A WEARABLE ROBOTIC FOREARM FOR HUMAN-ROBOT COLLABORATION

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Vighnesh Vatsal

December 2020

A WEARABLE ROBOTIC FOREARM FOR HUMAN-ROBOT COLLABORATION

Vighnesh Vatsal, Ph.D.

Cornell University 2020

The idea of extending and augmenting the capabilities of the human body has been an enduring area of exploration in fiction, research, and industry alike. The most concrete realizations of this idea have been in the form of wearable devices such as prostheses and exoskeletons, that replace or enhance existing human functions. With recent advances in sensing, actuation, and materials technology, we are witnessing the advent of a new class of wearable robots: Supernumerary Robotic (SR) devices that provide additional degrees of freedom to a user, typically in the form of extra limbs or fingers. The development, analysis, and experimental evaluation of one such SR device, a Wearable Robotic Forearm (WRF) for close-range collaborative tasks, forms the focus of this dissertation.

We initiated its design process through a basic prototype mounted on a user's elbow, and conducted an online survey, a contextual inquiry at a construction site, and an in-person usability study to identify usage contexts and functions for such a device, and formed guidelines for improving the design. In the next WRF prototype, we added two more degrees of freedom while remaining within acceptable human ergonomic load limits, and expanding its reachable workspace volume. We then developed the final prototype based on further feedback from a pilot interaction study, and found an analytical solution for its inverse kinematics. Going beyond static analyses with predefined

robot trajectories, we further addressed the biomechanical effects of wearing the WRF using a detailed musculoskeletal model, and developed a motion planner that minimizes loads on the user's muscles. Looking at the other side of the physical interaction between the user and WRF, we applied human motion prediction and feedback control for stabilizing the robot's end-effector position when subjected to disturbances from the wearer's body movements. Finally, we conducted a user study involving a collaborative pick-and-place task with the WRF acting in two conditions: responding to direct speech commands from the wearer, and predicting human intent using supervised learning models. We evaluated the quality of interaction in the two conditions through human-robot fluency metrics.

The WRF, and its associated systems described in this dissertation do have limitations, particularly in terms of ergonomics, feedback control performance, and fluency of interaction. However, as a prototype, the WRF shows that SR devices can be effective agents in human-robot collaboration when they possess capabilities for mutual adaptation while reducing the cognitive load on the user.

BIOGRAPHICAL SKETCH

Vighnesh Vatsal received his Bachelor's and Master's degrees in Mechanical Engineering in 2015 from the Indian Institute of Technology Bombay. He specialized in Computer Integrated Manufacturing, working on a multi-material Stewart Platform-based 3D printer for his Master's thesis. He joined the Ph.D. program at Cornell in Fall 2015, and commenced the Wearable Robotic Forearm project in Spring 2016 as part of the Human-Robot Collaboration and Companionship Lab. His research interests lie at the intersection of robot design, control systems, and human-robot interaction.

To Maitri, my beacon in stormy seas.

ACKNOWLEDGEMENTS

First and foremost, none of this would have been possible without Maitri, had it not been for her unwavering support through this long and sometimes painful process. The support system consisting of my family, lab mates, and friends in Ithaca and across the world are what kept me going.

In terms of the work itself, I owe immense gratitude to my advisor, Prof. Guy Hoffman, for being a true mentor. My journey at Cornell was somewhat serendipitous, as I did not know before coming here that Guy was joining the faculty at the same time. We started the Wearable Robotic Forearm project in 2016 as a spin-off of the drumming third arm from Georgia Tech, building up the lab in tandem with developing this robot. From Guy, I learned that it is important to not get bogged down by labels pertaining to fields of expertise, as creative output is often interdisciplinary. I would also like to thank my committee members, Prof. Hadas Kress-Gazit and Prof. Douglas MacMartin for their valuable guidance. Interactions with fellow students were just as helpful as coursework in the process of learning new things. For this I acknowledge fellow graduate students and members of the HRC^2 lab whom I had the privilege of working with: Rain Zhou, Jonah Mittler, Lauren Klein, Kevin Kruempelstaedter, Mei Zhang, Tatsuhiro Koshi and others. A special thanks to Michael Suguitan for being my sounding board for ideas on life and research all these years.

The administrators at Cornell, especially Dr. Marcia Sawyer, made the process of doing good research as smooth as possible, providing valuable guidance on almost every aspect of living and working in Ithaca. The sense of social community among students, staff, and faculty at Cornell is something that I will sorely miss.

This work was supported by the National Science Foundation under National Robotics Initiative Award no. 1734399 (NRI: FND: Collaborative Control for Wearable Robots).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Humans have long sought to go beyond the boundaries of their natural capabilities, especially in terms of motor and cognitive skills. Wearable devices for augmenting the human body have formed a part of this endeavor throughout history, with evidence of prostheses dating back to the Iron Age [1]. Until recently, actuated wearable devices have primarily been realized in the form of prostheses and exoskeletons, reaching considerable maturity in both research and commercial applications [20]. These wearable robots serve to replace human limb capabilities that have been lost, are used as rehabilitative tools, and can boost human musculoskeletal output [88, 19, 112].

Along with advances in prostheses and exoskeletons, we are witnessing the advent of another class of wearable robots: Supernumerary Robotic (SR) devices. These do not merely replace or support human limbs, but add degrees of freedom (DoFs) that are not naturally present in the human body, commonly in the form of additional limbs or fingers [33, 56].

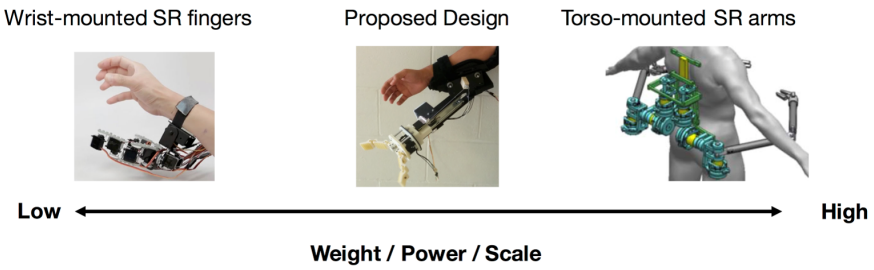


Figure 1.1: Spectrum of SR devices: ranging from lightweight and low-power [67], to high-power and large-scale [33] devices.

We explore a new wearable robot design which falls between torso-mounted arms [33] and wrist-mounted fingers [67, 118, 89] in terms of weight and power

(Figure 1.1): a lightweight supernumerary forearm attached at the elbow for close-range human-robot collaboration. This intermediate design can provide mobility and low weight—like SR fingers—allowing the user to quickly position it in a desired reference frame. Still, it is capable of increasing the user's reach and workspace beyond their natural limits—like torso-mounted SR arms.

We envision this Wearable Robotic Forearm (WRF) to be an autonomous agent, which the user can dynamically position and then collaborate with in a variety of ways. This is opposed to a fully user-controlled SR device with a similar form factor described in [29] which executes predefined motions selected through a button-operated interface. The WRF could pick-and-place objects which are out of the wearer's reach, aid human-human handovers when the wearer's hands are occupied, speed up repetitive tasks through self-handovers, and stabilize tools and objects within the wearer's workspace.

We describe the user-centered design process for the WRF in Chapter 2. Given the novelty of this human-wearable-robot configuration, we set out to explore the usage contexts and interaction scenarios that such a device may be deployed in, using an initial prototype (Model I) as a reference point. Through an online survey, we found that the WRF is seen as a functional tool in professional settings. This led to a contextual inquiry at a construction site, where we obtained need themes for the WRF based on workers' tasks. An in-person usability study was also conducted where participants wore the device and performed two open-loop controlled tasks, grounding the design guidelines for future prototypes in real-world interaction.

We analyze the design of the next prototype (Model II) in Chapter 3. Using guidelines obtained from the user-centered inquiries, we made structural

Figure 1.2: Successive wearable robotic forearm prototypes: Models I, II, and III (left to right).

changes to the WRF, adding two more degrees of freedom while reducing the weight from ˜2.0 kg to ˜1.5 kg through the use of lightweight structural components. We describe its physical structure, kinematics, and determine that the enhancement in a user's reachable workspace volume while wearing Model II is 246% over the volume spanned by the human arm's normal range of motion. A rigid-body analysis of the combined human and robot shows that the forces and moments experienced by a user at their elbow and shoulder joints lie well within ergonomic load limits. Following a pilot interaction study, we determined that another degree of freedom was desirable at the WRF's wrist joint, leading to the final prototype (Model III) weighing ˜2.0 kg with five DoFs and a gripper. The progression in WRF design is shown in Figure 1.2. We describe the forward and inverse kinematics (IK) for Model III, developing an analytical solution for the IK, which was applied to the real robot along with more common numerical approaches.

In Chapter 4, we further expand on the initial rigid-body biomechanics analysis of the WRF's effects on the wearer's arm. While we had earlier established that the robot exerts loads within ergonomic limits based on predetermined trajectories, we now focus on finding robot paths, given a start and end state, that minimize load on the wearer. Using a high-fidelity musculoskeletal model, we

determine the forces generated in the human arm muscle fibers due to a particular WRF trajectory, and construct a cost function for an optimizing stochastic motion planner. Through this process, we obtain WRF trajectories that minimize load on the user's arm muscles.

While the above analyses estimate the kinematic and biomechanical effects of the WRF on the user, there is another aspect of this interaction: the disturbance introduced in the robot's intended pose by virtue of it being mounted on the human's arm. In Chapter 5, we model the actuators of the WRF as linear systems and develop feedback control strategies for stabilizing its end-effector in light of disturbances due to human motion. Starting with a planar scenario, we develop a time series model for human arm motion prediction in 2D, and generate set points for the WRF's motors that compensate for sensing and actuation delays, leading to improved stabilization performance. We extend this approach to full 3D task scenarios, developing another time series model as well as adapting a recurrent neural network model for human motion prediction and end-effector stabilization.

Through a user study described in Chapter 6, we experimentally evaluate the quality of interaction with the device in two modes. As opposed to some other SR devices where the wearer directly controls the robot [67, 118, 55, 29], we explored a scenario where the robot would act autonomously in structured tasks in order to reduce the user's cognitive load. Users gave direct speech commands to the WRF in a collaborative assembly and pick-and-place task, while their body movements were recorded. This speech command-annotated movement data was used to train supervised learning models for determining the user's intent. Another round of trials was conducted with these trained models

in which the robot would act autonomously based on user intent prediction. We compared the objective and subjective fluency of interaction, where the WRF in its human intention predictive mode was found to improve upon some aspects, such as reducing the mean trial times and robot idle times, along with exhibiting positive traits as a teammate. Through the aforementioned studies and analyses, we found the WRF to be a moderately useful augmentation in close-range collaborative tasks.

Limitations are bound to arise in any novel application area of robotics, with SR devices in particular presenting additional challenges due to the human being an intrinsic part of the system. Starting with the physical design, while the biomechanical loads due to the WRF were theoretically within ergonomic limits, users found it to be cumbersome in terms of weight, and experienced fatigue on their upper bodies when worn for extended periods of time. For the biomechanical motion planning framework, major challenges need to be addressed before it may be deployed in a real SR device, primarily regarding its computational expense, as well as with guaranteeing its efficacy across human physiologies. While we achieved some improvement in accounting for the disturbances to the WRF's end-effector caused by the user's movements by using human motion prediction models in the feedback control loop, these results may not be satisfactory in a real setting, especially if the robot were holding heavier objects, or if the task did not involve repetitive human actions. In terms of fluency during the user study, the WRF was found be less trustworthy in the human intention predictive mode compared to the direct speech control mode. In some cases, owing to limitations in the predictive models, users had to expend even greater cognitive effort to rectify a situation where the robot acted on an incorrect prediction.

While having the human attached to the robot adds uncertainty in the planning of SR device behaviors and motions, it can also be beneficial in situations where the human and robot are able to mutually adapt in a manner that does not add to the user's cognitive load, and minimizes the physical load. To achieve this goal, various facets of robotics need to be accounted for: design, kinematics and biomechanics analysis, motion planning, control, and human intention prediction. This dissertation explores each of these aspects, and studies their relative importance in building a truly collaborative wearable robotic augmentation.

CHAPTER 2

USER-CENTERED DESIGN PROCESS

In this chapter, we describe the design process for the Wearable Robotic Forearm, which took into account inputs from potential users at various stages of development.

Existing SR designs include robotic arms mounted on the torso or shoulders, with form factors similar to human arms. They are typically used for supporting a worker's body in settings such as aircraft manufacturing [84] in standing or crawling-like positions [64], for bracing an object while the user works on it [71], or for more close-range tasks such as playing the drums [61]. A second common configuration is wrist-mounted robotic fingers, designed to perform two-handed tasks with a single hand [118, 56], or providing an eleventh finger in more precise tasks such as playing the piano [32].

Aside from these human-mimetic approaches, there have also been animal-inspired designs in the form of robotic tails for balance assistance [74, 78], and a wearable snake-like robot for assisting daily activities [13]. Advances in soft robotics have allowed for safer, lightweight SR devices with flexible linkages and pneumatic actuators [79].

These devices are commonly controlled using force sensing from the human hand, or by signals from a myoelectric armband [67]. Other interfaces include direct commands through dials [29], push-buttons [55], foot-control [94, 37, 9], and gloves with flex-sensors [82, 14].

The torso-mounted SR arms are capable of up to 50–70 Nm of torque, while the additional fingers typically weigh under 0.5 kg, with torques of up to 2 Nm.

The WRF is aimed at tasks with lower demands than the torso-mounted robots above, leading to a smaller footprint. At the same time, its configuration allows for extended reach and multi-location work capabilities, in contrast to wrist-worn robots. This is similar in terms of physical design to [29], but with a different paradigm for interaction: the WRF is aimed to be an autonomous assistive agent, as opposed to being directly controlled through a push-button interface. Interaction with SR devices has been automated in task-oriented settings through demonstration-based control [71], and user intent detection [79, 69]. We describe supervised-learning based approaches for achieving the same in Chapter 6.

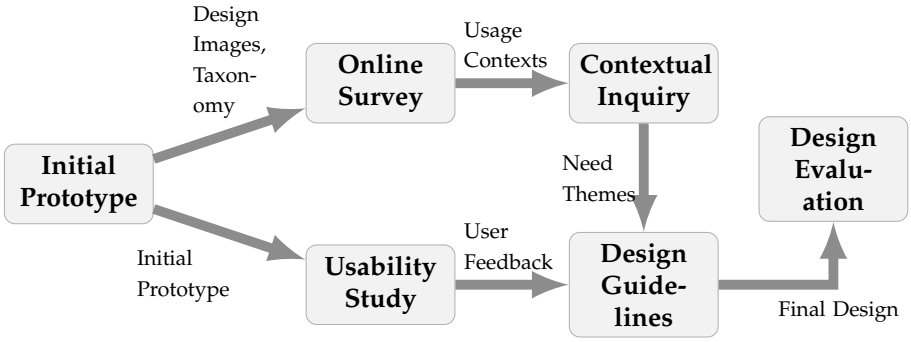


Figure 2.1: Steps involved in user-centred design of the wearable robotic forearm, going from initial concepts to an evaluated functional design.

Focusing on the user-centered design process in this chapter, its stages are shown in Figure 2.1. We followed study designs and methodologies from [21]: using an initial prototype (Model I) to conduct surveys and usability studies and generate design guidelines for the next prototype.

2.1 Model I Prototype

We developed an initial version of the device, in the spirit of low-fidelity ("paper") prototypes [103], to serve as a starting point for the design process. This prototype was a fully functional, albeit non-autonomous, wearable robot. The physical realization conveyed to study participants a functionality that is similar to that envisioned for the final design, and allowed users to experience the device in an embodied manner, enabling design backtalk [62].

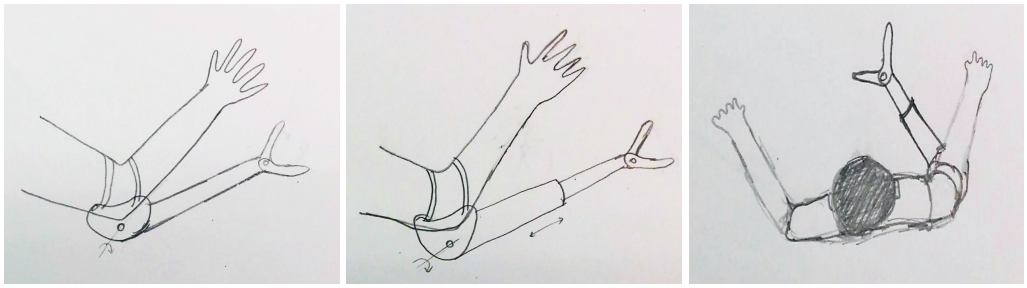


Figure 2.2: Initial concept sketches for an elbow-mounted third arm robot (from left): single vertical DoF, one vertical and one prismatic DoF, and one horizontal and one prismatic DoF (chosen design).

In the development of the prototype, we were not concerned with its formal appearance, but were focused on the degrees of freedom (DoFs). As a result, we built a skeletal version of the device which moves in the way we envision, but whose aesthetic design was rudimentary. Initial sketches exploring the DoFs replicated the flexion and extension of the human elbow on the robotic arm (Figure 2.2a). This would allow a user to reach objects below the level of a fully extended elbow, as well as enable self-handovers. For instance, when standing on a ladder, the wearable arm could reach down to bring a tool without requiring the wearer to step down. The addition of a prismatic joint further extends the wearer's reach (Figure 2.2b). We also envisioned using the arm around a workbench or a desk, where its vertical movement might interfere with the

workspace. We therefore explored a different degree of freedom in the form of a horizontal panning that is analogous to the horizontal adduction and abduction of the human shoulder (Figure 2.2c). This side-to-side design, in addition to the prismatic extension, effectively broadens the wearer's "wing span." A gripper was chosen as the end effector to enable grasping of objects and bracing. This choice of DoFs is also suitable for two people working side by side, enabling them to stand further apart from each other during handovers.

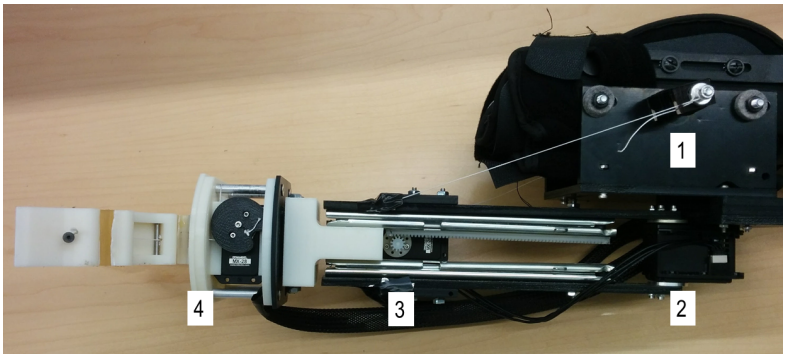


Figure 2.3: Structure of the prototype: 1) Mounting Platform, 2) Motor for panning, 3) Motor, rack and pinion for prismatic length extension, 4) Gripper

The assembled Model I prototype is shown in Figure 2.3. Weighing ˜2kg, its body was realized out of laser-cut acrylonitrile butadiene styrene (ABS) sheets (shown in black) and 3D printed ABS components (shown in white). Stainless steel rolling slides enabled the prismatic length extension. The horizontal panning DoF was direct-driven, and the prismatic DoF was actuated using a rack-and-pinion transmission. The gripper was based on the Yale OpenHand Model T42 [80], modified to constrain both fingers to move together by attaching their cables to a single motor. It had a compliant structure, with rubber connectors between finger segments. This design, along with rubberized fingertips allowed for reliable gripping of objects that are primarily prismatic. The hand itself was 3D printed in ABS, and the elastic components were cast out of

PMC 780 urethane rubber.

As shown in Figure 2.3, a mounting platform attached the robot to a two-joint medical support brace. The elbow adduction and abduction DoF (horizontal panning) was actuated by a ROBOTIS Dynamixel MX–64 motor [6] with its axis pointing downwards, and a range of motion of 120°. The pinion gear in the length extension DoF was actuated by a Dynamixel MX–28 motor, with ˜160mm range of extension of the drawer slides.

The initial prototype was tethered, receiving both power and control commands through cables. This design choice was made over powering it through a battery pack, as it would have added weight without significantly improving user experience in our trials. The motors were controlled using a Robotis CM-700 board [6].

2.2 Online Study: Contexts and Functions

After constructing the prototype, we conducted a brainstorming session [108] with colleagues to produce a large number of possible use cases for a wearable robotic arm. We also shared pictures of the prototype on social media and collected open-ended responses, through snowball sampling [15], to the question: "What would you use a wearable robotic third arm for?" We then categorized the collected responses into a taxonomy of usage contexts and functions (Figure 2.4).

Usage **contexts** are groupings of *where* a wearable robotic arm would be useful. Usage **functions** are groupings of *what* such a device would be useful for.

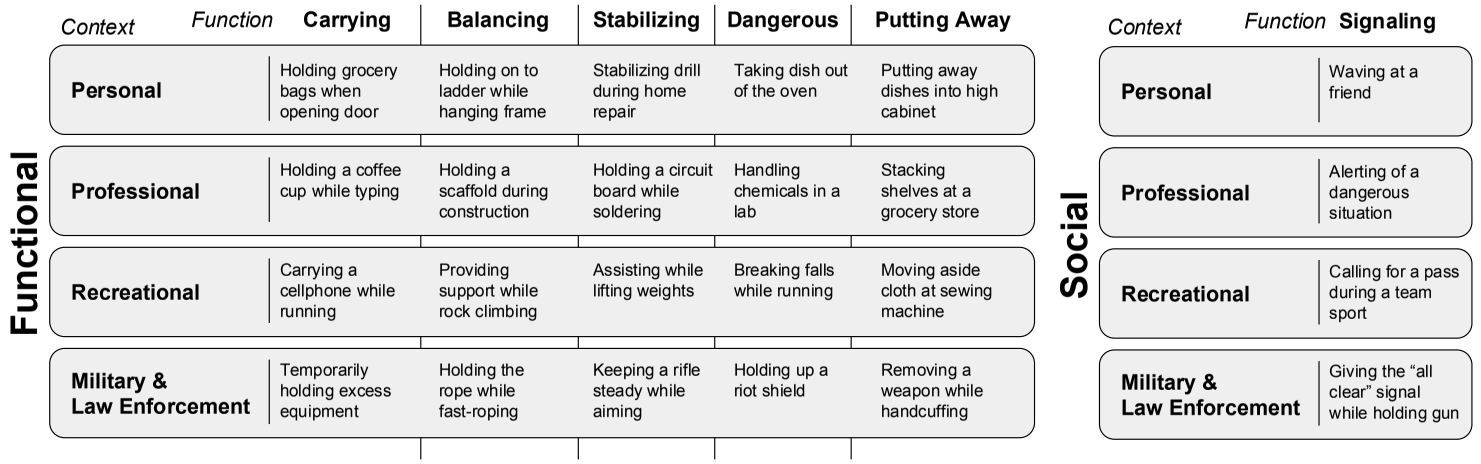


Figure 2.4: A taxonomy of contexts and functions of use for a wearable robotic arm, developed through brainstorming sessions, open-ended snowball sampled surveys, and affinity diagram clustering, along with non-exhaustive illustrative examples.

Clustering the usages via affinity diagrams led to the identification of four usage contexts:

- **Personal:** Error-tolerant tasks in environments familiar to the user, supporting daily activities.
- **Professional:** Tasks performed in office and industrial contexts, requiring more robustness from the robot.
- **Recreational:** Hobby or fitness related tasks, possibly outdoors or in unfamiliar environments.
- **Military and Law-Enforcement:** High-risk tasks in uncertain environments.

Orthogonal to these contexts, we identified five usage functions:

- **Carrying** objects and performing human-robot handovers.
- **Balancing** the user by grasping and bracing using objects in the environment.
- **Stabilizing** an object that the user is holding.
- Handling **dangerous** objects such as chemicals and hot plates.
- **Putting away** objects to outside the wearer's reach or while hands are occupied.
- **Signaling**, for example, using the robotic arm to gesture to a coworker for assistance.

Finally, we noted a high-level distinction between functions that are purely pragmatic or "functional" and those involving social interactions. This differentiation stems from the fact that beyond its physical assistive function, a third arm offers additional modalities for expressing non-verbal behavior. We refer to this functional/social dichotomy as usage "classes". This is similar to the distinction made between assistive robots and socially assistive robots [39].

2.2.1 Survey Procedure

To inform the design features and requirements of the device, we collected responses to an online survey gauging public opinion of potential usage contexts and functions based on the above taxonomy, in the spirit of [109].

We collected 105 responses from participants recruited using the Amazon Mechanical Turk platform [25]. The age distribution was: 18–25 (12.4%), 26–35 (45.7%), 36–50 (28.6%), 50 and above (13.3%).

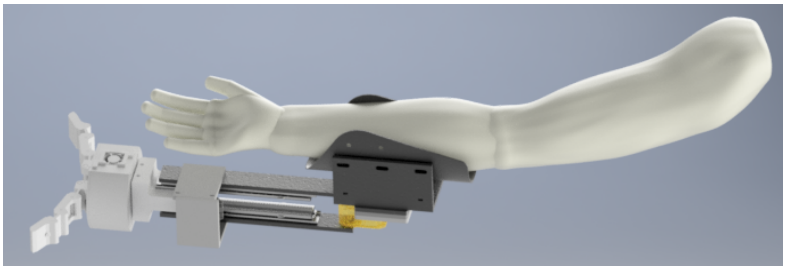


Figure 2.5: Image of Model I shown in the online survey.

The survey showed images of a 3D model of the arm and pictures of the physical prototype, as in Figures 2.5 and 2.3, followed by the text:

We are building a smart robotic "third arm" that attaches at your elbow.

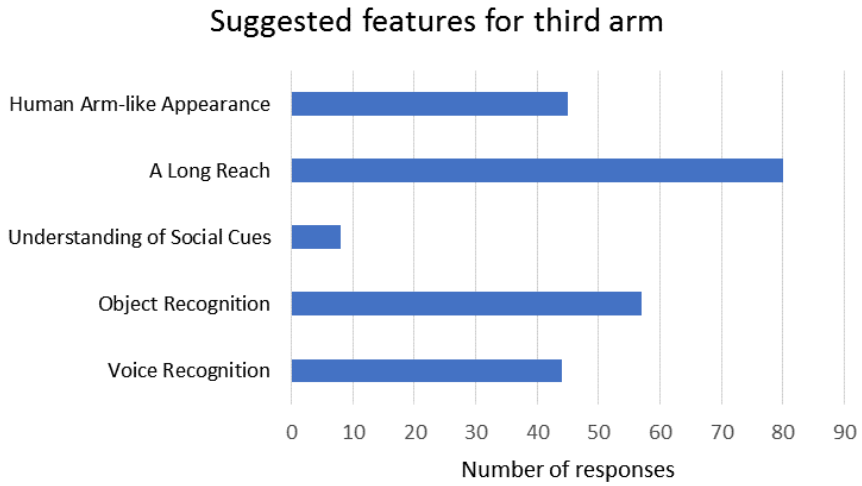


Figure 2.6: Trends in the responses to online survey questions about desirable features for the third arm.

> *The first prototype shown below has three motors: rotation at the elbow, arm extension, and gripping. The purpose of this survey is to gauge application areas and features, which will motivate development of future prototypes.*

Then, we presented participants with three sections, one for usage context and usage class, one for specific functions, and one for desired features, along with space for open-ended responses and demographic questions.

In each of the first two sections, we presented the context, class or function of use, alongside two activity examples. For each context, we chose one example from the functional list of activities, and one from the social list: "PERSONAL USE (e.g., holding grocery bags while opening a door, or shaking hands with multiple people)." For each of the use classes and functions, we selected activity examples from different contexts: "FUNCTIONAL USE (e.g., holding a coffee cup while typing, providing support while rock climbing, or stabilizing a firearm)." The list was preceded by the phrase: "Think about the following

Table 2.1: Online survey response scores about contexts and classes of use (mean, standard deviation and mode)

A robotic third arm is useful for..			
Usage context	$\bar{x}$	σ	mode
Personal	4.80	1.95	6
Professional	5.00	1.61	4
Recreational	4.51	1.70	4
Military/Law Enforcement	5.02	1.68	6
Usage class	$\bar{x}$	σ	mode
Functional	5.77	1.41	7
Social	2.82	1.60	1

I can see myself using a robotic third arm for..			
Usage context	$\bar{x}$	σ	mode
Personal	4.17	2.23	1
Professional	3.30	1.93	1
Recreational	3.63	2.13	1
Military/Law Enforcement	2.38	1.98	1
Usage class	$\bar{x}$	σ	mode
Functional	5.08	2.13	7
Social	2.23	1.60	1

contexts [functions] for a robotic third arm." We asked two questions per context and class, and two questions per function, each on a scale of 1–7 ("Not at all useful" to "Extremely useful"), shown in Tables 2.1 and 2.2. We also asked respondents to provide examples of usage contexts and features for the robot (Figure 2.6).

2.2.2 Results

Table 2.1 shows the means, standard deviations, and modes for each context and usage class. A wearable robotic third arm was considered more useful as a functional tool than for social uses by a wide margin. Moreover, people thought it was more useful in professional and military settings, and least in recreational contexts. Similarly, respondents could generally see themselves using a third

Table 2.2: Online survey response scores about functions (mean, standard deviation and mode) in decreasing order of means.

How useful is a robotic third arm for..			
Function	$\bar{x}$	σ	mode
Handling dangerous objects	5.87	1.58	7
Carrying things	5.47	1.70	7
Stabilizing an object	4.77	1.74	6
Putting things away	4.19	1.91	5
Signaling to others	3.56	1.72	4
Balancing the user	3.41	1.89	4
I can see myself using a robotic third arm for..			
Function	$\bar{x}$	σ	mode
Handling dangerous objects	5.06	2.14	7
Carrying things	5.02	2.20	7
Stabilizing an object	4.17	2.11	1
Putting things away	3.75	2.14	1
Balancing myself	2.73	1.97	1
Signaling to others	2.66	1.83	1

arm more for functional use, but chose the personal context as more likely for their own use. Overall, respondents saw themselves less likely to use such a robot compared to how useful they rated it to be overall.

Table 2.2 shows the means, standard deviations, and modes for each function. Handling dangerous objects and carrying were the highest rated functions for a third arm (mode 7 for both questions in Table 2.2), with stabilizing objects and putting things away being rated as generally useful (if not for own use). There was a low usage expectation for social functions.

This discrepancy between general usefulness and respondent-use could be explained by the fact that people thought of the robot more as a professional tool, and it is unlikely that they would have worked in particular settings in which a wearable robotic arm would be used. Also, it may be hard to imagine oneself using a device which is of a category that is unfamiliar to respondents.

While they could see the utility of the robot in general functional use, if not

their own, a third arm was decidedly not considered useful for social contexts or signaling functions, either by the users themselves or in general. This is also reflected in the number of responses for each of the suggested features (Figure 2.6), where functional aspects such as an enhanced reach were considered to be more important than understanding social cues.

These results were obtained from a text-based description and images of Model I displayed online. Discrepancies in perceived utility of the device arose between the online survey and physical studies, e.g. balancing a user was considered important in the contextual inquiry described in the next section, while online survey participants did not consider it to be a useful function. In order to proceed with the development of future prototypes, the design guidelines needed to be grounded in insights from a real usage scenario, as well as physical interaction with the device.

2.3 Contextual Inquiry: Building Construction

The findings from the online survey suggest that a wearable robotic arm can be a valuable tool in a professional setting, where carrying things, handling hot or dangerous objects, and stabilizing and putting objects away are key functions. Informed by these results, we proceeded to conduct a need-finding inquiry [49] to guide the design of specific capabilities. We chose the domain of building construction, which includes many of the above-mentioned functions. Importantly, this trade involves a range of activities that are strenuous, repetitive and present some degree of hazard to a worker. This is reflected in the fact that, in the United States, about 40% to 65% of worker's compensation costs in con-

struction result from musculoskeletal disorders and soft-tissue injuries which develop over time [91]. A robotic third arm may be deployed to reduce this risk from repetitive injuries in tasks classified as handovers, pick-and-place, and stabilization of a worker or tool.

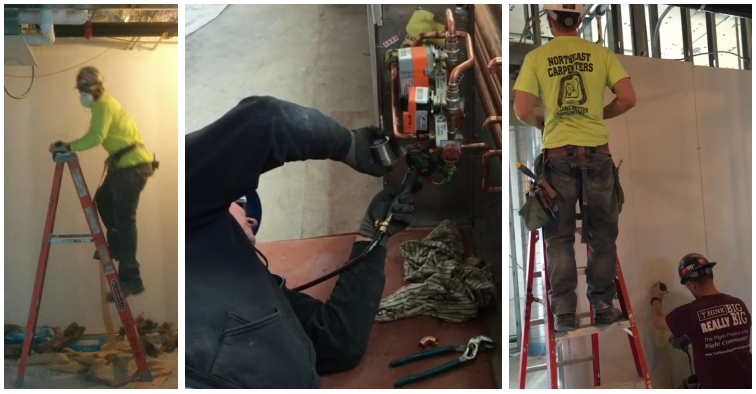


Figure 2.7: Observations from the construction site (left to right): a roof paneling installer climbing a ladder, plumber welding copper pipes, and collaborative drywall installtaion.

We conducted the contextual inquiry with a construction crew working on building renovations at the Cornell University campus. We were guided through the site by a supervisor who provided a brief description of each task and offered expert testimony during and after our observations of each worker. Since we are interested in building a physical augmentation device, we documented the body motions and ergonomic states of a worker while performing a task along with the hazards and loads associated with the task. We also elicited comments from the workers about the cognitive loads and common frustrations involved in their tasks.

Based on these observations, we identify three usability "need themes" informing promising functional requirements of the robot.

Reaching and Self-Handovers

We observed multiple instances where a worker would reach for a tool or work piece in a way that impedes their current task. For example, a worker installing roof panels while standing on a ladder would frequently bend down and pick up tools placed on the ladder or in his utility belt. He also had to step down from the ladder to consult plans. Once finished at a particular location, he would then carry the ladder after having to place those objects back onto the ladder or in his pockets.

Another instance involved an electrician who had to modify a control panel that had been installed in the midst of ventilation piping in the ceiling. To reach the panel in the tightly enclosed space, he had to climb up a ladder and take off his safety helmet to be able to unfasten the screws on the panel. He mentioned that the task would have been much easier if a tool could reach into the enclosed space and bring the screws to him after removal.

This suggests that a third arm would need to function as a tool for handovers and as a temporary storage space, while extending the reach of the wearer. Depending on the dexterity, it would allow a user to perform complex operations in constrained workspaces. The ability to bring nearby objects to a user would reduce the time and effort expended to bend, or to climb up and down to fetch objects.

Stabilization of Objects and Self

We found numerous tasks where adding another point of support for bracing a worker would enhance their safety and comfort. The supervisor mentioned

that safety regulations require a worker to have three points of contact with a ladder at all times (Figure 2.7a). In practice, this is difficult to achieve, especially in bi-manual tasks. When both hands of a worker were occupied, we noticed that they braced against the ladder with their stomach or hip.

A plumber installing copper piping for heating units described a challenge during soldering operations. He would lay down on the floor and have to hold his blowtorch in one hand, and attempt to feed more solder to the joint while also holding the pipe steady with his other hand (Figure 2.7b). Another instance was of a cement-layer who constantly had to brace against the floor while spreading a layer of cement.

This suggests that a wearable third arm should be able to stabilize objects and provide for additional contact points for workers when balanced in uncomfortable positions.

Coordination of Repetitive Actions

In tasks performed in pairs, workers tend to develop coordination strategies as a result of repetition. For example, one person would cut a gypsum board for drywall installation, while his colleague would hold it in place and brace it against a wall frame. The first person would then get up on a ladder and nail the board into place at positions marked out by the colleague (Figure 2.7c). They would perform this series of tasks fluently with minimal communication and acknowledgement from each other, as a result of having done these tasks together multiple times.

Another example of collaborative activity performed in pairs was by win-

dow installers: one worker would bring nails and ties from a bin to the worker installing the window panel. This repeated handover task proceeded with fluency and coordination to the point where the installer was able to anticipate the handover without even looking at the other worker.

This suggests the need for a robotic device deployed in such scenarios to not just be physically robust and capable, but also be able to coordinate fluently in repetitive activities, leading to a reduction in the cognitive load of a worker.

2.4 In-person Usability Study

The two studies described thus far provide us with a selective focus for contexts and functions, and with specific user needs in a potential application area. However, both the online study and the contextual inquiry were conducted on a purely conceptual basis, with participants imagining the use of a wearable third arm. To generate actionable design principles grounded in physical interaction with the device, we also conducted a user study with participants wearing and using the prototype arm, followed by semi-structured interviews.

2.4.1 Study Design

This usability study had three phases: a preliminary interview, interaction with the device, and debriefing. In the initial interview, participants began by describing a typical day in their lives. After identifying some activities at home, at work, and performed for recreation, we asked them to imagine if having a third arm attached to their body would affect these activities, following a semi-

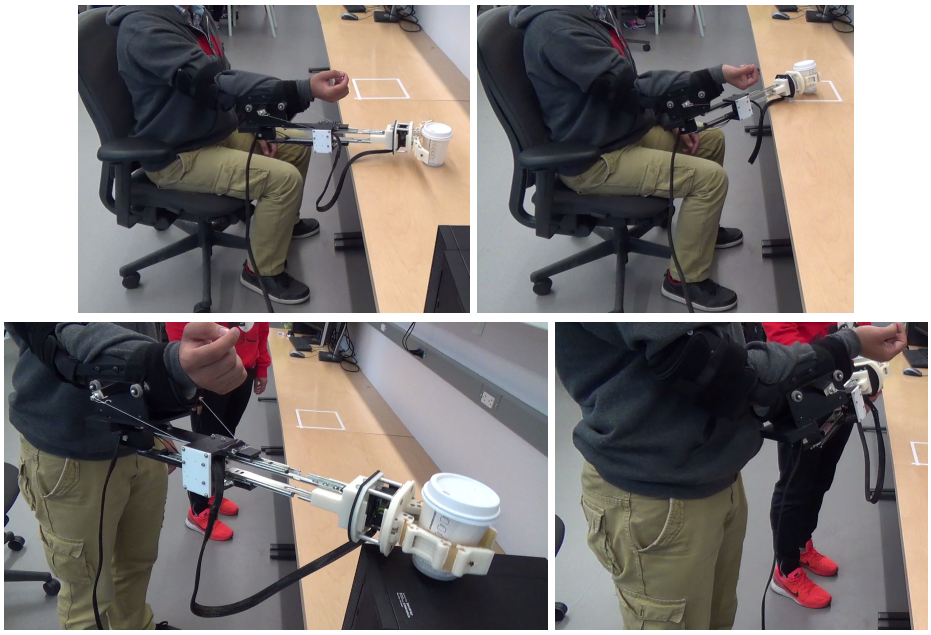


Figure 2.8: Tasks performed by users: moving a cup on a table while seated, and handing over a cup to the interviewer while standing.

structured approach [17]. We questioned them about the structure, appearance and capabilities of the hypothetical third arm. In order to narrow down their thought process towards forearm mounted devices, we showed the participants images of a 3D model of the prototype. We then repeated the questions about their daily activities and elicited suggestions for changes or improvements to the device at this stage.

After responding to the pictures, participants were shown the physical prototype. They proceeded to wear it and perform two scripted tasks: moving a coffee cup on a table while seated, and handing over the cup to the interviewer (Figure 2.8). During the tasks, the robot followed preset trajectories in an open-loop manner, i.e., without feedback, sensing, or adaptation. Finally, participants were debriefed and asked for improvements and suggestions that they would like to see in future prototypes, and features they would like to see in a com-

mercial product.

In debriefing, participants were asked about reactions to the prototype, and if they could think back now to the their daily activities, and if such a device might be useful.

The participants were 14 university students at the graduate or undergraduate level. We recruited participants by distributing fliers throughout the university campus, and sending out e-mails on a special interest forum for robotics. Each person received a $10 gift card for participating in the study.

2.4.2 Findings from Interviews

A qualitative analysis of audio and video recordings from the interviews revealed five recurring themes (Figure 2.9):

Weight and Balance

The weight of the robot was a major concern for participants. Users often struggled to perform the task, and in one case, even had to hold up the third arm with their free hand. Reactions included: *"It was very heavy, very cumbersome to use,"* and *"I could not imagine holding it up for more than ten minutes."* It was also felt that the weight would make it difficult to perform activities like typing with your own hands if the device was attached. Reducing weight was the most cited concern, and it is a natural requirement of any wearable device. However, even if material and actuation constraints require a heavier device, designers should consider distributing the weight along the user's body. We attached the pro-

totype using a two-joint medical support brace, intended for an average-sized adult male. As a result, some participants felt that it tended to slip off if they had smaller arms, and have their arms twist inside the harness when the robot swiveled at the elbow. Suggestions for the attachment point to the body highly depended on the use context: The arm should mount at the *"center [of the torso], to keep it symmetrical while running,"* or *"On the back, like a crane."* Generally, a robot directly attached to the forearm was often not considered to be a desirable location.

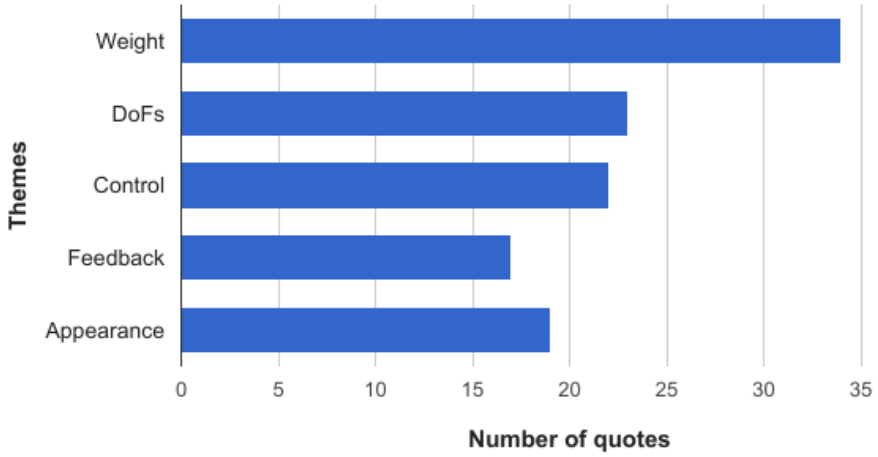


Figure 2.9: Number of instances of the five recurring themes in the oral feedback from user study participants.

Dexterity

Participants desired more dexterity than was presented to them, especially from the end effector: *"wrist rotation would be desirable," "maybe it should have more fingers."* Many users commented that it would have been easier if the third arm were able to *"pivot in the vertical direction."*, or have an *"axis where it can go up and down"*.

The particular DoFs desired by participants were strongly tied to the application domain. A participant who had worked in chemistry labs said that it would be helpful if the third arm could *"scoop out [small] amounts of powders with a small spatula."* Another user who worked at a library said that the end effector should be able to *"open the cover of a book."* A longer reach during arm extension was desirable for a user who had to regularly reach up well above their head to press buttons on cameras in their lab.

Users desired an added-value of functionality to that goes beyond the capabilities of their own arms and hands, e.g., dexterous manipulation of objects that are either too small or too delicate for human hands, higher load capacity, or a further reach. A longer reach during arm extension was desirable for a user who had to regularly reach up well above their head to press buttons on cameras in their lab.

Control and Autonomy

Many participants wanted voice-control of the robotic arm with at least some autonomy: *"I want to be able to tell it to do something, and it should just do it."* Others suggested implicit control based on movements or just intention: *"There should be some sensing so that I don't have to adapt to it, it can adapt to my position."* Some participants wanted a combination of intention-recognition and voice-override in case of errors. There was a sentiment that the arm could control itself better than the user could directly: *"I am not very good with remote controlled cars and helicopters, and would much rather have the arm control itself than have me crash it while using a joystick."* Another reason given for autonomy was the difficulty of multitasking: *"We can't really concentrate on that many things at*

once."

There were some in favor of direct control, *"like a joystick or [...] a motorcycle throttle"*, or *"small buttons near the fingertips"*, along with some kind of proximity sensor for gripping, allowing the arm to detect if it has successfully gripped the cup and try again if it has not.

Feedback

In this study, the robot was open-loop controlled, meaning that it went through a preset trajectory. Most participants commented that the arm's intentions were not clear throughout the trajectory: *"I had no sense of warning [when it was going] to pick up or drop objects."* They suggested ways for the robot to show its intention with a variety of feedback options. One participant suggested *"a light that turns on to indicate closing of the gripper,"* another *"a sequence of beeping sounds,"* and some suggested speech acknowledgments, such as *"I'm moving forward"* and *"I'm about to grip."*

Other users said that if they were to use the arm several times over the course of a few hours, they may be able to better predict its intentions, and *"get used to the distance and angle range"*, but adding indicators would be helpful for a new user.

Appearance

When speaking about the appearance of a wearable third arm device, users' imaginations were often informed by fictional characters. In a few cases, par-

ticipants suggested modeling the device on existing prosthetic devices. Most prominently, however, participants found the idea of another human-like arm attached to their bodies to be *"a bit creepy,"* especially *"if it were to look like human skin or flesh."* One participant said that this would be *"[...] scary. It would feel like the arm of another person, but no one's there."* Another user noted that the arm is not exactly discreet due to its bulky size, so it might stand out less if it had a humanoid appearance.

Finally, some users tied their notion of appearance with the capabilities of the third arm. One participant, who had worked at an oil rig, preferred a robust, futuristic look in that scenario. This appearance would have assured the user that the device was *"resistant to grease, water and dust"*.

2.5 Design Guidelines

The following design guidelines emerged from the online survey, contextual inquiry, and qualitative laboratory study described above:

- The robotic arm should be designed to work in **professional**, military, and law-enforcement contexts, mainly supporting **functional** rather than social uses.
- It should enable **reaching for, storing, and handing over** an object out of the reach for a worker, relieving them from the repetitive strain of these actions.
- The arm should also be able to **handle hot, toxic, or otherwise dangerous objects**. For this purpose, the materials for the body and end-effector of

the device should be resilient to electric currents, heat, and chemicals.

- The arm should **stabilize and brace** a user while they are working in precarious poses in terms of balance and ergonomics. For this purpose, the actuators and structural elements of the robot should be able to withstand forces and moments at scales produced by a human body.
- **Weight and balance** are key considerations for any wearable device. This, however, poses a trade-off with workspace and payload maximization. If material and actuation constraints require a heavier device, cable-driven systems for larger devices can help distribute the weight along the user's body by placing the heaviest components closer to the wearer's torso and thus reducing moments about the user's joints.
- Users desire high **dexterity** from the robot, especially at the end-effector, to be able to perform everyday tasks efficiently. This requirement for dexterous manipulation is at odds with the desired reduction in weight mentioned above, and the trade off is a substantial design challenge.
- Dexterity of physical design needs to be coupled with **feedback control** through different modalities: speech acknowledgments, visuals such as a screen or even a simple array of lights, or haptic feedback. The motion trajectories of the device itself should be designed to convey intent, even at the cost of task efficiency [38].
- Providing a wearable arm some **autonomy**, or at least a degreee of adjustable autonomy [44], is crucial for useful collaborative activities that reduce cognitive load on the user. This is especially important during **repetitive activities**. Given the close operation to the human body, autonomy must be limited to safe operations. Designers must therefore evaluate the kinds of autonomous tasks users will trust the device to perform safely.

For other tasks, they can include an option where the user can toggle to controlling it themselves, via a joystick or button.

- Finally, designers must take into account the prevailing social norms regarding robot and prosthesis **appearance**. Generally, we conclude from our interviews that a human-like appearance is undesirable, and might invoke a sense of the "uncanny valley" [77] and that a machine-like appearance would be more socially acceptable, as opposed to prostheses, where mimicking human appearance is more desirable.

2.6 Model II Prototype

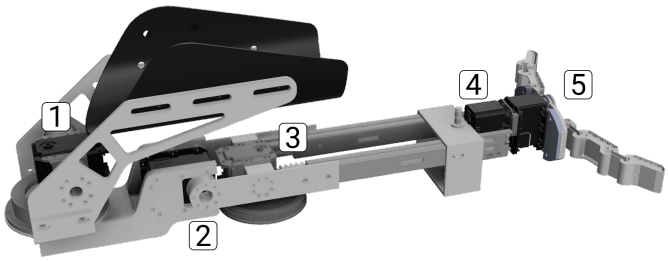


Figure 2.10: CAD rendering of the WRF Model II with four DOFs: 1) Horizontal Panning, 2) Vertical pitching, 3) Length extension, and 4) Wrist Rotation, along with a gripper (5).

To address the concerns in the design guidelines, two additional degrees of freedom (DoFs) were added in the redesigned WRF: vertical pitching of the arm, and wrist rotation before the gripper (Figure 2.10). Additionally, the robot's mounting point was shifted closer to the human elbow for improved ergonomic performance. These design improvements allowed the robot to perform tasks similar to those found in the contextual inquiry.

The vertical pitching, along with complete 360° panning, resulted in a full 3D

workspace, and allowed the robot to reach objects placed below as well as behind the user. The horizontal panning and vertical pitching DoFs were primarily responsible for bulk positioning of the WRF, while length extension further enhanced its reach.

In the design of the wrist and end-effector, the most important trade-off is between dexterity and ergonomics. Articulated spherical wrists, including serial and parallel mechanisms, have been extensively studied and deployed in commercial and research robots [93]. A fully articular 6-DoF parallel mechanism, similar to a Gough-Stewart platform [45, 105], was initially considered for the WRF's wrist. However, based on the indicative usage scenarios discovered in the studies, a wrist with a single rotational DoF similar to human wrist pronation and supination was found to suffice, along with rudimentary grasping capabilities with a two-fingered gripper. This is due to the fact that the WRF's dexterity is augmented by the user, as they can adjust its pose in real-time by moving the arm on which it is attached.

Adding more dexterity to the wrist, in the form of either serial or parallel mechanisms, would increase the load on a user's body due to the long moment arm for any mass added near the end-effector. Additionally, once we have narrowed down on a single actuated DoF for the wrist, a serial mechanism in the form of a connector directly mounted on the motor horn of the wrist actuator is preferred over a parallel mechanism such as a four-bar linkage [16]. This is due to the mechanical simplicity, larger workspace, and lack of singularities within the workspace of a serial mechanism [104], as well as the fact that the motor body can itself act as a structural element in the relatively low-load applications for the WRF.

Following the design guidelines, we describe the materials, actuators, and electronics architecture in the WRF Model II prototype's hardware implementation.

2.6.1 Materials

Material selection played a major role towards weight reduction in Model II compared to the initial prototype. The ABS mounting platform was replaced with a waterjet-machined sheet aluminum structure (Figure 2.11a). Aluminum sliders were used instead of stainless-steel ones in the length extension mechanism, serving as both actuation and structural elements.

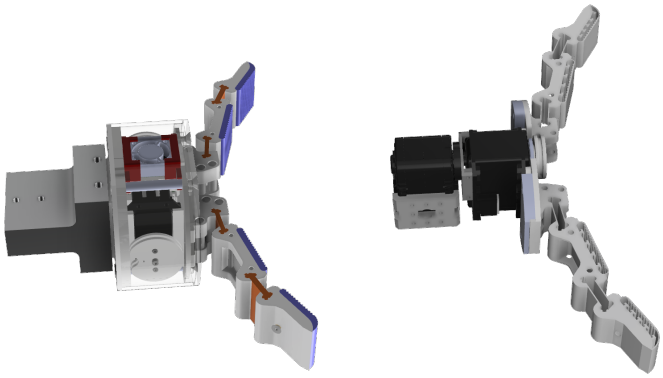
The initial gripper was designed after the Yale OpenHand Model T42 [80] but adapted to constrain both fingers to move together using a single motor for weight considerations. The gripper finger sizes in the WRF were reduced, and the motor housing and adaptor were removed, resulting in the motor body itself acting as a structural element connecting it to the previous DoF (Figure 2.11b).

The robot was mounted on a medical brace with metal hinges and molded plastic upper-arm and forearm supports, attached to the human arm using velcro straps. Most of the other structural elements were 3D-printed in ABS.

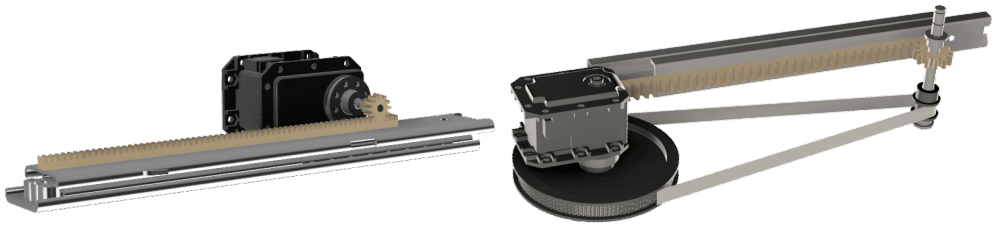
2.6.2 Actuation

The WRF Model II was actuated with ROBOTIS Dynamixel servo motors [6]. The horizontal panning and vertical pitching DoFs used MX-64T motors weigh-

(a) Mounting platform



(b) Gripper



(c) Length extension mechanism

Figure 2.11: Design changes in Model II over the initial prototype: (a) ABS mounting platform (left) replaced with sheet-aluminum (right); (b) Gripper size reduced and adaptor removed; (c) Direct-driven length extension mechanism (left) replaced with a belt drive (right).

ing 135 g each, with built-in proportional-integral-derivative (PID) feedback control for position and velocity, stall torque of 6.0 Nm at 12 V, and maximum speed of 63 rpm. These two DoFs required the most powerful motors since they were subject to the bulk of lifting and carrying loads. The length extension and gripper used smaller MX-28T motors, weighing 77 g with a stall torque of 2.5

Nm at 12 V, also with PID position and velocity control.

The wrist rotation motor was subject to the least loads during operation, being at the end of the robot's serial kinematic chain and not needing to generate contact forces for gripping. As a result, a lower-end AX-12A motor was used to actuate this DoF, with 1.5 Nm stall torque at 12 V, weighing 54.6 g, and with only proportional feedback controllers for position and velocity.

The rack-and-pinion length extension mechanism in the initial prototype was direct-driven, with the pinion gear mounted directly on the motor horn (Figure 2.11c, left). This design was updated to a belt-driven mechanism with a 7:1 transmission ratio and separated pinion gear, resulting in a faster extension speed and lower chance of slippage (Figure 2.11c, right).

Combined with aluminum sliders instead of steel, these design choices resulted in improved ergonomics and weight distribution, along with an overall reduction in weight by ˜0.5 kg.

2.6.3 Electronics

The motors in the WRF communicate at 1 mbps over a TTL protocol, attached serially in a daisy-chain fashion (Figure 2.12). The arm was tethered, receiving control commands from a PC, connected using a Xevelabs USB2AX v3.2a USB to TTL Dynamixel Servo Interface. It was powered by a 12 V, 5 A DC supply through an SMPS2Dynamixel Adapter. The MX-64T and MX-28T motors had onboard Cortex M3 CPUs, while the AX-12A had an Atmega8-16AU CPU.

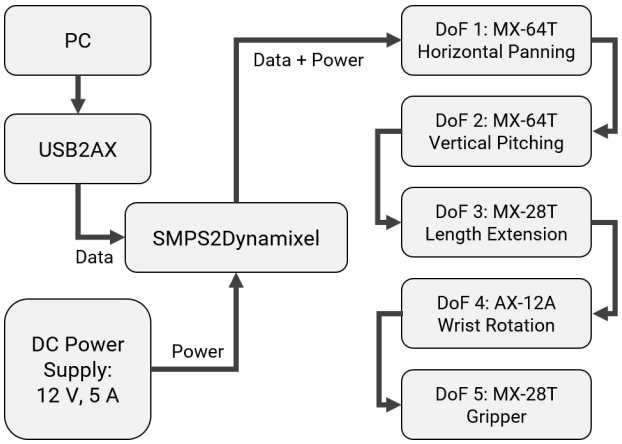


Figure 2.12: System architecture of WRF Model II electronics.

2.7 Conclusion

This chapter described the process of going from a rudimentary initial prototype (Model I) of the WRF to a lighter, more articular, and structurally stronger version (Model II) by following a user-centered approach, identifying need themes and design guidelines based on feedback from study participants.[1]

While the user-centered design process revealed multiple dimensions for improving successive prototypes, at this point in the project, we were still in the early stages of physical design. As a result, we prioritized the weight, balance, and dexterity for Model II as opposed to the feedback control, or autonomy of the robot. Also, modifying the appearance was not considered to be a priority at this stage.

In the next chapter, we analyze Model II in terms of its kinematics, enhancement in reachable workspace volume it affords, and the loads it exerts on the user's arm joints, and describe the design and kinematics of the final prototype, Model III.

[1]Portions of this chapter have been published in [113].

CHAPTER 3

DESIGN ANALYSIS: KINEMATICS, BIOMECHANICS, AND WORKSPACE VOLUME

Having realized an improved Model II WRF prototype based on an iterative user-centered process, in this chapter we present an analysis of its design.

Starting with the kinematic structure of the human arm with Model I attached, having a planar reachable workspace envelope, we showed that the structure of Model II allows for reaching in full 3D. We then analyzed the enhancement in reachable workspace volume when the combined human arm and WRF system simultaneously undergoes its full range of motion (RoM). Using a rigid body model of the human arm, we determined that while Model II includes two additional degrees of freedom (DoFs), the forces and moments it exerts on the user's elbow and shoulder joints while performing close-range assistive tasks are well within the acceptable limits for ergonomic use, and within motor torque limits.

Finally, we report on a pilot interaction study with the WRF Model II which informed the final prototype (Model III), and led to the development of a more detailed interaction scenario, described in Chapter 6.

3.1 Kinematic Structure

The WRF is a serial kinematic chain attached to the human forearm (Figure 3.1). Generally, the human arm can be represented as a 7-DoF chain [88]. However, since the robot's motion is unaffected by human wrist movements, we used a

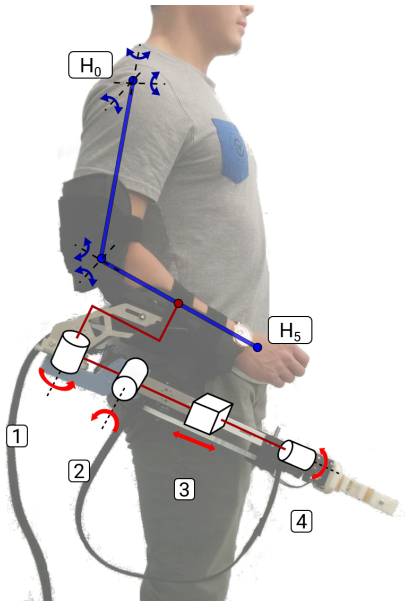


Figure 3.1: Kinematic architecture of the WRF attached to the human arm.

reduced 5-DoF model, with three joints at the shoulder and two at the elbow.

Table 3.1: D-H parameters for the 5-DoF human arm model

Degree of Freedom	α_i	a_i(m)	d_i(m)	θ_i
1) Shoulder circumduction	-90°	0	0	[0°, 180°]
2) Shoulder adduction	+90°	0	0	[-90°, 140°]
3) Shoulder flexion	0°	0.335	0	[-90°, 170°]
4) Elbow flexion	+90°	0	0	[80°, 235°]
5) Elbow pronation	+90°	0	0.263	[0°, 180°]

The forward kinematics of each of these serial chains was described with coordinate frames derived using the Denavit-Hartenberg (D-H) convention [47], resulting in a homogeneous transformation matrix T_0^n between the frame H_0 at the origin (human shoulder joint) and the frame H_5 at the human's hand:

$$T_0^n = \prod_{i=1}^{n} T_{i-1}^{i}(\alpha_i, a_i, d_i, \theta_i) \tag{3.1}$$

Here $n = 5$ is the number of joints, and $(\alpha_i, a_i, d_i, \theta_i)$ are the D-H parameters for human arms. In Table 3.1, the anthropometric parameters and ranges of motion

have been adapted from the model described in [88] and the NASA Man-System Integration Standards [4].

3.1.1 Model I

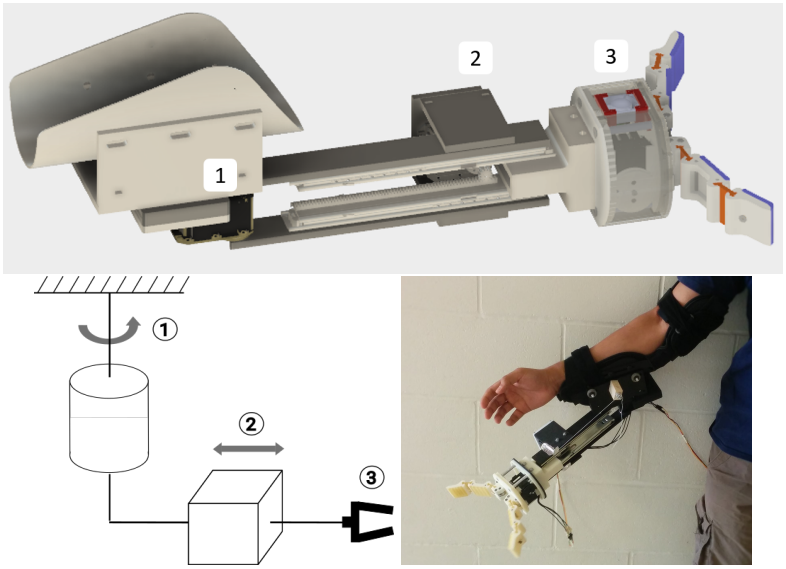


Figure 3.2: CAD model, kinematic diagram and physical realization of Model I with three DoFs: 1) Horizontal panning, 2) Length extension, 3) Gripping.

The initial prototype, Model I, had a DoF at the elbow for horizontal panning, a prismatic joint for length extension, and a two-fingered gripper as the end effector, shown in Figure 3.2. Its D-H parameters are listed in Table 3.2. This first prototype was designed with workbench operation in mind, broadening the wearer's effective "wing span" through planar movement with respect to a fixed human pose.

Table 3.2: D-H parameters for WRF Model I

Degree of Freedom	α_i	a_i(m)	d_i(m)	θ_i
1) Horizontal panning	-90°	0.254	0	[-60°, 60°]
2) Length extension	0	0	[0.07, 0.23]	0

3.1.2 Model II

As described in the Chapter 2, the main design shortcomings of Model I were its weight, limited dexterity with respect to grasping angle, and the speed of the prismatic joint. For example, in handover tasks, particularly involving two people, it was desirable for the workpiece to be oriented by the WRF in a manner that facilitated an easy grasping position for the receiver.

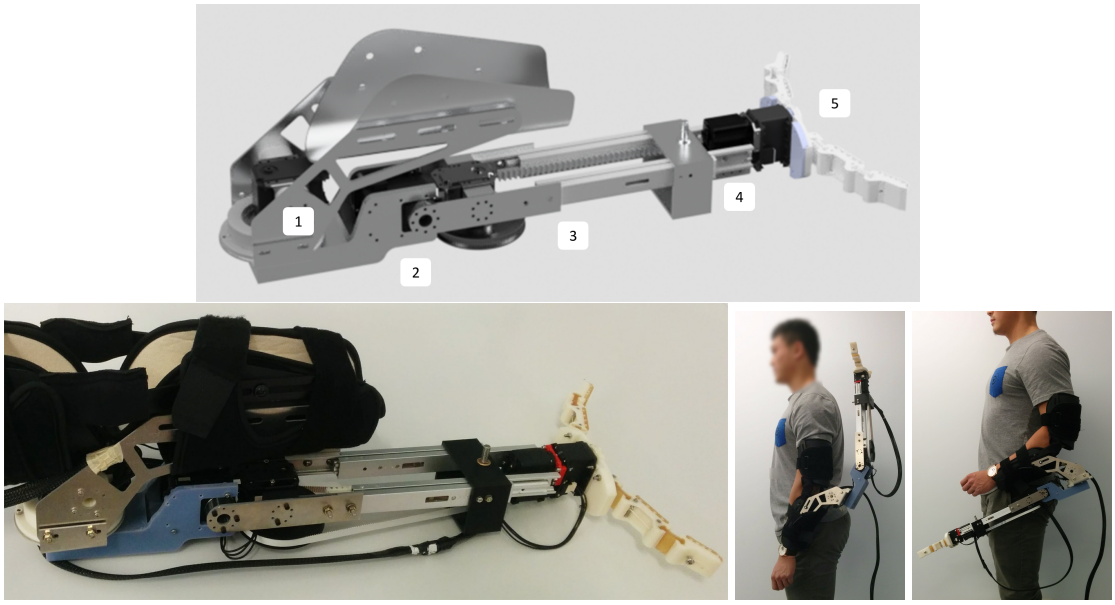
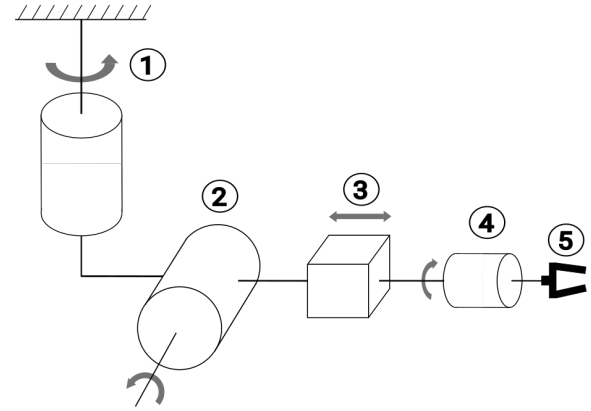


Figure 3.3: CAD model and physical implementation of Model II, allowing for reach below and behind the user.

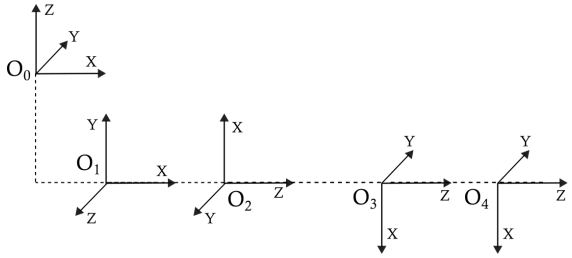
Table 3.3: D-H parameters for WRF Model II

Degree of Freedom	α_i	a_i(m)	d_i(m)	θ_i
1) Horizontal panning	+90°	-0.112	0	[-180°, 180°]
2) Vertical pitching	+90°	0	0	[-180°, 30°]
3) Length extension	0°	0	[0.28, 0.44]	180°
4) Wrist rotation	0°	0	0.106	[-180°, 180°]

To address these concerns, we added two additional DoFs in Model II: a vertical pitching of the arm and a wrist rotation before the gripper (Figure 3.3).



(a) Model II kinematic diagram



(b) D-H coordinate frames.

Figure 3.4: (a) Kinematic diagram of the WRF with the human arm attachment point as the base link, (b) Coordinate frames based on the D-H convention.

Figure 3.4 shows its kinematic structure and associated frames for its D-H parameters listed in Table 3.3.

The vertical pitching, along with complete 360° panning, resulted in a full 3D workspace (Figure 3.5), and allowed the user to reach objects placed below as well as behind the user. In contrast, the workspace in Model I was constrained to a planar region with respect to the wearer's arm.

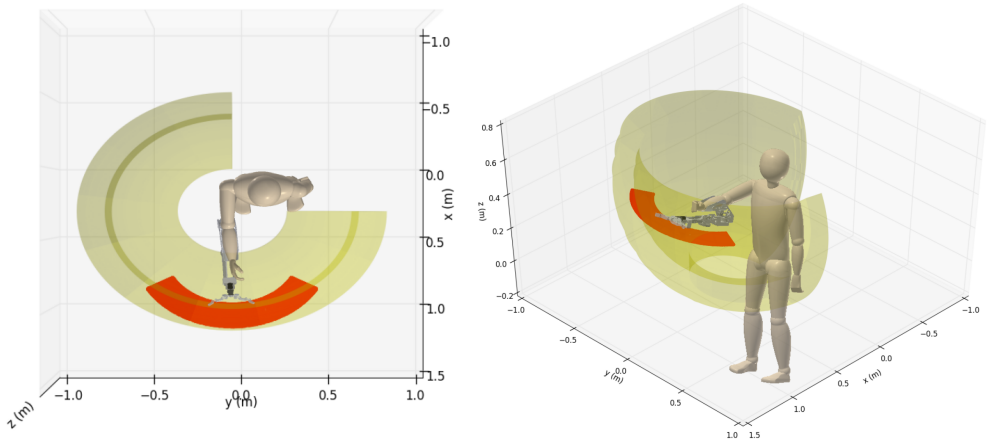


Figure 3.5: Model II offers a larger, full 3D workspace (yellow) compared to Model I (red) for the same fixed body configuration of the human.

3.2 Workspace Analysis

When the DoFs of the human arm and WRF are allowed to undergo their full range of motion (RoM), the resulting volume spanned by the set of end-effector positions in 3D forms the total reachable workspace. We computed these volumes for the human hand as end-effector, as well as for the WRF's end-effector while wearing Model I and Model II, and found that the robot enhanced the user's reach compared to the normal human range.

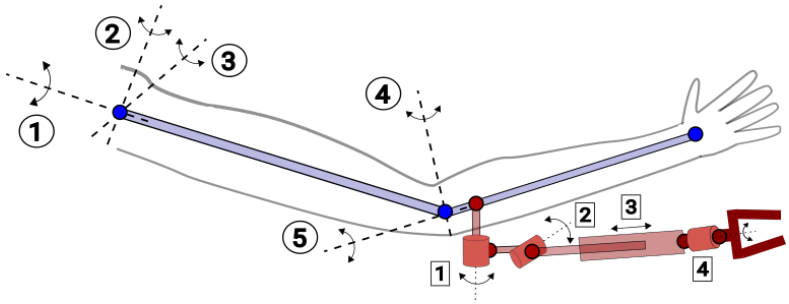


Figure 3.6: Schematic used in the workspace analysis of a human arm with the robotic arm attached at the elbow, illustrating the DoFs in Tables 3.1 and 3.3.

$$S_0^{n+m} = \prod_{i=1}^{n} T_{i-1}^{i} \prod_{j=1}^{m} P_{j-1}^{j} \tag{3.2}$$

To quantify this workspace enhancement, we employed a combined human-robot model (Figure 3.6). Similar to (3.1), we constructed transformation matrices P_0^m using the D-H parameters for the WRF (Tables 3.2 and 3.3), and concatenate them with T_0^n to get the transformation matrix S_0^{n+m} for the combined human-robot model. To account for the attachment point offset between the human and robot, parameters for the fifth DoF in T_0^n in Table 3.1 needed to be modified with $a_5 = 0.093$ m, $d_5 = 0.085$ m for Model I, and $a_5 = 0.075$ m, $d_5 = 0.016$ m for Model II.

3.2.1 Workspace Computation Results

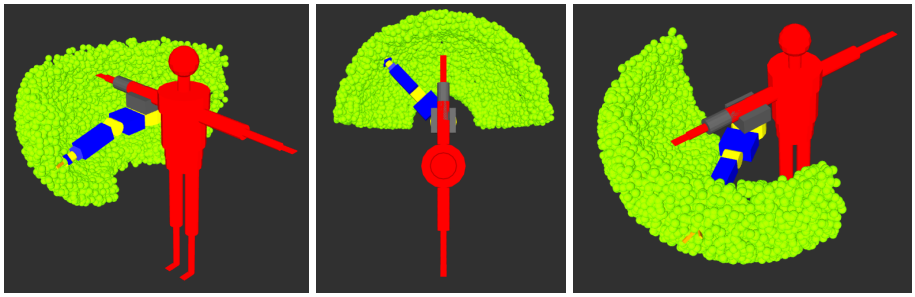


Figure 3.7: Sampling of end-effector positions for workspace volume computation.

The total reachable workspace volume is the union of workspaces generated by the end-effector when a mechanism undergoes its full RoM. The three mechanisms in this case were the combined human-robot systems for Model I and Model II, and the human arm alone. This set of end-effector positions was estimated using a Monte-Carlo sampling procedure as proposed in [28], illustrated

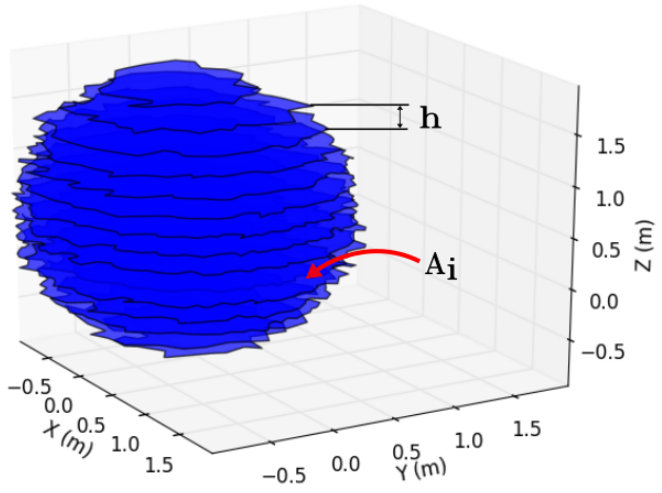


Figure 3.8: Generation of 2D slices from the workspace volume point cloud for Model II.

in Figure 3.7. Each joint variable, θ_i (or d_i for the prismatic joint) was drawn from a Beta random distribution, $\theta_i \sim Beta(\alpha, \beta)$, where the distribution parameters α and β were determined based on the RoM for each DoF. This reduced the sparsity of points at the ends of the joint space range. The resulting point clouds were sectioned into 2D slices along the Z-axis (Figure 3.8), and numerically integrated using a trapezoidal method.

$$V = h\,[\sum_{i=1}^{s} A_i - \frac{1}{2}(A_1 - A_s)] \tag{3.3}$$

Here $h = [z_{max} - z_{min}]/s$ is the step size with s steps, and A_i is the area of the i^{th} slice used to compute the Volume V. The simulations were performed using a Unified Robot Description Format (URDF) [7] model of the WRF and an articular 34-DoF model of the human body, adapted from [8]. The 3D convex hulls of the workspace for the human arm and combined human-robot systems were constructed by tessellating the outer boundary curves of the 2D slices.

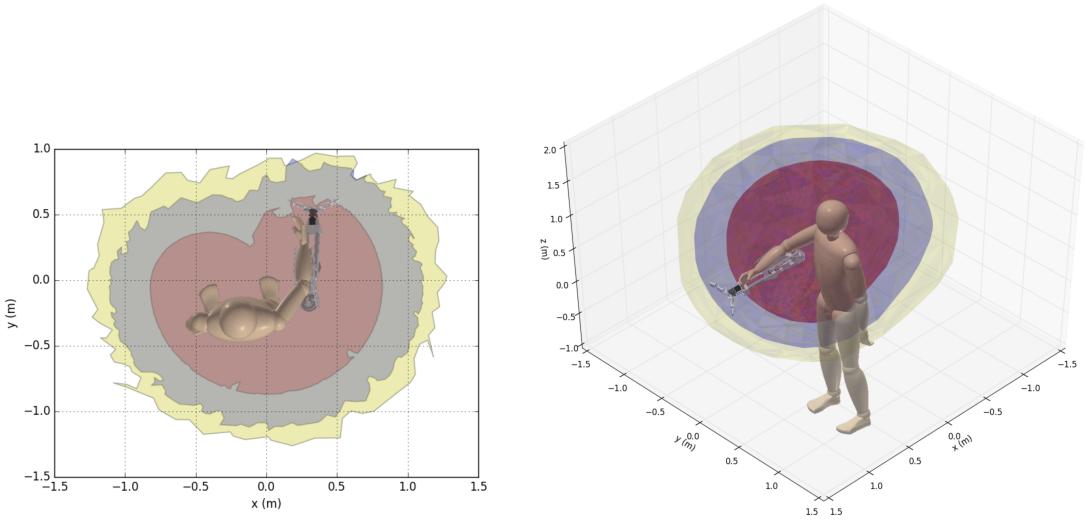


Figure 3.9: Convex hulls of end-effector point clouds showing total reachable workspace volume afforded by Model II (yellow), compared to Model I (blue), and the natural human arm range (red).

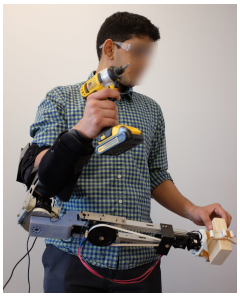
The total reachable workspace volume for the human arm alone was found to be 1.003 m^3. This was enhanced to 2.389 m^3 while wearing Model I, an improvement of 138%. Wearing Model II further increased the total reachable workspace volume to 3.467 m^3, an improvement of 246%, as illustrated in Figure 3.9.

3.3 Evaluation: Operating Loads, Interaction Modes

While the WRF enhances a user's reach, for a person to wear the robot for prolonged periods of time, we need to consider the biomechanical loads experienced by them. In this section, we evaluate these loads for the Model II prototype and compare with Model I. There is a direct trade-off between a low biomechanical load, and adequate power density in the actuators for rapid collaborative action. We ensured that the robot's motors were capable of performing rel-

evant tasks by measuring the torque loads in these tasks. We also conducted a pilot in-person study to compare interaction with the robot under direct speech control, and remote control presented as an "autonomous" mode to the participants.

(a) Fetching an object with user's hands occupied.



(b) Self-handover.

(c) Assisted two-person handover.

(d) Stabilizing and object for bi-manual manipulation.

Figure 3.10: Illustrative usage scenarios for the WRF.

To analyze the biomechanial and actuator loads, we considered three indicative collaboration scenarios based on the design goals and user study findings from Chapter 2: fetching an object from below the human's workspace for self-handover; human-to-human handover assisted by the robotic arm; and fixing an object in position while the human is operating on it. The WRF's trajectory was predefined and executed in an open-loop manner.

a) Fetching from Below This task involves a combination of horizontal panning, pitching down, and length extension to reach the object to be fetched. Once the object is grasped, it is brought over to the user's own hands (Figure 3.10a). Therefore it is equivalent to a traditional human-robot handover, with the added kinematic constraint between the robot and the arm it is worn on.

b) Assisted Human-Human Handover In this scenario, two humans are working back to back. The wearer has an object within their workspace that they want to transfer to another person while their own hands or attention is occupied. The robot grasps the object and hands it over to a person standing behind the user by panning outwards (Figure 3.10c). Thus the WRF acts as a functional tool to facilitate human-human collaboration. This two-person handover performed using a third arm robot is particularly interesting, because it adds another degree of uncertainty to a traditional human-robot handover, in the way that the robot is not an independent agent. It is guided by the pose of the wearer, whose actions are non-deterministic and based on their belief of the intentions of the receiver.

c) Fixing an Object in Position This scenario does not involve moving an object, but focuses on the robot stabilizing both itself and an object, for example holding a block of wood in place while the user drills into it (Figure 3.10d). This type of assistance in immobilizing an object is common in manufacturing or assembly settings [85], or in situations where the human body is not capable of providing the necessary force or moment.

3.3.1 Biomechanical Load Analysis

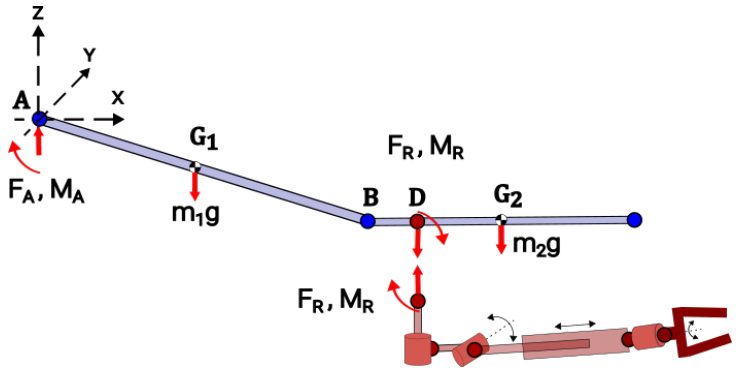


Figure 3.11: Free-body diagram of the human arm with the WRF as a point load and moment.

To evaluate the biomechanical load on the user in these scenarios, we modeled the human-robot system as two distinct bodies, as shown in Figure 3.11. The interaction between the robot and the human arm took the form of the support force $\vec{F}_R$ and moment $\vec{M}_R$.

For each usage scenario trajectory, $\vec{F}_R$ and $\vec{M}_R$ were computed using the iterative Newton-Euler dynamics algorithm [72] applied to the robot. As shown in Figure 3.6, the robot was considered to be a 4-DoF model with 5 links, based on the structure of Model II. For each task trajectory, the motors were assumed to rotate at constant angular speeds. The ramping up and down of angular velocities during the start and end of a trajectory, as well as sign changes during a trajectory, were assumed to be nearly instantaneous, taking between two and five time steps. The resulting accelerations of the links, and forces and moments at the joints were computed iteratively, going from link 1 to link 5, with a different reference frame attached to each link. The outward iteration equations below provide the linear acceleration $\dot{v}$ and angular acceleration $\dot{\omega}$ of link $i+1$, given these quantities for link i. The matrix $R_i^{i+1} \in SO(3)$ is the rigid-body trans-

formation between frames attached to link $i+1$ and link i. θ_i is the joint angle for each motor, d_i is the length of the prismatic joint, and $\hat{z}_i^i$ is the joint axis for joint i in frame i. P_i^{i+1} is the position vector going from the origin of frame i to the origin of frame $i+1$. The initial conditions for the outward iterations were specified at the ground link (link zero) of the serial chain as in [31]: $\omega_0^0 = 0$, $v_0^0 = 0$, $\dot{\omega}_0^0 = 0$ and $\dot{v}_0^0 = \vec{g}$, where $\vec{g}$ is acceleration due to gravity.

$$\omega_{i+1}^{i+1} = R_i^{i+1}\omega_i^i + \dot{\theta}_{i+1}\hat{z}_{i+1}^{i+1} \tag{3.4}$$

For rotary joints (all except DoF 3 in Model II):

$$\dot{\omega}_{i+1}^{i+1} = R_i^{i+1}\dot{\omega}_i^i + R_i^{i+1}\omega_i^i \times \dot{\theta}_{i+1}\hat{z}_{i+1}^{i+1} + \ddot{\theta}_{i+1}\hat{z}_{i+1}^{i+1} \tag{3.5}$$

$$\dot{v}_{i+1}^{i+1} = R_i^{i+1}[\dot{\omega}_i^i \times P_i^{i+1} + \omega_i^i \times (\omega_i^i \times P_i^{i+1}) + \dot{v}_i^i] \tag{3.6}$$

For DoF 3, which is prismatic:

$$\dot{\omega}_{i+1}^{i+1} = R_i^{i+1}\dot{\omega}_i^i \tag{3.7}$$

$$\begin{aligned}\dot{v}_{i+1}^{i+1} = R_i^{i+1}[\dot{\omega}_i^i \times P_i^{i+1} + \omega_i^i \times (\omega_i^i \times P_i^{i+1}) + \dot{v}_i^i] \\ +2\omega_{i+1}^{i+1} \times \dot{d}_{i+1}\hat{z}_{i+1}^{i+1} + \ddot{d}_{i+1}\hat{z}_{i+1}^{i+1}\end{aligned} \tag{3.8}$$

The linear accelerations $\dot{v}_{c,i}$ of the centers of mass (CoMs) of each link were computed as follows, along with the inertial forces F_i and moments M_i on each link. $P_{c,i}^i$ is the position of the CoM of link i in frame i, and $I_{c,i}$ is the moment of inertia about the CoM.

$$\dot{v}_{c,i} = \dot{\omega}_i^i \times P_{c,i}^i + \omega_i^i \times (\omega_i^i \times P_{c,i}^i) + \dot{v}_i^i] \tag{3.9}$$

$$F_i = m_i \dot{v}_{c,i} \tag{3.10}$$

$$M_i = I_{c,i}\dot{\omega}_i^i + \omega_i^i \times (I_{c,i}\omega_i^i) \tag{3.11}$$

The inward iterations, going from link 5 to link 1 in Model II, used the above quantities to compute the forces f_i^i and moments n_i^i exerted on link i by link $i+1$ as seen in frame i.

$$f_i^i = R_{i+1}^i f_{i+1}^{i+1} + F_i^i \tag{3.12}$$

$$n_i^i = M_i^i + R_{i+1}^i n_{i+1}^{i+1} + P_{c,i}^i \times F_i^i + P_i^{i+1} \times (R_{i+1}^i f_{i+1}^{i+1}) \tag{3.13}$$

The initial condition for inward iterations is the external loading at the end-effector of the robot. For the usage scenarios, we considered the gripper to be holding a plastic cup weighing ~30g, so that $f_6^6 = -0.29\hat{k}$ N, $n_6^6 = \vec{0}$. The required interaction loads between the robot and human arm are given by $\vec{F}_R = f_0^0$ and $\vec{M}_R = n_0^0$.

The human arm was assumed to remain stationary during the trajectories in scenarios (a) and (b) of the robot, as in [33]. This allowed us to use a simplified static arm model to estimate the biomechanical load, consisting of the force norms at the human shoulder and elbow: $\|\vec{F}_A\|$, $\|\vec{F}_B\|$, and corresponding moment norms: $\|\vec{M}_A\|$, $\|\vec{M}_B\|$.

These quantities can be computed for each scenario assuming static equilib-

rium. Using the notation in Figure 3.11,

$$\vec{F}_A = \vec{F}_R + m_1\vec{g} + m_2\vec{g} \tag{3.14}$$

$$\vec{M}_A = \vec{M}_R + \vec{r}_{G_1/A} \times m_1\vec{g} + \vec{r}_{G_2/A} \times m_2\vec{g} + \vec{r}_{D/A} \times \vec{F}_R \tag{3.15}$$

$$\vec{F}_B = \vec{F}_R + m_2\vec{g} \tag{3.16}$$

$$\vec{M}_B = \vec{M}_R + \vec{r}_{G_2/B} \times m_2\vec{g} + \vec{r}_{D/B} \times \vec{F}_R \tag{3.17}$$

Comparison Between Models

We evaluated the improvement in Model II compared with Model I in a scenario that can be performed by both devices, using the load analysis method discussed above. The scenario was to grasp an object that is just out of human reach while seated at a desk (Figure 3.12 left), with the following steps: *Pan outwards to +60°→ Fully extend arm → Grip → Pan inwards to –60° → Open gripper*. We find that the magnitudes of moments experienced by the wearer at their elbow and shoulder were reduced by 18 to 35% for Model II compared to Model I (Figure 3.12).

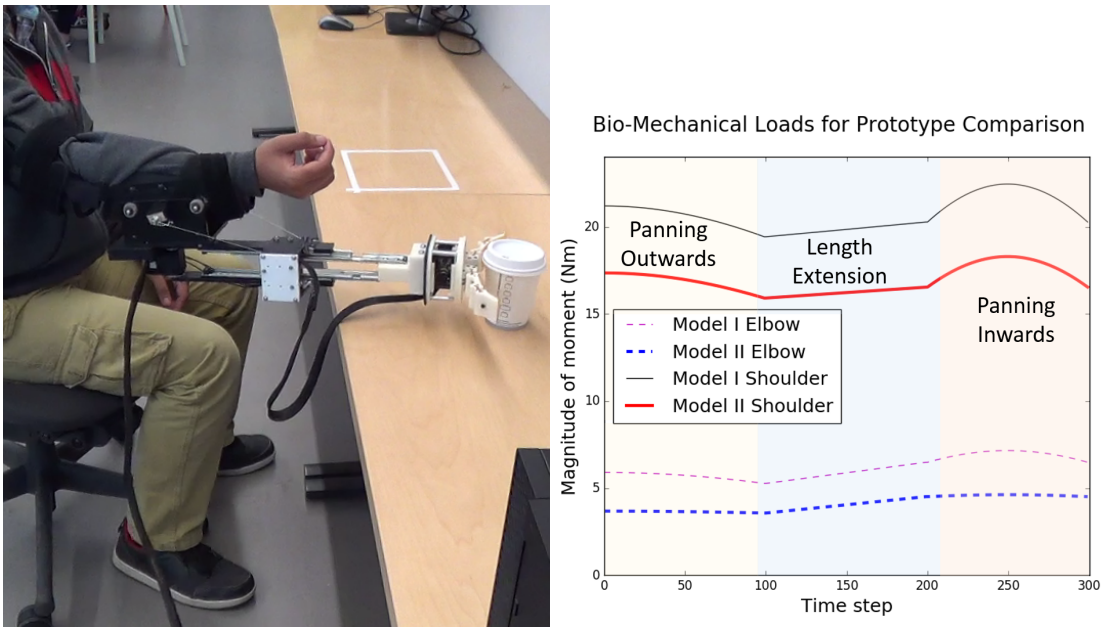


Figure 3.12: In moving an object across a table, Model II exerts a moment load at the user's shoulder and elbow that is 18–35% lower than Model I.

Model II Loads

For the scenarios shown in Figure 3.10, the magnitudes of biomechanical forces at the human shoulder and elbow, $\|\vec{F}_A\|$, and $\|\vec{F}_B\|$ remained almost constant, as the cetrifugal and Coriolis effects were negligible compared to the weights of the links. The peak force loads were ˜55.8 N and ˜31.3 N, at the shoulder and elbow respectively. For comparison, the peak force loads that a human can withstand are ˜100 to 500 N at the shoulder, and ˜50 to 400 N at the elbow, depending on the arm configuration [63].

Figure 3.13 shows the biomechanical moment loads $\|\vec{M}_A\|$ and $\|\vec{M}_B\|$ during scenarios (a) Fetching an object from below, and (b) Assisted human-human handover. In both tasks, the object being manipulated was a plastic cup weighing ˜30 g. The peak moment loads on the wearer's shoulder and elbow during these tasks were ˜24.8 Nm and ˜11.6 Nm respectively. For comparison, the hu-

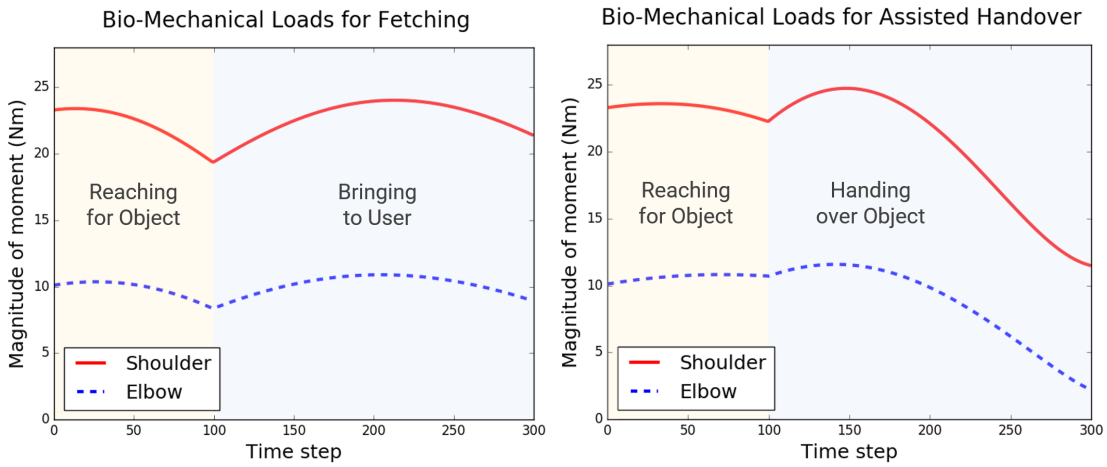


Figure 3.13: The biomechanical moment loads during the fetching and handover tasks were well within human ergonomic limits for the elbow and shoulder.

man shoulder can withstand moments of magnitude ˜85 to 130 Nm, while the elbow can withstand ˜40 to 80 Nm [63, 83]. Biomechanical load evaluations were omitted for scenario (c) Fixing and object in position, due to its statically indeterminate nature.

3.3.2 Motor Torque Loads

We measured motor torque loads directly from the built-in sensors on the Dynamixel servo motors used in the WRF. In all three scenarios, non-negligible torque loads were present in DoFs 1 and 2. In the bracing operation, the wrist and gripper were isolated from the vibration loads of the power screwdriver due to the compliant finger structure.

Figure 3.14 shows motor torque loads for all three scenarios. These loads remained below 50% of the motor's peak rating of 6.0 Nm for the first two scenarios, and below 20% for bracing.

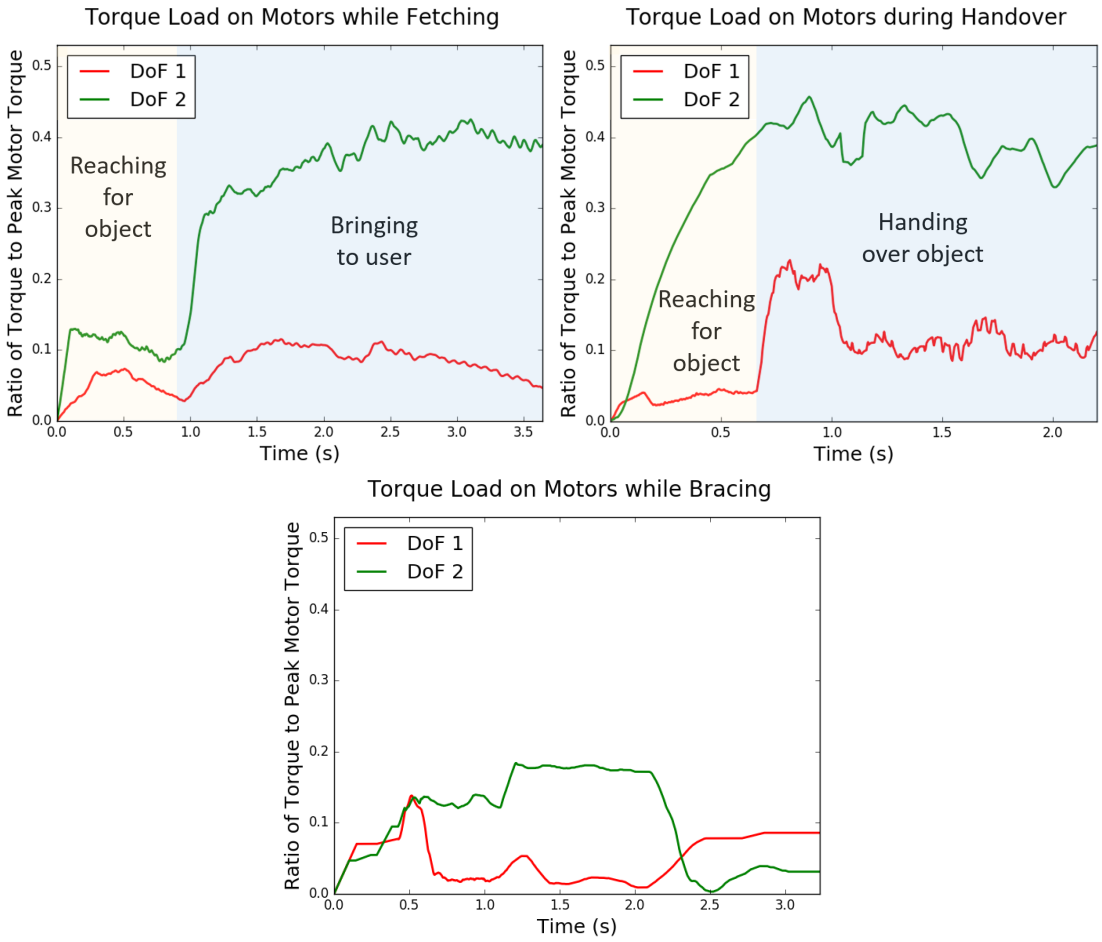


Figure 3.14: The torque loads on the servo motors used for the horizontal panning and vertical pitching DoFs never exceeded 50% of their peak rating of 6.0 Nm

3.3.3 Pilot Interaction Study

We found our design to be ergonomic enough to stay within human biomechanical force and moment limits throughout the above-mentioned scenarios. In addition, the low motor torque loads we measured suggest the possibility of using even more lightweight motors in subsequent design iterations, further reducing loads.

Apart from the physical design aspect, an important area of exploration is

the level of autonomy, and desired mode of interaction between the user and the robot. In the design studies in Chapter 2, participants had expressed the following insights regarding control and autonomy:

- There should be some mode of communication between the human and robot, in the form of visual, haptic, or verbal cues.
- At the same time, the kind of autonomy provided to the robot should reduce cognitive load on the user. The user should not be confused by the robot's actions.
- Participants expressed interest in being able to give high-level commands verbally, and the robot being able to handle the implementation of those commands autonomously.

These present a well-known interaction trade-off or gradient between full autonomy of the robot, and the user giving commands to the robot to complete a given task [44]. To explore this, we conducted a pilot study for a pick-and-place task with the WRF Model II.

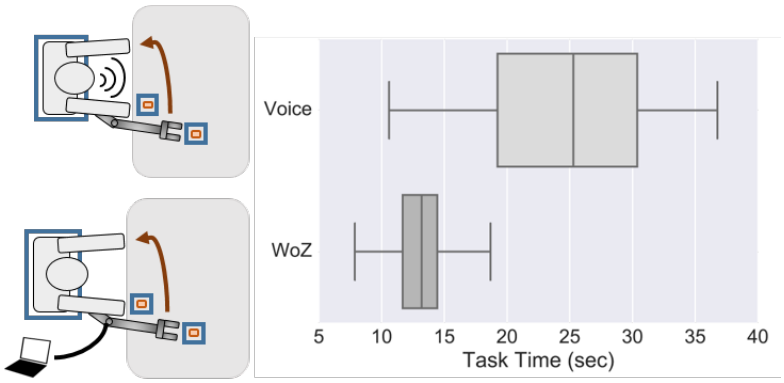


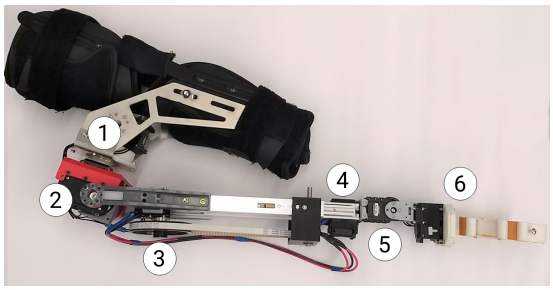
Figure 3.15: Setup of the pilot study: the robot brings the orange objects on the table to the user's other arm (left). "Autonomous" WoZ behavior of the robotic arm resulted in shorter task times than direct voice control (right), as well as being preferred by users.

In the study, participants wore the robot and used both of their hands, and the WRF for a self-handover task involving two objects, one of which was out of reach (Figure 3.15). They performed the task under two conditions. In one, participants used voice commands (e.g. "left", "right", "grasp") to directly steer the robot. The second interface was presented as an autonomous arm, but was in fact a Wizard-of-Oz ("WoZ") setup [60] where a remote operator directly controlled the robot's motion, unbeknownst to the user. Eight participants wore the WRF and completed the same task three times with each interface, counterbalanced for order effects.

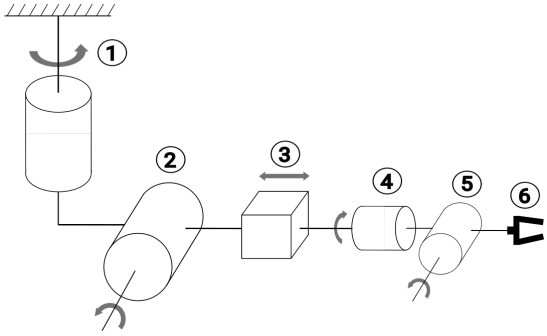
The mean completion time was faster and less variable using WoZ (mean M_w = 12.97s, std. dev. S_w = 3.31) compared to the voice interface (mean M_v = 24.60s, std. dev. S_v = 8.79; $M_v - M_w$ = 11.64s; $t(7) = 4.39, p < 0.01, d = 1.55$, Figure 3.15). This suggests that even though participants had no control over the arm's movement, they were able to complete the task faster. In questionnaires, users also rated the desirability of the "autonomous" movement higher than the direct voice control. While these were only preliminary results and required more rigorous testing, they indicated that providing a wearable robot with autonomy could reap task efficiency and usability benefits.

3.4 Model III Prototype

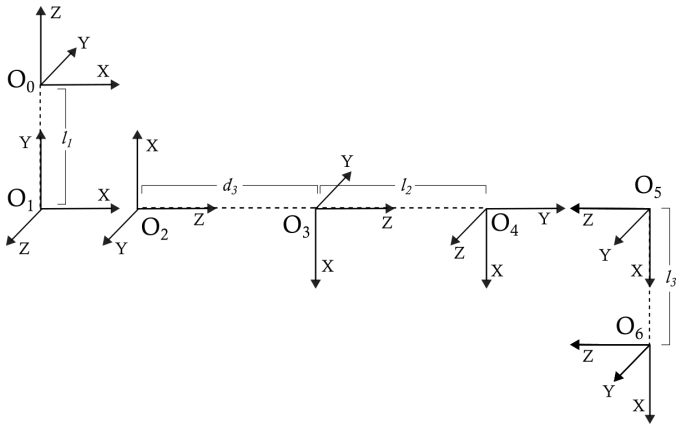
While the pilot interaction study shed some light onto the performance of the WRF in the two control modes, users also mentioned that a more dexterous wrist would be desirable, especially in pick-and-place tasks. For instance, while reaching for an object placed on a tabletop, rotating the gripper to align the fin-



(a) WRF design with six actuators



(b) Kinematic structure



(c) Coordinate frames in D-H parameterization

Figure 3.16: The final design of the WRF has five degrees-of-freedom (DoFs) including prismatic length extension, along with a two-fingered gripper.

gers parallel to the surface would lead to a more stable grasp. A more articulate wrist would also be advantageous in self handovers and assisted human-human handovers, for orienting the grasped object to a pose more amenable to the receiver.

This led to the development of the final WRF prototype: Model III, which has an added degree of freedom at the wrist (vertical pitching) before the gripper (Figure 3.16a). It has a total of five articular degrees-of-freedom (DoFs) including prismatic length extension along with a gripper,and has a maximum reach of 0.63 m from the base of the first DoF (Figure 3.16b). While this version

is heavier than Model II, weighing ˜2 kg due to the addition of wrist pitching, it has has carbon-fiber reinforced plastic components to replace the aluminum connectors between DoFs 2 and 3, as well as a lighter connection between DoF 1 and 2, reducing the length of the effective moment arm about the human's forearm due to the WRF's weight.

In a real setting, for the WRF to be able to perform tasks such as the ones shown in Figure 3.10, we need sensing, motion planning, and motor control. For the WRF, sensing was achieved through off-board cameras, as described in the following chapters. Motion planning generally involved finding linear paths in the WRF's configuration space, or through an optimizing planner described in Chapter 4. The final aspect, motor control, involves converting the Cartesian paths generated by a motion planner into reference signals given as inputs to a robot's actuators. Achieving this for the WRF requires an analysis of its forward and inverse kinematics.

3.4.1 Forward Kinematics

The kinematics for Model III are described using the D-H convention, with the parameters listed in Table 3.4 for the coordinate frames shown in Figure 3.16c.

Table 3.4: D-H parameters for WRF Model III

Degree of Freedom	α_i	a_i(m)	d_i(m)	θ_i
(1) Horizontal panning	$\pi/2$	0	-0.08	$(-\pi, \pi)$
(2) Vertical pitching	$\pi/2$	0	0	$(0, \pi/2)$
(3) Length extension	0	0	[0.33, 0.45]	π
(4) Wrist rotation	$\pi/2$	0	0.045	$(-\pi, \pi)$
(5) Wrist pitching	$\pi/2$	0	0	$(0, \pi)$
(6) End-effector	0	0.135	0	0

Forward Kinematics (FK) provides the mapping from a robot's joint variables to the pose of the end-effector. As shown in Figure 3.16c, the base frame O_0 of the WRF lies at the top of horizontal panning DoF, and the final frame O_6 is attached to the end-effector, lying at the mid point of the gripper's fingers. Using the D-H parameters (α, a, d and θ), we specify the relationship between frames through homogeneous transformation matrices T_i^{i+1}, defined for the pose of frame $i+1$ as seen in frame i:

$$T_i^{i+1} = \begin{bmatrix} cos\theta_i & -sin\theta_i cos\alpha_i & sin\theta_i sin\alpha_i & a_i cos\theta_i \\ sin\theta_i & cos\theta_i cos\alpha_i & -cos\theta_i sin\alpha_i & a_i sin\theta_i \\ 0 & sin\alpha_i & cos\alpha_i & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.18}$$

The joint variables ($\vec{\theta}$) for the WRF are θ_i for the four revolute joints, and d_3 for the prismatic joint. The end-effector is fixed with respect the wrist pitching DoF. However, it is listed separately in order to decouple it from wrist pitching for ease of solution of the analytical inverse kinematics, as in [119]. The transformation T_0^6, between the base frame O_0 and the end-effector frame O_6 can be found by through the following product with successive terms right-multiplied:

$$T_0^6 = \prod_{i=0}^{5} T_i^{i+1} \tag{3.19}$$

Expanded out in terms of individual elements:

$$T_0^6 = \begin{bmatrix} R_{1x} & R_{2x} & R_{3x} & P_x \\ R_{1y} & R_{2y} & R_{3y} & P_y \\ R_{1z} & R_{2z} & R_{3z} & P_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.20}$$

The rotation between these frames is described by $R = [\vec{R}_1, \vec{R}_2, \vec{R}_3] \in SO(3)$, where $SO(3)$ is the 3D rotation group consisting of 3×3 orthogonal matrices with unit positive determinant. $\vec{P} = [P_x, P_y, P_z]^T \in \mathbb{R}^3$ describes the translation between these frames. We denote $cos\theta_i$ as c_i and $sin\theta_i$ as s_i for ease of notation. The length parameters for the WRF are l_1, l_2 and l_3, which correspond to d_1, d_4 and a_6 respectively from Table 3.4. There are twelve equations that constitute the forward kinematics:

$$R_{1x} = c_1 s_2 s_5 - c_5(s_1 s_4 + c_1 c_2 c_4) \tag{3.21}$$

$$R_{1y} = c_5(c_1 s_4 - c_2 c_4 s_1) + s_1 s_2 s_5 \tag{3.22}$$

$$R_{1z} = -c_2 s_5 - c_4 c_5 s_2 \tag{3.23}$$

$$R_{2x} = -c_4 s_1 - c_1 c_2 s_4 \tag{3.24}$$

$$R_{2y} = -c_1 c_4 - c_2 s_1 s_4 \tag{3.25}$$

$$R_{2z} = -s_2 s_4 \tag{3.26}$$

$$R_{3x} = -c_1 s_2 c_5 - s_5(s_1 s_4 + c_1 c_2 c_4) \tag{3.27}$$

$$R_{3y} = s_5(c_1 s_4 - c_2 c_4 s_1) - s_1 s_2 c_5 \tag{3.28}$$

$$R_{3z} = c_2 c_5 - s_2 c_4 s_5 \tag{3.29}$$

$$P_x = l_2 c_1 s_2 - l_3(c_5 s_1 s_4 \ + \ c_5 c_1 c_2 c_4) + l_3 c_1 s_2 s_5 + d_3 c_1 s_2 \tag{3.30}$$

$$P_y = l_2 s_1 s_2 + l_3(c_5 c_1 s_4 \ + \ c_5 s_1 c_2 c_4) + l_3 s_1 s_2 s_5 + d_3 s_1 s_2 \tag{3.31}$$

$$P_z = l_1 - l_2 c_2 - l_3 c_2 s_5 - l_3 c_4 c_5 s_2 - d_3 c_2 \tag{3.32}$$

3.4.2 Inverse Kinematics

The inverse kinematics (IK) problem involves finding the values of the joint variables for a desired position and orientation (pose) of the end-effector. The robot is over-constrained, having five articulated DoFs instead of six, resulting in no guaranteed solutions to the general position and orientation IK problem [111]. Here, we describe first an analytical approach for finding IK solutions for a fixed orientation of the WRF wrist, followed by a general solution strategy that involves certain constraints.

Position-only IK

Most close-range tasks with the WRF can be performed with its wrist in the upright orientation shown in Figure 3.16a, corresponding to fixing $\theta_4 = 0$, and $\theta_5 = \pi/2$. An object grasped by the WRF with this wrist configuration remains fairly accessible even with slight tilts of the gripper due to the motions of the other DoFs. In this situation, it is sufficient to determine the values for joint variables that specify just the position of the end-effector. The position-only IK

problem has an analytical solution in the first three DoFs. Using the forward kinematics equations,

$$\vec{p} = \begin{bmatrix} P_x \\ P_y \\ P_z \end{bmatrix} = \begin{bmatrix} (l_2 + l_3 + d_3)cos\theta_1 sin\theta_2 \\ (l_2 + l_3 + d_3)sin\theta_1 sin\theta_2 \\ l_1 - (l_2 + l_3 + d_3)cos\theta_2 \end{bmatrix} \tag{3.33}$$

The joint variables for the first three DoFs can be computed for a given position $\vec{P}$ in terms of these parameters for the fixed straight wrist orientation:

$$\begin{bmatrix} \theta_1 \\ d_3 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} tan^{-1}(P_y/P_x) \\ \sqrt{{P_x}^2 + {P_y}^2 + (P_z - l_1)^2} - (l_2 + l_3) \\ cos^{-1}((l_1 - P_z)/(l_2 + l_3 + d_3)) \end{bmatrix} \tag{3.34}$$

Approximate method: Jacobian Pseudoinverse

One approach to solve the position-only IK problem with variable wrist orientation involves approximating the change in joint variables ($\Delta\vec{\theta}$) required for a small change in end-effector position ($\Delta\vec{P}$). This involves determining the Jacobian matrix, J for the transformation between $\vec{P}$ and $\vec{\theta}$, following by computing its Moore-Penrose inverse (pseudoinverse) to find the change in joint angles [31]. Each element of the Jacobian matrix J is defined as:

$$J_{i,j} = \frac{\partial P_i}{\partial \theta_j} \tag{3.35}$$

For the WRF, J is a 3×5 matrix such that:

$$\Delta\vec{P} \approx J\Delta\vec{\theta} \tag{3.36}$$

This leads to the following approximate solution for $\Delta\vec{\theta}$, involving J^+, the pseudoinverse of J:

$$\Delta\vec{\theta} = (J^T J)^{-1} J^T \Delta\vec{P} = (J^+)\Delta\vec{P} \tag{3.37}$$

This approach can also be extended to find approximate solutions for both variable position and orientation, for instance by using Euler angles for the orientation, leading to a 6×5 Jacobian. It results in a fast computational method to implement IK for the WRF, used in Chapter 6 for the human-robot collaboration task.

General Analytical Solution

We now describe an analytical approach for solving the general IK problem, with some constraints to check if a solution is within the subspace of the WRF. We started with the solution to the horizontal panning angle θ_1 using the geometric projection method described in [119] and [43]. As shown in Figure 3.17, θ_1 is the angle between the projection of $\vec{P'} = \vec{O_0 O_5}$ onto the XY-plane, and the X-axis of the base frame O_0. $\vec{P'}$ can be obtained by using the transformation between O_5 and O_0:

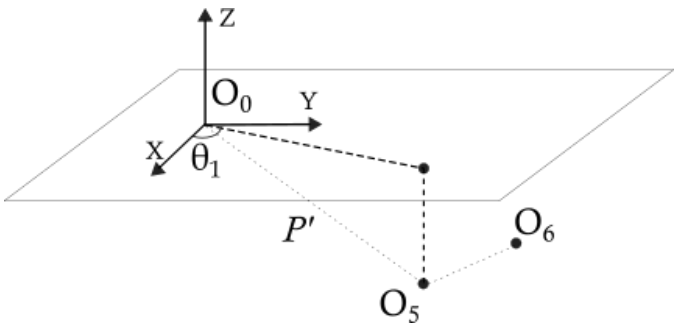


Figure 3.17: Projecting the position of O_5 onto the XY-plane of the base frame O_0 to solve for θ_1.

$$T_0^5 = T_0^6 \, [T_5^6]^{-1} = \begin{bmatrix} \vec{R'}_1 & \vec{R'}_2 & \vec{R'}_3 & \vec{P'} \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.38}$$

The components P'_y and P'_x of $\vec{P'}$ can be found from the corresponding components in the last column of T_0^5. The relation between θ_1 and the projections of $\vec{P'}$ is as follows:

$$tan\theta_1 = \frac{P'_y}{P'_x} = \frac{-l_3 R_{1y} + P_y}{-l_3 R_{1x} + P_x} \tag{3.39}$$

θ_1 can be found using the four-quadrant inverse tangent function, as it lies in the desired range of (-180°, 180°). Also note that θ_2 is never fully vertical, since we have restricted the range of DoF-2 in the open interval (0°, 90°). This ensures that a solution for θ_1 always exists, since O_0, O_1, O_2 and O_5 are never collinear.

The solution strategies for the other joint variables can be grouped into two cases, depending on whether $R_{2z} \neq 0$ or $R_{2z} = 0$, which correspond to whether the links of the robot become coplanar.

Case 1: $R_{2z} \neq 0$ From the forward kinematics, we know that:

$$R_{2z} = -s_2 s_4 \tag{3.40}$$

$R_{2z} \neq 0$ implies that both $s_2 \neq 0$ and $s_4 \neq 0$. Then we can perform the following substitutions:

$$s_4 = \frac{-R_{2z}}{s_2} \tag{3.41}$$

This results in a pair of linear equations in $cos\theta_4$ and $cot\theta_2$:

$$\begin{bmatrix} s_1 & c_1 R_{2z} \\ -c_1 & s_1 R_{2z} \end{bmatrix} \begin{bmatrix} cos\theta_4 \\ cot\theta_2 \end{bmatrix} = \begin{bmatrix} R_{2x} \\ R_{2y} \end{bmatrix} \tag{3.42}$$

We can find $tan\theta_2$ by solving this pair of linear equations, and constrain the solution to lie in the range (0°, 90°). These equations also give us $cos\theta_4$, with which we can find $tan\theta_4$:

$$tan\theta_4 = \frac{-R_{2z}}{s_2\ cos\theta_4} \tag{3.43}$$

θ_4 is computed using the four-quadrant inverse tangent function $atan2(x, y)$ as it lies in the range (-180°, 180°). An inverse tangent function is preferred over an inverse cosine function due to greater numerical stability. Once θ_1, θ_2 and θ_4 are known, equations (3.23) and (3.29) reduce to linear expressions in terms of $sin\theta_5$ and $cos\theta_5$:

$$\begin{bmatrix} -c_2 & -c_4 s_2 \\ -c_4 s_2 & c_2 \end{bmatrix} \begin{bmatrix} sin\theta_5 \\ cos\theta_5 \end{bmatrix} = \begin{bmatrix} R_{1z} \\ R_{3z} \end{bmatrix} \tag{3.44}$$

θ_5 can be obtained using the inverse tangent function $atan2(x, y)$, and constrained to lie in the range (0°, 180°). Once all joint angles are known, equations (3.30-3.32) yield three candidate solutions for the length of the prismatic joint d_3:

$$d_3^1 = (P_x - l_2 c_1 s_2 + l_3(c_5 s_1 s_4 \ + \ c_5 c_1 c_2 c_4) - l_3 c_1 s_2 s_5)/c_1 s_2 \tag{3.45}$$

$$d_3^2 = (P_y - l_2 s_1 s_2 - l_3(c_5 c_1 s_4 \ + \ c_5 s_1 c_2 c_4) - l_3 s_1 s_2 s_5)/s_1 s_2 \tag{3.46}$$

$$d_3^3 = (P_z - l_1 + l_2c_2 + l_3c_2s_5 + l_3c_4c_5s_2)/c_2 \tag{3.47}$$

The candidate for d_3 that lies within the range [0.33 m, 0.45 m] is chosen as the solution.

Case 2: $R_{2z} = 0$ A different strategy must be employed in situations when R_{2z} is zero. Since $R_{2z} = s_2s_4$, and we have restricted θ_2 to lie in the open interval (0°, 90°), $R_{2z} = 0 \Rightarrow s_4 = 0$. Additionally, since θ_4 is also restricted to be in the open interval (-180°, 180°), we obtain $\theta_4 = 0$.

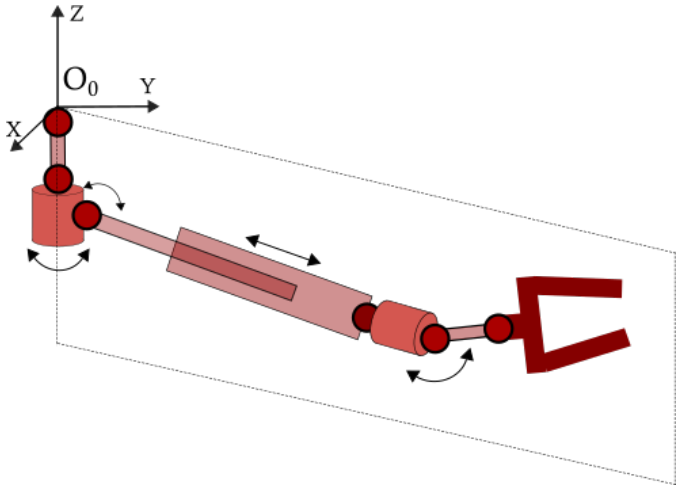


Figure 3.18: All links of the robotic arm become co-planar when $\theta_4 = 0$.

This places the robot in a configuration similar to the one shown in Figure 3.18, where all the links of the arm become co-planar. With $\theta_4 = 0$, equations (3.21, 3.22) and (3.27 - 3.29) simplify to:

$$R_{1x} = -c_1cos(\theta_2 + \theta_5) \tag{3.48}$$

$$R_{1y} = -s_1cos(\theta_2 + \theta_5) \tag{3.49}$$

$$R_{3x} = -c_1 sin(\theta_2 + \theta_5) \tag{3.50}$$

$$R_{3y} = -s_1 sin(\theta_2 + \theta_5) \tag{3.51}$$

$$R_{3z} = cos(\theta_2 + \theta_5) \tag{3.52}$$

This gives us the sum $(\theta_2 + \theta_5)$:

$$tan(\theta_2 + \theta_5) = \frac{R_{3x}}{R_{1x}} = \frac{R_{3y}}{R_{1y}} \tag{3.53}$$

To find the the joint angle θ_2 separately, we look at the triangle formed by the points O_0, O_2 and O_5, as shown in Figure 3.19.

The edge lengths t_1, t_2 and t_3 of this triangle are known quantities, since the vector $\vec{O_0O_2}$ is fixed, and the vector $\vec{O_0O_5}$ can be found using the last column of the matrix T_0^5 in equation (3.38). We can apply the cosine formula to this triangle:

$$t_1 = \left\| \vec{O_0O_2} \right\| \tag{3.54}$$

$$t_2 = \left\| \vec{O_0O_5} \right\| \tag{3.55}$$

$$t_3 = \left\| \vec{O_0O_2} - \vec{O_0O_5} \right\| \tag{3.56}$$

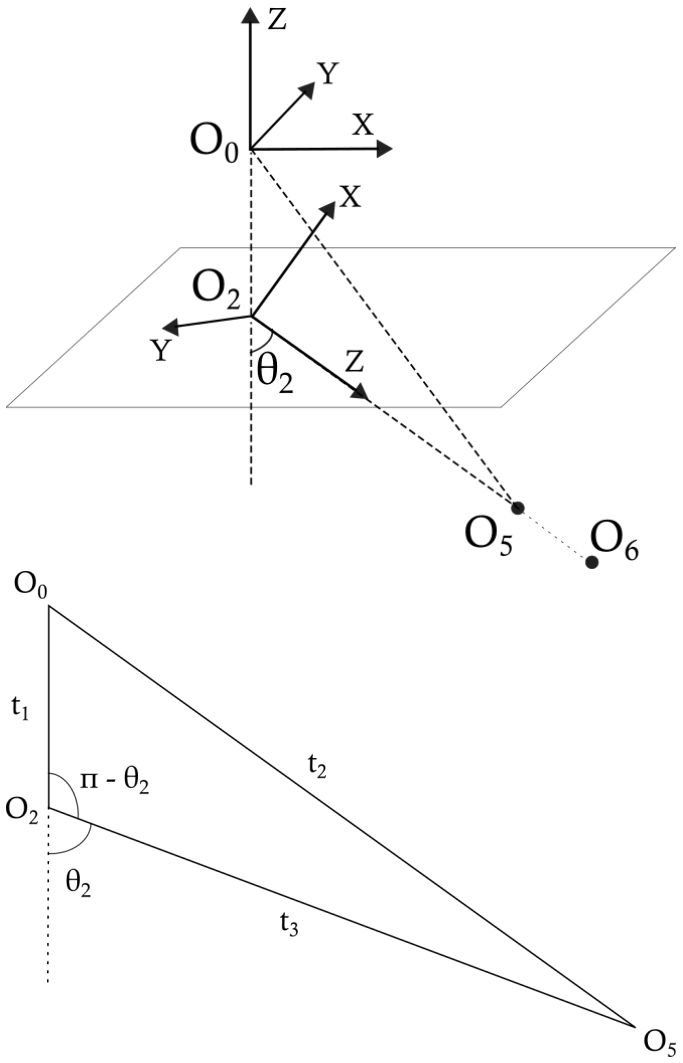


Figure 3.19: O_0, O_2 and O_5 are coplanar. The triangle formed by these points is used to find θ_2.

$$cos(\pi - \theta_2) = \frac{t_1^2 + t_3^2 - t_2^2}{2t_1t_3} \tag{3.57}$$

We can now solve for θ_2:

$$\theta_2 = \pi - acos([\frac{t_1^2 + t_3^2 - t_2^2}{2t_1t_3}]) \tag{3.58}$$

Once θ_2 is known, we get two possible candidates for θ_5:

$$\theta_5 = tan^{-1}\frac{R_{3x}}{R_{1x}} - \theta_2 = tan^{-1}\frac{R_{3y}}{R_{1y}} - \theta_2 \tag{3.59}$$

Having found all the joint angles, we can proceed as earlier, using equations (3.45 - 3.47) to find d_3, the length of the prismatic joint.

Constraints on Solutions We chose between multiple candidate solutions by applying the condition that the variables must lie within the intervals specified in Table 3.4. Apart from these, there are also conditions on the revolute joint angles θ_1 - θ_5 that arise from the forward kinematic equations. These conditions make use of the individual transformation matrices T_i^{i+1} for each row of Table 3.4.

Let us consider the product $T_0^1 T_1^2 T_2^3$:

$$T_0^1 T_1^2 T_2^3 = \begin{bmatrix} -c_1c_2 & -s_1 & c_1s_2 & d_3c_1s_2 \\ -c_2s_1 & c_1 & s_1s_2 & d_3s_1s_2 \\ -s_2 & 0 & -c_2 & l_1 - d_3c_2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.60}$$

The entry in row 3, column 2 of this matrix is zero. Using equation (3.19), we can compute this matrix in terms of the final transformation T_0^6 :

$$T_0^1 T_1^2 T_2^3 = T_0^6 [T_3^4 T_4^5 T_5^6]^{-1} \tag{3.61}$$

Equating corresponding matrix entries at row 3, column 2 on both sides, we

obtain the following condition on θ_4 and θ_5:

$$R_{1z}c_5s_4 - R_{2z}c_4 + R_{3z}s_4s_5 = 0 \tag{3.62}$$

Similarly, in the product $T_3^4T_4^5T_5^6$, the element at row 3, column 2 is also zero:

$$T_3^4T_4^5T_5^6 = \begin{bmatrix} c_4c_5 & s_4 & c_4s_5 & l_3c_4c_5 \\ s_4c_5 & -c_4 & s_4s_5 & l_3s_4c_5 \\ s_5 & 0 & -c_5 & l_2 + l_3s_5 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.63}$$

This can also be written in terms of the final transformation matrix:

$$T_3^4T_4^5T_5^6 = [T_0^1T_1^2T_2^3]^{-1}T_0^6 \tag{3.64}$$

Similar to equation (3.62), we obtain the following condition on θ_1 and θ_2:

$$R_{2x}c_1s_2 - R_{2z}c_2 + R_{2y}s_1s_2 = 0 \tag{3.65}$$

Candidate IK solutions that do not satisfy conditions (3.62) and (3.65) are rejected.

3.5 Conclusion

This chapter addresses the analysis and evaluation of successive prototypes (Models I and II) of the WRF in terms of kinematics, enhancement in total reachable workspace volume, and biomechanical loads on the user's elbow and shoulder joints.[1] We also ensured that the actuators used to build the WRF were

[1]Portions of this chapter have been published in [114].

up to the task in illustrative usage scenarios, by measuring the motor torques at the most loaded DoFs.

Some of the major design guidelines from Chapter 2 were regarding the weight, balance, and dexterity of the WRF. Along with these factors, we explored the types of control modes and levels of robot autonomy desirable by users through a pilot interaction study. While adding more dexterity to the robot through an additional wrist pitching DoF increased the weight, we offset some of the loads by reducing the distance between the first two motors. This led to the final WRF prototype: Model III, whose kinematics were described in detail, including an analytical solution for its inverse kinematics.

While the static analysis of biomechanical loads showed the loads to be within ergonomic limits, it was performed purely in simulation with simplifying assumptions. The real forces in the human shoulder and elbow joints were not measured directly using orthopedic sensors. Also, the skin effects of the interface between the human and robot system (through the medical brace), which play a role in the subjective ergonomic experience, were not accounted for. These limitations were reflected in users' reactions of feeling fatigued after wearing the WRF for longer than 40-45 minutes. One strategy for improving ergonomics by reducing the worn weight while maintaining functionality, is to place the control and actuation systems off-board with cable-driven [33] or hydraulic [116] mechanisms. These designs present another trade-off, this time between the functionality and overall footprint of the device.

The reachable workspace volume analysis performed here included point clouds of the end-effector's reachable positions. When executing tasks, it may also be important to determine if a point in the workspace is reachable with a

specific orientation of the gripper, for instance when holding an object that must not be tilted. While an approximate solution to a desired position and orientation of the gripper may be found using numerical IK methods, owing to the dimension subspace mismatch between the WRF (5-DoF) and Cartesian space (6-DoF), a specific desired pose may be unreachable. However, for the tasks described in this dissertation, we considered the trade-off of adding another DoF to be too large against the additional weight it would impose on the user.

Through the interaction pilot, we found that while giving the robot direct speech commands was intuitive, users were quickly able to adapt to and perform well with an "autonomous" robot (controlled here through WoZ). This led to the formulation of a more extensive user study, described in Chapter 6, where the robot was truly autonomous in a collaborative pick-and-place scenario, operating with supervised learning models trained using data from speech controlled trials.

In the next chapter, we take a more detailed look at the impact of the robot's motion on a user's musculoskeletal system, and describe a framework for motion planning that minimizes these biomechanical loads.

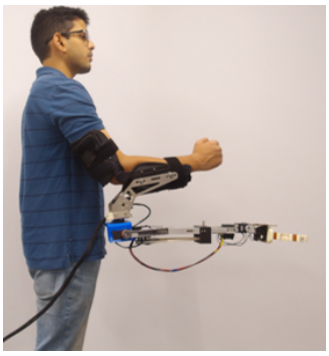
CHAPTER 4

BIOMECHANICS-AWARE MOTION PLANNING

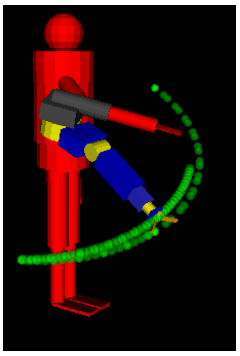
As a user must support a Supernumerary Robotic (SR) device with their own body, one of the major challenges in their design is to reduce the ergonomic load on the human while maintaining functionality. For the WRF, we showed in Chapter 3 that the force and moment loads on a user's shoulder and elbow joints were within the human ergonomic ranges in scenarios involving fetching and assisted handovers. This analysis assumed rigid body models for both the human arm and robot, and considered the robot's effect on the user during pre-specified trajectories.

In this chapter, we extend the biomechanics analysis by incorporating, in simulation, a detailed muscle model of the human arm into a motion planning framework that finds robot trajectories given a start and goal state, while minimizing human muscle loads (Figure 4.1). Henceforth, all studies and analyses described in this work refer to the final WRF Model III prototype.

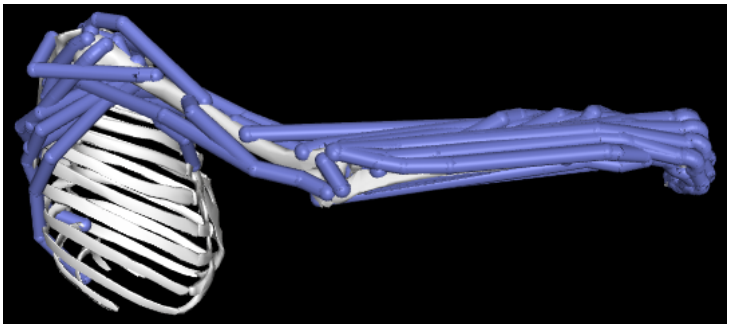
SR devices span a wide range of sizes, functions, and mounting styles on the user. Larger devices in the form of additional arms are typically supported by the user's shoulders [85], mounted as a backpack [94], or at the hip [79]. The WRF is mounted on a user's arm through a medical brace, with the first motor vertically aligned with their elbow (Fig. 4.1a). For the purpose of generating trajectories with low biomechanical loads for the WRF, we drew upon two bodies of work:



(a) WRF mounted on a user



(b) Illustrative robot path



(c) OpenSim muscle model

Figure 4.1: Trajectories are generated for the Wearable Robotic Forearm (WRF) that minimize human muscle loads determined from the model shown in (c).

Biomechanics Simulation

OpenSim [35] has been widely used in the biomechanics community for dynamic simulations of human motion, with an API accessible through MATLAB [66]. Specifically for wearable robotics, combined human-robot models in OpenSim have been used for rehabilitation exoskeletons, to study their effect on the user [34], as well as to parameterically improve robot designs [11]. We addressed the biomechanics of interaction between the human and robot after the design phase, during usage in close-range tasks.

Manipulator Motion Planning

The motion planning problem for robotic arms has been studied extensively, with sampling-based planners being the most common approach [65]. In motion planning for the WRF, a feasible trajectory may often be found by direct interpolation in the configuration space between the start and goal states (Figure 4.1b). To improve upon this initial guess and minimize muscle loads, we used Stochastic Trajectory Optimization for Motion Planning (STOMP) [59], a planner that does not require gradients for its cost function. In each iteration, STOMP samples noisy trajectories drawn from a Gaussian distribution around an initial guess, and updates the result by minimizing the expected value of a state-dependent cost function. To further improve the performance of STOMP, we applied a computationally inexpensive local search approach to generate better initial trajectories than the linear interpolations in configuration space.

We developed a full simulation pipeline for determining the effects of a wearable robot's trajectory on the human arm, as well as a muscle force-based approach for planning optimal robot trajectories. We then applied this framework to ten start and goal state pairs. The combined OpenSim and STOMP model, on initialization with locally optimal paths, resulted in trajectories with up to 23.47% lower mean muscle fiber forces compared to interpolated shortest paths in the robot's configuration space.

4.1 Biomechanics Model

4.1.1 WRF Dynamics

The kinematics of the WRF can be described using the Denavit-Hartenberg (D-H) parameters $[\alpha, a, d, \theta]$, listed in Table 3.4 for each DoF. In all trajectories considered for optimization, the wrist joint and gripper were assumed to be fixed in the pose shown in Figure 4.1a, with joint angles $\theta_4 = 0$ and $\theta_5 = \pi/2$.

The start and goal states for the WRF were specified in terms of a configuration space vector $\boldsymbol{\theta} = [\theta_1, \theta_2, d_3]^T$, where θ_1 and θ_2 are the joint angles for horizontal panning (DoF-1) and vertical pitching (DoF-2), and d_3 is the length of the prismatic joint (DoF-3).

A trajectory $\boldsymbol{\Theta} = [\boldsymbol{\theta_1}, \boldsymbol{\theta_2}, ..., \boldsymbol{\theta_N}]$ is a set of N poses, going from the start state $\boldsymbol{\theta_1}$ to the goal state $\boldsymbol{\theta_N}$.

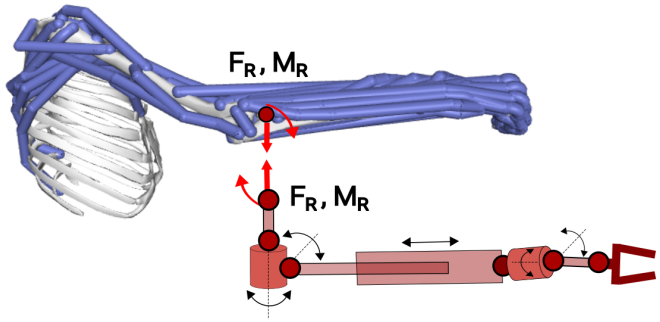


Figure 4.2: The reaction force $\mathbf{F_R}$ and moment $\mathbf{M_R}$ are assumed to act on the human forearm near the elbow joint.

For a given trajectory $\boldsymbol{\Theta}$ of the WRF, we can compute the force $\mathbf{F_R}$ and moment $\mathbf{M_R}$ applied at the base of the first DoF due to the robot's motion. The WRF has five links, one between between each pair of successive DoFs. As a simpli-

fying assumption, they were considered to be cylinders with masses, lengths, and diameters as listed in Table 4.1.

Table 4.1: Inertial parameters for the WRF

Link	Mass (kg)	Length (mm)	Diameter (mm)
(1)	0.470	150	55
(2)	0.425	300	68
(3)	0.457	50	35
(4)	0.248	45	25
(5)	0.350	50	35

Since most of the structural components were mounted on the user's forearm, we assumed that the corresponding reaction loads $\mathbf{F_R}$ and $\mathbf{M_R}$ were applied on the radius and ulna bones near the elbow joint, as shown in Figure 4.2.

An iterative Newton-Euler algorithm [72] was applied to compute $\mathbf{F_R}$ and $\mathbf{M_R}$, as described in Chapter 3, with zero external forces and moments at the end-effector as boundary conditions.

In each trajectory, the speeds of the DoFs were held constant at $\dot{\theta}_1 = 1.5$ rad/s, $\dot{\theta}_2 = 1.0$ rad/s, $\dot{d}_3 = 0.1$ m/s. These values were found to be the upper limits for comfortable operation of the WRF during the pilot interaction study. For simplicity, the velocities were considered to start and stop instantaneously at the initial and goal poses. Keeping velocities constant reduced the trajectory optimization problem to search only within the robot's configuration space.

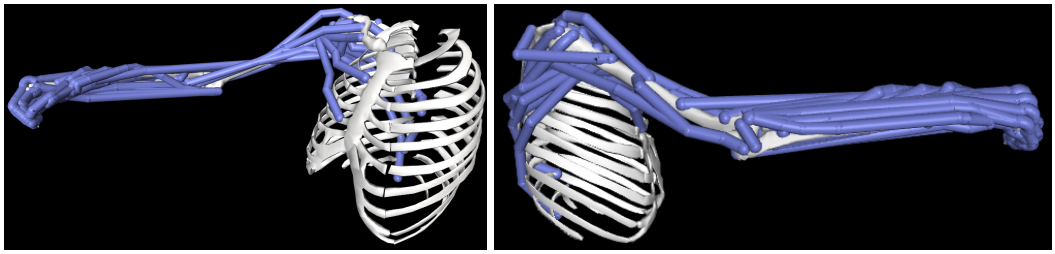


Figure 4.3: OpenSim model of the human right arm from [95].

4.1.2 Human Arm Model

Along each trajectory $\boldsymbol{\Theta}$, the WRF applies reaction forces $\mathbf{F_R}$ and moments $\mathbf{M_R}$ on the user's forearm. The muscle fibers in their arm are activated in response to the external loads. We aimed to generate trajectories for the WRF that minimize the active forces in these muscle fibers.

An OpenSim-based human upper limb musculoskeletal model from [95] was used for this purpose (Figure 4.3). It contains fifty Hill-type muscle-tendon actuators [50] with their dynamic properties as described in [97]. This model includes seven articular DoFs for the human shoulder, elbow, and wrist.

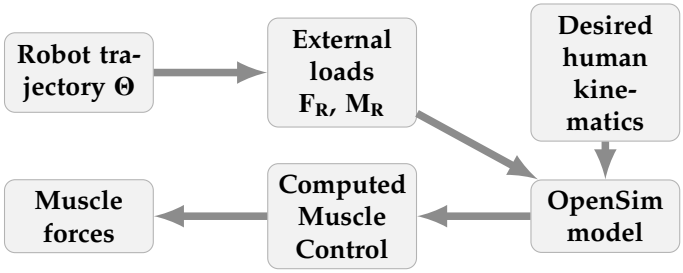


Figure 4.4: Schematic of biomechanics simulation.

We followed a procedure similar to [11], using the MATLAB-OpenSim interface for conducting the biomechanics simulations, as shown in Figure 4.4. For a given robot trajectory, the computed external loads $\mathbf{F_R}$ and $\mathbf{M_R}$ were applied to the human radius and ulna bones (Fig 4.2).

In the OpenSim model, these external loads were considered to act on the user's joints, necessitating appropriate muscle reactions to track a desired human kinematic trajectory. For all robot trajectories in this chapter, the muscle efforts for the human arm aimed to keep it static in the pose shown in Figure 4.3. The human joints of interest in this case were the three DoFs at the shoulder,

and two at the elbow. The muscles controlling wrist flexion and deviation were considered to be unaffected by the WRF's motion.

The desired joint angles for the shoulder and elbow are listed in Table 4.2. The desired velocities and accelerations for all human arm joints were zero throughout.

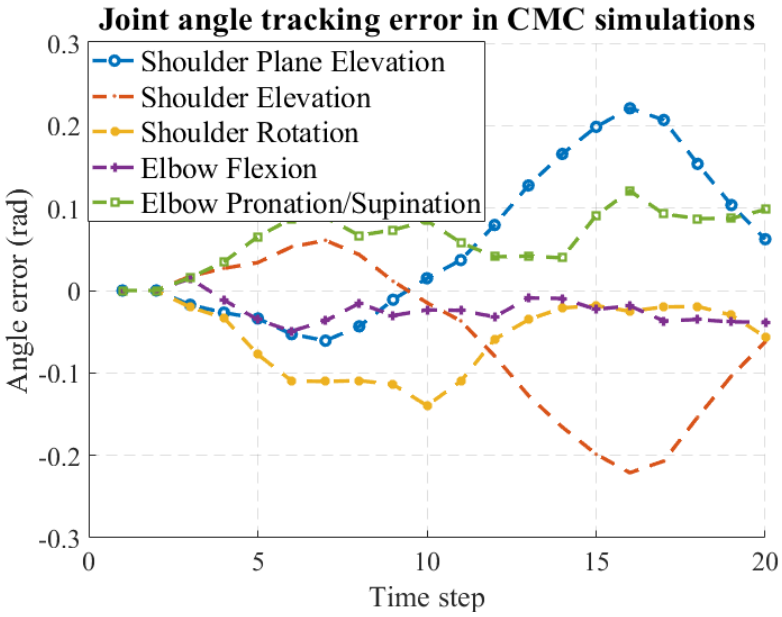
Table 4.2: Desired human kinematics

Degree of Freedom	Joint Angle (rad)
(1) Shoulder Plane Elevation	1.0
(2) Shoulder Elevation	1.2
(3) Shoulder Rotation	0.3
(4) Elbow Flexion	0.7
(5) Elbow Pronation/Supination	0.0

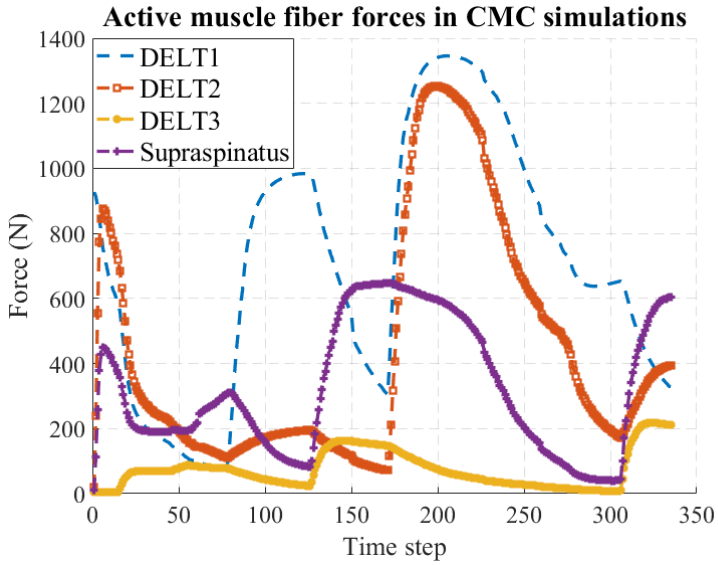
These desired kinematics were tracked using Computed Muscle Control (CMC) [110], a simulation technique that combines static optimization with feedforward control and proportional-derivative (PD) feedback control to determine the necessary muscle excitations to generate a human trajectory. The forward dynamics model described in [97] was used to determine the active muscle fiber lengths and active fiber forces based on these excitations.

To illustrate this simulation process, consider a trajectory $\mathbf{\Theta_r}$, going from $\boldsymbol{\theta_1} = [-1.42, 0.77, 0.43]$ to $\boldsymbol{\theta_N} = [1.0, 1.57, 0.40]$ with $N = 25$. This trajectory represents the WRF lifting an object placed below and to the right of the user, and bringing it to their left hand, as shown in Figure 4.1b. The intermediate poses $\boldsymbol{\theta_i}$, $i \in [2, N-1]$, are linearly interpolated between $\boldsymbol{\theta_1}$ and $\boldsymbol{\theta_N}$.

The OpenSim model is initialized with the human kinematics listed in Table 4.2, and external loads $\mathbf{F_R}$ and $\mathbf{M_R}$ are computed for $\mathbf{\Theta_r}$ using the WRF's dynamics model. The CMC simulation divides this trajectory into 20 time steps for



(a)



(b)

Figure 4.5: Illustrative CMC results for the trajectory $\mathbf{\Theta}_r$: (a) tracking errors in human joint angles are below 0.25 rad, (b) active muscle forces in four fibers.

which the kinematic errors are determined after each iteration. The PD feedback control gains for tracking each human joint are set to $K_p = 400$ and $K_v = 40$, in order to achieve critically-damped error dynamics [110] with $K_v = 2\sqrt{K_p}$. The resulting muscle excitations were interpolated, in this case to 335 steps, and used

to determine the corresponding active human muscle fiber forces $\mathbf{f}_{h,i}$, $i \in [1, 50]$ along the trajectory.

Fig 4.5 illustrates the results of this CMC simulation for $\mathbf{\Theta_r}$. The joint angle errors were below 0.25 radians in the five human kinematic DoFs of interest. Fig 4.5b shows the estimated active fiber forces in the first four shoulder muscles as listed in [95]: three muscles forming the deltoid group, and the supraspinatus.

We considered the first 32 muscles listed in [95], corresponding to the human shoulder and elbow joints to be relevant, and assumed that the WRF's motion has no effect on the wrist and hand muscles. The fiber forces generated in these muscles formed the basis of trajectory optimization and planning described in the next section.

4.2 Trajectory Optimization

The planning framework uses the biomechanics model described above to generate ergonomic WRF trajectories. While the overall goal is to plan trajectories that minimize the muscle fiber forces in the user's arm, in the interest of faster convergence, the cost function for the planner also includes other factors that positively correlate with this muscle force, e.g. reaction force and moment norms, and smoothness of the trajectory. We also describe an approach for improving the initialization of the planner.

4.2.1 STOMP with CMC

The backbone for planning WRF trajectories in this chapter is STOMP, chosen since it allows for the specification of a customized state-dependent cost function without the requirement of gradients for the costs [59]. The optimization objective is specified as follows:

$$\begin{aligned} Q(\boldsymbol{\Theta}) &= \mathbf{w}_a^T \boldsymbol{\Theta}^T \mathbf{R} \boldsymbol{\Theta} \; + \; \mathbf{w}_q^T \mathbf{q}(\boldsymbol{\Theta}) \\ \boldsymbol{\Theta}^* &= \arg\min_{\boldsymbol{\Theta}} \; Q(\boldsymbol{\Theta}) \end{aligned} \tag{4.1}$$

Starting with an initial trajectory $\boldsymbol{\Theta}^0$, the optimal trajectory $\boldsymbol{\Theta}^*$ minimizes the weighted sum of squared accelerations $\boldsymbol{\Theta}^T \mathbf{R} \boldsymbol{\Theta}$ and state-dependent costs $\mathbf{q}(\boldsymbol{\Theta})$. The matrix $\mathbf{R} = \mathbf{A}^T \mathbf{A}$ is derived from a second-order finite differencing matrix $\mathbf{A}$, such that $\mathbf{A}\boldsymbol{\Theta} = \ddot{\boldsymbol{\Theta}}$ gives the accelerations of the joints. This term incentivizes the generation of smooth trajectories.

The STOMP algorithm generates K noisy trajectories drawn from a Gaussian distribution around the initial guess $\boldsymbol{\Theta}^0$, and weighs them based on the cost $Q(\boldsymbol{\Theta}_K)$ to estimate a stochastic gradient update, and repeats this process over successive iterations.

The state-dependent trajectory cost function contained five individual terms, $\mathbf{q}(\boldsymbol{\Theta}) = [q_h, q_f, q_m, q_{df}, q_{dm}]^T$, defined as follows:

Human Muscle Cost q_h

For each noisy trajectory $\boldsymbol{\Theta}_K$, a CMC simulation was performed to obtain the muscle fiber forces generated in the user's right arm due to the WRF's motion

along $\mathbf{\Theta_K}$.

$$q_h = \sum_{j=1}^{M} F_{h,j}(\mathbf{\Theta_K}) \tag{4.2}$$

The total fiber force $F_{h,j}$ for muscle j is the L_1 norm of the force vector $\mathbf{f}_{h,j}$ (Figure 4.5b) interpolated for that muscle along $\mathbf{\Theta_K}$. q_h was computed for the first $M = 32$ out of 50 muscles in the OpenSim model, with the wrist and hand muscles excluded.

Reaction Load Costs q_f and q_m

The costs q_f and q_m accounted for the total reaction force $\mathbf{F_R}$, and reaction moment $\mathbf{M_R}$ experienced by the user. The norms of $\mathbf{F_R}$ and $\mathbf{M_R}$ were summed for each pose $\boldsymbol{\theta_i}$, $i \in [1, N]$ along a noisy trajectory $\mathbf{\Theta_K}$:

$$q_f = \sum_{i=1}^{N} ||\mathbf{F_R}(\boldsymbol{\theta_i})|| \quad , \quad q_m = \sum_{i=1}^{N} ||\mathbf{M_R}(\boldsymbol{\theta_i})|| \tag{4.3}$$

Load Deviation Costs q_{df} and q_{dm}

In addition to direct costs q_f and q_m for the reaction loads, q_{df} and q_{dm} penalized positive deviations in the reaction loads compared to the initial trajectory $\mathbf{\Theta^0}$, and reward negative deviations. This allowed for a more aggressive optimization towards a low-cost path.

The deviations were computed based on the relative difference in force and moment norms at each pose between a candidate trajectory $\mathbf{\Theta_K}$ and the initial

trajectory $\boldsymbol{\Theta^0}$. The relative difference in force ΔF_i at step i was defined as:

$$\Delta F_i = \frac{\|\mathbf{F_R}(\boldsymbol{\theta_i})\| - \left\|\mathbf{F_R}(\boldsymbol{\theta_i^0})\right\|}{\left\|\mathbf{F_R}(\boldsymbol{\theta_i^0})\right\|} \tag{4.4}$$

Depending on the sign of ΔF_i, a corresponding deviation metric $d_{f,i}$ was obtained as follows:

$$d_{f,i} = \begin{cases} e^{(\lambda_1 \Delta F_i)} - 1, & \Delta F_i \leq 0 \\ e^{(\lambda_2 \Delta F_i)} - 1, & \Delta F_i > 0 \end{cases} \tag{4.5}$$

where $\lambda_1 < \lambda_2$, resulting in a higher positive penalty for large increases in the reaction force norm (Figure 4.6). A similar procedure was applied to compute the deviation metric $d_{m,i}$ for the reaction moment norm.

$$\Delta M_i = \frac{\|\mathbf{M_R}(\boldsymbol{\theta_i})\| - \left\|\mathbf{M_R}(\boldsymbol{\theta_i^0})\right\|}{\left\|\mathbf{M_R}(\boldsymbol{\theta_i^0})\right\|} \tag{4.6}$$

$$d_{m,i} = \begin{cases} e^{(\lambda_1 \Delta M_i)} - 1, & \Delta M_i \leq 0 \\ e^{(\lambda_2 \Delta M_i)} - 1, & \Delta M_i > 0 \end{cases} \tag{4.7}$$

Using these metrics, the total deviation costs for a trajectory were defined as follows:

$$q_{df} = \sum_{i=1}^{N} d_{f,i} \ , \ q_{dm} = \sum_{i=1}^{N} d_{m,i} \tag{4.8}$$

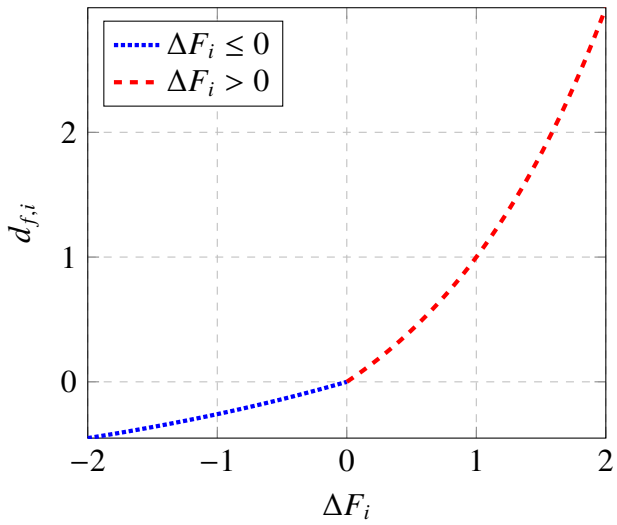


Figure 4.6: Deviation metric $d_{f,i}$ as a function of ΔF_i, with higher penalties for positive deviations.

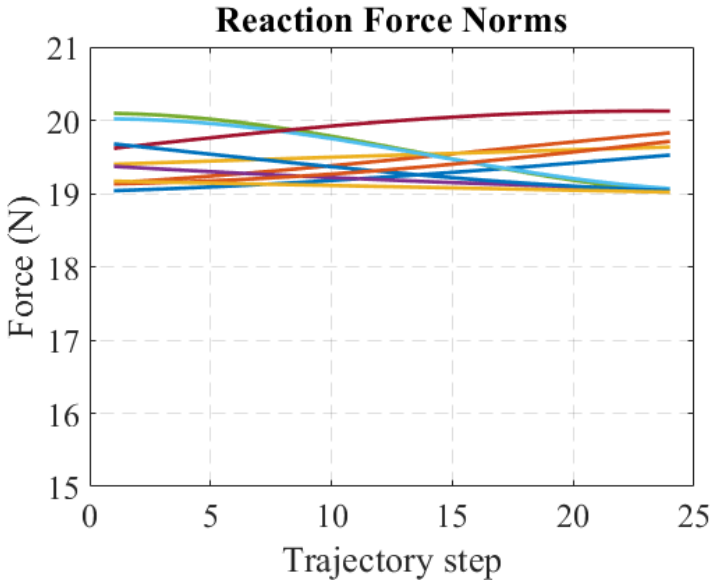
There is a trade-off between the range of exploration for the noisy trajectories and the rate of convergence. We set the number of noisy trajectories sampled in each iteration, $K = 8$, and found that STOMP converged at around 25 iterations for each of the ten pairs of start and goal states considered in this work. However, even with a relatively small number of iterations, the task of finding a trajectory with low muscle fiber forces was computationally expensive due to the CMC simulations.

4.2.2 Local Search for Initialization

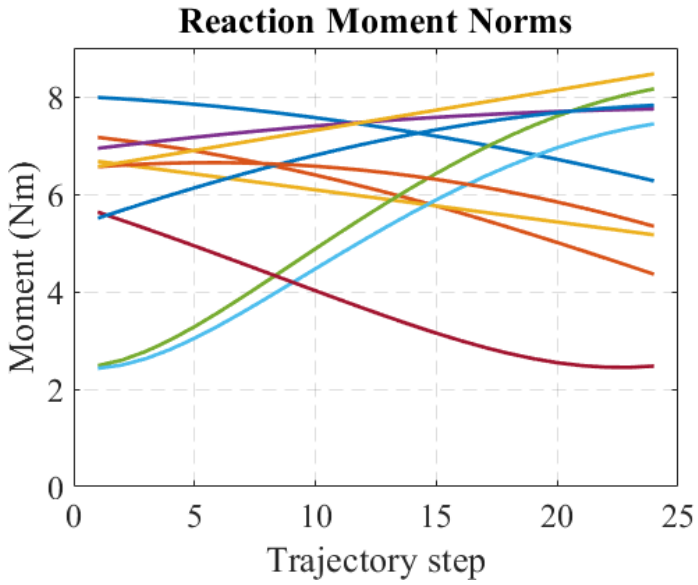
Within the constraint of computational costs for CMC, we improved the convergence of STOMP by generating better trajectories for its initialization while keeping the number of iterations fixed at 25.

Starting with the linearly interpolated shortest paths in configuration space between the start and goal states, we found a trajectory that minimizes the norm

of the reaction moment $\mathbf{M_R}$. This results in the trajectory $\boldsymbol{\Theta}^*_{M'}$, consisting of poses $\boldsymbol{\theta}^*_i$, determined by local minimization in a greedy manner along successive steps from the start state to the goal state.



(a)



(b)

Figure 4.7: (a) Reaction force norms $\|\mathbf{F_R}\|$, varied much less than (b) moment norms $\|\mathbf{M_R}\|$, along the linearly interpolated configuration space paths for the ten example start and goal states.

This strategy emerged from an observation of the trends for reaction forces and moments. Given a trajectory $\mathbf{\Theta}$ consisting of N steps $\boldsymbol{\theta}_i$, $i \in [1, N]$, the corresponding reaction forces $\mathbf{F_R}$ and moments $\mathbf{M_R}$ on the human arm were computed using the WRF's dynamics model. Across all linearly interpolated configuration space trajectories considered here, relatively little variation was seen in $\|\mathbf{F_R}\|$ (between ˜19 and ˜21 N) compared to the variation in $\|\mathbf{M_R}\|$ (between ˜2 and ˜9 Nm), as shown in Figure 4.7. This suggests that minimizing $\|\mathbf{M_R}\|$ alone would be a more suitable objective, leading to the following constrained nonlinear optimization problem:

$$\boldsymbol{\theta}_i^* = \arg\min_{\boldsymbol{\theta}_i} \|\mathbf{M_R}(\boldsymbol{\theta}_i)\|^2 \tag{4.9}$$

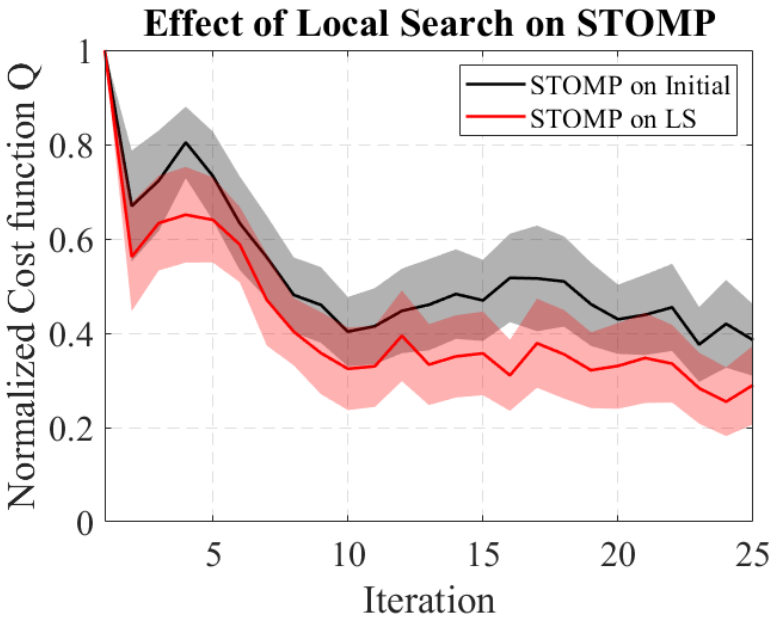


Figure 4.8: On average, local search (red) leads to a relative improvement in the STOMP cost function over linearly interpolated initial trajectories (black).

An interior-point algorithm [27] was used to compute $\boldsymbol{\theta}_i^*$. The lower and upper bounds for feasible configurations were determined by the speeds of each DoF, $\dot{\theta}_1$, $\dot{\theta}_2$, and $\dot{d}_3$, considered to be constant as described in the WRF dynamics

model. For instance, the horizontal panning joint angle θ_1 was bounded by $[\theta_1 - \dot{\theta}_1 \Delta t, \theta_1 + \dot{\theta}_1 \Delta t]$, where Δt is the maximum trajectory time ΔT (taken to be 3s), divided by the number of steps N. The algorithm was initialized with the trajectory $\mathbf{\Theta^0}$, consisting of linearly interpolated poses between $\boldsymbol{\theta}_\mathbf{1}^\mathbf{0}$ and $\boldsymbol{\theta}_{\boldsymbol{N}}^\mathbf{0}$.

Although the STOMP cost function $Q(\mathbf{\Theta})$ already contains a term for $\|\mathbf{M_R}\|$, explicitly optimizing for it in the local search before starting the STOMP iterations led to an improvement over initial trajectories obtained through linear interpolation. Figure 4.8 shows the means and standard errors for the cost functions, normalized by the cost of the first iteration, for each of the ten start and goal state pairs, resulting from STOMP performed on the linearly interpolated configuration space trajectories, and on the results from the local search. As discussed in the next section, this approach consistently generated trajectories with reduced muscle fiber forces compared to the other methods.

4.3 Results

In this section, we compare WRF trajectories computed using all four approaches described in this chapter, for ten start and goal state pairs.

The biomechanical STOMP-based optimization was performed twice for each pair, first using an initial guess for the trajectory, and then using the result from the local search. The initial guess for a valid trajectory was the linear interpolation between the start and goal states in configuration space. The local search-based approach found a trajectory around this initial guess that minimized the reaction moment norm $\|\mathbf{M_R}\|$.

The results for mean total muscle fiber forces from CMC simulations on the outputs from each of these four approaches are listed in Table 4.3. Trajectories (1) and (2) correspond to the two stages of the task shown in Figure 3.10b: a self-handover (SH-1 & 2), and (3) represents the assisted two-person handover (A2PH) shown in Figure 3.10c. The other start and goal state pairs were randomly generated in the WRF's configuration space (R-1 to R-7) with path lengths of at least 0.8 m.

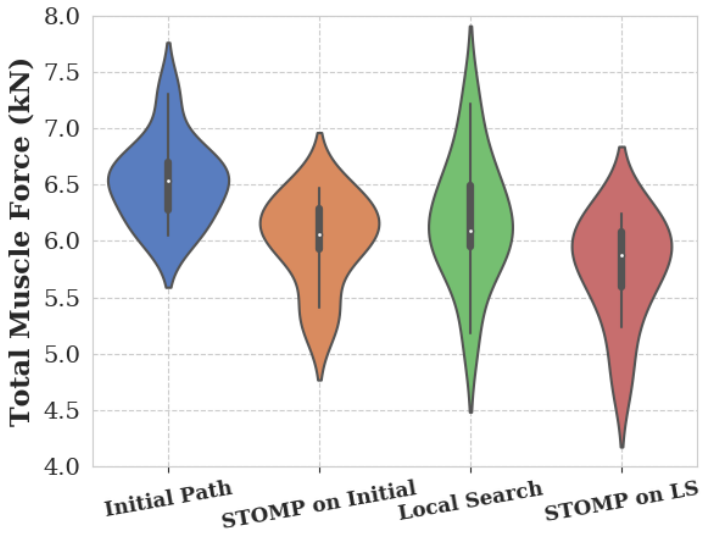


Figure 4.9: Violin plots (box plots with kernel density estimates) of mean total muscle fiber force data from Table 4.3 for all trajectories.

As listed in Table 4.3, STOMP performed on the local search output consistently resulted in WRF trajectories with lower mean total muscle fiber forces in the human arm compared to the other approaches.

The median improvement for STOMP on local search, compared to the initial interpolated paths is by 11.00 %, compared to STOMP on the initial trajectories by 2.61 %, and compared to local search alone by 5.26 %. The degree of improvement varies with the start and goal state pairs (Table 4.4). Among all

Table 4.3: Mean total muscle fiber forces for ten start and goal state pairs

Trajectory	Start State $\boldsymbol{\theta_1}$	Goal State $\boldsymbol{\theta_N}$	Mean Total Muscle Fiber Force (N)			
			Initial Path	STOMP on Initial Path	Local Search	STOMP on Local Search
(1) Reaching for object (SH-1)	[0.00, 1.57, 0.33]	[-1.42, 0.77, 0.43]	6105.84	5411.82	5925.77	5235.18
(2) Handing over object (SH-2)	[-1.42, 0.77, 0.43]	[0.00, 1.57, 0.33]	7298.53	6301.19	7215.32	6246.42
(3) Assisted handover (A2PH)	[-1.33, 0.89, 0.41]	[2.41, 0.03, 0.45]	6052.42	5970.94	6033.73	5950.68
(4) Random-1 (R-1)	[3.08, 1.08, 0.43]	[0.32, 1.52, 0.40]	6425.37	6357.45	5832.79	5720.25
(5) Random-2 (R-2)	[1.00, 1.57, 0.41]	[-3.00, 0.95, 0.43]	6521.30	5910.62	6016.74	5802.18
(6) Random-3 (R-3)	[1.85, 1.43, 0.38]	[-2.62, 0.62, 0.41]	6554.13	6036.47	6141.92	5954.56
(7) Random-4 (R-4)	[2.34, 1.08, 0.42]	[-1.61, 0.82, 0.39]	6721.73	6245.51	6372.70	6123.26
(8) Random-5 (R-5)	[2.55, 0.08, 0.44]	[-1.35, 1.55, 0.42]	6626.21	6464.93	6668.06	6227.89
(9) Random-6 (R-6)	[1.23, 0.06, 0.39]	[-1.06, 1.56, 0.38]	6227.89	5266.09	5176.58	4765.93
(10) Random-7 (R-7)	[1.98, 1.42, 0.34]	[-2.15, 1.52, 0.44]	6770.48	6074.39	6524.06	5546.47

Table 4.4: Percentage reductions in mean q_h for STOMP initialized with local search (LS) compared to other methods

Trajectory	Percentage reduction for STOMP+LS versus: (%)		
	Initial Path	STOMP on Initial	Local Search
(1) SH-1	11.03	1.83	3.57
(2) SH-2	9.15	1.36	3.05
(3) A2PH	8.90	1.96	3.91
(4) R-1	10.97	10.02	1.93
(5) R-2	6.01	3.67	6.60
(6) R-3	23.47	9.50	7.93
(7) R-4	1.68	0.34	1.38
(8) R-5	14.41	0.87	13.43
(9) R-6	14.26	3.26	11.65
(10) R-7	18.08	8.69	14.98

approaches, the maximum improvement for STOMP on local search is over the initial trajectory (23.47 % for R-3), followed by over local search alone (14.98 % for R-7), and finally over STOMP on the initial trajectory (10.02 % for R-1).

These results are also illustrated in Figure 4.9, with the local search outputs improving upon the initial linearly interpolated paths, but not consistently outperforming STOMP on the linear interpolation. The local search tended to outperform STOMP on the linear interpolation in cases where the noisy trajectories generated by STOMP were unable to explore the configuration space widely enough. Box plots and kernel density estimates on the data from Table 4.3 show that STOMP with either initialization has a lower variance than the local search output alone, and skews towards lower mean total muscle fiber forces.

Looking at a specific start and goal state pair, we illustrate the effects of each of the trajectory generation approaches. Consider the start state $\boldsymbol{\theta_1} = [3.08, 1.08, 0.43]$ and goal state $\boldsymbol{\theta_N} = [0.32, 1.52, 0.40]$, with $N = 25$ (Trajectory

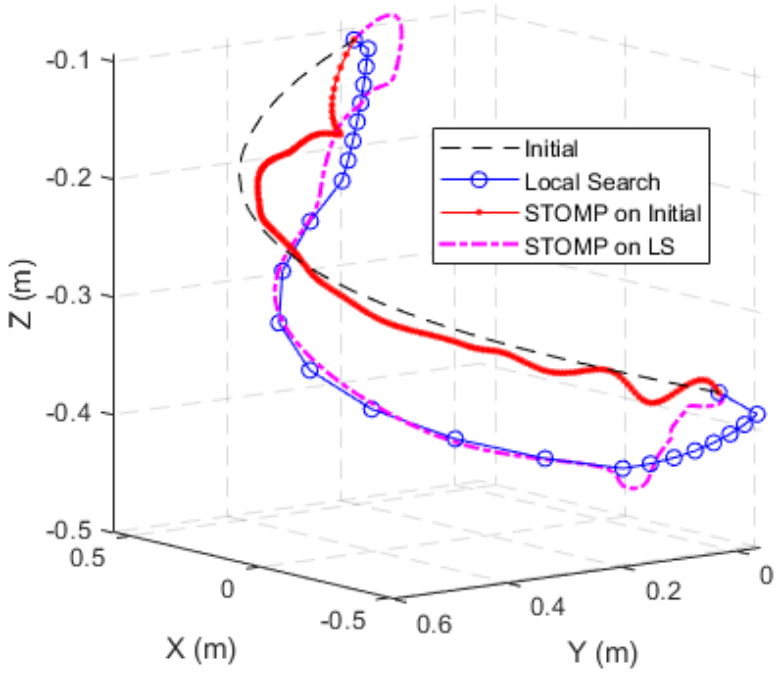
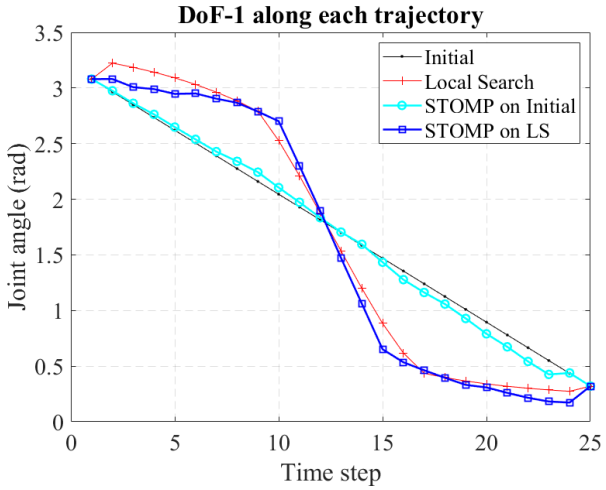


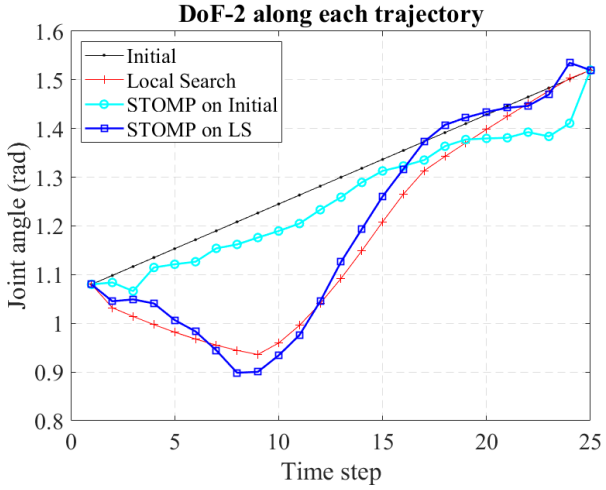
Figure 4.10: End-effector positions in 3D for trajectory R-1 from Table 4.3 showing the initial path (black), STOMP optimized on initial path (red), local search optimum (blue), and STOMP optimized on local search (magenta).

(4) R-1, from Table 4.3). Figure 4.10 shows the 3D Cartesian position of the end-effector along trajectories from each of the four approaches.

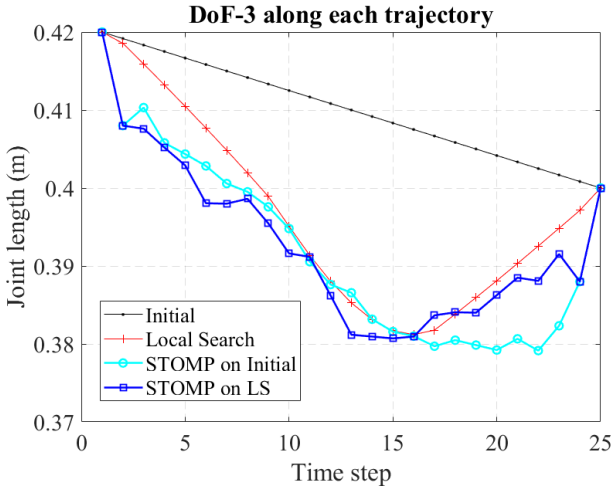
The STOMP cost component q_h, the total muscle fiber force along a trajectory, is shown in Figure 4.12 for each of these approaches, interpolated to 300 steps during CMC simulations. The STOMP optimal trajectory initialized on the linearly interpolated configuration space results in a lower mean total muscle load, but does not explore far enough to reduce it significantly (from ˜6.42 kN to ˜6.36 kN). The local search is able to quickly find a trajectory that pitches lower in the Z-direction than the initial (Figure 4.11b), reducing the WRF's moment arm about the user's elbow. In this case, local search outperforms STOMP on the initial guess, with a result of ˜5.83 kN. The final STOMP iterations performed on the local search result further reduce this load to ˜5.72 kN.



(a)



(b)



(c)

Figure 4.11: DoFs 1–3 along all approaches for trajectory R-1.

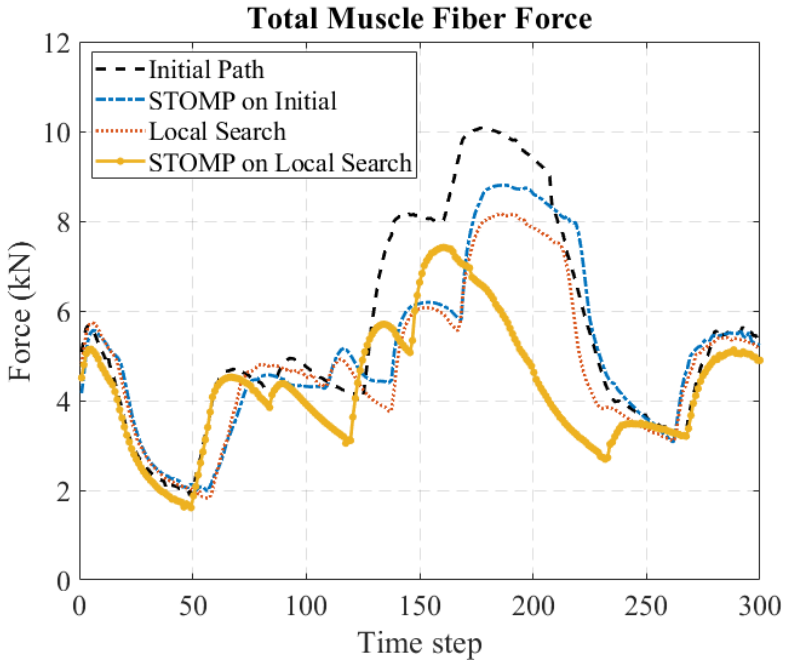


Figure 4.12: Comparison of total muscle fiber force q_h for trajectory R-1.

4.3.1 STOMP Ablation Study

We performed an ablation study to determine the effect that each component of the state-dependent cost $\mathbf{q}(\boldsymbol{\Theta})$ has on the overall performance of STOMP. We computed the total cost, $Q(\boldsymbol{\Theta})$ with STOMP performed using subsets of the terms in $\mathbf{q}(\boldsymbol{\Theta})$, initialized on linearly interpolated configuration-space trajectories $\boldsymbol{\Theta}^0$ for 25 iterations. To compare the effects of the components, the total costs were normalized, dividing them by $Q(\boldsymbol{\Theta}^0)$ to provide the relative improvements in performance, as shown in Figure 4.13.

For all ten start and goal state pairs, we first considered only the effect of the direct reaction load costs q_f and q_m. This resulted in a marginal relative improvement over $Q(\boldsymbol{\Theta}^0)$ (mean reduction by ˜6%) . Adding the load deviation costs q_{df} and q_{dm} improved this performance to a ˜16% mean reduction. Finally, adding the cost term q_h from the muscle fiber forces computed through CMC

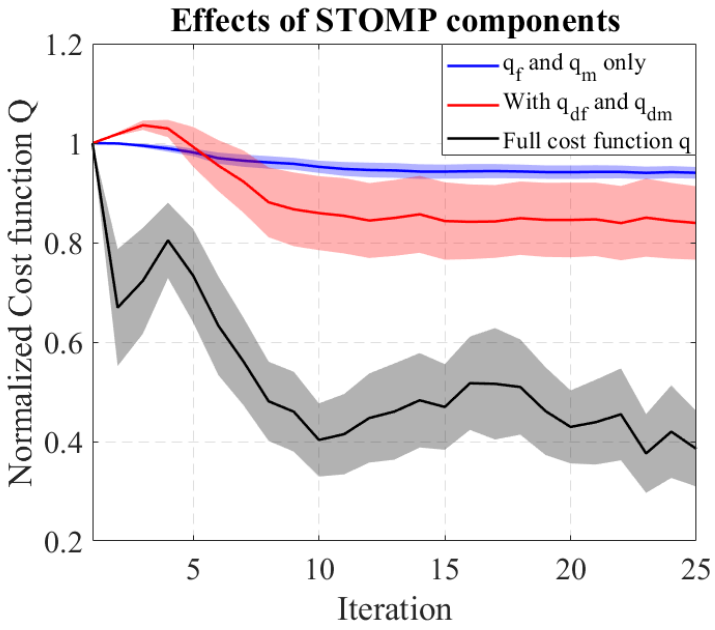


Figure 4.13: Mean normalized STOMP cost functions with standard errors for subsets of the cost terms (red, blue), and for the full cost function (black).

resulted in a ˜61% mean reduction over $Q(\mathbf{\Theta^0})$.

Including all of the components in the cost function for STOMP helped generate trajectories with reduced biomechanical loads even with a relatively small number of iterations.

4.4 Conclusion

We presented a simulation framework based on OpenSim for determining the effect of a supernumerary wearable robot's motion on the user, and finding trajectories that minimize human muscle loads. Aside from motion planning, this framework can also be used for improving the design of a wearable robot, as in [11], by accounting for biomechanical effects.

We described a STOMP-based approach for finding optimal trajectories that include human muscle forces determined by CMC in their cost. They were initialized with paths that linearly interpolate between the start and goal states in the configuration space, as well as with a locally-optimal paths, consistently finding trajectories that reduce human muscle loads.

While these are promising results, our analysis included key assumptions:

- Keeping the speeds of the motors constant simplified the dynamics analysis and allowed the local search to be performed rapidly. However, this would prevent dynamic obstacle avoidance.
- We assumed that the human keeps their arm static as the robot moves. While local search is independent of the human's motion aside from kinematic constraints, this assumption sped up the OpenSim CMC simulations.
- The assumption of absence of collisions between the human and robot was justified based on the human's static pose considered here, chosen such that most WRF motions remained away from the human, resulting in a simplified STOMP algorithm, which would otherwise require an explicit obstacle cost as described in [59].

As the STOMP iterations with CMC simulations are far from real-time, requiring several hours to compute on typical hardware associated with robotic arms, approximate methods are needed for addressing some of the above assumptions and deploying this framework on a physical system. The local search, which can be optimized for real-time performance, would be viable in applications where a coarse level of improvement over an initial guess is accept-

able. One approach for further incorporating human muscle loads could be a supervised learning model, trained on a large database of combined human and robot motions, and used to augment a local search-based planner, or one of the more common sampling-based planners.

As further limitations, these results were obtained from numerical simulations that do not capture the full objective and subjective experience of wearing the WRF, which includes factors such as shear on the human skin from the mounting interface, and jerks due to accelerations of the motors. We also do not account for musculoskeletal fatigue after wearing the device for extended periods of time, an aspect of interaction that was often cited by study participants to be a limiting factor in its usability. Within the simulations themselves, the improvement afforded by STOMP over the initial linear interpolation or local search output, while consistent, did not have large effects for every start and goal state considered here.

Despite these limitations, the framework presented in this chapter could assist in the design, biomechanical analysis, and motion planning for SR devices, allowing them to be effective agents in human-robot collaboration tasks.

Having analyzed the effects of the WRF on the human's arm, in the next chapter we look at the reverse side of this two-way interaction, and study how the user affects the WRF's motion. We discuss the WRF's control systems through linear analysis, and apply human motion prediction to improve the feedback control of the end-effector when subject to disturbances due to the user's arm movements.

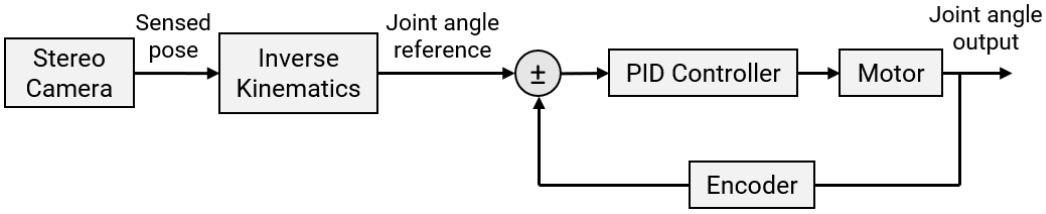
CHAPTER 5

STABILIZATION AND CONTROL

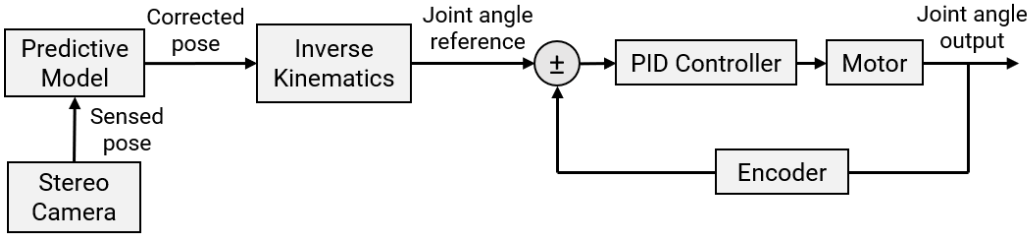
In order for the WRF to be an effective augmentation, it needs to be able to counteract the disturbances introduced in its motion plan due to the user's independent activities. In this chapter, we present a method for stabilizing the end-effector of the WRF, which is subject to disturbances arising from movements of the user's arm to which it is attached.

We initially consider stabilizing the end-effector in the simplified case of small, planar movements using time series forecasting of human arm motion. This approach is extended to include a recurrent neural network (RNN) model for human motion prediction, and applied for 3-D stabilization in five common tasks such as wiping a desk and stowing items into drawers. The goal in all of these approaches is to keep the WRF's end-effector static at an initial position.

The challenge of wearable arm compensation and stabilization can be viewed as a PID control problem. The direct feedback control strategy is outlined in Figure 5.1a. The joint angle reference signals for each motor are determined from the poses of the human and WRF detected by an optical motion capture system, as well as the desired pose of the end-effector. However, the performance of this approach is affected by delays in sensing and actuation, where the reception of joint angle reference signals lags the real state of the system. This drop in performance can be accounted for, even with stock motors, by adding a predictive model of human motion before set point generation (Figure 5.1b).



(a) Overview of direct feedback control.



(b) Feedback control with prediction model incorporated.

Figure 5.1: The predictive models generate motor joint angle references over a finite horizon. We compared the end-effector stabilization (b) without human motion prediction, and (c) with human motion prediction.

5.1 Related Work

This challenge of stabilizing a wearable robotic arm is somewhat similar to that of manipulators rigidly mounted onto a mobile base. In typical mobile manipulators, however, both robotic components are controllable and as a result external disturbances from loads on either of them can be compensated for by generating combined cooperative motion plans for a specified task [99]. Force-based approaches, such as potential functions tracking the projection of the robot's equilibrium point on the ground (Zero Moment Point) [53], and frequency-domain methods for vibration suppression [10] have been used to stabilize the mobile base and manipulator end-effector. A PID feedback controller with exponential position error weighting for the manipulator arm has been applied in scenarios where the mobile base is considered to be an independent system [90].

In our case, since the wearable robot is attached to a human arm which has structured but uncontrollable movements, the stabilization of the end-effector's position needs to be performed by the manipulator alone.

This situation bears some similarity to the domain of hand-held robotic surgery devices, where physiological hand tremors have to be compensated for. One approach estimates tremors online by modeling them as truncated Fourier series, which are fed as an input to an open-loop controller for tool-tip position compensation [92]. Another approach involves modeling hand tremors as linear dynamic systems, with Kalman filtering for state estimation [18]. The hand tremor velocity estimate is considered to be a load disturbance, leading to a feed-forward control term added to a standard PID feedback controller. In our case, the actuators do no have feed-forward control capability, motivating a set-point prediction strategy to compensate for latency in sensing and actuation.

The most closely related work was presented in backpack mounted Supernumerary Robotic Limbs (SRL) involving human disturbance rejection through bracing [86]. Bracing involves the SRL grasping a structure in the environment to stabilize its user in a standing position. Human disturbances, in the form of postural sway, are modeled as external forces and torques acting on the SRL's base. The pose estimation of the SRL base is improved using a bracing strategy generated by shaping the stiffness of the manipulator. Another scenario involves a force sensing-based drilling task in which the SRL holds a wooden piece steady while the wearer drills into it [71]. This approach used human-human demonstrations of the task, where the effects of the robot's motions were condensed into force inputs, and the humans's motions into tool-tip positions. An autoregressive moving-average process with exogenous inputs (ARMAX)

model was used to predict the force required by the SRL to hold the workpiece stable during the task. Our scenario involves human arm movements on a similar scale as the leader's hand motion with the SRL. However, instead of applying grasping forces to a static workpiece or the environment, the end-effector is stabilized in free space. We adopted a hybrid approach, combining prediction models (autoregressive time series and recurrent neural network) for the motion of the human arm, with its effects on the end-effector compensated through feedback control.

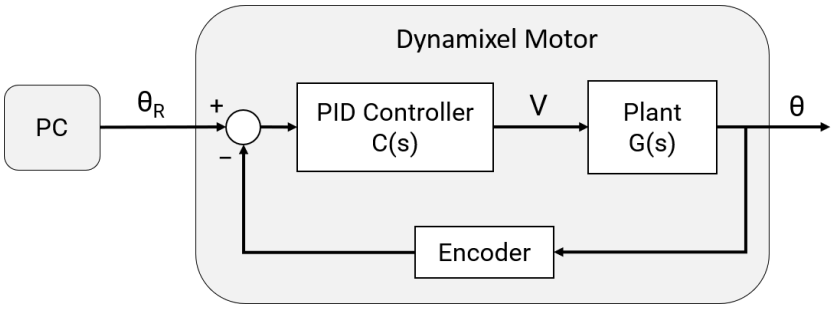
5.2 System Identification of Motors

In order to better understand the WRF's assistive capabilities and develop appropriate controllers, we first performed system identification tests to recover the in-situ motor parameters.

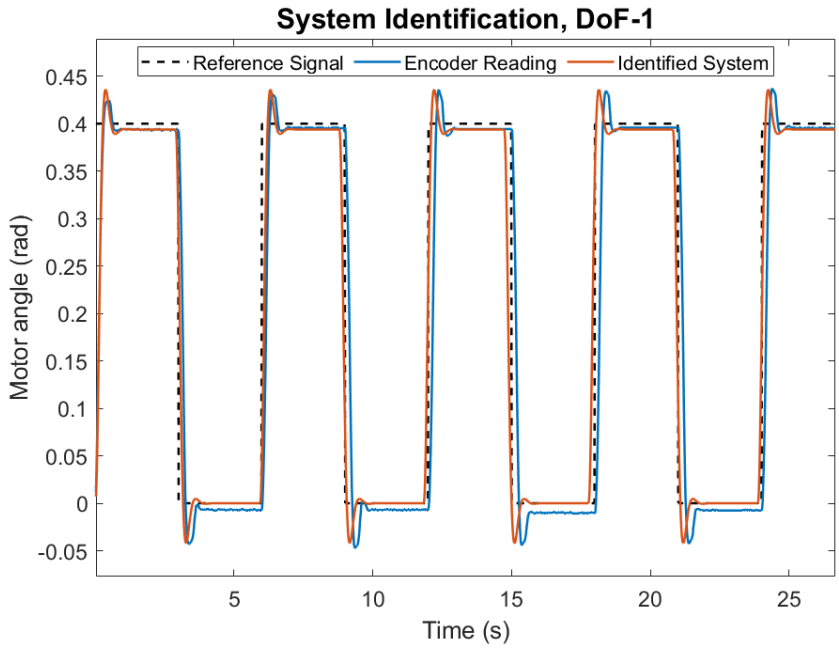
Each of the Dynamixel motors used in the WRF have built-in PID controllers, apart from the AX-12A motor for wrist rotation that only has P control. Each motor receives a reference angle θ_R as input from the PC, driving a DC motor plant, with output angle θ measured using built-in encoders (Figure 5.2a).

The plant transfer function $G(s)$ between voltage V and output angle θ is based on an L-R circuit DC motor model [22], resulting in a third order system in terms of parameters α_0, γ_0 and γ_1:

$$G(s) = \frac{\theta(s)}{V(s)} = \frac{\alpha_0}{s^3 + \gamma_0 s^2 + \gamma_1 s} \tag{5.1}$$



(a)



(b)

Figure 5.2: (a) Control system schematic for the Dynamixel motors used in the WRF. (b) The responses of identified system models (green) are compared with measured motor responses (black) to the same reference signal (blue), shown here for the horizontal panning motor.

$$C(s) = K_p + \frac{K_i}{s} + K_d s \tag{5.2}$$

During system identification, the PID controller's transfer function $C(s)$ uses manufacturer supplied values for the gains $K_p = 4$, $K_i = 0$, and $K_d = 0$

The closed-loop transfer function $P(s)$ between the motor output angle θ and reference signal θ_R is a third-order system with no zeros:

$$P(s) = \frac{\theta(s)}{\theta_R(s)} = \frac{G(s)C(s)}{1 + G(s)C(s)} = \frac{B_0}{A_0 + A_1 s + A_2 s^2 + s^3} \tag{5.3}$$

We fit the closed-loop model parameters A_{0-2} and B_0 to the measured output signals using the Simplified Refined Instrumental Variable method for Continuous-time model identification (SRIVC) [117, 120]. The plant parameters α_0, γ_0 and γ_1 were then obtained from A_{0-2} and B_0. Each DoF was identified individually, keeping all other motors fixed, and the magnitudes of the step reference input signals were determined from the usage scenarios (e.g. steps of 0.7 rad over 2 seconds for DoF-1 as shown in Figure 5.2b).

The accuracy of the identified system models was evaluated by computing the Normalized Root Mean Squared Error (NRMSE) goodness of fit between the output signals measured by the encoders and the simulated motor model outputs, for the same reference input signals. The plant parameters and model fitting metrics for each DoF are listed in Table 5.1.

Table 5.1: Identified motor model parameters

DoF	Motor	α_0	γ_0	γ_1	Fit (1-NRMSE)
1) Horizontal Panning	MX-64T	1078.49	572.86	48.38	0.961
2) Vertical Pitching	MX-64T	463.53	408.31	31.31	0.909
3) Length Extension	MX-28T	1134.06	674.49	29.88	0.899
4) Wrist Rotation	AX-12A	1098.43	733.68	98.76	0.919
5) Wrist Pitching	AX-12A	4041.57	1439.19	64.49	0.869
6) Gripper	MX-28T	173.35	1633.97	272.87	0.926

5.3 Planar End-Effector Stabilization

In this section, we consider a planar collaborative usage scenario (Figure 5.3), where the wearable robotic arm would retrieve workpieces and bring them to the user's workspace, assist the user in an assembly task, and stow workpieces at a desired location. The robotic arm would grasp objects that are either handed over by the human in free space or under-constrained on a surface, preventing the adoption of a bracing strategy that requires a rigid connection with the environment. Once the object is grasped, if the user were to then perform other tasks with their hands, it would lead to disturbances in the position of the end-effector, and consequently the object. We aim to reduce these disturbance, leading to more stable grasp for objects such as cup filled with a liquid, or a camera.

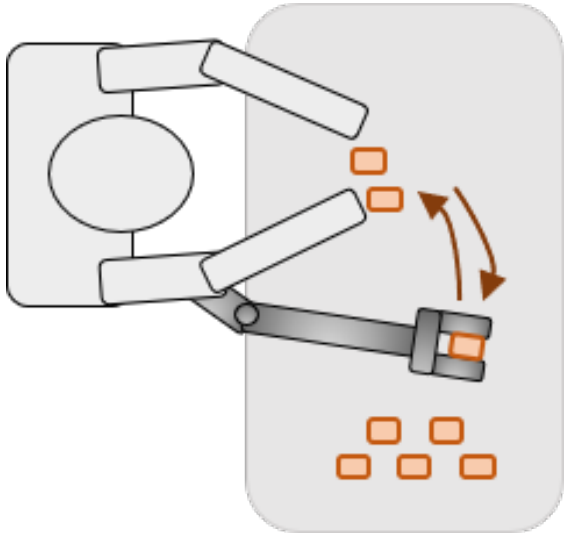


Figure 5.3: Planar collaborative usage scenario.

In order to study the stabilization of the end-effector in the planar (2D) domain, we used the setup shown in Figure 5.4. AprilTag fiducial markers [81] were placed on the robot's end-effector, and above the axis of DoF-1 (robot base) on the attachment point between the robot and the human arm. They were

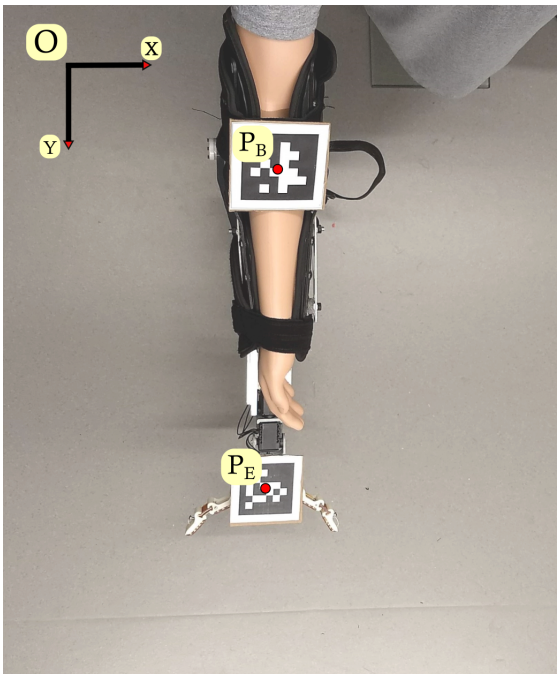


Figure 5.4: Planar system setup: the robot's base and end-effector positions are tracked using visual fiducial markers sensed by a stereo camera.

sensed with an Orbbec Astra stereo camera for vision-based position tracking of these two points. Figure 5.4 shows the camera coordinate system $O = (x, y)$, with base coordinates $P_B = (x_B, y_B)$, and end-effector coordinates $P_E = (x_E, y_E)$. Given a detected deviation in the base position $(\Delta x_B, \Delta y_B)$, the new joint reference angle Θ_1 for DoF 1, and length reference L_3 for DoF-3, are computed using the 2-D Inverse Kinematics (IK) equation:

$$\begin{bmatrix} \Theta_1 \\ L_3 \end{bmatrix} = \begin{bmatrix} tan^{-1}((y_E - y'_B)/(x_E - x'_B) \\ \sqrt{(x_E - x'_B)^2 + (y_E - y'_B)^2} \end{bmatrix} \tag{5.4}$$

where $(x'_B, y'_B) = (x_B + \Delta x_B, y_B + \Delta y_B)$.

In the WRF, length extension is achieved with a rack-and-pinion mechanism, so that L_3 is linearly related to Θ_3, the motor joint angle for DoF-3. Given the

current robot morphology, deviations in L_3 can be compensated to within ˜12 cm. This also served as a decision variable for whether or not a compensatory motion should be attempted.

Joint angle set points for each motor are generated from the fiducial marker positions after applying equation (5.4), which were then tracked by built-in closed loop controllers in the servo motor hardware. Deviations in x and y coordinates of the end-effector were considered separately, since for the initial pose shown in Figure 5.4, the x coordinate was predominantly affected by DoF-1, and y coordinate by DoF-3, and there was a marked difference in the performance characteristics of the two motors.

Table 5.2: Motor step response characteristics

Parameter	DoF 1	DoF 3
Rise time (s)	0.133	0.294
Settling time (s)	0.539	0.489
Overshoot (%)	10.79	0.487
Peak time (s)	0.293	0.650
Bandwidth (Hz)	2.726	1.174

Following the system identification procedure described earlier, the step response characteristics for these motors were estimated (Table 5.2). The motors had reasonably fast responses, with bandwidths of ˜2.7 Hz and ˜1.2 Hz respectively, and settling within ˜0.5 s to the maximum joint angle deviations likely to be encountered in our scenario. The difference in performance was due to a higher-end Dynamixel MX-64 motor used for DoF-1, compared to a lower power MX-28 motor for DoF-3.

5.3.1 Human Arm Motion Data

Given the characterization of the motors, we now turn to the human arm motions that need to be compensated for, analyzing the spectral properties of a sample motion dataset.

Using the setup shown in Figure 5.4, a dataset of N = 3604 points (x_B, y_B) was collected, of a wearer moving their arm about a mean position within an amplitude of ~15 cm, with the motors of the robotic arm held static. The camera in our setup had a frame rate of ~26 Hz. Figure 5.5 shows an illustrative data sample.

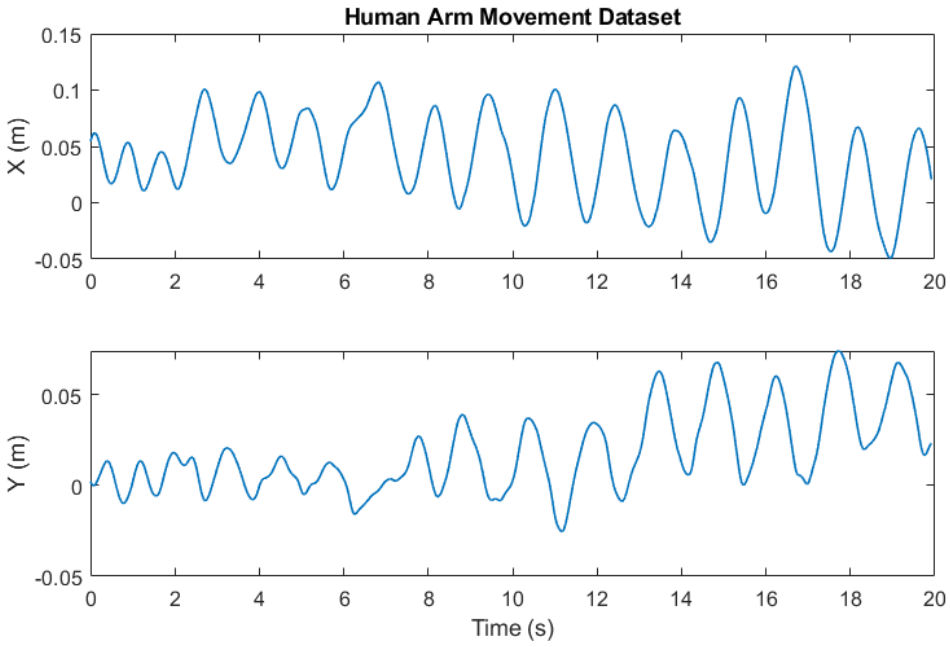


Figure 5.5: Illustrative plots of the planar human arm movement dataset: most motions were restricted to within ~15 cm from the starting position.

The Fourier transforms of the x and y coordinates from the dataset show that the human arm motion while wearing the robotic arm was composed of frequencies largely below 1 Hz (Figure 5.6). The bandwidth of the robot's actu-

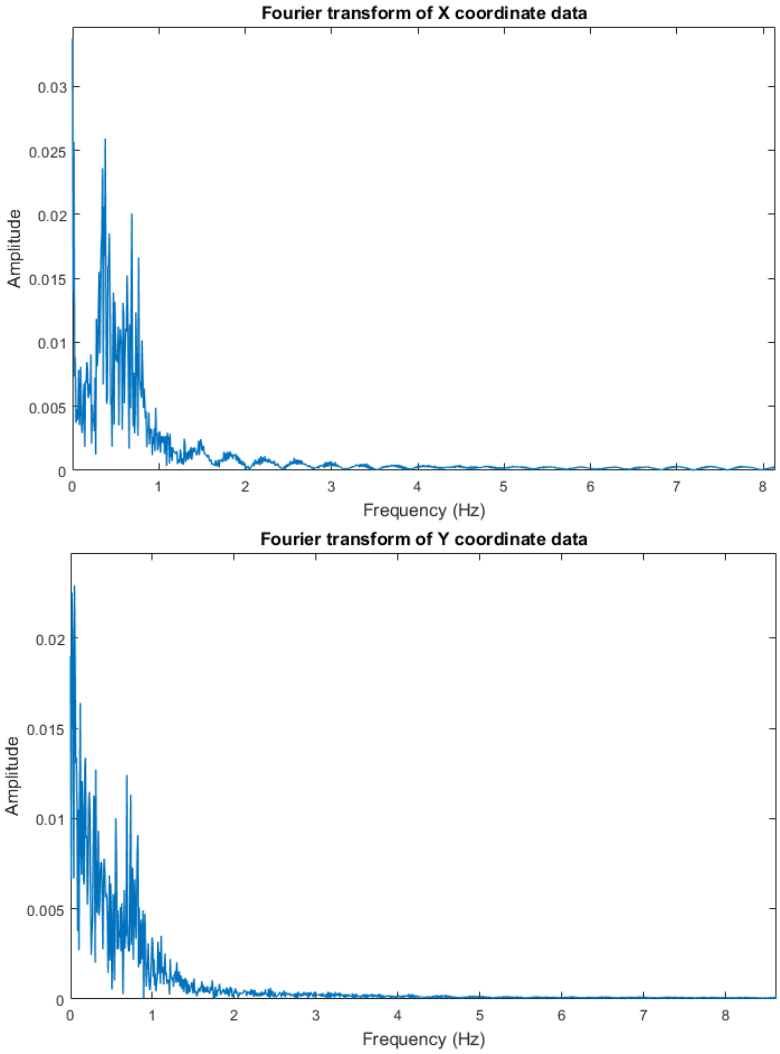


Figure 5.6: Fourier transform plots of robot base position (x_B, y_B) data.

ators is greater than 1 Hz (Table 5.2). This indicates that the actuators should be able to compensate for most of the disturbances introduced at the robot's base by the human arm in a feedback control system as shown in Figure 5.1a. The effect of physiological tremors with frequencies of 10 Hz and higher is negligible, since their typical amplitudes are of the order of 0.1 mm [92].

5.3.2 Delay Estimation

While the above characterization of the planar human movement and motor properties suggests that a feedback control system would be able to compensate for human-generated disturbances, in practice the system had a mean error of about 5.9 cm in x and 3.7 cm in y using feedback control alone (Table 5.3).

This performance degradation is affected by delays stemming from a combination of latency in communication with the motors (rate of ˜50 Hz), vision tracking system (rate of ˜26 Hz), mechanical transmission delays, as well as nonlinear effects in the motors due to Coulomb and viscous friction [98]. Improving stabilization performance requires a predictive approach to account for these delays.

The aggregate effect of these delays can be identified using a closed-loop step response procedure. Since the motor response for DoF-3 has higher rise time and settling time, as well as lower bandwidth (Table 5.2), it will act as the limiting factor in terms of time delay. Therefore the third-order system for DoF-3 was augmented with a time delay term:

$$P(s) = \frac{\theta(s)}{\theta_R(s)} = \frac{e^{-\tau_d s} B_0}{A_0 + A_1 s + A_2 s^2 + s^3} \tag{5.5}$$

The step response was determined by tracking the visual markers for the same reference inputs given in Figure 5.7, and computing the joint angle Θ_3 by applying the IK equation (5.4). This allowed for the estimation of the time delay τ_d between the motor's internal response and detection of the same movement by the camera. As before, the SRIVC system identification procedure was applied to step response tests for DoF-3 (Figure 5.7), and the delay τ_d was found to

be 76.48 ms, with an NRMSE model fit of 0.8479. Compensating for this delay required a prediction of two time steps into the future for a sampling rate of ˜26 Hz of the camera (sampling period of ˜38.5 ms).

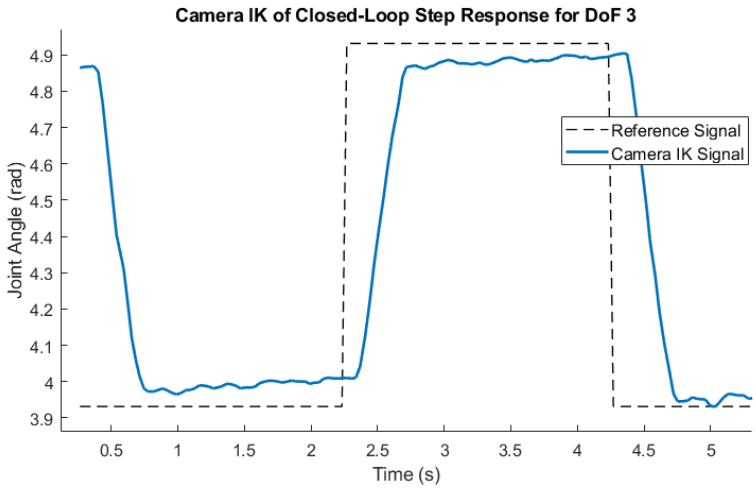


Figure 5.7: Closed-loop step response of DoF-3 measured by applying IK equations to data from the stereo camera. The delay τ_d is estimated by fitting a linear model to this data.

5.3.3 Autoregressive Model

After estimating the delay present in our system, we aimed to improve end-effector stabilization performance by predicting the robot's base position movements over the time horizon equal to the delay. We did so by learning an autoregressive predictive model for the human arm motion.

As a first assumption, similar to [71], the deviations in the x and y coordinates of the human arm movement were modeled as discrete univariate time series composed of a combination of Auto Regressive (AR) and Moving Average (MA)

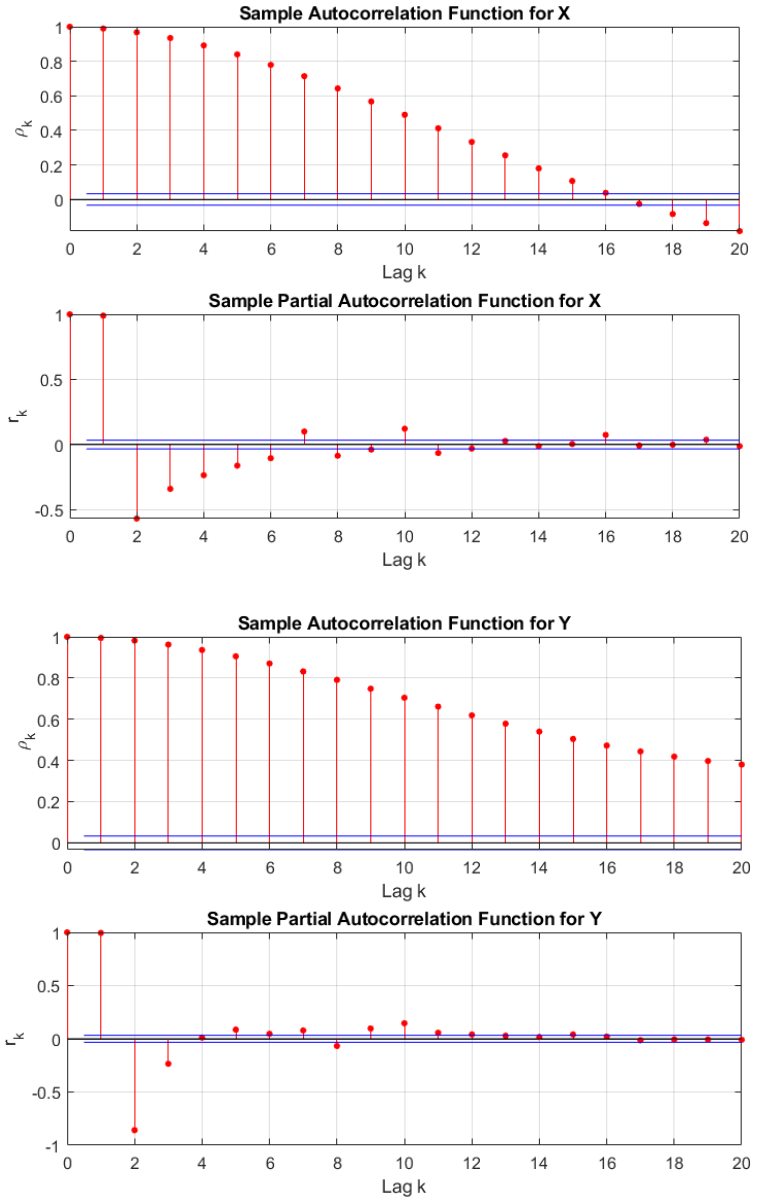


Figure 5.8: Sample autocorrelation and partial autocorrelation functions, ρ_k and r_k, for x and y coordinate data with lags $k \in [1, 20]$ and confidence bounds of twice the standard error. r_k drops off much faster than ρ_k, indicating a predominance of AR terms.

terms ("ARMA"), with no exogenous inputs:

$$x_t = C + \sum_{i=1}^{p} A_i x_{t-i} + \sum_{j=1}^{q} B_j \epsilon_{t-j} \tag{5.6}$$

Here x_t is either the x or y position coordinate error, ϵ_t is a white noise series with zero mean and σ^2 variance, assumed to be normally distributed. Past coordinates are included in this model up to p time steps, which is the order of the autoregression (AR) part, with coefficients A_i. Past white noise terms in the moving average (MA) part are included up to an order q, with coefficients B_j. The constant term C is assumed to be zero, since deviations are measured from an initial position. This model makes the assumption that x_t is stationary, i.e. its expected value and variance are independent of time. This property was verified to hold true for our dataset using the augmented Dickey-Fuller hypothesis test [36].

In the next step, we investigated which of the terms, AR or MA, were dominant in the data. The dominance of either term would indicate that a simplified model could be adopted instead of a full ARMA model. For this, we compared the sample autocorrelation function, ρ_k with the sample partial autocorrelation function, r_k. ρ_k is the correlation between sample points x_t and x_{t-k}:

$$\rho_k = Corr(x_t, x_{t-k}) = \frac{Cov(x_t, x_{t-k})}{Var(x_t)} \tag{5.7}$$

The partial autocorrelation r_k measures the correlation between x_t and x_{t-k} after discounting for the effects of all intermediate lags $x_{t-1}, ..., x_{t-k+1}$. This measures the direct dependence between a data point and its lagged value, as opposed to ρ_k, which encodes the dependence of intermediate terms as well.

$$r_k = Corr(x_t - \hat{x}_t, x_{t-k} - \hat{x}_{t-k}) \tag{5.8}$$

Here $\hat{x}_t$ and $\hat{x}_{t-k}$ are the best linear projections of x_t and x_{t-k} onto the interme-

diate terms using least-squares regression.

Figure 5.8 shows that ρ_k decays at a much lower rate than r_k for both coordinates, indicating that a data point x_t depends more heavily on past points x_{t-k} rather than past white noise terms ϵ_{t-k} [23]. This allowed for the simplification of the ARMA(p,q) model into a purely autoregressive AR(p) model:

$$x_t = \sum_{i=1}^{p} A_i x_{t-i} + \epsilon_t \tag{5.9}$$

To find the order p of this AR(p) model, we applied the Bayesian Information Criteria (BIC), which optimizes a log-likelihood goodness of fit while penalizing more complex models [23]:

$$BIC = -2log(\mathrm{L}) + (p+1)log(N) \tag{5.10}$$

Here L is the likelihood function, p is the order of the AR process, and N is the number of samples. The order $p = 11$ is found to have minimum BIC for both coordinates. The model parameters A_i and white noise variance σ^2 for these AR(p=11) processes were estimated using the Yule-Walker method [42].

This model was tested on a validation dataset of N=1050 points collected in the same setup (Figure 5.4), for predictions k=2 time steps into the future, which should compensate for system delays of ˜77 ms. This two-step prediction matched the ground truth data well (Figure 5.9), having mean absolute errors of 1.53 cm and 0.73 cm respectively in x and y coordinates, with standard deviations of 1.87 cm and 0.89 cm.

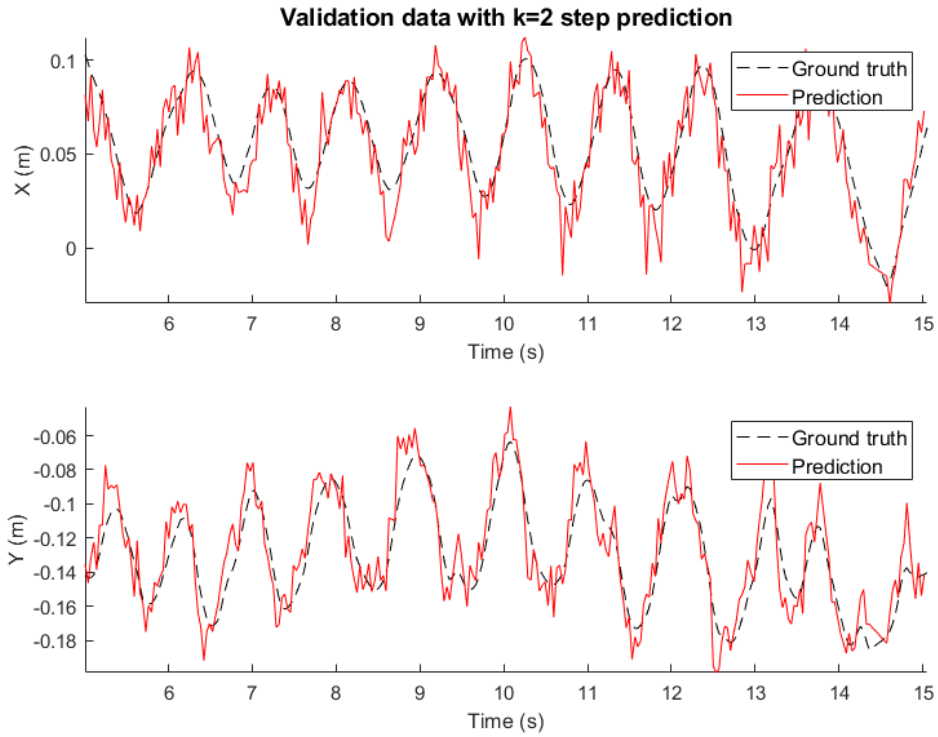


Figure 5.9: Sample of human arm motion prediction data (red), aligned with ground truth (black).

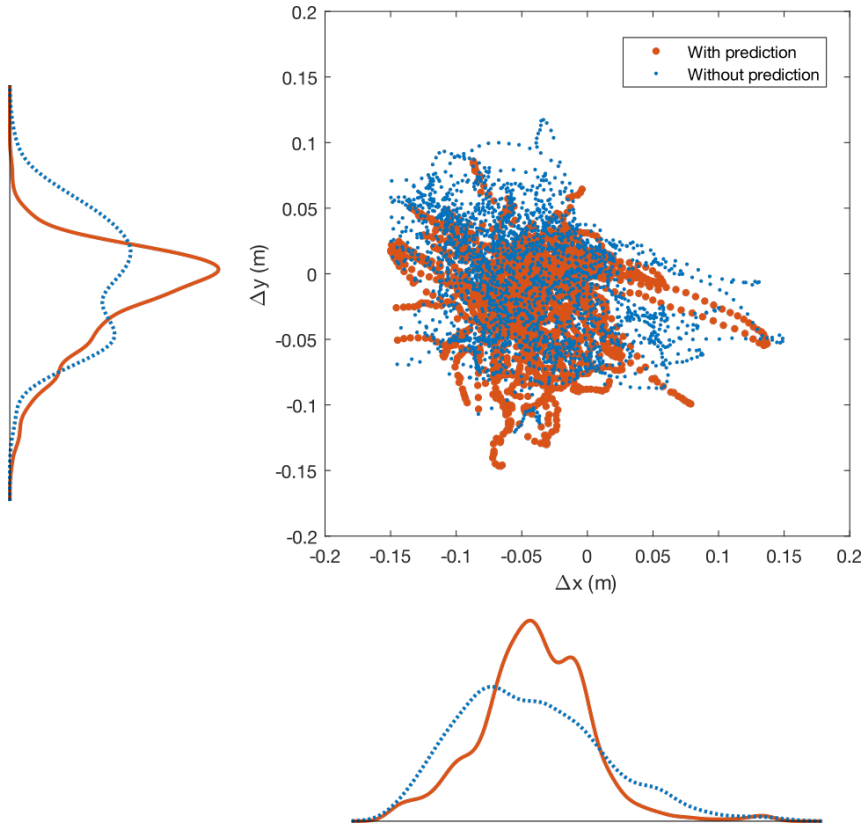
5.3.4 Results

Having obtained an autoregressive predictive model for planar human arm motion, its performance was compared to the case without prediction when applied to end-effector stabilization, as per the approaches shown in Figure 5.1.

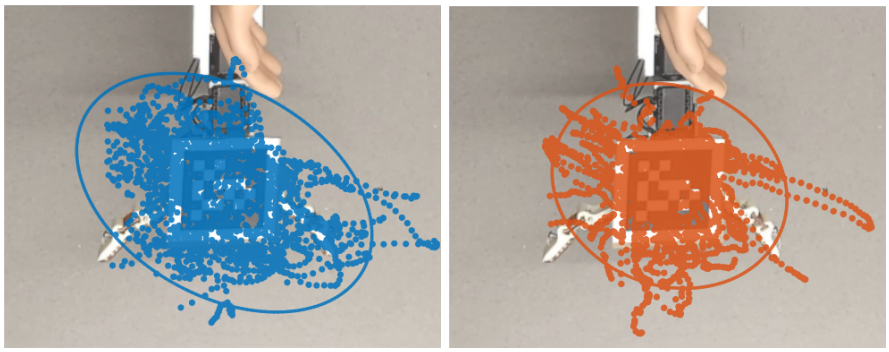
Table 5.3: Deviation of end-effector with and without prediction in 2D

Quantity	Mean (cm)	Std. Dev. (cm)
$\|\Delta x\|_{nopred}$	5.921	3.732
$\|\Delta x\|_{pred}$	4.765	3.398
$\|\Delta y\|_{nopred}$	3.706	2.402
$\|\Delta y\|_{pred}$	2.959	2.768

Both these strategies were implemented independently, collecting N=2312 data points without prediction, and N=2750 points with the AR predictive model in use. The resulting scatter plots of end-effector pose along with their



(a) Scatter plots and Gaussian kernel density estimates of 2D end-effector position error



(b) Without prediction (c) With prediction

Figure 5.10: Comparison of planar end-effector position errors with (orange) and without (blue) the predictive AR model. (a) The predictive model results in lower mean position error in x and y coordinates; (b), (c) show the individual error scatter plots with 95% confidence regions.

kernel density estimates are shown in Figure 5.10. Table 5.3 reports mean and standard deviations of the absolute errors in displacement from the starting pose with only feedback control ($|\Delta x|_{nopred}$ and $|\Delta y|_{nopred}$), and feedback control along with the predictive AR model ($|\Delta x|_{pred}$ and $|\Delta y|_{pred}$). Position errors were reduced by about 1.15 cm (19.4%) in the x coordinate and 0.75 cm (20.1%) in the y coordinate.

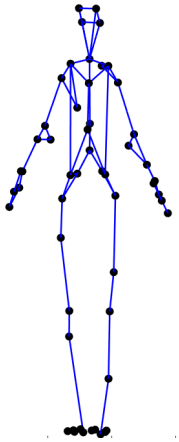
5.4 End-Effector Stabilization in 3D

Having obtained improvements in end-effector stabilization in the planar setup, we extend this approach to more general tasks in full 3D space. As before, linear system identification techniques were applied to estimate the delays and determine the prediction horizon for human motion.

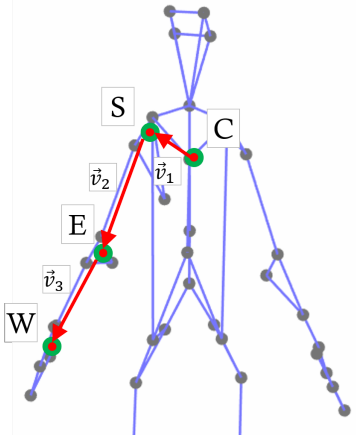
Two approaches were considered for predicting human motion over this horizon: an autoregressive (AR) time series model as before, and a recurrent neural network (RNN) model adapted from [76]. These models take in pose data for the WRF and human, and generate a sequence of joint angle references over a specified time horizon. Owing to the requirement of large datasets for RNNs, both of these approaches were trained offline using the KIT Whole-Body Human Motion Database [75] and adapted for online predictive control through the framework shown in Figure 5.1b. The performance of these offline models transferred well to the real 3D tasks, which were slightly different from the training examples in the database. Common training data also allowed for a more direct comparison between the two approaches.

5.4.1 Human Motion Prediction

The criteria for human motion prediction models for end-effector stabilization were real-time (or close to real-time) prediction with optical motion capture data, and good performance over the required controller time horizon in close-range tasks.



(a) Marker set for the entire skeleton.



(b) Relevant positions for human motion prediction.

Figure 5.11: Of the full set of (a) 56 markers from the KIT Whole-Body Human Motion Database, (b) we extract the relevant points on the human right arm—C: clavicle, S: shoulder, E: elbow, W: wrist.

Table 5.4: Relevant task motions from KIT Whole-Body Motion Database

Task	No. of trials	Total no. of data points
1) Drying right arm	15	7918
2) Washing right arm	6	4034
3) Taking a book from a shelf	6	825
4) Opening pants	6	672
5) Bringing hand to mouth	5	586
6) Pouring from a cup	7	2758
7) Wiping a table	6	2411
8) Stirring a bowl	7	2157
9) Pouring from a bottle	5	1421

Two predictive models were developed: an autoregressive (AR) time series

as before, and a single-layered gated recurrent unit (GRU) adapted from [76] and modified for real-time performance.

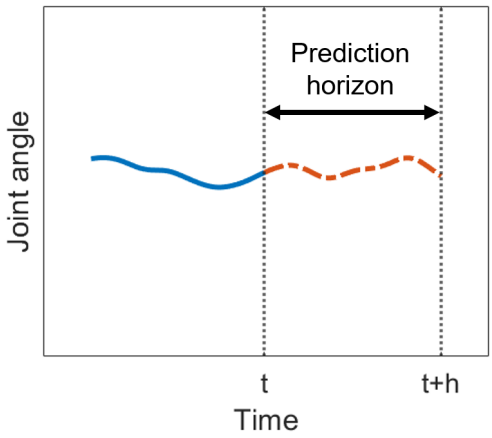
Both of these models were trained offline using the KIT Whole-Body Human Motion Database [75], available at [3]. It consists of a wide selection of task and motion scenarios, with annotated recordings from optical motion capture systems, raw video, as well as auxiliary sensors (e.g. force plates). For this work, we utilized labeled human skeleton marker data (Figure 5.11) from nine tasks in the database that involved periodic movement of the right arm. They are listed in Table 5.4 along with the number of trials performed for each task, and the total number of data points with human right arm movements extracted from all trials.

The full-body skeleton marker set consists of 56 points, out of which 10 are relevant for prediction of human right arm motion, with the positions on the body determined by a weighted sum of the individual 3-D positions of the markers (Figure 5.11b): 3 for the clavicle (C), 3 for the shoulder (S), 3 for the elbow (E), and 4 for the wrist (W).

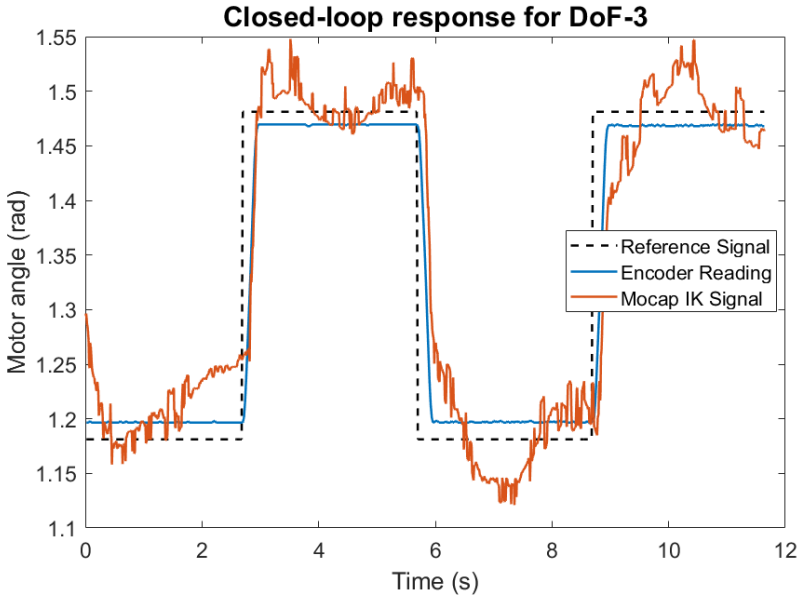
We generated three relative position vectors from the four body points: $\vec{v}_1 = \vec{CS}$, $\vec{v}_2 = \vec{SE}$, and $\vec{v}_3 = \vec{EW}$. This allowed for prediction of movements of a particular body segment independent of its previous neighbor, and improved the training accuracy of the models.

Delay Estimation

The first step in developing these models was to estimate the time horizon for predictions over which the WRF's motors need to be controlled to compensate



(a) Joint angle references need to be predicted over a finite horizon h.



(b) Response of DoF-3 measured using motor encoders (blue) and IK from optical tracking (orange) to a reference step signal (black).

Figure 5.12: Estimating sensing and actuation delays for determining the time horizon for human motion prediction.

for sensing and actuation delays. As earlier, the time horizon h (Figure 5.12a) was determined by system identification of a linear model with a delay of τ_d

included:

$$\hat{P}(s) = \frac{\hat{\theta}(s)}{\theta_R(s)} = \frac{e^{-\tau_d s} B_0}{A_0 + A_1 s + A_2 s^2 + s^3} \tag{5.11}$$

$\hat{\theta}$ is the motor response to an input step signal θ_R, reconstructed though inverse kinematics using data from the motion capture system (Figure 5.12b). The other parameters in the transfer function, A_{0-2} and B_0, were obtained from the system identification performed earlier.

The delay τ_d was estimated to be ˜86 ms using the same SRIVC method as earlier, averaged across DoFs 1-3 which showed relatively slower responses due to larger loads. This corresponds to a prediction time horizon h of about 10 time steps for the OptiTrack motion capture system with frame rate 120 Hz [5] used for 3D stabilization.

Autoregressive Time Series Model

As before, for the prediction model, we started with the initial assumption of an Autoregressive Moving-Average (ARMA) process:

$$x_t = C + \sum_{k=1}^{p} A_k x_{t-k} + \sum_{j=1}^{q} B_j \epsilon_{t-j} \tag{5.12}$$

Here x_t is a discrete univariate series, composed of a constant term C, past terms x_{t-k} weighted by coefficients A_k for lag k (AR term), and past white noise terms ϵ_{t-j} weighted by the coefficients B_j. The number of past terms, p and q determine the orders of the AR and MA parts, respectively.

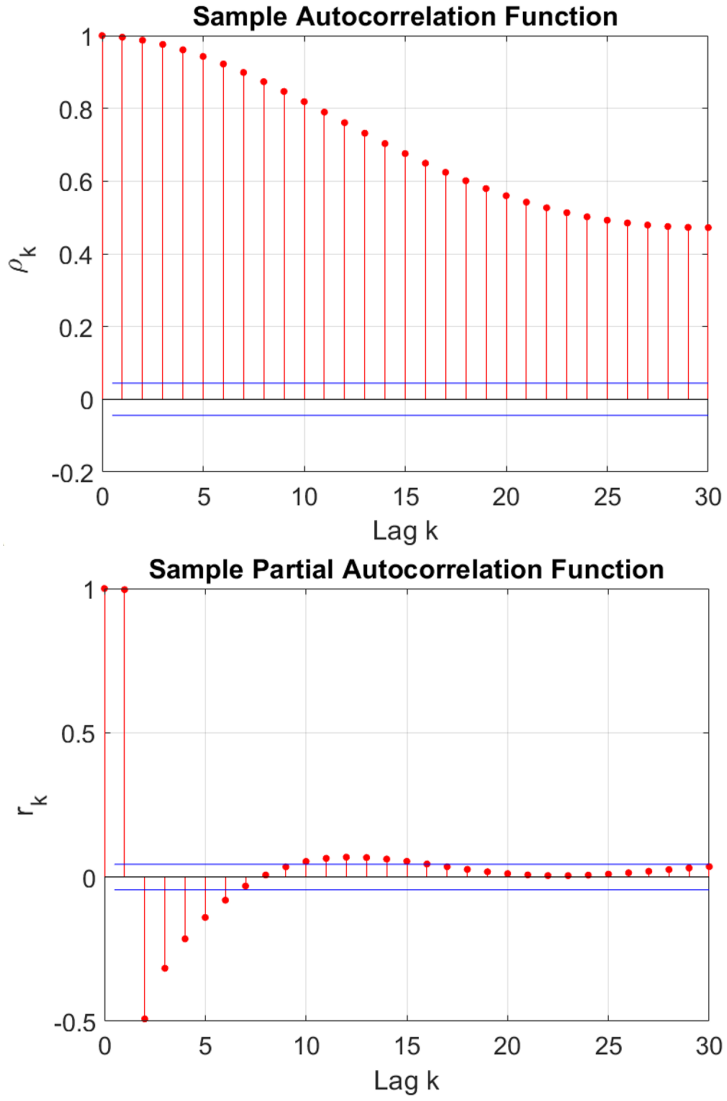


Figure 5.13: Sample autocorrelation and partial autocorrelation functions for the component $v_{2,x}$ from relevant motions in the KIT Database for lags $k \in [1, 30]$.

Each component of the relevant body vectors v_1, v_2 and v_3, was considered to be an independent univariate series. The stationarity of these series was verified with augmented Dickey-Fuller hypothesis tests [36].

As before, we analyzed the autocorrelation (ρ_k) and partial autocorrelation (r_k) functions at lags k for these series. There were sharp drop-offs in r_k com-

pared to ρ_k over successive lags for each component of the body vectors, illustrated Figure 5.13 for the X component of v_2. This indicated that the ARMA processes could be simplified into purely autoregressive (AR) models [23]:

$$x_t = C + \sum_{k=1}^{p} A_k x_{t-k} + \epsilon_t \tag{5.13}$$

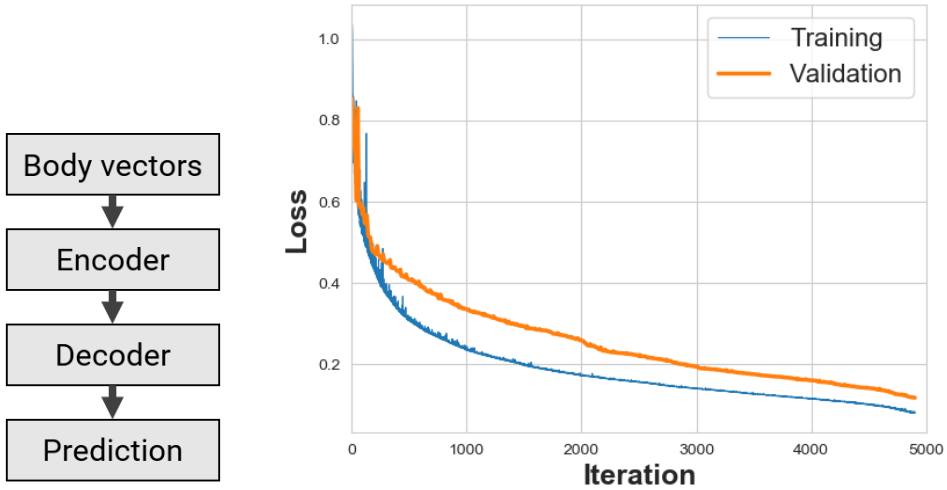
The model order p for each of the nine components in the body vectors was determined using the Akaike Information Criterion (AIC), a maximum-likelihood measure of the goodness of fit [12]. The AIC was computed for model orders up to 30 for each of the nine series, and the one with minimum AIC was selected as p for that series. The minimum AIC values were obtained at different model orders for each series, ranging from p=18 to p=25. The model parameters A_k, C, and ϵ_t were determined using the Yule-Walker method [42], trained on the task motions listed in Table 5.4.

Recurrent Neural Network Model

While an AR model is able to forecast human motions through local predictions, it does not capture dependencies over a longer time period, or encode structural information about the correlations between body components over time. To account for these factors and improve on the predictions from the AR models, we used a recurrent neural network (RNN) model for human arm motion prediction, and compared the performance between the methods.

Independent of robotics, RNNs have been applied extensively for human motion prediction, including architectures with Long-Short Term Memory (LSTM) cells [40], and structural RNNs that encapsulate semantic knowledge

through a spatio-temporal graph [57]. These approaches include multiple recurrent layers as they are aimed at offline prediction of the entire human skeleton, and task classification in general motion scenarios. As the task scenarios for WRF stabilization involve periodic motions and require prediction of only the wearer's arm, we used a simpler model with a sequence-to-sequence architecture [107] and a single gated recurrent unit (GRU), as proposed in [76], which also includes a residual connection for modeling velocities. Compared to an AR model, this resulted in higher prediction accuracy of human arm motion, and improved the end-effector stabilization in most task scenarios.



(a) Structure of RNN model. (b) Training and validation losses for the RNN model.

Figure 5.14: (a) The RNN model consists of an encoder, GRU cell, and spatial decoder. (b) It was trained for 5000 iterations on the KIT Database motions.

The schematic of the RNN model is shown in Figure 5.14a. It consists of an encoder network that takes in a 9-dimensional input of the body vectors, [$\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$], 50 frames at a time from the KIT database or motion capture system, and a decoder network that converts the output from a single GRU cell with 1024 units into 9-dimensional predictions over k steps. Based on the estimated system delay, we set k=10, and the learning rate to be 0.05 for batch sizes of 16,

as specified in [76] for predictions up to 400 ms. This RNN model was trained on the KIT Database motions listed in Table 5.4, and converged at about 5000 iterations, as shown in Figure 5.14b with mean-squared error (MSE) loss.

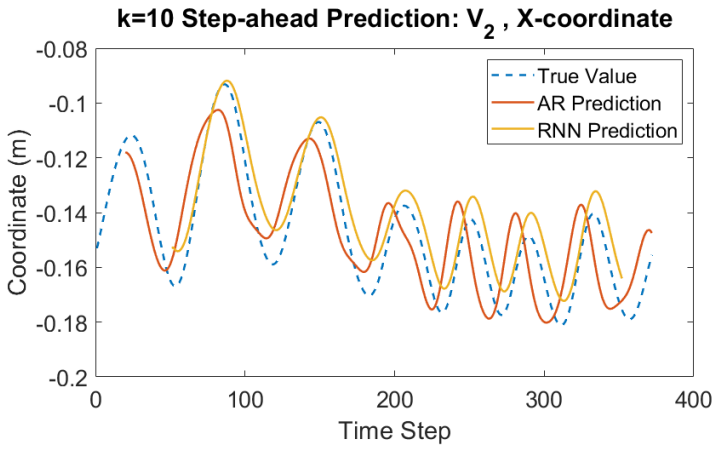
Model Evaluation

Both predictive models were evaluated on the relevant motions from the KIT Database listed in Table 5.4. They were trained offline using all but two trials for each task, with the one of remaining trials serving as the validation set, and the other as the test set. The training set was expanded to four times its original size by adding Gaussian white noise with standard deviation 1 cm to each of the nine components of the body vectors, leading to 89864 data points for training. The test and validation sets had 18922 and 15042 data points respectively.

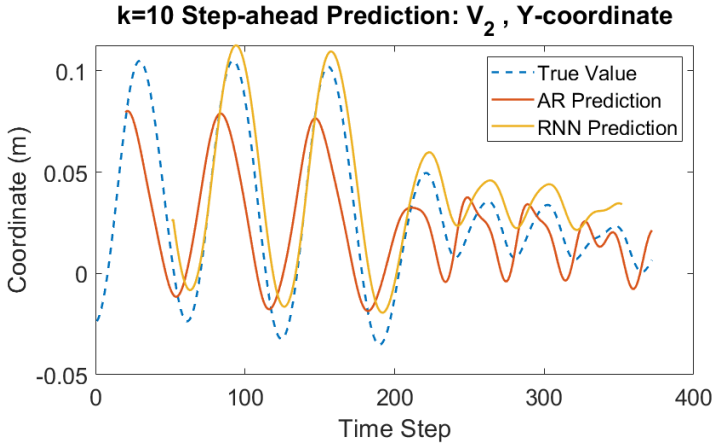
Table 5.5: Predicting human body vectors from KIT Database, RMS Errors (cm)

	Shoulder			Elbow			Wrist		
	$v_{1,x}$	$v_{1,y}$	$v_{1,z}$	$v_{2,x}$	$v_{2,y}$	$v_{2,z}$	$v_{3,x}$	$v_{3,y}$	$v_{3,z}$
AR Model	0.16	0.23	0.37	1.47	2.61	0.36	3.38	2.02	0.64
RNN Model	0.53	0.42	0.43	0.90	1.42	1.41	2.08	0.57	0.31

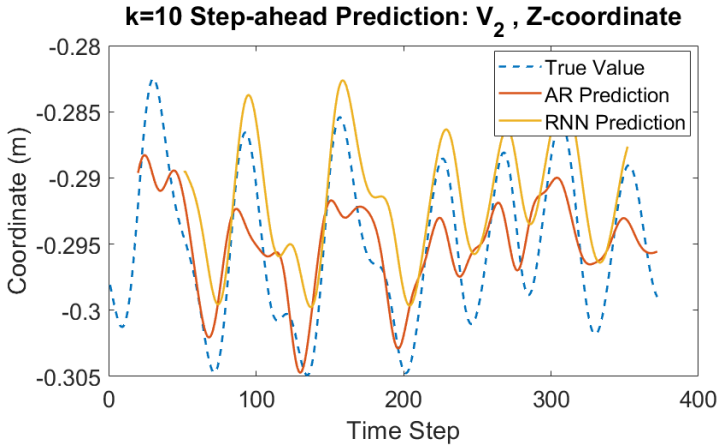
The Root-Mean-Square (RMS) prediction errors were computed on the test set for both models, and are listed in Table 5.5. While the RNN model did not improve upon the AR model for every component, it reduced the prediction errors in the components with the worst performance using AR (Figure 5.16). The RNN model also performed better overall, with an average RMS error of ˜0.90 cm, compared to ˜1.25 cm for the AR model. Figure 5.15 shows that while the RNN model tended to overshoot the ground truth, and be offset from it, the



(a) v_2 : X-component



(b) v_2 : Y-component



(c) v_2 : Z-component

Figure 5.15: Predictions from the AR model (red) and RNN model (yellow), k=10 time steps ahead for body vector v_2, realigned with ground truth (blue).

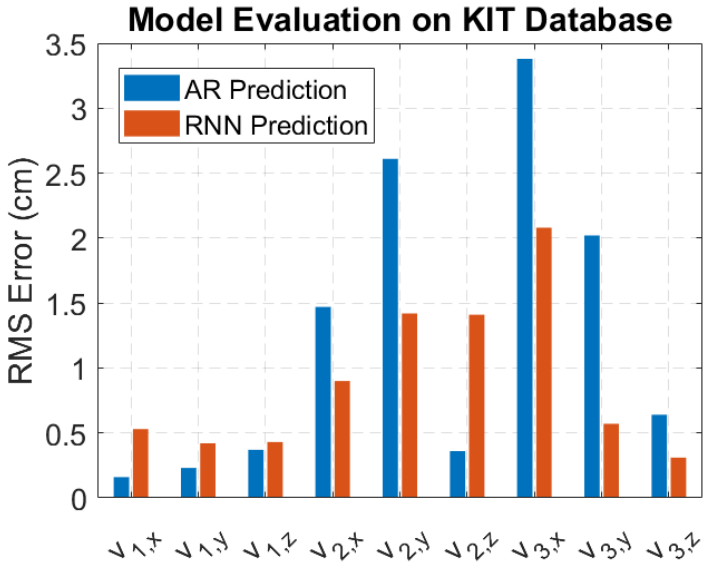
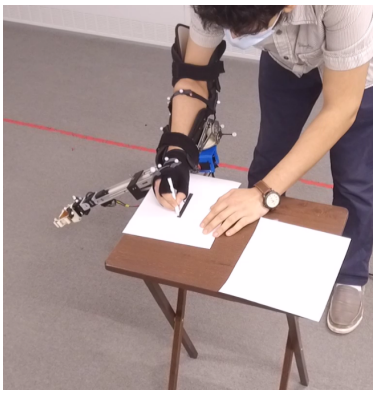


Figure 5.16: RMS errors in the AR and RNN models for prediction on the KIT Database motions.

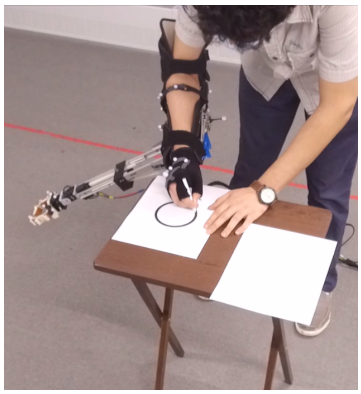
tracking of overall motion trends was better than the AR model.

5.4.2 Implementation on the WRF

Having obtained two predictive models for human arm motion that performed well on the KIT Database, we applied them for stabilization of the WRF's end-effector at an initial pose when subjected to disturbances due to movement of the user's right arm. We considered five task scenarios, shown in Figure 5.17, that involved periodic arm movements of relatively small magnitude—(a) tracing a line of length 10 cm, (b) tracing a circle of diameter 10 cm, (c) wiping a desk top, (d) painting with small brush strokes on a canvas, and (e) placing ten objects into shelves of a table-top drawer unit. Each task was performed for ˜5 minutes. The initial end-effector pose was selected to be on the right of the user and below them, so as to not impede the task.

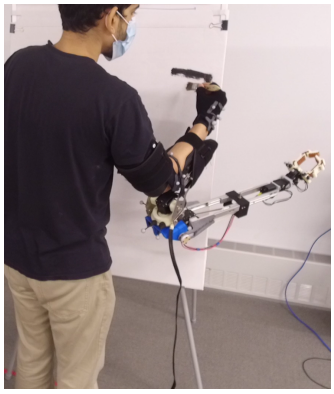


(a) Tracing a line.



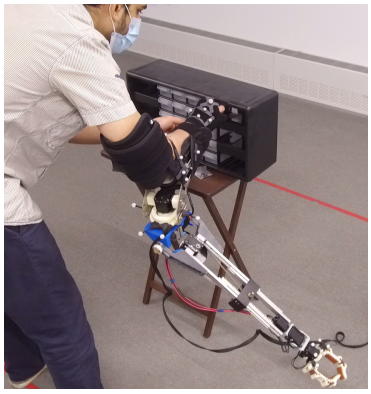
(b) Tracing a circle.

(c) Wiping a desk.



(d) Painting on a canvas.

(e) Setup for stowing ten items into a drawer unit.



(f) Stowing items into the drawers.

Figure 5.17: Scenarios in which the WRF's end-effector was stabilized while the user performed a task.

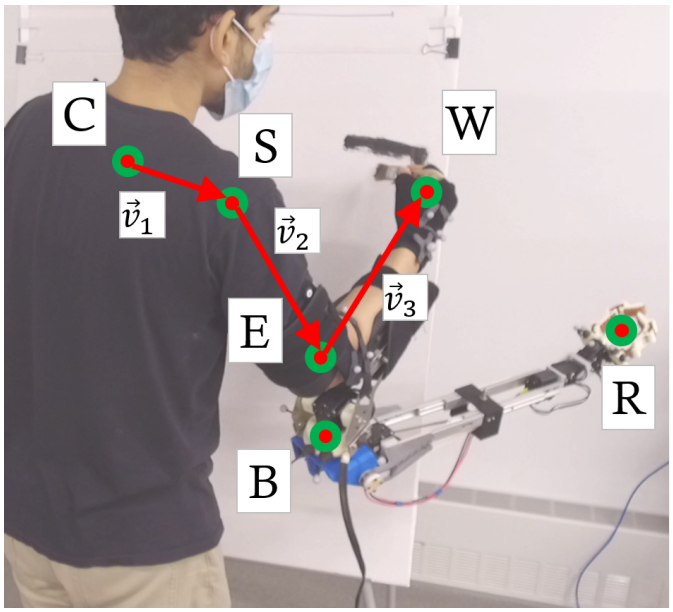


Figure 5.18: Tracked markers and positions for determining the WRF's motor joint angles.

Optical markers were placed on the user's right hand and elbow, as well as on the WRF's end-effector and near the DoF-1 motor (Figure 5.18). These markers were tracked at 120 Hz using an OptiTrack motion capture system [5]. The raw marker position data was smoothed and filtered using an IIR low-pass digital filter with transfer function coefficients for 6 Hz normalized cutoff frequency [2], following the techniques discussed in [102, 101].

In all the scenarios shown in Figure 5.17, we assumed the body vector v_1 to be constant in each task, as the human shoulder and torso remained almost stationary at their initial positions. The other relevant points, B (base position of the WRF), and R (position of the end-effector), to be tracked are shown in Figure 5.18, We aimed to keep the end-effector static at the initial point $R = R_0$ at the start of each task. If the user's arm were to move, the end-effector would also move by an amount $\Delta\vec{P} = R_t - R_0$ at time t. To generate appropriate setpoints for the WRF's motors, $\Delta\vec{P}$ is converted from the a global frame G (fixed lab frame)

to the robot's base frame B. In terms of homogeneous transformation matrices, we need to convert T_G^R into T_B^R using the same notation as Chapter 3. Using the elbow frame E as an intermediate,

$$T_G^R = T_G^E T_E^B T_B^R \tag{5.14}$$

$$\therefore T_B^R = (T_E^B)^{-1}(T_G^E)^{-1} T_G^R \tag{5.15}$$

The transformation between the robot base B and elbow E is constant, while the transformation T_G^E consists of two variable parts: the rotation matrix R_G^E between the elbow and ground frames, and the position of the elbow, $\vec{P_E}$ which is tracked directly by the motion capture system. R_G^E is the rotation matrix that takes the unit vector along the local X-axis, $\hat{i} = [0, 0, 1]^T$, and aligns it with the unit vector along the human forearm, $\hat{v}_3$, in the ground frame. Using the approximate method for inverse kinematics (Jacobian pseudoinverse) discussed in Chapter 3, we can determine the change in WRF joint variables:

$$\Delta\theta = J^{+}(-R_B^R \Delta\vec{P}) \tag{5.16}$$

At time t, this gives the desired setpoint reference for each motor used for direct feedback control:

$$\hat{\theta}_d[t] = \theta_d[t = 0] + \Delta\theta \tag{5.17}$$

Following the procedure shown in Figure 5.1b, the predictive models were used to generate setpoint references over a time horizon of ˜83 ms for each motor in

the WRF:

$$\theta_d[t + \Delta t_i] = \hat{\theta}_d[t] \; + \; \Delta\tilde{\theta}_i \quad , \quad i \in [1, k] \tag{5.18}$$

For a stereo camera frame received at time t, a sequence of k=10 joint angle references θ_d were sent to each motor, with $\Delta t_i \in [0, 83]$ ms, $i \in [1, k]$. As described above, $\hat{\theta}_d[t]$ is the desired joint angle in direct feedback control, computed using inverse kinematics for the detected human and robot poses at time t. The predictions from the AR and RNN models are represented as residuals $\Delta\tilde{\theta}_i$ added to $\hat{\theta}_d[t]$.

Table 5.6: End-Effector position errors in 3D

Task	Mean error in end-effector position (cm)		
	Direct Control	AR Prediction	RNN+AR Prediction
Line Tracing	5.02	4.32 (-13.94 %)	4.54 (-9.56 %)
Circle Tracing	6.92	5.87 (-15.17 %)	5.09 (-26.44 %)
Wiping	11.01	8.91 (-19.07 %)	8.74 (-20.62 %)
Painting	14.30	12.04 (-15.82 %)	11.45 (-19.95 %)
Stowing	15.87	14.74 (-7.12 %)	12.65 (-20.28 %)

During implementation, it was found that while the AR model could generate predictions nearly in real time, the RNN model had lags of up to ˜50 ms when predicting over the specified time horizon. To account for these lags, the RNN model was executed in parallel with the AR model. Until a prediction was received from the RNN model, the AR prediction was used for computing θ_d. Depending on the amount of lag, determined through time stamps, a corresponding number of RNN predictions were discarded (typically the first 5–6 terms), and the remaining ones were added to the sequence $\Delta\tilde{\theta}_i$ to be sent to the

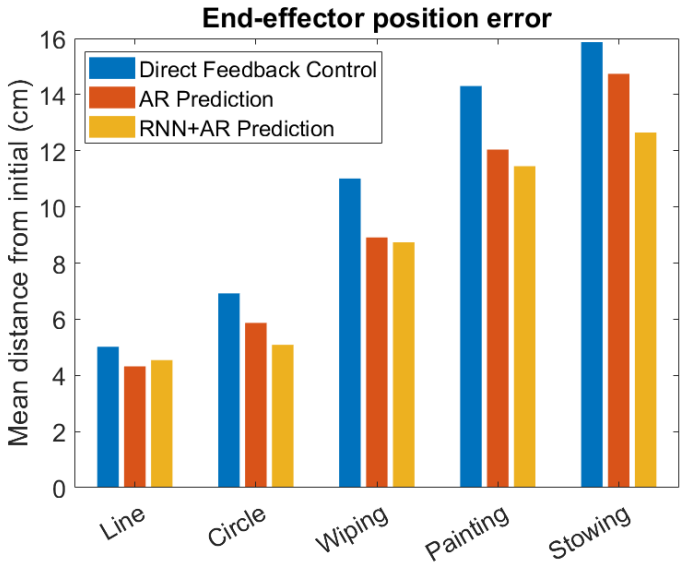


Figure 5.19: Summary of end-effector stabilization position errors.

motors.

This implementation of human motion prediction (RNN + AR) reduced the mean error in end-effector position by up to ˜26 % over direct feedback control, while the AR model alone was able to improve upon direct feedback control by up to ˜19 %, as listed in Table 5.6. Figure 5.19 shows that the performance of all three control methods varied according to the task, with more structured and mostly planar motions such as tracing a line and circle showing better stabilization performance compared to full 3D motions with less structured or periodic behavior such as stowing items into a drawer.

5.5 Conclusion

We presented an approach involving human motion prediction for stabilizing the WRF's end-effector, first in a planar scenario[1], and then in full 3D tasks. We performed linear system identification, both to determine the characteristics of the actuators, and to estimate the delays in both these scenarios. The predictive models took the form of autoregessive time series in planar and 3D, and a recurrent neural network in the 3D case.

While these methods resulted in lower mean position errors of the end-effector compared to direct feedback control, the improvements in performance were marginal for the tasks described in this chapter. Stabilizing over a trajectory in free space, and over bulk human motions such as walking, would be even more challenging. Aside from the uncertainties in human motion, these scenarios would also require sensing and actuation with minimal delay. Along with vision-based sensing, systems with wearable sensors such as Inertial Measurement Units mounted on the human and robot might help improve performance. Having the actuation and control systems off-board [116] would, for a limited workspace, sidestep the trade-off between the motor power and weight of an SR device.

Thus far we have addressed the WRF's impact on the user, and seen how the user can affect and disturb the WRF, along with methods for mitigating both. In the next chapter, we analyze how users interacted with the WRF while it assisted them in a close-range collaborative pick-and-place task.

[1]Portions of this chapter have been published in [115].

CHAPTER 6

HUMAN-ROBOT COLLABORATION STUDY

Thus far, we have described the design process, biomechanics, and control of the WRF. In this chapter, we present an interaction study where the WRF predicts the user's intention and autonomously assists them in a collaborative pick-and-place task.

Other SR devices have generally included modes of controlling the robot directly and intentionally by the user, such as a sixth finger with switch-operated grasping [55], control based on muscle electromyography (EMG) [67], hand-pose synergies by measuring finger joint angles [118], foot-based control of SR arms [94, 37], and a brain-machine interface [87].

These interaction modes allow for fine-grained command over an SR device, and may eventually also lead to the incorporation of the device into a user's body schema [96]. At the same time, these modes require the user to concentrate on the robot's state and motion, adding to their cognitive load while performing tasks with the robot. In more general human-robot collaboration scenarios, not necessarily with wearable robots, there have been efforts to ease some of this cognitive load by predicting the user's intent and acting appropriately, through approaches such as Bayesian filtering [58], recurrent neural networks [26], and changing the robot's strategy based on a Partially Observable Markov Decision Process (POMDP) model of the mutual trust between the human and robot [30]. Among wearable robots, learning-from-demonstration (LfD) approaches have been applied, for instance with shoulder-mounted SR limbs using Colored Petri Nets [70] for collaborative action in a structured assembly task. Another example of LfD applied to wearable robots was a supervised learning model (K-

nearest neighbors) used to map surface EMG signals to desired poses of a hand exoskeleton to provide assistance in grasping tasks [100].

From the pilot study described in Chapter 3, we found that a wearable robot with some level of autonomy was desirable to users. In the study described in this chapter, we applied supervised learning approaches to provide these autonomous capabilities to a mid-scale SR device that can be re-positioned much more readily than large-scale SR arms, with the user's body movements serving as the input. As with the pilot, there was also a direct speech-based interaction condition. We explored the relative performance and user preferences in both of these broad cognitive frameworks: direct, intentional control of the WRF, and the robot acting after predicting human intent.

The study involved a planar task, without vertical pitching of the arm, and a static wrist orientation. The user performed an assembly sub-task with a relatively higher cognitive load, and handed over the assembled object to the WRF for stowing into bins. Users familiarized themselves with the system in the initial trials, and provided direct speech commands to the robot. Human and robot motion data was annotated with the speech commands from these trials to train supervised learning models for predicting the user's intention. This was followed by two test conditions, counterbalanced across users—another round with direct speech control, and a predictive robot mode.

In the latter test condition, we aimed to predict both the user's intended commands, as well as the intended target bin for placing the assembled object. This was expected to reduce the cognitive load on the user compared to direct speech commands. The user study procedure is illustrated in Figure 6.1. We evaluated the performance of the human-robot team in these two conditions

using objective and subjective human-robot fluency metrics.

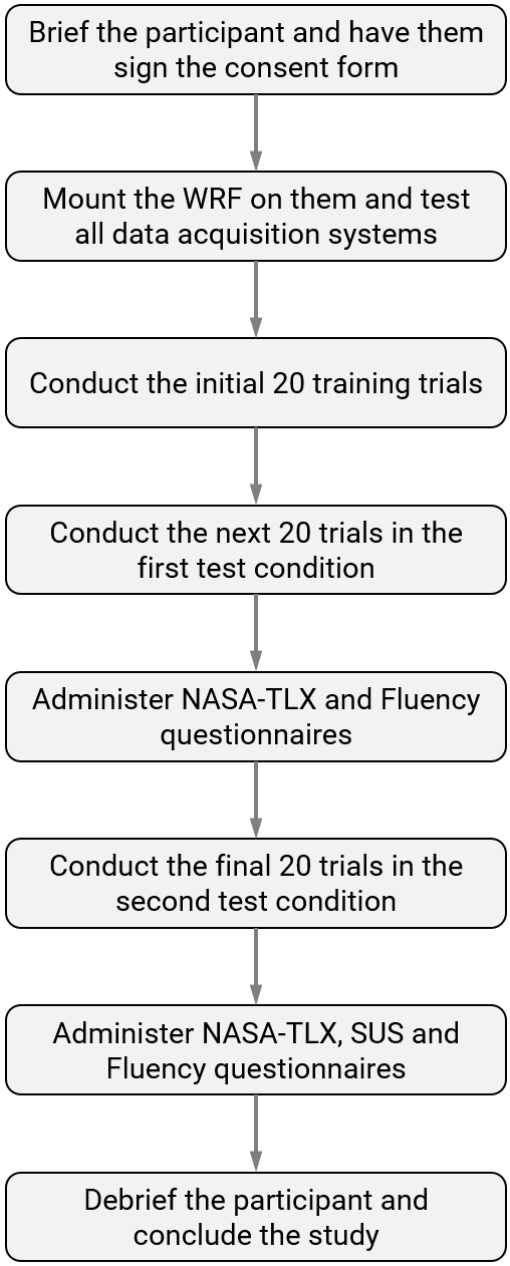
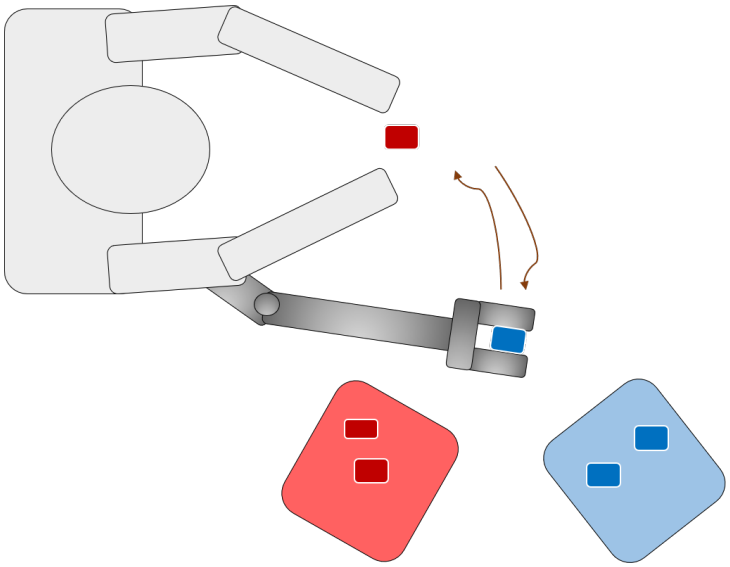
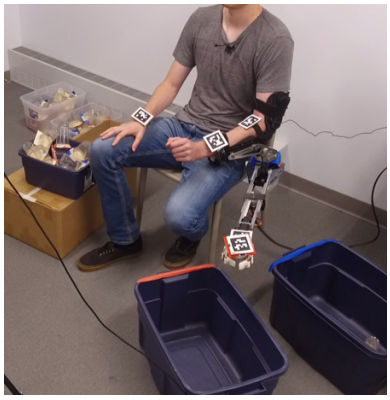


Figure 6.1: Flow diagram for the user study. The robot control modes in the first and second test conditions were counterbalanced across users.



(a) Planar collaboration task.



(b) Physical system setup.

(c) View from stereo camera.

Figure 6.2: Schematic of the human-robot collaboration task: the human hands over a cup to the robot, which places it in color-coded bins.

6.1 Human-Robot Collaboration Task

The collaboration task in this study involved a seated user performing an assembly activity close to their laps, with the WRF mounted on their left arm operating between the user's working region and drop-off locations on one side (Figure 6.2). It included three sub-tasks: assembly of an object, handover to the robot, and placing the object in a target bin. The objects in this case were red

and blue color-coded cups, with the user assembling a lid onto a cup as the first sub-task.

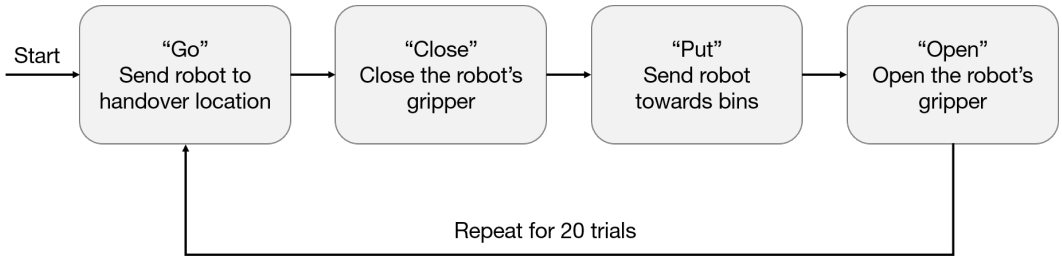


Figure 6.3: Overview of training stage: user provides speech commands to the WRF.

The user provided the following speech commands to the WRF related to each sub-task (Figure 6.3):

- **Go:** The user gives this command to bring the robot close to their working region, and prepare for the handover. Users can provide this command at any stage of completion of the lid assembly sub-task.
- **Close:** After assembling the lid, the user places the cup in the robot's gripper, and gives this command to close the gripper and complete the handover.
- **Put:** This command sends the robot towards the color-coded target bins. The exact locations of the bins is not known *a-priori*, only their general direction from the user (in this case towards their left side). The user then guides the WRF with their left arm and places the end-effector above the correct target bin.
- **Open:** This command opens the WRF's gripper and drops the cup into the bin.

During trials, the user's speech commands were recognized using the Pock-

etsphinx system [54], and the user's arm motions, as well as motion of the robot's base and end-effector were tracked using AprilTag markers [81] through an Orbbec Astra stereo camera mounted overhead (Figure 6.2c).

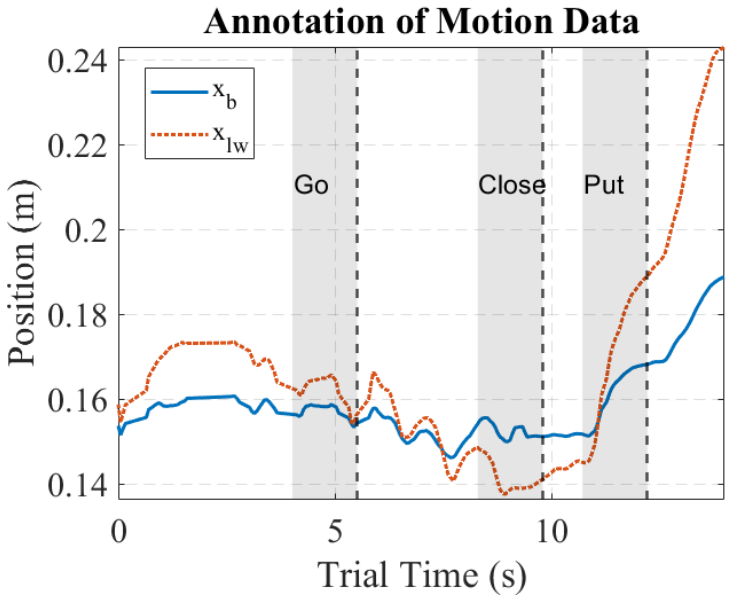
The tracked poses of interest were the x and y coordinates of markers placed on the human and robot, with the human's wrist markers denoted collectively by the vector $X_H = [x_{lw}, y_{lw}, x_{rw}, y_{rw}]$, with subscripts lw denoting the left wrist, and rw for the right wrist. The robot's pose is tracked by two markers: one placed on the human's elbow directly above the robot's base (DoF-1), and the other on the robot's end-effector, collectively denoted as $X_R = [x_b, y_b, x_{ee}, y_{ee}]$, with subscripts b for the coordinates of the base, and ee for the end-effector.

Each user performed 20 trials in the initial stage, with the collected data used for training two predictive models for the human-robot team—one for predicting the sub-task, and the other for predicting the intended target for drop-off. This was followed by two test stages—with the robot first in the predictive mode, or direct speech control mode, counterbalanced for order effects. Each set of models was trained and applied independently to every user, allowing for customized predictions based on individual interaction patterns.

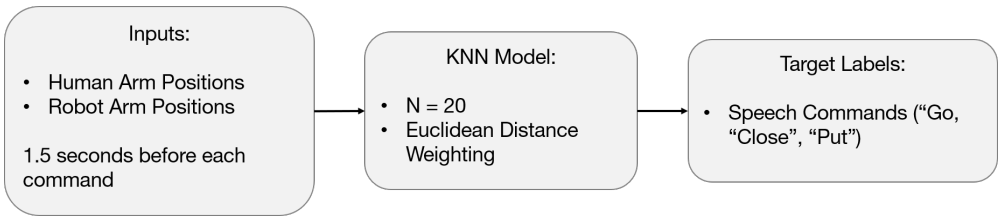
6.1.1 Sub-task Prediction

Using the speech commands and motion data of the human-robot team during the initial 20 trials, we trained K-Nearest Neighbors (KNN) classifiers [48] to predict the human's intended speech command based on their pose. As seen in [100], KNNs provide a computationally inexpensive and non-parametric method for determining intent based on body movement alone. As the tasks in

this study were fairly repetitive, the motion data was expected to be sufficiently distinct for each voice command. Also, as a KNN model is trained individually for each user, it was expected that repetitions of similar motion patterns would lead to good prediction performance.



(a) Motion data annotated with speech commands.



(b) Schematic of KNN classifier.

Figure 6.4: KNN classifier for predicting the intended speech command for a sub-task.

As shown in Figure 6.4, the input to the KNN model was a sequence of 8-dimensional vectors X, a collection of human and robot poses, $X = [X_R, X_H]$, that preceded each speech command by 1.5 seconds. Every such input sequence was annotated with the corresponding speech command as the target label. This predictive model was built for the "Go", "Close", and "Put" commands. During

training, for each of these commands, a set of 39 or 40 vectors X (number of frames in 1.5 seconds for a camera frame rate of ˜26 Hz) were annotated with the command (Figure 6.4a).

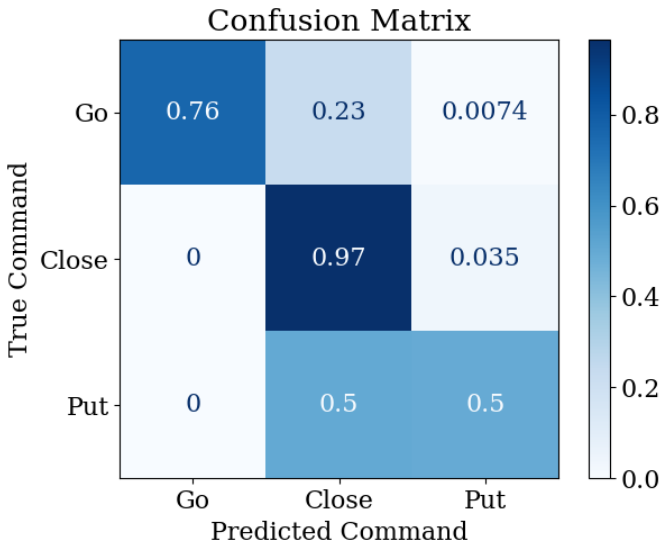


Figure 6.5: Confusion matrix of KNN performance for one participant (P13) based on their first 20 training trials.

We set the number of neighbors (K) to be 20 in the KNN classification model, with Euclidean distance weighting for each label. During predictive robot mode stage, the intended speech command was determined at each instant of detected human and robot poses X, with the transition between commands constrained to follow the sequence shown in Figure 6.3. The "Open" command was triggered when the WRF's end-effector came within 10 cm of a target location.

The performance of the KNN classifiers varied with the participants depending on their initial 20 training trials, e.g. as shown in Figure 6.5, the classifier for Participant-13 had good predictions for "Go" and "Close", while "Put" was confused with "Close".

6.1.2 Target Prediction

During the sub-task of placing a cup into a bin, neither the number of target bins nor the color of the cup were known to the system *a-priori*. We applied a target prediction model that determined the number of target bins, then formed clusters of target bins, and finally predicted the user's intended target based on classification of human and robot pose data. Although in the present study we had two classes of cups (red and blue), this made the system robust to the addition of more classes and more target bins at unspecified locations. For these models, the human and robot pose sequence, $X = [X_R, X_H]$, was recorded from the utterance of the "Put" command spoken by the user until the "Open".

Determining the number of target bins is analogous to determining the number of clusters in two-dimensional data. The pose of the robot's end-effector $[x_{ee}, y_{ee}]$ relative to its base $[x_b, y_b]$ was recorded at the utterance of every "Open" command during the first 20 trials of the training stage. An information-theoretic approach was used to find the number of clusters, based on a rate-distortion "jump" metric [106].

The distortion $\hat{d}_K$ determines the amount of dispersion within a cluster, defined in this case as the mean-squared error between a cluster center c_K and the sample vector $P = [x_{ee}, y_{ee}]^T - [x_b, y_b]^T$. K-means clustering [73] is used to determine the cluster centers c_K, for a variable number of clusters K. The the "jump" metric J_K is computed as follows:

$$J_K = \hat{d}_K^Y - \hat{d}_{K-1}^Y \tag{6.1}$$

We considered K to be between 1 and 8, with the transformation exponent $Y = 0.4$. The target number determination was performed at the end of every trial

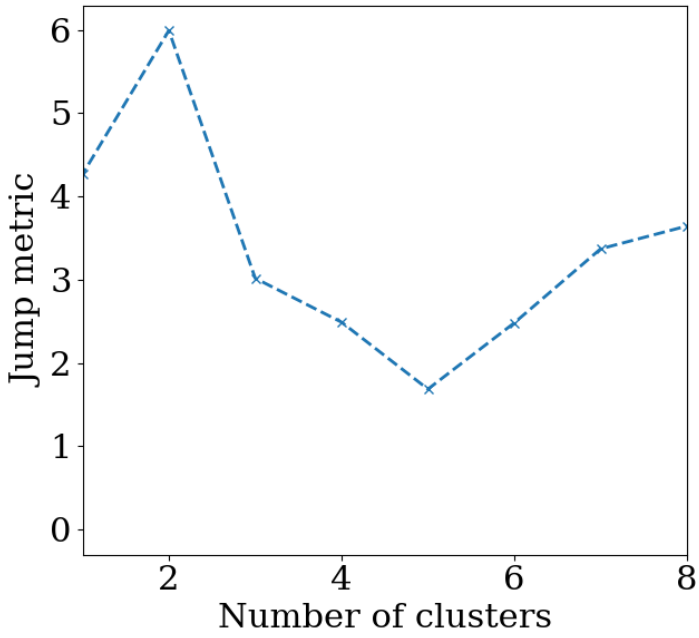


Figure 6.6: Jump metric for determining number of clusters in data.

in the training stage. Figure 6.6 shows an example of this method after 20 trials, with $K = 2$ having the largest jump metric.

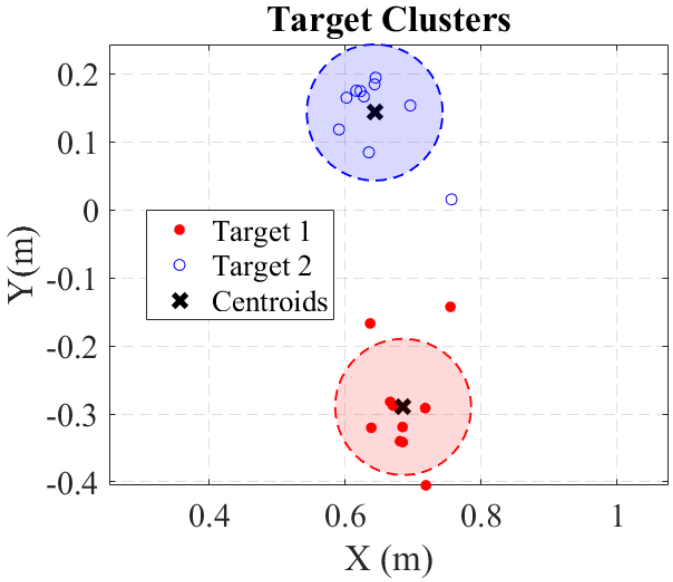
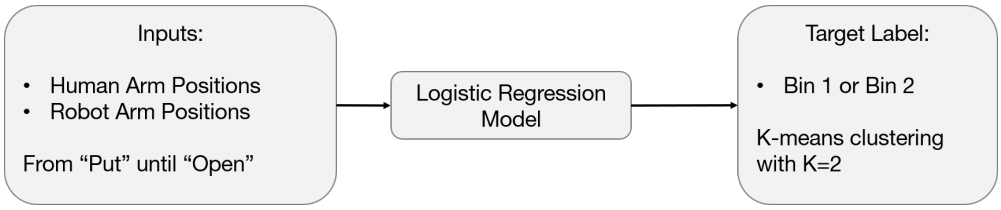


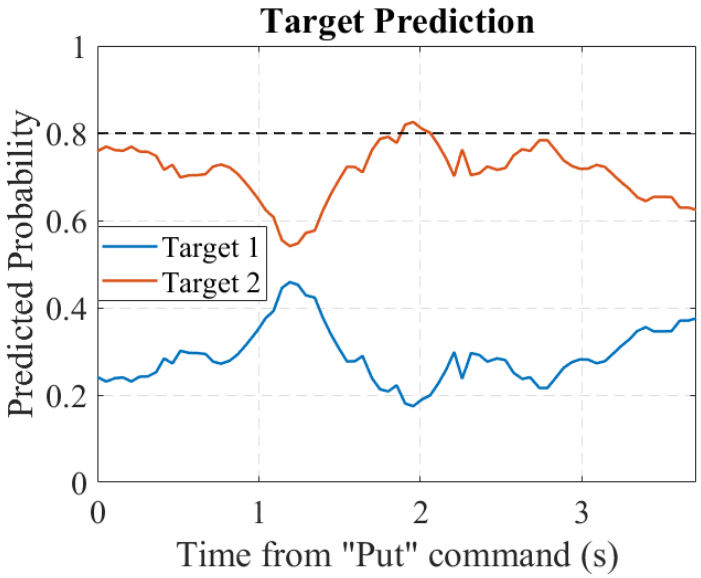
Figure 6.7: K-means target clustering with K=2. During predictive mode, the gripper opens when it reaches within 10 cm of a cluster centroid.

The number of clusters K with the largest jump were then chosen for K-means clustering of the target positions, labeling them from 1 to K (Figure 6.7).

For each target in a particular cluster, the human-robot pose sequences X between the "Put" and "Open" commands that led to the drop locations in that cluster were annotated with the target label number. Finally, a logistic regression classifier was trained with the pose sequences X as input and cluster label number as the target (Figure 6.8a).



(a) Schematic of Logistic Regression classifier.



(b) Predicted probability of a target bin during testing.

Figure 6.8: Logistic regression classifier for predicting the intended target bin.

During the predictive robot mode stage, we assumed that the number of target bins, object classes, and target bin locations remained unchanged from the training stage. When the "Put" command was predicted during a trial, the lo-

gistic regression model began predicting the intended target bin based on the instantaneous pose X of the human-robot system. Once the predicted probability of a particular target exceeded 0.8, it was considered to be the intended target (Figure 6.8b). A closed-loop control sequence then placed the WRF's end-effector at the mean pose for that target cluster. The "Open" command was triggered when the end-effector was within 10 cm of the target cluster centroid.

6.2 User Study

In the user study, we recruited N=24 participants, consisting of graduate and undergraduate University students. They were vetted to preclude neuromuscular conditions that could have led to adverse effects from the task involving wearing the WRF for over 30 minutes. Each participant was given a $ 20 gift card on successful completion. This study was approved by the Cornell University Institutional Review Board for Human Participant Research, under protocol number 1706007253.

We analyzed objective and subjective human-robot fluency metrics, as described in [51], for both conditions of robot control in this study.

The objective metrics were defined on the basis of the durations of human activity (H-ACT) and robot activity (R-ACT) during each trial. The video recordings from these trials were coded using BORIS [41], with time-stamp triggers set up for the start and end of a human activity—defined as intentional body movements made by participants which would aid in the completion of the task. For instance, the time from when a participant would begin reaching for a cup, until they finished assembling it, would count as valid human activity. After as-

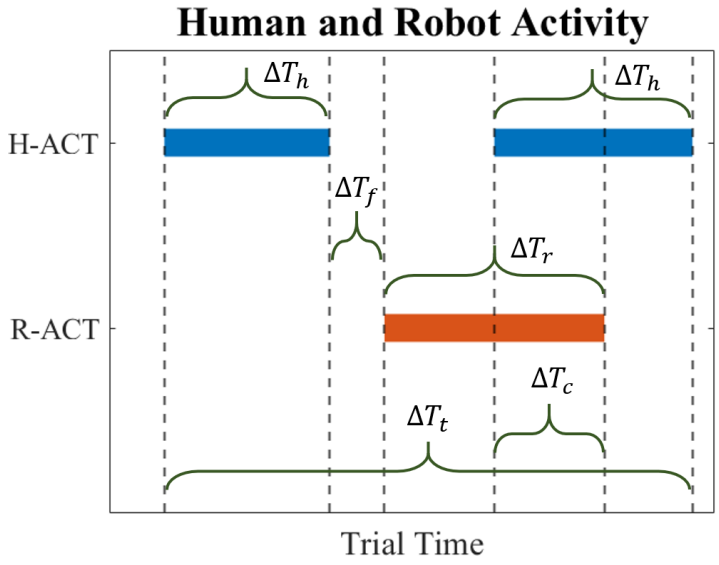


Figure 6.9: Illustrative plot showing times of human activity (H-ACT) and robot activity (R-ACT) during each trial.

sembly, the time spent waiting for the robot to approach the handover location was not considered human activity, even if there were small movements of the participant's arms. Robot activity was determined using data from the motor encoders.

Objective Metrics As illustrated in Figure 6.9, ΔT_t is the time taken for a trial starting from a cup assembly to placing it in a bin. Let $\Delta T_{t,i}$ be the time taken for the i^{th} trial by a participant. Similarly, define $\Delta T_{h,i}$ as the period of human activity, and $\Delta T_{r,i}$ as the period of robot activity. $\Delta T_{f,i}$ is the gap between when the human finishes their subtask and when the robot starts acting. $\Delta T_{c,i}$ is the duration for which both the human and robot are active. With this data we determined the following objective metrics:

- Mean Trial Time (T_M): the mean time taken, over $n = 20$ trials, of complet-

ing the full assembly and pick-place task:

$$T_M = \frac{1}{n}\sum_{i=1}^{n} \Delta T_{t,i} \tag{6.2}$$

- Human Idle Time (T_{HI}): the ratio of time for which the human either remained static or did not move in a way beneficial for performing the task, to the total time spent in each condition.

$$T_{HI} = 1 - \frac{\sum_{i=1}^{n} \Delta T_{h,i}}{\sum_{i=1}^{n} \Delta T_{t,i}} \tag{6.3}$$

- Robot Idle Time (T_{RI}): the ratio of the time for which the robot remained static (zero motor velocities), divided by the total time spent in each condition. Any agent spending too much time idle is indicative of poor fluency.

$$T_{RI} = 1 - \frac{\sum_{i=1}^{n} \Delta T_{r,i}}{\sum_{i=1}^{n} \Delta T_{t,i}} \tag{6.4}$$

- Functional Delay (T_{FD}): this is the delay experienced by the user immediately after completing their part of the sub-task, while waiting for the robot to act, ideally to be minimized for fluent interactions. It is the ratio of the sum of durations after human actions are completed and prior to the beginning of robot actions, to the total time per condition.

$$T_{FD} = \frac{\sum_{i=1}^{n} \Delta T_{f,i}}{\sum_{i=1}^{n} \Delta T_{t,i}} \tag{6.5}$$

- Concurrent Acitivity(T_{FD}): this is the proportion of the task for which both the human and robot are acting simultaneously. A high concurrent activity is indicative of better fluency [51].

$$T_{CA} = \frac{\sum_{i=1}^{n} \Delta T_{c,i}}{\sum_{i=1}^{n} \Delta T_{t,i}} \tag{6.6}$$

Table 6.1: Human-robot fluency questionnaire

1. The human-robot team worked fluently together	15. The robot and I understand each other.
2. The human-robot team's fluency improved over time.	16. I believe the robot likes me.
3. The robot contributed to the fluency of the interaction.	17. The robot and I respect each other.
4. I had to carry the weight to make the human-robot team better. (R)	18. I am confident in the robot's ability to help me.
5. The robot contributed equally to the team performance.	19. I feel that the robot appreciates me.
6. I was the most important team member on the team. (R)	20. The robot and I trust each other.
7. The robot was the most important team member on the team.	21. The robot perceives accurately what my goals are.
8. I trusted the robot to do the right thing at the right time.	22. The robot does not understand what I am trying to accomplish. (R)
9. The robot was trustworthy.	23. The robot and I are working towards mutually agreed upon goals.
10. The robot was intelligent.	24. I find what I am doing with the robot confusing. (R)
11. The robot was committed to the task.	25. The robot's had an important contribution to the success of the team.
12. The human-robot team improved over time.	26. The robot was committed to the success of the team.
13. The robot's performance improved over time.	27. I was committed to the success of the team.
14. I feel uncomfortable with the robot.	28. The robot was cooperative.

Subjective Metrics Through the questionnaire listed in Table 6.1, we also measured subjective fluency metrics. Each question was rated by the participants on a 7-point Likert item [68] going from "Strongly disagree (1)" to "Strongly agree (7)". Questions marked with (R) indicate a reversed scale where a lower score is

better. This questionnaire resulted in seven scales comprising of multiple questions, as well as five single-item measures from individual questions:

- Human-robot Fluency: This scale consisted of questions 1, 2, and 3 from Table 6.1. It measured the perceived fluency of the robot, and its evolution over time.
- Robot Relative Contribution: This scale, consisting of questions 4, 5, 6, and 7 measured the levels of perceived contribution that the user felt while performing the trials.
- Trust in Robot: This scale measured the level of trust the human reported to have in the robot in terms of its abilities and effectiveness (questions 8 and 9).
- Positive Teammate Traits: This scale asked the user if they thought of the robot as an effective teammate (questions 9, 10, and 11).
- Improvement: This scale measured the temporal aspect the task, in terms of perceived improvement (questions 2, 12, and 13).
- Working Alliance, Bond: This scale was adapted to human-robot teams in [51] from the Working Alliance Inventory [52], a measure of the effectiveness of an agent providing an intervention or treatment to another agent. Covering questions 14–20, the Bond subscale looked at the level of comfort that the user had with the WRF.
- Working Alliance, Goal: Another subscale of the Working Alliance, the Goal measure (questions 21, 22, and 23) looked at the user's perception of consensus between them and the robot regarding the task.
- Single-item Measures:

- Working Alliance, Confusion: This subscale (question 24) asked the user how confusing they found the shared task.
- Robot's Contribution: Question 25 examined whether the user considered the robot to have contributed, in absolute terms, to the task.
- Robot's Commitment: Question 26 asked how much the user perceived the robot to be committed to the task.
- Human's Commitment: Question 27 asked the user for their level of commitment, to gauge whether they may have found the task unpleasant or cumbersome.
- Robot's Cooperation: Question 28 asked the user to rate the how cooperative they found the robot to be during the shared collaborative task.

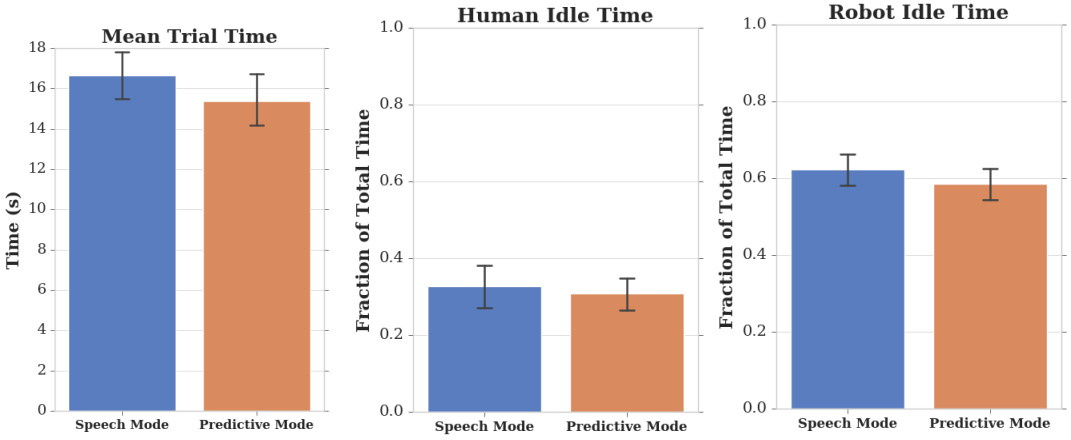
We also collected NASA-TLX questionnaires from each participant [46], after each of the direct speech control and robot predictive modes, along with an overall System Usability Scale (SUS) [24] at the end of the study.

Hypotheses For these metrics, we formulated the following hypotheses:

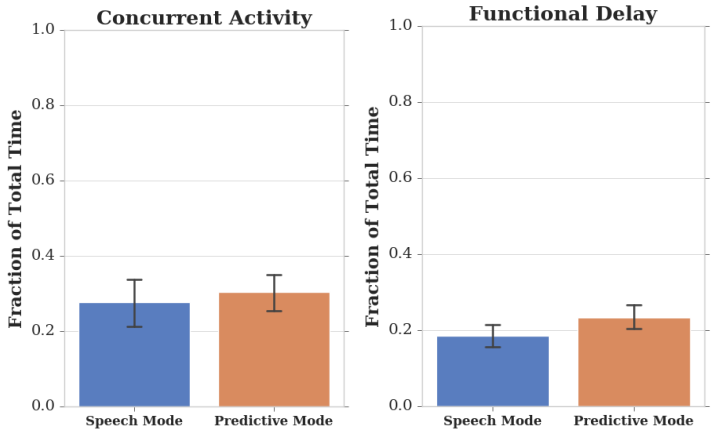
- **H1**: The means of the following objective metrics would be lower in the predictive robot condition compared to the speech-controlled condition:
  - H1a: Total Task Time
  - H1b: Human Idle Time
  - H1c: Robot Idle Time
  - H1d: Robot Functional Delay

- **H2**: The mean of the Concurrent Activity metric would be higher in the predictive robot condition compared to the speech-controlled condition.
- **H3**: For subjective metrics, the mean Likert scale scores would be higher for each metric in the predictive robot condition compared to the speech-controlled condition, indicating greater perceived fluency.

6.2.1 Results: Objective Metrics



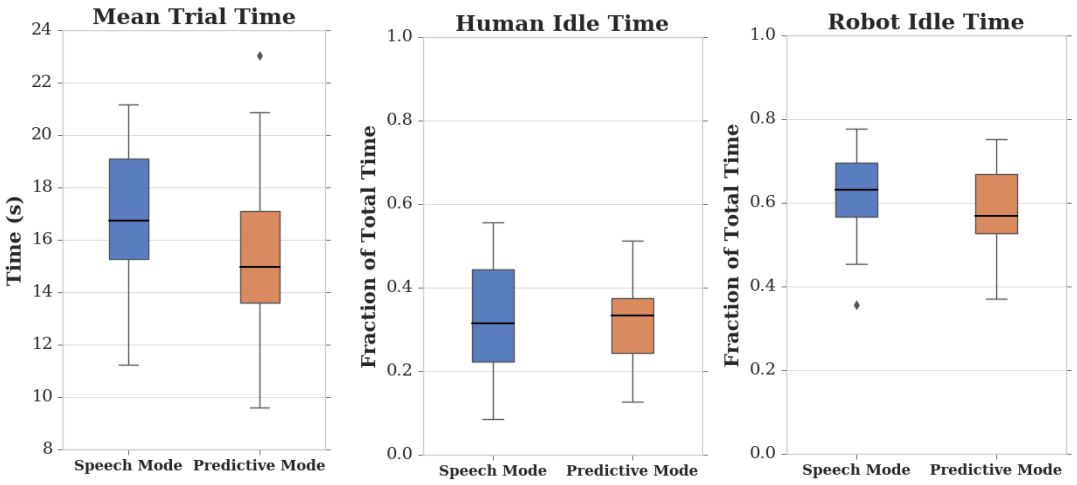
(a) Mean Trial Time (b) Human Idle Time (c) Robot Idle Time



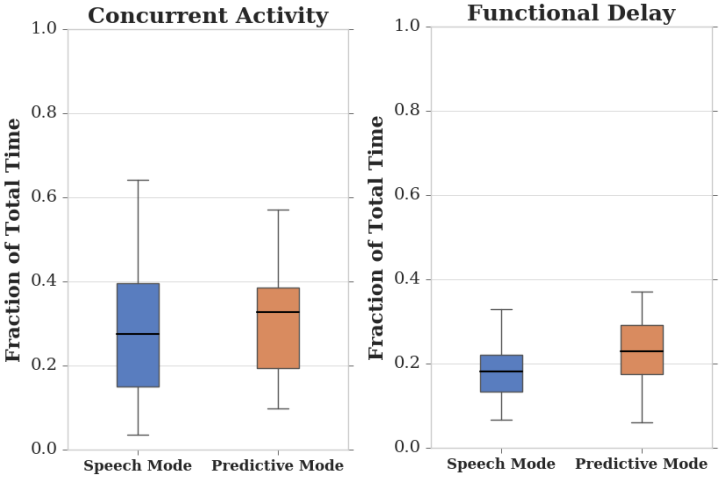
(d) Concurrent Activity (e) Functional Delay

Figure 6.10: Means and standard errors for the objective fluency metrics in the two conditions

The means and standard errors for objective fluency metrics are summarized in Figure 6.10. The distribution of this data is shown in Figure 6.11, with large variances observed across participants. Each sub-hypothesis in **H1** was tested using left-tailed paired t-tests, to determine whether the metric for predictive mode had a lower value than speech control. For **H2**, concurrent activity was tested using a right-tailed paired t-test, to determine whether the metric for predictive mode had a higher value than speech control.



(a) Mean Trial Time (b) Human Idle Time (c) Robot Idle Time



(d) Concurrent Activity (e) Functional Delay

Figure 6.11: Box plots for the objective fluency metrics.

The mean time taken to complete one trial was reduced in predictive mode

Table 6.2: Summary of objective fluency metrics

Metric	Speech		Predictive		Paired t-test, p	Effect size, Cohen's d
	Mean	Std. Dev.	Mean	Std. Dev.		
T_M (s)	16.64	2.88	15.36	3.32	0.066	0.41
T_{HI}	0.33	0.14	0.31	0.11	0.183	0.16
T_{RI}	0.62	0.11	0.58	0.09	0.013	0.37
T_{FD}	0.18	0.07	0.23	0.07	$>$0.5	-0.64
T_{CA}	0.28	0.16	0.30	0.12	0.087	0.19

(M = 15.36 s, SD = 3.32 s) compared to speech mode (M = 16.64 s, SD = 2.88 s), with t(23) = -2.08, p = .025, and a moderate effect size of Cohen's d = .41.

There was a marginal reduction in human idle time for predictive mode (M = 0.31, SD = 0.11) compared to speech mode (M = 0.33, SD = 0.14), however not significant enough to reject the null hypothesis with t(23) = -0.92, p $>$.05.

The idle time for the robot was lower in predictive mode (M = 0.58, SD = 0.09) compared to speech mode (M = 0.62, SD = 0.11), with t(23) = -2.38, p = 0.013, and a moderate effect size of Cohen's d = .37.

Functional delay showed an opposite effect in predictive mode (M = 0.07, SD = 0.11) compared to speech mode (M = 0.18, SD = 0.07), with t(23) = 3.76, and Cohen's d = -0.64. This suggests that functional delay may increase in predictive mode since the human is not directly giving commands to the WRF, and must instead wait for the predictive model to respond and actuate the robot.

There was marginal improvement in concurrent activity for predictive mode (M = 0.30, SD = 0.12) compared to speech mode (M = 0.28, SD = 0.16), with t(23) = 1.40, p = .087, and Cohen's d = 0.19.

6.2.2 Results: Subjective Metrics

For subjective fluency, we summed up the items in the seven Likert scales, and compared them using right-tailed paired sample t-tests for the hypothesis **H3**—that the predictive robot mode would be perceived as more fluent than direct speech control.

The results from these scales are summarized in Table 6.3. Six of the seven grouped subjective fluency metrics were not significantly higher for the predictive mode compared to the speech mode at the $p < .05$ level. Only the Working Alliance, Goal metric was higher in the predictive mode, with $p = .045$, and a moderate effect size of Cohen's $d = 0.40$. Trust in the robot showed a moderate opposite effect, with Cohen's $d = -0.61$, indicating that a user may trust the robot more in speech mode due to a sense of enhanced agency and control over it.

Table 6.3: Summary of subjective fluency metric scales

Subjective Scale	Speech		Predictive		Paired t-test, p	d
	Mean	Std. D.	Mean	Std. D.		
Human-robot Fluency	16.21	3.06	16.21	4.05	0.500	0.0
Robot Rel. Contribution	12.79	3.97	13.67	3.93	0.133	0.22
Trust in Robot	10.54	2.67	8.75	3.15	>0.5	-0.61
Positive Team Traits	15.21	3.73	15.46	4.16	0.305	0.06
Improvement	16.04	3.60	16.83	4.29	0.096	0.20
Working Alliance, Bond	30.25	7.68	29.96	7.91	>0.5	-0.04
Working Alliance, Goal	13.71	4.70	15.38	3.46	0.045	0.40

Non-parametric tests were performed for the single-item measures, considered to be ordinal. We used right-tailed Wilcoxon signed-rank tests for this part of **H3**, to see if the predictive robot mode was more fluent than the speech control mode in each individual item. Table 6.4 lists the median, mode, range, and

skewness for each item.

Table 6.4: Summary of individual subjective fluency metric items

Subjective Item	[Median, Mode, Range]		Skewness	
	Speech	Predictive	Speech	Predictive
Confusion	[6 , 7, 4]	[6, 7, 6]	-0.91	-1.83
Robot's Contribution	[6, 6, 5]	[6, 6, 6]	-1.20	-1.45
Robot's Commitment	[5.5, 6, 6]	[6, 6, 6]	-0.89	-1.10
Human's Commitment	[7, 7, 5]	[7, 7, 5]	-2.51	-2.40
Robot's Cooperation	[6, 6, 5]	[6, 6, 5]	-2.30	-1.40

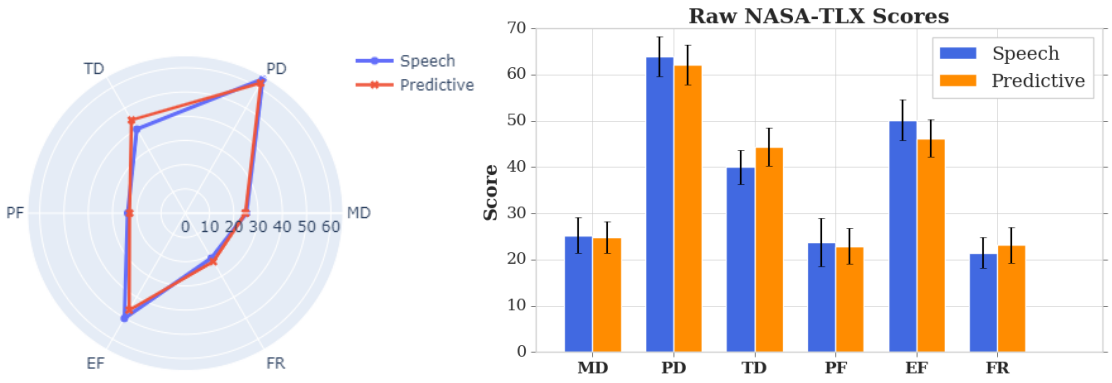
The Working Alliance item for human-robot teams measuring confusion had a median = 6 for speech and median = 6 for predictive, with p=.576, W = 37 using a right-tailed Wilcoxon signed-rank test, with effect size of rank-biserial correlation r=0.03.

The item measuring the robot's level of contribution had a median = 6 for speech and median = 6 for predictive, with p=.426, W = 19.5 using a right-tailed Wilcoxon signed-rank test, with effect size of rank-biserial correlation r=0.04.

The item measuring the robot's level of commitment had a median = 5.5 for speech and median = 6 for predictive, with p=.088, W = 35 using a right-tailed Wilcoxon signed-rank test, with effect size of rank-biserial correlation r=0.11.

The item measuring the human's level of commitment had a median = 7 for speech and median = 7 for predictive, with p=.500, W = 16 using a right-tailed Wilcoxon signed-rank test, with effect size of rank-biserial correlation r=0.06.

The item measuring the robot's level of cooperation had a median = 6 for speech and median = 7 for predictive, with p=.742, W = 18 using a right-tailed Wilcoxon signed-rank test, with effect size of rank-biserial correlation r=0.04.



(a) (b)

Figure 6.12: Raw NASA-TLX scores for the two interaction conditions.

Table 6.5: Summary of raw NASA-TLX scores

Subscale	Speech		Predictive	
	Mean	Std. Dev.	Mean	Std. Dev.
Mental Demand	25.21	19.31	24.79	17.35
Physical Demand	63.96	21.77	62.08	21.56
Temporal Demand	40.00	18.12	44.38	21.08
Performance	23.75	25.80	22.92	18.99
Effort	50.21	21.74	46.25	20.44
Frustration	21.46	16.71	23.12	19.10

The raw NASA-TLX scores for the the predictive robot mode and direct speech control modes were quite similar, as shown in Figure 6.12, and summarized in Table 6.5.

6.2.3 Post-study Semi-structured Interviews

After completing all trials with the WRF, each participant was debriefed through a short semi-structured interview [17] with the follow questions as prompts:

- Do you feel like you were able to do the task well?
- How easy or difficult was the robot to use?
- Did you have a preference between the two conditions?
- Did the robot seem like a competent teammate/augmentation?
- What changes or improvements would you suggest to this system?
- Do you see it being useful in a real setting?

Most participants reported that they were able to perform the task moderately well with minimal errors on their part. Some shortcomings of the WRF, such as *"the robot didn't understand me at times"*, and *"[the task] was a little too rushed when automatically closing [the gripper]"*, were cited as reasons for lower performance.

Nearly every participant reported that the weight the robot was a challenge, with reactions ranging from *"[the robot] got a little heavy to use by the end"*, to *"[the robot] is extremely heavy and very difficult to use"*, *"[the robot] was very easy to use mentally, but physically cumbersome"*. Other factors affecting usability were inconsistencies in the predictive mode for some participants, where they had to expend cognitive effort to correct the robot while also assembling the cup.

These shortcomings were also reflected in the WRF's score on the SUS [24] questionnaire (M = 60.52, SD = 12.16), indicating average perceived usability.

6.3 Conclusion

In this chapter, we described an interaction study between the WRF and a user in a planar, collaborative task involving assembly and pick-and-place. We considered two interaction modes: the user providing direct speech commands to the robot; and trained sub-task and target prediction models for the other mode: the robot predicting the user's intentions and acting accordingly in this structured task.

We measured the performance and preferences of the study participants using objective and subjective human-robot fluency metrics, and found that users performed the task marginally faster, and with more concurrent activity and lower robot idle times in the predictive mode. However, users had to wait longer in the predictive mode for the robot to act once they had done their part. In terms of subjective metrics, we found the users to perceive the robot as being less ambiguous in terms of mutual understanding of the task in the predictive mode, while having less trust in the robot due to loss of direct control over it in this condition.

Even for the metrics that showed improvements in predictive mode, the increase in performance was not consistent across users, as evidenced by small to moderate differences in the statistics. Aside from the limitations of the predictive models themselves, this may stem from strong personal preferences of wanting to retain intentional control over a device.

The task chosen for this study, a planar interaction where the user remains seated in one position, is well-structured, with successful predictions relying on the repeatability of the human motions involved. As seen in the previous

chapter, real-world deployments, say in a factory or warehouse setting, would involve more uncertainty in human motion, such as those induced by walking. These scenarios would also require the robot to be able to assist in a wider range of tasks, necessitating more robust and scalable prediction models.

In summary, despite these limitations, the user study described in this chapter represents an important step towards exploring the design of autonomous interactions with SR devices, and uncovers some of the challenges involved. The following chapter discusses the considerations that future developers of autonomous SR devices need to tackle for their wider adoption, and reflects on insights for their development obtained from the studies and analyses performed with the WRF.

CHAPTER 7
DISCUSSION

Supernumerary Robotic (SR) devices are emerging as a new class of wearable robots, owing to advances in lightweight materials, sensing, and actuation technologies. They have taken various forms, ranging from human-mimetic large-scale additional limbs and low-powered extra fingers, to snake-like robots and prehensile tails. We studied one such device, the Wearable Robotic Forearm (WRF), aimed for a novel paradigm—possessing sufficient power to assist in close-range collaborative tasks, while being lightweight enough to be freely readjusted by the user without placing too much strain on them. In this chapter, we reflect on the design, control, and interaction with the WRF, and aim to extrapolate our findings to make a broader comment on SR devices.

Design Analysis We started by looking at possible use cases and interaction scenarios, leading to the development of an initial prototype of the WRF mounted on the user's arm, with the base of the robot supported near their elbow joint. This mounting position was chosen so that the user would be able to perform fine corrections in the robot's trajectory.

For the WRF, while being mounted on the forearm helped perform close-range tasks, it affected the usability of the device due to physical strain on the wearer's arm—a factor that we were unable to fully mitigate through design iterations. Selection of a mounting position is largely determined by the intended usage of an SR device. As far as possible, designers should select positions closer to the core of the human body—the torso, back, shoulders, or hips. Even in the case of devices placed at the extremities, such as additional fingers,

they should aim to move the control and actuation modules closer to the core through the use of transmission systems (cable-based, pneumatic, or hydraulic).

The methods described in Chapter 2 stress upon the importance of including potential users in the design process, to find the right balance between ergonomics and usability. Each successive prototype of the WRF improved along both fronts, with the theoretical static loads on the human body remaining within prescribed limits, and an enhancement in a user's reachable workspace, as seen in Chapter 3. However, there were discrepancies between these analyses and the subjective experiences of users. Even after three design iterations, the WRF was restricted to short-term interactions due to ergonomics and robustness limitations. A commercially viable SR device would require mitigation of these discrepancies in design and user experience through numerous prototypes tested in wider *in-situ* studies.

Biomechanics After the design, through a simulation framework described in Chapter 4, we determined the detailed biomechanical effects of the WRF's motion on the user's arm muscles, and developed a strategy for finding trajectories that minimize these muscle loads. Aside from motion planning, accounting for biomechanical effects plays a large role in improving the design of a wearable robot, and is closely related with its ergonomics.

Care must also be taken in extrapolating laboratory results such as these to real-world interactions. Even if the simulation outcomes positively correlate with the subjective experiences of some users, this trend may not generalize well to all users, since the biomechanics model was based on a specific average human physiology, applied in the reduced scenario of a static human pose.

These limitations may be mitigated by accounting for the physiologies of individual users in the biomechanics analysis, along with ergonomic factors such as skin interface effects and muscle fatigue.

Robot Control and Stabilization Following the biomechanics analysis, we considered the other end of the consequences of wearing an SR device—how the human affects the robot's motion, and how these disturbances may be compensated for.

Human motion prediction approaches applied to robot control can be used to improve the performance of low-power SR devices with minimal off-board sensing. We describe methods for achieving end-effector stabilization in Chapter 5, first in a reduced 2D setting, and then in full 3D tasks. Although such predictions may not be necessary in mobile manipulation settings with powerful actuators and high-speed cameras, the performance of the WRF was limited due to the trade-off between motor power and weight, as described in Chapter 2. However, replacing the WRF's actuators with more powerful ones would lead to a significant increase in the overall weight of the device, considering the power densities of presently available motors. This trade-off between power and weight is ubiquitous in SR devices, and motivates the wider adoption of predictive models in their deployment.

While a high power density may be the most important factor in the choice of actuators, aside from weight, they present another trade-off—between performance and safety. In situations where there is clarity on the expected loads that would be applied on the robot, choosing actuators with much higher power ratings than necessary could adversely affect user experience, for instance through

increased jerks on the wearer's body. Typical commercial motors should also be augmented with compliant structures to mitigate the impacts of possible collisions with the wearer. Soft robotic structural elements and actuators may play a significant role in boosting the wider adoption of SR devices through enhanced safety, and should be incorporated in their design whenever possible.

Another factor restricting the improvement in stabilization performance of the WRF, aside from motor power, could also be the limitations imposed by the prediction models themselves. In this work, we did not aim to extend the state-of-the-art in terms of machine learning applied to human motion prediction, nor can we claim to have implemented the temporal models used for this purpose in the most efficient way possible. Even so, advances in motion prediction, occurring alongside improvements in other aspects of SR devices, would greatly enhance their usability. The designers of SR devices should remain informed of the latest developments in this research area, especially if they aim to extend the stabilization over robot trajectories, and across a range of task scenarios.

Usage in Collaborative Tasks Estimation of the real-world efficacy of an SR device during development is challenging, and design limitations impose restrictions on the range of tasks that can be used for testing. In Chapter 6, we examined how well a user performed with the WRF in a close-range collaborative task consisting of assembly and pick-and-place operations.

Users interacted with the WRF in two conditions: a direct speech control mode where they gave commands to the robot at each stage, and a predictive robot mode where the WRF acted autonomously after inferring their intent. Based on an analysis of human-robot fluency and interview responses, we

found that that while users were able to perform better in the predictive mode for some metrics, they were not unanimous in their preference for it. Some users trusted the robot more in the direct control mode, as it afforded a greater sense of agency. This lack of trust is to be expected when interacting with SR devices in general, since they are not as ubiquitous as other forms of wearable devices, such as prostheses. As with social robots, long-term interaction studies need to be conducted with SR devices, especially those with some degree of autonomous capability. This would help in finding ways to accustom potential users to these devices, while also mitigating novelty effects.

The study task emulated a factory or warehouse setting, with the application space determined through the user-centered design process. While the WRF was found to be moderately useful in this situation, the study was conducted in a highly structured laboratory environment, for a task with relatively low inherent physical and cognitive demands. As a result, its conclusions may not generalize to real settings involving greater uncertainties in every aspect. With SR devices, it is important to both, apply the device in an appropriate task scenario, and to continuously iterate on its design through *in-situ* studies. This is especially true for devices with autonomy, as aspects other than physical design, such as predictive controllers, also need to to be updated according to the task. For instance, a follow up study with a device such as the WRF could be conducted in a more unstructured environment, e.g. a warehouse, where the user would walk between various stations and perform tasks together with the robot.

The larger goal in adding autonomy to the WRF was reducing the cognitive load on a user. While a human is generally able to adapt in interactions with

an autonomous robot, we aimed for the robot to be able to mutually adapt to the human as well. It was expected that this teaming would lead to a situation where the human would not need to think intentionally while performing structured tasks with the robot, or at least not expend as much effort as is needed for direct control. The process of predicting human intent for deciding robot actions was based on users' body movements during a task. Aside from improving the predictive models, sensing additional information, such as gaze tracking, would help in determining user intent. Advances in brain-computer interfaces may also play a role in allowing for seamless interaction with SR devices, and allow for their wider adoption.

Conclusion Through the process of designing the WRF, and formulating the associated systems for control and interaction, we identified various aspects of general robotics that need to be considered in the development of such devices.

Findings from the studies and analyses with the WRF motivate future research efforts that would also be applicable to other SR devices—system design that allows for rapid compensatory robot motion, improved real-time human motion prediction models, adaptive user intent prediction, and the eventual incorporation of a robotic augmentation into the user's body schema.

While these are some of the long-term goals in this area of research, we may also draw recommendations from the WRF for designers of SR devices within the purview of currently available technology—the focus needs to be on ergonomics, usability, and providing the most appropriate and desirable levels of autonomy for the robot, in a way that attempts to reduce both, physical, and cognitive loads on a user.

BIBLIOGRAPHY

[1] BBC News, Cairo toe earliest fake body bit. Available online (Accessed 02-Oct-2020). http://news.bbc.co.uk/2/hi/health/6918687.stm.

[2] Filtering MoCap data for real-time applications – project page, S. A. Skogstad, University of Oslo. Available online (Accessed 04-Sep-2020). https://tinyurl.com/y2box9jn.

[3] KIT Whole-Body Human Motion Database. Available online (Accessed 05-Sep-2020). https://motion-database.humanoids.kit.edu/.

[4] NASA-STD-3000 Man-System Integration Standards. Available online (Accessed 02-Feb-2017). https://msis.jsc.nasa.gov/.

[5] OptiTrack Motion Capture, Flex 13 camera. Available online (Accessed 05-Sep-2020). https://www.optitrack.com/products/flex-13/.

[6] ROBOTIS Dynamixel motor datasheets. Available online (Accessed 13-Mar-2019). http://en.robotis.com/.

[7] ROS Wiki: Unified Robot Description Format. Available online (Accessed 10-Sep-2017). http://wiki.ros.org/urdf.

[8] URDF ROS Human model. Available online (Accessed 01-Feb-2017). https://github.com/baxter-flowers/human_moveit_config.

[9] Elahe Abdi, Etienne Burdet, Mohamed Bouri, and Hannes Bleuler. Control of a Supernumerary Robotic Hand by Foot: An Experimental Study in Virtual Reality. *PLOS ONE*, 10(7):1–14, 07 2015.

[10] Pradeep KW Abeygunawardhana and Toshiyuki Murakami. Vibration Suppression of Two-Wheel Mobile Manipulator Using Resonance-Ratio-Control-Based Null-Space Control. *IEEE Transactions on Industrial Electronics*, 57(12):4137–4146, 2010.

[11] Priyanshu Agarwal, Pei-Hsin Kuo, Richard R Neptune, and Ashish D Deshpande. A novel framework for virtual prototyping of rehabilitation exoskeletons. In *2013 IEEE 13th International Conference on Rehabilitation Robotics (ICORR)*, pages 1–6. IEEE, 2013.

[12] Htrotugu Akaike. Maximum likelihood identification of gaussian autoregressive moving average models. *Biometrika*, 60(2):255–265, 1973.

[13] Mohammed Al-Sada, Thomas Höglund, Mohamed Khamis, Jaryd Urbani, and Tatsuo Nakajima. Orochi: Investigating Requirements and Expectations for Multipurpose Daily Used Supernumerary Robotic Limbs. In *Proceedings of the 10th Augmented Human International Conference 2019*, AH2019, New York, NY, USA, 2019. Association for Computing Machinery.

[14] Mochammad Ariyanto, Rifky Ismail, Joga Dharma Setiawan, and Zainal Arifin. Development of low cost supernumerary robotic fingers as an assistive device. In *2017 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, pages 1–6. IEEE, 2017.

[15] Rowland Atkinson and John Flint. Accessing Hidden and Hard-to-Reach Populations: Snowball Research Strategies. *Social Research Update*, 33(1):1–4, 2001.

[16] Neil M Bajaj, Adam J Spiers, and Aaron M Dollar. State of the Art in Artificial Wrists: A Review of Prosthetic and Robotic Wrist Design. *IEEE Transactions on Robotics*, 35(1):261–277, 2019.

[17] K Louise Barriball and Alison While. Collecting data using a semi-structured interview: a discussion paper. *Journal of Advanced Nursing*, 19(2):328–335, 1994.

[18] Brian C Becker, Robert A MacLachlan, and Cameron N Riviere. State estimation and feedforward tremor suppression for a handheld micromanipulator with a kalman filter. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5160–5165. IEEE, 2011.

[19] Joseph T Belter and Jacob L Segil. Mechanical design and performance specifications of anthropomorphic prosthetic hands: a review. *Journal of Rehabilitation Research and Development*, 50(5):599–618, 2013.

[20] Massimo Bergamasco and Hugh Herr. Human–Robot Augmentation. In *Springer Handbook of Robotics*, pages 1875–1906. Springer, 2016.

[21] H Russell Bernard. *Social Research Methods: Qualitative and Quantitative Approaches*. Sage, 2012.

[22] William Charles Bolton. *Mechatronics: Electronic Control Systems in Mechanical and Electrical Engineering*. Pearson UK, 2018.

[23] George EP Box, Gwilym M Jenkins, Gregory C Reinsel, and Greta M Ljung. *Time Series Analysis: Forecasting and Control*. John Wiley & Sons, 2015.

[24] John Brooke. SUS-a 'quick and dirty' usability scale. *Usability Evaluation in Industry*, pages 189–194, 1996.

[25] Michael Buhrmester, Tracy Kwang, and Samuel D Gosling. Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspectives on Psychological Science*, 6(1):3–5, 2011.

[26] Judith Bütepage and Danica Kragic. Detect, anticipate and generate: Semi-supervised recurrent latent variable models for human activity modeling. *arXiv preprint arXiv:1809.07075*, 2018.

[27] Richard H Byrd, Mary E Hribar, and Jorge Nocedal. An Interior Point Algorithm for Large-Scale Nonlinear Programming. *SIAM Journal on Optimization*, 9(4):877–900, 1999.

[28] Yi Cao, Ke Lu, Xiujuan Li, and Yi Zang. Accurate Numerical Methods for Computing 2D and 3D Robot Workspace. *International Journal of Advanced Robotic Systems*, 8(6):76, 2011.

[29] D. Carter-Davies, J. Chen, F. Chen, M. Li, and C. Yang. Mechatronic Design and Control of a 3D Printed Low Cost Robotic Upper Limb. In *2018 11th International Workshop on Human Friendly Robotics (HFR)*, pages 1–6, 2018.

[30] Min Chen, Stefanos Nikolaidis, Harold Soh, David Hsu, and Siddhartha Srinivasa. Trust-Aware Decision Making for Human-Robot Collaboration: Model Learning and Planning. *ACM Transactions on Human-Robot Interaction (THRI)*, 9(2), January 2020.

[31] John J Craig. *Introduction to Robotics: Mechanics and Control*, volume 3. Pearson Prentice Hall, 2005.

[32] James Cunningham, Anita Hapsari, Pierre Guilleminot, Ali Shafti, and A Aldo Faisal. The Supernumerary Robotic 3^{rd} Thumb for Skilled Music

Tasks. In *2018 7th IEEE International Conference on Biomedical Robotics and Biomechatronics (Biorob)*, pages 665–670. IEEE, 2018.

[33] Clark Davenport, Federico Parietti, and H Harry Asada. Design and Biomechanical Analysis of Supernumerary Robotic Limbs. In *ASME Dynamic Systems and Control Conference (DSCC), joint with JSME Motion and Vibration Conference (MoViC)*, pages 787–793, 2012.

[34] Bas J de Kruif, Emilio Schmidhauser, Konrad S Stadler, and Leonard W O'Sullivan. Simulation Architecture for Modelling Interaction Between User and Elbow-articulated Exoskeleton. *Journal of Bionic Engineering*, 14(4):706–715, 2017.

[35] Scott L Delp, Frank C Anderson, Allison S Arnold, Peter Loan, Ayman Habib, Chand T John, Eran Guendelman, and Darryl G Thelen. OpenSim: Open-source Software to Create and Analyze Dynamic Simulations of Movement. *IEEE Transactions on Biomedical Engineering*, 54(11):1940–1950, 2007.

[36] David A Dickey and Wayne A Fuller. Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association*, 74(366a):427–431, 1979.

[37] Zachary Dougherty and Ryder C Winck. Evaluating the Performance of Foot Control of a Supernumerary Robotic Limb. In *ASME Dynamic Systems and Control Conference (DSCC)*, volume V003T16A003. American Society of Mechanical Engineers, 2019.

[38] Anca D Dragan, Kenton CT Lee, and Siddhartha S Srinivasa. Legibility and Predictability of Robot Motion. In *2013 8th ACM/IEEE International Conference onHuman-Robot Interaction (HRI)*, pages 301–308. IEEE, 2013.

[39] David Feil-Seifer and Maja J Mataric. Defining Socially Assistive Robotics. In *2005 IEEE 9th International Conference on Rehabilitation Robotics (ICORR)*, pages 465–468. IEEE, 2005.

[40] Katerina Fragkiadaki, Sergey Levine, Panna Felsen, and Jitendra Malik. Recurrent Network Models for Human Dynamics. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 4346–4354, 2015.

[41] Olivier Friard and Marco Gamba. Boris: a free, versatile open-source

event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution*, 7(11):1325–1330, 2016.

[42] Benjamin Friedlander and Boaz Porat. The modified Yule-Walker Method of ARMA Spectral Estimation. *IEEE Transactions on Aerospace and Electronic Systems*, (2):158–173, 1984.

[43] John Q Gan, Eimei Oyama, Eric M Rosales, and Huosheng Hu. A complete analytical solution to the inverse kinematics of the Pioneer 2 robotic arm. *Robotica*, 23(1):123–129, 2005.

[44] M.A. Goodrich, D.R. Olsen, J.W. Crandall, and T.J. Palmer. Experiments in adjustable autonomy. In *Proceedings of IJCAI Workshop on Autonomy, Delegation and Control: Interacting with Intelligent Agents*, pages 1624–1629. Citeseer, 2001.

[45] VE Gough. Universal Tyre Test Machine. *Proceedings of 9th FISITA International Technical Congress*, pages 117–137, 1962.

[46] Sandra G Hart and Lowell E Staveland. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, 52:139–183, 1988.

[47] Richard S Hartenberg and Jacques Denavit. A kinematic notation for lower pair mechanisms based on matrices. *Journal of Applied Mechanics*, 77(2):215–221, 1955.

[48] John A Hartigan. *Clustering Algorithms*. John Wiley & Sons, Inc., 1975.

[49] Rex Hartson and Pardha S Pyla. *The UX Book: Process and guidelines for ensuring a quality user experience*. Elsevier, 2012.

[50] Archibald Vivian Hill. The heat of shortening and the dynamic constants of muscle. *Proceedings of the Royal Society of London. Series B-Biological Sciences*, 126(843):136–195, 1938.

[51] Guy Hoffman. Evaluating Fluency in Human–Robot Collaboration. *IEEE Transactions on Human-Machine Systems*, 49(3):209–218, 2019.

[52] Adam O Horvath and Leslie S Greenberg. Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology*, 36(2):223, 1989.

[53] Qiang Huang and Shigeki Sugano. Manipulator motion planning for stabilizing a mobile-manipulator. In *Proceedings of the 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, volume 3, pages 467–472. IEEE, 1995.

[54] David Huggins-Daines, Mohit Kumar, Arthur Chan, Alan W Black, Mosur Ravishankar, and Alexander I Rudnicky. Pocketsphinx: A Free, Real-Time Continuous Speech Recognition System for Hand-Held Devices. In *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*, volume 1. IEEE, 2006.

[55] Irfan Hussain, Leonardo Meli, Claudio Pacchierotti, Gionata Salvietti, and Domenico Prattichizzo. Vibrotactile haptic feedback for intuitive control of robotic extra fingers. In *Proceedings of the 2015 IEEE World Haptics Conference (WHC)*, pages 394–399, 2015.

[56] Irfan Hussain, Gionata Salvietti, Monica Malvezzi, and Domenico Prattichizzo. Design guidelines for a wearable robotic extra-finger. In *2015 IEEE 1st International Forum on Research and Technologies for Society and Industry Leveraging a better tomorrow (RTSI)*, pages 54–60, 2015.

[57] Ashesh Jain, Amir R Zamir, Silvio Savarese, and Ashutosh Saxena. Structural-RNN: Deep Learning on Spatio-Temporal Graphs. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5308–5317, 2016.

[58] Siddarth Jain and Brenna Argall. Probabilistic Human Intent Recognition for Shared Autonomy in Assistive Robotics. *ACM Transactions on Human-Robot Interaction (THRI)*, 9(1), 2019.

[59] Mrinal Kalakrishnan, Sachin Chitta, Evangelos Theodorou, Peter Pastor, and Stefan Schaal. STOMP: Stochastic Trajectory Optimization for Motion Planning. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 4569–4574, 2011.

[60] John F Kelley. An Iterative Design Methodology for User-Friendly Natural Language Office Information Applications. *ACM Transactions on Information Systems (TOIS)*, 2(1):26–41, 1984.

[61] R. Khodambashi, G. Weinberg, W. Singhose, S. Rishmawi, V. Murali, and E. Kim. User oriented assessment of vibration suppression by command shaping in a supernumerary wearable robotic arm. In *2016 IEEE-RAS 16th*

International Conference on Humanoid Robots (Humanoids), pages 1067–1072, 2016.

[62] SR Klemmer, B Hartmann, and L Takayama. How bodies matter: five themes for interaction design. In *Designing Interactive Systems (DIS 2006)*, pages 140–149, 2006.

[63] Kornelia Kulig, James G Andrews, and James G Hay. Human Strength Curves. *Exercise and Sport Sciences Reviews*, 12(1):417–466, 1984.

[64] Daniel A Kurek and H Harry Asada. The MantisBot: Design and impedance control of supernumerary robotic limbs for near-ground work. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5942–5947. IEEE, 2017.

[65] Steven M LaValle. *Planning Algorithms*. Cambridge university press, 2006.

[66] Leng-Feng Lee and Brian R Umberger. Generating optimal control simulations of musculoskeletal movement using OpenSim and MATLAB. *PeerJ*, 4:e1638, 2016.

[67] Sang-Won Leigh and Pattie Maes. Body Integrated Programmable Joints Interface. In *Proceedings of the 2016 ACM CHI Conference Extended Abstracts*, pages 3719–3722.

[68] Rensis Likert. A technique for the measurement of attitudes. *Archives of Psychology*, 1932.

[69] Baldin Llorens-Bonilla and H Harry Asada. Control and coordination of supernumerary robotic limbs based on human motion detection and task petri net model. In *ASME Dynamic Systems and Control Conference (DSCC)*, volume V002T27A006. American Society of Mechanical Engineers, 2013.

[70] Baldin Llorens-Bonilla and H Harry Asada. A robot on the shoulder: Coordinated human-wearable robot control using Coloured Petri Nets and Partial Least Squares predictions. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 119–125, 2014.

[71] Baldin Llorens-Bonilla, Federico Parietti, and H Harry Asada. Demonstration-based control of supernumerary robotic limbs. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 7–12, 2012.

[72] John YS Luh, Michael W Walker, and Richard PC Paul. On-Line Computational Scheme for Mechanical Manipulators. *Journal of Dynamic Systems, Measurement, and Control*, 102(2):69–76, 1980.

[73] James MacQueen. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 281–297, 1967.

[74] Azumi Maekawa, Kei Kawamura, and Masahiko Inami. Dynamic Assistance for Human Balancing with Inertia of a Wearable Robotic Appendage. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020.

[75] Christian Mandery, Ömer Terlemez, Martin Do, Nikolaus Vahrenkamp, and Tamim Asfour. The KIT Whole-Body Human Motion Database. In *2015 IEEE International Conference on Advanced Robotics (ICAR)*, pages 329–336. IEEE, 2015.

[76] Julieta Martinez, Michael J Black, and Javier Romero. On human motion prediction using recurrent neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2891–2900, 2017.

[77] Masahiro Mori. The uncanny valley. *Energy*, 7(4):33–35, 1970.

[78] Junichi Nabeshima, MHD Yamen Saraiji, and Kouta Minamizawa. Arque: Artificial Biomimicry-Inspired Tail for Extending Innate Body Functions. In *ACM SIGGRAPH 2019 Posters*, SIGGRAPH 2019, 2019.

[79] Pham Huy Nguyen, Curtis Sparks, Sai G Nuthi, Nicholas M Vale, and Panagiotis Polygerinos. Soft Poly-Limbs: Toward a New Paradigm of Mobile Manipulation for Daily Living Tasks. *Soft Robotics*, 6(1):38–53, 2019.

[80] Lael U Odhner, Raymond R Ma, and Aaron M Dollar. Open-loop precision grasping with underactuated hands inspired by a human manipulation strategy. *IEEE Transactions on Automation Science and Engineering*, 10(3):625–633, 2013.

[81] Edwin Olson. AprilTag: A robust and flexible visual fiducial system. In *2011 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3400–3407. IEEE, 2011.

[82] Teddy Ort, Faye Wu, Nicholas C Hensel, and H Harry Asada. Supernumerary Robotic Fingers as a Therapeutic Device for Hemiparetic Patients. In *ASME Dynamic Systems and Control Conference (DSCC)*, volume V002T27A010. American Society of Mechanical Engineers, 2015.

[83] James C Otis, Russell F Warren, Sherry I Backus, Thomas J Santner, and Jay D Mabrey. Torque production in the shoulder of the normal young adult male the interaction of function, dominance, joint angle, and angular velocity. *The American Journal of Sports Medicine*, 18(2):119–123, 1990.

[84] Federico Parietti. *Design and Control of Supernumerary Robotic Limbs*. PhD thesis, Massachusetts Institute of Technology, 2016.

[85] Federico Parietti and H Harry Asada. Supernumerary Robotic Limbs for Aircraft Fuselage Assembly : Body Stabilization and Guidance by Bracing. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1176–1183, 2014.

[86] Federico Parietti and Harry H Asada. Dynamic Analysis and State Estimation for Wearable Robotic Limbs Subject to Human-Induced Disturbances. In *2013 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3880–3887. IEEE, 2013.

[87] Christian I Penaloza and Shuichi Nishio. BMI control of a third arm for multitasking. *Science Robotics*, 3(20):eaat1228, 2018.

[88] Jose L Pons. *Wearable Robots: Biomechatronic Exoskeletons*. John Wiley & Sons, 2008.

[89] Domenico Prattichizzo, Monica Malvezzi, Irfan Hussain, and Gionata Salvietti. The Sixth-Finger: A modular extra-finger to enhance human hand capabilities. In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 993–998. IEEE, 2014.

[90] RV Ram, PM Pathak, and SJ Junco. Trajectory control of a mobile manipulator in the presence of base disturbance. *Simulation*, 2018.

[91] Knut Ringen, J Seegal, and A England. Safety and health in the construction industry. *Annual Review of Public Health*, 16(1):165–188, 1995.

[92] Cameron N Riviere, Jacques Gangloff, and Michel De Mathelin. Robotic

Compensation of Biological Motion to Enhance Surgical Accuracy. *Proceedings of the IEEE*, 94(9):1705–1716, 2006.

[93] Mark E Rosheim. *Robot Wrist Actuators*. John Wiley & Sons, Inc., 1989.

[94] MHD Yamen Saraiji, Tomoya Sasaki, Kai Kunze, Kouta Minamizawa, and Masahiko Inami. Metaarms: Body remapping using feet-controlled artificial arms. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology (UIST)*, pages 65–74, 2018.

[95] Katherine R Saul, Xiao Hu, Craig M Goehler, Meghan E Vidt, Melissa Daly, Anca Velisar, and Wendy M Murray. Benchmarking of dynamic simulation predictions in two software platforms using an upper limb musculoskeletal model. *Computer Methods in Biomechanics and Biomedical Engineering*, 18(13):1445–1458, 2015.

[96] Michael Schaefer, Hans-Jochen Heinze, and Michael Rotte. My Third Arm: Shifts in Topography of the Somatosensory Homunculus Predict Feeling of an Artificial Supernumerary Arm. *Human Brain Mapping*, 30(5):1413–1420, 2009.

[97] Lisa M Schutte, Mary M Rodgers, FE Zajac, and Roger M Glaser. Improving the Efficacy of Electrical Stimulation- Induced Leg Cycle Ergometry: An Analysis Based on a Dynamic Musculoskeletal Model. *IEEE Transactions on Rehabilitation Engineering*, 1(2):109–125, 1993.

[98] Max Schwarz and Sven Behnke. Compliant Robot Behavior Using Servo Actuator Models Identified by Iterative Learning Control. In *RoboCup: Robot Soccer World Cup*, pages 207–218. Springer, 2013.

[99] Homayoun Seraji. Configuration control of rover-mounted manipulators. In *1995 IEEE International Conference on Robotics and Automation (ICRA)*, volume 3, pages 2261–2266. IEEE, 1995.

[100] Ho Chit Siu, Ana M Arenas, Tingxiao Sun, and Leia A Stirling. Implementation of a Surface Electromyography-Based Upper Extremity Exoskeleton Controller Using Learning from Demonstration. *Sensors (MDPI)*, 18(2):467, 2018.

[101] Ståle A Skogstad, Sverre Holm, and Mats Høvin. Digital IIR filters with minimal group delay for real-time applications. In *2012 International Conference on Engineering and Technology (ICET)*, pages 1–6. IEEE, 2012.

[102] Ståle A Skogstad, Kristian Nymoen, Mats Høvin, Sverre Holm, and Alexander Refsum Jensenius. Filtering motion capture data for real-time applications. In *Proceedings of the International Conference on New Interfaces For Musical Expression*, pages 142–147, 2013.

[103] Carolyn Snyder. *Paper prototyping: The fast and easy way to design and refine user interfaces*. Morgan Kaufmann, 2003.

[104] Mark W Spong and Mathukumalli Vidyasagar. *Robot Dynamics and Control*. John Wiley & Sons, 2008.

[105] Doug Stewart. A Platform with Six Degrees of Freedom. *Proceedings of the Institution of Mechanical Engineers*, 180(1):371–386, 1965.

[106] Catherine A Sugar and Gareth M James. Finding the number of clusters in a dataset: An information-theoretic approach. *Journal of the American Statistical Association*, 98(463):750–763, 2003.

[107] Ilya Sutskever, Oriol Vinyals, and Quoc V Le. Sequence to Sequence Learning with Neural Networks. In Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 27, pages 3104–3112. Curran Associates, Inc., 2014.

[108] Robert I Sutton and Andrew Hargadon. Brainstorming groups in context: Effectiveness in a product design firm. *Administrative Science Quarterly*, pages 685–718, 1996.

[109] Leila Takayama, Wendy Ju, and Clifford Nass. Beyond Dirty, Dangerous and Dull: What Everyday People Think Robots Should Do. In *Proceedings of the 3rd ACM/IEEE International Conference on Human-Robot Interaction*, pages 25–32, 2008.

[110] Darryl G Thelen, Frank C Anderson, and Scott L Delp. Generating dynamic simulations of movement using computed muscle control. *Journal of Biomechanics*, 36(3):321–328, 2003.

[111] Deepak Tolani, Ambarish Goswami, and Norman I Badler. Real-Time Inverse Kinematics Techniques for Anthropomorphic Limbs. *Graphical Models*, 62(5):353–388, 2000.

[112] HF Machiel Van der Loos and David J Reinkensmeyer. Rehabilitation and

Health Care Robotics. In *Springer Handbook of Robotics*, pages 1223–1251. Springer, 2008.

[113] Vighnesh Vatsal and Guy Hoffman. Wearing your arm on your sleeve: Studying usage contexts for a wearable robotic forearm. In *The 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pages 974–980, 2017.

[114] Vighnesh Vatsal and Guy Hoffman. Design and Analysis of a Wearable Robotic Forearm. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5489–5496. IEEE, 2018.

[115] Vighnesh Vatsal and Guy Hoffman. End-Effector Stabilization of a Wearable Robotic Arm Using Time Series Modeling of Human Disturbances. In *ASME Dynamic Systems and Control Conference (DSCC)*, volume V001T05A001. American Society of Mechanical Engineers, 2019.

[116] Catherine Véronneau, Jeff Denis, Louis-Phillipe Lebel, Marc Denninger, Vincent Blanchard, Alexandre Girard, and Jean-Sébastien Plante. Multifunctional Remotely Actuated 3-DOF Supernumerary Robotic Arm Based on Magnetorheological Clutches and Hydrostatic Transmission Lines. *IEEE Robotics and Automation Letters*, 5(2):2546–2553, 2020.

[117] Takashi Wada, Masato Ishikawa, Ryohei Kitayoshi, Ichiro Maruta, and Toshiharu Sugie. Practical modeling and system identification of R/C servo motors. In *2009 IEEE Control Applications (CCA) & Intelligent Control (ISIC)*, pages 1378–1383. IEEE, 2009.

[118] Faye Wu and Harry H Asada. Supernumerary robotic fingers: an alternative upper-limb prosthesis. In *ASME Dynamic Systems and Control Conference (DSCC)*, volume V002T16A009. American Society of Mechanical Engineers, 2014.

[119] De Xu, Carlos A Acosta Calderon, John Q Gan, Huosheng Hu, and Min Tan. An Analysis of the Inverse Kinematics for a 5-DOF Manipulator. *International Journal of Automation and Computing*, 2(2):114–124, 2005.

[120] Peter Young and Anthony Jakeman. Refined instrumental variable methods of recursive time-series analysis Part III. Extensions. *International Journal of Control*, 31(4):741–764, 1980.