACO arm identified mode switching as a key problem, both in terms of time and cognitive load. We further confirmed objectively that mode switching consumes about 17.4% of execution time even for able-bodied users controlling the JACO.

The potential for robotics autonomy to ease control burden within assistive domains has been recognized for decades. While full autonomy is an option, it removes all control from the user. When this is not desired by the human, the assistive technology in fact has made them *less* able. It also discards useful input the human might provide, leveraging for example their superior situational awareness, that would add to system robustness.

Challenge 2: How to share control in a way that doesn't lead to rejection?

Control sharing is a way to offload some control burden, without removing all control authority, from the human [Dragan and Srinivasa, 2013, 2012, Yanco, 2000, Philips et al., 2007, Vanhooydonck et al., 2003, Bourhis and Sahnoun, 2007]. The most common paradigms augment or adjust control signals from the human (e.g. to bridge the gap in control signal dimensionality), or partition the control problem (e.g. high-level decisions like which task to execute lie with the human, and low-level execution decisions lie with the robot).

Here, we propose an alternative role for the autonomy: to assist the user in transitioning between different subsets of the control space—that is, to autonomously remap signals from the user interface to different control *modes* (e.g. subsets of the control dimensions).

Figure 1.3: Users controlling the Kinova JACO arm through a hand-operated (left) and foot-operated (middle) joystick. Our platform (right), the Kinova MICO arm, is kinematically identical but smaller, lighter and cheaper, with a two-fingered hand.

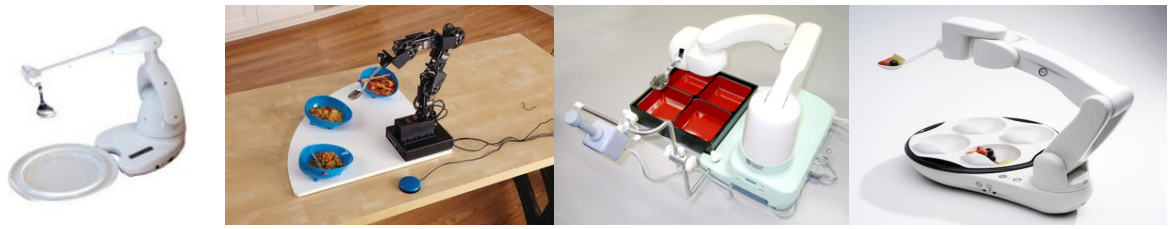
The operation of an assistive device via different control modes is reminscient of upper-limb prosthesis control [Ajiboye and ff. Weir, 2005, Chu et al., 2006, Nishikawa et al., 1999, Scheme and Englehart, 2011, Simon et al., 2011, Tenore et al., 2008, 2009]. In this case control is diverted between different "functions" (e.g. elbow, wrist). The parallel for a robotic arm is to divert control between different subsets of the joint-control space. (Modes that operate subsets of the end-effector control space of course are equally viable.) Within the field of prosthetics, function switching is known to be cumbersome, and the opportunity for autonomous switching to ease this burden has been identified [Pilarski et al., 2012] (though it is not yet feasible to implement on today's prosthetic hardware).

We introduce a formalism section 2.4 for assistive mode-switching in which the user issues low-dimensional control signals, and these signals are not augmented or modified by the autonomy. What the autonomy *is* changing is what these low-dimensional signals map to.

Mode-switching assistance has the key benefit that it is low-bandwidth and intermittent, and hence possibly easier to learn. We also hypothesize that it acheives a sweet-spot of assistance: helping them with tedious parts of the task while still giving them full control over continuous motion. Furthermore, we believe that because of these benefits, users will be more tolerant to errors in mode-switching assistance.

Feeding oneself has a high impact on quality of life [Stanger et al., 1994], which we confirmed in our own interviews with people with disabilities who use an assistive robot arm section 2.1. Therefore we have decided to tackle the challenges of controlling an assistive robot arm listed above, in the context of eating.

This thesis takes an in-depth dive into how to add autonomy to



(a) Meal-Mate™ (b) Meal Buddy (c) My Spoon (d) Obi

Figure 1.4: Four assistive feeding devices currently on the market.

an assistive robot arm in the specific application of eating, to make it faster and more enjoyable for people with disabilities to feed themselves. While we are focused on this specific application, the tools and insights we gain can generalize to the fields of deformable object manipulation, selection from behavior libraries, intent prediction, robot teleoperation, and human-robot interaction. The nature of the physical proximity and heavy dependence on the robot arm for doing daily tasks creates a very high-stakes HRI interaction.

1.1 *Feeding Devices on the Market*

Several specialized feeding devices for people with disabilities have come onto the market in the past decade, some of which are shown in fig. 1.4, and they all work in a similar manner.

The Meal Buddy Assistive Feeder is made by Patterson Medical and features three bowls of food that are rigidly attached to the base of a robot arm. A button is used to select which bowl, and to initiate bite collection.

More information on the Meal Buddy Assistive Feeder at `http://www.pattersonmedical.com/`

My Spoon is made by Secom, and features a robot arm with four bowls for food and an actuated silverware end effector, that will open and close to grasp food before presenting it to the user.

More information on My Spoon at `http://www.secom.co.jp/english/myspoon/`

The Obi is made by Desin and features four bowls of food that are rigidly attached to the base of a robot arm. The Obi also has a "teach" mode where a caregiver can place the robot at a position near the mouth that will be remembered when the user triggers a bite via button press.

More information on the Obi at `https://meetobi.com/`

Meal-Mate™ Eating Device is made by RBF Healthcare uses a spoon attached to a robot arm, which will move down to a plate and back up with the press of a button, but while the spoon is on the plate, a joystick or arrow buttons can be used to move the spoon along the plate.

More information on Meal-Mate™ at `http://rbfindustries.co.uk/healthcare/`

All these feeding devices are designed only for food manipulation, and require specialized food containers to function effectively. With the exception of Meal-Mate™, all the feeding devices are controlled with buttons to select which food bowl and then to trigger a bite

action, with control over the robot's motion. Similarly, none of these devices have any way to sense success or failure of taking bites, nor a way to automatically time when to provide bites to the user.

This thesis builds on commercially available devices by using a robot arm that will proactively estimate the appropriate bite timings to the user and will intelligently acquire food that does not require partitioning into separate bins or low-level user control of the spoon's motion. In addition, by keeping the user "in the loop", the user has the ability to intervene in any failure cases.

1.2 *Approach*

Prior work strongly supports that the most important factor in the design of a home health care interface was a clear understanding of the user population and their needs, capabilities, and limitations [Delano and Hartman, 2001]. We are following this approach by working iteratively with our target population through interviews and studies.

We consider the problem of an assistive robot that is teleoperated by a person with disabilities to eat a meal. In this scenario, the person is seated at a table with a plate or bowl of food in front of them. We are assuming that there is a way to detect the color and depth values for the food on the table. This can be in the form of an RGBD camera mounted on the robot's wrist, or other sensing mounted in the environment. Mounting the sensor on the user's wheelchair allows for more flexibility as it can extend to situations outside the sensor-equipped home location. For the purposes of the proposed experiments, we have both setups, a Microsoft Kinect mounted above the table, and a PrimeSense StructureIO mounted on the wrist of the robot.

We propose implementing a fully autonomous feeding system (section 3.2) with which we can then vary the level of shared control and the corresponding interface to find the best combination in terms of task performance and usability.

2
Completed Work

2.1 *Kinova Interviews*

Our own interviews with current users of the Kinova arm pinpointed that the struggles with modal control relate back to the need to constantly change modes. Our users found switching between the various control modes, seen in fig. 1.2, to be slow and burdensome, noting that there were "a lot of modes, actions, combination of buttons". Each of these mode changes requires the user to divert their attention away from accomplishing their task to consider the necessary mode change [Tijsma et al., 2005a,b]. The cognitive action of shifting attention from one task to another is referred to as *task switching.* Task switching slows down users and can lead to increased errors regardless of the interface they are using to make this switch [Monsell, 2003, Wylie and Allport, 2000, Meiran et al., 2000, Strobach et al., 2012, Arrington and Logan, 2004]. Simply the need to change modes is a harmful distraction that impedes efficient control.

2.2 *Inventory of OT Manipulation Metrics*

There are a wide variety of evaluation techniques used for wheelchair-mounted assistive robotic manipulators. A literature review of end-user evaluations [Chung and Cooper, 2012] revealed that task completion time and task completion rate were the most used evaluation metric, but the tasks being used were different across researchers and platforms leading to difficult comparisons. In one Manus arm evaluation, there were no predefined set of tasks, but rather the user was allowed to use the arm naturally about their environment. Afterward, the common tasks across users were inventoried and evaluated[Eftring and Boschian, 1999]. Another study with the Manus was evaluated by having users grasp the following objects: remote controller, a cereal box, two jars, a soda can, and a water bottle [Kim et al., 2012]. A different study with the Manus was tested with the

(a) 9 Hole Peg Test (b) Purdue Pegboard Test (c) Box and Block Test (d) Minnesota Test

Figure 2.1: Functional tests that repeat specific motion primitives.

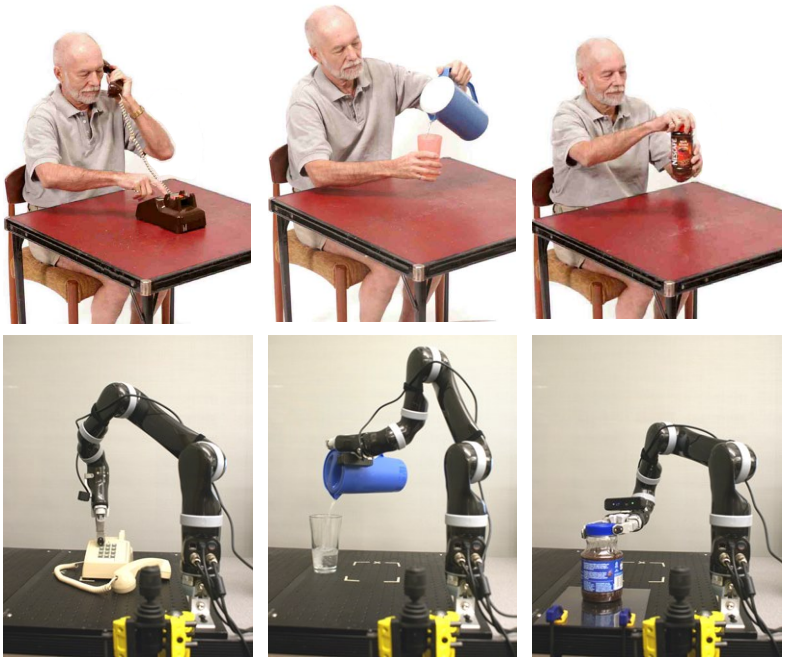
task of grasping a foam ball that was hanging on a string [Tsui and Yanco, 2007]. While each task is reasonable, it is very difficult to draw comparisons between the studies, even though all are done on the same system.

There are advantages to evaluating our assistive robot arm system with the same standardized tests that Occupational Therapists (OT) use with human patients. Standard OT tests are already validated, have years of application, provide a metric that can be used across robots and people, and makes our results more transparent to a larger community. Additionally, this would allow for a larger interaction with insurance companies or healthcare economists [Mahoney, 1997].

Manipulation tests can be divided into strength tests, which measure the muscular ability of the human hand and upper limbs, and functional tests, which measure the ability to perform particular tasks or motions. Because the robot's hardware will exclusively determine its performance on strength tests and will not change drastically over time, evaluations using functional tests make more sense for comparing across robot and control platforms.

A survey of functional tests shows that functional tests fall into two categories: repeated tasks that test specific motion primitives, and tasks that are based on daily living and cover a several motion primitives per task. The Purdue Pegboard test [Tiffin and Asher, 1948], 9 Hole Peg Test [Sunderland et al., 1989], Box and Block Test [Mathiowetz et al., 1985], and Minnesota rate of manipulation test, shown in fig. 3.4, ask the user to provide a number of short repeated tasks and use the overall task time as the primary evaluation metric. Tests based on daily activities include the Motor Activity Log (MAL) [Taub et al., 2011, Uswatte et al., 2005], the Jebsen Taylor Hand Function Test [Tipton-Burton, 2011], the Action Research Arm (ARA) test [McDonnell, 2008], Sollerman hand function test [Sollerman and Ejeskär, 1995], and the Chedoke Arm and Hand Activity Inventory (CAHAI) [Barreca et al., 2004].

Our prior work has used tasks from the CAHAI test, which has the advantage of including both qualitative and quantitative metrics



(a) (b) (c)

Figure 2.2: Three modified tasks from the Chedoke Arm and Hand Activity Inventory, which able-bodied users performed through teleoperating the MICO robot.

for evaluation and having tasks that can be easily modified to be performed with one hand. It is optimized for evaluating stroke recovery, which is not the target population for this work. We argue that spinal cord injury patients *with an assistive robot manipulator* is more similar to the situation of a stroke patient, who has varying levels of control over their limb.

2.3 *Chedoke Tasks with the Robot with Able-bodied Users*

To objectively measure the impact of mode switching, we ran a study with able-bodied users performing household tasks with the Kinova MICO arm using a joystick interface.

Experimental Setup Users sat behind a table on which the MICO arm was rigidly mounted. They used the standard Kinova joystick to control the arm.

Tasks The tasks we chose are modified from the Chedoke Arm and Hand Activity Inventory (CAHAI), a validated, upper-limb measure to assess functional recovery of the arm and hand after a stroke [Barreca et al., 2004]. The CAHAI has the advantage of drawing tasks from instrumental activities of daily living, which are representative of our desired use case, rather than occupational therapy assessment tools such as the 9-Hole Peg Test [Kellor et al., 1971] or the Minnesota Manipulation Test [Cromwell et al., 1960] that also evaluate upper extremity function, but do not place the results into context.

Manipulated Factors We manipulated which task the user performed. The three tasks we used were: calling 911, pouring a glass of water from a pitcher, and unscrewing the lid from a jar of coffee. These three tasks were chosen from the CAHAI test because they could easily be modified from a bimanual task to being performed with one arm. The three tasks are shown in fig. 2.2.

Procedure After a five minute training period, each user was given a maximum of ten minutes per task. The order of tasks was counterbalanced across the users. The joystick inputs and the robot's position and orientation were recorded throughout all trials. After all the tasks were attempted, we asked the users to rate the difficulty of each task on a 7-point Likert scale and what aspects make performing tasks difficult with this robot.

Participants and Allocation We recruited 6 able-bodied participants from the local community (4 male, 2 female, aged 21-34). This was a within subject design, and each participant performed all three tasks with a counterbalanced ordering.

Analysis On average, 17.4 $\pm$ 0.8% of task execution time is spent changing control modes and not actually moving the robot. The mode changing times were calculated as the total time the user did not move the joystick before and after changing control mode. The fraction of total execution time that was spent changing modes was fairly consistent both across users and tasks as seen in fig. 2.3. If time spent changing mode could be removed, users would gain over a sixth of the total operating time.

The tasks the users performed were not of equal difficulty. Users responded that the pitcher pouring was the most difficult task (M=5.5,SD=0.7), followed by unscrewing the jar (M=5.2,SD=0.7), and the easiest task was dialing 911 (M=4,SD=0.6). The total execution time shown in fig. 2.3 mirrors the difficulty ratings, with harder tasks taking longer to complete. Difficulty could also be linked to the number of mode switches, mode switching time, or ratio of time spent mode switching, as shown in fig. 2.5. The hardest and easiest tasks are most easily identified when using switching time as a discriminating factor. The pitcher and jar tasks both rated as significantly more difficult than the telephone task, which may be due to the large number of mode changes and small adjustments needed to move the robot's hand along an arc — as one user pointed out: "Circular motion is hard."

One might argue that we are basing our findings on novice users, and their discomfort and hesitation switching modes will diminish over time. However, over the course of half an hour using the arm, and an average of more than 100 mode switches, users did not show any significant decrease in the time it takes to change mode (fig. 2.4).

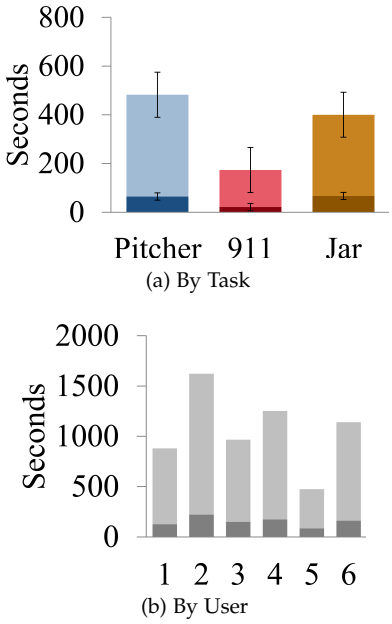


Figure 2.3: Execution time, with time spent mode switching shown in the darker shades.

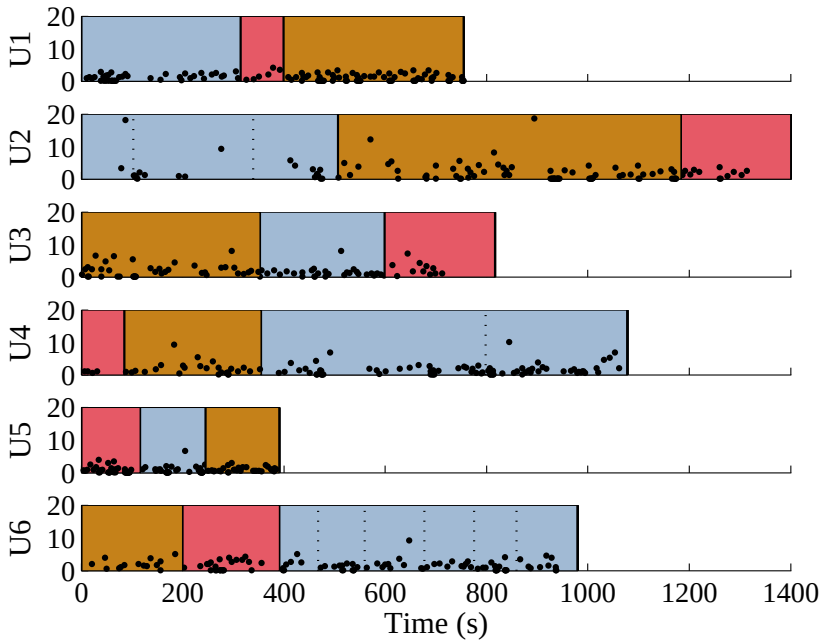


Figure 2.4: Each point is a mode switch, with the y-value indicating the mode switching time, and the x-value indicating when the mode switch occurred. The colors correspond to the task (blue: pouring pitcher, brown: unscrew jar of coffee, red: dial 911), and the order of tasks can be seen for each user in arranged from left to right. Dashed lines in the pitcher task identify locations where the user dropped the pitcher and the task had to be reset.

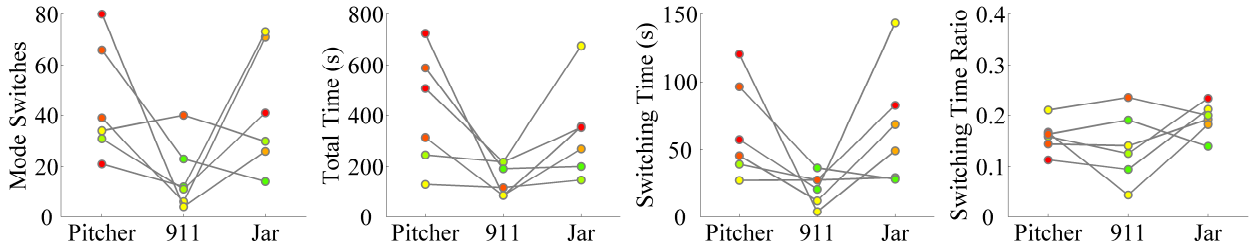


Figure 2.5: The connected components are a single user, and the colors represent the difficulty that user rated each task with red being most difficult, yellow being second most difficult, and green being least difficult.

The continued cost of mode switching is further supported by our interviews, in which a person using the JACO for more than three years stated "it's really hard with the JACO because there are too many mobilizations and too many movements required."

The users had three possible modes and used two buttons on the top of the joystick to change between them. The left button started translation mode, the right button started orientation mode, and pressing both buttons simultaneously started gripper mode. Changing into gripper mode was particularly since the timing between the two buttons had to be very precise lest the user accidentally press the left or right button when releasing and switch to translation or wrist mode. The cost to change from one mode to another was not constant across the modes; table 2.1 shows the average time it took to change from the mode in the row to the mode in the column. While in this case the difference can be explained by the chosen interface, it could be important to consider if switching from one particular

	Translation	Wrist	Finger
Translation	-	1.98 ± 0.15s	1.94 ± 0.16s
Wrist	2.04 ± 0.51s	-	3.20 ± 1.85s
Finger	1.30 ± 0.13s	0.98 ± 0.24s	-

Table 2.1: Mean mode switching times

control mode to another causes a larger mental shift in context. Such differences would require the cost of mode switches to be directional, which we leave for future work.

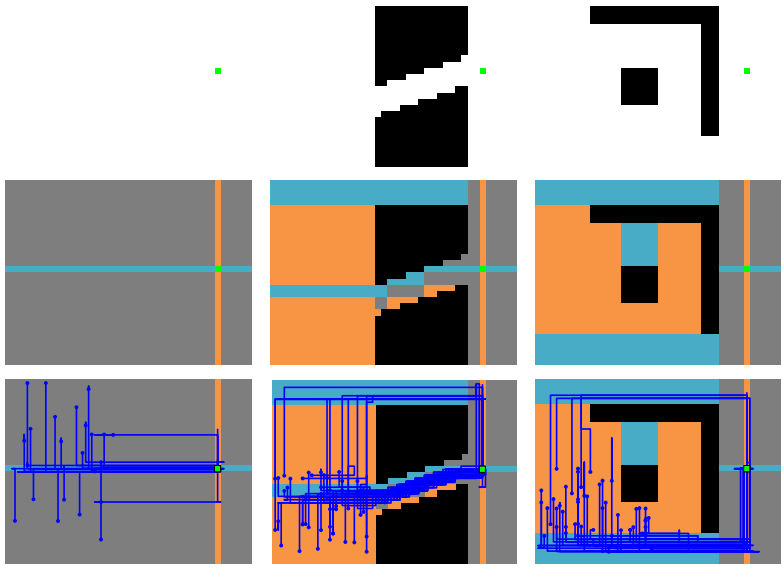
2.4 *Time-Optimal Mode Switching*

The users of the JACO arm identified that frequently changing modes was difficult. We objectively confirmed the difficulty of mode changing by having able-bodied users perform everyday tasks with the MICO arm. Having identified mode switching as a problem in this complex scenario, we tried to model the problem in a much simpler scenario and provide the foundations for scaling the solution back up to the full space of the MICO arm.

2.4.1 *Study 2: 2D Mode Switching Task*

Study 1 demonstrated that people using modal control spend a significant amount of their time changing modes and not moving the robot. The next step is to model when people change modes so that the robot can provide assistance at the right time. We identified certain behaviors from Study 1 that could confound our ability to fit an accurate model. We observed that different people used very different strategies for each of the tasks, which we postulated is because they were performing multi-step tasks that required several intermediate placements of the robot's gripper. In some trials, users changed their mind about where they wanted to grab an item in the middle of a motion, which we could detect by the verbal comments they made. To gather a more controlled set of trajectories under modal control, we ran a second study in which we more rigidly defined the goal and used only two modes. To fully constrain the goal, we used a simulated robot navigating in two dimensions and a point goal location. We kept all the aspects of modal control as similar to that of the robot arm as possible. Using a 2D simulated robot made it simpler to train novice users and removed confounds, allowing us to more clearly see the impacts of our manipulated factors as described below.

Experimental Setup In this study, the users were given the task



(a) Task 1 (b) Task 2 (c) Task 3

Figure 2.6: Top row: Three tasks that the users performed with a 2D robot. The green square is the goal state and the black polygons are obstacles. Middle row: regions are colored with respect to the optimal control mode; in blue regions it is better to be in x mode, in orange regions it is better to be in y mode, and in gray regions x and y mode yield the same results. Bottom row: user trajectories are overlaid on top of the decision regions, illustrating significant agreement.

of navigating to a goal location in a planar world with polygonal obstacles. We had each user teleoperate a simulated point robot in a 2D world. There were two control modes: one to move the robot vertically, and one to move it horizontally. In each mode, the users pressed the up and down arrow keys on the computer keyboard to move the robot along the axis being controlled. By using the same input keys in both modes, the user is forced to re-map the key's functionality by pressing the spacebar. This is a more realistic analogy to the robot arm scenario, where the same joystick is being used in all of the different control modes to control different things.

Manipulated Factors We manipulated two factors: the delay when changing modes (with 3 levels) and the obstacles in the robot's world (with 3 levels). To simulate the cost of changing modes, we introduced either no delay, a one second delay, or a two second delay whenever the user changed modes. Different time delays are analogous to taking more or less time to change mode due to the interface, the cognitive effort necessary, or the physical effort. We also varied the world the robot had to navigate in order to gather a variety of different examples. The three tasks are as follows: (1) an empty world, (2) a world with concave and convex polygons obstacles, and (3) a world with a diagonal tunnel through an obstacle, and are shown in the top row of fig. 2.6.

Procedure This was a within users study design. Each user saw only one task, but they saw all three delay conditions. Each user had a two trial training period with no delay to learn the keypad controls, and then performed each of the three delay conditions twice. Five

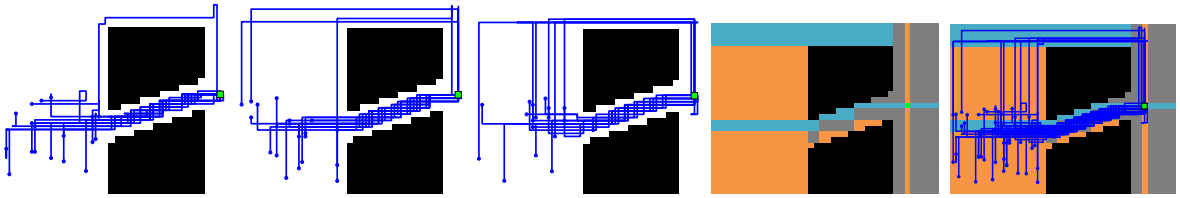


Figure 2.7: User strategies for Task 2 are shown via their paths colored in blue. As the delay increases, some users choose to go around the obstacles rather than through the tunnel, to avoid switching mode. There is still significant agreement with the time-optimal strategy.

users performed each task. The goal remained constant across all the conditions, but the starting position was randomly chosen within the bottom left quadrant of the world. We collected the timing of each key press, the robot's trajectory, and the control mode throughout each of the trials.

Measures To measure task efficiency, we used three metrics: the total execution time, the number of mode switches, and the total amount of time switching modes. We also recorded the path the user moved the robot and which control mode the robot was in at each time step.

Participants We recruited 15 participants through Amazon Mechanical Turk (aged 18-60).

Analysis When the cost of changing modes increases, people choose different strategies in particular situations. This is best seen in Task 2, where there were two different routes to the goal, whereas in Task 1 and Task 3 the map is symmetrical. When there was no mode delay, nearly all users in Task 2 navigated through the tunnel to get to the goal, fig. 2.7. When the delay was one second, some users began to navigate around the obstacles completely, and not through the tunnel. While navigating the tunnel was a shorter Euclidean distance, it required more mode changes than navigating around the tunnel entirely. Therefore we saw that when the cost of mode changes increased, people optimized for fewer mode changes.

We noticed that the user trajectories could be very well modeled by making the assumption that the next action they took was the one that would take them to the goal in the least amount of time. Since switching modes is one of the possible actions, it becomes possible to use this simple model to predict mode switches. The next section discusses the time-optimality model in more detail.

2.5 *Time-Optimal Mode Switching*

The time-optimal policy was found by assigning a cost to changing mode and a cost to pressing a key. These costs were found by empirically averaging across the time it took the users from Study 2 to perform these actions. Using a graph search algorithm, in our case Dijkstra's algorithm [Dijkstra, 1959], we can then determine how

much time the optimal path would take. By looking at each (x,y) location, we can see if the optimal path is faster if the robot is in x-mode or y-mode. The time-optimal mode for each particular (x,y) location is the mode which has a faster optimal path to the goal. A visualization of the optimal mode can be seen in fig. 2.6 for each of the tasks. Time-optimal paths change into the optimal mode as soon as the robot enters one of the x-regions or one of the y-regions. By plotting the user trajectories over a map of the regions, we can see where users were suboptimal. If they were moving vertically in the x-region or horizontally in the y-region, they were performing sub-optimally with respect to time.

In Task 1, users were in the time-optimal mode 93.11% of the time. In Task 2, users were in the time-optimal mode 73.47% of the time. In Task 3, users were in the time-optimal mode 90.52% of the time. Task 2 and Task 3 require more frequent mode switching due to the presence of obstacles.

2.5.1 *Study 3 : 2D Automatic Mode Switching*

Once we determined that people often switch to be time-optimal, we tested how people would react if the robot autonomously switched modes for them. Using the same tasks from Study 2, we used the time-optimal region maps (fig. 2.6), to govern the robot's behavior.

Manipulated Factors We manipulated two factors: the strategy of the robot's mode switching (with 3 levels) and the delay from the mode switch (with 2 levels). The mode switching strategy was either manual, automatic or forced. In the manual case, changing the robot's mode was only controlled by the user. In the automatic case, when the robot entered a new region based on our optimality map, the robot would automatically switch into the time-optimal mode. This change would occur only when the robot first entered the zone, but then the user was free to change the mode back at any time. Within each region in the forced case, by contrast, after every action the user took, the robot would switch into the time-optimal mode. This meant that if the user wanted to change to a suboptimal mode, they could only move the robot one step before the mode was automatically changed into the optimal mode. Hence the robot effectively forces the user to be in the time-optimal mode.

Similar to Study 2, we had a delay condition, however we considered the following three cases: (1) no delay across all assistance types, (2) a two second delay across all assistance types, and (3) a two second delay for manual switching but no delay for auto and forced switching. The purpose of varying the delay was to see if the users' preference was impacted equally by removing the imposed cost of

changing mode (delay type 3), and by only removing the burden on the user to decide about changing mode (delay type 1 and 2).

Hypotheses

H1: People will prefer when the robot provides assistance.

H2: Forced assistance will frustrate users because they will not be able to change the mode for more than a single move if they do not accept the robot's mode switch.

H3: People will perform the task faster when the robot provides assistance.

Procedure After giving each user two practice trials, we conducted pairs of trials in which the user completed the task with the manual mode and either the forced or automatic mode. Testing the automatic assistance and forced assistance across the three delay conditions led to six pairs. For each pair, users were asked to compare the two trials on a forced choice 7-point Likert scale with respect to the user's preference, perceived task difficulty, perceived speed, and comfort level. At the conclusion of the study, users answered how the felt about the robot's mode switching behavior overall and to rate on a 7-point Likert scale if the robot changed modes at the correct times and locations.

Participants We recruited 13 able-bodied participants from the local community (7 male, 6 female, aged 21-58).

Analysis People responded that they preferred using the two types of assistance significantly more than the manual control, $t(154) = 2.96$, $p = .004$, supporting **H1**. The users' preference correlated strongly with which control type they perceived to be faster and easier (R=0.89 and R=0.81 respectively).

At the conclusion of the study, users responded that they felt comfortable with the robot's mode switching (M=5.9, SD=1.0), and thought it did so at the correct time and location (M=5.7, SD=1.8). Both responses were above the neutral response of 4, with t(24)=4.72, $p < .001$ and t(24)=2.34, $p = .028$ respectively. This supports our finding that mode switching can be predicted by following a strategy that always places the robot in the time-optimal mode.

Since this was an experiment, we did not tell participants which trials the robot would autonomously changing modes in. As a result, the first time the robot switched modes automatically, many users were noticeably taken aback. Some users immediately adjusted, with one saying "even though it caught me off guard that the mode automatically switched, it switched at the exact time I would have switched it myself, which was helpful". While others were initially hesitant, all but two of the participants quickly began to strongly prefer when the robot autonomously changed for them, remarking

that it saved time and key presses. They appreciated that the robot made the task easier and even that "the program recognized my intention".

Over time people learned where and when the robot would help them and seemed to adjust their path to maximize robot assistance. People rarely, if ever, fought against the mode change that the robot performed. They trusted the robot's behavior enough to take the robots suggestions [Mead and Matarić, 2009, Baker and Yanco, 2004, Kubo et al., 2009]. We found no significant difference between the forced and automatic mode switching in terms of user preference $t(76) = 0.37, p = 0.71$. Some users even stated that there was no difference at all between the two trials. Therefore we did not find evidence to support **H2**.

Task efficiency, measured by total execution time and total time spent changing modes (as opposed to moving the robot), was not significantly different between the manual control, auto switching, and forced switching conditions. Therefore we were not able to support **H3**. However, this is not surprising as the assistance techniques are choosing when to switch modes based on a model that humans already closely follow. It follows that the resulting trajectories do not differ greatly in terms of path length nor execution time.

2.6 *Discussion*

While this work proposes a useful intervention for controlling assistive robotic arms, there remain many avenues to explore.

Generalization While our studies involved exclusively able-bodied subjects, we want to see how these results, particularly those relating to acceptance, generalize to people with disabilities. Study 3, as described in section 2.5.1, was restricted to a 2D point robot and we are looking to reconduct this experiment on the MICO robot arm. The optimal mode regions will become optimal mode volumes, and the assisted mode switching will occur when the robot enters a new volume.

Priority List of Modes How best to improve modal control is an open question. Here we presented one technique: having the robot perform mode switches automatically. An alternative form of assistance would be to have a priority ordered list of modes. With a single button press the user could transition to the most likely next mode, as estimated by the robot. In the case of an incorrect ordering the user would cycle through the list. This kind of assistance would complement Kinova's LCD screen as described in section 2.1.

Extramodal Control We have suggested a method for more easily switching modes. An alternative approach is to remove this prob-

lem entirely by having the user only control one mode. The remaining modes would be controlled by the robot, in a type of assistance known as *extramodal assistance*. While this eliminates the burden on the user in general, mistakes made by the robot become much more costly. In the case of a robot mistake, the user would have to change modes manually, correct the mistake and then revert back to their original task.

Goal uncertainty Often the exact goal the user wants to achieve is not known in advance. In these situations, the system must simultaneously predict the user's goals and assist them to complete the task [Dragan and Srinivasa, 2012, 2013]. Robot policies for each goal are key to most formalisms that address goal uncertainty. Our framework provides that policy and can be naturally plugged into a policy blending formalism for goal uncertainty.

By investigating modal control and helpful interventions we strive to close the gap between what users want to do with their assistive arms and what they can achieve with ease.

Figure 2.8: Feeding Demo at DARPA

2.7 *Feeding Demo*

We have implemented a pipeline for basic autonomous feeding. In the current system, a bite is acquired by (1) the user pressing a button, (2) the robot looking at the plate to gather sensor data of the food, (3) the robot identifies possible bite locations, (4) the robot selects a bite at random, and then (5) the robot moves to position a fork above the chosen bite, moves straight up and down to skewer the bite, and then returns to the user's programmed mouth location.

2.8 *Preliminary Social Dining Study*

Between 3 and 4 people were asked to sit around a table and eat lunch while having a conversation. During the study, there was a camera on the table in front of each person, pointing at their face, recording video and audio. From the audio, we could identify who was speaking, and from the video, we could identify head pose, gaze direction, and mouth activity. We ran 6 sessions, with a total of 22 participants (9 male, 13 female, age 20-61), which generated 11.5 hours of video recordings.

From the raw audio and video recordings, we used IntraFace [De la Torre et al., 2015] to detect eyegaze and head pose of each person. We could also use the motion of the mouth and face tracking uncertainty to automatically detect when the participants took bites. We observed a number of interesting behaviors that we will explore more empirically in future steps. We observed that while speaking,

participants would "queue" their food, by loading a fork full of food, then holding it a resting position near the plate until there was a an appropriate pause in their speech, or until their turn to talk was completed. Only at a natural break did the participant proceed to put the food in their mouth. This queuing behavior was not observed if the participant was not speaking; in that case, the participant would load the fork and then directly bring it to their mouth.

3
Proposed Work

We have taken the first step towards an intelligent robot to assist with feeding.We propose to address the following key issues: 1) Coordinating robot behavior (selection, parameterization, and timing) with the human user's behavior. 2) Handling a variety of robot and human behaviors (in this case driven by the type of food and how it is eaten). 3) Providing natural and unobtrusive user controls (making human robot interaction less painful and perhaps even enjoyable).

First, the intent prediction does not take into account what food the user wants or the timing of the bites. We propose to predict bite-timing based on conversational queues or user input, and to incorporate eye gaze into intent prediction (section 3.1).

Second, the movement to collect food morsels is a single skewering motion, which works well for individual pieces of fruit, but fails completely when applied to rice, soup, or other food items. We propose a dynamic method of food collection that takes into account the type of food via an RGBD camera, and then applies one of several motions that is the most likely to work for that food type (section 3.2).

Finally, the existing system is fully autonomous aside from a single button to trigger bite collection. We would like to add more user control and allow the robot to plan in a more user-centric way considering the unique HRI challenges of an assistive feeding robot (section 3.3).

3.1 Bite Intent and Timing Prediction

We want to create a feeding device that takes the initiative and offers bites proactively during the meal at times when a bite is likely to be desired. However, the timing becomes important. Imagine the unpleasantness of being interrupted constantly by a robot with a full fork waiting to be unloaded, or the frustration of having to wait on

an seemingly inattentive or slow robot in order to take the next bite.

RQ 1. *How do we predict appropriate bite timing in social situations?*

We propose taking insight from how able-bodied people eat in social settings to learn a model for bite probability as a function of social cues. To that end, we have completed a study in which able-bodied users have a meal together in a social setting while being recorded to capture social cues such as eye contact and gaze direction, and conversational pauses and turn-taking. Lower level features such as speaking volume, audio features, and time since the last bite will also be collected. Human social interactions are complicated and multifaceted, making the choice of features to use non-trivial. More details on this study can be found in section 2.8.

We propose fitting a model that predicts the probability of taking a bite based on the input features. Many algorithms for this type of supervised learning exist, and we will apply several such algorithms and compare the effectiveness of each.

Figure 3.1: Social cues are automatically extracted from videos on each participant as they are eating and conversing.

Figure 3.2: Experimental setup for gathering data to build a model for bite prediction. A table-top camera is pointed at each participant to gather gaze direction, bite timing, and detect who is speaking.

3.1.1 *Evaluation*

We will evaluate the bite timing prediction both objectively and subjectively. Objectively, the correct bite prediction rate on a test set would indicate how well the model matches with human bite timing. Subjectively, we will run a study in which a user is fed by the robot in a social setting, shown in fig. 3.2. The test condition will be the model used by the robot to choose when to present a bite of food. One baselines for comparison will be using a constant time between bites. The gold standard will be allowing the user to trigger bites. After each trial we will ask survey questions aimed to determine the naturalness and fluency of the timing and interaction.

3.2 *Adaptive Food Collection Techniques*

At present, the robot's goal is defined as a single fork pose above a potential bite. The robot then plans a straight downward motion for the fork to impale the bite, and then plans to a programmed location in front of the user's mouth. The downward motion has been tested to be effective for certain foods, such as cut fruit and marshmallows, which are sticky and remain attached to the fork after being impaled. This is a very small subset of possible food items, and even if we restrict ourselves to discrete bites of solid food, some items such as heavy strawberries and banana slices fall right off the fork tines when being lifted straight up. For example, a more effective strategy would be to angle the fork after impact, to better resist gravity's pull.

RQ 2. *How can we generate a database of effective food procurement actions?*

We propose that instead of hand-coding a number of food procurement actions — e.g. skewer, scoop, etc. — we learn a set of actions formalized as dynamic movement primitives (DMPs) from user demonstrations. The advantage of using learned actions is that it will already capture the nuanced motions that humans learn over a lifetime of using silverware.

To collect these demonstrations, we will record fork trajectories of people as they are eating using visual fiducial placed on the fork or physical sensors attached to the fork in such a way as to not impede the range of motion. Once the trajectories are gathered, we will either perform trajectory clustering to identify classes of motion, or hand labeling into logical categories. We will then fit a DMP to each type of motion.

DMPs are formalized as a set of differential equations which can be interpreted as a linear spring system perturbed by an external forcing term. In a 1-D system, the DMP is governed by the following equations.

$$\tau \dot{v} = K(g - x) - Dv + (g - x_0)f \tag{3.1}$$

$$\tau \dot{x} = v \tag{3.2}$$

Here x and v represent the position and velocity of the robot. In the feeding application, we are interested in the end-effector positions and velocities, so that is the representation we will use as opposed to the more common representation of joint states. The non-linear function f can be learned to generate arbitrarily complex motions. Using Gaussian basis functions with center c_i and width h_i, f is defined as:

$$f(s) = \frac{\sum_i w_i \psi_i(s) s}{\sum_i \psi_i(s)} \tag{3.3}$$

Here, s is a phase variable, which monotonically decreases from 1 to 0 during the movement in order to remove time-dependency, and is obtained by the equation:

$$\tau \dot{s} = -\alpha s \tag{3.4}$$

Once we have a database of possible actions to apply to collect food, we need a way of determining which action should be applied and which goal pose on the plate should be selected. Since we are using a DMP representation for the actions, and will be using the same basis functions, each action can be represented only be a set of weighting coefficients w_i.

RQ 3. *How can we autonomously choose an action and bite location based on features of the food?*

We propose using features of the food computed from the raw RGBD data and learning a mapping from these features to the DMP actions. There is a wide choice of features available to describe the food. We first present what features have been used with food in the past, then we present the strengths and weaknesses of possible features in this domain, and finally we discuss learning techniques that would could use.

The vast majority of research using computer vision with food detection and classification has been done in the context of quality assurance of a particular item. Algorithms have been designed for nuts, apples, oranges, strawberries, meats, cheese, and even pizza [Brosnan and Sun, 2002]. Each algorithm is tailored to the specific domain. A summary of the type of features that have been successfully used in these targeted applications is given in [Du and Sun, 2006] and summarized in table 3.1. While these features were successful in targeted applications, they may not be sufficient to distinguish *between* types of food.

Some recent work has tried to classify images of food into types of dishes, usually for the purpose of determining calorie counts [Bossard et al., 2014, Jiang et al., Sudo et al., 2014, Oliveira et al., 2014]. What follows are potential features or feature-extraction techniques that have been used in this domain. While we want to manipulate food, food classification is a complementary task, since the type of food is likely correlated with the way in which to manipulate it.

Just as there are several options for food features, there are many learning algorithms that have been applied to the food quality as-

Characterization	Products
Area	Apple
Hinge	Oyster
Color	Apple, Citrus, Lemon, Mandarin, Barley, Oat, Rye, Wheat, Bell pepper, Muffin
Morphological Features	Apple, Corn, Edible bean, Rye, Barley, Oat, Wheat
Textural features	Barley, Oat, Rye, Wheat, Edible bean
Spectral Images	Tomato, Poultry carcass
Hue histograms	Apple, Potato
Gradient magnitude	Raisin, Asparagus
Curvature	Carrot
Edges	Asparagus

Table 3.1: Summary of features used for quality assurance of different food products.

surance problem [Du and Sun, 2006]. We will formulate this as a supervised learning problem, where the labels are obtained from hand-labeling of sensor data through a GUI interface described in section 3.2.1.

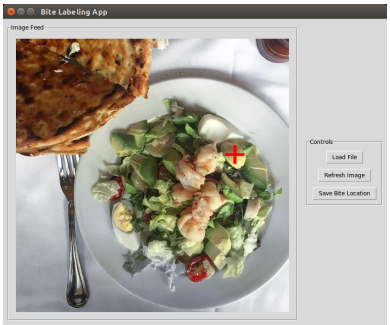
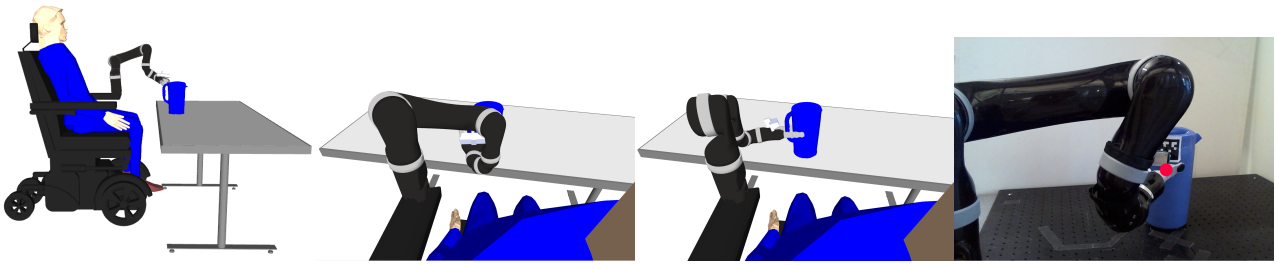


Figure 3.3: GUI for performing standardized hand labeling of bite locations in 2D.

3.2.1 *Collecting Training Data*

To collect training data to learn models to answer both RQ. 2 and RQ. 3, we will a way to track fork movements, as well as to provide useful bite location labels in bulk on raw gathered sensor data. To that end, we will develop a GUI that will display the sensor data and allow the human labeler to select several candidate bite locations. In 2D, the labeler can provide an (x, y) coordinate based on an RGB image and select from a list of motion descriptors (e.g. skewer, scoop). In 6D, the labeler can drag and rotate a fork model, then play one of the learned DMP motions to see how that motion would play out before submitting the final bite candidate location and primitive labels.

We will explore what databases are available that contain RGBD or RGB data for plates of food, with particular interest in databases that may contain partially eaten meals. Candidate datasets include the TST Intake Monitoring database [Gasparrini et al., 2015], Cornell Activity Datasets CAD-120 [Koppula et al., 2013], the RGBD-HuDaAct: A Color-Depth Video Database for Human Daily Activity Recognition [Ni et al., 2011], and the Leeds Activity Dataset [Aldoma et al., 2014]. If needed, we will supplement existing datasets with our own collection of RGBD images of plates of partially eaten food.



(a) Operating Robot Arm (b) Occluding the Pitcher (c) Showing the Pitcher (d) Eye Tracking

Figure 3.4: By generating paths that specifically minimizing occlusion, we can create better user focused manipulation.

3.3 *HRI Considerations*

Combining the results of section 3.1 and section 3.2, we will obtain a reliable autonomous method for collecting an arbitrary bite of food as indicated by a user. This system is not used in isolation, however. The robot is attached to the user's wheelchair and as such is working in very close physical proximity with the operator. This presents several unique human-robot interaction research questions, and amplifies others.

3.3.1 *Limited Viewpoint*

People who use a wheelchair have a restricted ability to change their viewpoint. Aside from moving the wheelchair base, which is presently impossible to do while moving the robot arm, they are limited to small torso and head movements to change their field of view while doing manipulation tasks. In addition, people with cervical SCI or multiple sclerosis (MS) may experience neck weakness, leading to a significantly decreased range of motion [LoPresti et al., 2000], and therefore a significant decreased range of viewpoint. Due to the limited range of viewpoint of the robot operator, the burden falls on the robot to ensure that important objects or parts of objects remain visible to the operator.

Task-oriented studies on eye gaze revealed that the eyes are centered on a point that is not the most visually salient, but is the best for the spatio-temporal demands of the particular job that needs to be done [LoPresti et al., 2000]. Some work has been done on determining what these spactio-temporal demands are in the context of object manipulation [Johansson et al., 2001], and their findings are that landmarks at which physical contacts take place are obligatory gaze targets. In their studies, the grasp site, the object's target location, and the support surface for placing the object at rest, are all obligatory gaze landmarks.We propose collecting eye gaze data from able-bodied and disabled users as they teleoperate the robot, to see if the same eye gaze findings hold for teleoperating a robot to perform

manipulation tasks. Preliminary data collection already supports the finding that certain parts of objects are consistently important, shown in fig. 3.5 for a pitcher pouring task.

The challenge will be in identifying relevant objects or object parts for a given task. We propose using a spactio-temporal gaze saliency score $s_g(t)$ for each object, or for parts of an object that is calculated using collected user data for that particular task.

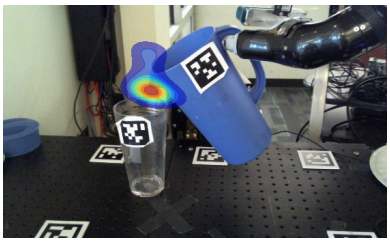


Figure 3.5: A heatmap of eye-gaze during the direct teleoperation task of pouring a glass of water from a pitcher.

RQ 4. *How can we determine the spactio-temporal gaze saliency of an object (or part of an object) for a given task?*

We will be collecting the gaze direction of users as they teleoperate the robot arm to perform the modified Chedoke tasks, and while using the robot arm to eat. During the trials, we will be tracking the location of all the objects, the user's head pose, and the robot's configuration. From this information, we can reconstruct the scene in order to calculate $s_g(t)$ for each point in the scene.

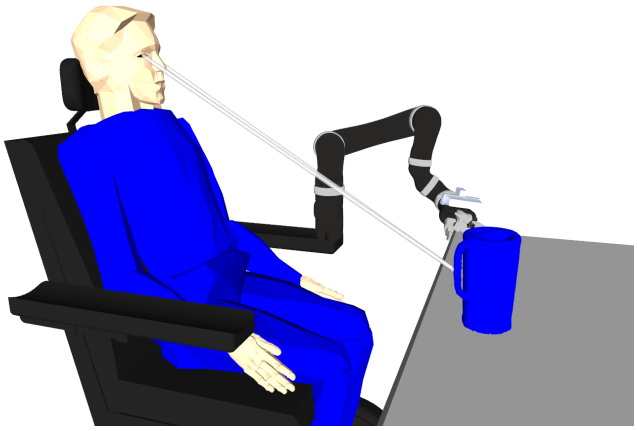


Figure 3.6: Visualization of a single moment in time of gaze and gaze target.

We assume for a moment that the spactio-temporal gaze saliency does not depend on time, but rather is intrinsic to the features of each object (for the given task). Take the example of pouring a glass of water from a pitcher. This would mean that the importance of seeing the pitcher's spout is independent of what phase of the task is being completed (grasping, lifting, pouring, or placing the pitcher). This is almost definitely not the case, but by ensuring visibility of the spout throughout the whole task, we ensure that it is visible during the relevant phase of the task.

We can think of the operator's eyes as a spotlight that is pointed at relevant parts of the scene (fig. 3.6). We accumulate the amount of light that is absorbed by each object's surface over the course of task execution to tell how salient that part of the object is. But the eye gaze is not a laser beam penetrating one object at a single point,

it diffuses with distance forming a cone . As it spreads, we will decrease the amount being added to the salience score as a function of the distance to the center of the gaze direction.

Since the gaze saliency is now a property of each object surface, we can compute the gaze saliency off-line from collected gaze data as a texture for the object mesh (RQ. 4). Finally, we can use the saliency to inform the robot's motion planning to optimize visibility of salient parts of the scene.

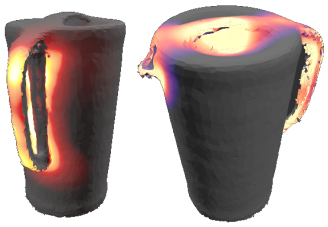


Figure 3.7: The pitcher textured with the gaze saliency calculated using simulated gaze data.

3.3.2 *Modal Control*

As we have thoroughly covered in chapter 2, using a modal control system presents teleoperation challenges. We have a method for performing automatic mode switching which we have tested with able-bodied users and in a simplified 2D setting. We propose extending this work and implementing mode switching with a full robot arm and running a study with users with disabilities to determine the effectiveness of this assistance strategy.

RQ 5. *Does time-optimal mode switching on an assistive robot arm improve task performance and human perceptions of the robot?*

Through collaborations with Kinova Robotics, the Rehabilitation Institute of Chicago, and UPMC, we have access to a disabled population with both novice and expert users of the JACO assistive robot arm. We will ask them to perform the modified Chedoke tasks in full teleoperation and with time-optimal automatic mode switching. The evaluation metrics will be the same as those in the completed study on time-optimal mode switching in the 2D setting chapter 2. To ensure that the time-optimal mode switching can be performed in realtime on the actual robot for the purposes of this study, we propose either modifying the search algorithm to use non-uniform discretization and other algorithmic speed-ups or pre-computing the time-optimal mode for each 6-D state offline.

3.3.3 *Level of Shared Autonomy*

In this proposal, we have outlined the many steps needed to create a robot autonomous feeding robot. We have motivated the need for autonomy by the difficulty of modal control and through interviews with expert users. But from the literature, there is evidence to support that even when control assistance has beneficial impact on performance, people with disabilities prefer manual control to feel more independent . To reconcile increasing performance while maintaining independence, we propose a study that will evaluate the

automated feeding system using different levels of shared autonomy.

RQ 6. *What is the best way to share control between the operator and the robot during feeding?*

There are many ways the robot and user could share control during feeding, but we will restrict this exploratory study to 6 conditions which encompass all of the research questions of this thesis, shown in table 3.2.

Study Condition		Bite Timing	Bite Location	Which DMP Action	Velocity Commands	Mode Switching
Full Autonomy	User					
	Robot					
Button-Triggered Bites	User					
	Robot					
2D Bite Location Selection	User					
	Robot					
2D Bite Location and Action Selection	User					
	Robot					
Direct Teleop with Mode-Switching	User					
	Robot					
No Autonomy	User					
	Robot					

Table 3.2: Study conditions for varying levels of shared control between the human operator and the robot. The blue boxes show when the human will be controlling that particular aspect of the task, and the orange boxes show when the robot will be controlling that aspect.

4
Summary of Proposed Work

Timeline

Topic	Section	Questions	Deadline
Autonomous Mode Switching (2D)[1]	2	RQ. 5	Oct. 2015 (HRI)
Thesis Proposal	–	–	Aug. 2016
Predicting Bite Timing for Assistive Feeding	3.1	RQ. 1	Oct. 2016 (HRI)
Viewpoint Optimization	3.3.1	RQ. 4	Oct. 2016 (HRI)
Learning feeding DMPs from demonstration	3.2	RQ. 2	Jan. 2017 (RSS)
DMP Selection for Food Acquisition	3.2	RQ. 3	Sep. 2017 (ICRA)
Indicating Goals in Shared Autonomy Feeding	3.3.2	RQ. 6	Oct. 2017 (HRI)
Autonomous Mode Switching (6D) with target population	3.3.2	RQ. 5	Oct. 2017 (ICRR)
Thesis Defense	–	–	Apr. 2018

Table 4.1: Proposed Timeline.
[1][Herlant et al., 2016]

Acknowledgements

This work was supported by the DARPA SIMPLEX program through ARO contract number 67904LSDRP, National Institute of Health R01 (award R01EB019335), National Science Foundation CPS (award 1544797), the Office of Naval Research, the Richard K. Mellon Foundation, the Google Hertz Foundation Fellowship Award, and the National Science Foundation Graduate Fellowship Program.

Bibliography

Abidemi Bolu Ajiboye and Richard F. ff. Weir. A heuristic fuzzy logic approach to emg pattern recognition for multifunctional prosthesis control. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 13(3), 2005.

Aitor Aldoma, Thomas Fäulhammer, and Markus Vincze. Automation of âĂIJground truthâĂİ annotation for multi-view rgb-d object instance recognition datasets. In *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5016–5023. IEEE, 2014.

Catherine M Arrington and Gordon D Logan. The cost of a voluntary task switch. *Psychological Science*, 15(9):610–615, 2004.

Michael Baker and Holly A Yanco. Autonomy mode suggestions for improving human-robot interaction. In *SMC (3)*, pages 2948–2953, 2004.

Susan Barreca, Carolyn Gowland, Paul Stratford, Maria Huijbregts, Jeremy Griffiths, Wendy Torresin, Magen Dunkley, Patricia Miller, and Lisa Masters. Development of the chedoke arm and hand activity inventory: theoretical constructs, item generation, and selection. *Topics in stroke rehabilitation*, 11(4): 31–42, 2004.

Lukas Bossard, Matthieu Guillaumin, and Luc Van Gool. Food-101–mining discriminative components with random forests. In *European Conference on Computer Vision*, pages 446–461. Springer, 2014.

G Bourhis and M Sahnoun. Assisted control mode for a smart wheelchair. In *Rehabilitation Robotics, 2007. ICORR 2007. IEEE 10th International Conference on*, pages 158–163. IEEE, 2007.

Tadhg Brosnan and Da-Wen Sun. Inspection and grading of agricultural and food products by computer vision systemsâĂŤa review. *Computers and electronics in agriculture*, 36(2):193–213, 2002.

Jun-Uk Chu, Inhyuk Moon, and Mu-Seong Mun. A real-time emg pattern recognition system based on linear-nonlinear feature projection for a multifunction myoelectric hand. *IEEE Transactions on Biomedical Engineering*, 53(11):2232–2239, 2006.

CS Chung and RA Cooper. Literature review of wheelchair-mounted robotic manipulation: user interface and end-user evaluation. In *RESNA Annual Conference [Internet]*, 2012.

Florence S Cromwell, United Cerebral Palsy Association, et al. *Occupational Therapist's Manual for Basic Skills Assessment Or Primary Pre-vocational Evaluation*. Fair Oaks Print. Company, 1960.

Fernando De la Torre, Wen-Sheng Chu, Xuehan Xiong, Francisco Vicente, Xiaoyu Ding, and Jeffrey Cohn. Intraface. In *Automatic Face and Gesture Recognition (FG), 2015 11th IEEE International Conference and Workshops on*, volume 1, pages 1–8. IEEE, 2015.

Kristine Turville Delano and Mary Hartman. Home health care: A needs assessment with design implications. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 45, pages 1289–1292. SAGE Publications, 2001.

Edsger W Dijkstra. A note on two problems in connexion with graphs. *Numerische mathematik*, 1(1): 269–271, 1959.

A. Dragan and S. Srinivasa. Formalizing assistive teleoperation. In *Robotics: Science and Systems*, July 2012.

Anca D Dragan and Siddhartha S Srinivasa. A policy-blending formalism for shared control. *The International Journal of Robotics Research*, 32(7):790–805, 2013.

Cheng-Jin Du and Da-Wen Sun. Learning techniques used in computer vision for food quality evaluation: a review. *Journal of food engineering*, 72(1):39–55, 2006.

Hå kan Eftring and Kerstin Boschian. Technical results from Manus user trials. *ICORR*, pages 136–141, 1999.

S Mohammad P Firoozabadi, M Asghari Oskoei, and Huosheng Hu. A human-computer interface based on forehead multi-channel bio-signals to control a virtual wheelchair. In *Proceedings of the 14th Iranian conference on biomedical engineering (ICBME)*, pages 272–277, 2008.

Ferran Galán, Marnix Nuttin, Eileen Lew, Pierre W Ferrez, Gerolf Vanacker, Johan Philips, and J del R Millán. A brain-actuated wheelchair: asynchronous and non-invasive brain–computer interfaces for continuous control of robots. *Clinical Neurophysiology*, 119(9):2159–2169, 2008.

Samuele Gasparrini, Enea Cippitelli, Ennio Gambi, Susanna Spinsante, and Francisco Flórez-Revuelta. Performance analysis of self-organising neural networks tracking algorithms for intake monitoring using kinect. In *Technologies for Active and Assisted Living (TechAAL), IET International Conference on*, pages 1–6. IET, 2015.

Laura Herlant, Rachel Holladay, and Siddhartha Srinivasa. Assistive teleoperation of robot arms via automatic time-optimal mode switching. In *Human-Robot Interaction*, March 2016.

Michael Hillman, Karen Hagan, Sean Hagen, Jill Jepson, and Roger Orpwood. The weston wheelchair mounted assistive robot - the design story. *Robotica*, 20:125–132, 2002.

Alberto Jardón Huete, Juan G Victores, Santiago Martinez, Antonio Giménez, and Carlos Balaguer. Personal autonomy rehabilitation in home environments by a portable assistive robot. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 42(4):561–570, 2012.

Yue-Min Jiang, Ho-Hsin Lee, and Cheng-Chang Lien. Automatic meal region extraction based on superpixel mapping.

Roland S Johansson, Göran Westling, Anders Bäckström, and J Randall Flanagan. Eye–hand coordination in object manipulation. *the Journal of Neuroscience*, 21(17):6917–6932, 2001.

Marjorie Kellor, J Frost, N Silberberg, I Iversen, and R Cummings. Hand strength and dexterity. *The American journal of occupational therapy: official publication of the American Occupational Therapy Association*, 25(2):77–83, 1971.

Dae-Jin Kim, Zhao Wang, and Aman Behal. Motion Segmentation and Control Design for UCF-MANUSâĂŤAn Intelligent Assistive Robotic Manipulator. *IEEE/ASME Transactions on Mechatronics*, 17(5):936–948, October 2012. ISSN 1083-4435. DOI: 10.1109/TMECH.2011.2149730.

Hema Swetha Koppula, Rudhir Gupta, and Ashutosh Saxena. Learning human activities and object affordances from rgb-d videos. *The International Journal of Robotics Research*, 32(8):951–970, 2013.

Kazuya Kubo, Takanori Miyoshi, and Kazuhiko Terashima. Influence of lift walker for human walk and suggestion of walker device with power assistance. In *Micro-NanoMechatronics and Human Science, 2009. MHS 2009. International Symposium on*, pages 525–528. IEEE, 2009.

Edmund LoPresti, David M Brienza, Jennifer Angelo, Lars Gilbertson, and Jonathan Sakai. Neck range of motion and use of computer head controls. In *Proceedings of the fourth international ACM conference on Assistive technologies*, pages 121–128. ACM, 2000.

T Luth, Darko Ojdanic, Ola Friman, Oliver Prenzel, and Axel Graser. Low level control in a semi-autonomous rehabilitation robotic system via a brain-computer interface. In *Rehabilitation Robotics, 2007. ICORR 2007. IEEE 10th International Conference on*, pages 721–728. IEEE, 2007.

Richard Mahoney. Robotic products for rehabilitation: Status and strategy. In *Proceedings of ICORR*, volume 97, pages 12–22, 1997.

Christian Mandel, Thorsten Luth, Tim Laue, Thomas Rofer, Axel Graser, and Bernd Krieg-Bruckner. Navigating a smart wheelchair with a brain-computer interface interpreting steady-state visual evoked potentials. In *Intelligent Robots and Systems, 2009. IROS 2009. IEEE/RSJ International Conference on*, pages 1118–1125. IEEE, 2009.

Virgil Mathiowetz, Gloria Volland, Nancy Kashman, and Karen Weber. Adult norms for the box and block test of manual dexterity. *American Journal of Occupational Therapy*, 39(6):386–391, 1985.

Michelle McDonnell. Action research arm test. *Australian journal of physiotherapy*, 54(3):220, 2008.

Ross Mead and Maja J Matarić. The power of suggestion: teaching sequences through assistive robot motions. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*, pages 317–318. ACM, 2009.

Nachshon Meiran, Ziv Chorev, and Ayelet Sapir. Component processes in task switching. *Cognitive psychology*, 41(3):211–253, 2000.

Stephen Monsell. Task switching. *Trends in cognitive sciences*, 7(3):134–140, 2003.

Bingbing Ni, Gang Wang, and Pierre Moulin. Rgbd-hudaact: A color-depth video database for human daily activity recognition. In *IEEE Workshop on Consumer Depth Cameras for Computer Vision in conjunction with ICCV*, 2011. URL `http://www.adsc.illinois.edu/demos.html`.

Daisuke Nishikawa, Wenwei Yu, Hiroshi Yokoi, and Yukinori Kakazu. EMG prosthetic hand controller discriminating ten motions using real-time learning method. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '99)*, 1999.

Marnix Nuttin, Dirk Vanhooydonck, Eric Demeester, and Hendrik Van Brussel. Selection of suitable human-robot interaction techniques for intelligent wheelchairs. In *Robot and Human Interactive Communication, 2002. Proceedings. 11th IEEE International Workshop on*, pages 146–151. IEEE, 2002.

Luciano Oliveira, Victor Costa, Gustavo Neves, Talmai Oliveira, Eduardo Jorge, and Miguel Lizarraga. A mobile, lightweight, poll-based food identification system. *Pattern Recognition*, 47(5):1941–1952, 2014.

Johan Philips, J del R Millan, Gerolf Vanacker, Eileen Lew, Ferran Galán, Pierre W Ferrez, Hendrik Van Brussel, and Marnix Nuttin. Adaptive shared control of a brain-actuated simulated wheelchair. In *Rehabilitation Robotics, 2007. ICORR 2007. IEEE 10th International Conference on*, pages 408–414. IEEE, 2007.

Patrick M Pilarski, Michael R Dawson, Thomas Degris, Jason P Carey, and Richard S Sutton. Dynamic switching and real-time machine learning for improved human control of assistive biomedical robots. In *Biomedical Robotics and Biomechatronics (BioRob), 2012 4th IEEE RAS & EMBS International Conference on*, pages 296–302. IEEE, 2012.

Oliver Prenzel, Christian Martens, Marco Cyriacks, Chao Wang, and Axel Gräser. System-controlled user interaction within the service robotic control architecture massive. *Robotica*, 25(02):237–244, 2007.

S. D. Prior. An electric wheelchair mounted robotic arm-a survey of potential users. *Journal of medical engineering & technology*, 14(4):143–154, 1990.

Erik Scheme and Kevin Englehart. Electromyogram pattern recognition for control of powered upper-limb prostheses: State of the art and challenges for clinical use. *Journal of Rehabilitation Research & Development*, 48(6):643–660, 2011.

Joris Sijs, Freek Liefhebber, and Gert Willem RBE Romer. Combined position & force control for a robotic manipulator. In *Rehabilitation Robotics, 2007. ICORR 2007. IEEE 10th International Conference on*, pages 106–111. IEEE, 2007.

Ann M. Simon, Levi J. Hargrove, Blair A. Lock, and Todd A. Kuiken. Target achievement control test: Evaluating real-time myoelectric pattern-recognition control of multifunctional upper-limb prostheses. *Journal of Rehabilitation Research & Development*, 48(6):619–628, 2011.

Tyler Simpson, Colin Broughton, Michel JA Gauthier, and Arthur Prochazka. Tooth-click control of a hands-free computer interface. *Biomedical Engineering, IEEE Transactions on*, 55(8):2050–2056, 2008.

Christer Sollerman and Arvid Ejeskär. Sollerman hand function test: a standardised method and its use in tetraplegic patients. *Scandinavian Journal of Plastic and Reconstructive Surgery and Hand Surgery*, 29(2): 167–176, 1995.

Carol A Stanger, Carolyn Anglin, William S Harwin, and Douglas P Romilly. Devices for assisting manipulation: a summary of user task priorities. *IEEE Transactions on rehabilitation Engineering*, 2(4): 256–265, 1994.

Tilo Strobach, Roman Liepelt, Torsten Schubert, and Andrea Kiesel. Task switching: effects of practice on switch and mixing costs. *Psychological Research*, 76(1):74–83, 2012.

Kyoko Sudo, Kazuhiko Murasaki, Jun Shimamura, and Yukinobu Taniguchi. Estimating nutritional value from food images based on semantic segmentation. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, pages 571–576. ACM, 2014.

Alan Sunderland, Deborah Tinson, Lesley Bradley, and R Langton Hewer. Arm function after stroke. an evaluation of grip strength as a measure of recovery and a prognostic indicator. *Journal of Neurology, Neurosurgery & Psychiatry*, 52(11):1267–1272, 1989.

Edward Taub, Karen McCulloch, Gitendra Uswatte, David M Morris, Mary Bowman, Jean Crago, Danna Kay King, Staci Bishop, Francilla Allen, and Sherry Yakley. Motor activity log (mal) manual. 2011.

Francesco Tenore, Robert S. Armiger, R. Jacob Vogelstein, Douglas S. Wenstrand, Stuart D. Harshbarger, and Kevin Englehart. An embedded controller for a 7-degree of freedom prosthetic arm. In *Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society*, 2008.

Francesco V. G. Tenore, Ander Ramos, Amir Fahmy, Soumyadipta Acharya, Ralph Etienne-Cummings, and Nitish V. Thakor. Decoding of individuated finger movements using surface electromyography. *IEEE Transactions on Biomedical Engineering*, 56(5):1427–1434, 2009.

Joseph Tiffin and ESTON J Asher. The purdue pegboard: norms and studies of reliability and validity. *Journal of applied psychology*, 32(3):234, 1948.

Hylke A Tijsma, Freek Liefhebber, and J Herder. Evaluation of new user interface features for the manus robot arm. In *Rehabilitation Robotics, 2005. ICORR 2005. 9th International Conference on*, pages 258–263. IEEE, 2005a.

Hylke A Tijsma, Freek Liefhebber, and J Herder. A framework of interface improvements for designing new user interfaces for the manus robot arm. In *Rehabilitation Robotics, 2005. ICORR 2005. 9th International Conference on*, pages 235–240. IEEE, 2005b.

Michelle Tipton-Burton. Jebsen–taylor hand function test. *Encyclopedia of Clinical Neuropsychology*, pages 1365–1365, 2011.

Katherine Tsui, Holly Yanco, David Kontak, and Linda Beliveau. Development and evaluation of a flexible interface for a wheelchair mounted robotic arm. In *Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction*, pages 105–112. ACM, 2008.

Katherine M Tsui and Holly A Yanco. Simplifying wheelchair mounted robotic arm control with a visual interface. In *AAAI Spring Symposium: Multidisciplinary Collaboration for Socially Assistive Robotics*, pages 97–102, 2007.

Gitendra Uswatte, Edward Taub, David Morris, Mary Vignolo, and Karen McCulloch. Reliability and validity of the upper-extremity motor activity log-14 for measuring real-world arm use. *Stroke*, 36(11): 2493–2496, 2005.

Ravi Vaidyanathan, Monique Fargues, Lalit Gupta, Srinivas Kota, Dong Lin, and James West. A dual-mode human-machine interface for robotic control based on acoustic sensitivity of the aural cavity. In *Biomedical Robotics and Biomechatronics, 2006. BioRob 2006. The First IEEE/RAS-EMBS International Conference on*, pages 927–932. IEEE, 2006.

Diana Valbuena, Marco Cyriacks, Ola Friman, Ivan Volosyak, and Axel Graser. Brain-computer interface for high-level control of rehabilitation robotic systems. In *Rehabilitation Robotics, 2007. ICORR 2007. IEEE 10th International Conference on*, pages 619–625. IEEE, 2007.

Dirk Vanhooydonck, Eric Demeester, Marnix Nuttin, and Hendrik Van Brussel. Shared control for intelligent wheelchairs: an implicit estimation of the user intention. In *Proceedings of the 1st international workshop on advances in service robotics (ASERâĂŹ03)*, pages 176–182. Citeseer, 2003.

Glenn Wylie and Alan Allport. Task switching and the measurement of âĂIJswitch costsâĂİ. *Psychological research*, 63(3-4):212–233, 2000.

H. A. Yanco. *Shared user-computer control of a robotic wheelchair system*. PhD thesis, Citeseer, 2000.

Holly A Yanco. Wheelesley: A robotic wheelchair system: Indoor navigation and user interface. In *Assistive technology and artificial intelligence*, pages 256–268. Springer, 1998.