Haptic Perception, Decision-making, and Learning for Manipulation with Artificial Hands

by

Randall Blake Hellman

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved July 2016 by the
Graduate Supervisory Committee:

Veronica J. Santos, Chair
Panagiotis K. Artemiadis
Spring Berman
Stephen I. Helms Tillery
Georgios Fainekos

ARIZONA STATE UNIVERSITY

August 2016

ABSTRACT

Robotic systems are outmatched by the abilities of the human hand to perceive and manipulate the world. Human hands are able to physically interact with the world to perceive, learn, and act to accomplish tasks. Limitations of robotic systems to interact with and manipulate the world diminish their usefulness. In order to advance robot end effectors, specifically artificial hands, rich multimodal tactile sensing is needed. In this work, a multi-articulating, anthropomorphic robot testbed was developed for investigating tactile sensory stimuli during finger-object interactions. The artificial finger is controlled by a tendon-driven remote actuation system that allows for modular control of any tendon-driven end effector and capabilities for both speed and strength. The artificial proprioception system enables direct measurement of joint angles and tendon tensions while temperature, vibration, and skin deformation are provided by a multimodal tactile sensor. Next, attention was focused on real-time artificial perception for decision-making. A robotic system needs to perceive its environment in order to make decisions. Specific actions such as "exploratory procedures" can be employed to classify and characterize object features. Prior work on offline perception was extended to develop an anytime predictive model that returns the probability of having touched a specific feature of an object based on minimally processed sensor data. Developing models for anytime classification of features facilitates real-time action-perception loops. Finally, by combining real-time action-perception with reinforcement learning, a policy was learned to complete a functional contour-following task: closing a deformable ziplock bag. The approach relies only on proprioceptive and localized tactile data. A Contextual Multi-Armed Bandit (C-MAB) reinforcement learning algorithm was implemented to maximize

cumulative rewards within a finite time period by balancing exploration versus exploitation of the action space. Performance of the C-MAB learner was compared to a benchmark Q-learner that eventually returns the optimal policy. To assess robustness and generalizability, the learned policy was tested on variations of the original contour-following task. The work presented contributes to the full range of tools necessary to advance the abilities of artificial hands with respect to dexterity, perception, decision-making, and learning.

DEDICATION

To my family, for their love and support.

To the memory of friend and colleague Kevin Bair.

ACKNOWLEDGMENTS

I am extremely grateful to be in this position. There are many people to whom I owe a debt of gratitude as they have helped shape me into the scientist and engineer that I am today.

First to Dr. Veronica Santos, this would not have been possible without your vision and guidance. You have been a wonderful mentor throughout my studies and I am truly grateful.

I would like thank my committee members Dr. Panagiotis Artemiadis, Dr. Spring Berman, Dr. Georgios Fainekos, and Dr. Stephen Helms Tillery. I am very appreciative of your guidance. I am grateful for the classes, seminars, research, and your overall energy that you have put into shaping my experiences during my studies.

I would like to thank Dr. Mihaela van der Schaar and Dr. Cem Tekin for their guidance and discussions on reinforcement learning during the work developed in Chapter 4.

The material presented herein was supported by multiple funding agencies. I would like to thank the National Science Foundation under award numbers 0932389, 0954254/ 1461547, and 1208519/ 1463960, and the Fulton Schools of Engineering University Graduate Fellowship.

I would like to thank my fellow lab members Dr. Michael De Gregorio, Ryan Manis, Dr. Ruben Ponce Wong, and Ismael Reveles. I have many fond memories of our time both in and out of the Biomechatronics Lab at Arizona State University. I would also like to thank the lab members from the Biomechatronics Lab at the University of California, Los Angeles for discussing the work developed in Chapter 4.

TABLE OF CONTENTS

Page

CHAPTER Page

CHAPTER Page

CHAPTER Page

CHAPTER Page

CHAPTER Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

CHAPTER 1

INTRODUCTION

Motivation

Robot end effectors, artificial hands, and prosthetics are commonly benchmarked against the abilities of the human hand. The human hand is able to perceive, act and learn seamlessly in its environment through the integration of multiple sensory systems such as auditory, visual, proprioceptive and especially tactile stimulation. When learning a new manipulation task, the primary feedback modality is most often visual; one observes a deviation from a goal along with the success or failure of the task. While this is happening, visual cues are also being related to tactile and proprioceptive stimulation. Once the manipulation task has been successfully learned, it can often be completed with only tactile and proprioceptive feedback (Klatzky & Lederman, 1999). The task becomes almost automatic, requiring minimal attention (Johansson & Flanagan, 2009). This reduction of cognitive load happens only once we learn the nuances of a task and it no longer requires additional sensory oversight. These learned nuances of a manipulation task are then stored as experience with expected results. This allows us to learn similar and overlapping tasks quickly as our manipulation abilities improve.

While the kinematic capabilities of prosthetic and artificial hands are important, without sensory, or afferent, signals their usefulness becomes questionable. The majority of currently available prosthetic hands lack rich tactile sensing and the user must rely on visual and aural feedback. For a thorough review of current capabilities of common anthropomorphic prosthetic hands see (Belter & Dollar, 2011). It is important to note that while we are discussing the kinematic and sensory abilities of prosthetic devices, the

aesthetics are also of great importance although not the focus of this work. The necessity of sensory feedback is shown from the results of a study on the motor function of a man deafferented by severe peripheral sensory neuropathy (Rothwell et al., 1982). While the subject still has full motor function of his arm and hand, he is able to complete certain dexterous tasks primarily through visual feedback. However, without the use of vision, dexterous tasks quickly degrade: he is unable to hold a cup, button his shirt, or even judge the weight of objects. This shows the importance of afferent sensory signals in the functionality of our hands and reinforces the importance of sensory feedback for capable and effective anthropomorphic prosthetic hands. Additionally, it is important not to overwhelm the amputee with unintuitive sensory feedback. If sensory stimuli are non-intuitive, it can become distracting or even habituated (Jimenez & Fishel, 2014). Preprocessing sensory feedback in an effort to recreate reflex responses or simplify sensory perception to the user of a prosthetic device could reduce the cognitive load and improve the functionality and usefulness of the device.

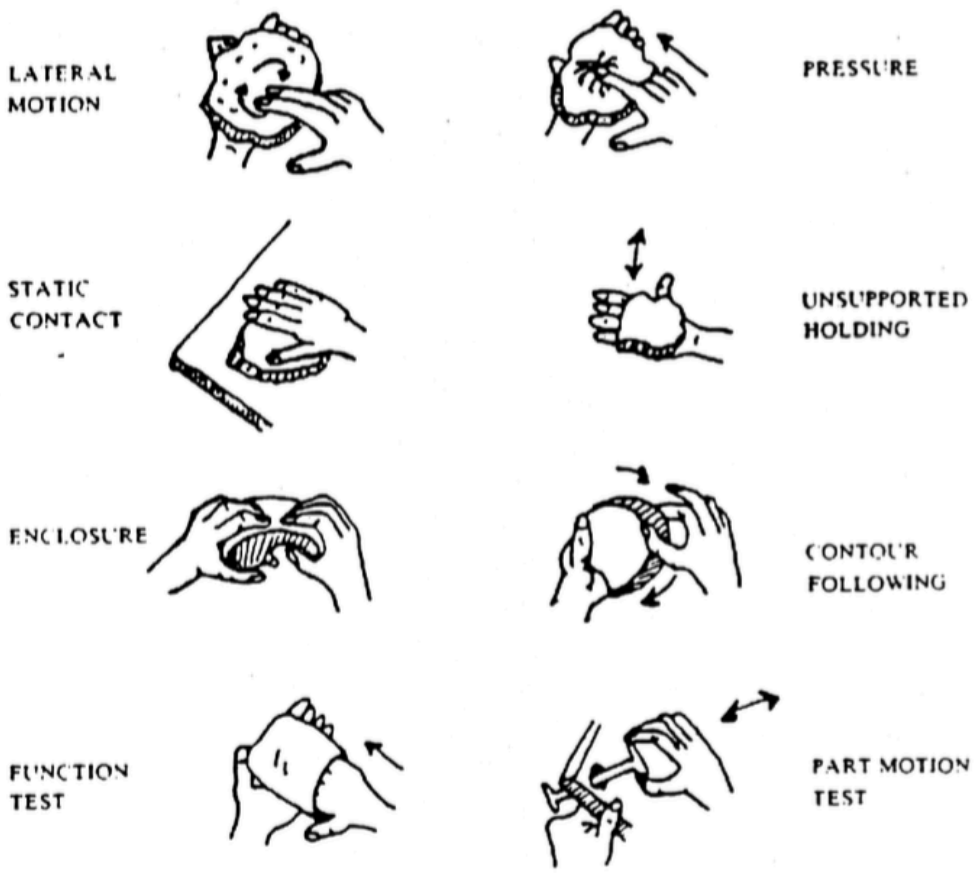

Figure 1.1 Typical set of "exploratory procedures" for haptically determining object properties. Excerpt from (Lederman & Klatzky, 1987).

When learning to haptically perceive object features, humans rely on "exploratory procedures" (EPs) that are sets of motions used to determine specific properties (Lederman & Klatzky, 1987). Even while using vision as the primary feedback modality for learning, there are properties of objects that remain ambiguous through vision alone. This ambiguity arises because properties such as hardness, texture, or other visually obscure features cannot be ascertained through vision alone. In order to more accurately determine these features and reduce uncertainty, "exploratory procedures" are used. Based on the feature of interest, an appropriate exploratory action or set of actions will be

chosen. For example, pressure is used to determine hardness while lateral motions along a small region of the object are used to determine texture (*Figure 1.1*). These action choices are based on previous knowledge developed from years of learning action-perception relationships.

Current robotic systems have a limited ability to interact with the world. Developing robotic systems that can haptically interact and learn from the environment will advance the capabilities of the field. By advancing the abilities of artificial perception, learning, and manipulation, the functionality and usefulness of robotic systems and prosthetic devices will progress.

Contributions

The work presented is focused on advancing the abilities of artificial hands with respect to dexterity, perception, and decision-making. This dissertation is comprised of three separate studies and a discussion on future work. **Chapter 2** is focused on the development of a highly dexterous anthropomorphic testbed with rich tactile sensing. The testbed is tendon-driven by a remote actuation system. Two actuation systems are presented, one with zero-backlash (Hellman & Santos, 2012) and the second with a significantly streamlined design for cost reduction and manufacturability (Hellman, Chang, Tanner, Helms Tillery, & Santos, 2015). Both systems are characterized and shown to be capable of human-like speed and strength. Applications of the testbed are then discussed for the rehabilitation of upper-limb impairment or loss.

Chapter 3 explores the use of probabilistic models for real-time haptic perception of geometric object features. Prior work was done to develop models for perceiving object features, but complete fingertip trajectories were needed (R. D. Ponce Wong, Hellman, & Santos, 2014). The work presented here resulted in anytime classification of edge orientation on partial information with minimally processed tactile data (Hellman & Santos, 2015).

Chapter 4 describes the development of capabilities for a robotic system to perform a functional contour-following task. Perception and decision-making are needed to successfully complete the task. Perception during short end effector trajectories is useful for identifying deviations from successful task execution and taking frequent corrective actions (Hellman & Santos, 2016). Trained state classifiers are robust to novel contour-following tasks that were not used during training. Reinforcement learning is

used to determine the best policy with both Contextual Multi-Armed Bandits and Q-learning algorithms.

Chapter 5 summarizes the contributions of this dissertation and discusses applications and future directions.

CHAPTER 2

ROBOTIC HAND TESTBED TO AID IN THE DEVELOPMENT OF HAPTIC ALGORITHMS FOR ARTIFICIAL PERCEPTION AND MANIPULATION

Introduction

Commercially available multi-articulating artificial hands use an intrinsic actuation system in which motors reside in the palm or at finger joints themselves (Belter, Segil, Dollar, & Weir, 2013; Controzzi, Cipriani, & Carrozza, 2014). As a result, these hands are either limited in movement speed or grip strength. For the purposes of a testbed, the "BairClaw" presented in this chapter along with multiple actuation systems was designed to have an extrinsic, tendon-driven, remote actuation system to enable human-like speeds and grip strengths while maintaining the small volume and form factor of a human-sized hand. The testbed was designed for a maximum fingertip force of 44.5 N (10 lbf) and maximum individual tendon tensions of 111 N (25 lbf). The maximum fingertip force was selected to be consistent with human capabilities for opposition pinch and single-digit force production against a surface (Keenan & Massey, 2012; Swanson, Matev, & De Groot, 1970). The maximum individual tendon tension was estimated based on planned BairClaw kinematics and overall dimensions. As with the human hand and its extrinsic muscles, the BairClaw's extrinsic actuation system resides proximal to the wrist and transmits multi-articular joint torques using tendons.

Currently there are multiple robotic hands that have similar kinematic capacities to that of human hands but are limited in their sensing capabilities. Chapter 1 discussed the currently available robot systems for manipulation while highlighting their capabilities. To advance the area of robot manipulation, a bench-mounted testbed capable of dexterous manipulation and rich tactile sensing has been developed.

The robot testbed is comprised of an actuation system and a highly dexterous robot finger. The robot finger is extrinsically actuated by tendons that are connected through a custom built transmission system. Extrinsic actuation allows the size and weight of the robot finger to remain anthropomorphic.

The majority of this chapter is from published work (Hellman et al., 2015; Hellman & Santos, 2012).

Robot Testbed

Actuation Systems

The design of human-sized artificial hands is challenged by the relatively small volume available for sensor, actuator, and plant components. An intrinsic actuation approach (with motors in the fingers and palm) results in a compact and elegant form (e.g., (Liu et al., 2008)), but often at the expense of finger motion and force production capabilities. When both strength and speed are desired, an extrinsic actuation approach can be taken similar to that of the human hand. In addition to smaller intrinsic muscles in the palm, powerful extrinsic muscles in the forearm transmit torques to finger joints through tendons and extensor hoods that pass over the joints (Tubiana, 1981). For practical reasons, many robot hands have actuators located proximal to the wrist and rely on cables to transmit torques about finger joints. Such tendon-driven designs have been used for underactuated (Dollar & Howe, 2010) and robotic hands (Lotti et al., 2005; Loucks, Johnson, Boissiere, Starr, & Steele, 1987; Lovchik & Diftler, 1999; Nahvi, Hollerbach, Yangming Xu, & Hunter, 1994; Shadow Robot Company, 2011), prosthetic hands (Dalley, Wiste, Withrow, & Goldfarb, 2009), and complex, anthropomorphic systems (Gialias & Matsuoka, 2004; Grebenstein, Chalon, Hirzinger, & Siegwart, 2010).

Here we present a novel design for a rotary motor-based actuation system for a single revolute joint of a tendon-driven robotic hand. The actuation system was designed for a double-acting actuation scheme ("N-type") in which one motor can control both rotation directions of a single revolute joint (Jacobsen, Ko, Iversen, & Davis, 1989). This "push-pull" capability reduces hardware and control complexity. However, the modular actuation system could be duplicated and used in a single-acting actuation scheme ("2N-

type") in which two motors actuate a single joint using an agonist/antagonist control scheme if co-contraction is desired. While the actuation system can be applied to any tendon-driven mechanism, the design criteria were created with a human-sized, anthropomorphic robot hand in mind. The actuation system is designed for use in a robotic hand testbed that allows for seamless transitions between different tendon-driven end effectors.

First Generation Zero Backlash, Back-Driveable Actuation System

I. Methods

A. Design Criteria

Strength, speed, precision, and robustness are key features to consider in the design of remote actuation systems (Melchiorri & Kaneko, 2008) for tendon-driven robotic hands. These features affect the ability of the fingertips to achieve the responsiveness necessary for implementation of artificial grip reflexes and the dexterity required for object manipulation during grasp. In particular, the use of flexible-link transmission systems (e.g., pulley- and sheath-routed tendons) adds to the design challenge (Melchiorri & Kaneko, 2008).

The actuation system must have sufficient bandwidth and minimal phase delay so as to produce a fingertip force with minimal lag. Furthermore, proprioception capabilities are especially important. Motor encoders can provide tendon displacement and velocity measurements that, coupled with a known Jacobian matrix for the robot hand, could supplement joint angle and velocity data provided by the robot hand itself. In some cases, motor controllers can also provide motor current measurements as a method to determine

joint torques. However, the desired method to determine joint torques will be through the use of uniaxial load cells in series with each tendon. In the actuation system presented here, the compound pulley reduction produces a desired high tendon force output along with zero backlash back-drivability (*Figure 2.1*). The tendon preload and stiffness mechanisms, as well as the tendon tension sensors, are all located alongside a single Maxon EC-max 30 (30 mm diameter), 60 watt motor (Maxon Precision Motors, Inc.).

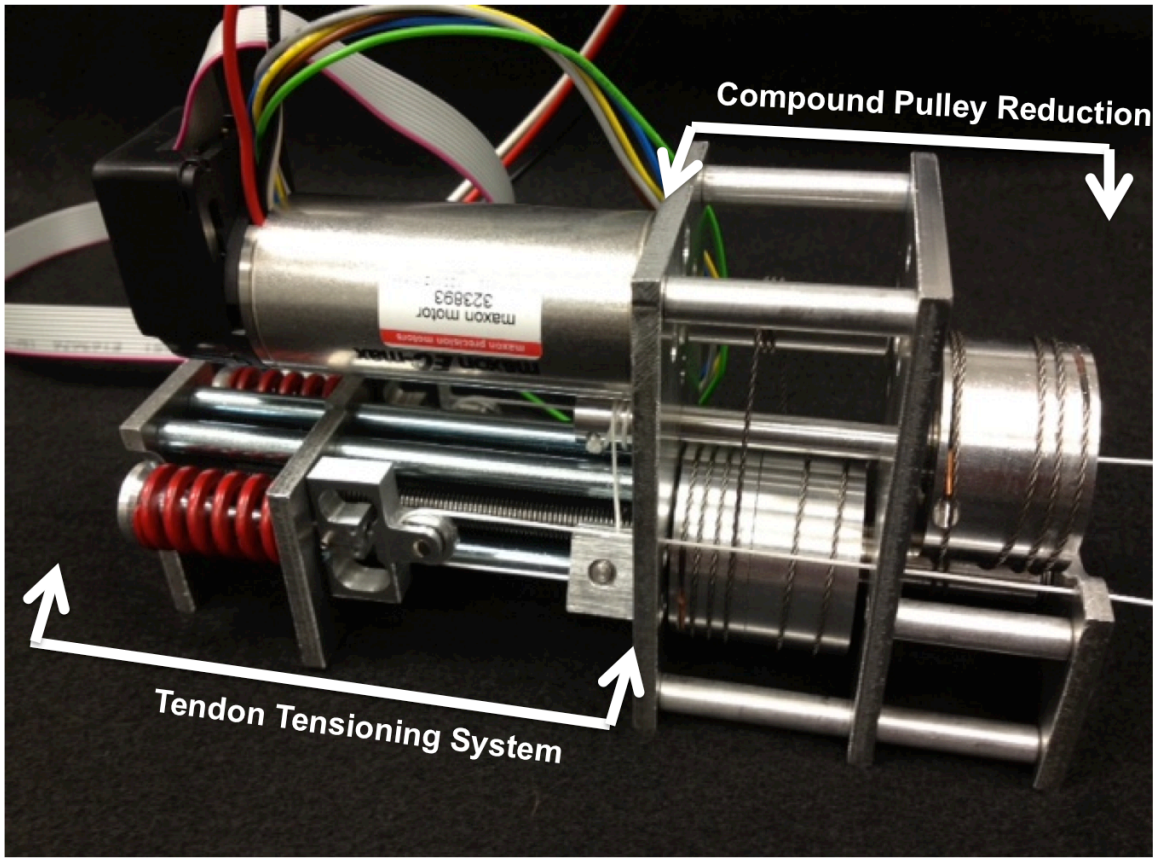

Figure 2.1 Actuation system. The right half of the assembly contains a two-stage, zero backlash, 12:1 compound pulley reduction. Below the motor on the left are the tendon stiffness and preload mechanisms along with the inline load cells for measuring tendon tension.

High factors of safety were used in the design of the actuation system in order to minimize performance limitations that might prohibit fast, reflex-like movements or functionally meaningful fingertip forces. Importantly, we are able to reduce the maximum motor current in order to accommodate different tendon-driven end effector designs without the actuation system itself limiting the performance of the overall robotic system. The motor bank is capable of achieving high forces that exceed the suggested maximum tendon tension of the Shadow Dexterous Robot Hand (Shadow Robot Company). To prevent damage to the end effector, a mechanical tendon fuse can be used as a secondary precaution in addition to software limitations on motor current. The mechanical fuse will be located serially in between the input tendons of the robotic hand and the output tendons of the actuation system. If the system were unexpectedly overloaded and the springs in the actuation system were unable to absorb the load, the mechanical fuse would fail, thereby sparing the robot hand from damage.

1) Compound pulley reduction

Table 2.1 Design specification for actuation system.

Actuation System Specifications	
Tendon Force	
Dynamic Operating Range	58 N (13 lbf)
Motor Stall Torque (Max Tendon Force)	533 N (120 lbf)
Tendon Displacement	5 cm
Tendon Tensioner Adjustment	14 cm
Module	
Dimensions	161 mm x 87 mm x 50 mm
Weight	834 g

The motor module has a two-stage, 12:1 zero backlash compound pulley reduction from the motor shaft to the final output shaft (*Figure 2.2*). Based on the 60 watt specifications, stall torque of the Maxon EC-max motor will sustain tendon tensions at the output shaft of 533N (120 lbf) with a nominal dynamic load of 58N (13 lbf) (*Table 2.1*). Although pulleys are mounted on ball bearing-supported shafts, the actual tendon forces will be slightly less due to frictional losses in the system. The fingertip force will vary depending on the actuation Jacobian that is unique to the design of the end effector that is in use.

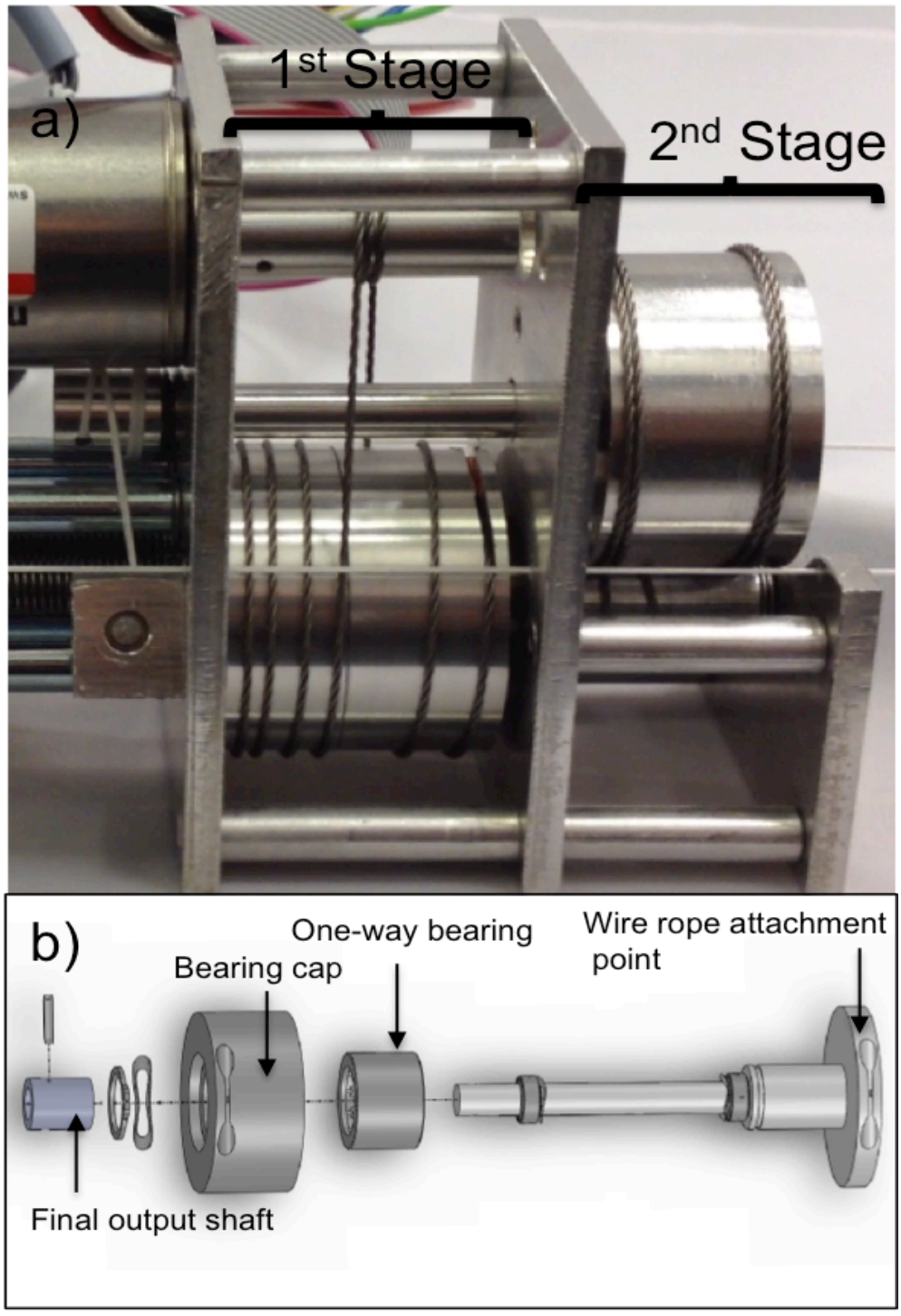

Figure 2.2 Compound pulley reduction. a) Close up photo of 1st and 2nd stages of the compound pulley reduction. b) Exploded view of the "secondary pulley," in reference to the second reduction of the compound reduction).

Each stage of the independent pulley reduction consists of a double-wrapped wire rope (1/32" diameter, 3x7 hollow-core, McMaster-Carr #3458T51) that has both of its attachment points on the main pulley and is double-wrapped around the pinion. The double-wrap around the pinion ensures sufficient friction such that the wire rope will not slip over the pinion as it "walks" during pinion rotation. Inspired by tendon tensioning mechanisms in the Barrett WAM (Barrett Technology), tension in each independent loop of wire rope is maintained with a one-way bearing housed inside of a larger pulley, which maintains a desired preload in the wire rope (*Figure 2.2*b). Crimped copper stop sleeves (McMaster-Carr #3897T31) at each end of the wire rope are held in place by grooves at wire rope attachment points on each half of the two-part pulley.

From the motor output shaft to the final output shaft, the system has zero backlash and is back-driveable. Zero backlash allows angular position/velocity and torque of the motor output shaft to be transformed according to the 12:1 reduction ratio with little loss at the final output shaft. Zero backlash also allows for back-driveable control of individual joints which simplifies the design of the actuation system as well as its modeling and control. The zero backlash pulley reduction and tensioning provided by a one-way bearing in each pulley serve to maintain a preload in the wire rope and prevent slack and unspooling. The relative rotation between each stage of the pulley reduction is kept fixed through the use of inextensible wire rope that is robust to creep. When first assembling, it takes a few cycles before the wire rope attains a stable position and rewraps in the same location. Once the desired wrap around the pulley is achieved, the rewrap position stays constant.

2) Tendon tensioning system

After the 12:1 compound pulley reduction, the final output shaft transitions to the tendons that actuate the robot hand. The tendon is comprised of a 0.7 mm diameter, monofilament fishing line (200 lbf Spectra line, Power Pro), line will be routed through low friction sheathing having an inner diameter of 1.2 mm and wall thickness of 0.3 mm. From the output shaft, the tendons are routed across two pulleys to enable the use of mechanisms for pre-tensioning the tendons, maintaining tendon tension and absorbing unexpected loads during use, and directly measuring tendon tension (*Figure 2.3*). Due to the routing configuration of each tendon across a single pulley that attaches to a custom uniaxial load cell, the load cell measurements will be equal to twice the tendon tension. The load cell measurements can be used to monitor tendon tension directly and calculate finger joint torques.

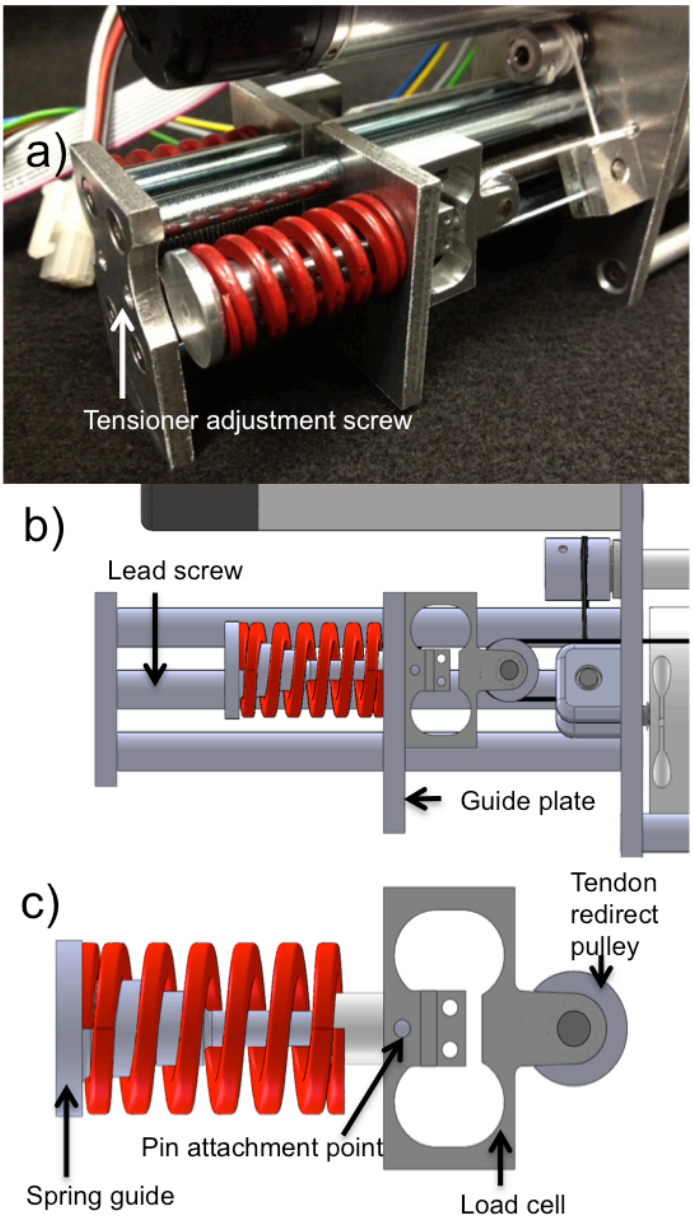

Figure 2.3 Tendon tensioner, stiffness, and load cell. Tendon stiffness is determined by the modular spring (shown in red). Tendon preload and tension is adjusted by moving the guide plate along the lead screw with the tension adjustment screw (a-b). A custom uniaxial load cell is attached to the spring guide with a 1/16" diameter spring pin that constrains the assembly c) to the guide plate (a-b) for direct measurement of tendon tension.

Custom uniaxial load cells were designed specifically for the expected tendon force range of 0-530 N (0-120 lbf) with a resolution of 0.01 N (0.04 lbf). Each load cell has a full Wheatstone bridge with all strain gauges located on the interior of the load cell. The signal is amplified locally with an op-amp mounted to the load cell and a shield will cover all electrical components. Shielding on the load cell is required because the Maxon motor is in close proximity and will otherwise introduce significant noise into the load cell data. Each load cell serves as the main rigid link between the tendon and a die spring that maintains tendon stiffness.

The load cell is mounted on the opposite side of a guide plate and pinned to a die spring guide shaft that determines the tendon response to unanticipated loads and sets/maintains initial tendon preload and stiffness. The load cell itself is secured to an adjustable guide plate with two guide rails and a lead screw (*Figure 2.3*b). Left-right adjustment of the guide plate (and spring-loaded tensioning mechanism) via rotation of the lead screw allows for changes in tendon lengths of up to 14 cm and tendon preload once tendon slack has been taken up. This tendon preload feature reduces overall setup time when an end effector is newly attached and simplifies adjustments that may be necessary due to the development of slack in tendons within the robot hand. Tendon tension will be incorporated into the control structure for determining finger joint torques and as a safety feature to prevent joint failure or system damage.

An early design of the tendon tensioning assembly included a simple mechanism for achieving non-linear stiffness (as in (Petit et al., 2010)) where the tendon passed over a spring and, as the spring compressed, the angle of action over the spring changed. For compactness of design and to keep the spring tensioning and tendon routing in a single

plane (reducing hardware needs for redirecting tendons) we created the design presented here. The die spring in the modular tendon tensioning assembly can be changed to suit the tendon stiffness requirements of the user. This will affect the end effector's response to unanticipated loads. For instance, the user may want different rates of compliance for flexion compared to that of extension. The prototype described here uses a 60 lbf/in die spring (McMaster-Carr #9584K15) for each of the flexor and extensor tendons. Alternate setups are also possible with the spring guide to achieve a composite, non-linear spring stiffness. A non-linear response can be achieved by nesting a preloaded spring and disc washer within the main spring. The interior spring would be held in place by disc washers held against the stepped features of the spring's guide shaft (*Figure 2.3*c).

B. Key Features

Both flexion and extension are achieved with a single actuation unit, with one actuator per revolute joint. Tendon stiffness and pretension can vary depending on the assigned joint and desired response to unanticipated loads. The zero-backlash back driveable motor module simplifies the design along with reducing the size of the motor bank necessary to control all degrees of freedom of the end effector. Use of a one-way bearing in the pulley allows the motor module to be compact and robust. The desired preload maintained with the one-way bearing ensures that the wire wrapped pulley will not unspool even in the case of an unanticipated load or rapid rotation.

The custom uniaxial loads cells are designed specifically for tendon loads for robotic grasp applications. Load cells are attached to a spring guide with a 1/16" diameter spring pin. The large diameter cap of the spring guide shaft constrains the die spring

against the guide plate, allowing the load cell to move as a unit under high tendon loads. When one tendon, say a flexor, is loaded, a net torque will develop at the joint, the flexor's die spring will compress, and the extensor's die spring will extend and take up the slack in the extensor tendon.

The ease of setup between various end effectors is a major advantage of the actuation system. Such a modular setup will enable experiments with different end effectors, each with unique proprioceptive and tactile sensing capabilities, and the testing of similar control policies on different plants.

Coupling between the actuation system and the robot hand is done by linking tendons via fishing-style swivel clips (100 lbf SPRO Power Swivels). Attachments are made with the system in slack and then each individual tendon is brought to a desired preload with the lead screw tendon tensioning mechanism.

C. Control Setup

The motor module allows for multiple control structures including position, velocity, force, and impedance (by varying the motor current). The Maxon EPOS2 24/5 Motor Controller has sampling rates for position control of 1 kHz and current control of 10 kHz. With a quad count encoder the motor position can theoretically be controlled with a resolution of 1/2000 counts per revolution (1.3 μm of tendon displacement). However, in practice, the position control resolution is on the order of 1/1000 counts per revolution and actual tendon displacements are not likely to be on the order of microns.

Depending on the context of the grasp or manipulation task, the control structure can be switched on the fly. For automated closure of the hand around a novel object, the motor module can be set to a compliance mode which uses limits on joint torques for safety, as in commercially-available prosthetic hands such as the Touch Bionics i-limb ("Touch Bionics i-limb hand," 2009). Furthermore, by controlling motor current one can implement impedance control for each digit (Hogan, 1985).

Maxon EPOS2 24/5 controllers also have built-in position profile modes with limited variability. Position profile mode allows the user to set the minimum and maximum motor acceleration and uses a trapezoidal velocity profile. The motor controllers communicate through an internal CANOpen (Control Area Network) network between the individual controllers. The master controller on the CAN network is capable of being connected to the central computer by either USB or the more robust industrial standard of RS-232.

The central computer interfaces with the motors through the use of a Maxon command library in C++. There are multiple options for communication with Maxon devices, but C++ was selected as it the main language used for additional pieces of equipment, such as the Shadow Dexterous Robot Hand and multimodal BioTac tactile sensors (Syntouch, LLC) (Wettels, Santos, Johansson, & Loeb, 2008; Wettels et al., 2008).

We can also measure compliance indirectly by comparing joint-based Hall effect sensor measurements and motor position encoder measurements. Such comparative information can be used in the control loop or as a safety fault in case the actuation and hand system experiences tendon slip or creep.

D. Preliminary Dynamic Testing

The dynamic response of a single motor actuation system was tested with 500 g masses attached to each of the two output tendons. The actuation system was clamped to the edge of a workbench and the masses were allowed to hang freely over the edge of the workbench. The motor was controlled in position control mode to produce sinusoidal waves having an amplitude of 180° of rotation at the motor output shaft, which is equivalent to 1.3mm of tendon excursion. A tendon excursion of 1.3mm is approximately equal to 15° of rotation at a joint having a 10 mm diameter actuation pulley.

The sine wave was implemented for the following frequencies: 3.33 Hz, 4 Hz, 5 Hz, 6.67 Hz, 10 Hz, and 20 Hz. Due to sensor limitations in the current experimental setup, encoder position sampled at 100 Hz was subjected to a forward and reverse low pass filter that does not distort phase (MATLAB's "filtfilt" function) and then differentiated twice to enable comparisons between the acceleration of the motor output shaft and that of the 500 g mass hanging from the end of a single tendon. Acceleration of the 500 g mass was recorded at 150 Hz using a six degree-of-freedom inertial measurement unit (MEMSense IM05-0600C050T00).

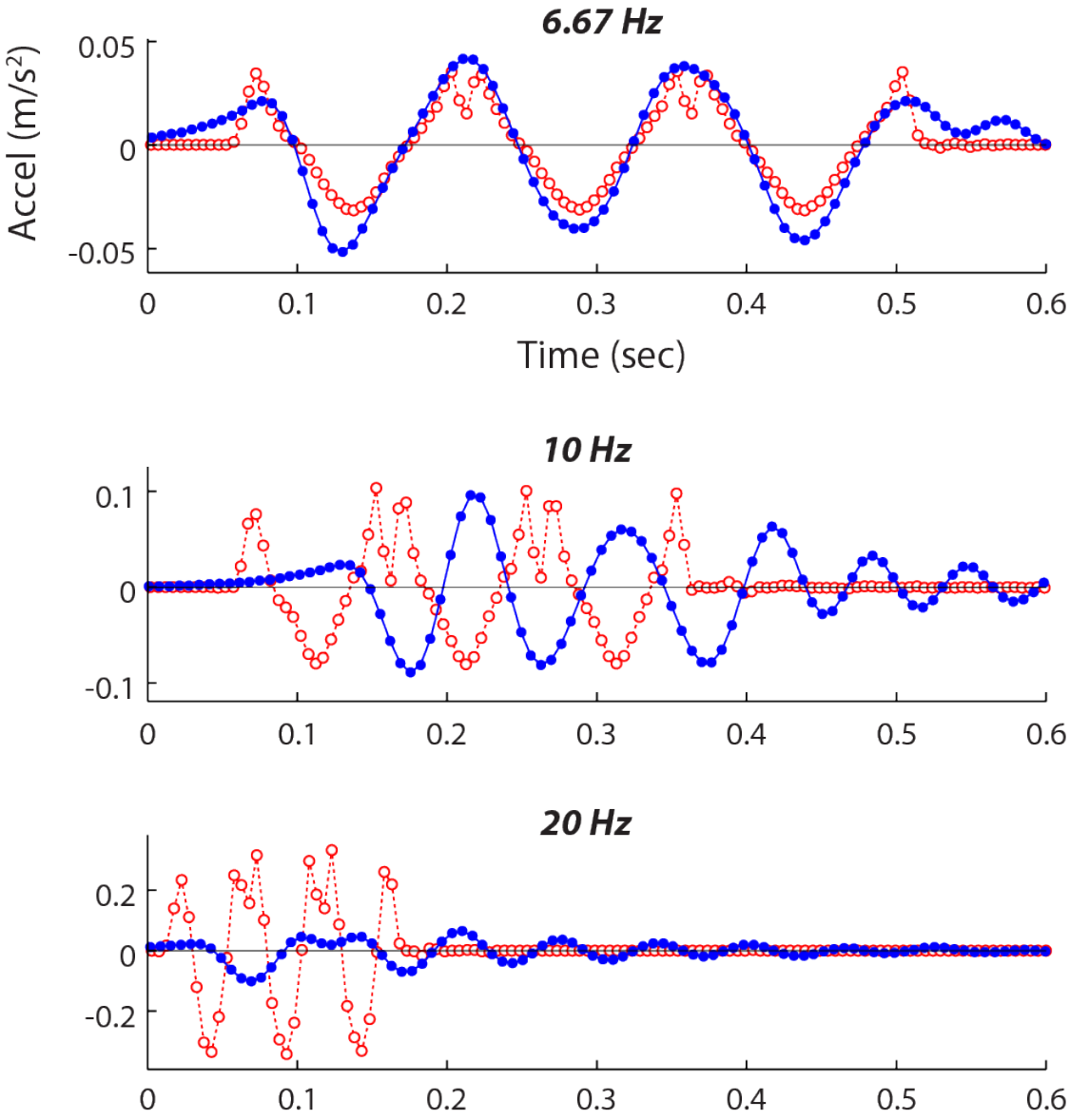

Figure 2.4 Dynamic response of compound pulley reduction and tendon tensioning system. At a frequency of 6.67 Hz (top), the magnitude of the acceleration of the 500 g free-hanging mass (closed circles, solid line) is comparable to that of the motor shaft output (open circles, dotted line). Decreases in the system gain are shown for 10 Hz (middle) and 20 Hz (bottom). Phase shift is estimated for visualization purposes only.

Initial dynamic characterization was conducted on a single actuation unit, specifically the dynamic response of the compound pulley reduction and tendon tensioning system. Due to limitations of the experimental setup, we can only reliably report on input and output magnitudes at this time (*Figure 2.4*). At a frequency of 6.67 Hz, the gain of the compound pulley reduction and tendon tensioning system is approximately one. At a frequency of 10 Hz, a decrease in system gain occurs until the gain drops severely around 20 Hz.

The actuation system was also used to directly actuate a single joint in the 18 degree-of-freedom Shadow Dexterous Robot Hand (*Figure 2.5*). For initial testing, a single actuation unit was used to actuate the proximal interphalangeal (PIP) joint of the index finger directly. The distal interphalangeal (DIP) joint, coupled to the PIP joint, also moved under actuation.

The attachment of the Shadow Dexterous Robot Hand tendons and removal of tendon slack were simple due to the lead screw tendon preloading mechanism. Using position control mode, the actuation unit flexed and extended the index finger at both low (1 Hz) and high (4 Hz) frequencies. Flexion and extension were smooth and highly sensitive to small motor displacements. *Figure 2.5* shows the rapid cycling of flexion and extension at a rate of 415 mm/min. The motion is shown in 0.25 sec increments. The wire rope “walked” smoothly along the pinion of the first stage of the pulley reduction without interfering with itself (see white arrows in *Figure 2.5*).

III. Discussion

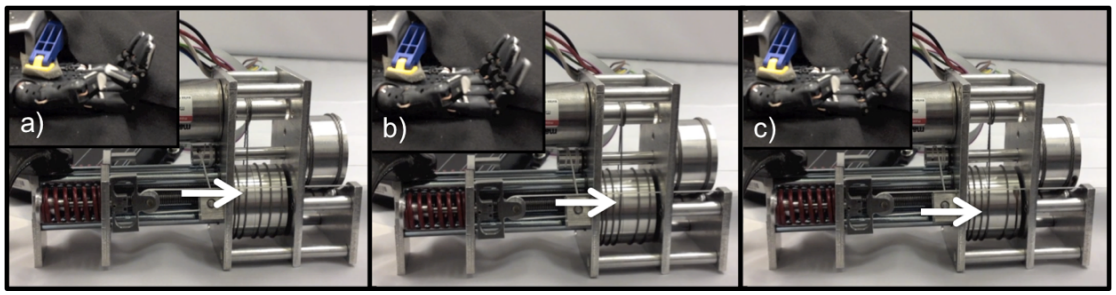

Figure 2.5 Direct actuation of a single joint of the index finger of the Shadow Dexterous Robot Hand (inset). The a, b, and c snapshots show increments of 0.25 seconds (tendon rate of 415mm/min).

A. Advantages of Design

The main advantages of the actuation system are its modularity, ease of setup, and ability to control different tendon-driven mechanisms. A single actuation unit is capable of controlling both flexion and extension while maintaining a tendon preload and stiffness for unanticipated loading of a fingertip, for example. The compression springs, in particular, afford deformation in the overall system that provides safety from unexpected loading of the finger. Various springs with different spring constants can be installed on a joint-specific basis depending on desired behaviors. In the two tendon back-driveable configuration (N-type) each tendon can have user-defined tension preloads due to the modular compression spring while the motor output shaft is traveling to or maintaining a desired joint angle.

Attachment between a remote bank of actuation units and the robot hand will be made via a bundle of low friction tendon sheaths. This will allow for a flexible, kink-resistant, and low-friction attachment of tendons from the actuation system to the robot hand as the hand is moved in the workspace.

Actuation units can be mounted such that each additional unit's controller can be linked to the existing EPOS CAN network and assigned a node value directly on the controller. The EPOS2 24/5 CAN network has a maximum of 128 separate nodes which is more than sufficient for an anthropomorphic robot hand.

The Shadow Dexterous Robot Hand requires 18 modules for complete double-acting, N-type actuation. Two actuation units can be operated with only a single output tendon to accomplish co-contraction in a 2N-type actuation scheme. Individual joints on a robot hand can be manually changed between single-acting actuation and double-acting actuation depending on the control architecture. To reduce control complexity, multiple tendons can be actuated simultaneously in order to produce synergistic joint movements or control underactuated manipulators. Preliminary tests suggest that the actuation system will enable high frequency responses that can match or exceed those of the human hand. As such, the actuation system presented here is appropriate for a robotic hand testbed that is capable of producing fast, reflex-like grip responses.

B. Disadvantages of Design

Due to size and weight issues, the bank of actuation units cannot be located on the robot arm itself without severely limiting the payload of the robot arm. The motor bank will be placed alongside the base of the robotic arm and will require lengths of tendon sheathing to route tendons flexibly up to the robot hand. However, this is acceptable for current purposes of creating a tendon-driven robot hand testbed as opposed to a body-worn neuroprosthesis, for example. Efforts will be made to reduce the length of the tendon sheaths as much as possible in order to minimize frictional losses between the remote actuation system and the end effector. Weight and power consumption are both issues that require further attention in the development of a compact and efficient motor control bank, although they are not immediate concerns for the present outlet-powered, workbench-mounted testbed design. Since the design of the robot hand testbed is intended for use as a research tool to advance the functional capabilities of highly dexterous robot hands, the power consumption of the motor units and supporting systems has not been considered as a significant design criterion.

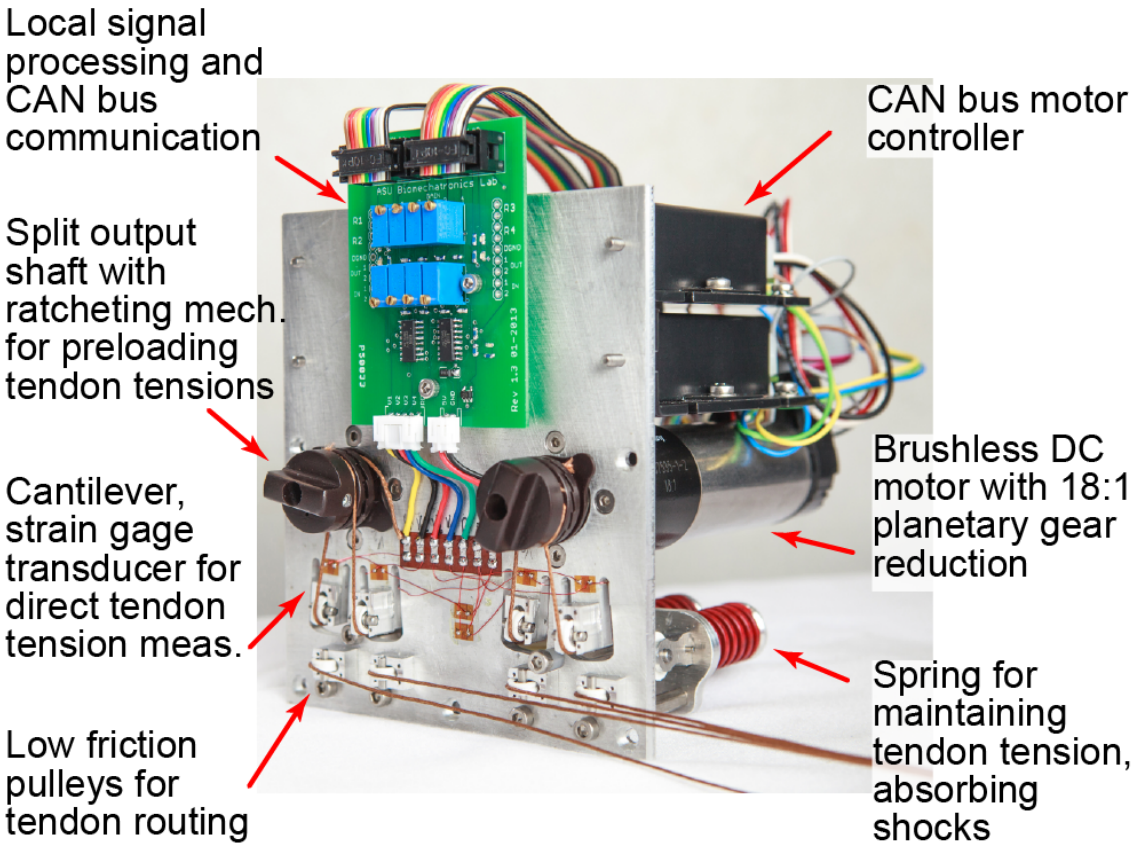

Figure 2.6 The modular remote actuation system is shown for two degrees of freedom in an "N-type" configuration: one motor each for PIP/DIP and MCP flexion/extension. A customizable circuit board locally amplifies and samples tendon tensions for transmission on the CAN bus.

A modular motor bank was used to apply tendon tensions and/or excursions (*Figure 2.6*). The design of the motor bank allows for either a "2N-type" or "N-type" set-up. With a 2N-type arrangement, there are two motors per joint allowing for independent control of a flexor and extensor tendon and enabling co-contraction and joint stiffness

control. With an N-type arrangement, a single motor is used at each joint in a "push-pull" fashion such that rotation of the motor shaft in one direction flexes the joint and rotation in the other direction extends the joint (Jacobsen et al., 1989). It was desired that the motor bank allow for the actuation of any tendon-driven mechanism. Thus, each motor has a split output shaft with a spring-loaded ratcheting mechanism to allow for quick setup and adjustment of tendon preloads (*Figure 2.6*).

Each motor (EC-max 30, 60 watt, Maxon Precision Motors, Inc.) is controlled by an EPOS 24/5 controller, which is connected to a CAN bus. Built-in microcontrollers allow for the offloading of low-level processes during position or current control, which aids in reducing the bandwidth of the bus and enables fast communication rates. Since the BairClaw is intended as a testbed, features such as weight, size, and power consumption of the actuation system were not optimized.

Tendon Tension Calculations

Each tendon was routed over multiple low friction pulleys (*Figure 2.6*). A spring-loaded pulley was used to maintain tendon tension and to provide passive compliance for unexpected loads or impacts during operation. For each tendon, another pulley was placed on the end of a cantilever beam cut into the motor plate. In order to measure tendon tensions, the base of each cantilever beam was instrumented with strain gages. A half Wheatstone bridge configuration was used for temperature compensation with reference gages located centrally on the plate. All gage measurements were amplified and sampled locally using a custom circuit board. Trimpots on the board allow for customization of baseline values, amplification, and resolution according to each tendon's range of operation (*Table 2.2*). It was desired that tendon tensions be measured with a resolution of ≤ 1 N (0.22 lbf).

Tendon tensions are sampled via the EPOS 24/5 motor controllers and transmitted over the central CAN bus. By using the motor controllers to sample tendon tensions, we can scale the entire testbed by simply daisy chaining more controllers onto the CAN bus with little to no modification of the low-level communication scheme.

Table 2.2 Specifications for the artificial proprioception, tactile sensor, and remote actuation subsystems

	Sampling rates (Hz)	Design range [min, max]	Design resolution
BairClaw index finger			
Joint angle sensors			
DIP flex/ext	100 – 1000	[-30°, 90°]	0.12°
PIP flex/ext	100 – 1000	[-10°, 90°]	0.10°
MCP flex/ext	100 – 1000	[-30°, 90°]	0.12°
MCP add/abd	100 – 1000	[-15°, 15°]	0.03°
Multimodal tactile sensor (BioTac, SynTouch LLC)			
Electrode impedance (19 electrodes total)	100 – 200 (per elec.)	[0, 3.3 V]	3.2 mV
Internal fluid pressure	100 – 200	[0, 100 kPa]	36.5 Pa
Vibration	2200 – 4400	± 0.76 kPa	0.37 Pa
Temperature	100 – 200	[0, 75 °C]	0.1 °C
Thermal flux	100 – 200	[0, 1 °C/s]	0.001 °C/s
Remote actuation system			
Tendon excursion	200	---	0.9 μm
Tendon tension	200	[0, 111 N (25lbf)]	0.11 N (0.025lbf)

The tendon tension measurement system was calibrated using a multi-step process that accounted for interactions between neighboring strain gages on the motor plate. In general, the change in resistance of a strain gage is linearly dependent on the internal strain and stress at the location of the sensor. Due to highly sensitive gages and the close proximity of pulleys on the motor plate, each half Wheatstone bridge sensed strain caused by tendons routed over nearby pulleys.

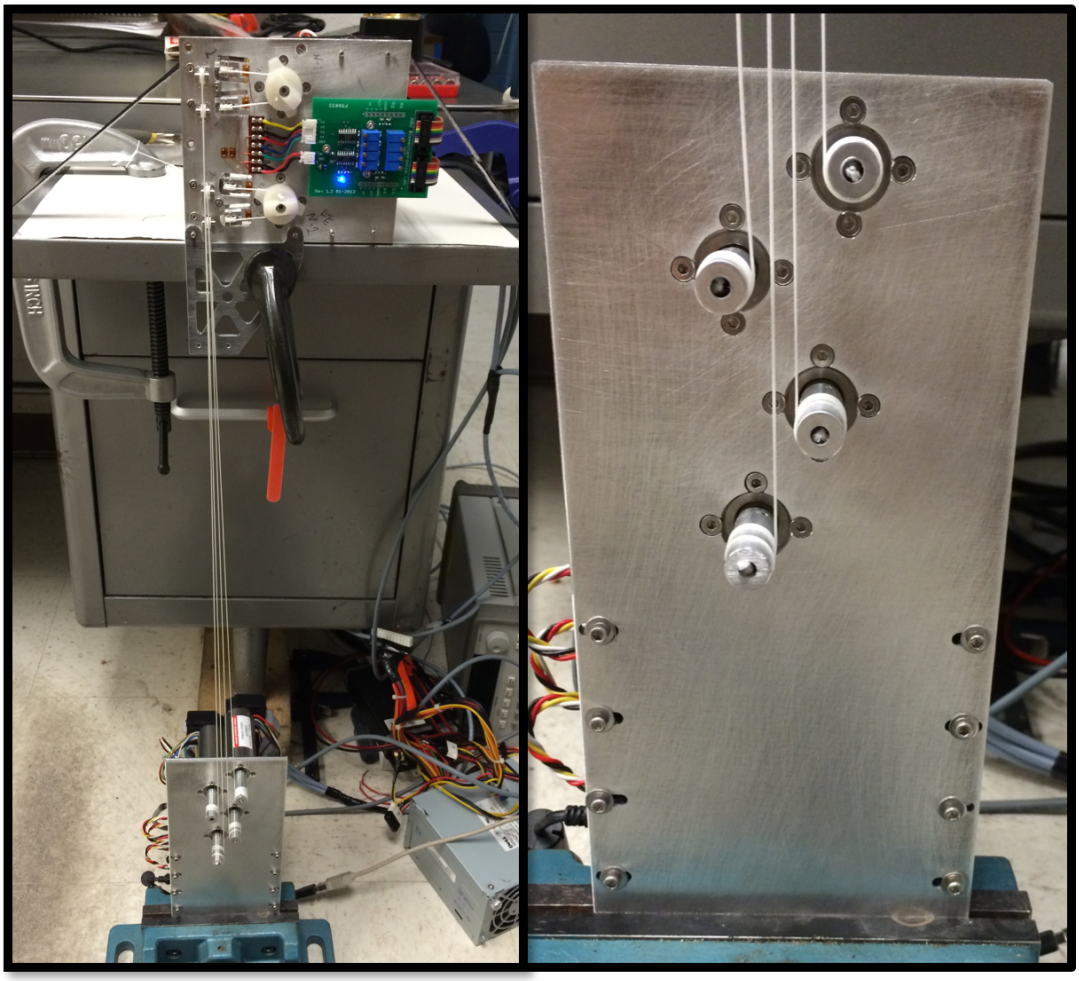

Figure 2.7 Strain gage calibration with custom calibration rig. The calibration rig was used to apply random tensions ranging from 0-50 N on each half bridge. Recorded strain measurement was then used to train a multivariate tendon tension model.

A custom calibration rig (*Figure 2.7*) was built that randomly applied a known force to all four tendons simultaneously. The Maxon EC-max 30 motors were initially calibrated using a single axis force transducer. The motors were fit to static motor current torque curves to determine the force applied on each of the tendons. One thousand trials of randomly selected tendon tension combinations were applied, with each individual

tendon tension ranging from 0-50 N. Each tendon tension sensor was fit to a multivariate linear model comprised of a sum of scaled tensions of tendons mounted nearby on the motor plate (*Figure 2.8*). The calibration models were cross-validated using a Lasso method to minimize mean squared error. The Lasso method returned comparable models and confirmed the statistical significance of the additive terms associated with nearby tendons (Tibshirani, 1996).

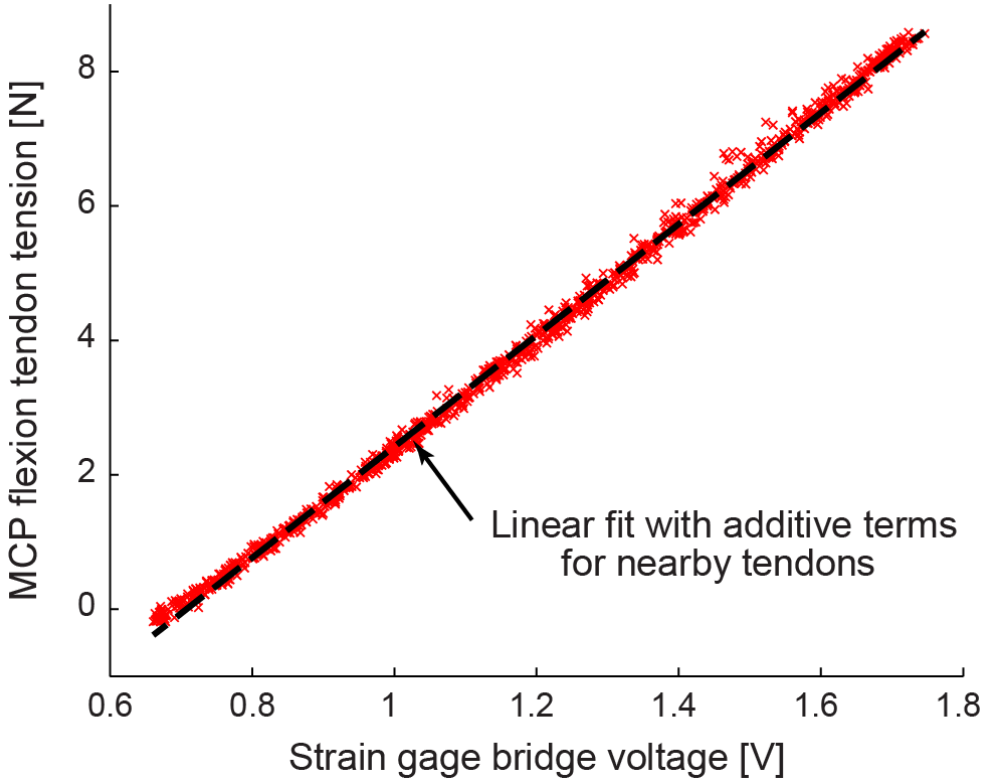

Figure 2.8 A representative tension calibration is shown for the MCP flexion tendon. Using 1000 trials of randomly selected tendon tension combinations, each tendon tension sensor was fit to a linear model comprised of a sum of scaled tensions of tendons mounted nearby on the motor plate.

Highly Dexterous Robot Finger with Rich Tactile Sensing: The "BairClaw"

In this section, we present the "BairClaw," a highly sensorized, multi-articulating, anthropomorphic robot hand testbed with rich proprioceptive and tactile sensing capabilities (*Figure 2.9*). The BairClaw was originally conceived for the advancement of artificial grasp, manipulation, and haptic exploration. We posit that the system could also be used for the neurorehabilitation of somatosensory disorders due to upper limb impairment or loss.

Thus far, efforts to enhance an amputee's sense of embodiment with a prosthesis have focused on cosmetic appearance via the development of five-digit, multi-articulating prosthetic hands and attempts to design realistic, skin-like cosmeses (Marasco, Kim, Colgate, Peshkin, & Kuiken, 2011). We believe that embodiment additionally requires the development of consistent action-perception relationships and their encoding in the nervous system. The BairClaw testbed was designed to enable the development of consistent action-perception relationships that enhance one's sense of embodiment for robotic or human-in-the-loop use. Mirror visual feedback techniques that facilitate neural plasticity can be further enhanced through the provision of rich proprioceptive and tactile feedback in synchrony with action. By providing amputees with the ability to control, visualize, and feel physical finger-object interactions in a controlled clinical setting, it may be possible to extend current therapies that focus on visualizations of posture alone.

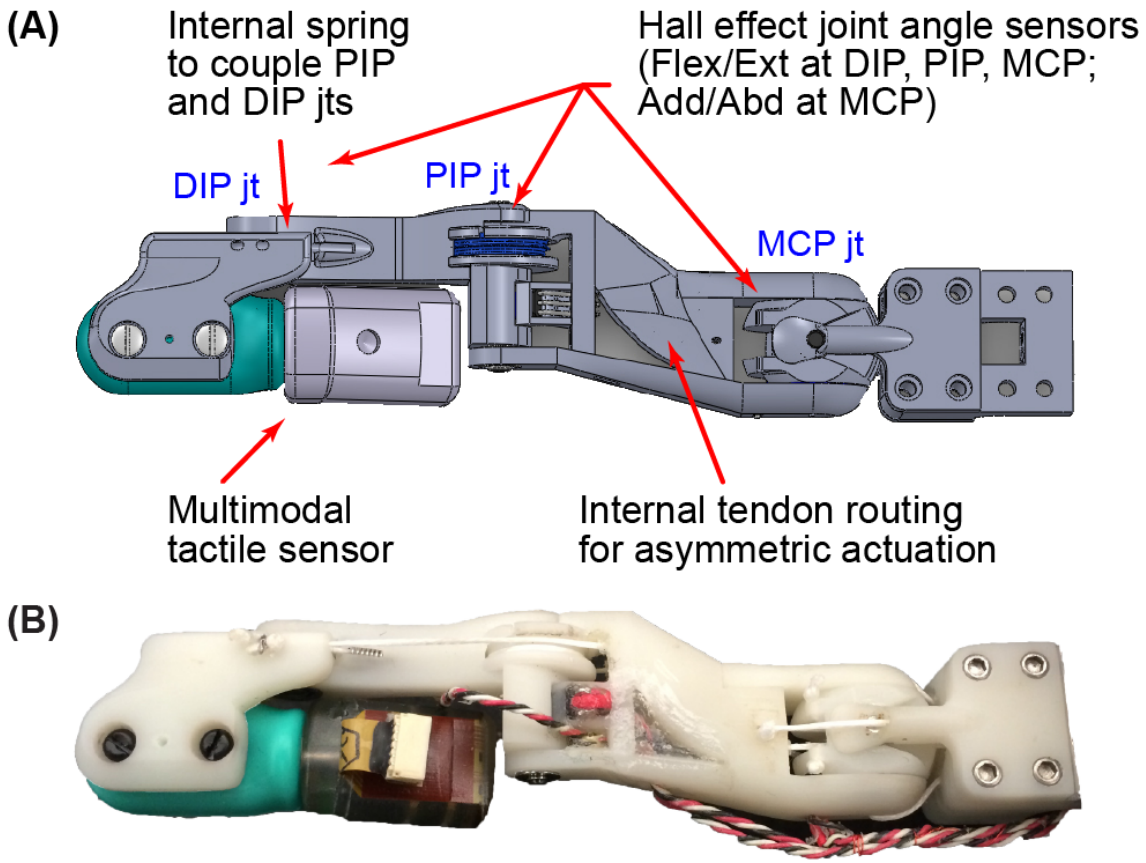

Figure 2.9 The BairClaw index finger has four degrees of freedom: DIP, PIP, and MCP flexion/extension and MCP adduction/abduction. The DIP and PIP joints are passively coupled by a spring. Joint angles are measured by Hall effect sensors while temperature, vibration, and skin deformation are provided by a multimodal tactile sensor. Dorsal views of the (A) design schematic and (B) prototype are shown.

Asymmetric Finger Design

The BairClaw was designed to accommodate a multimodal tactile sensor (BioTac, SynTouch LLC) that has an immobile distal interphalangeal joint ("SynTouch BioTac biomimetic tactile sensor," 2007). In order to perform complex, human-inspired motions

such as stroking and rolling of objects between the fingertips, the distal interphalangeal joint was restored through the use of an asymmetric finger design. The four degree-of-freedom index finger features flexion/extension and adduction/abduction at the metacarpophalangeal joint (MCP) and coupled flexion/extension of the proximal and distal interphalangeal joints (PIP and DIP, respectively). The proximal end of the BioTac can rotate toward and slightly through the dorsum of the hand, thereby allowing the distal joint to function normally (*Figure 2.9B*).

An embedded spring in the middle phalanx controls the flexure of the DIP joint during PIP and DIP joint flexion. The spring slightly increases the torque required to flex the DIP joint, which causes the PIP joint to flex first. Motion at the DIP joint begins when the PIP joint has reached its full range of motion or if an object impedes PIP joint motion. The spring and an internal PIP joint pulley allow for a low friction, passive compliance of the finger. To minimize friction, all joints were supported by ball bearings, and a PTFE-lined internal channel within the proximal phalanx was used to route tendons.

Eventually, an elastomeric cosmesis could be used to hide the slight protrusion of the BioTac through the dorsal surface of the hand, so as not to break the anthropomorphic illusion for embodiment purposes. However, such illusions are fragile and strengthening of the body schema may be better served by accurate motions and the provision of sensory feedback consistent with actions as opposed to an anthropomorphic appearance only.

Tendon Routing and Sheath Design

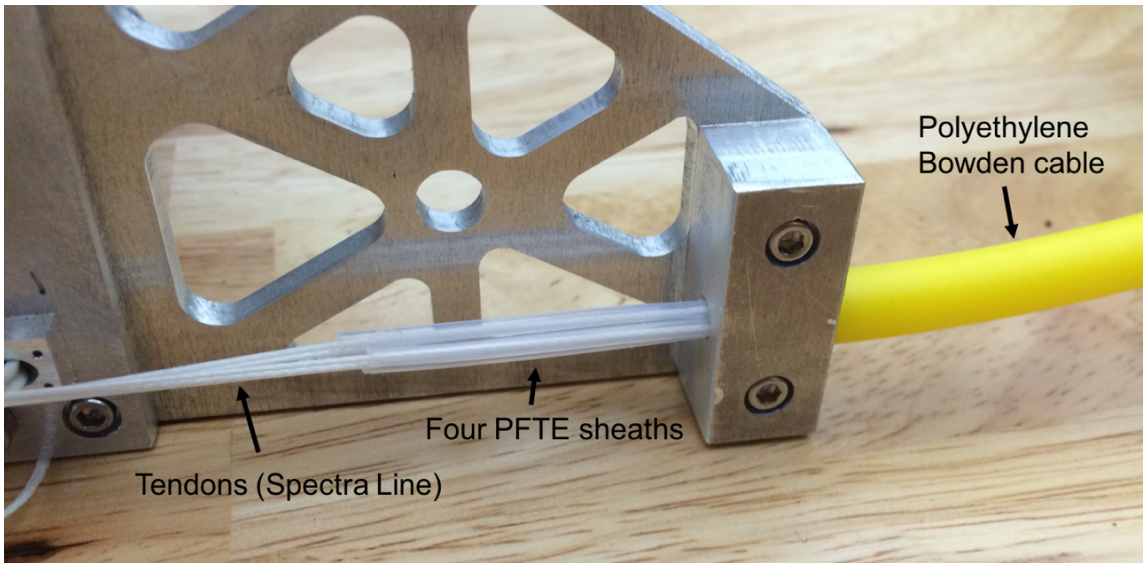

Figure 2.10 Tendon routing from 2nd actuation system (*Figure 2.6*) to the BairClaw. Four PTFE sheaths were routed through a yellow vacuum line that acts as a Bowden cable.

Under load, friction in the tendon routing system can significantly influence the dynamics of the system and cannot be overlooked (Nahvi et al., 1994). Thus, each tendon (200 lbf Spectra line, Power Pro) was routed through a low friction polytetrafluoroethylene (PTFE) sheath, four of which were additionally bundled within a polyethylene vacuum line that serves as the supporting structure of the Bowden cable design (*Figure 2.10*). Bowden cables consist of an outer sheath that is constrained at both ends while the internal cable transmits a pulling force. The outer cable is flexible and constant in length allowing for force transmission. Additionally, tendon paths were rerouted using small, ball bearing-mounted pulleys.

Joint Angle Sensing

In the index finger prototype, Hall effect sensors were used to measure four sets of joint angles: flexion/extension and adduction/abduction at the MCP joint, flexion/extension at the PIP joint, and flexion/extension at the DIP joint. Each joint angle sensor was comprised of a Hall effect sensor that measured the change in magnetic field induced by the rotation of a diametrically magnetized ring magnet. Various Hall effect sensor and ring magnet combinations were used to optimize the resolution of each sensor over the full, joint-specific range of motion (*Table 2.2*), and were designed to measure joint angles with a resolution of ≤ 1°. It was desired that all proprioceptive sensors be sampled at rates of at least 100 Hz.

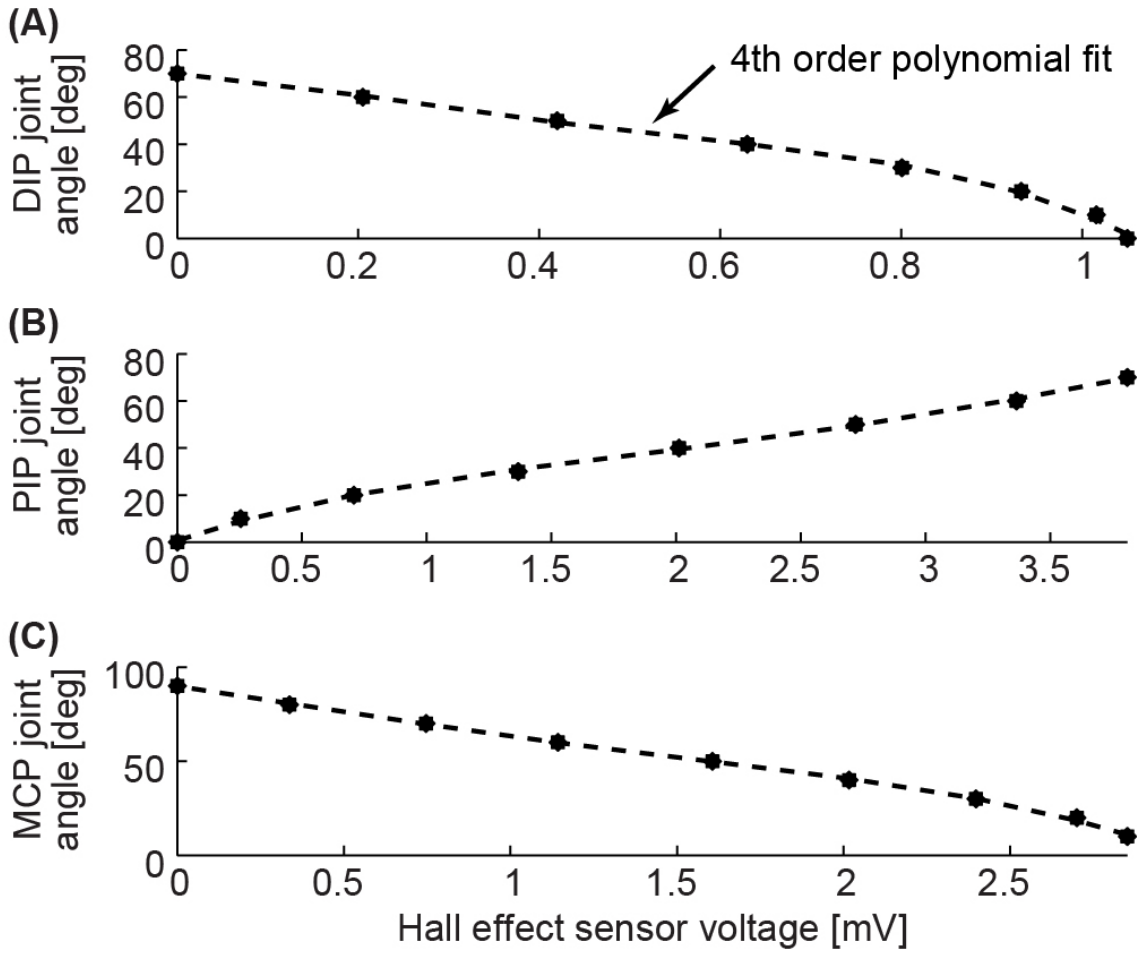

Figure 2.11 Joint angle measurement calibrations are shown for the (A) DIP, (B) PIP, and (C) MCP flexion/extension degrees of freedom. Fourth order polynomial fits performed on calibration data collected in 10° increments resulted in $R^2 > 0.99$ for each joint. Positive angles indicate joint flexion from a neutral position at 0°.

The joint angle measurement system was calibrated for each joint's range of motion in 10° increments using a goniometer. The Hall effect sensors were designed to respond linearly to changes in the magnetic field. Due to small variations of the magnetic field near the ends of the range of motion, it was necessary to fit a 4th order polynomial

model to the sensor response (*Figure 2.11*). Each model had a coefficient of determination R^2 value greater than 0.99. The slope of the calibration curves depended on space requirements within the limited volume of the BairClaw finger and the Hall effect sensor configuration at each joint.

Communication and Scalability

The BairClaw testbed is controlled by a central, Linux computer running Ubuntu 12.04 that has been modified with a Xenomai kernel patch for hard real-time operation. All communication is performed on a CAN bus, a standard in industrial automation and motor vehicles that ensures real-time communication with simple message packets and a node-based communication structure. Data transmitted via CAN and USB (for the multimodal tactile sensor) are recorded and logged in real-time. Since CAN uses a simple two-line bus, the entire system can easily be scaled by daisy chaining additional digits and motor controllers onto the original bus.

Controllers

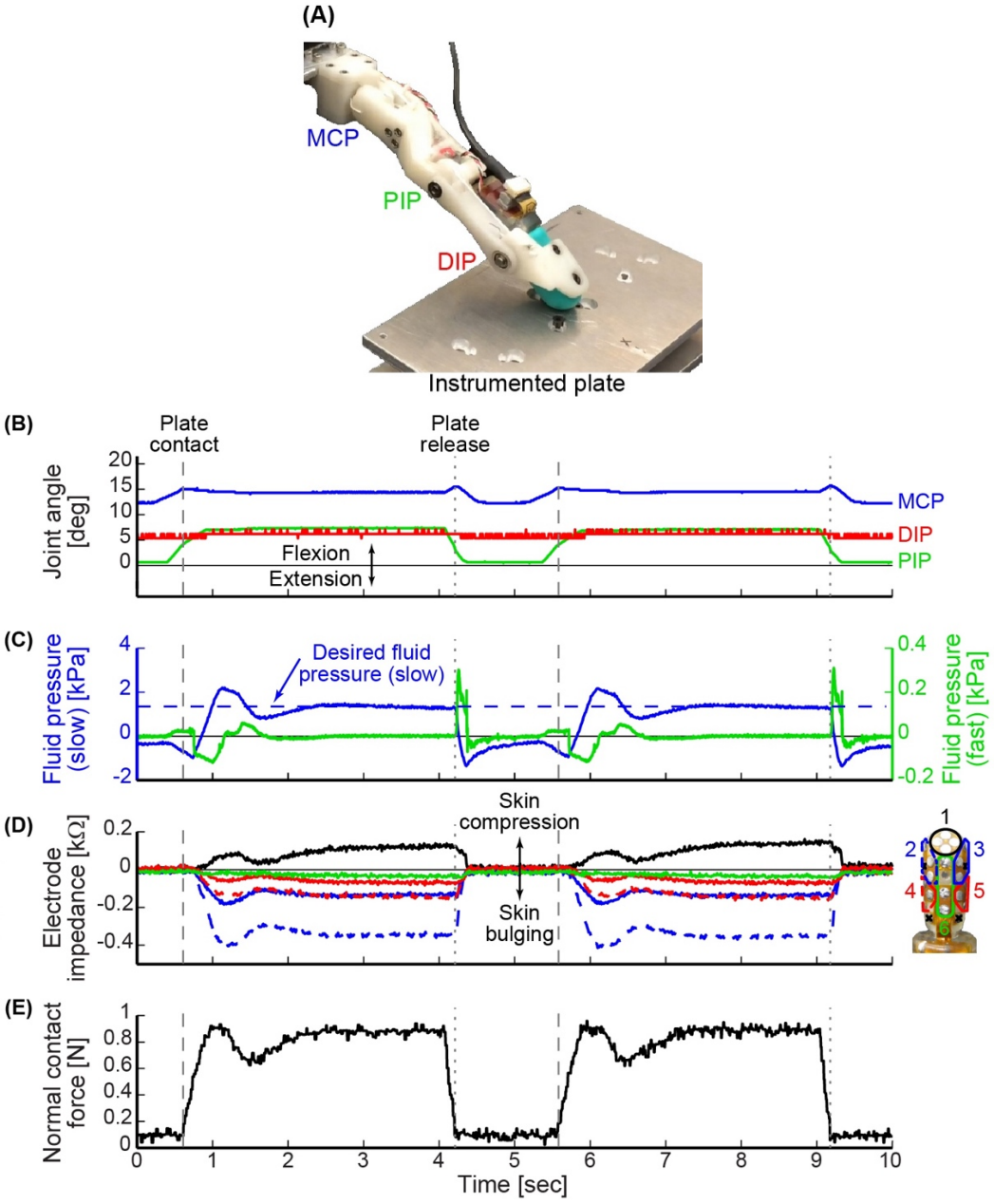

Figure 2.12 (A) The BairClaw was used to perform a tap-and-hold experiment against an instrumented plate. (B) Joint angles, (C) tactile sensor internal fluid pressure (left y-axis) and microvibration (right y-axis), (D) tactile sensor skin deformation, and (E) normal contact force data are shown for two cycles of motion and force production. As the finger

flexed and the tactile sensor's internal fluid pressure exceeded a threshold value, the position controller switched to a force controller designed to achieve and maintain a desired fluid pressure value (horizontal dashed line in (C)). Fingertip contact with and release from the plate are indicated by dashed and dotted vertical lines, respectively.

A variety of control schemes could be devised for the testbed. A position controller could use proprioceptive sensor data from joint angle sensors and motor encoders as control signals. A force controller could use proprioceptive sensor data, such as tendon tensions and motor current, or tactile sensor data as control signals. Tendon tensions and moment arms, known from design schematics, can be used to calculate joint torques created by the multi-articular tendons. Standard robotics equations can be applied to relate joint motion to fingertip motion, or joint torque to fingertip forces and torques in three dimensions (Murray, Li, Sastry, & Sastry, 1994).

For demonstration purposes, we illustrate the use of a hybrid position and force feedback controller for a cyclic tap-and-hold task (*Figure 2.12B*). The controller was designed to function as a state machine that initially operates in position control and moves at a set rate to achieve a commanded posture unless the finger pad comes into contact with an object. In this example, once the tactile sensor's internal fluid pressure exceeded a threshold, the position controller switched to a force feedback controller designed to achieve and maintain a desired fluid pressure value (as a proxy for fingertip contact force). Specifically, the fluid pressure signal was used in a proportional-integral-derivative feedback controller to quickly achieve and maintain the desired reference value with zero steady-state error.

Hybrid Position and Force Feedback Controller

The tap-and-hold example demonstrates the speed with which the BairClaw testbed can switch control modes, the versatility of control using a variety of feedback control signals, and the stability of the overall mechatronic system. *Figure 2.12B-E* shows two cycles of a tap-and-hold experiment in which joint angles and angular velocities were tracked as the BairClaw flexed to a pre-specified posture. Before the final posture could be achieved, the BairClaw fingertip contacted a plate instrumented with a six degree-of-freedom load cell (Nano-17, ATI Industrial Automation). Once the internal fluid pressure of the tactile sensor reached a threshold, the position controller switched to a force controller to achieve and maintain a desired reference fluid pressure value. The pre-specified posture and fluid pressure threshold and reference values were selected arbitrarily for demonstration purposes, but could be set according to the context of the experimental task.

Figure 2.12B-D show the joint angles, and tactile sensor internal fluid pressure, microvibration, and skin deformation data. As with any higher order, underdamped system, a slight overshoot occurred in the internal fluid pressure control signal, but was quickly corrected (*Figure 2.12C*). Trends in the normal contact force measured by the instrumented plate aligned with the internal fluid pressure signal used for force feedback control (*Figure 2.12E*).

The tactile sensor's electrode impedance values provide information on skin deformation caused by the BairClaw's forceful interaction with the plate. Individual electrode impedance values were grouped into anatomically meaningful clusters for visualization purposes. Increases in impedance indicate that the skin is being compressed

toward the sensor's rigid core while decreases in impedance indicate bulging away from the core. As expected, the electrode impedance data indicate compression of the skin at the fingertip and bulging of the skin on other regions of the finger during the tap-and-hold phase of the trial.

In order to relate one's voluntary actions to resulting stimuli (visual or otherwise), there must be minimal delay between the action and the perceived stimuli. Previous work has shown that delays for myoelectric prosthetics should be kept below 125 ms (Farrell & Weir, 2007). The majority of the delay found in myoelectric controllers is due to processing of the myoelectric signals. The BairClaw is capable of processing and reacting to various inputs within a single sampling period (10 ms). Mechanical and computational delays, estimated from the delay between the switching of the controller and a measurable change in system response, was approximately 65 ms in the tap-and-hold example. Any additional delays for a human-in-the-loop configuration would be specific to the human-machine interface and signal processing, such as pattern recognition that may be performed on the human command signals.

Discussion

Contrasting First and Second Generation Actuation Systems

Each actuation system has similar functionality although accomplished through different designs. The main differences between the two actuation systems are in their reduction drive, tensioning mechanism, and force transducers.

The first design presented has a 12:1 compound pulley reduction that has zero backlash (*Figure 2.1*). Zero backlash is a convenient feature that will improve high frequency tracking response of the actuation system and also contribute to an aesthetic feel. However, the additional cost and assembly time of the zero backlash actuation system did not justify its added benefit. In order to reduce both cost and time of production, we explored alternate reduction drives. Maxon offers a low backlash 18:1 planetary gearhead designed to fit the EC-max 30 motors used on both actuation systems. Use of the planetary gearhead introduced small amounts of backlash into the system, but the majority of the backlash from the planetary gear head is damped by the compression springs and any tendon creep. The second actuation system (*Figure 2.6*) presented centers around a main mounting plate in which the planetary gearhead is directly mounted. Mounting the planetary gearhead directly to a central fixture plate reduces cost of both assembly and machining.

The functionality of the tensioning system on the first actuation system (*Figure 2.1*) is similar to that of the split output shaft with ratcheting mechanism on the second actuation system (*Figure 2.6*). The split output shaft has additional benefits. For instance, slots are cut into each pulley allowing a stop knot to be easily slid into the pulley without any cutting of tendons or retying of knots. The split output shaft is not limited in the

allowed adjustments while the lead screw tensioning mechanism (*Figure 2.3*) is limited to 14 cm of overall adjustment. If the split output shaft fills will excess tendon you can simply unspool and trim the tendon and then slide the stop knot back on the pulley. Releasing tension on the split output shaft is easily done by compressing the spring between the motor output and the back pulley (*Figure 2.6*). Once the ratchet is disengaged, the two pulleys on the split output shaft will counter-rotate, thereby loosening the tendons for swapping of end effectors. Reapplying the tension to the tendons is quickly done by impeding rotation of the tab on the backside pulley and rotating the thumb tab on the front pulley. While the output shaft rotates the ratchet will click incrementally to increase the tension in the tendons. To accomplish similar tension adjustments with the first actuation system a screwdriver is required to adjust a centrally mounted lead screw. The use of the split output shaft provides multiple benefits over the lead screw of the first design by reducing overall weight, parts and machining and assembly time.

The force transducer for the second actuation system is designed directly onto the central mounting plate shown in *Figure 2.6*. By having all the major features on the central mounting plate we are able to reduce the CNC machining to a single face of the plate. The benefit of machining only one of the six faces of the plate is that there is no need adjust or index the CNC mill between tool changes or the production of multiple plates. For instance, with the first transducer design, the strain gages required additional support to correctly epoxy them on the tight inner contours of the inline force transducer (*Figure 2.6*) while the second design simplified the installation by utilizing a simple cantilever beam to measure strain. Furthermore, by mounting the gages to the center plate

we can remove any possible failure modes introduced by additional wiring and connections necessary for communicating with the moving gages of the first actuation system. Given the simplified installation of the strain gages there is always a tradeoff. The strain gages of the second actuation systems are now in close proximity to one another. As a result, all gages are now sensitive to loads on any specific cantilever beam. The resulting cross-talk between the gages was managed in a straightforward way; a multivariate linear model was trained to accurately determine each individual gage response.

Overall, the design changes in the second actuation system reduced cost and assembly time while improving performance and ease of use compared to the first actuation system.

Applications of the Robot Testbed

Upper limb impairment or loss can be caused by a multitude of factors including disease, trauma, surgery, and brain infarction (Dickstein & Deutsch, 2007; Harris, 1999). There is a 50-80% chance that when one loses a limb an incessant pain called “phantom limb pain” will remain after amputation (Nikolajsen & Jensen, 2001). The pain can occur immediately after trauma or may take months to years to develop. The root cause of phantom limb pain is not well understood and may be due to an irritation of nerve endings, a central remapping of sensations that results in misinterpreted activations of pain neurons, or the mismatch of motor commands and visual feedback that are then interpreted as pain (Ramachandran & Hirstein, 1998). Even in the absence of severe pain, amputees often refer to their missing limbs as feeling paralyzed in an uncomfortable or cramped position. Patients often experience depression due to the pain and discomfort that is degrading their quality of life. In the remainder of this chapter we will discuss potential clinical benefits to upper limb amputees that could arise from the confluence of known concepts such as mirror visual feedback and the "rubber hand" illusion, and new technologies such as neural interfaces for artificial sensory feedback and highly sensorized robot hand testbeds, such as the “BairClaw”.

Non-Invasive, Vision-Based Therapies for Pain Disorders and Paralysis

Mirror visual feedback (MVF) was introduced in 1992 as a non-invasive technique to treat phantom limb pain due to amputation and paralysis due to stroke (Ramachandran & Altschuler, 2009). A mirror or virtual environment is used to provide a visualization of one's missing or hidden impaired limb by reflecting the movement of the contralateral unimpaired limb. Despite inaction by the impaired limb, this technique results in activation of regions of the brain corresponding to the lost or impaired limb. When MVF was first examined over 20 years ago, pain disorders and paralysis were believed to be untreatable. Since then, MVF has been used to treat complex regional pain syndrome and peripheral nerve damage. Even though MVF is not a panacea, it has been shown to be an effective form of therapy for phantom limb pain (Darnall, 2009; Stevens & Stoykov, 2003). Any positive treatment can have a large impact considering the high occurrence of phantom limb pain in amputees, and the fact that strokes are the leading cause of long-term disability (Go et al., 2014).

Graded Motor Imagery (GMI) is a variation of MVF that has had success in reducing pain and discomfort associated with pain and movement problems. GMI consists of three steps: left/right discrimination, motor imagery exercises, and mirror therapy (Johnson et al., 2012; Moseley, 2006). When first starting GMI treatment, left/right discrimination is the primary focus because it has been shown that individuals with chronic pain are less accurate and/or slower in determining whether an image is of a left or right limb compared to healthy individuals (Schwoebel, Friedman, Duda, & Coslett, 2001). Difficulty in determining laterality reflects the lack of a strong body schema. Motor imagery exercises such as imagining hand movements aid in increasing

activity of motor cortical neurons that are involved with observed, imagined, or executed movements (Rizzolatti & Craighero, 2004). Strengthening of the body schema through left/right discrimination and explicit motor imagery creates a foundation upon which subsequent mirror therapy can be most effective (Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Priganc & Stralka, 2011; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). Through the use of GMI and sensory feedback to the phantom limb, it should be possible for a neuroprosthetic or robotic system to be incorporated into one's body schema, which could aid in the treatment of phantom limb pain and improve functional performance with a prosthesis.

The "Rubber Hand" Illusion

Studied often, the rubber hand illusion phenomenon illustrates the interactions between vision, touch, and proprioception as they relate to the body's self-identification (Botvinick & Cohen, 1998). The illusion is created by hiding the subject's hand out of view and then placing a rubber hand in its place. Both the subject's hand and the rubber hand are brushed simultaneously. After some time, the subject can develop a sense of ownership with the rubber hand and disassociate from his/her native hand, reporting the feeling of brush strokes when only the rubber hand is brushed (Botvinick & Cohen, 1998). Subjects also experience "proprioceptive drift," which describes the phenomenon in which subjects report the location of his/her native hand as being closer to the rubber hand than the native hand's actual location. That is, the proprioceptive percept of the subject's native hand has "drifted" towards the rubber hand.

Neuroplasticity and the ability to incorporate an artificial limb into one's body schema date back to studies from 1937. Early work on the Aristotle illusion examined localization errors in perceived tactile stimuli when an object was touched simultaneously by the outer regions of two crossed fingers (Tastevin, 1937). For example, simultaneous contact of one's nose with the radial aspect of the index finger and ulnar aspect of the middle finger can result in the perception that one has two noses. Recent studies have further demonstrated a link between one's body schema and the physiological self, and how expressions of this link manifest themselves in measurable physiological changes. It has been hypothesized that increased ownership of an artificial limb disrupts regulation of certain aspects of the native limb. Interestingly, as an artificial upper limb becomes accepted into one's body schema, the temperature of the native limb decreases (Moseley et al., 2008). Other experiments have shown that, through the rubber hand illusion, subjects' immunological responses can be altered. The immune system's primary goal is to discriminate self from non-self in order to protect the body from foreign organisms. In one such experiment, the response to a topically applied histamine was altered during the rubber hand experiment; welt size was larger for the hidden native limb when the illusion was in effect (Barnsley et al., 2011). The ability to manipulate the physiological response of the body through a visual illusion leads one to believe that the addition of congruent proprioceptive and touch feedback could accelerate the incorporation of a neuroprostheses into one's body schema.

In a 2012 rubber hand illusion experiment, biological fingerpads were subjected to vibrotactile stimulation while a rubber hand was stroked or tapped (D'Alonzo & Cipriani, 2012). When vibrotactile stimuli were synchronized with the visual feedback, subjects developed a sense of ownership of the fake hand, despite the sensory substitution, modality mismatched nature of the sensory feedback to the biological hand. Recently, a single digit version of the rubber hand illusion was conducted with a tactile sensor (Hartmann, Dosen, & Farina, 2014). The fingertip sensor was stroked and pressed in different regions in each subject's view while the subject's native hidden forearm was electrocutaneously stimulated according to changes in the tactile sensor data. Preliminary findings showed that subjects' skin temperature decreased slightly in the native limb and proprioceptive drift resulted, as would be expected when the illusion is successful. Interestingly, even though some subjects indicated a lack of embodiment of the green-colored fingertip sensor, and sensory substitution methods were employed, researchers still observed physiological signs of a subconscious incorporation of the artificial finger into the body schema.

Proprioceptive and Exteroceptive Feedback for Amputees

Prior work suggests that tactile and proprioceptive inputs are encoded simultaneously in unimpaired individuals (Rincon-Gonzalez, Naufel, Santos, & Helms Tillery, 2012). For instance, non-weight bearing contact of the fingertip against a surface can improve the accuracy of perceived posture. In turn, limb posture can significantly change the cortical response to identical tactile stimuli (Rincon-Gonzalez, Warren, Meller, & Helms Tillery, 2011). Sensory feedback mappings are clearly a function of both proprioceptive and exteroceptive information. Ongoing efforts to artificially produce conscious perceptions of phantom limb posture, motion, and contact with objects could be accelerated if proprioceptive and exteroceptive information could be provided to an amputee simultaneously and in an intuitive manner.

Natural Sensory Feedback from the Residual Limb

The body-powered, cable-driven prosthesis is still a popular choice for many amputees. While rejection rates remain somewhat high for powered, myoelectric prostheses (35% and 23% for children and adults, respectively) (E. A. Biddiss & Chau, 2007), many amputees prefer the speed of control and immediate natural sensory feedback obtained via extended physiological proprioception. During operation, cable excursion and stiffness can be sensed through the prosthesis socket as well as through the body harness (e.g. standard figure eight harness or cutaneous anchor adhered directly to the skin) (Williams III, 2011). Although direct joint movement information is not available, body-powered prosthesis users are able to learn how to use this extended form

of proprioception for the grasp and manipulation of objects. It has been shown that functional performance with body-powered prostheses can be further improved when the prosthesis is designed with extended physiological proprioception in mind (Doubler & Childress, 1984).

Methods for the Provision of Artificial Sensory Feedback

Communicating proprioceptive and exteroceptive information to amputees remains a grand challenge. Non-invasive sensory substitution methods using vibrotactile or electrotactile stimuli can be used to provide feedback, but the feedback is typically non-intuitive or difficult to scale to a multitude of simultaneous signals (Kaczmarek, Webster, Bach-y-Rita, & Tompkins, 1991). For example, vibrations can be applied to a residual limb in relation to prosthesis grip force, but the amputee must learn this non-intuitive mapping. This may suffice for a single channel of information, but additional non-intuitive vibratory feedback that is simultaneously applied to other regions of the body will likely increase the cognitive burden on the user. In different studies, subjects often reported that feedback provided via sensory substitution methods was distracting (Jimenez & Fishel, 2014; Pylatiuk, Kargov, & Schulz, 2006; Uellendahl, Mandacina, & Ramdial, 2006).

Significant progress has been made toward the development of non-invasive and invasive peripheral and cortical neural interface technologies for providing multiple channels of sensory feedback in a more intuitive manner. Tactors have been used to non-invasively vibrate regions of skin covering tissue that has under gone targeted muscle reinnervation. Impressive subject-specific mappings have been published that show the

regions of the chest, for example, that can be stimulated to induce percepts on the phantom limb (Kuiken, Marasco, Lock, Harden, & Dewald, 2007). Subjective and objective outcome measures have shown that the use of tactors reinforces one's sense of embodiment of the artificial limb (Marasco et al., 2011).

Peripheral neural interfaces, such as nerve cuff electrodes, have been used to stimulate the median, ulnar, and radial nerves in the residual limb (Navarro et al., 2005). Such electrodes have recently been used to provide simultaneous proprioceptive and tactile feedback to different regions of the phantom limb (D. Tan et al., 2014; D. W. Tan et al., 2014, p. -). Interestingly, when two distinct channels on the electrode were stimulated simultaneously, subjects reported percepts in regions of the phantom limb that were not previously reported after stimulation by any individual channel. Although further subject-specific characterization of this phenomenon is necessary, it is clear that the provision of simultaneous tactile and proprioceptive feedback will be possible for many more regions of the phantom limb than there are physical neural interface channels. After stimulation sessions, amputees have reported changes in their previously paralyzed phantom limb postures and, importantly, a reduction in phantom limb pain. Some subjects even reported that they were practically pain free (D. W. Tan et al., 2014).

Intracortical microstimulation has been used in brain-machine interfaces for the provision of tactile and proprioceptive feedback (Rincon-Gonzalez et al., 2012; Velliste, Perel, Spalding, Whitford, & Schwartz, 2008). In non-human primate studies, electrical stimulation in somatosensory cortex has been used to convey limb movement, although the provision of absolute limb position remains a challenge, irrespective of the neural interface method being used (London, Jordan, Jackson, & Miller, 2008). While

intracortical microstimulation has been shown to be an adequate tool for influencing the perception of limb motion, stimulation in area 3a also elicits detectable changes in electromyograms in associated musculature (Witham & Baker, 2011). However, the sensations that are elicited by area 3a stimulation remain unknown.

Researchers have also vibrated tendons in the residual limb to provide proprioceptive feedback to amputees. It is hypothesized that the vibrations excite muscle spindles such that a muscle lengthening is perceived. For example, vibration of an extensor tendon can create the sense that the associated joint is being flexed. Amputees have been able to sense joint motion in the phantom limb, as when opening or closing the hand (Marasco, 2014b).

Availability of Sensor Technology in Artificial Hands

To date, the availability of proprioceptive sensors in commercially available prosthetic hands has been limited (Controzzi et al., 2014). Joint angle encoders are not available in commercially available myoelectric prostheses, rather motor encoders are used to estimate grip aperture. However, motor encoders cannot be used to estimate the posture of multi-articulating digits when underactuated finger designs are used. For instance, the tendon-driven, conformal grasp of prosthetic hands such as the Touch Bionics i-limb prosthesis ("Touch Bionics i-limb product range," 2014) reduces the degrees of freedom that an amputee must consider for control, but specific hand configurations cannot be measured or conveyed to an amputee in real-time. One recent study with the i-limb reports the use of motor current monitoring and timing of finger movements from the i-limb's neutral, fully opened position as a way to estimate joint angles (Kyranou, 2014). Future work is required to overcome limitations resulting from assumptions about finger velocity, battery power, and object rigidity. The VINCENT hand prosthesis ("Vincent Systems VINCENTevolution 2," 2014) was designed with less joints per digit than the biological hand and a spring is used to couple joints in each digit. In a research model of the VINCENT hand called the "Bionic Hand," flex/bend sensors were placed at the metacarpal joints only ("Vincent Systems Bionic Hand," 2014). The rigid link design of the RSL Steeper Bebionic hand prosthesis (Medynski & Rattray, 2011; "RSL Steeper Bebionic3 Hand," 2014) facilitates the use of motor encoders to track digit posture, but the system does not currently include tactile sensors.

Tactile sensor technology is also scarce in commercially available prosthetic hands, and remains unimodal in nature (Controzzi et al., 2014). The one degree of

freedom Otto Bock SensorHand Speed hand prosthesis (Puchhammer, 2000) uses fingertip sensors to detect the slip of a grasped object. Some commercially available myoelectric hands have been modified with multimodal tactile sensors as research tools (Jimenez & Fishel, 2014). Advanced multi-articulating prosthetic hands produced by the DARPA Revolutionizing Prosthetics Initiatives 2007 and 2009, such as the DEKA "Luke Arm" and the Johns Hopkins University Applied Physics Lab "Modular Prosthetic Limb" (Otto, 2013), are highly sensorized, but access to these systems remains limited.

Since affordable, off-the-shelf solutions were unavailable, the BairClaw testbed was designed to be highly sensorized for both proprioception and multimodal tactile sensation from the ground up. Joint angles are measured directly at each joint and with minimal drift. As a result, the BairClaw proprioception system enables accurate joint angle tracking without having to cycle the hand through neutral postures to reset postural baselines, as with commercially available prosthetic hands. The proprioception system utilizes inexpensive ring magnets and Hall effect sensors to achieve joint angle resolution of just over a tenth of a degree at 100-200 Hz sampling rates. The BioTac is capable of measuring multiple types of graded tactile information at data rates of 100-4400 Hz. The BairClaw testbed is limited in that the system requires tethered power and users cannot don the bulky actuation system. While the entire testbed was not designed to be donned by subjects, the hand itself can be mounted to a lightweight test socket for whole arm experiments. The inclusion of rich proprioceptive and tactile sensing will enable the study of action-perception relationships, the development of new feedback control schemes, and the ability to provide amputees with simultaneous proprioceptive and tactile sensory feedback via cutting edge neural interface techniques.

Applications of the BairClaw Testbed to Neurorehabilitation of Body Schema

The hope that a mechanical or robotic system could become part of one's body schema is hardly a new idea. Mirror visual feedback and graded motor imagery techniques are established methods for the manipulation of body schema through visual feedback alone. The use of MVF is an important paradigm shift in the treatment of neurological damage to the brain and peripheral nervous system, as the technique seeks to take advantage of the dynamic restructuring capabilities of the brain to manipulate body schema. It is believed that the illusory influence of visual feedback can be further enhanced by simultaneous proprioceptive and tactile feedback that is congruent with what is being seen.

Closing the Somatosensory Loop with Amputees

While the development of neural interfaces for proprioceptive and tactile feedback is hardly a solved problem, promising new techniques and highly encouraging findings are being reported. For instance, intrafascicular multichannel electrodes inserted into median and ulnar nerves have been used to provide real-time sensory feedback of a bidirectional prosthetic hand (Raspopovic et al., 2014). Force sensors at prosthetic fingertips were used to drive electrode stimulation currents. The amputee subject was able to exploit features of the dynamic, graded tactile feedback, such as rates of change of current amplitude and differential timing of contact across the hand, to distinguish between objects based on stiffness and shape, respectively.

More recently, researchers have used selective, nonpenetrating peripheral nerve cuff electrodes to stimulate residual upper limb nerves in unilateral amputees (D. W. Tan et al., 2014). Using a systems identification approach, they were able to elicit long-term stable, graded, natural percepts including tapping, constant pressure, vibration, and even light moving touch, all of which could be driven by a highly sensorized testbed, such as the BairClaw. Percept area and intensity could be modulated via stimulation intensity and frequency, respectively. Percept sites were numerous, independent, well-defined, and even included sites on fingertips. Proprioceptive percepts remain to be systematically explored and mapped.

Using closed-loop feedback, subjects were able to accomplish dexterous tasks while blindfolded. In addition to the functional benefits enabled by the sensory system, there were positive embodiment-related and therapeutic effects as well. According to page 9 of (D. W. Tan et al., 2014), "When sensation was active, both subjects perceived the hand and prosthetic hand to be nearly perfectly colocated in space. When sensation was not active, the prosthesis was viewed by the subjects as a tool that extended beyond their hands." Although further investigation is required, it is exciting that both subjects reported the elimination of phantom limb pain with the use of the sensory feedback system.

Other researchers have recently used a non-invasive tendon vibration technique to elicit percepts of joint-specific movement via the "Kinesthetic Illusion" effect (Marasco, 2014a, 2014b). It is well-known that tendon vibration creates an illusion of muscle lengthening (Lackner, 1988). Working with amputees who had undergone targeted sensory reinnervation, researchers were able to vibrate reinnervated muscle to produce the percepts of different, gross hand postures, including a cylinder grip, precision pinch, and opening of the hand.

While methods for communication between artificial hands and the human nervous system continue to improve, further investigation is needed to address gaps that remain. For instance, a myriad of high resolution joint angles can easily be obtained from highly sensorized artificial testbeds, such as the BairClaw. However, it is unclear how to convey this detailed postural information via coarse methods for artificial proprioceptive feedback such as tendon vibration, which has been recently used to convey a small number of gross hand postures. Furthermore, the lack of validated, objective functional

outcome measures for upper-limb myoelectric prosthesis use, let alone bidirectional prosthesis use, makes it difficult for researchers to relate the quality of artificial sensory feedback to improvements in quality of life (Hill et al., 2009; Wright, 2006).

Extending Vision-Based Therapies with a High-Tech Rubber Hand Illusion

The high occurrence of phantom limb pain and proprioceptive disorders may be due to the lack of embodiment, or a disrupted sense of ownership due to mismatches between different modalities of sensory feedback, such as touch and vision (Harris, 1999). Prior efforts to improve the embodiment of prosthetic devices have focused on appearance. Although visual appearance is extremely important to amputees for both embodiment and interactions with others, it is a fragile illusion. Surveys have shown that sensory feedback is often ranked as a higher priority than life-like appearance for powered prostheses (E. Biddiss, Beaton, & Chau, 2007; E. Biddiss & Chau, 2007). Through artificial proprioceptive and exteroceptive feedback to a phantom limb, an amputee could develop and maintain an internal model of a neuroprostheses as part of his/her body schema.

It has already been reported that phantom limb pain could be reduced when sensory substitution via electrotactile stimulation was used to provide feedback on grip force to transradial amputees. Clinically relevant improvements were observed even after a short two week training period (Dietrich et al., 2012). It has been postulated that, based on the increased functionality and decreased phantom limb pain that was observed, a cortical reorganization likely occurred ("Touch Bionics i-limb hand," 2009). Thus, it may

pbe possible to rehabilitate body schema and reduce chronic pain through the dynamic restructuring of neuronal pathways in the brain, even with sensory substitution techniques.

In this chapter, we demonstrated the use of the BairClaw testbed for a tap-and-hold experiment that resembles a task that might be done in a mirror visual feedback therapy session. The BairClaw testbed has also been used to produce stroking motions for haptic exploration of surfaces (supplemental video) and to track fingertip forces from a nonhuman primate precision grip task (Hellman, Chang, Tanner, Helms Tillery, & Santos, 2014). Preliminary results suggest that the BairClaw is capable of fine fingertip force control at physiologically meaningful magnitudes and timescales.

Summary

Established therapeutic techniques such as mirror visual feedback and graded motor imagery rely purely on visual feedback or imagined action, respectively. The rubber hand illusion demonstrates the power of visual feedback combined with somatosensory feedback. Advances in artificial hand technology and techniques for providing proprioceptive and exteroceptive feedback now make it possible to combine mirror visual feedback with a high-tech version of the rubber hand illusion. The efficacy of visual manipulations for neurorehabilitation could be enhanced if coupled with proprioceptive and tactile feedback in a controlled therapy environment.

Even at this nascent stage of neural interface development, reports of natural proprioceptive and tactile percepts from upper limb amputees are highly encouraging. It will soon be possible to provide graded, natural percepts to amputees that could be driven

directly by joint angle encoders, tactile sensors, skin stretch sensors, thermistors, etc. There are numerous potential benefits of enhanced embodiment by way of congruent multisensory feedback that is spatiotemporally consistent with commanded actions: reduction of phantom limb pain, a renewed sense of ownership, stronger connections to others and to society, and increased use of a prosthesis due to improved functionality and reduced cognitive burden. We believe that highly sensorized testbeds such as the BairClaw can be used to enhance the embodiment of a neuroprosthesis into one's body schema, and be used to probe the complex relationships between sensory feedback, illusion-based percepts, and body schema manipulation.

CHAPTER 3

DEVELOPMENT OF PROBABILISTIC MODELS FOR REAL-TIME PERCEPTION OF GEOMETRIC FEATURES WITH A SENSORIZED ARTIFICIAL FINGER

Introduction

Humans use haptic "exploratory procedures" (EPs) to extract object properties using the sense of touch (Lederman & Klatzky, 1987). Geometric features such as curvature, aspect ratio, and edge orientation can be extracted using EPs such as whole hand enclosure and contour following (*Figure 1.1*). Haptic perception often enables object identification through touch alone (Plaisier, Bergmann Tiest, & Kappers, 2009).

Previously, we developed a support vector regression model to perceive edge orientation with respect to a body- fixed, fingertip reference frame (Plaisier et al., 2009). With multimodal data from a bio-inspired tactile sensor (BioTac, SynTouch LLC), the support vector regression model was able to perceive edge orientation with accuracies similar to that of the human finger (R. D. Ponce Wong et al., 2014). However, one limitation of the model was that it required some degree of post-processing; carefully constructed temporal "windows" of sensor data had to be extracted and further processed to yield model inputs. Post-processing of the sensor data makes it difficult to implement these models in real-time decision-making applications, such as contour-following of object features at human-like speeds.

In this work, we develop Hidden Markov Models (HMMs) that can be used as a foundation for real-time decision-making. HMMs have been applied extensively in speech recognition applications and have been successfully used to model temporal data

(Rabiner, 1989). Importantly, HMMs enable the modeling of observations generated by stochastic processes where the states are hidden. HMMs also enable perception updating as new observations are collected. Thus, one might be able to identify the orientation of an edge, before an EP has been completed and without the need for post-processing the sensor data.

The majority of this chapter is an extension of previously published work (Hellman & Santos, 2015; R. D. Ponce Wong et al., 2014).

Methods

Experimental Apparatus

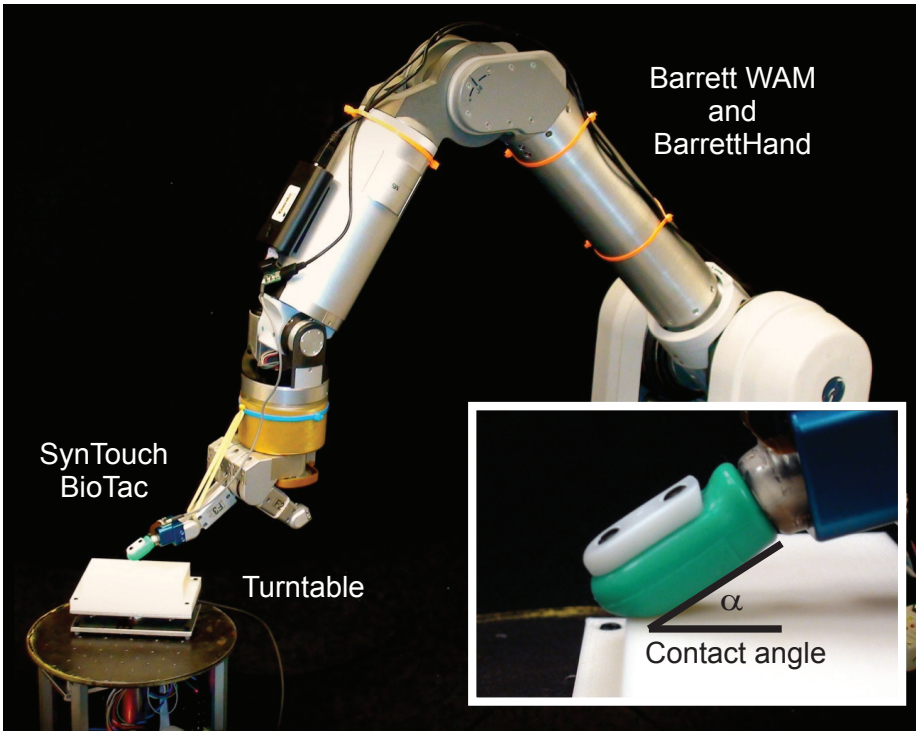

Figure 3.1 From (R. D. Ponce Wong et al., 2014)The Barrett WAM, BarrettHand, and BioTac were used to explore edge stimuli presented at random orientation with respect to the fingertip reference frame by a motor-driven turntable.

A subset of multimodal sensor data from (R. D. Ponce Wong et al., 2014) were used in this study. Briefly, the data were collected using a robot testbed (*Figure 3.1*) consisting of a Whole Arm Manipulator and BarrettHand (Barrett Technology, Cambridge MA) in which a single digit had been outfitted with a BioTac sensor. The artificial fingertip was used to haptically explore a 0.4 cm thick blade using pre-planned trajectories. The edge was randomly oriented at angles ranging from 90° to -90° in 1°

increments using the angle definition shown in *Figure 3.2*a. A radial to ulnar stroke (*Figure 3.2a,* was conducted at fingertip speeds of 2 and 4 cm/s, and at two contact pressures that can be considered low and high.

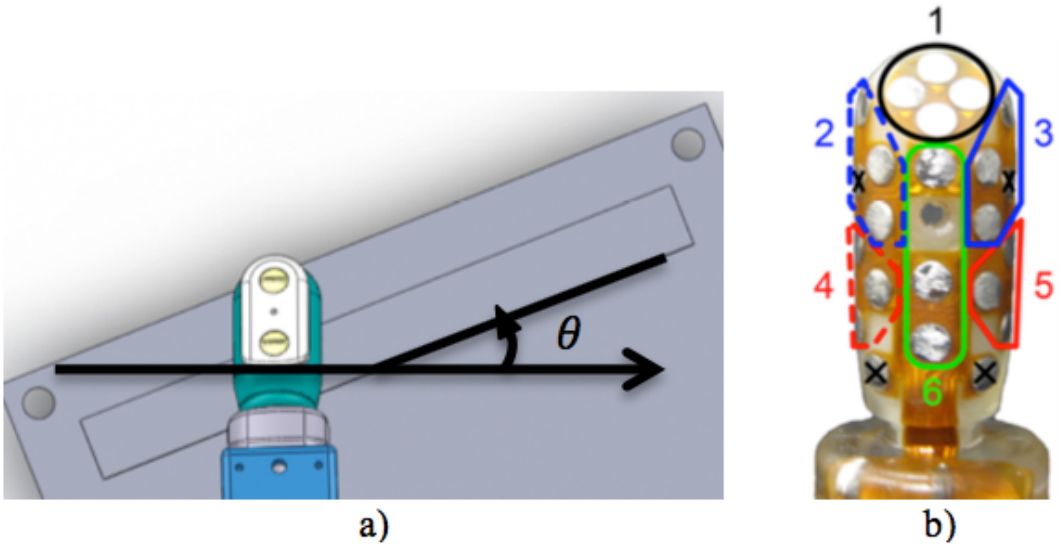

Figure 3.2 From (R. D. Ponce Wong et al., 2014), a) The arrow indicates the radial to ulnar motion of a right-handed index finger across a thick blade that is at an angle of Θ with respect to a body-fixed fingertip reference frame. b) The BioTac sensor electrodes were clustered into 6 regions of the fingertip.

Processing of Tactile Sensor Data

The BioTac contains 19 electrodes, sampled at 100 Hz, which can be used to extract information about the elastomeric skin deformation relative to the rigid sensor core. As in (R. D. Ponce Wong et al., 2014), the electrodes were clustered into 6 groups based on fingertip location (*Figure 3.2b*). In this study, we did not use specific temporal windows of data to extract model inputs. Rather a 12-element vector was constructed for each timepoint (in 10 ms increments) that consisted of mean impedance and differential changes in mean impedance for each of the six clusters.

Hidden Markov Models

Hidden Markov Models (HMM) have been used extensively in speech recognition and are statistical methods of representing temporal sequences of complex data. They have been successfully applied to stochastic signal modeling (Rabiner, 1989). HMMs are a special form of a Markov chain; A Markov chain is conditionally independent of other observations given the current state. This means that the probability of being in a state is only dependent on its current state. HMM model parameters $\lambda = (A, B, \pi)$ are specified by a transition matrix A, belief matrix B, and an initial state distribution π. The transition matrix A is the probability of transitioning to a state given the current state. The belief matrix B is the probability of an observation given the current state, and the initial state distribution π is the probability of starting in a given state.

The fundamental idea behind using a statistical model is to determine the probability of an object feature given a sequence of observations X Eq. (3.1). By using

Bayes' Theorem Eq. (3.2), we can consider the impact of an observation and the belief of possible features. The posterior is the left side of Bayes' Theorem and is proportional to the likelihood times the prior Eq. (3.3). Eq. (3.4) shows how we can determine the likelihood of a given model where $p(X|E)$ is considered the sensor model and $p(E)$ is the salient feature model. By using Eq. (3.4) in place of Eq. (3.1) we return the results of the probability of observing a given feature from information that is available through prior knowledge and the observation sequence.

$$E^* = arg\max_E p(E|X) \quad 3.1$$

$$p(E|X) = \frac{p(X|E)p(E)}{p(X)} \quad 3.2$$

$$p(E|X) \propto p(X|E)\ p(E) \quad 3.3$$

$$E^* = arg\max_E p(X|E)\ p(E) \quad 3.4$$

There are many distinct types of HMMs, such as an ergodic model in which every state can be reached from every other state (transition matrix shown in *Figure 3.3*a). Other models include n-state left-right models where the number of states *n* determines the size of the forward state jump that can occur (2-state left-right model transition matrix shown in *Figure 3.3*b). For the models considered here we use a left-right state model (*Figure 3.3*), which has been shown to fit stochastic sequential models well (Gales & Young, 2007; Rabiner, 1989).

a) Ergodic model transition coeff.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

b) Left-right model transition coeff.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{23} & 0 \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{bmatrix}$$

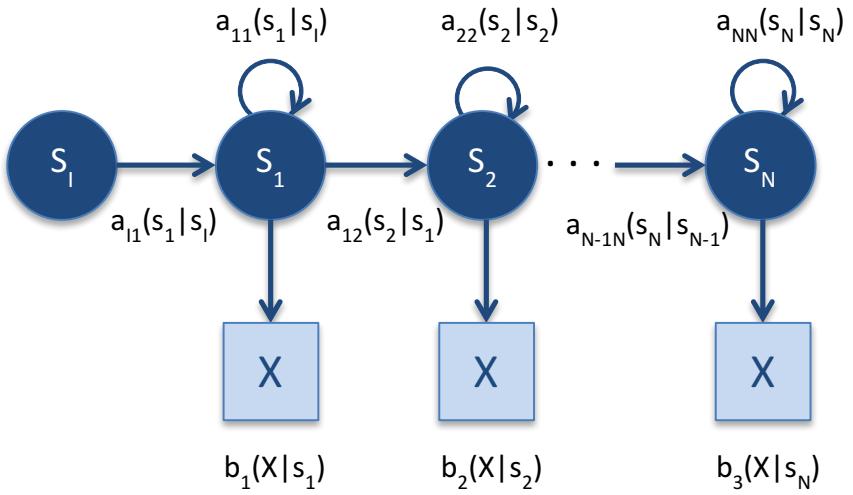

Figure 3.3 Continuous observation HMM. Circles are discrete hidden nodes and squares are continuous observations. *X* is the observation vector. Transition matrices show the structure for an (a) ergodic and (b) 2-state left-right HMM.

HMMs can have discrete or continuous observations. Discrete observation HMMs require a code book for a given observation and have shown success in speech recognition with a large corpus (Rabiner, 1989). The continuous emission model can be either a single multivariate Gaussian Eq. (3.5) or an M-component Gaussian mixture model Eq. (3.6). The M-component Gaussian mixture is the one used in the preliminary study presented here due to its ability to model multiple observation sequences for each state belief matrix.

$$b_j(x) = p(x|s_j) = \mathcal{N}(x; \mu^j, \Sigma^j) \quad 3.5$$

$$b_j(x) = p(x|s_j) = \sum_{m=1}^{M} c_{jm} \mathcal{N}(x; \mu^{jm}, \Sigma^{jm}) \quad 3.6$$

To successfully apply Hidden Markov Models, we need to first build the signal model which can then be used to determine the likelihood that a signal fits a specific model. Model likelihood will be described first as it is necessary to drive the tuning of the model parameters.

Likelihood

The likelihood of a given model $p(X|\lambda)$ can be determined recursively. The forward probability is denoted as α and is the probability of observing the observation sequence *X* at time *t*. The forward algorithm Eq. (3.7) can determine the probability of an observation sequence on the order of N^2T which is computationally feasible for large values of the number of states *N* and length of the datastream *T*. After initialization Eq. (3.7-3.9) the forward probability is solved through induction Eq. (3.9) and then terminated Eq. (3.10) at the end of the observation sequence of length *T*. The forward algorithm is necessary to solve the probability of an observation sequence in a tractable manner. If we wanted to directly solve the probability of an observation sequence we would need to enumerate over all possible state sequences, which is intractable for even small values of *N* and *T* on the order of $2T * N^T$. The forward algorithm is visualized in *Figure 3.4*, which shows the sequence of operations based on the lattice or trellis structure (Rabiner, 1989).

Initialization:

$$\propto_0 (s_i) = 1 \tag{3.7}$$

$$\propto_0 (s_j) = 0 \quad if\ s_i \neq s_j \tag{3.8}$$

Induction:

$$\propto_t (s_j) = \left[\sum_{i=1}^{N} \propto_{t-1} (s_i) a_{ij}\right] b_j(x_t) \tag{3.9}$$

Termination:

$$p(X|\lambda) = \sum_{i=1}^{N} \propto_T (s_i) \tag{3.10}$$

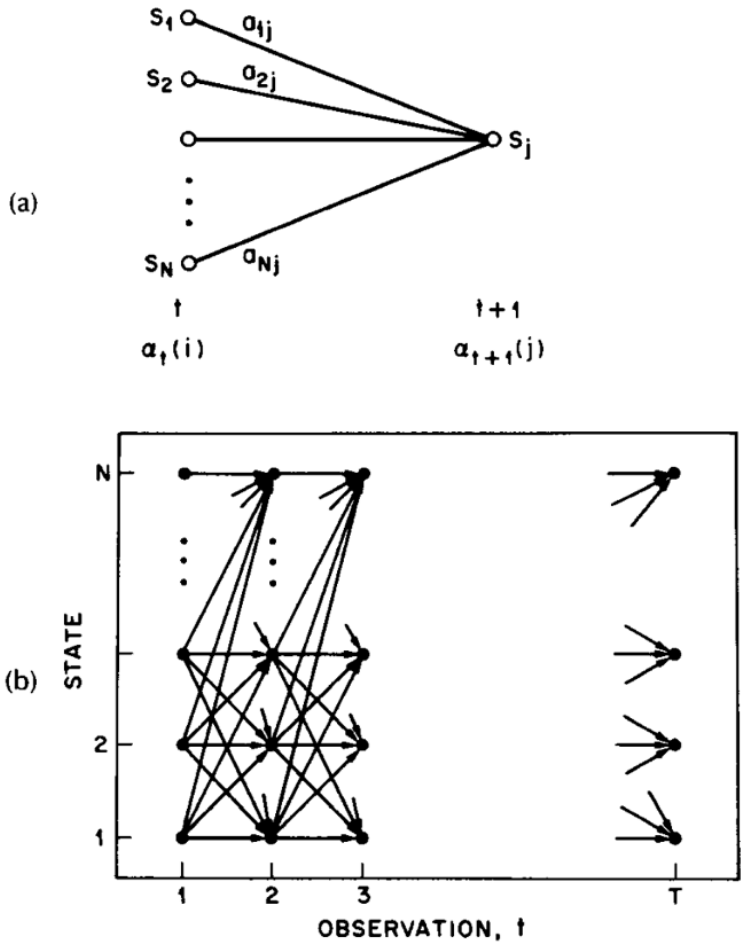

Figure 3.4 a) Shows the sequence of operations required to compute the forward variable $\alpha_{t+1}(j)$. b) Illustration of the complete process for the forward algorithm. Excerpt from (Rabiner, 1989).

We introduce the backward probability here due to its similarity to the forward algorithm. The backward algorithm is primarily used in the solution to model training and will return the same result as the forward algorithm for an observation sequence. The backward algorithm Eq. (3.11-3.12) is the backward probability of the partial observation sequence from t+1 to the end (*Figure 3.5*).

$$\beta_t(i) = \sum_{i=1}^{N} a_{ij} b_j(x_{t+1}) \beta_{t+1}(j) \tag{3.11}$$

$$t = T-1, T-2, \cdots, 1, \qquad 1 \leq i \leq N \tag{3.12}$$

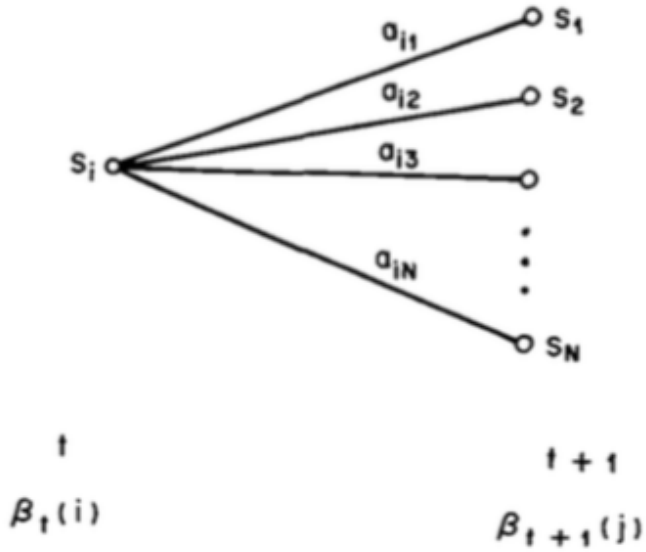

Figure 3.5 Visualization of the calculation of the backward variable β. Excerpt from (Rabiner, 1989).

Sequence

There are a few options when determining the best sequence of states for a given observation sequence. One could possibly choose the state sequence where each state is individually most likely given the observations, although that might not be the best considering the structure of the HMM and transition matrix. The Viterbi algorithm is a method to find the single best state sequence for a given observation sequence (Viterbi, 1967). We define the quantity V_t in Eq. (3.16) as the highest probability along a single path at time *t*. The algorithm is similar to that of the forward algorithm but determines the maximum likelihood for each time step and then keeps an array Eq. (3.17) of backpointers to store the most probable path. The back trace variable *bt* Eq. (3.15) is the array of backpointers that record the best state sequence given by the Viterbi algorithm.

Initialization:

$$V_0(s_i) = 1 \tag{3.13}$$

$$V_0(s_j) = 1 \quad if\ s_i \neq s_j \tag{3.14}$$

$$bt_0(s_j) = 0 \tag{3.15}$$

Recursion:

$$V_t(s_j) = \max_{i=1} V_{t-1}(s_i) a_{ij} b_j(x_t) \tag{3.16}$$

$$bt_t(s_j) = \operatorname*{argmax}_{i=1} V_{t-1}(s_i) a_{ij} b_j(x_t) \tag{3.17}$$

$$P^* = \max_{1 \leq i \leq N} V_T(s_i) \tag{3.18}$$

$$s_T^* = \operatorname*{argmax}_{1 \leq i \leq N} V_T(s_i) \tag{3.19}$$

The Viterbi algorithm can be very helpful in determining the ability of a model to fit a complex signal. The algorithm can provide insight into the performance of an HMM by showing the sequence of most likely state transitions from the trained model and then comparing it to how well it fits a known observation state sequence. Essentially, by inputting a known observation sequence through a trained HMM and comparing the known and predicted best sequence of states side-by-side we can gain an empirical estimate of model accuracy.

Tuning

Tuning, or training of model parameters, is the hardest of the three steps when developing an HMM. There is no analytical solution to solve this problem, although we can use an iterative procedure such as the Baum-Welch method, which is an expectation maximization (EM) algorithm (Rabiner, 1989). Baum-Welch allows the selection of model parameters $\lambda = (A, B, \pi)$ such that $P(X|\lambda)$ is locally maximized. The E-step (expectation) is an estimate of the state occupation probability. The M-step (maximization) is a re-estimation of the HMM parameters based on the estimate of the state occupation probability. The state occupation probability $\gamma_t(s_i)$ Eq. (3.20) is the probability of occupying state *i* at time *t* with respect to both the forward and backward probabilities. The EM equations presented will be for a mixture of Gaussians Eq. (3.6). To obtain the same equations for a single Gaussian model Eq. (3.5), the number of Gaussians *m* is set to 1.

$$\gamma_t(s_i) = \frac{\alpha_t(s_i)\beta_t(s_i)}{P(X|\lambda)} = \frac{\alpha_t(s_i)\beta_t(s_i)}{\sum_{i=1}^{N} \alpha_t(s_i)\,\beta_t(s_i)} \quad 3.20$$

We also need to define $\xi_t(s_i, s_j)$ Eq. (3.21), which is the probability of being in state *i* at time *t* and state *j* at time *t+1*.

$$\xi_t(s_i, s_j) = P(\, S(t) = s_i,\ S(t+1) = s_j | X,\ \lambda\,) \quad 3.21$$

$$\xi_t(s_i, s_j) = \frac{\alpha_t(s_i) a_{ij} b_j(x_{t+1}) \beta_{t+1}(s_j)}{\alpha_T(s_E)} \quad 3.22$$

The sum of the occupation probabilities through time is used to re-estimate the Gaussian parameters mean Eq. (3.23) and covariance Eq. (3.24).

$$\hat{\mu}^{jm} = \frac{\sum_{t=1}^{T} \gamma_t\,(s_j, m) x_t}{\sum_{t=1}^{T} \gamma_t\,(s_j, m)} \quad 3.23$$

$$\hat{\Sigma}^{jm} = \frac{\sum_{t=1}^{T} \gamma_t\,(s_j, m)(x_t - \hat{\mu}^{jm})(x_t - \hat{\mu}^{jm})'}{\sum_{t=1}^{T} \gamma_t\,(s_j, m)} \quad 3.24$$

The mixture coefficients Eq. (3.25) are re-estimated in a similar way to the Gaussian parameters and are needed so that the beliefs remain stochastic.

$$\hat{c}_{jm} = \frac{\sum_{t=1}^{T} \gamma_t\,(s_j, m)}{\sum_{l=1}^{M} \sum_{t=1}^{T} \gamma_t\,(s_j, l)} \quad 3.25$$

Using the state occupation probability we can also re-estimate the transition probabilities Eq. (3.26-3.27).

$$\hat{a}_{ij} = \frac{expected\ number\ of\ transitions\ from\ s_i\ to\ s_j}{expected\ number\ of\ transitions\ from\ s_i} \quad 3.26$$

$$\hat{a}_{ij} = \frac{\sum_{t=1}^{T} \xi_t\,(s_i, s_j)}{\sum_{k=1}^{N} \sum_{t=1}^{T} \xi_t\,(s_i, s_k)} = \frac{\sum_{t=1}^{T} \xi_t\,(s_i, s_j)}{\sum_{t=1}^{T} \gamma_t\,(s_i)} \quad 3.27$$

Using the EM Baum-Welch equations shown here (sometimes called forward-backward algorithm) we can iteratively solve for $\bar{\lambda}$, which is the re-estimation of the HMM model parameters λ that improve the probability of the observation sequence from the model until the model likelihood converges.

Statistical Feature Models to Predict Edge Orientation

When applying HMMs to tactile data the belief matrix is a mixture of Gaussians that contains the probability of an observation (6 mean impedance and 6 differential changes in mean impedance values) given the current state.

The state of an HMM is an abstraction of the feature space of the data. To avoid overfitting of the model, it is necessary to limit the number of states. In this work, we investigated the effects of the number of states on model performance. Nine different classification models were examined, with states ranging in number from 2-10.

A total of 1445 data samples were used from (R. D. Ponce Wong et al., 2014). A threshold on the internal fluid pressure of the BioTac sensor was used to define the start and end of contact between the fingertip and the edge stimulus. The [-90°, 90°] range of edge orientation angles was split into 36 equally sized 5° increment bins. Thus, for each classification model, a total of 36 independent HMMs were built (one HMM per bin). Within each bin, 90% of the data were randomly selected for a training set, and the remaining 10% of the data were reserved for testing of the classification model.

Each of the HMM's belief and transition matrices were initialized with a random training observation from each bin. The initial state distributions were set to one for the first state and zero for all others. The models were then run through the Baum-Welch

algorithm, which adjusts model parameters based on expectation-modification (Rabiner, 1989). At each incremental time point for a given datastream that includes all data from initial contact, the classifier determined the most likely bin based on which bin-specific HMM yielded the maximum log-likelihood value. Each of the 9 classification models was tested with a novel dataset.

Results and Discussion

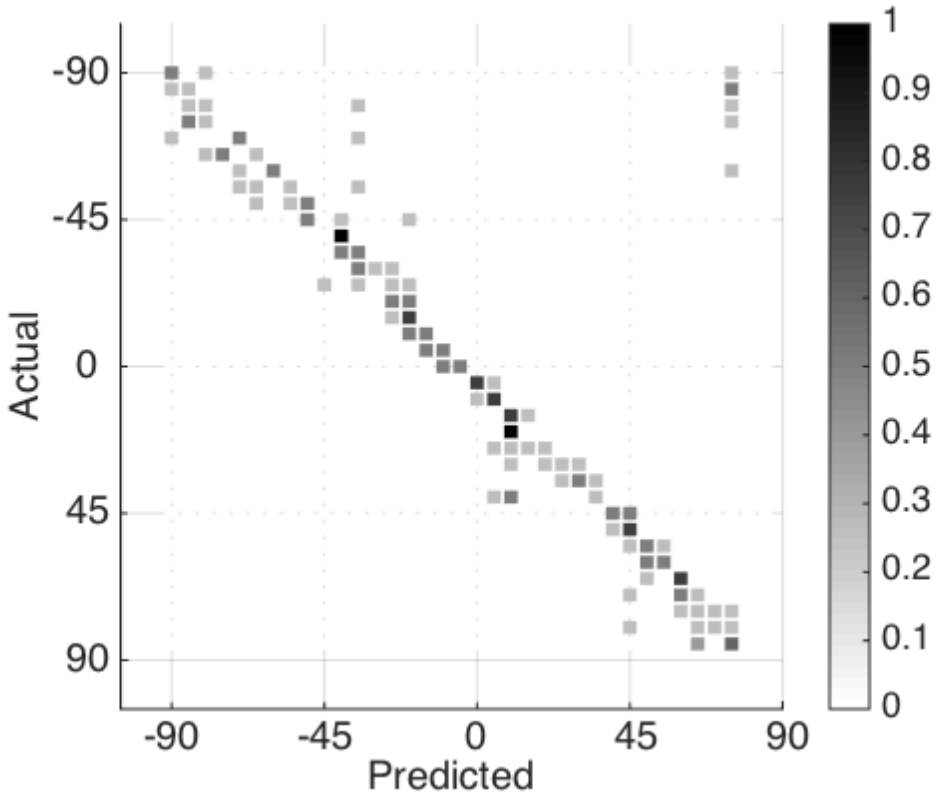

Figure 3.6 Confusion matrix for an 8-state classification model applied to a test dataset.

The 8-state model performed the best with probabilities of 0.35 and 0.81 for accuracies of ±2.5° (one bin) and ±12.5° (one bin ± 2 bins), respectively (*Figure 3.6*). The 6-state model produced the highest probability of 0.90 for an accuracy of ±12.5°. Large prediction errors tended to be near -90° and 90° (*Figure 3.6*). It may be that a radial to ulnar stroke against the thick blade generated similar sensor data for the -90° and 90° orientations, which would make these angles appear similar to the classification model.

While these preliminary probabilistic models could be improved, the results constitute a promising first step toward real-time decision-making based on haptic perception. Real-time calculations of log-likelihood for a given model are on the order of N^2T where N is the number of states and T is the length of the datastream (R. D. Ponce Wong et al., 2014). Modern processors are more than capable of determining log likelihood in real-time. It should be noted that the bin-specific HMMs were robust to differences in fingertip speed and contact pressure. Using a single HMM per bin reduces the computational expense of the log-likelihood calculations and speeds up the classification process. Prediction accuracy would likely improve with a more comprehensive training dataset, additional tuning of model parameters such as observation size and sliding window spacing, and modification of model prior distributions.

Conclusions and Future Work

In this study, HMMs were used to develop models that can classify edge orientation relative to a body-fixed, fingertip reference frame in real-time. Such models could be used as a foundation for contour following at human-like speeds. Future work includes expanding the probabilistic models beyond edge orientation and adding models for other salient geometric features, such as bumps and pits (Ruben D. Ponce Wong, Hellman, & Santos, 2014). By leveraging probabilistic HMM techniques used in speech recognition for grammar models, we can develop the grammar of haptic exploration.

CHAPTER 4

FUNCTIONAL CONTOUR-FOLLOWING VIA HAPTIC PERCEPTION AND REINFORCEMENT LEARNING

Introduction

Contour-following is often necessary for tasks requiring fine manipulation skills such as tying shoelaces, searching pockets, and buttoning shirts. The ability to perform haptic contour-following enables specific dexterous abilities that are necessary for capable robotic systems. In this chapter, we present a functional form of contour-following in which the contour is the zipper on a deformable ziplock bag to be closed.

Early work in contour-following used contact forces to track edges (Chen, Zhang, & Rink, 1995). Scene and tactile images of the contacted interface were filtered to determine edges for vision-based contour-following (Nakhaeinia, Payeur, & Laganiere, 2014). Techniques integrating vision, force sensing, and accelerometers have also been used to track contours (Koch, Konig, Weigl-Seitz, Kleinmann, & Suchy, 2013). The effectiveness of these computer vision approaches can be limited if target objects are occluded, transparent, deformable, or otherwise optically ambiguous (Irani, Rousso, & Peleg, 1994). In such scenarios, it becomes necessary to rely on haptic and proprioceptive feedback in order to perform functional tasks.

Robotic systems have various levels of proprioceptive precision. For instance, lash in tendon-driven systems or static friction in joints can lead to noisy torque estimates. Compliance in soft robot arms can mask finger-object interactions that would otherwise be observable with precise, direct-drive robot arms (Pratt & Williamson, 1995). As robotic devices become more common place in our surroundings they will begin to

interact in ways similar to their human counterparts. This interaction will involve incidental and planned contact with not only the end effector but other areas of the robotic system as well (Bhattacharjee, Shenoi, Park, Rehg, & Kemp, 2015). From these incidental contacts, proprioceptive data that could have been used to calculate localized contact information will become indeterminate. This will happen as robots start to manipulate objects in pockets and bags due to incidental and/or planned contacts away from the manipulator, such as on the forearm or upper regions of the arm and hand. In this work, we rely heavily upon tactile percepts that are localized near the point of finger-object contact, as visual and proprioceptive feedback can be noisy or intermittent.

Tasks that require contour-following are not limited to manipulation of rigid objects. The manipulation of deformable objects such as rope, cloth, and sponges have been a focus of extensive research due to their applications in surgical and service robotics (Lazher, Belhassen-Chedli, Sabourin, & Youcef, 2014; Salzmann, Pilet, Ilic, & Fua, 2007; Schulman, Lee, Ho, & Abbeel, 2013). Both physics-based and model-free approaches have been successfully implemented (Berenson, 2013; Bretl & McCarthy, 2013; Maitin-Shepard, Cusumano-Towner, Lei, & Abbeel, 2010; Schulman et al., 2013; Shah & Shah, 2016). In the work presented here, we use a model-free approach in which the stochastic nature of physical interactions with a deformable object is learned from action-driven haptic and proprioceptive sensory feedback. Using reinforcement learning we explore the benefits of Contextual Multi-Armed Bandits and compare against a Q-learning benchmark. Learning algorithms are shown to be capable of intelligently selecting actions to maximize cumulative rewards.

Importance of Local Contact

Manipulating objects based purely on proprioceptive and visual information is often challenging due to compliance in robotic systems and/or objects being manipulated. It is unlikely that a rigid robotic system could perfectly implement planned motions as there is typically uncertainty in joint angles (Klingensmith, Sirinivasa, & Kaess, 2016). By sensing at the point of contact we remove the possibility of error propagation in forward calculations from both visual and proprioceptive systems. Having a compliant fingertip sensor at the point of the grasp allows it to conform to object features. Localized sensing at the point of contact is robust to uncertainty due to incidental contact on the robot system, proprioceptive sensor errors, or visual occlusions.

The use of vision for the complete execution of a manipulation task will be challenging as objects become occluded. While invaluable for task planning, vision is not a panacea. Occlusions caused by the arm and hand are inevitable and will hamper the robustness of systems that rely purely on vision for feedback control.

Real-Time Perception During Task Execution

In order to complete a functional manipulation task, the robotic system has to perceive the current state and then take the best action based on previous experiences or a learned policy. By perceiving the state of the system in real-time, corrective actions can be taken to achieve the goal of the manipulation task. The method of real-time perception of features differs from previous work presented in Chapter 2. Work presented in Chapter

2 used known exploratory procedures (EPs) (*Figure 1.1*) and tactile data from complete fingertip trajectories to perceive edge orientation (Lederman & Klatzky, 1987).

The EPs, which are mostly used to determine object properties, rely on known trajectories with large displacements of the finger relative to the object size. These known trajectories simplify the ability of an algorithm to successfully classify object features. The classification of these longer trajectories or EPs is often easier than classification of shorter trajectories because the signals generated are more predictable and distinguishable from noise and transients for a complete trajectory. It is typical to start a pre-planned fingertip trajectory with the sensor in free space and not in contact with the object. By beginning a trajectory with the sensor in free space, strain in the compliant fingerpad is mechanically reset to baseline conditions. However, when the goal is to complete a manipulation task, the larger displacement trajectories may not be an option and may actually interfere with the successful completion of a task. In contrast, shorter motions are more realistic by requiring the classifier to more frequently determine the state of the system. With more frequent real-time perception, however, accurate classification is more difficult than that of longer EP trajectories. With a compliant sensor, the signal response and initial state may be completely different given the set of prior actions taken. This difference in the raw sensor baseline signals during each motion adds to the difficulty in classifying salient object features during real-time perception.

For the development of real-time manipulation capabilities, it is necessary to update percepts after relatively small windows of time and displacements. By updating percepts at a higher rate, after smaller displacements, corrective actions can be taken sooner and more often.

Task Description

The closure of a plastic ziplock bag is an everyday task that presents unique challenges for robotic systems. The task requires the manipulation of a transparent, deformable object whose geometric features are visually occluded by artificial fingertips. Computer vision could be useful for planning, but will likely be insufficient for task execution. While the closure of a plastic bag is typically a bimanual task, we simplified the experimental set-up to focus on the control of a single sensorized robot hand, wrist, and arm.

To successfully complete the task, we will need to determine the current orientation of the zipper in the grasp and then take an appropriate action. The action chosen should either maintain the current state or bring the system into a new desired state which will have the highest probability of a successful zipper closure.

Pilot Study

A preliminary study that employed Q-learning was conducted to determine the validity of applying reinforcement learning to a functional contour-following task. The preliminary work presented in this section was used to refine many of the system parameters that will be presented in the following sections of this chapter.

Methods

Robot Testbed

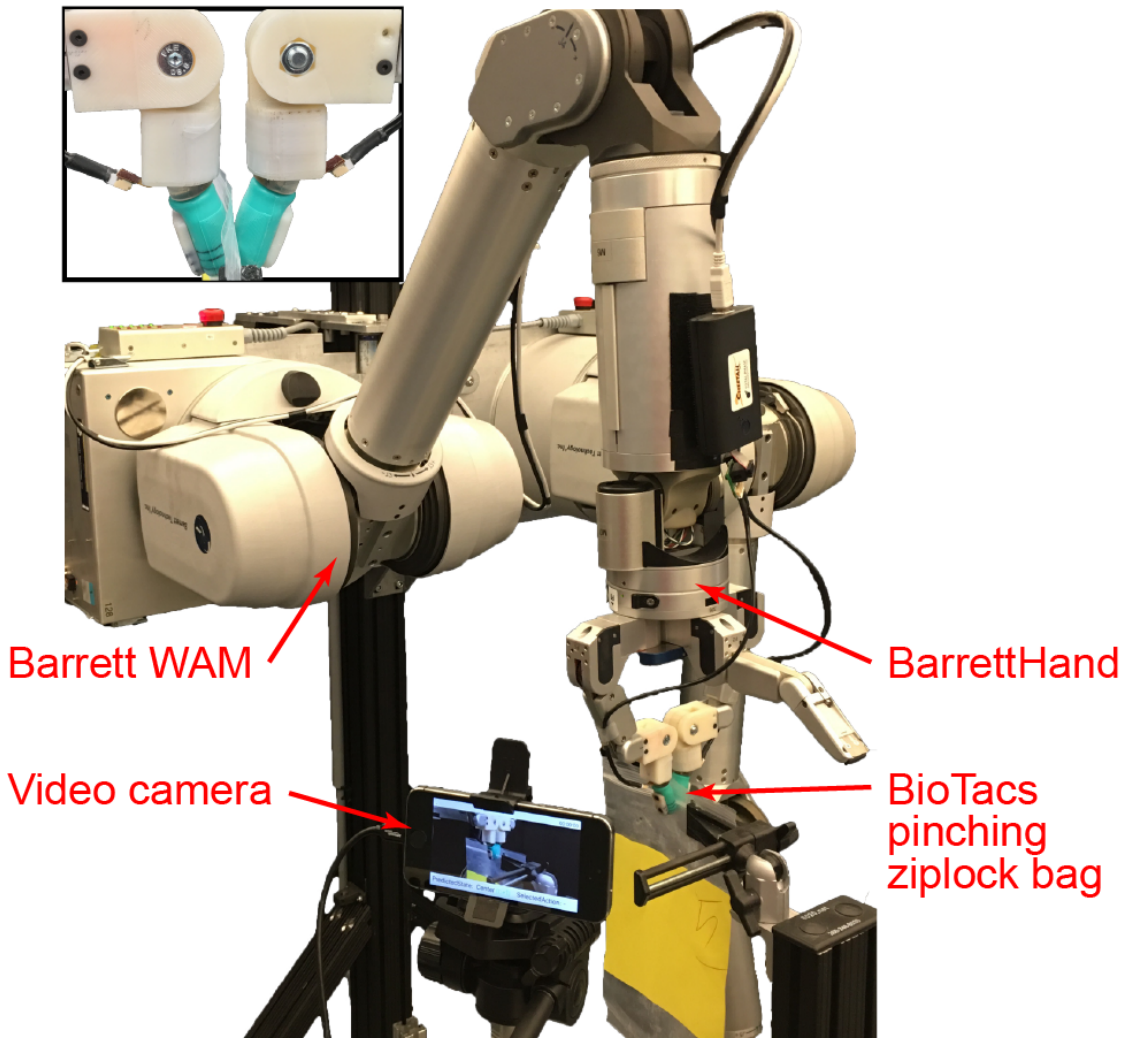

Figure 4.1 Experimental set-up used to develop online haptic perception and decision-making capabilities for a functional contour-following task.

The experimental set-up required the integration of multiple robotic systems, sensors, and processors. Processing was separated into hard and soft real-time modules. Hard real-time processes were controlled from a Linux computer running Ubuntu 12.04

modified with a Xenomai kernel. For reliable control and sensing, command signals and sensor sampling were handled by hard real-time modules. Computational tasks such as haptic perception, decision-making, and reinforcement learning were implemented by soft real-time processes on a 2013 MacBook Pro. By isolating the soft real-time processes, the load variability on the hard real-time threads was reduced since certain algorithms and tasks did not have guaranteed execution times.

The robot testbed consisted of a 7 DOF Barrett Whole Arm Manipulator (WAM) with a 4 DOF BarrettHand (Barrett Technology, Cambridge, MA). The BarrettHand can independently flex/extend three digits and adduct/abduct the outer two digits. BioTac sensors (SynTouch LLC, Los Angeles, CA) were attached to two opposing digits on the hand (*Figure 4.1*) for the recording of internal fluid pressure, vibration, skin deformation, and temperature (Wettels et al., 2008). Due to the redundant information from the two BioTacs in this particular experiment, only a single BioTac was used for analysis. Electrode impedance (*Figure 4.2*) was sampled at 100 Hz and internal fluid pressure was sampled at 100 and 2,200 Hz.

To achieve a precision pinch with ample fingerpad contact areas for bag closure, a custom adapter was used to reduce the BarrettHand-BioTac system to a sensorized parallel gripper (*Figure 4.1*). The variable adapter enabled the selection of human-inspired, task-appropriate fingertip contact angles, such as 20° per digit in this experiment.

A custom iOS application (*iOS Apple Inc.*, 2014) was built for video and image capture during task execution. The application recorded incremental video files from each action to be used for offline analysis and labeling. A local http server on an iPhone 4s

acted as an access point for the other systems that send start and stop triggers along with a filename to the application.

Motion Planning

An online motion planner was necessary to select one of an infinite number of robot poses and desired end states, and then determine a feasible trajectory while maintaining the desired end effector orientation. Using the Robot Operating System ("Robot Operating System (ROS)," 2011), we communicated between computational nodes and the motion planner (Quigley et al., 2009).

We used MoveIt! for motion planning and collision avoidance (Sucan & Chitta, 2015). MoveIt! incorporates various planning algorithms; we selected the commonly used sampling-based Optimally Rapidly-exploring Random Trees (RRT*) planner from the Open Motion Planning Library ("Open Motion Planning Library (OMPL)," 2015). Once a path is found with RRT* the algorithm will continue to optimize the path until a specified planning time has expired. If the planning time expires without returning a feasible trajectory, the system will recursively add additional time to the algorithm and replan until a successful trajectory is returned.

Communication and Data Logging

MATLAB (MathWorks) was used for the online analysis of tactile (BioTac) and proprioceptive (WAM) data. A TCP/IP connection was used for communication between MATLAB and the hard real-time robotic system running in C++. The sensory data were sent to an external soft real-time node for processing. For training and data labeling, we recorded close-up videos of the artificial fingertips and zipper. The iPhone video camera received start/stop triggers and filenames from the hard real-time node. After each trial, we had a recording of all tactile sensor data, joint positions, joint velocities, joint torques and a close-up video of each individual action taken. All data were synchronized on the hard real-time thread and recorded only during the motion of the end effector. Having recordings that are the length of each motion provides benefit by reducing the need for any thresholding or preprocessing of the data recordings.

Reinforcement Learning via Q-Learning

Reinforcement learning is a well-established approach for learning a policy to select actions in order to maximize the expected rewards. During reinforcement learning, state and action spaces are explored and appropriate rewards are given after each action. The reward is then used to update the expected reward from the action taken.

Q-learning was used as a baseline to learn which action to take in order to maximize expected rewards. As a temporal difference method, Q-learning results in delayed rewards, and will converge to the optimal policy as time goes to infinity. The policy is in the form of a "Q matrix" that contains the expected rewards for each "state-

action pair" (Christopher J. C. H. Watkins & Dayan, 1992; Christopher John Cornish Hellaby Watkins, 1989). Each row of the Q-matrix corresponds to a state and the columns are the available actions. A single entry in the Q-matrix corresponds to the expected reward of taking an action given the current state which is referred to as the "state-action pair." It is necessary for the Q-learning algorithm to comprehensively explore the state space in order for rewards to back-propagate through the system. Given enough exploration, the expected reward of each state will converge to the optimal solution. However, exploration and exploitation must be balanced; excessive exploration to reduce uncertainty will waste time and increase regret, which is the difference in rewards earned by an agent acting optimally versus those earned by an agent that receives random rewards.

Designers of Q-learning algorithms must set three parameters: learning rate α, discount rate γ, and exploration rate ε. The *learning rate* determines how quickly expected rewards are updated based on the difference in the expected reward and actual reward observed. If the learning rate is too high then stochastic responses can excessively influence expected rewards. The *discount rate* can be used to specify a decrease in value of future rewards. A low discount rate causes the learner to behave in a short-sighted and greedy manner by favoring immediate rewards. The discount rate is typically below one because future rewards can be troublesome, as they have the ability to propagate large stochastic errors. The discount rate was set to 1 in this work such that future rewards were not discounted. This seemed reasonable given the small state space of this system. The *exploration rate* is used to adjust the probability of further exploring the state space to reduce uncertainty versus exploiting current knowledge and taking the action expected

to return the maximum reward based on the current state. While Q-learning does not provide any guarantees on regret via an exponentially decaying exploration rate, it is possible to tune performance and empirically improve the rate of convergence to the optimal policy. However, such manual tuning of the Q-learning parameters could be sidestepped through the use of more advanced learners, such as Contextual Multi-Armed Bandits, to be discussed later. Eq. (4.1) shows the main update to the Q-matrix that happens at every iteration when a reward is received. The Q-matrix in Eq. (4.1) shows the estimated reward *r* at time *t* of taking an action *a* given the current state *s*.

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[\, r_{t+1} + \gamma max_a Q(s_{t+1}, a) - Q(s_t, a_t)] \quad 4.1$$

An interesting benefit of Q-learning is that the robotic system can learn from incorrect actions. All tuning of the learning rate α, discount rate γ, and exploration rate ε are done manually. While Q-learning will converge to the optimal policy, it does not provide any guarantees of optimality in balancing exploration versus exploitation.

State Space and Action Space

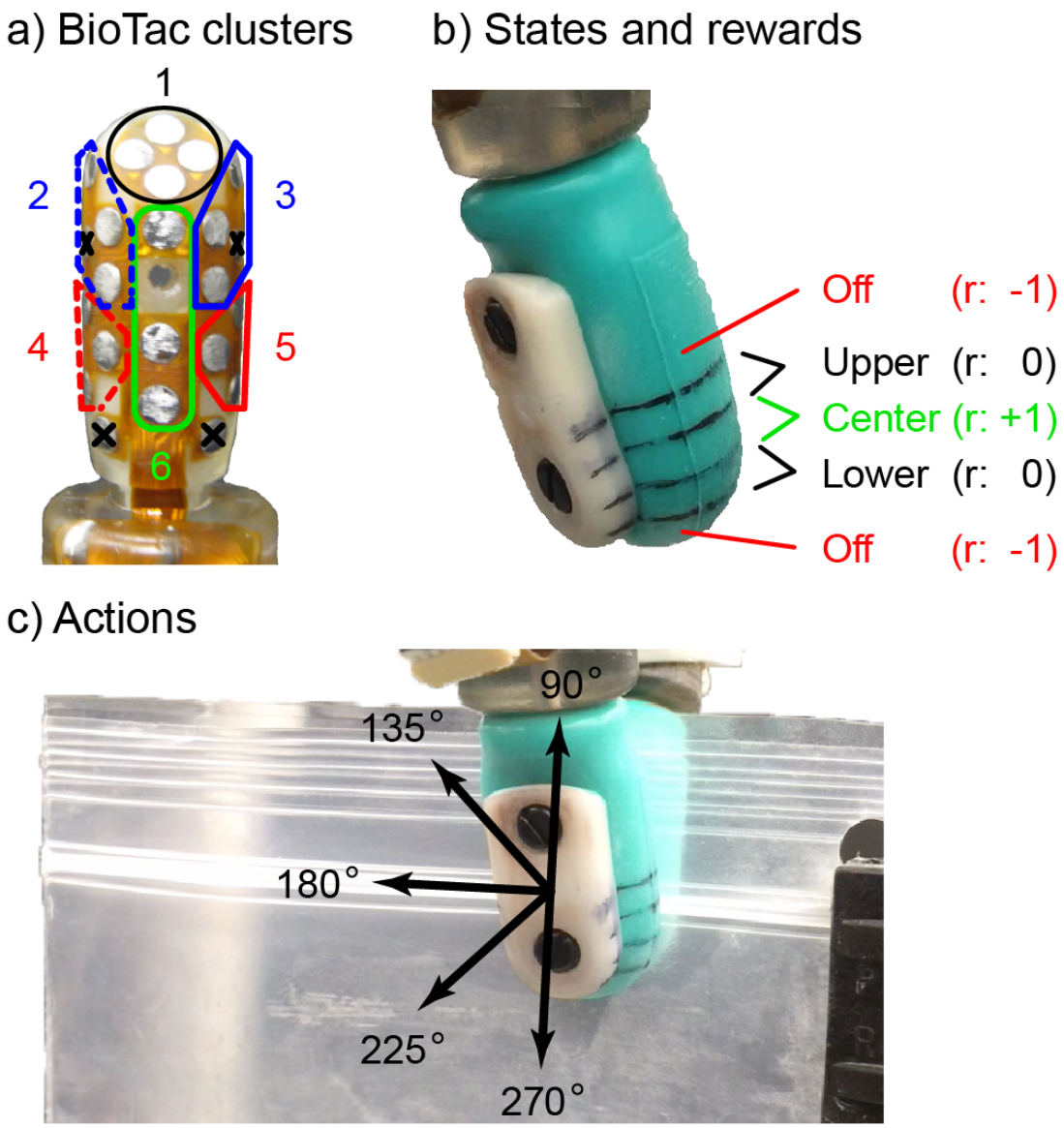

Figure 4.2 Pilot study state, actions, and rewards. Excerpt from (Hellman & Santos, 2016). (a) BioTac electrodes were clustered into six regions according to spatial location. (b) Sensor data were labeled and rewards assigned using the regions indicated on the sensor skin. (c)The action space consists of 0.5 cm fingertip movements in five directions.

Contour-following requires the fingertips to be moved relative to the ziplock bag, which requires perception of the fingertips relative to the contour (zipper) to be followed. We defined a discretized state space that would represent the longitudinal position of the zipper relative to the artificial fingerpad. The state space classes were “Off”, “Lower” (distal), “Center”, or “Upper” (proximal) (*Figure 4.2*b).

Once the fingerpad-zipper relationship had been perceived, a decision could be made about which action the robot should take next. For this proof-of-concept work, we considered a simple action space: 0.5 cm fingertip movement in a 90°, 135°, 180°, 225°, or 270° direction from the current fingertip location (*Figure 4.2*c). A displacement of 0.5 cm was implemented such that the fingertips would remain in contact with the zipper during exploration even if the worst action were selected from an ideal state. By remaining in contact with the zipper, the robot can estimate the state, determine error, take an appropriate corrective action, and learn from the action-state pair. Each sequential action maintains the initial (constant) fingertip orientation and constrains motions to the plane of the ziplock bag.

Training of State Classifier

Baseline BioTac data were collected prior to each trial under no-contact conditions. Upon initial contact with the ziplock bag, grasp pressure was increased until a pre-specified internal fluid pressure was achieved by both BioTacs. This grasp pressure was maintained throughout the experiment. Random exploration of the action space provided training data for building classifier models for the state of the system. We collected 3,000 trials for offline training of classifiers to estimate the location of the zipper along the longitude of the artificial fingertip. Each training trial consisted of a random series of 7-10 consecutive actions after a pinch grasp had been established on the bag. A custom MATLAB GUI was used to extract video snapshots for the manual labeling of states according to *Figure 4.2*b.

For analysis, the 19 BioTac electrodes were grouped into six clusters according to spatial location on the finger (*Figure 4.2*a). It has been shown previously that the reduction of the spatial resolution of the electrodes does not significantly degrade machine learning performance (R. D. Ponce Wong et al., 2014; Ruben D. Ponce Wong et al., 2014; Su, Fishel, Yamamoto, & Loeb, 2012). The reduction in spatial resolution reduces the complexity of trained models, speeds up online inferences, and ultimately, aids in the real-time execution of the classifier.

Two commonly used classification techniques were applied to the labeled training data: Hidden Markov Models (HMMs) and Support Vector Machines (SVMs). Due to the dynamic nature of the tactile sensor data, HMMs were considered for their anytime classification capabilities (i.e. while the fingertip is moving and before a complete pre-planned fingertip trajectory has been completed). SVMs were considered for their broad

success in classifying nonlinear data. The simplicity of the SVM feature vector allows for fast inference, although the action (fingertip movement) must be completed prior to classification.

Hidden Markov Models

As described in Chapter 3 a Hidden Markov Model is a Markov chain in which observations are conditionally independent of other observations given the current state. HMMs are specified by a transition matrix, belief matrix, and an initial state distribution. The transition matrix is the probability of transitioning to a state given the current state. The belief matrix is a mixture of Gaussians that contains the probability of an observation (six cluster-based observations per BioTac, as shown in *Figure 4.2*a) given the current state. Each of the HMM's belief and transition matrices were initialized with a random training observation from each class ("Off," "Lower," "Center," "Upper"). The initial state distributions were set to one for the first state and zero for all others. The models were then run through the Baum-Welch algorithm, which adjusts model parameters based on expectation-modification (Rabiner, 1989). At each incremental timepoint for a given datastream that includes all data from the start of an action, the classifier determined the most likely state based on which state-specific HMM yielded the maximum log-likelihood value.

The state of an HMM is an abstraction of the feature space of the data. To avoid overfitting of the model, it is necessary to limit the number of states. We investigated the effects of the number of states (2-10) on model performance. For additional details on HMMs see Chapter 3.

Support Vector Machines

Support Vector Machines provide a robust and well-established way to classify data. SVMs are robust to outliers and will converge to a global minimum (P. N. Tan, Steinbach, & Kumar, 2006). SVMs were trained using the mean value from each of six BioTac clusters over the entire action. MATLAB was used to train the SVM and 10-fold cross-validation was used to verify the model (P. N. Tan et al., 2006).

The SVM was tuned to improve performance and reduce overfitting of the data. We investigated the effect of changes in the box constraint ($[1x10^{-5} :10: 1x10^{6}]$) on model performance. Tuning of the box constraint adjusts how tightly the model is forced to fit the training data; too large of a box constraint results in model overfitting.

Results

Classifier Training

Each state classifier was trained independently from 3,000 trials. The offline training of each HMM and SVM classifier took approximately 20 min. and 2 min., respectively (2013 MacBook Pro, MATLAB). Model performance was assessed using the best 10-fold, cross-validated, out-of-sample classification. The HMM with 6-states and the SVM having a box constraint of 0.1 performed comparably. Cross-validated out-of-sample classification for the best HMM and SVM models were 85% and 89%, respectively. The SVM was selected for online haptic perception during the test trials described later.

Reinforcement Learning in Simulation

The success of the Q-learning algorithm depends on the classifier's ability to provide an accurate state estimate. Without an accurate model of the location of the zipper with respect to the fingerpad, the Q-matrix will be useless, as it will be unable to develop a consistent model of the expected reward for each action-state pair.

In simulation, the Q-learning algorithm successfully built a Q-matrix with the 3,000 training trials (*Figure 4.3*). The four possible states of the system are shown alongside pie charts that show the five actions that can be taken from each state. Specifically, each pie chart consists of five slices that represent the five directions [90°, 135°, 180°, 225°, 270°] that the fingertips could be moved relative to their current location. The green color indicates the recommended actions based on expected rewards learned from the training data.

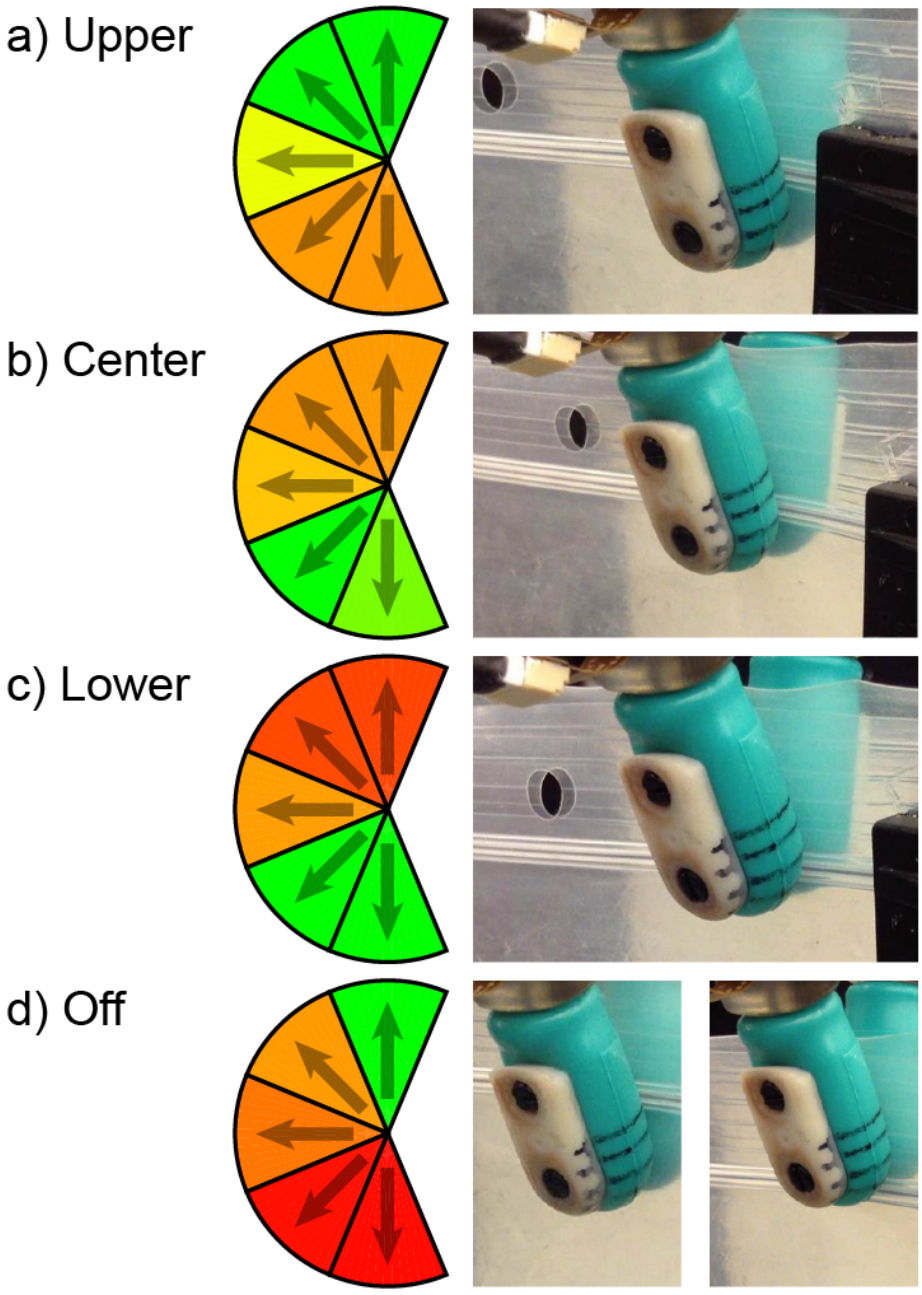

Figure 4.3 State-action policy from 3,000 exploratory motions. (a-d) Four possible states (*Figure 4.2*b) were manually labeled. Each slice in the pie charts represents an action (*Figure 4.2*c). Green and red represent the scaled maximum and minimum expected reward of each action-state pair.

Test Trials

Five test trials were run with the trained SVM classifier (with 0.1 box constraint) and Q-learner. Each test trial began by moving the WAM fingers close to the pinned corner of the ziplock bag (*Figure 4.4*a). Baseline BioTac data were collected in a non-contact state and then the gripper was closed to a pre-specified contact pressure. The classifier would make its first state estimate and then begin the iterative action-perception process by selecting an action that exploited the previously learned policy that is the Q matrix.

The system continued until criteria were met, thereby signaling completion of the task (*Figure 4.4*(a-e)). Completion was specified as either three consecutive "Off" states or by traversing 1 cm longer than the length of the ziplock bag determined by the straight line distance from the original grasp point.

Four out of five trials were concluded with three consecutive "Off" states. Due to misclassification of the "Off" state near the end of the ziplock bag, the fifth trial concluded by exceeding the pre-specified zipper length threshold. The zipper was successfully tracked in all test trials and the bag was successfully closed in three out of five trials.

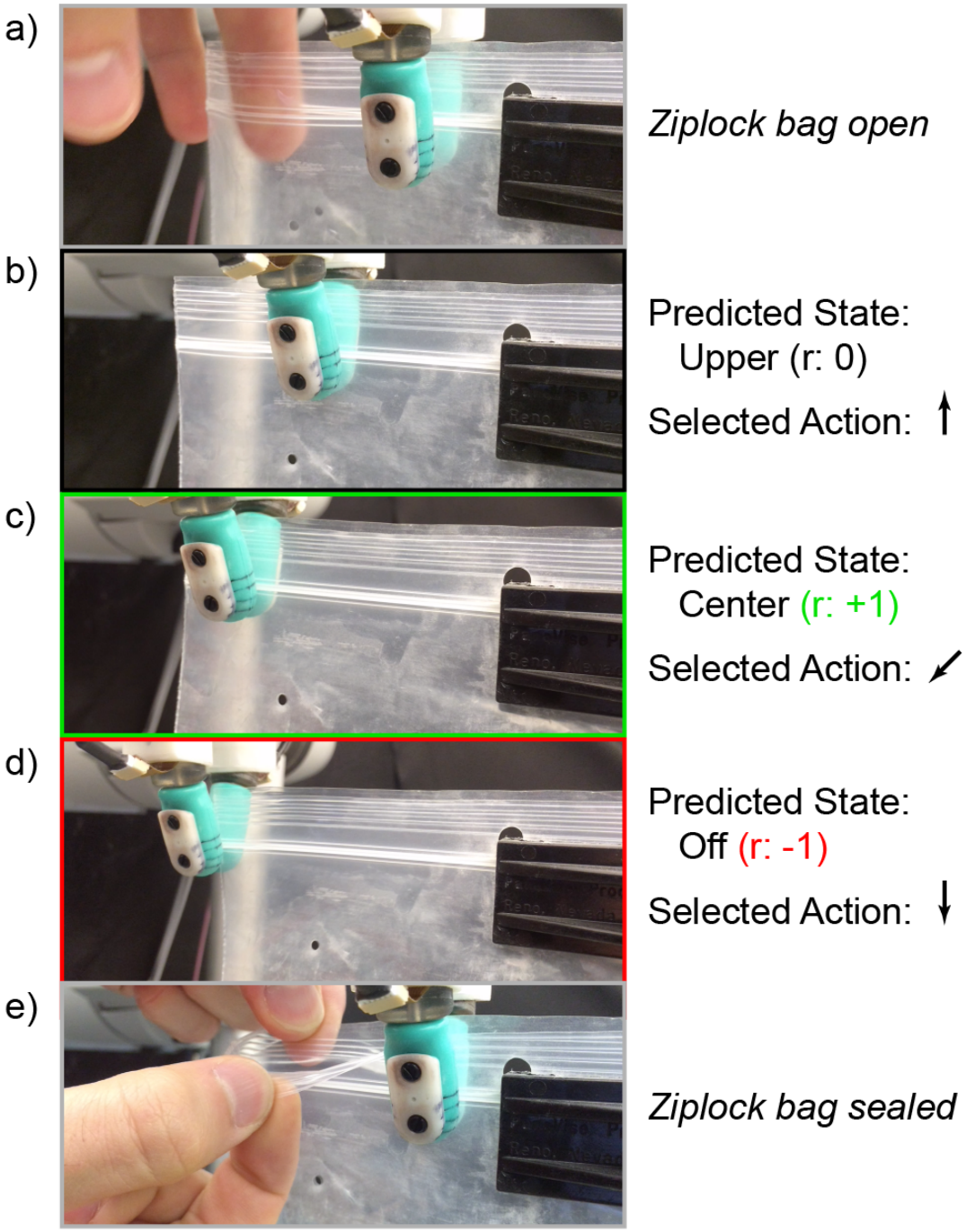

Figure 4.4 Representative trial of successful haptic perception, reinforcement learning, and decision-making. (a) Initial state with open bag. (b-c) Examples of SVM state estimates, rewards, and actions selected using Q-learning. (d) Completed trial with fingers reset and the sealed ziplock bag.

Discussion

Effects of Classifier Model Type

HMMs are inherently able to provide model likelihoods during action execution and have been shown to model complex temporal data well (Rabiner, 1989). This is useful for classifying states prior to action completion, which could enable more frequent inferences, earlier decision-making, and earlier corrective actions. SVMs do not inherently support anytime inference, but have the advantage that they can be used to classify high-dimensional nonlinear spaces.

Both the SVM and HMM had similar cross-validated, out-of-sample classification errors. The HMM approach enabled classification while the fingertips were moving. However, the SVM resulted in a more reliable classification rate. This slightly superior performance of the SVM may reflect that the nonlinear polynomial kernel is more generalizable than that of the HMM mixed Gaussian model, especially given the nonlinearity of the tactile sensor data.

Assumption of Observation Independence May Not Apply to Data from Tactile Sensors

Task success critically depends upon the accuracy of the state classifier. Classifiers were trained on random sequences of actions but applied to tactile sensor data from non-random sequences of actions. In this sense, the classifier was disadvantaged when it was trained, although random exploration of the state space was important.

Our prior work showed that the deformation of the tactile sensor skin is greatly affected by the conditions of the initial contact, especially contact area (R. D. Ponce Wong et al., 2014). In this work, the "initial contact" conditions for each state-action iteration were dependent upon the previous sequence of actions for a given trial. Due to the bladder design of the tactile sensor, there is valuable information away from the region of contact. This information is encoded in skin stretch, which depends upon prior contact conditions and actions that may be irrelevant. Since a recent strain history is encoded in the skin stretch, it may be valuable to implement reinforcement learning with classifiers that rely only on the current state for context, as is presented in the following sections.

In addition, the SVM classifier model was built under the assumption of action independence and, thus, observation independence. An approach to improve the current classifier could be to group the training data according to the previous action in order to account for variation in "initial contact" conditions. One could also build separate action-specific models, but the number of models would increase with the size of the action space. The possible sequences of actions that would lead to successful bag closure are infinite and cannot possibly be haptically experienced a priori. Furthermore, additional training data would be required and tactile sensor data are costly to obtain, with respect to

time and wear, as compared to less physically interactive sensing modalities such as computer vision.

Stochastically Robust Techniques for Haptic Perception and Manipulation

Reinforcement learning is a promising approach to develop methods to learn new haptic perception and manipulation skills. The stochastic nature of physical systems can cause issues when developing new skills. Fortunately, Q-learning can account for the stochastic interactions of soft, deformable tactile sensors with the environment. The Q-learning algorithm has been extremely successful in the advancement of reinforcement learning techniques primarily due to its relative simplicity and ability to support stochastic feature sets (Christopher J. C. H. Watkins & Dayan, 1992). Q-learning serves as a benchmark for the future development of haptic decision-making and learning algorithms.

Summary of Pilot Study

While we were successful in closing a ziplock bag, some actions were inefficient. During task execution, the policy cycled between “center” and “upper” states. This cycling was mostly due to the coarse nature of the action space and the absence of a penalty for unnecessary moves. To improve time to completion, more efficient and aggressive actions could be used along with a discounted reward for excessive actions. New actions could include adjustments to grip contact pressure and variable fingertip travel based on state confidence.

During the pilot study, many of the hardware systems were fine-tuned. Yet, some aspects of the system required that a human be present to manually label data and oversee other aspects of the system. Extending the incremental work of this pilot study, the following sections will describe how the system was fully automated with the ability to label and learn without human supervision.

Learning a Functional Task with Optimal Cumulative Rewards

From the pilot study we successfully learned a policy to track a zipper and close a ziplock bag. While Q-learning was successful, a significant downside is that it does not provide guarantees on optimality. From lessons learned during the pilot study, we made changes to the state and action spaces and implemented new reinforcement learning algorithms that minimize regret by optimizing the cumulative rewards received.

The same hardware shown for the pilot study was employed for the following work. State labeling and classification were automated to remove the need for human supervision and minimize the introduction of human error. Additionally, changes to the processing of the tactile data greatly improved the robustness of the state classification.

Methods

Reinforcement Learning via Contextual Multi-Armed Bandits

Multi-Armed Bandit (MAB) reinforcement learning algorithms are often explained using the analogy of a bank of slot machines that return unknown random rewards. MABs are used to determine the best sequence and choice of slot machine arm "pulls" (or action choices) in order to maximize rewards. The slot machines all return a random reward from a unique unknown probability distribution. The terms "arm" and "action" will be used interchangeably for the remainder of this Chapter. The question of what is the appropriate choice of arms requires probability theory to determine which arm and sequence of arms will result in an optimal policy that maximizes cumulative rewards. Early work such as the Gittins Index derived the optimal policy for maximizing the expected rewards of MABs (Gittins, Glazebrook, & Weber, 2011). MABs can be designed to be optimal and maximize the total rewards received for both a known and unknown number of finite actions (number of arm pulls, or data collection trials). Each action is defined by a single pull of an arm which then receives a reward based on the outcome (i.e. new state) resulting from the action. Simply put, MABs balance the exploration versus exploitation of the arms to maximize the total reward. MABs track the number of pulls per arm (i.e. how frequently each action has been explored) and estimate the expected reward of each arm. Given a sufficient amount of exploration of all arms the MAB will then begin choosing the arm with the highest expected reward.

Contextual Multi-Armed Bandits (C-MABs) are a particularly useful variant of MABs. C-MABs receive a feature vector (context) that is associated with each iteration or action. The feature vector can contain a wide range of information that the algorithm can then use to make better choices during exploration and exploitation. The contextual bandit problem is well known and has been an area of comprehensive study (Bubeck, Munos, Stoltz, & Szepesvari, 2011; Dudik et al., 2011; Kleinberg, Slivkins, & Upfal, 2008; Langford & Zhang, 2008; Slivkins, 2014). The algorithms listed all assume a known similarity metric between neighboring contexts and achieve sublinear-in-time regret with respect to the complete knowledge benchmark. Sublinear regret implies that the policy will converge to the optimal reward.

The variant of contextual bandits used in this chapter is Contextual Learning with Uniform Partition (CLUP) applied to a single agent setting (Cem Tekin & van der Schaar, 2015). CLUP is optimized to maximize the cumulative rewards given the number of current trials. Other algorithms may optimize on a finite number of trials which then becomes a problem of optimizing the total exploration. Given haptic interactions and an indeterminate number of trials we believe that maximizing the cumulative rewards is the best course of action.

We now define the parameters of CLUP before discussing the algorithm. Let $f \in \mathcal{F}$ be the set of arms available to the learner. Given uniform partitioning, let $p(t)$ be the set in $\mathcal{P}$ to which the $x(t)$ context belongs. Let $N_{f,p}(t)$ be the number of "context arrivals" (i.e. how many times a context has been visited) for each arm given the uniform partitioning of the context.

From the set of arms $\mathcal{F}$ and the fact that the learner operates in discrete time space $t = 1, 2, \ldots, T$ the following events happen sequentially for each time step: (i) receive a context $x(t)$; (ii) from $x(t)$, the learner will choose an arm $f \in \mathcal{F}$. The choice of arm f depends on D, which is of the form $D(t) = t^z \log t$. The learner will determine the set of under-explored arms (Eq. 4.2).

$$\mathcal{F}^{ue} := \left\{ f \in \mathcal{F} : N_{f,p}(t) \leq D(t) \right\} \tag{4.2}$$

The parameter D is a deterministic, increasing function of t called the control function. For a rigorous proof of CLUP along with the control function D see (Cem Tekin & van der Schaar, 2015). If the set of $\mathcal{F}^{ue}$ is non-empty, the learner enters the exploration state and will select a random arm from $\mathcal{F}^{ue}$ to explore. Once the set $\mathcal{F}^{ue}$ is empty, the learner will enter the exploitation phase where it selects the choice with the highest estimated reward (Eq. 4.3).

$$a(t) \in \underset{f}{\operatorname{argmax}}\, \bar{r}_{f,p}(t) \tag{4.3}$$

The mean estimated reward is updated for each arm as context arrivals occur and a reward r is received (Eq. 4.4). The pseudocode for CLUP for a single agent learner is shown in *Figure 4.5*. All context arrivals occur before an arm is selected, as shown on line 6 of *Figure 4.5*. The parameters z and m_T define the similarity and size of the metric space, respectively, that are used to calculate the control function at a given time.

$$\bar{r}_{f,p} = \frac{\bar{r}_{f,p}\, N_{f,p} + r}{N_{f,p} + 1} \tag{4.4}$$

CLUP adapted for single agent learner:

Input: $D(t)$, T, m_T
Initialize sets: Create partition $\mathcal{P}$ of context
Initialize counter: $N_{f,p} = 0,\ \forall f \in \mathcal{F}, \rho \in \mathcal{P}$
Initialize estimates: $\bar{r}_{f,p} = 0,\ \forall f \in \mathcal{F}, \rho \in$
While $t \geq 1$ **do**
  Find the set in $\mathcal{P}$ that $x(t)$ belongs to
  Let $p = p(t)$
  Compute set of under explored arms $\mathcal{F}^{ue}(t)$ from Eq. 4.2
  if $\mathcal{F}^{ue}(t) \neq 0$ **then**
    Select a randomly from $\mathcal{F}^{ue}(t)$
  else
    Select a randomly from $\underset{f}{\operatorname{argmax}}\, \bar{r}_{f,p}(t)$
  end
  $\bar{r}_{f,p} = \frac{\bar{r}_{f,p} N_{f,p} + r}{N_{f,p} + 1}$
  $N_{f,p} + +$
  $t = t + 1$
end while

Figure 4.5 Pseudocode for the Adapted CLUP algorithm for a single agent learner (Cem Tekin & van der Schaar, 2015).

State Space and Action Space

There are various ways in which to represent the states of the fingerpad such as the mean longitudinal position of the zipper relative to the fingerpad, or possibly the direction of motion of the zipper relative to the fingerpad during each action. The state space used in this experiment was defined according to empirical results from the pilot study. The system state was discretized into fingerpad states of "High," "Center," and "Low" that represent the location of the zipper along the length of the fingerpad (*Figure 4.8*).

During the pilot study with the four state system (*Figure 4.2*) a sawtooth trajectory of the fingertips was observed during task execution. We believe that the learned policy was oscillating between the "Upper" and "Lower" states in an attempt to maximize the occurrence of the "Center" state while also avoiding the "Off" state. The reward of the "Off" state may have been overweighted in the model, which would cause the policy to act overly cautious. In an attempt to avoid the "Off" state the system would learn to oscillate around "Center." Since the policy will optimize a system according to its defined state and action spaces, both spaces were refined for the follow-on study.

In defining the appropriate context vector for the C-MAB, two techniques were explored: (i) using raw sensor data as the context to the bandits, and (ii) using machine learning to pre-process the sensor data and classify the context. In order to input raw sensor data as the context in technique (i), the tactile data were partitioned. Using the six cluster-based observations per BioTac from the pilot study shown in *Figure 4.2*a, each cluster was assigned a value indicating a compressed, neutral, or bulged state of the fingerpad skin relative to the sensor core. Partitioning of the data was done through k-means clustering (Hartigan, 1979). The mean value of each cluster from the full 3,000 training trials of the pilot study are shown in *Figure 4.6*. The selected thresholds are shown as red vertical bars that indicate the threshold for each individual cluster (rows) and contact pressure (columns: low, medium and high). Global thresholds, shown as the green regions in *Figure 4.6* for each cluster, were determined through k-means clustering of the full set of training data that includes all contact pressures.

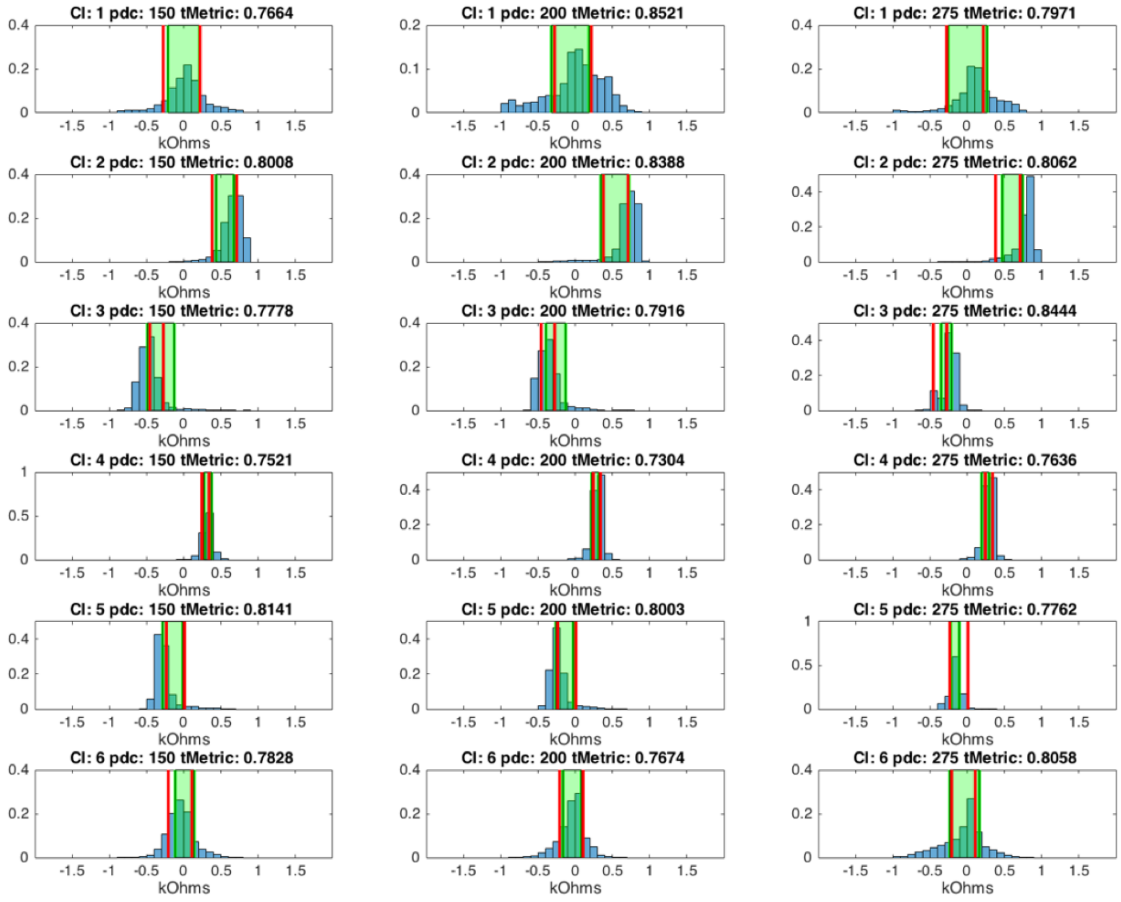

Figure 4.6 Tactile data from the pilot study were partitioned for the labeling of three regions: compressed (left of green region), neutral (green region), or bulged (right of green region). Rows indicate clusters 1-6 and columns indicate the three contact pressures (low, medium, and high). Thresholds shown on the histograms evenly distribute the context labeling by minimizing in-sample variance.

The total number of possible contexts was 3^6 (729), where there are 3 possible states for each of the six clusters. The context with the most arrivals was consistent with what would be expected during a simple perpendicular sliding motion along the fingertip: the radial side of the finger was compressed, the ulnar side was bulged, and the center of the finger and the fingertip were in a neutral state (*Figure 4.7* inset). While the training data were fairly distributed along the possible contexts there were still areas with disproportionate context arrivals (*Figure 4.7*). Having 729 possible contexts requires significantly more training data then the 3,000 available trials, and there is no guarantee that all contexts are physically meaningful or possible. Due to the large volume of training data necessary for this technique (i), a machine learning algorithm was instead used to interpret sensor data and reduce the dimensionality of the context.

To reduce the size of the context space and keep the required exploration tractable, a classifier was implemented (technique (ii) mentioned previously). The classifier takes normalized sensor data from each action and determines the state of the fingerpad-object interaction. The details of the state classifier are expanded in the following section, "State Classification via Deep Neural Nets."

For the remainder of the work presented we used a simple reward structure founded upon previous experience with models developed in the pilot study, artificial fingertips, and the need to remain in contact with the zipper for bag closure (Hellman & Santos, 2016). The state space in *Figure 4.8* consists of three states "High", "Center", and "Low". The actual position and labeling of the states will be discussed in the "Data labeling" section of this chapter. Having only three states helps to keep the necessary exploration of the state space tractable. Rewards were assigned based on the classified

state: “Center” received a +1 reward while all other states received a reward of 0. The “Off” state was removed because it is assumed that if the current state is predicted as “Off” then the robotic system has already failed the task and the appropriate action would be to restart the task. Given that the appropriate course of action from the “Off” state is to restart, any exploration from there is unproductive.

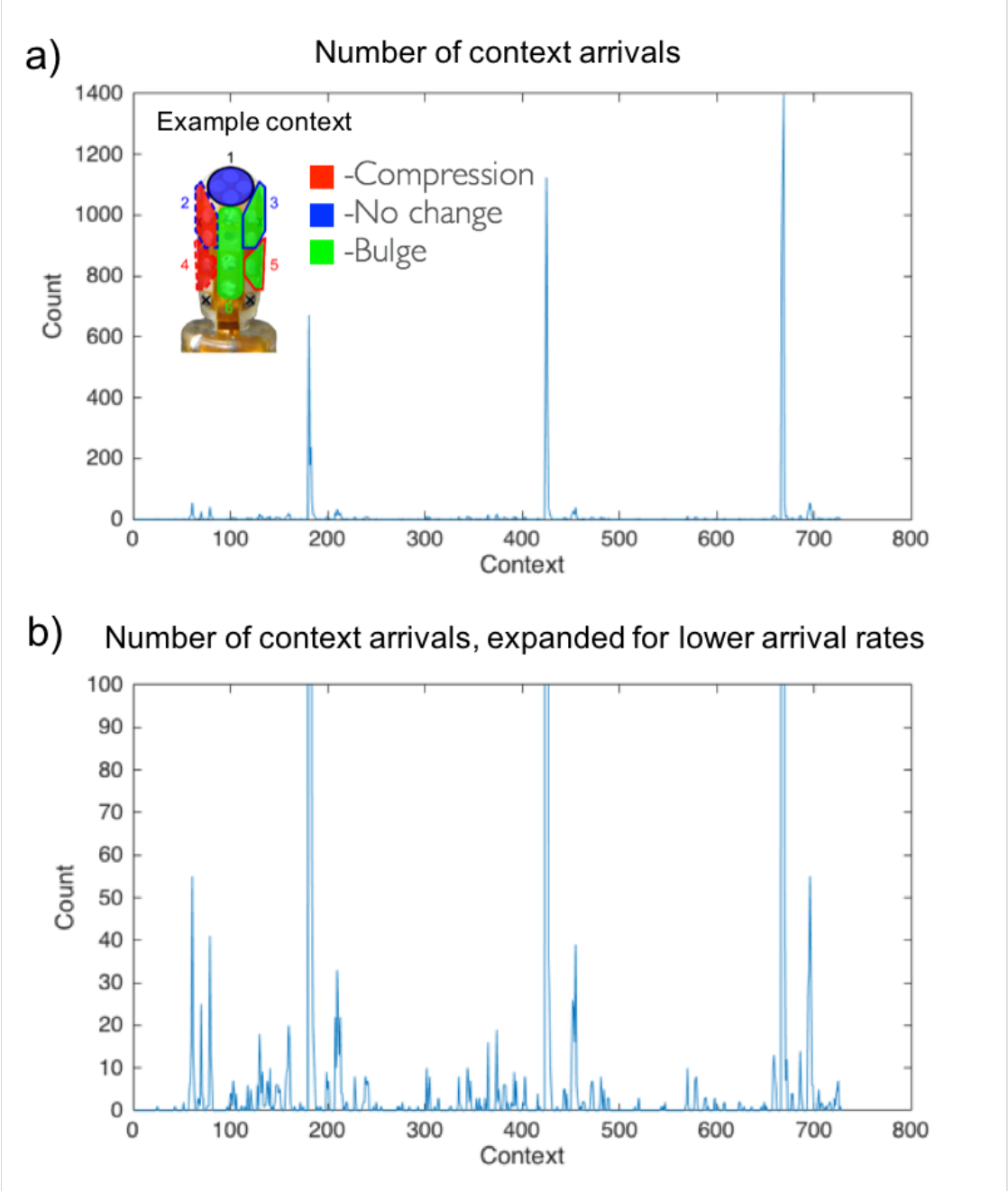

Figure 4.7 a) Number of context arrivals shown for all 729 variations of the 6 clusters in compression, no change (neutral), or bulging. b) Shows expanded view of context that have a lower arrival rate. Inset image is the context with the highest number of occurrences (1,398) during 3,000 trials.

When the fingerpad-zipper relationship is perceived, a decision is made about which action the robot should take next. It is necessary to define an appropriate set of actions that the robotic system can take to either maintain or correct the state of the system.

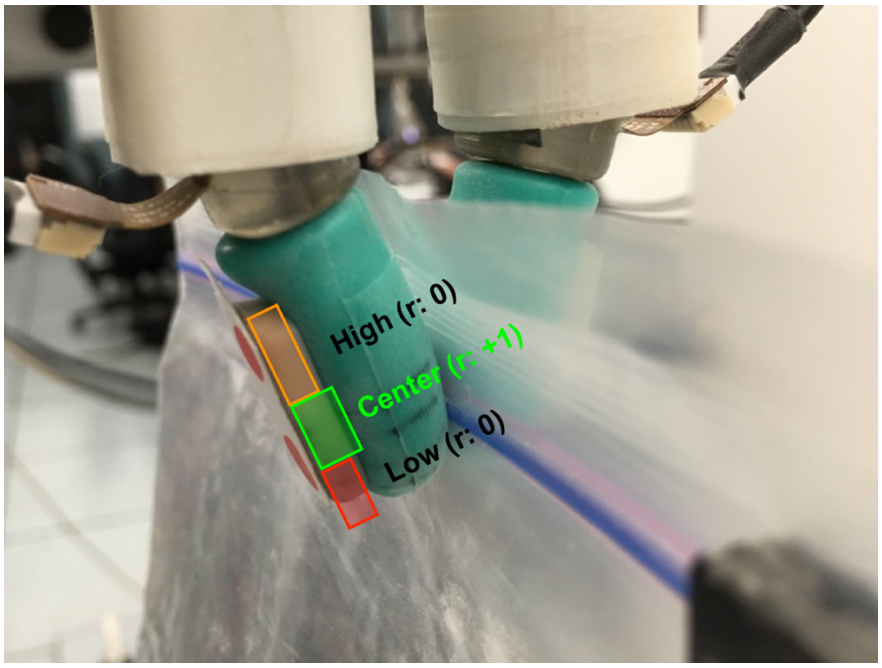

Figure 4.8 Fingerpad state representation for Q-Learner and Contextual Multi-Armed Bandits. The reward *r* is shown for each state.

For the trials presented in this follow-on study, action displacements were increased to 0.75 cm per trajectory from the previous 0.5 cm of the pilot study (Hellman & Santos, 2016). The increase was due to the elasticity of the silicone fingerpad and the desire for the fingerpads to slide along the zipper; each motion requires enough skin stretch before sliding begins and the friction in the contact region transitions from static to kinetic. The reduced coefficient of friction during sliding helps to reduce the stored

energy in the elastic stretch of the skin between pauses in actions, and during motion planning of successive fingertip trajectories. All motions had a trapezoidal velocity profile with a speed of 0.5 cm/s. The action space consisted of trajectories in a 135°, 155°, 180°, 225°, or 270° direction from the current fingertip location (*Figure 4.9*). The action space has asymmetric trajectories for proximal versus distal corrective actions. The asymmetric action space is an empirical result from many trials of learning during the pilot study. The geometric shape of the round fingertip results in large normal forces for distally directed corrective actions that tend to push the zipper out of the grasp. Thus, the action space was defined with more aggressive angles of attack in the distal directions.

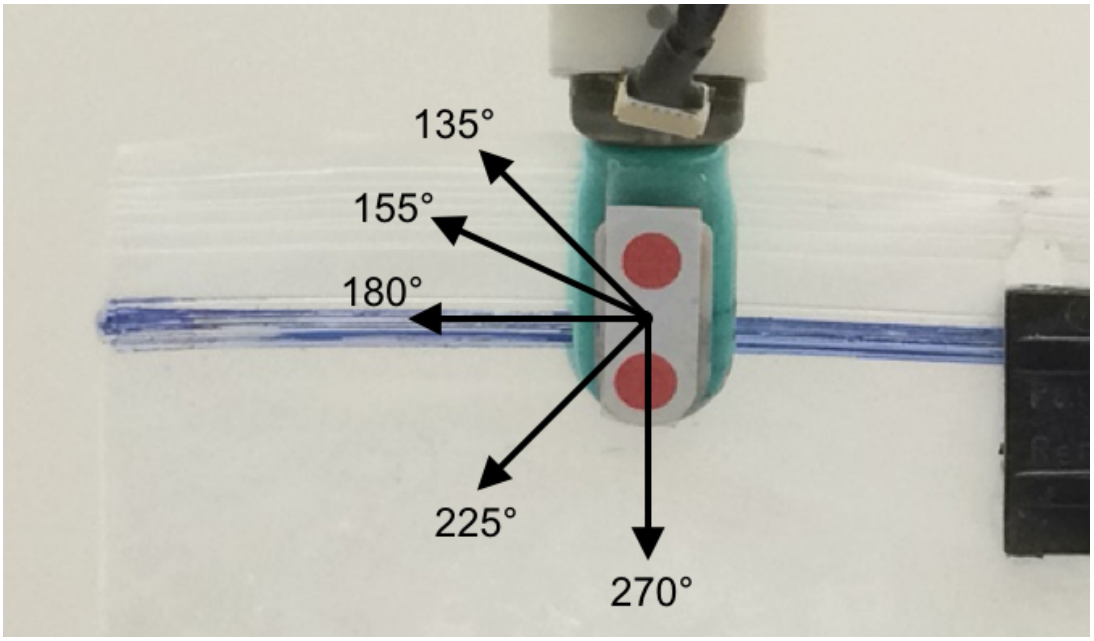

Figure 4.9 The action space consisted of 0.75 cm fingertip movements in five directions. The asymmetric action space was due to the large normal forces generated by corrective actions in the distal direction that tended to push the zipper out of the grasp.

Robust Conditioning of Tactile Data

Tactile data processing was kept to a minimum in order to maintain the speed and robustness of classification during online trials. Minimizing this processing was a challenge because many tactile sensors, including the BioTac (Yamamoto, Wettels, Fishel, Lin, & Loeb, 2012) used in this experiment, are susceptible to sensor drift . For the BioTac, sensor drift most likely stems from the variability in the silicone skin which is inflated by a weakly conductive fluid away from the sensor's rigid core. All electronic sensing elements of the BioTac are located on the rigid core of the finger. Variations in the internal fluid volume, external temperature, and any damage or overuse to the skin of the BioTac may be the cause of this instability. Sensor drift is also apparent from day-to-day classification of trained algorithms (Hui, Block, Taylor, & Kuchenbecker, 2016).

To best regularize classifier inputs, we focused on relative movements of the finger. In the pilot study, the baseline reference was set for each trial prior to grasp initialization. Each trial is defined by the set of actions from initial grasp to either successful zipper closure or failure (*Figure 4.4*). Referencing the baseline for multiple actions can introduce sensor variations and stochastic noise from the robot system.

In order to minimize all possible sources of noise and ensure the robustness of the classifier, we refined the baseline to be action-dependent during each trial. Prior to each action the BioTac records 100 ms (10 batches of data) that are used to baseline each individual action. The processing was done to ensure the robustness and accuracy of the classifier. Each action coincided with approximately 1.5 seconds of recorded data.

Automated Labeling of States

In order to train classifiers to predict the state of the zipper within the pinch grasp, it was necessary to label the data for supervised learning. During the pilot study, labeling of states was done manually based on lines marked on the skin of the BioTac. This approach ignored the fact that the skin is elastic and, depending on the sequence of actions and strain stored in the skin, the actual location of the zipper relative to the core of the finger could be very different. It was necessary to develop a new method that would be robust to skin strain and representative of the zipper position relative to the core of the tactile sensor. It is important to maintain consistency relative to the core because this is also the reference location of the sensing electrodes of the BioTac (*Figure 4.2*a). Automated labeling would also reduce the introduction of uncertainty from human error during labeling.

In the pilot study, images were captured with a view of the radial aspect of the fingerpad in order to view the center of the pinch grasp similar to that of *Figure 4.8.* By repositioning the plane of the camera lens to be parallel to the plane of the ziplock bag, as in *Figure 4.10*, both the nail of the BioTac and the full contour of the zipper could be observed. The nail of the BioTac is rigidly attached to the core with two screws that serve as convenient reference points. In order to quickly identify the location of the nail on the BioTac, red circles were placed over each fingernail screw. The distance between the centers of the two red circles was 1 cm (*Figure 4.10*). The zipper was colored blue for ease of automated visual identification. By marking the inside of the transparent zipper, the coloring withstood thousands of trials.

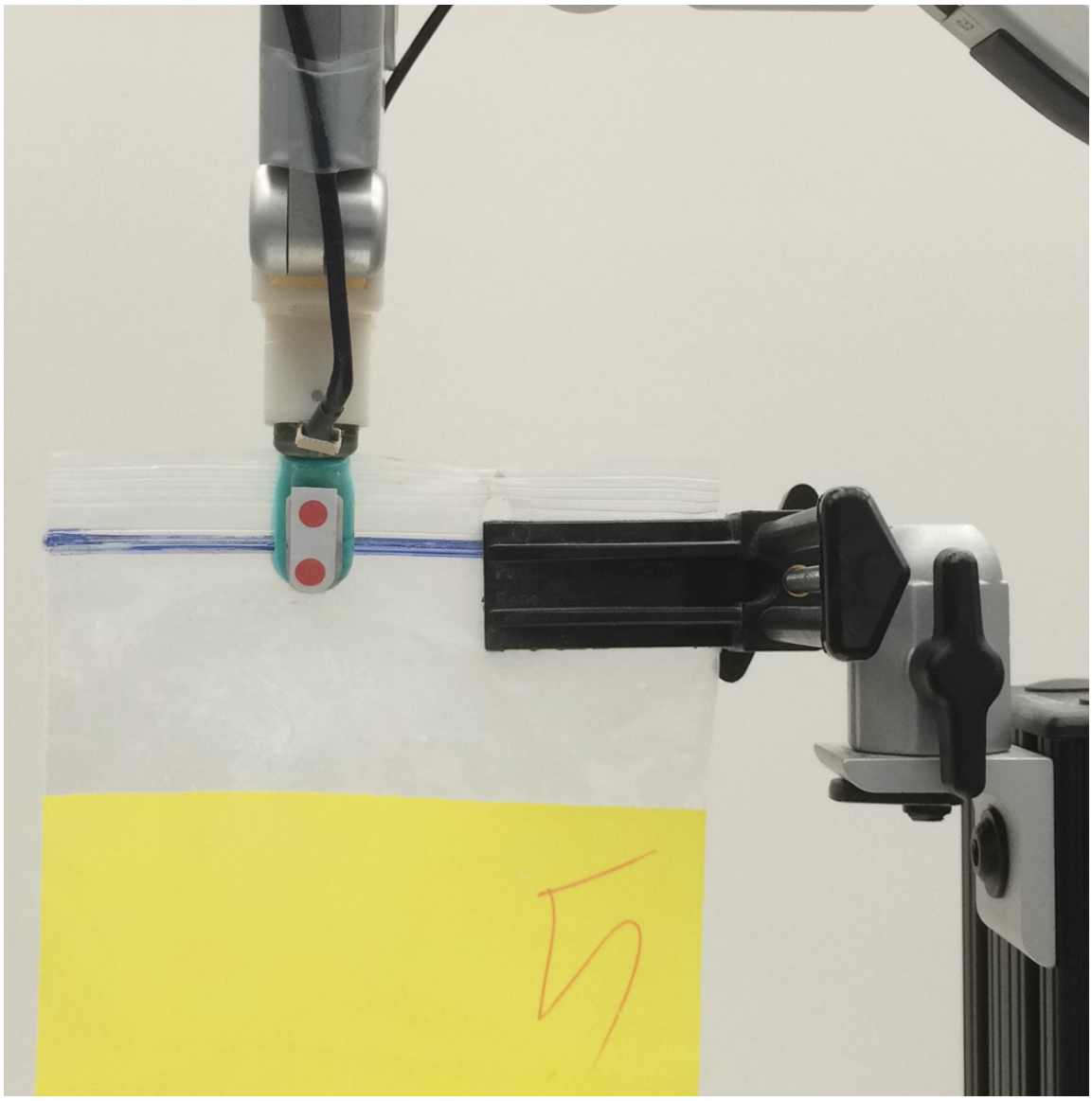

Figure 4.10 Image captured after completion of action. Red circles were placed over the fingernail attachment screws to determine the location of the BioTac. The zipper was colored blue in order to determine zipper offset relative to the center of the two red circles.

The ability of a trained model to successfully classify data is significantly influenced by the accuracy of the label. When manually labeling thousands of data trials, as in the pilot study, it is possible for human error to introduce errors in labeling. In order to reduce errors and speed up the labeling process, an automated labeling system was developed using OpenCV ("Open Computer Vision (OpenCV) Library (SourceForge.net)," 2015).

Each image was processed through a custom Python node that received an image from a custom iOS application (*Figure 4.1*). The output of the OpenCV processing was a measurement of the offset (shortest distance) between the center of the two red fingernail circles and the contour of the blue zipper along the center line of the finger shown in *Figure 4.11*. The image processing steps are listed below. Once processing was complete, the Python nodes saved the edited image file and appended the offset error in mm to the image name.

Image Processing

1. Load image through Python http server.
2. Prepare image for processing
    a. Apply Gaussian blur to reduce any high frequency noise.
    b. Convert image to HSV. HSV is a more intuitive representation of color that allows for easier thresholding compared to that of BGR (OpenCV defaults to BGR not RGB for historical reasons).
3. Locate red BioTac nail markers.
    a. Threshold image for HSV red.
    b. Dilate mask to smooth any misses in the thresholding.
    c. Locate the center of the two red clusters with k-mean clustering. The center of the two red circles are indicated by green dots in *Figure 4.11*.
4. Locate blue Zipper.
    a. Threshold image for blue.
    b. The color mask will occasionally be inaccurate due to some noise and the blue palm of the BarrettHand. Since we know where the BioTacs are we built a rectangle around that region to shrink the area where the bag should be and update the threshold. By narrowing the blue threshold around the BioTacs we remove any possible interference with the BarrettHand.
    c. Fit two rectangular contours that match the blue bag on both sides of the finger and combine into one contour.
    d. Fit line to contour (yellow line *Figure 4.11*).
5. Calculate the offset value as the distance between the center of the two red circles (green dot) and the line (blue dot on the yellow line *Figure 4.11*) at the current location of the BioTac. Convert from pixel offset to mm by using known separation distance of 1 cm between the two red dots on the BioTac (- for below center and + for above center).

Classifying Region of Fingerpad

Using the offset value from the processed images each state could be labeled according to predefined thresholds. The state for an offset value greater than 0 mm (indicated at the center of the two red circles) was labeled “High” (*Figure 4.11*a). The state for an offset value less than or equal to 0 mm and greater than -2.5 mm was labeled “Center,” (*Figure 4.11*b) while all other states less than or equal to -2.5 mm (*Figure 4.11*c) were labeled “Low”. Once the zipper was considered to be out of contact with the BioTacs based on the predefined offset thresholds, the data were discarded and excluded from model training.

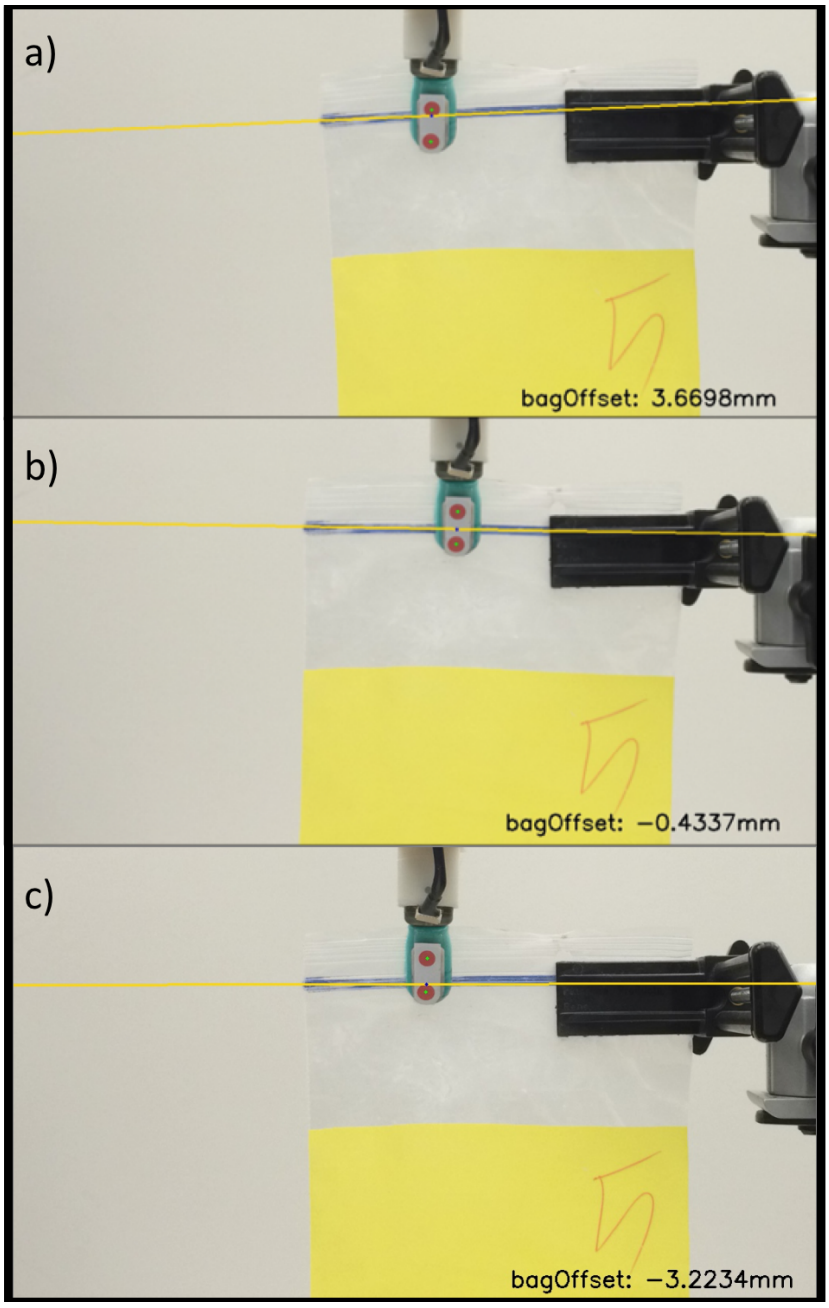

Figure 4.11 Automated labeling of grasp offset. a) State for an offset value greater than 0 mm was labeled “High.” b) The state for an offset value less than or equal to 0 mm and greater than -2.5 mm was labeled “Center.” c) All other states less than or equal to -2.5 mm were labeled “Low”. Green dots in the center of the red circles are found through k-means clustering of the red image threshold. Blue color threshold determines the contour;

a straight-line fit to the contour is indicated by yellow line. The intersection of the blue contour and line connecting the green dots is indicated by a blue dot on the yellow line. The offset value is measured as the pixel/mm difference between the center of the finger (middle of two green dots) and the location of the zipper in the grasp (blue dot).

State Classification via Deep Neural Nets

A Deep Neural Net (DNN) was selected for state classification of the tactile data. DNNs consist of multiple neural net layers in which each layer has a rectified linear unit (ReLU) applied to the outputs. By using ReLUs with DNNs we are able to fit nonlinear data, thereby allowing for robust classification of the nonlinear robot testbed system. DNNs require a feature vector as inputs and return a classification from a predefined set of outputs. The feature inputs selected for the trained DNN were the 2-norm of the mean value of the final 100 ms of recorded tactile data from each action. The sensor data were initially baselined for the action-dependent movements, as described in the previous section on “Robust Conditioning of Tactile Data”.

Classification and training were achieved with TensorFlow (Abadi et al., 2016). Over 8,000 labeled data trials were used in training. Training of the DNN was accomplished through stochastic gradient descent (SGD) with an exponentially decaying learning rate, regularization, and dropout. SGD is a method of adjusting the weights and biases of each layer of the DNN to minimize a specified loss function. Instead of using the full set of training data for every update to the model, SGD randomly selects a small set of data. By using SGD, training time is significantly reduced. Even though each training step does not use the full set of training data, by stochastically varying the

training set the full set is represented. In this work we use L_2 Regularization (Ng, 2004) which adds additional terms to the loss in order to penalize large weights. Penalizing large weights helps to prevent overfitting of the model to training data. Dropout is a technique of stochastically turning weights on and off during training (Srivastava, Hinton, Krizhevsky, Sutskever, & Salakhutdinov, 2014). Dropout is an extremely useful development for neural nets that prevents overfitting and distributes activation weights across the nodes of the DNN. By randomly turning a set percentage of the weights on and off during training, the model no longer relies on a single activation for classification. Randomly removing weights ensures that the activation of nodes are distributed though out the network. The DNN used three hidden layers with 512 nodes per hidden layer.

Results and Discussion

State Classification

The DNN classifier was trained with 7,200 (90%) data samples; 800 (10%) data samples were used to validate the model. Approximately 15,000 iterations were taken to train the model with an average training time of less than 10 minutes (2013 MacBook Pro, MATLAB). The DNN performed with an 89% classification accuracy on the training data and an 86% classification accuracy on the validation set.

Robustness of the model to daily environmental variations is critical to the success of the learner. The model's ability to correctly classify the state of the zipper within the pinch grasp will minimize stochasticity of rewards during reinforcement learning. Correct classification will allow both Q-learning and CLUP to more efficiently converge to the optimal policy for each state-action pair.

Prior to action classification, each trial is initialized by grasping the ziplock bag and taking a 180° action. This initializing action is necessary to set the direction of strain in the skin of the BioTac. Initial variations of the baselined electrode values fluctuate during the first action as seen in *Figure 4.12*a show why the initialization action is necessary prior to classification. The changes in signs of the electrode values are due to the initial state of the skin during grasp initialization.

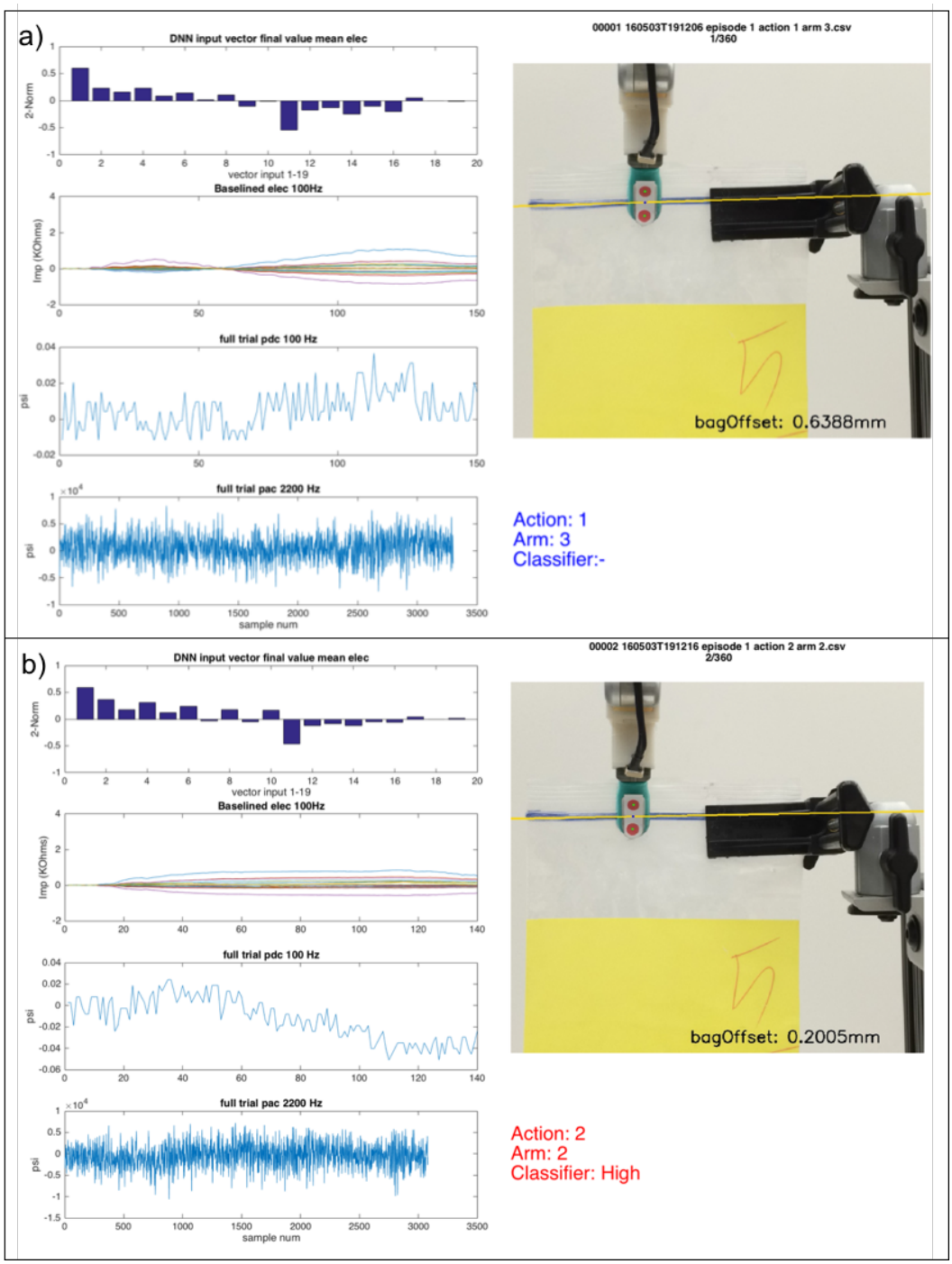

Figure 4.12 Full action dataset and classification for sequential actions. a) shows the initial start of a trial where a 180° action always initializes the system and does not receive a classification. The second action taken (b) is correctly classified as "high" with an offset value of 0.2 mm. The plot for each action shows the sensor's response during

movement. Plot descriptions from the top down (all tactile data initialized with action-dependent baseline): histogram shows mean value of the last 100 ms for all 19 electrodes, complete electrode response from the full action (100 Hz), low frequency fluid pressure of the BioTac (100 Hz), and high frequency pressure response of BioTac (2200 Hz). DNN input vector values are shown as a histogram of the 19 normalized electrode values. The x-axis for the baselined impedance electrode (100Hz), low frequency pressure (100Hz), and high frequency pressure (2200Hz) data is the sample number within the full trial.

Reinforcement Learning and Expected Reward

Reinforcement learning is used to determine the expected reward for each of the state-action pairs. When learning, rewards are stochastically received and the learner must appropriately update the expected reward based on the difference between an observed reward and the expected reward. Q-learning will converge to the optimal expected reward, but only as time goes to infinity. Q-learning also does not provide any finite guarantees on optimality. C-MABs provide an optimal search of the state-action space to maximize the cumulative rewards given the current number of actions taken.

Both Q-learning and C-MAB were run for just over 750 actions to compare each algorithm's expected reward. *Figure 4.13* shows the expected rewards after 757 actions taken with Q-learning and 758 actions taken with C-MABs. All reinforcement learning was done online and independently of one another. For a true comparison between both algorithms, it is necessary to run them online because each algorithm determines actions sequentially.

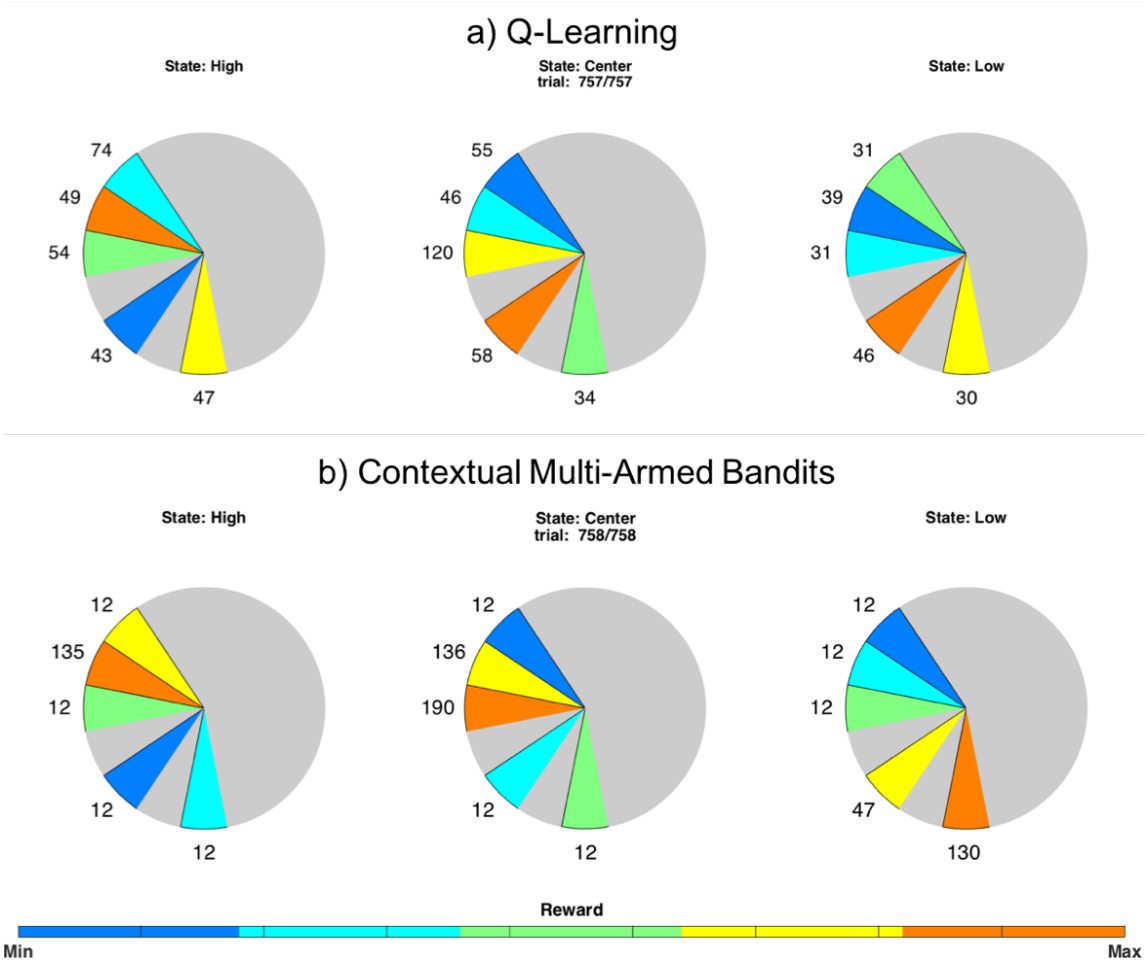

Figure 4.13 Expected rewards from each state-action pair. Each pie chart represents the state "high", "center," and "low" (*Figure 4.8*). Slices of each pie chart correspond to available actions (*Figure 4.9*). a) Expected reward from Q-learning. b) Expected reward from C-MABs. Numbers next to each slice of the pie chart correspond to the number of context arrivals. Actions are ranked based on the expected reward, as indicated by the color bar shown below the pie charts.

Of the maximum expected reward from both Q-learning and C-MABs, only one of the three states ("high") have the matching action with the highest expected reward. The two states ("center" and "low") do not have matching actions for the highest expected maximum rewards, but do have similar actions and expected maximum rewards. The maximum expected reward for the "low" state is received for a downward action for CLUP and a downward diagonal action for Q-Learning. If Q-learning were to continue for additional trials it is likely that it would converge to same maximum expected reward of C-MABs. C-MABs employ a controlled exploration of the state-action space and are able to hone in on the actions that it learns to have the maximum expected reward. By selecting actions to reduce uncertainty, C-MABs are able to increase the confidence of the expected rewards. For example, in *Figure 4.13*b the C-MAB "low" state has spent the majority of its exploitation efforts on actions 225° and 270°. C-MABs are receiving stochastic rewards and alternate between the belief that either action 225° or 270° has the maximum expected reward. By alternating between the two likely correct actions, the learner will converge to the optimal action that returns the highest expected reward.

Efficient exploration during learning will help to reduce uncertainty for the actions suspected of returning larger rewards. In *Figure 4.13*, based on the number of context arrivals, indicated next to each pie slice, it is apparent that C-MABs exhibit a more controlled exploratory behavior. If a state-action pair is not believed to return high rewards, it will only be explored to reduce uncertainty. The robotic system does not waste effort on exploration once the control parameter D is met. In contrast, Q-learning

randomly chooses arms during exploration and does not consider the uncertainty of the expected rewards.

In order to demonstrate that the learned policy was necessary to accomplish the task, naive trajectories of the fingerpad along the zipper were tested. The zipper was placed at a known angle and assumed to follow a straight line. The fingerpads were then moved in a straight line without stopping and without the use of haptic perception or decision-making. All combinations of three trajectories and three zipper orientations (5°, 0°, and -5°, relative to the horizontal) were tested twice. With the exception of one trial (zipper and fingerpad trajectory aligned with the horizontal), the pinch grasp failed to robustly track and seal the zipper bag during these 18 trials. The single successful trial occurred when the preplanned trajectory and orientation of the zipper were initialized perfectly.

Comparison of Reinforcement Learning Algorithms

To compare the Q-learning and CLUP reinforcement learning algorithms, cumulative rewards during learning were evaluated. Cumulative rewards are the sum of all rewards received during the full set of trials. Both reinforcement learning algorithms have the identical state and action space available to them. What differentiates each learner is how actions are chosen during exploration and how expected rewards are updated.

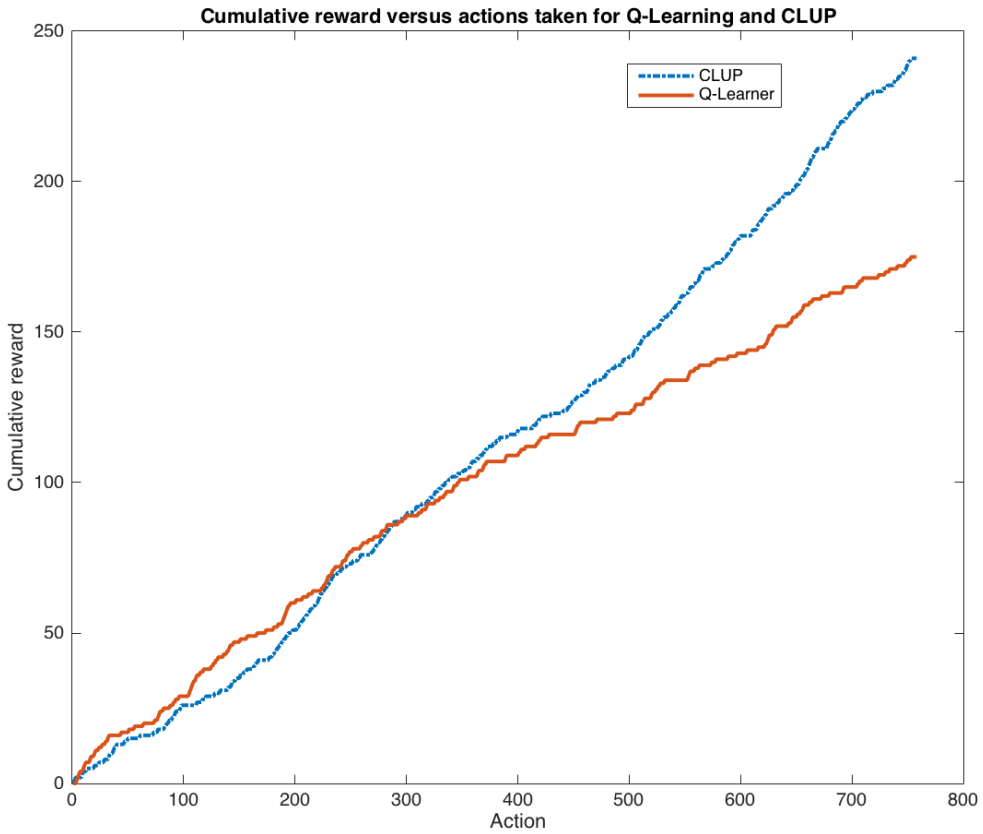

Figure 4.14 Cumulative rewards for over 750 consecutive actions of CLUP (dash-dotted blue) and Q-learning (solid red). CLUP outperforms Q-learning during learning by optimizing exploration versus exploitation.

Q-learning initially outperforms CLUP until approximately 240 trials have elapsed. This is because, based on the control function D, CLUP is purely exploring to gather information about the distribution of expected rewards. Once CLUP has sufficient information about expected rewards, it begins to exploit the policy it is learning. There are also short segments in *Figure 4.14* where the CLUP cumulative reward line appears to be flat. These flat segments correspond to periods when the time-dependent control parameter D changes to the next integer value of context arrivals required for the exploration of the state-action pairs (Cem Tekin & van der Schaar, 2015). This single agent version of CLUP significantly outperforms Q-learning for cumulative rewards for the entire 750-action duration of learning (*Figure 4.14*).

Robustness of the Learned Policy

Robust state classification is critical to the success of any reinforcement learning algorithm. Improving accuracy for compliant sensors such as the BioTac continues to be a challenge (Hui et al., 2016). By taking the action-dependent baseline of the tactile data during each action and then the L2-norm of the baselined data for model features, we observed only minor decreases in classification accuracy for a model trained with data that were 84 days old.

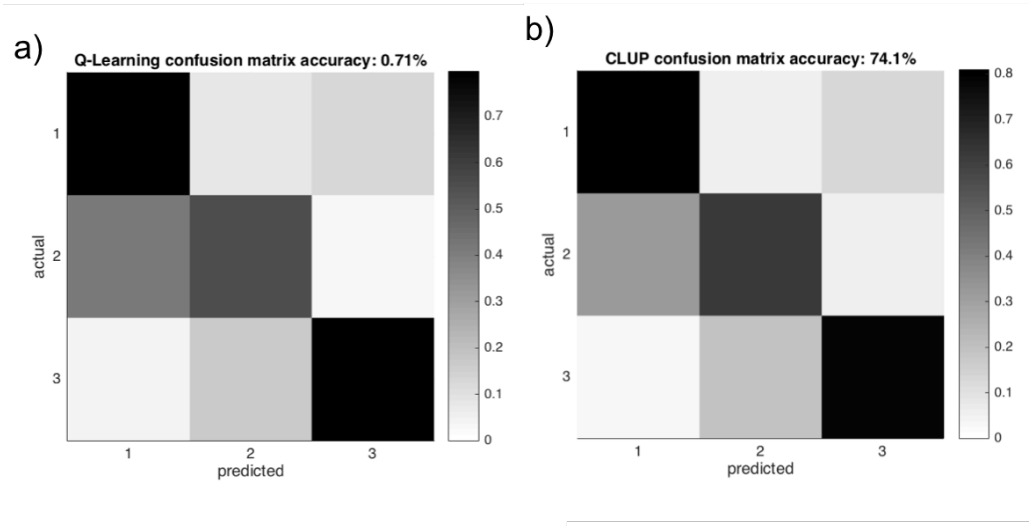

Figure 4.15 Confusion matrix showing classification accuracy during training for (a) Q-learning and (b) CLUP. Both learning algorithms used the same classifier that was trained with data recorded 84 days prior.

The trained DNN had an initial accuracy of 86% on out-of-sample data not seen during training. This trained model was used to classify states during learning for both the Q-learner and CLUP presented in the previous section. State classification during learning was 71% for Q-learning and 74% for CLUP (*Figure 4.15*). The state classification accuracy for both Q-learning and CLUP should be identical, as the trained classifier should degrade uniformly due to environmental changes of the sensor. An average loss of 13% of classifier accuracy over an 84 day period shows the robustness of the classifier.

The average time required to close the ziplock bag was approximately 2.5 minutes. This time fluctuated due to length of the contour along with the efficiency of the actions chosen. While execution time was not the primary focus of this work, it is

important for the future functionality of the system. Significant delays were due to the motion planning time required for the robot arm. Calculations for the DNN and C-MABs required minimal processor time. Future work includes optimizing the motion planning and communication system to reduce the time for task completion. Implementing a graphical processing unit and parallelized planning during movements would significantly reduce total task execution time.

Automated Identification of Task Completion (Zipper Closure)

The robot system has learned a policy for a functional contour-following task, but still needs to learn to identify when the task has been completed (i.e. when the end of the contour has been reached). Completion of the contour-following task is defined as when the pinch grasp has successfully reached the end of the zipper after maintaining contact during all prior actions.

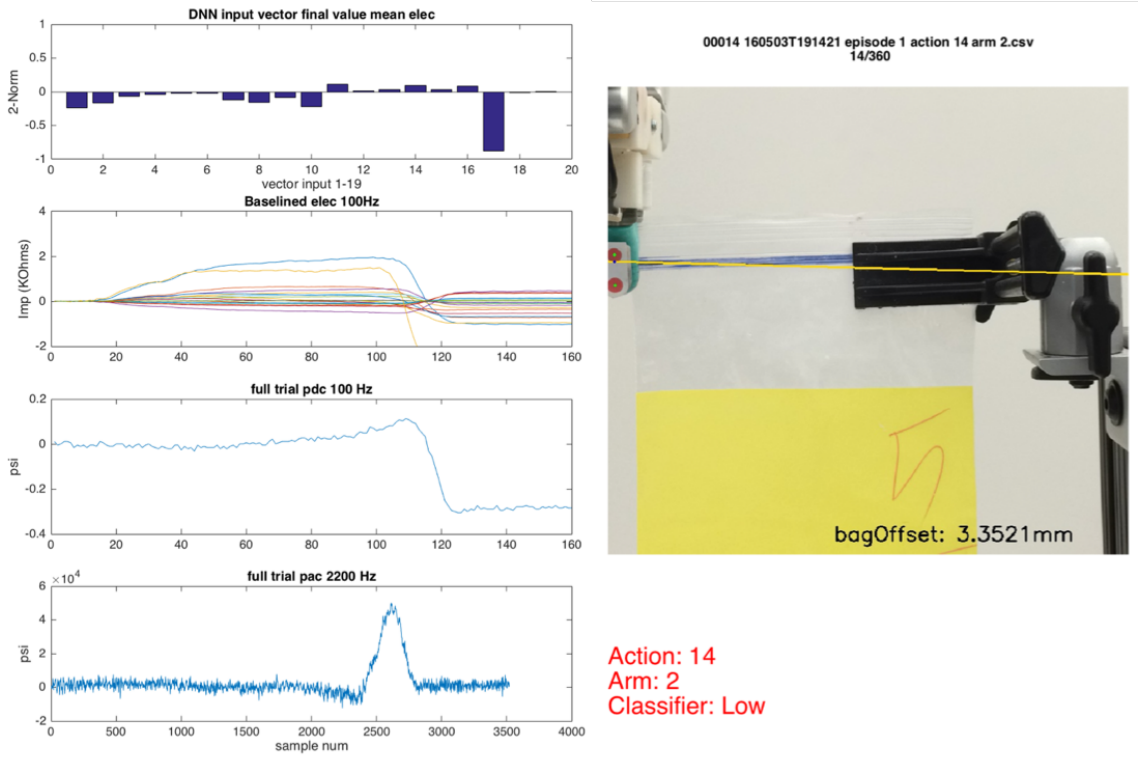

Figure 4.16 BioTac sensor data from an action in which the ziplock bag starts in contact with the fingertips and is lost. Both low (100 Hz) and high frequency (2200 Hz) fluid pressure signals, third and fourth rows, respectively, show clear signs of loss of contact.

In order to analyze what features are important for task completion, we ran 20 trials of exactly 18 actions each, resulting in a total of 360 actions. During these 20 training trials, even if the finger was the no longer in contact we continued executing until the full 18 actions were completed. For the 20 trials, the pinch grasp averaged 14 consecutive actions before a loss of contact (*Figure 4.16*, example loss of contact). BioTac fluid pressure measurements from the low frequency pressure sensor of the BioTac were robust indicators of loss of contact. *Figure 4.17* shows all 360 actions; the blue lines with diamond markers represent actions that resulted in loss of contact with the ziplock bag (*Figure 4.16*). By using a simple baselined threshold on the low frequency (100 Hz) pressure signal, we were able to accurately determine task completion. All trials conducted after the implementation of such a pressure threshold were able to identify loss of contact and task completion.

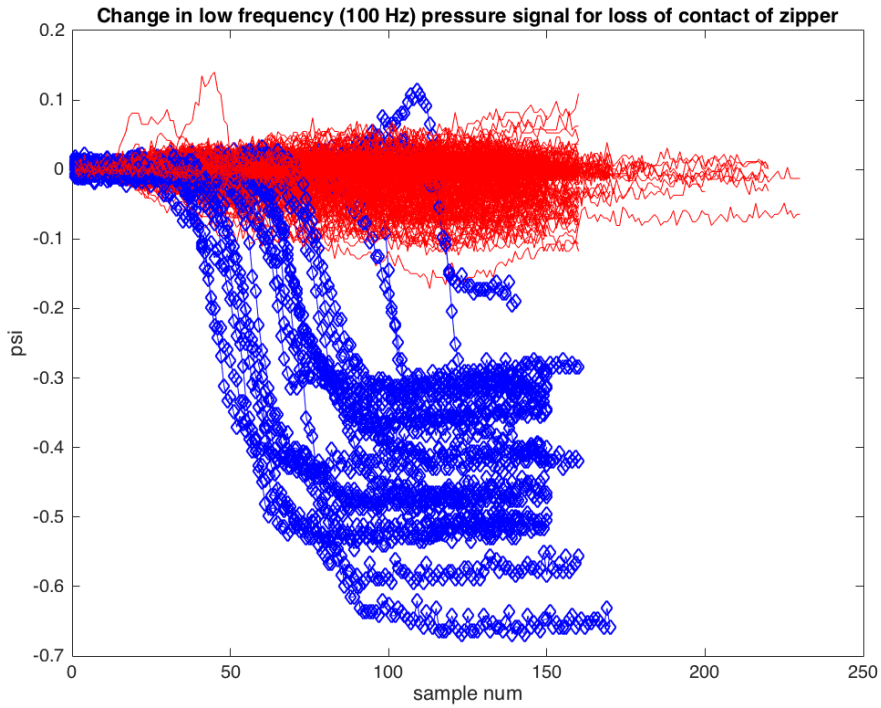

Figure 4.17 Analysis of low frequency (100 Hz) pressure during actions shown in *Figure 4.9*. Blue lines with diamond markers represent the 20 action instances that resulted in loss of contact with the ziplock bag. Red lines represent the 340 other actions.

Of the 20 trials, the first 10 trials were conducted with the ziplock bag open (*Figure 4.18*, red) while the last 10 trials were conducted with the ziplock bag sealed (*Figure 4.18*, blue). All trials successfully followed the contour of the zipper; all 10 trials in which the ziplock bag was initally open were successfully sealed. We postulate that the vibration sensed by the BioTac will be different when the bag is initially open as compared to when the bag is already closed. The power spectral density of the red trials (*Figure 4.18*), for which the ziplock bag was initially open and then successfully sealed, have a higher distribution up to 10 Hz. In contrast, the blue trials (*Figure 4.18*), for which

the bag was already closed, have a clear low frequency peak in the power spectral density. This suggests that additional vibration is generated from the zipper engagement during the last action that is made while in contact with the ziplock bag. The difference in the vibration signature generated from sealing the zipper could be used to determine the successful closure of a ziplock bag, and is left for future work.

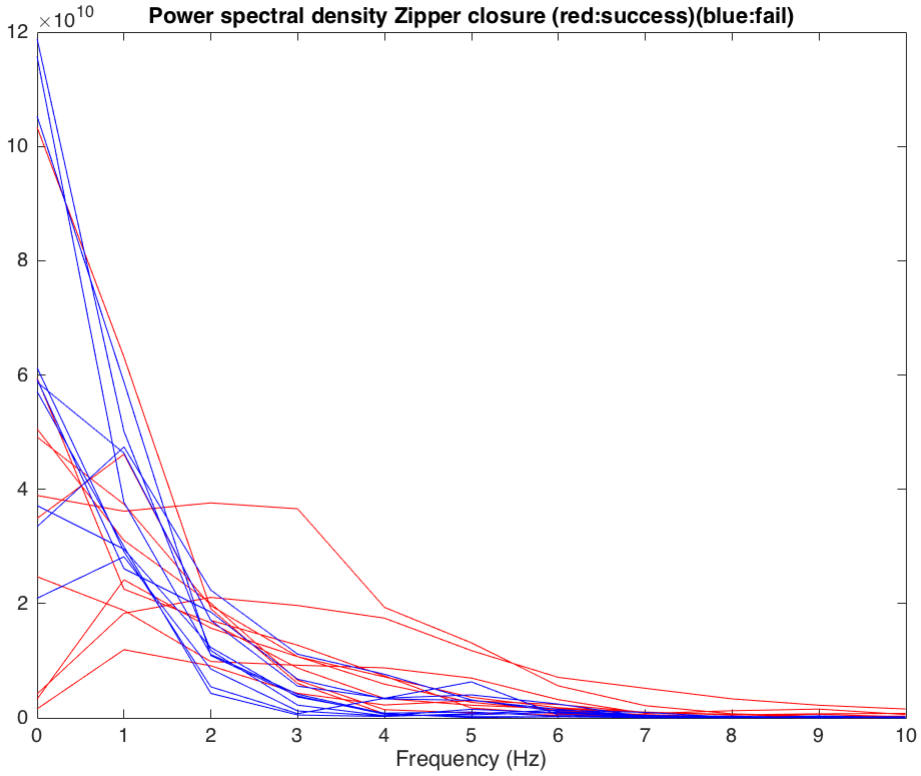

Figure 4.18 Power spectral density of the last action made while in contact with the ziplock bag during contour following. Red and blue trials correspond to initially open and closed bags, respectively.

Generalizability to Novel Contours

Now that the system is trained to close a particular ziplock bag, the ability of the system to generalize to new deformable contours is explored. New ziplock bags not used for training were tested with the trained classifier and learned policies. Bags were also filled with various distributions of 10 grams of cereal to test system robustness.

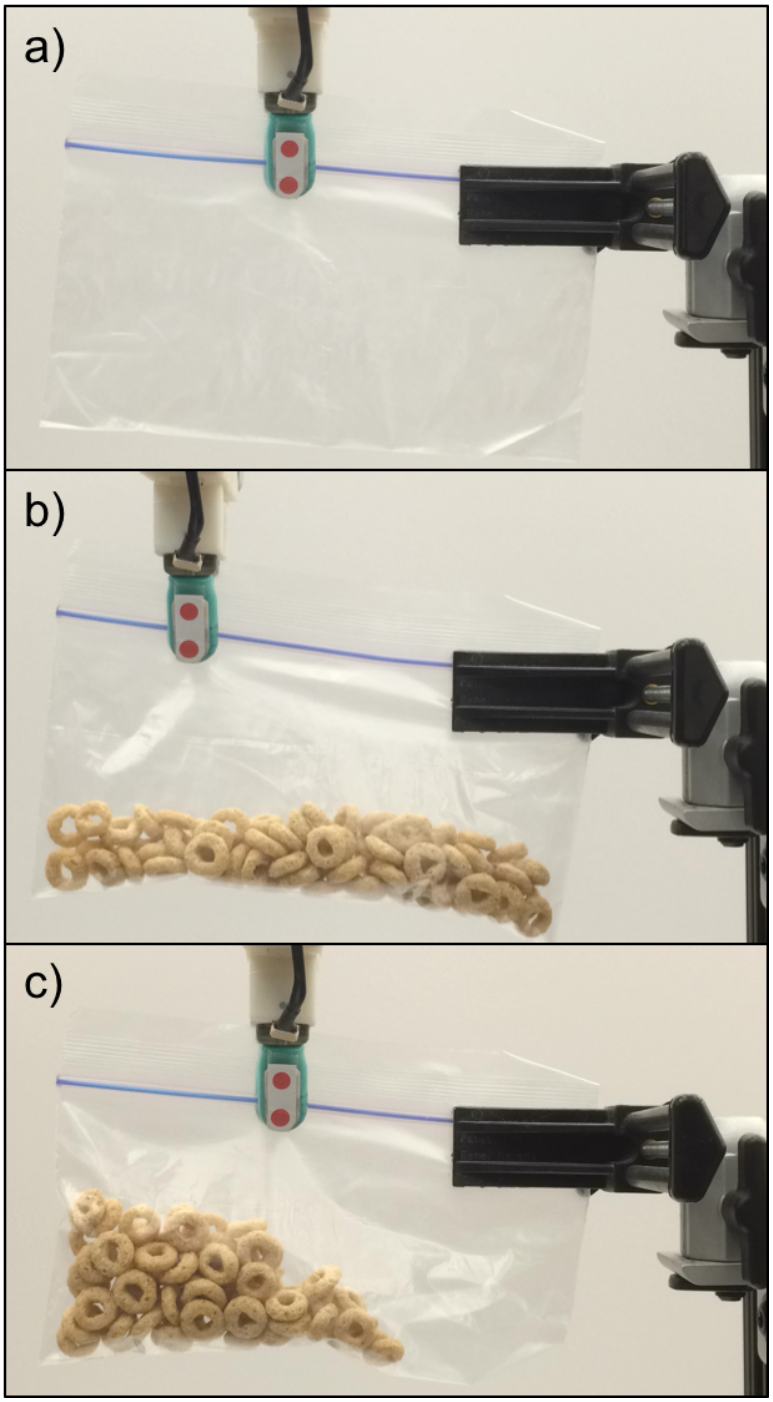

Figure 4.19 System tested on a novel ziplock bag not used during classifier training or policy learning. a) Ziplock bag with narrower, more deformable zipper than that of the bag used for learning the policy. b) Bag with 10 grams of cereal evenly distributed. c) Bag with 10 grams of cereal positioned toward the end of the bag.

The ziplock bag used for testing is shown in *Figure 4.19* and was chosen as the most generic thin-walled ziplock bag available at the supermarket. Actual training data were collected with an industrial-style bag used for storing small machine parts, as shown in *Figure 4.10*. While the industrial bag lends itself to thousands of training trials, the thin-walled ziplock bag is more realistic as an everyday household bag. The classifier and policy were able to generalize and successfully close all five trials for each of the ziplock bag configurations shown in *Figure 4.19*.

Two weight distributions were used for the trials in which the bag contained 10 grams of cereal. The first configuration was with the weight of the cereal evenly distributed and the second where the weight was distributed near the end of the bag. Having the weight distributed near the end of the bag caused the contour to have an exaggerated downward slope. The classifier was able to recognize that the zipper was in the "low" state and take corrective actions to correctly track the zipper. It should also be noted that the system was able to aggressively correct the "low" state (*Figure 4.8*) due to the maximum expected reward of the CLUP policy (*Figure 4.13*) with the asymmetric actions.

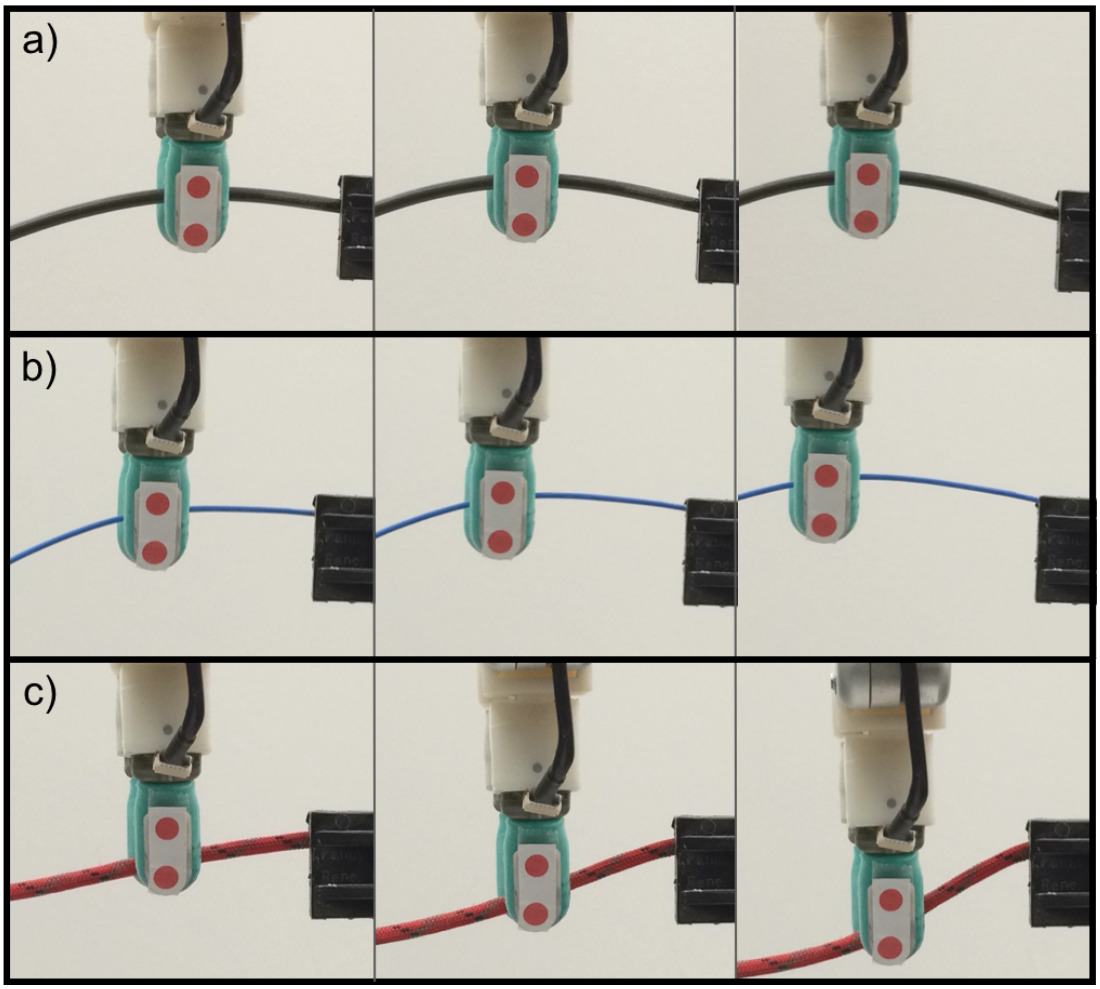

Figure 4.20 Policy tested on three non-zipper deformable contours, (a) thick electrical wire (3.5 mm), (b) thin electrical wire (1.5mm), and (c) nylon rope (4 mm). Each row of the figure shows sequential actions. None of the contours shown were used during classifier training or policy learning.

Additional non-zipper contours were tested to determine the robustness and generalizability of the system. *Figure 4.20*a-c shows three novel contours, a thick electrical wire (3.5 mm), thin electrical wire (1.5 mm), and nylon rope (4 mm). The policy was unable to maintain the contours in the desired "center" state of the grasp but

was robust in classifying the current state of the contour in the grasp. Since the classifier correctly identified the state of the contour, the correct action was robustly taken in an attempt to return the contour to the "center" of the grasp. Each row of *Figure 4.20* show sequential actions. The robot was able to follow the contour for approximately three sequential actions before loss of contact. Low friction between the silicone skin and that of the electrical wire caused the contour to easily slip out the grasp. Grasp pressure was held constant for all actions, all trials, and all contours (training and testing). Modulating the grasp pressure during each action could improve performance for round contours. The modulation of the grasp pressure may enable the learning of fingerpad-contour interactions that are robust to differences in the size and shape of novel contours. Normal forces generated from the shape of the fingerpad along with the round shape of the wire and rope contours made it difficult for the system to maintain the contour in the "center" state. The actions chosen by the policies attempted to correct the state of the contours within the grasp, but were limited by the policies that were trained for a ziplock bag and its comparatively flat zipper. The state classifier was generalizable across different contour sizes and shapes. However, the policy for the flat zipper did not generalize to the round wires and rope, as the learned policy was unsuccessful in returning those contours to the "center" state after multiple fingertip movements along the contour. Adjustments to the policy through additional learning may allow it to perform more robustly. By initializing a new policy for the contours in *Figure 4.20* with the expected rewards of the previously trained model for zippers, a reduction in total training/exploration time may be achieved.

Summary

The work presented in the chapter represents the development of a robotic system capable of learning a haptics-intensive functional manipulation task. A ziplock bag was used for manipulation task because it introduced many interesting challenges, such as being deformable, transparent, and visually occluded by the artificial fingertips. The approach presented was model-free and the policy was learned directly from exploration. Once the system learned a policy, the task was accomplished through the use of only localized tactile data and proprioception. By leveraging localized tactile data, we were able to reduce potential sources of error. As robots advance and begin to interact with cluttered environments, incidental contacts will reduce the usefulness of proprioceptive feedback for the perception of end effector interactions with the environment.

When defining the state and action space it was necessary to define the spaces in a way that allows the algorithm to find interesting solutions that might not be apparent to the designer. Non-intuitive solutions can be found by exploring all areas of the state space and minimizing any uncertainty in the distribution of expected rewards. Although, the additional freedom of exploration comes at a price, as it was important to keep the required training tractable given the high cost of collecting tactile data. Tactile data is expensive because it is difficult to simulate, is time consuming to collect, and causes wear of the entire hardware system.

The system developed was capable of learning a functional task while optimizing the cumulative rewards received. Policies trained in this chapter were also able to generalize to novel zippers and bags with various weight distributions. A reduction of

training time could result from initializing policies for similar tasks with previously trained model parameters.

New reinforcement learning algorithms could also be considered in order to enhance performance. Adaptive algorithms that can zoom in on regions of higher state occupancy (context arrivals) can reduce the uncertainty of expected rewards (C. Tekin & Van der Schaar, 2015). By adaptively zooming in and further partitioning the context space, exploration remains tractable and focuses exploratory effort on contexts with higher state occupancy.

Accurate state classification is of critical importance as inaccurate labeling will cause undesired stochasticity in the rewards. The state classification of tactile data is something that will continue to be developed as new techniques for real-time perception are explored. Future work will explore additional actions such as online adjustments to grasp pressure, confidence-based actions, and out of plane movements and rotation.

CHAPTER 5

SUMMARY AND CONCLUSIONS

The work presented in this dissertation has advanced the abilities of artificial hands in areas of dexterity, perception, decision making, and learning. The general theme of much of this work focused on the localized contact between an object and a sensorized fingerpad. By developing the abilities of artificial hands and haptic perception at the location of contact we remove multiple sources of error and create robust systems capable of online perception and manipulation.

Major Contributions

Novel design of robot testbed with rich tactile sensing and high dexterity: In order to perceive the environment through touch, a robotic system must be capable of executing dexterous movements. The robot testbed has human-like dexterity and is capable in both speed and strength. The actuation system measures proprioceptive tendon tension and allows for the actuation of any tendon-driven end effector. Modular design of the actuation system allows for quick setup of an end effector and the ability to control additional degrees of freedom. The asymmetric design of the finger allows for multimodal tactile sensing with the BioTac. The range of capacities of the testbed were shown in a simple tap-and-hold experiment. Applications of the testbed include the development of algorithms for haptic perception and manipulation, and aiding in rehabilitation for upper-limb impairment or loss.

Online haptic perception to aid in the design of manipulation algorithms: By developing methods to haptically perceive the features of an object during motion we can

take corrective actions to accomplish a task or to better guide the exploration of an object. In this work we explored the ability of HMMs to provide anytime predictions of object features.

Development of task-driven haptic perception, decision-making, and learning: In order to build a system capable of completing a functional contour-following task, perception, decision-making and learning all need to work in unison. The closure of a plastic ziplock bag was chosen as the task, as it is an everyday task that presents unique challenges for robotic systems. The task requires the manipulation of a transparent, deformable object whose geometric features are visually occluded by artificial fingertips. Novel contributions to the online perception of the grasp state focused on action-dependent baselines. Performance comparison of two reinforcement learning algorithms were explored to learn and generalize a functional contour-following task. We compared the use of minimally processed tactile signals as context versus a classifier to determine context. The learned policy and classifier were able to generalize to novel zipper contours and bags with various weight distributions.

Directions for Future Work

Expand State-Action Space Toward Continuous Representation

The state and action space presented in Chapter 4 represent a small combination of the possible representations of the state of the fingerpad and the available actions that an unconstrained robot hand could perform. The state space used in Chapter 4 was defined according to the task. To improve the functionality for other tasks, it is necessary to expand the state representation on the fingerpad. Incorporating state confidence and considering that multiple state representations can occur simultaneously will add to the generalizability of the model. An example of a "simultaneous state representation" could include the perception that the zipper is in the "low" position of the fingerpad and the perception that the zipper is moving in the proximal direction along the fingerpad. Instead of learning all possible combinations of the policy, a blending of states could be explored similar to the forward algorithm Eq. (3.10) in Chapter 3. The forward algorithm is able to produce a model likelihood while not iterating over the full set of states. By blending the expected rewards, we may be able to provide an approximation of the actual rewards while keeping the training of the system tractable. This technique of blending state spaces and policies together could additionally improve the generalizability of the system.

There are infinitely many actions that the robot end effector can take. Determining the best set of motions to keep training tractable is often a challenge because one needs to consider the amount of training time required for a given task. For the function contour-following task of Chapter 4, fingertip motion was constrained to the plane of the zipper and maintained at a constant orientation. Future work could expand upon the available action set. By considering the level of confidence of the expected

reward based on the number of occurrences in the C-MAB, the action space can be modified. Allowing the system to explore and expand the available actions in areas of high expected rewards may improve performance. Additionally, the expansion of the action space can be based on object affordances or the desired goal of the task (De Granville, Wang, Southerland, Fagg, & Platt, 2009).

Incorporate New Learning Models

The Distributed Context Zooming Algorithm (DCZA) can be used to adaptively partition the context based on the number of arrivals (C. Tekin & Van der Schaar, 2015). The assumption of CLUP used in Chapter 4 was that each partition of the context had an equal probability of arrival. With DCZA, the context will adaptively zoom in on areas of higher arrival to more evenly explore and discretize the context space. Context arrivals will determine the partitioning of the state space. The action space should also be adaptively modified in areas of high uncertainty. By adaptively, purposefully zooming on areas on interest, training can remain tractable.

Anytime State Predictions

Anytime haptic prediction developed in Chapter 3 was a major advancement in the abilities of the perception system utilized in this work. This advancement influenced greatly the work in Chapter 4. Having a system capable of perceiving features while executing an "exploratory motion" allows the system to make the most informed decisions for corrective actions or planning. For the work presented in Chapter 4, we did not use anytime algorithms mainly due to the robustness and ability of DNN to fit nonlinear data. However, the classifiers in Chapter 4 relied on 0.5 cm and 0.75 cm displacement trajectories to build the feature inputs to a DNN. These short displacement trajectories might result in similar outcomes as the results from anytime algorithms. What is important in the development of artificial perception capabilities is that systems are able to perceive more frequently and based on shorter sensory periods. In order to expand the capabilities developed in this dissertation, a larger haptic and manipulation library will need to be built that is continuously searching for features and evaluating additional inputs from all sensory modalities, including vision and sound. Context from multiple sensory modalities will help to reduce uncertainty and allow for the development of effective robotic systems that can learn how to perform novel functional tasks.

REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., … Zheng, X. (2016). TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. *arXiv:1603.04467 [cs]*. Retrieved from http://arxiv.org/abs/1603.04467

Barnsley, N., McAuley, J. H., Mohan, R., Dey, A., Thomas, P., & Moseley, G. L. (2011). The rubber hand illusion increases histamine reactivity in the real arm. *Current Biology*, *21*(23), R945–R946. http://doi.org/10.1016/j.cub.2011.10.039

Belter, J. T., & Dollar, A. M. (2011). Performance characteristics of anthropomorphic prosthetic hands. In *Proc IEEE Intl Conf Rehab Robotics* (pp. 1–7). Zurich, Switzerland.

Belter, J. T., Segil, J. L., Dollar, A. M., & Weir, R. F. (2013). Mechanical design and performance specifications of anthropomorphic prosthetic hands: A review. *Journal of Rehabilitation Research & Development*, *50*(5), 599–617. http://doi.org/10.1682/JRRD.2011.10.0188

Berenson, D. (2013). Manipulation of deformable objects without modeling and simulating deformation. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 4525–4532). http://doi.org/10.1109/IROS.2013.6697007

Bhattacharjee, T., Shenoi, A. A., Park, D., Rehg, J. M., & Kemp, C. C. (2015). Combining Tactile Sensing and Vision for Rapid Haptic Mapping. In *Proc IEEE/RSJ Intl Conf Intelligent Rob Sys*. Hamburg, Germany.

Biddiss, E. A., & Chau, T. T. (2007). Upper limb prosthesis use and abandonment: A survey of the last 25 years. *Prosthetics and Orthotics International*, *31*(3), 236–257. http://doi.org/10.1080/03093640600994581

Biddiss, E., Beaton, D., & Chau, T. (2007). Consumer design priorities for upper limb prosthetics. *Disability & Rehabilitation: Assistive Technology*, *2*(6), 346–357. http://doi.org/10.1080/17483100701714733

Biddiss, E., & Chau, T. (2007). Upper-Limb Prosthetics: Critical Factors in Device Abandonment. *American Journal of Physical Medicine & Rehabilitation*, *86*(12), 977–987. http://doi.org/10.1097/PHM.0b013e3181587f6c

Botvinick, M., & Cohen, J. (1998). Rubber hands `feel' touch that eyes see. *Nature*, *391*, 756. http://doi.org/10.1038/35784

Bretl, T., & McCarthy, Z. (2013). Mechanics and Quasi-Static Manipulation of Planar Elastic Kinematic Chains. *IEEE Transactions on Robotics*, *29*(1), 1–14. http://doi.org/10.1109/TRO.2012.2218911

Bubeck, S., Munos, R., Stoltz, G., & Szepesvari, C. (2011). *X-Armed Bandits* (Post-Print). HAL. Retrieved from http://econpapers.repec.org/paper/haljournl/hal-00450235.htm

Chen, N., Zhang, H., & Rink, R. (1995). Edge tracking using tactile servo. In *Proc IEEE/RSJ Intl Conf on Intelligent Robots and Systems, "Human Robot Interaction and Cooperative Robots"* (Vol. 2, pp. 84 –89). http://doi.org/10.1109/IROS.1995.526143

Controzzi, M., Cipriani, C., & Carrozza, M. C. (2014). Design of Artificial Hands: A Review. In R. Balasubramanian & V. J. Santos (Eds.), *The Human Hand as an Inspiration for Robot Hand Development* (Vol. 95, pp. 219–246). Cham: Springer International Publishing. Retrieved from http://link.springer.com/10.1007/978-3-319-03017-3_11

Dalley, S. A., Wiste, T. E., Withrow, T. J., & Goldfarb, M. (2009). Design of a Multifunctional Anthropomorphic Prosthetic Hand With Extrinsic Actuation. *IEEE/ASME Transactions on Mechatronics*, *14*(6), 699–706. http://doi.org/10.1109/TMECH.2009.2033113

D'Alonzo, M., & Cipriani, C. (2012). Vibrotactile Sensory Substitution Elicits Feeling of Ownership of an Alien Hand. *PLoS ONE*, *7*(11), 1–9. http://doi.org/10.1371/journal.pone.0050756

Darnall, B. D. (2009). Self-Delivered Home-Based Mirror Therapy for Lower Limb Phantom Pain: *American Journal of Physical Medicine & Rehabilitation*, *88*(1), 78–81. http://doi.org/10.1097/PHM.0b013e318191105b

De Granville, C., Wang, D., Southerland, J., Fagg, A. H., & Platt, R. (2009). Grasping Affordances: Learning to Connect vision to Hand Action. In *The Path to Autonomous Robots* (pp. 1–22). Retrieved from http://www.springerlink.com/index/k8321g4581427r28.pdf

Dickstein, R., & Deutsch, J. E. (2007). Motor Imagery in Physical Therapist Practice. *Physical Therapy*, *87*(7), 942–953. http://doi.org/10.2522/ptj.20060331

Dietrich, C., Walter-Walsh, K., Preißler, S., Hofmann, G. O., Witte, O. W., Miltner, W. H. R., & Weiss, T. (2012). Sensory feedback prosthesis reduces phantom limb pain: Proof of a principle. *Neuroscience Letters*, *507*(2), 97–100. http://doi.org/10.1016/j.neulet.2011.10.068

Dollar, A. M., & Howe, R. D. (2010). The highly adaptive SDM Hand: Design and performance evaluation. *The International Journal of Robotics Research*, *29*(5), 585 –597. http://doi.org/10.1177/0278364909360852

Doubler, J. A., & Childress, D. S. (1984). Design and evaluation of a prosthesis control system based on the concept of extended physiological proprioception. *Journal of Rehabilitation Research and Development*, *21*(1), 19–31.

Dudik, M., Hsu, D., Kale, S., Karampatziakis, N., Langford, J., Reyzin, L., & Zhang, T. (2011). Efficient Optimal Learning for Contextual Bandits. *arXiv:1106.2369 [cs, Stat]*. Retrieved from http://arxiv.org/abs/1106.2369

Farrell, T. R., & Weir, R. F. (2007). The Optimal Controller Delay for Myoelectric Prostheses. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *15*(1), 111–118. http://doi.org/10.1109/TNSRE.2007.891391

Gales, M., & Young, S. (2007). The Application of Hidden Markov Models in Speech Recognition. *Found. Trends Signal Process.*, *1*(3), 195–304. http://doi.org/10.1561/2000000004

Gialias, N., & Matsuoka, Y. (2004). Muscle actuator design for the ACT Hand. In *Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on* (Vol. 4, pp. 3380–3385 Vol.4).

Gittins, J., Glazebrook, K., & Weber, R. (2011). *Multi-armed Bandit Allocation Indices*. John Wiley & Sons.

Go, A. S., Mozaffarian, D., Roger, V. L., Benjamin, E. J., Berry, J. D., Blaha, M. J., … American Heart Association Statistics Committee and Stroke Statistics Subcommittee. (2014). Heart disease and stroke statistics--2014 update: a report from the American Heart Association. *Circulation*, *129*(3), 1–267. http://doi.org/10.1161/01.cir.0000441139.02102.80

Grebenstein, M., Chalon, M., Hirzinger, G., & Siegwart, R. (2010). Antagonistically driven finger design for the anthropomorphic DLR Hand Arm System. In *2010 10th IEEE-RAS International Conference on Humanoid Robots (Humanoids)* (pp. 609–616). IEEE. http://doi.org/10.1109/ICHR.2010.5686342

Harris, A. J. (1999). Cortical origin of pathological pain. *The Lancet*, *354*(9188), 1464–1466. http://doi.org/10.1016/S0140-6736(99)05003-5

Hartigan, J. A. (1979). Algorithm AS136 : A K-means clustering algorithm. *Applied Statistics*, *28*, 100–108. http://doi.org/10.2307/2346830

Hartmann, C., Dosen, S., & Farina, D. (2014). Sense of Embodiment of an Artificial Finger Following Sensory Substitution by Electrocutaneous Stimulation of the Forearm. In *Proc IEEE Engin Med Biol Soc*. Chicago, IL.

Hellman, R. B., Chang, E., Tanner, J., Helms Tillery, S. I., & Santos, V. J. (2014). Tendon-driven testbed for haptic exploration and sensory event-driven grasp and manipulation. In *Proc DARPA MTO Sensorimotor Prosthetics Workshop*. Scottsdale, AZ.

Hellman, R. B., Chang, E., Tanner, J., Helms Tillery, S. I., & Santos, V. J. (2015). A Robot Hand Testbed Designed for Enhancing Embodiment and Functional Neurorehabilitation of Body Schema in Subjects with Upper Limb Impairment or Loss. *Frontiers in Human Neuroscience*, *9*. http://doi.org/10.3389/fnhum.2015.00026

Hellman, R. B., & Santos, V. J. (2012). Design of a back-driveable actuation system for modular control of tendon-driven robot hands. In *Proc of IEEE RAS/EMBS Intl Conf on Biomedical Robotics and Biomechatronics* (pp. 1293–1298). Roma, Italy.

Hellman, R. B., & Santos, V. J. (2015). Development of probabilistic models for real-time perception of geometric features with a sensorized artificial finger. In *Workshop on "Robotic Hands, Grasping and Manipulation," IEEE/RAS Intl Conf on Robotics and Automation*. Seattle, WA.

Hellman, R. B., & Santos, V. J. (2016). Haptic perception and decision-making for a functional contour-following task. In *Proc IEEE Haptics Symposium, Work-in-progress paper*. Philadelphia, PA.

Hill, W., Kyberd, P., Hermansson, L. N., Hubbard, S., Stavdahl, Ø., & Swanson, S. (2009). Upper limb prosthetic outcome measures (ULPOM): A working group and their findings. *JPO: Journal of Prosthetics and Orthotics*, *21*(9), P69–P82.

Hogan, N. (1985). Impedance control - An approach to manipulation. I - Theory. II - Implementation. III - Applications. *ASME Transactions Journal of Dynamic Systems and Measurement Control B*, *107*, 1–24.

Hui, J. C. T., Block, A. E., Taylor, C. J., & Kuchenbecker, K. J. (2016). Robust tactile perception of artificial tumors using pairwise comparisons of sensor array readings. In *2016 IEEE Haptics Symposium (HAPTICS)* (pp. 305–312). http://doi.org/10.1109/HAPTICS.2016.7463194

iOS Apple Inc. (2014). Retrieved from www.apple.com

Irani, M., Rousso, B., & Peleg, S. (1994). Computing occluding and transparent motions. *International Journal of Computer Vision*, *12*(1), 5–16. http://doi.org/10.1007/BF01420982

Jacobsen, S., Ko, H., Iversen, E., & Davis, C. (1989). Antagonistic control of a tendon driven manipulator. In *Robotics and Automation, 1989. Proceedings., 1989 IEEE International Conference on* (pp. 1334–1339).

Jimenez, M. C., & Fishel, J. . (2014). Evaluation of force, vibration and thermal tactile feedback in prosthetic limbs. In *2014 IEEE Haptics Symposium (HAPTICS)* (pp. 437–441). http://doi.org/10.1109/HAPTICS.2014.6775495

Johansson, R. S., & Flanagan, J. R. (2009). Coding and Use of Tactile Signals from the Fingertips in Object Manipulation Tasks. *Nature Reviews Neuroscience*.

Johnson, S., Hall, J., Barnett, S., Draper, M., Derbyshire, G., Haynes, L., … Goebel, A. (2012). Using graded motor imagery for complex regional pain syndrome in clinical practice: failure to improve pain. *European Journal of Pain (London, England)*, *16*(4), 550–561. http://doi.org/10.1002/j.1532-2149.2011.00064.x

Kaczmarek, K. A., Webster, J. G., Bach-y-Rita, P., & Tompkins, W. J. (1991). Electrotactile and vibrotactile displays for sensory substitution systems. *IEEE Transactions on Bio-Medical Engineering*, *38*(1), 1–16. http://doi.org/10.1109/10.68204

Keenan, K. G., & Massey, W. V. (2012). Control of Fingertip Forces in Young and Older Adults Pressing against Fixed Low- and High-Friction Surfaces. *PLoS ONE*, *7*(10), 1–7. http://doi.org/10.1371/journal.pone.0048193

Klatzky, R. L., & Lederman, S. J. (1999). The haptic glance: A route to rapid object identification and manipulation. In D. Gopher & A. Koriats (Eds.), *Attention and Performance XVII: Cognitive Regulation of Performance: Interaction of Theory and Application*.

Kleinberg, R., Slivkins, A., & Upfal, E. (2008). Multi-Armed Bandits in Metric Spaces. *arXiv:0809.4882 [cs]*. Retrieved from http://arxiv.org/abs/0809.4882

Klingensmith, M., Sirinivasa, S. S., & Kaess, M. (2016). Articulated Robot Motion for Simultaneous Localization and Mapping (ARM-SLAM). *IEEE Robotics and Automation Letters*, *1*(2), 1156–1163.

Koch, H., Konig, A., Weigl-Seitz, A., Kleinmann, K., & Suchy, J. (2013). Multisensor Contour Following With Vision, Force, and Acceleration Sensors for an Industrial Robot. *IEEE Transactions on Instrumentation and Measurement*, *62*(2), 268–280. http://doi.org/10.1109/TIM.2012.2214934

Kuiken, T. A., Marasco, P. D., Lock, B. A., Harden, R. N., & Dewald, J. P. (2007). Redirection of cutaneous sensation from the hand to the chest skin of human amputees with targeted reinnervation. *Proceedings of the National Academy of Sciences*, *104*(50), 20061–20066.

Kyranou, I. (2014). *Improving Robotic Prosthetic Hand Performance Through Grasp Preshaping* (MS Thesis). University of Edinburgh.

Lackner, J. R. (1988). Some proprioceptive influences on the perceptual representation of body shape and orientation. *Brain: A Journal of Neurology*, *111 (Pt. 2)*, 281–297.

Langford, J., & Zhang, T. (2008). The Epoch-Greedy Algorithm for Multi-armed Bandits with Side Information. In J. C. Platt, D. Koller, Y. Singer, & S. T. Roweis (Eds.), *Advances in Neural Information Processing Systems 20* (pp. 817–824). Curran Associates, Inc. Retrieved from http://papers.nips.cc/paper/3178-the-epoch-greedy-algorithm-for-multi-armed-bandits-with-side-information.pdf

Lazher, Z., Belhassen-Chedli, B., Sabourin, L., & Youcef, M. (2014). Modeling and analysis of 3D deformable object grasping. In *2014 23rd International Conference on Robotics in Alpe-Adria-Danube Region (RAAD)* (pp. 1–8). http://doi.org/10.1109/RAAD.2014.7002259

Lederman, S. J., & Klatzky, R. L. (1987). Hand Movements: A Window Into Haptic Object Recognition. *Cognitive Psychology*, *19*(3), 342–368.

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

London, B. M., Jordan, L. R., Jackson, C. R., & Miller, L. E. (2008). Electrical Stimulation of the Proprioceptive Cortex (Area 3a) Used to Instruct a Behaving Monkey. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *16*(1), 32–36. http://doi.org/10.1109/TNSRE.2007.907544

Lotti, F., Tiezzi, P., Vassura, G., Biagiotti, L., Palli, G., & Melchiorri, C. (2005). Development of UB Hand 3: Early results. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation, 2005. ICRA 2005* (pp. 4488– 4493). IEEE. http://doi.org/10.1109/ROBOT.2005.1570811

Loucks, C., Johnson, V., Boissiere, P., Starr, G., & Steele, J. (1987). Modeling and control of the Stanford/JPL hand. In *1987 IEEE International Conference on Robotics and Automation. Proceedings* (Vol. 4, pp. 573– 578). IEEE. http://doi.org/10.1109/ROBOT.1987.1088031

Lovchik, C. S., & Diftler, M. A. (1999). The Robonaut hand: a dexterous robot hand for space. In *1999 IEEE International Conference on Robotics and Automation, 1999. Proceedings* (Vol. 2, pp. 907–912 vol.2). IEEE. http://doi.org/10.1109/ROBOT.1999.772420

Maitin-Shepard, J., Cusumano-Towner, M., Lei, J., & Abbeel, P. (2010). Cloth grasp point detection based on multiple-view geometric cues with application to robotic towel folding. In *2010 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 2308–2315). http://doi.org/10.1109/ROBOT.2010.5509439

Marasco, P. D. (2014a). Vibration of reinnervated muscle induces perceptual illusion of joint movement in upper limb amputees with targeted sensory reinnervation. Presented at the Cold Spring Harbor Laboratory 79th Symposium - Cognition, Cold Spring Harbor, NY.

Marasco, P. D. (2014b, August). *Engineered Perception*. Presented at the Myoelectric Controls/Powered Prosthetics Symp, Fredericton, New Brunswick, Canada.

Marasco, P. D., Kim, K., Colgate, J. E., Peshkin, M. A., & Kuiken, T. A. (2011). Robotic touch shifts perception of embodiment to a prosthesis in targeted reinnervation amputees. *Brain*, *134*(3), 747–758. http://doi.org/10.1093/brain/awq361

Medynski, C., & Rattray, B. (2011). Bebionic prosthetic design. In *Proc MyoElectric Controls/Powered Prosthetics Symp*. Fredericton, New Brunswick, Canada. Retrieved from http://dukespace.lib.duke.edu/dspace/handle/10161/4733

Melchiorri, C., & Kaneko, M. (2008). Robot hands. In B. Siciliano & O. Khatib (Eds.), *Springer Handbook of Robotics* (pp. 345–360). Berlin Heidelberg: Springer-Verlag. Retrieved from http://dx.doi.org/10.1007/978-3-540-30301-5_29

Moseley, G. L. (2006). Graded motor imagery for pathologic pain A randomized controlled trial. *Neurology*, *67*(12), 2129–2134. http://doi.org/10.1212/01.wnl.0000249112.56935.32

Moseley, G. L., Olthof, N., Venema, A., Don, S., Wijers, M., Gallace, A., & Spence, C. (2008). Psychologically induced cooling of a specific body part caused by the illusory ownership of an artificial counterpart. *Proceedings of the National Academy of Sciences*, *105*(35), 13169–13173. http://doi.org/10.1073/pnas.0803768105

Murray, R. M., Li, Z., Sastry, S. S., & Sastry, S. S. (1994). *A Mathematical Introduction to Robotic Manipulation*. CRC Press.

Nahvi, A., Hollerbach, J. M., Yangming Xu, & Hunter, I. W. (1994). An investigation of the transmission system of a tendon driven robot hand. In *Proceedings of the IEEE/RSJ/GI International Conference on Intelligent Robots and Systems '94. "Advanced Robotic Systems and the Real World", IROS '94* (Vol. 1, pp. 202–208 vol.1). IEEE. http://doi.org/10.1109/IROS.1994.407390

Nakhaeinia, D., Payeur, P., & Laganiere, R. (2014). Adaptive Robotic Contour Following from Low Accuracy RGB-D Surface Profiling and Visual Servoing. In *2014 Canadian Conference on Computer and Robot Vision (CRV)* (pp. 48–55). http://doi.org/10.1109/CRV.2014.15

Navarro, X., Krueger, T. B., Lago, N., Micera, S., Stieglitz, T., & Dario, P. (2005). A critical review of interfaces with the peripheral nervous system for the control of neuroprostheses and hybrid bionic systems. *Journal of the Peripheral Nervous System*, *10*(3), 229–258. http://doi.org/10.1111/j.1085-9489.2005.10303.x

Ng, A. (2004). Feature selection, L_1 vs. L_2 regularization, and rotational invariance. In *Proceedings of the International Conference on Machine Learning (ICML)*.

Nikolajsen, L., & Jensen, T. S. (2001). Phantom limb pain. *British Journal of Anaesthesia*, *87*(1), 107–116. http://doi.org/10.1093/bja/87.1.107

Open Computer Vision (OpenCV) Library (SourceForge.net). (2015). Retrieved October 14, 2011, from http://sourceforge.net/projects/opencvlibrary/

Open Motion Planning Library (OMPL). (2015, November). Retrieved from http://ompl.kavrakilab.org/

Otto, J. P. (2013, February). DARPA's RP Arms Progress, Nurture Related Wonders. *The O&P Edge*, *12*(2), 30–38.

Pellegrino, G. di, Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental Brain Research*, *91*(1), 176–180. http://doi.org/10.1007/BF00230027

Petit, F., Chalon, M., Friedl, W., Grebenstein, M., Schäffer, A. A., & Hirzinger, G. (2010). Bidirectional antagonistic variable stiffness actuation: Analysis, design & Implementation. In *2010 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 4189–4196). IEEE. http://doi.org/10.1109/ROBOT.2010.5509267

Plaisier, M. A., Bergmann Tiest, W. M., & Kappers, A. M. L. (2009). Salient Features in 3-D Haptic Shape Perception. *Attention, Perception, & Psychophysics*, *71*(2), 421–430.

Ponce Wong, R. D., Hellman, R. B., & Santos, V. J. (2014). Haptic exploration of fingertip-sized geometric features using a multimodal tactile sensor. In D. Popa & M. B. J. Wijesundara (Eds.), *Proc SPIE Defense, Security and Sensing / Sensing Technology and Applications "Sensors for Next-Generation Robotics" Conf* (Vol. 9116, pp. 911605–1 to 911605–15). Baltimore, MD. http://doi.org/10.1117/12.2058238

Ponce Wong, R. D., Hellman, R. B., & Santos, V. J. (2014). Spatial asymmetry in tactile sensor skin deformation aids perception of edge orientation during haptic exploration. *IEEE Trans on Haptics, Special Issue on "Haptics in Rehabilitation and Neural Engineering," 7*(2), 191–202. http://doi.org/10.1109/TOH.2013.56

Pratt, G. A., & Williamson, M. M. (1995). Series elastic actuators. In *Prof IEEE/RSJ Intl Conf on Intelligent Robots and Systems, "Human Robot Interaction and Cooperative Robots"* (Vol. 1, pp. 399–406). http://doi.org/10.1109/IROS.1995.525827

Priganc, V. W., & Stralka, S. W. (2011). Graded Motor Imagery. *Journal of Hand Therapy*, *24*(2), 164–169. http://doi.org/10.1016/j.jht.2010.11.002

Puchhammer, G. (2000). The Tactile Slip Sensor: Integration of a Miniaturized Sensory Device on a Myoelectric Hand. *Orthopadie-Technik Quarterly*, 7–12.

Pylatiuk, C., Kargov, A., & Schulz, S. (2006). Design and Evaluation of a Low-Cost Force Feedback System for Myoelectric Prosthetic Hands. *Journal of Prosthetics and Orthotics*. Retrieved from http://journals.lww.com/jpojournal/Fulltext/2006/04000/Design_and_Evaluation_of_a_Low_Cost_Force_Feedback.7.aspx

Quigley, M., Conley, K., Gerkey, B., Faust, J., Foote, T., Leibs, J., … Ng, A. (2009). ROS: an open-source Robot Operating System. Retrieved from http://www.ros.org/

Rabiner, L. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*(2), 257–286. http://doi.org/10.1109/5.18626

Ramachandran, V. S., & Altschuler, E. L. (2009). The use of visual feedback, in particular mirror visual feedback, in restoring brain function. *Brain*, 1–18. http://doi.org/10.1093/brain/awp135

Ramachandran, V. S., & Hirstein, W. (1998). The perception of phantom limbs. The D. O. Hebb lecture. *Brain*, *121*(9), 1603–1630. http://doi.org/10.1093/brain/121.9.1603

Raspopovic, S., Capogrosso, M., Petrini, F. M., Bonizzato, M., Rigosa, J., Di Pino, G., … Micera, S. (2014). Restoring Natural Sensory Feedback in Real-Time Bidirectional Hand Prostheses. *Science Translational Medicine*, *6*(222), 222ra19–222ra19. http://doi.org/10.1126/scitranslmed.3006820

Rincon-Gonzalez, L., Naufel, S. N., Santos, V. J., & Helms Tillery, S. (2012). Interactions Between Tactile and Proprioceptive Representations in Haptics. *Journal of Motor Behavior*, *44*(6), 391–401.

Rincon-Gonzalez, L., Warren, J. P., Meller, D. M., & Helms Tillery, S. (2011). Haptic Interaction of Touch and Proprioception: Implications for Neuroprosthetics. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *19*(5), 490–500. http://doi.org/10.1109/TNSRE.2011.2166808

Rizzolatti, G., & Craighero, L. (2004). The Mirror-Neuron System. *Annual Review of Neuroscience*, *27*(1), 169–192. http://doi.org/10.1146/annurev.neuro.27.070203.144230

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, *3*(2), 131–141. http://doi.org/10.1016/0926-6410(95)00038-0

Robot Operating System (ROS). (2011). Retrieved October 14, 2011, from http://www.ros.org/wiki/

Rothwell, J. C., Traub, M. M., Day, B. L., Obeso, J. A., Thomas, P. K., & Marsden, C. D. (1982). Manual Motor Performance in a Deafferented Man. *Brain*, *105*(3), 515–542. http://doi.org/10.1093/brain/105.3.515

RSL Steeper Bebionic3 Hand. (2014). Retrieved November 27, 2014, from http://bebionic.com/the_hand

Salzmann, M., Pilet, J., Ilic, S., & Fua, P. (2007). Surface Deformation Models for Nonrigid 3D Shape Recovery. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *29*(8), 1481–1487. http://doi.org/10.1109/TPAMI.2007.1080

Schulman, J., Lee, A., Ho, J., & Abbeel, P. (2013). Tracking deformable objects with point clouds. In *2013 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 1130–1137). http://doi.org/10.1109/ICRA.2013.6630714

Schwoebel, J., Friedman, R., Duda, N., & Coslett, H. B. (2001). Pain and the body schema Evidence for peripheral effects on mental representations of movement. *Brain*, *124*(10), 2098–2104. http://doi.org/10.1093/brain/124.10.2098

Shadow Robot Company. (2011). Shadow Dexterous Hand. Retrieved from http://www.shadowrobot.com/hand/

Shah, A. J., & Shah, J. A. (2016). Towards manipulation planning for multiple interlinked deformable linear objects. In *2016 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 3908–3915). http://doi.org/10.1109/ICRA.2016.7487580

Slivkins, A. (2014). Contextual Bandits with Similarity Information. *Journal of Machine Learning Research*, *15*, 2533–2568.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, *15*, 1929–1958.

Stevens, J. A., & Stoykov, M. E. P. (2003). Using Motor Imagery in the Rehabilitation of Hemiparesis. *Archives of Physical Medicine and Rehabilitation*, *84*(7), 1090–1092. http://doi.org/10.1016/S0003-9993(03)00042-X

Sucan, I. A., & Chitta, S. (2015, October). "MoveIt!" Retrieved from http://moveit.ros.org

Su, Z., Fishel, J. A., Yamamoto, T., & Loeb, G. E. (2012). Use of Tactile Feedback to Control Exploratory Movements to Characterize Object Compliance. *Frontiers in Neurorobotics*, *6*. http://doi.org/10.3389/fnbot.2012.00007

Swanson, A. B., Matev, I. B., & De Groot, G. (1970). The strength of the hand. *Bulletin of Prosthetics Research*, *10*(14), 145–153.

SynTouch BioTac biomimetic tactile sensor. (2007). Retrieved November 27, 2014, from http://www.syntouchllc.com/Products/BioTac/

Tan, D., Schieffer, M. A., Graczyk, E., Keith, M., Anderson, J. R., & Tyler, D. J. (2014). Restoring Sensation in Amputees: Chronic Stability of Implanted Cuff Electrodes. In *Proc Myoelectric Controls/Powered Prosthetics Symp*. Fredericton, New Brunswick, Canada.

Tan, D. W., Schiefer, M. A., Keith, M. W., Anderson, J. R., Tyler, J., & Tyler, D. J. (2014). A neural interface provides long-term stable natural touch perception. *Science Translational Medicine*, *6*(257), 1–11. http://doi.org/10.1126/scitranslmed.3008669

Tan, P. N., Steinbach, M., & Kumar, V. (2006). *Introduction to Data Mining*. Pearson Addison Wesley.

Tastevin, J. (1937). En partant de l'expérience d'Aristote les déplacements artificiels des parties du corps ne sont pas suivis par le sentiment de ces parties ni par les sensations qu'on peut y produire. [Starting from Aristotle's experiment the artificial displacements of parts of the body are not followed by feeling in these parts or by the sensations which can be produced there.]. *L'Encéphale: Revue de Psychiatrie Clinique Biologique et Thérapeutique*, *32*, 57–158.

Tekin, C., & van der Schaar, M. (2015). Distributed online learning via cooperative contextual bandits. *Signal Processing, IEEE Transactions on*, *63*(14), 3700–3714.

Tekin, C., & Van der Schaar, M. (2015). Distributed Online Learning via Cooperative Contextual Bandits. *IEEE Transactions on Signal Processing*, *63*(14), 3700–3714. http://doi.org/10.1109/TSP.2015.2430837

The Hand by Frank R. Wilson. (1998). Retrieved from http://www.penguinrandomhouse.com/books/191866/the-hand-by-frank-wilson/9780679740476

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, *58*(1), 267–288.

Touch Bionics i-limb hand. (2009). Retrieved July 21, 2009, from http://www.touchbionics.com/professionals.php?section=5

Touch Bionics i-limb product range. (2014). Retrieved November 27, 2014, from http://www.touchbionics.com/prosthetists-therapists/product-range

Tubiana, R. (1981). Architecture and functions of the hand. In R. Tubiana (Ed.), *The Hand* (Vol. 1). Philadelphia, PA: WB Saunders.

Uellendahl, J. E., Mandacina, S., & Ramdial, S. (2006). Custom silicone sockets for myoelectric prostheses. *JPO: Journal of Prosthetics and Orthotics*, *18*(2), 35–40.

Velliste, M., Perel, S., Spalding, M. C., Whitford, A. S., & Schwartz, A. B. (2008). Cortical control of a prosthetic arm for self-feeding. *Nature*, *453*(7198), 1098–1101. http://doi.org/10.1038/nature06996

Vincent Systems Bionic Hand. (2014). Retrieved November 27, 2014, from http://vincentsystems.de/en/research/bionic-hand/

Vincent Systems VINCENTevolution 2. (2014). Retrieved November 27, 2014, from http://vincentsystems.de/en/prosthetics/vincent-evolution-2/

Viterbi, A. J. (1967). Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *IEEE Transactions on Information Theory*, *13*(2), 260–269. http://doi.org/10.1109/TIT.1967.1054010

Watkins, C. J. C. H. (1989, May). *Learning From Delayed Rewards*. King's College, Cambridge, UK. Retrieved from http://www.cs.rhul.ac.uk/~chrisw/new_thesis.pdf

Watkins, C. J. C. H., & Dayan, P. (1992). Technical Note: Q-Learning. *Machine Learning*, *8*(3-4), 279–292. http://doi.org/10.1023/A:1022676722315

Wettels, N., Santos, V. J., Johansson, R. S., & Loeb, G. E. (2008). Biomimetic Tactile Sensor Array. *Advanced Robotics*, *22*, 829–849. http://doi.org/10.1163/156855308X314533

Williams III, T. W. (2011). Progress on stabilizing and controlling powered upper-limb prostheses. *The Journal of Rehabilitation Research and Development*, *48*(6), ix. http://doi.org/10.1682/JRRD.2011.04.0078

Witham, C. L., & Baker, S. N. (2011). Effects of Area 3a stimulation on muscle activity in the awake monkey. In *Proc Physiol Soc 23* (p. C70). Oxford, UK.

Wright, F. V. (2006). Measurement of functional outcome with individuals who use upper extremity prosthetic devices: current and future directions. *JPO: Journal of Prosthetics and Orthotics*, *18*(2), 46–56.

Yamamoto, T., Wettels, N., Fishel, J. A., Lin, C.-H., & Loeb, G. E. (2012). BioTac -- Biomimetic Multi-modal Tactile Sensor. *Journal of the Robotics Society of Japan*, *30*(5), 496–498.