

Slip Classification for Dynamic Tactile Array Sensors

Barrett Heyneman¹

Mark R. Cutkosky¹

December 2, 2014

Abstract

The manipulation of objects held in a robotic hand or gripper is accompanied by events such as making and breaking contact and slippage, between the fingertips and the grasped object and between the grasped object and external surfaces. Humans can distinguish among such events, in part, because they excite the various mechanoreceptors in the hands differently. As part of an effort to provide robots with a similar capability, we propose two features that can be extracted from dynamic tactile array data and used to discriminate between hand/object and object/world slips. Both features rely on examining how slippage affects an array of dynamic tactile sensors compared to the way it affects individual elements of the array. In comparison to approaches that require extensive training with particular combinations of objects and skin, the features work for a wide range of frequencies and grasp conditions. The performance and generalizability of the features are verified with testing on three different kinds of sensors and for a range of object textures, grasp forces and slip conditions. Both features demonstrate greater than 85% accuracy in identifying the location of slip.

1 Introduction

Human skill in grasping and manipulation relies on remarkable tactile sensing capabilities. We can quickly identify textures and fine features using our fingertips, and while holding objects we subconsciously control grip forces and prevent slippage. When using tools we act as if they are extensions of

our limbs, which requires not only dexterous manipulation but sophisticated and implicit understanding of how the sensations at our fingertips relate to the interactions between the tool and the environment.

In comparison, robotic tactile sensing is impoverished. A sophisticated hand may have dozens or even a hundred tactile sensors, but not tens of thousands; and the sensors are typically of only one or two types. Practical challenges include wiring and the need to embed sensors in compliant skin materials which experience frequent impacts and deformation. In consequence, the use of tactile information in robotics has proceeded slowly in comparison to vision.

Notwithstanding these challenges, the need for tactile sensing in robotics is steadily increasing. Robots are being pushed to perform increasingly complex tasks, often using human tools in unstructured environments. New developments in tactile sensing are exploiting advances in flexible electronics with shrinking sensors and microprocessors (often driven by the smartphone industry) for local signal processing and communication. With these advances we are seeing robust tactile sensor arrays capable of delivering dynamic data into the kilohertz range (Ascari et al, 2009; Aukes et al, 2014; Jentoft et al, 2013; Lin et al, 2009; Mittendorfer and Cheng, 2011; Schmitz et al, 2011), with recent reviews of the state of the art in (Cutkosky and Ulmen, 2014; Dahiya et al, 2010; Lucarotti et al, 2013; Tiwana et al, 2012; Yousef et al, 2011).

The work presented here is part of an effort to use these improved tactile sensing capabilities to identify and distinguish between different kinds of contact events, particularly those that occur between a hand and a grasped tool or between a grasped tool and the

¹Stanford University, Stanford, CA.

environment. In routine tasks, autonomous robotic systems will encounter both types of events. In identifying these events, we seek algorithms that do not require extensive training, which work robustly for a variety of sensor types, and which are not sensitive to variations in grasp force and object texture.

The following sections start with a brief review of relevant prior work, including studies of human mechanoreception and manipulation event detection. We then present two features that are based on how events such as hand/object and object/world slippage affect arrays of tactile sensors versus individual elements within an array. The corresponding algorithms are suited for real-time implementation on hardware with limited computational power, allowing for distributed computation to identify and classify these events. The features are compared for three different types of sensors and fingers, and a variety of grasp forces and object textures. The results show better than 85% identification of slip type across these widely varying conditions. We conclude with a discussion of extensions that could further improve the event detection and classification approach and with areas for future work.

2 Related Prior Work

The work presented here draws inspiration from human mechanoreception and event detection and builds upon prior work in robotic slip detection and texture identification.

2.1 Human Mechanoreception in Manipulation

An extensive body of research on human tactile sensing has elucidated the roles of fast-acting and slow-acting mechanoreceptors for detecting and distinguishing among various types of events during manipulation (Johansson and Flanagan, 2009). Fast-acting mechanoreceptors are of particular importance for detecting the formation and breaking of contacts and the onset of slip. Meissner corpuscles are superficial fast-acting mechanoreceptors mechanically coupled to the dermal papillae (Fig. 1). They respond

strongly to events in the 10-60 Hz range that deform the papillae, including deformations that occur as fingers press upon an object and “plucking” of the fingerprint ridges as an object starts to slide. Sensing these events is important to human grasp force modulation (Johansson and Flanagan, 2009; Johansson and Westling, 1984; Witney et al, 2004), because they indicate *incipient slip* which occurs before gross slip between the hand and a held object.

In contrast, when a held object vibrates, the many individual Meissner corpuscles are not significantly excited. As shown in (Hunt, 1961), the Pacinian corpuscles instead register significant neuronal activity. More specific research by (Brisben et al, 1999) showed that human perception of held object vibrations follows the response curves from individual Pacinian corpuscles, indicating that these receptors are responsible for our perception of transmitted vibration. Considering that the receptive fields of the Pacinian corpuscle are large (Fig. 1) it is not surprising that they are primarily responsible for sensing unlocalized vibrations in the 50-300 Hz range.

The coding of tactile signals involves more than the response spike-rate of certain mechanoreceptors. Various studies have indicated that relative timing, phasing, and even coherence of the responses of different mechanoreceptors may play a pivotal role in conveying tactile information. In humans, Johansson and Birznieks (2004) demonstrated that the relative timings of the first intra- and inter-mechanoreceptor neural spikes at the onset of stimulation are sufficient to predict tangential force direction as a fingertip comes into contact with an object. Other studies of mechanoreceptors in cats’ paws (Greenstein et al, 1987) and the halteres of flying insects (Fox and Daniel, 2008; Fox et al, 2010) indicate that relative phasing within populations of neurons carries important information. Collectively, these results suggest avenues for robotic sensing and perception which do not rely simply on the amplitude of a given input signal but rather on more complicated, time-dependent measures.

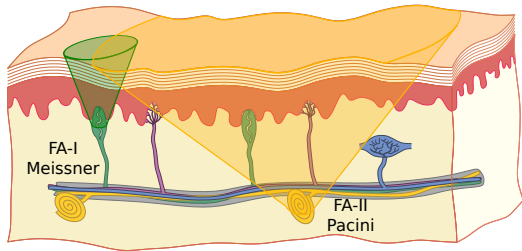


Figure 1: The depth of human mechanoreceptors and their mechanical coupling to surrounding tissue gives the fast-acting type I and II receptors very different receptive fields. Receptive fields for Meissner corpuscles and Pacinian corpuscles are shown here, with the area scaled to reflect actual receptive field measurements from (Knibestöl, 1973). *Image adapted from (Cutkosky and Ulmen, 2014).*

2.2 Robotic Slip Detection

Slip detection is one of the earliest tactile sensing problems discussed in robotics. Many of the demonstrated solutions have focused on detecting the onset of slippage, or even incipient slippage. Some of the earliest work (Howe and Cutkosky, 1989) utilized accelerometers embedded in the skin to measure the small vibrations that accompany the onset of sliding between an object and a finger. Embedded dynamic sensors, such as polyvinylidene-fluoride stress rate sensors, have continued to be a popular option with a variety of working implementations (Fujimoto et al, 2003; Howe and Cutkosky, 1993; Yahud et al, 2010; ?). These approaches directly measure vibrations in the skin or finger medium that are produced by the stick-slip phenomena occurring at the surface.

Other investigators have used less direct measurements of these micro-vibrations, for example examining the variation of the center of pressure with respect to the resultant forces to determine the onset of sliding (Gunji et al, 2008; Ho and Hirai, 2014; Melchiorri, 2000). Still others have applied computer vision techniques to the output of tactile sensing arrays, treating them as grayscale images, in order to track tactile features moving across the array (Ascari et al, 2009; ?).

In addition, a wide variety of processing techniques

have been used to characterize and identify the dynamic force variations associated with slip, including high-pass filtering, fast fourier transforms, extraction of predefined features, and artificial neural networks (Holweg et al, 1996; Jamali and Sammut, 2012; Romano et al, 2011; Teshigawara et al, 2011; Zhang and Liu, 2012; ?). In most cases these require training against a specific set of objects and grasp conditions, or assumptions about the frequencies of vibrations that occur as specific features of the robotic skin slide on an object. However, there has been little investigation of strategies to distinguish among slips at different interfaces (e.g. finger/object versus object/world) for a variety of grasping conditions.

2.3 Robotic Texture Recognition

Like slip detection, texture identification has been a common application of tactile sensors in robotics. For the present case, the interest is primarily to consider how the skin and tactile sensors are affected by a textured object so that slip detection can be made relatively insensitive to object texture.

With a few exceptions (e.g. (Johnson and Adelson, 2009; Kim et al, 2005a; Maheshwari and Saraf, 2006)) most texture sensing uses dynamic signals generated as a skin with sensors passes over the surface of an object. Indeed, the same sensors used for detecting slip can generally be used for texture identification. Approaches to texture discrimination vary widely, from modeling of the sensor/skin response (Howe and Cutkosky, 1993; Mazid and Russell, 2006), to utilizing statistical or machine learning techniques on features extracted from the time- or frequency-space representations of the sensor signals (Decherchi et al, 2011; Hosoda et al, 2006; Jamali and Sammut, 2011; Kim et al, 2005b; Lepora et al, 2010; Sinapov et al, 2011). In a few cases, researchers have identified signal features that are relatively insensitive to other aspects of the surface interaction (e.g. sliding speed or pressure). For example, Oddo et al (2011) used relative phasing extracted from signal cross-spectra as a speed-invariant feature for texture classification.

3 Classification Features

As noted in Section 1, there are increasing examples of tactile array sensors that, like fast-acting mechanoreceptors, can respond to dynamic changes in surface stresses, with bandwidths of 50 Hz or greater. In this section we present two methods to use signals from such arrays to distinguish between slips at two distinct interfaces: between the sensor and a held object or between the held object and the environment. The features are intended to be relatively insensitive to variations in grasp force and object characteristics such as texture.

We start with some definitions to frame the problem and limit its scope. The tactile array sensors are modeled as N elements, or *taxels*, which are embedded in the grasping surfaces of a robotic hand or gripper. Often, the array elements are sampled by microprocessors located within or near the array and communicating over a digital line. The n th sample from the discretized array signal is given as

$$\mathbf{p}_n = [p_1(nT + \delta_1), \dots, p_N(nT + \delta_N)] \quad (1)$$

where p_i sample from the i th sensor. The array of taxels is sampled at period T , while individual taxels may be sampled with an individual constant delay δ_i . Depending on the architecture, they may be sampled sequentially, where each element is read one-by-one, or simultaneously, where all elements are read at essentially the same time:

$$\delta_i = \begin{cases} \frac{T(i-1)}{N} & \text{sequential sampling} \\ 0 & \forall i, \text{ simultaneous sampling} \end{cases} \quad (2)$$

Of the N taxels of the array, there is an identifiable subset, S_{contact} , which are in contact with the held object. For this work we assume that the sensor's surface texture does not vary significantly over this contact patch and that slip, if it occurs, results in similar relative velocity across the entire patch. In our experiments this is true of the entire contact patch and we therefore utilize all the taxels physically beneath it. In more complex situations this might not be the case; however this initial work does not explore how to incorporate knowledge concerning areas with significantly varying contact conditions.

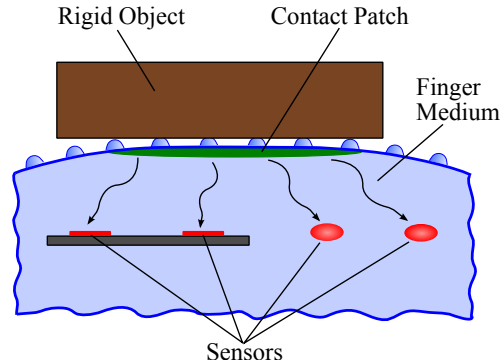


Figure 2: A cross section view of a contact patch, including M individual contact points, and a set of nearby sensors. The skin and finger medium form a transduction path to the sensors, represented by $\mathbf{F}(t)$. The sensors will pick up vibrations due to both object vibrations and slip between it and the finger.

The taxels are assumed to be embedded in a linearly viscoelastic material forming the gripper skin or body. It can be modeled as a linear, time-invariant system transforming surface tractions (normal stress, $\sigma_{zz}(t)$, and the two shear stresses, $\tau_{zx}(t)$ and $\tau_{zy}(t)$) to sensor outputs, similar to the model used in the soft finger analysis by (Fearing and Hollerbach, 1985).

Most tactile sensors designed for dynamic sensing applications have a textured surface. Textures have been shown to produce greater high-frequency signals and more consistent friction, and they have been investigated by a variety of researchers including (Fearing and Hollerbach, 1985; Howe and Cutkosky, 1993; Oddo et al, 2011). We assume the grasping surfaces of the hand make contact with a held object in a finite number of locations due to the surface features and that surface traction inputs occur only at these M locations. Therefore, the finger/sensor system maps the traction input vector $\boldsymbol{\sigma}(t) \in \mathbb{R}^{3M}$ to the sensor signal via the $N \times 3M$ linear system $\mathbf{F}(t)$:

$$\mathbf{p}(t) = (\mathbf{F} * \boldsymbol{\sigma})(t) \quad (3)$$

A schematic of an example sensor in contact with a rigid object is shown in Figure 2 to illustrate these assumptions.

Both features are extracted from the signals in fre-

quency space by computing spectral quantities over a given frequency band. While there are many techniques for spectral estimation from discretely sampled signals, in this work we will rely on the Discrete Fourier Transform, implemented via the Fast Fourier Transform algorithm, and Welch’s algorithm to improve estimation accuracy at the expense of frequency resolution.

The frequency space representation of a signal will be denoted with non-script text; e.g. the time signal $x(t)$ becomes the spectral coefficients $x(k)$; the time or frequency component arguments t and k are dropped where unambiguous. These spectral coefficients are complex quantities and $*$ is used to represent the complex-conjugate, as in x^* . The power spectral density, generally a matrix quantity, is denoted $\mathcal{P}_{x,y}$, where the second subscript is omitted when they are both the same.

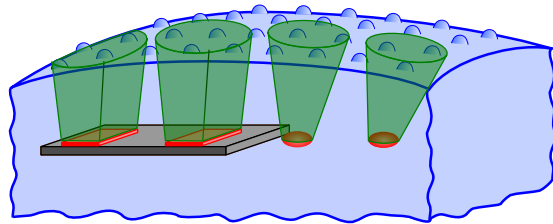
3.1 Signal Power Ratio

As noted in Section 2.1, humans use different fast-acting mechanoreceptors with small and large receptive fields to characterize and distinguish among different contact events. In a similar way, the signals from individual taxels of an array can be used to measure local vibrations while the sum of the taxel signals measures vibrations affecting the entire array as an ensemble (Figure 3).

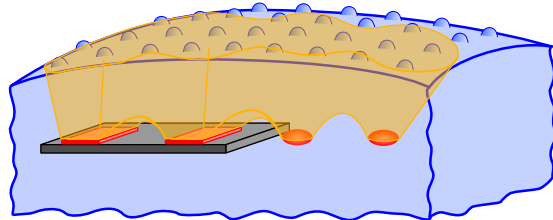
We hypothesize that classification can be made based on the ratio of ensemble to local signal power. When slip occurs between the object and the environment we expect to see comparatively more signal power in the ensemble signal than the local signal, just as Pacinian corpuscles are more excited by such events in humans. In contrast, when slip occurs between the hand and the object we expect to see comparatively more signal power in the local signal than the ensemble signal, analogous to the increased activity seen in Meissner corpuscles.

3.1.1 Definition

We begin by developing a measure of the local effect of vibrations on the taxel array. The spectral coefficients, $p_i(k)$, are estimated with a resolution of f_0



(a) Receptive fields for each individual sensor



(b) Receptive field for the sensors as an ensemble

Figure 3: The data from tactile array sensors can be processed to provide local and distributed receptive fields as found in human biology. In (a) each taxel is sampled and processed individually, producing smaller localized responses to stimuli at the surface. When the taxels are sampled together, as if they were one larger sensor, the response to surface stimuli spreads out, as shown in (b).

for each discrete signal. The power contained in a frequency band centered at f with width $2w$ is calculated for each taxel, and then summed across all taxels in contact with the held object

$$L(f, w) \triangleq \sum_{i \in S_{contact}} \left(\sum_{k=(f-w)/f_0}^{(f+w)/f_0} |p_i(k)|^2 \right) \quad (4)$$

By summing power rather than the time-domain signals $p_i(t)$ (or spectral coefficients) directly there is no cancellation due to differences in phasing. The signal at each taxel contributes to $L(f, w)$ in the same way, independent of the signals seen at any other taxel. For this reason the measure more directly reflects how strongly the vibrations in the given frequency range are affecting each of the sensor elements locally.

In order to compute a similar indicator for the effect of vibrations on the array as an ensemble, the

signals from all the taxels in contact with the object are combined in order to mimic a single *virtual taxel* that covers the entire grasping surface and therefore has a large receptive field. The power in this virtual sensor signal is computed for the same frequency range as in Equation (4)

$$E(f, w) \triangleq \sum_{k=(f-w)/f_0}^{(f+w)/f_0} \left| \sum_{i \in S_{contact}} \mathbf{p}_i(k) \right|^2 \quad (5)$$

Here the formation of the single virtual taxel is performed by summing the spectral coefficients, which is equivalent to summing the individual signals in the time domain. For some sensors, a more specific method may be appropriate. For example, an approach for the BioTac[®] sensor presented in (Wettels and Loeb, 2011) uses an *artificial neural network* to extract the point-of-application and vector of applied force from the array of signals. With any method it will be possible for the signals from different taxels to cancel one another at some frequencies while adding constructively at others.

Finally, this biologically inspired feature is defined as the ratio of the ensemble power measure to the local power measure:

$$\Gamma(f, w) \triangleq \frac{1}{N_c} \frac{E(f, w)}{L(f, w)} \quad (6)$$

where N_c is the number of taxels in contact and normalizes $\Gamma(f, w)$ such that it always lies in the range $[0, 1]$.

3.1.2 Sampling Delays

In order to compute the ensemble power $E(f, w)$ some additional steps must be taken. Because the intention is to mimic a virtual sensor covering the grasping surface, the relative timing of the taxel samples, given by δ_i , must be taken into account. If the array is simultaneously sampled the individual signals can simply be averaged as described above. For example, if all sensors are experiencing the same input signal, the relative phasing of the sequentially sampled signals will vary, causing cancellation in Equation (5). However, if the same input were to be applied and

the sensors simultaneously sampled, there would be no such cancellation.

If the samples p_i of the *i*th taxel have been delayed by δ_i relative to the ideal simultaneous sample time, a *resampled* signal \hat{p}_i must be formed which estimates the discrete sequence that would have been generated by sampling the taxel with no delay. Because a simple summation is used in Equation (5) to combine the individual signals, the pure delay can be corrected by applying a time *advance* to the original sample signal directly in the frequency domain. With this step added, Equation (5) becomes

$$E(f, w) = \sum_{k=(f-w)/f_0}^{(f+w)/f_0} \left| \sum_{i \in S_{contact}} \mathbf{p}_i(k) e^{2\pi j k \delta_i f_0} \right|^2 \quad (7)$$

where f_0 is the frequency resolution of the discrete spectral estimate. We have used the fact that a pure time delay manifests in frequency space as a pure phase delay; the complex exponential term above. This formula can be applied generally. In the simultaneous sampling case when $\delta_i = 0$ the exponential term from the pure delay becomes unity and may be dropped.

This method can also be used to provide the resampled time series in cases where a more complicated procedure is necessary to combine individual taxels into an ensemble, e.g. the artificial neural network process mentioned above. The original signals can be transformed to the frequency domain, the time advance applied, and then transformed back to the time domain.

3.2 Coherence

The second candidate classification feature is motivated by considering the source of vibrations and how those vibrations are transmitted to the individual taxels. Consider the case in which the finger is in contact with a rigid object, as shown in Figure 2. The surface tractions, $\boldsymbol{\sigma}(t)$, are now due to the combination of the motion of the held object and any slip that occurs between the object and the finger. We assume that over the sample duration, the ratios of the components of linear and angular velocity

of the object with respect to the contact patch on the finger remain constant. Therefore, the velocities can be described with a single time sequence, $v(t)$. In addition, the stick-slip phenomena at each contact location are assumed to generate surface tractions with the same constraint: the three components are not independent and can be characterized by a single time sequence, $s_i(t)$. Under these assumptions the surface tractions can be defined as

$$\boldsymbol{\sigma}(t) = (\mathbf{B} * (\mathbf{b}v))(t) + \mathbf{T}s(t) \quad (8)$$

where $\mathbf{b} \in \mathbb{R}^6$ describes how the single degree of freedom of the body's vibration is described relative to the finger, $\mathbf{B}(t) \in \mathbb{R}^{3M \times 6}$ represents how that motion creates surface tractions at each contact location, and $\mathbf{T} \in \mathbb{R}^{3M \times M}$ maps the single degree of freedom of each stick-slip phenomenon to the produced surface tractions. Note that \mathbf{T} is block diagonal, reflecting the fact that the stick slip at each contact location creates surface tractions at that location. An expanded system block diagram with these relationships is shown in Figure 4. Combining this model with Equation (3) and transforming to frequency space we have

$$\mathbf{p} = \mathbf{F}(\mathbf{b}v + \mathbf{T}s) \quad (9)$$

When the object vibrates, there is a common source of vibrations being transmitted through the finger medium to each sensor. Assuming those transmissions paths can be modeled as linear, the signals at each sensor element should be coherent with one another.

However, when the object slips against the finger the vibrations at each contact region are due to the local stick-slip behavior of the surface features. If these stick-slip events are *independent* they will not have consistent phase relationships. As indicated in Section 3, we assume that the contact conditions do not vary significantly across the contact patch, and for these reasons we can assume each stick-slip event has a similar power spectrum. Therefore the following holds for the power spectrum of the set of all signals:

$$\mathcal{P}_s(k) = \mathbf{c}_k \mathbf{I} \quad (10)$$

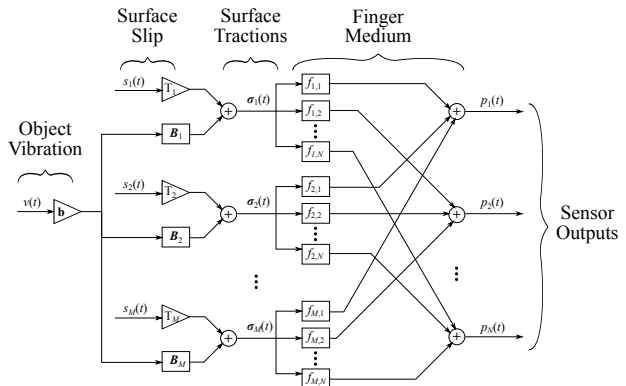


Figure 4: A block diagram representation of the object/gripper/sensor model in Figure 2. Object vibration is transmitted to the system at each of the M contact points where the individual signals due to slip between the object and the gripper are also injected.

for some constants $\mathbf{c}_k \in \mathbb{C}$, and the identity matrix \mathbf{I} .

For purposes of this initial development we will additionally assume that the sensor is designed to minimize crosstalk between the taxels or that the surface tractions at each contact location are sensed by a single taxel. This implies that for an appropriate ordering of contact locations, \mathbf{F} also has a block diagonal structure and many of the $f_{i,j}$ blocks in Figure 4 are identically zero.

3.2.1 Applying Signal Coherence: Two Sensors

The complex coherence between two signals, denoted $\mathcal{C}_{x,y}$, is defined as the normalized cross-spectral density of those two signals with respect to the spectral density of the individual signals, as in Equation (11). The mean square or magnitude squared coherence, $\mathcal{C}_{x,y}^2$ is also often used. As the name implies, it is the magnitude of the complex coherence squared, and due to the relationships between cross-spectral and auto-spectral density it lies in the range $[0, 1]$

$$\mathcal{C}_{x,y} \triangleq \frac{\mathcal{P}_{x,y}}{\sqrt{\mathcal{P}_x \mathcal{P}_y}} \quad (11)$$

There are two physical interpretations of complex and mean square coherence that are relevant to this work. Firstly, $\mathcal{C}_{x,y}^2$ measures the consistency of the relative phase between the two signals at each frequency; when the relative phase is constant $\mathcal{C}_{x,y}^2 = 1$, and as it varies $\mathcal{C}_{x,y}^2 \rightarrow 0$. Secondly, the complex coherence values are related to the linear filter coefficients which minimize the power density of the estimated signal when used to estimate one signal from the other. Additionally, the mean square coherence measures the fraction of the original signal power that is present in the estimated signal using this filter; if it is unity the signals can be perfectly reconstructed from one another. A more thorough treatment of coherence and its interpretations can be found in most spectral analysis texts, e.g. Marple (1987).

If we consider the filtering interpretation of coherence, we can demonstrate how to apply it to the slip interface classification problem using a simple two sensor example ($N = 2$). In the case where the object slips against the environment and not against the finger the surface tractions are due only to the vibration of the object. Each taxel signal is a linear filtered version of this underlying signal, and therefore of each other, and we expect them to be completely coherent. Equation (9) reduces to

$$\mathbf{p} = \mathbf{F}\mathbf{b}\mathbf{v} \quad (12)$$

and the power spectrum of \mathbf{p} computed as

$$\mathcal{P}_{\mathbf{p}} = (\mathbf{F}\mathbf{b})^* \mathcal{P}_v (\mathbf{F}\mathbf{b})^T \quad (13)$$

Because $\mathcal{P}_v \in \mathbb{C}$ we know that $\text{rank}(\mathcal{P}_{\mathbf{p}}) = 1$. When it is normalized per Equation (11) the mean square coherence matrix is

$$\mathcal{C}_{\mathbf{p}}^2 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad (14)$$

This indicates that the two signals are perfectly coherent when the surface tractions are due only to vibration of the held object.

In the second case we assume the object only slips against the finger, and the surface tractions are due only to the interaction between the finger and object. In this case Equation (9) reduces to

$$\mathbf{p} = \mathbf{F}\mathbf{T}\mathbf{s} \quad (15)$$

with the power spectrum of \mathbf{p} given by

$$\begin{aligned} \mathcal{P}_{\mathbf{p}} &= \mathbf{F}^* \mathbf{T}^* \mathcal{P}_s \mathbf{T}^T \mathbf{F}^T \\ &= \mathbf{c} \mathbf{F}^* \mathbf{T}^* \mathbf{T}^T \mathbf{F}^T \end{aligned}$$

using the fact that the signals are *identical* and *incoherent* (Equation (10)). The *no-crosstalk* sensor assumption, and the corresponding block-diagonal structure of \mathbf{F} and \mathbf{T} , can be used to show that the off-diagonal terms of $\mathcal{P}_{\mathbf{p}}$ are identically zero. The mean square coherence in this case becomes

$$\mathcal{C}_{\mathbf{p}}^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (16)$$

When we compare the results for the two cases we see that the sensor signals in the object/world slip case are completely coherent while the signals in the object/hand slip case are completely incoherent. In a real physical system this contrast is unlikely to be as stark for a variety of reasons. First, it is possible that there will be aspects of both types of slip at any given time, even in cases where only one is dominant. It is also likely that the surface tractions at the M locations due to object/hand slip will not be completely incoherent, or will have slightly different power spectra due to local differences in texture, curvature, contact pressure, etc.

In practice, few tactile array sensors achieve zero crosstalk, largely due to physical coupling in the finger medium. However, any practical array sensor array must limit the crosstalk or coupling between taxels in order to generate useful information across the array. As we relax the no-crosstalk assumption, there will be an increase in the coherence of signals from the taxels during object/hand slip, due to the fact that each is partially sensing the vibrations occurring in the other's active area.

Nonetheless, despite anticipated violations of our assumptions in practice, we expect to see significantly higher coherence values between two sensors when they experience object/world slip as compared to object/hand slip. As seen in Section 4, this expectation is borne out in the empirical results.

3.2.2 Extending Coherence

The example above used only two taxels and the coherence between them. In practice an array of N taxels will generate an $N \times N$ complex coherence matrix. In order to simplify the feature space used in classification we extend the concept of mean square coherence to cover groups of N signals. Motivated by the problem at hand, this will be accomplished by extending the optimal filtering interpretation to define the *group square coherence*, or GSC. The resulting definition is similar to that found in Ramírez et al (2008). However in this case there is a specific physical interpretation of the derivation which provides insight into how and why it is appropriate for this application.

The GSC, $\mathcal{G}_{\mathbf{x}}^2(\omega)$, measures the maximum normalized fraction of signal power across the group that can be estimated using a single underlying source signal and a set of optimal linear filters. In this section we show that the GSC thus defined can be efficiently calculated via an eigenvalue analysis of the complex coherence matrix.

We begin the derivation by calculating the normalized total power in the error between the measured signals $\mathbf{x}(t)$ and the signals produced by filtering a single input signal $u(t)$ through a filter bank $\mathbf{h}(t)$. The error vector $\mathbf{e}(t)$ is given by

$$\mathbf{e}(t) = \mathbf{x}(t) - (\mathbf{h} * u)(t) \quad (17)$$

The power in each error signal is normalized by the power of the original signal in order to prevent signals with higher power from dominating the optimization. If the errors are not normalized the optimal filters and source signal, \mathbf{h}_{opt} and u_{opt} respectively, will skew towards recreating the signals with more power more accurately, even if the majority of the signals are coherent but much lower power

$$\bar{\mathcal{P}}_{\mathbf{e}} = \sum_{i=1}^N \frac{\mathcal{P}_{e_i}}{\mathcal{P}_{x_i}} \quad (18)$$

$$= \text{Tr} \left(\mathcal{D}_{\mathbf{x}}^{-1/2} \mathcal{P}_{\mathbf{e}} \mathcal{D}_{\mathbf{x}}^{-1/2} \right) \quad (19)$$

where $\mathcal{D}_{\mathbf{x}}$ is a matrix consisting of only the diagonal entries of $\mathcal{P}_{\mathbf{x}}$. We can then expand $\mathcal{P}_{\mathbf{e}}$ by using the

definition in Equation (17) to get

$$\mathcal{P}_{\mathbf{e}} = \mathcal{P}_{\mathbf{x}} + \mathbf{h}^* \mathcal{P}_u \mathbf{h}^T - \mathbf{h}^* \mathcal{P}_{u\mathbf{x}} - \mathcal{P}_{\mathbf{x}u} \mathbf{h}^T. \quad (20)$$

Finally, $\bar{\mathcal{P}}_{\mathbf{e}}$ is minimized with respect to the spectral coefficients of the optimal filter bank, \mathbf{h} , by taking partial derivatives and setting equal to zero:

$$\begin{aligned} \frac{\partial}{\partial \mathbf{h}} \bar{\mathcal{P}}_{\mathbf{e}} &= \frac{\partial}{\partial \mathbf{h}} \text{Tr} \left(\mathcal{D}_{\mathbf{x}}^{-1/2} \mathcal{P}_{\mathbf{e}} \mathcal{D}_{\mathbf{x}}^{-1/2} \right) \\ &= \frac{\partial}{\partial \mathbf{h}} \text{Tr} \left(\mathcal{D}_{\mathbf{x}}^{-1} \mathcal{P}_{\mathbf{e}} \right) \\ &= \frac{\partial}{\partial \mathbf{h}} \mathcal{D}_{\mathbf{x}}^{-1} \mathcal{P}_{\mathbf{e}} \\ &= \mathcal{D}_{\mathbf{x}}^{-1} \frac{\partial}{\partial \mathbf{h}} \mathcal{P}_{\mathbf{e}} \\ &= 0 \end{aligned}$$

where we have used the properties that $\text{Tr}(\mathbf{A}\mathbf{B}) = \text{Tr}(\mathbf{B}\mathbf{A})$ and $\frac{\partial}{\partial x} \text{Tr}(\mathbf{A}) = \frac{\partial}{\partial x} \mathbf{A}$. Substituting $\frac{\partial}{\partial \mathbf{h}} \mathcal{P}_{\mathbf{e}}$ computed from Equation (20) and then solving for \mathbf{h} gives the optimal set of filters to minimize the sum of normalized error powers.

$$\mathbf{h}_{opt} = \frac{\mathcal{P}_{u\mathbf{x}}^T}{\mathcal{P}_u} = \frac{\mathcal{P}_{\mathbf{x}u}^*}{\mathcal{P}_u} \quad (21)$$

$$\min_{\mathbf{h}} \bar{\mathcal{P}}_{\mathbf{e}} = N - \frac{\mathcal{P}_{u\mathbf{x}} \mathcal{P}_{\mathbf{x}u}}{\mathcal{P}_u} \quad (22)$$

In order to minimize $\bar{\mathcal{P}}_{\mathbf{e}}$ with respect to u we first note that the space of all square-integrable functions, L^2 , can be represented by a *Hilbert space*, $\mathbb{H}_{\mathcal{P}}$ with the cross spectral density between two signals as the inner product.

We can define the space $\mathbb{X} \subset \mathbb{H}_{\mathcal{P}}$ as the subspace spanned by the individual signals, i.e. the rows of $\mathbf{x}(t)$. Any source signal $u(t)$ can be decomposed into a linear combination of the signals which span \mathbb{X} , $\mathbf{a}^T \mathbf{x}(t)$, and an additional signal which is orthogonal to the subspace, $u_{\perp}(t)$, such that

$$u(t) = \mathbf{a}^T \mathbf{x}(t) + u_{\perp}(t). \quad (23)$$

Using this decomposition into orthogonal subspaces one can compute the power spectral density matrices in Equation (22):

$$\mathcal{P}_u = \mathbf{a}^T \mathcal{P}_x \mathbf{a} + \mathcal{P}_{u_\perp} \quad (24)$$

$$\mathcal{P}_{u,x} = \mathbf{a}^T \mathcal{P}_x \quad (25)$$

$$\mathcal{P}_{x,u} = \mathcal{P}_x \mathbf{a} \quad (26)$$

Combining these results with Equation (22) results in

$$\min_{\mathbf{h}} \bar{\mathcal{P}}_e = N - \frac{\mathbf{a}^T \mathcal{P}_x \mathcal{D}_x^{-1} \mathcal{P}_x \mathbf{a}}{\mathbf{a}^T \mathcal{P}_x \mathbf{a} + \mathcal{P}_{u_\perp}} \quad (27)$$

from which we can see that $\bar{\mathcal{P}}_e$ is minimized if and only if $\mathcal{P}_{u_\perp} = 0$

Finally, using Equation (11) yields a simplified expression for $\bar{\mathcal{P}}_e$:

$$\min_{\mathbf{h}} \bar{\mathcal{P}}_e = N - \left(\mathcal{D}_x^{1/2} \mathbf{a} \right)^T \mathcal{C}_x \mathcal{C}_x \left(\mathcal{D}_x^{1/2} \mathbf{a} \right) \quad (28)$$

where \mathcal{D}_x is a matrix with only the diagonal entries of \mathcal{P}_x .

Equation (23) indicates that minimizing over the source signal $u(t)$ is equivalent to minimizing over the coefficients \mathbf{a} . This is an eigenvalue decomposition problem which is minimized when $\mathcal{D}_x^{1/2} \mathbf{a} = kv_1$, where v_1 is the first eigenvector of \mathcal{C}_x associated with the largest eigenvalue, λ_1 . The minimized value can be expressed as

$$\min_{\mathbf{h}, u} \bar{\mathcal{P}}_e = N - \lambda_1^2 k^2. \quad (29)$$

Because $u(t)$ and $\mathbf{h}(t)$ are convolved in Equation (17) the optimal solution is only unique up to a scaling factor between the two, i.e. if $u_{opt}(t)$ and $\mathbf{h}_{opt}(t)$ minimize $\bar{\mathcal{P}}_e$, so do $2u_{opt}(t)$ and $1/2\mathbf{h}_{opt}(t)$. To find a unique solution by fixing k , we restrict $\mathcal{P}_u = 1$ at all frequencies:

$$\begin{aligned} \mathcal{P}_u &= \mathbf{a}^T \mathcal{P}_x \mathbf{a} \\ &= \left(\mathcal{D}_x^{1/2} \mathbf{a} \right)^T \mathcal{C}_x \left(\mathcal{D}_x^{1/2} \mathbf{a} \right) \\ &= \lambda_1 k^2 \\ &= 1 \end{aligned}$$

Applying this final scaling factor yields the results

of the optimization.

$$\min_{\mathbf{h}, u} \bar{\mathcal{P}}_e = N - \lambda_1 \quad (30)$$

$$\mathbf{h}_{opt} = \mathcal{P}_x \mathbf{a}_{opt} \quad (31)$$

$$\mathcal{C}_x \mathcal{D}_x^{1/2} \mathbf{a}_{opt} = \lambda_1 (\mathcal{C}_x) \mathcal{D}_x^{1/2} \mathbf{a}_{opt} \quad (32)$$

Equipped with the optimal source signal and filters we can define \mathcal{G}_x^2 and then simplify it. Let the group square coherence be defined as the average of the normalized power in each estimated signal.

$$\mathcal{G}_x^2(\omega) \triangleq \left(\frac{\left(\sum_{i=1}^N \frac{\mathcal{P}_{x_i}(\omega)}{\mathcal{P}_{x_i}(\omega)} \right) - 1}{N - 1} \right)^2 \quad (33)$$

We normalize to $N - 1$, including the -1 term in the numerator, instead of N to reflect the fact that the underlying signal $u(t)$ can, at worst, perfectly estimate one of our measured signals $x_i(t)$.

Following similar simplifications steps as for $\bar{\mathcal{P}}_e$ above and substituting the optimal filters \mathbf{h}_{opt} and signal u_{opt} , this definition can be reduced to

$$\begin{aligned} \mathcal{G}_x^2(\omega) &= \left(\frac{\mathbf{h}_{opt}^T \mathcal{D}_x^{-1} \mathbf{h}_{opt}^* - 1}{N - 1} \right)^2 \\ &= \left(\frac{\lambda_1 - 1}{N - 1} \right)^2 \end{aligned} \quad (34)$$

In Appendix A we demonstrate the following useful properties that stem from this definition: that $\mathcal{G}_x^2(\omega) = \mathcal{C}_x^2(\omega)$ for $N = 2$ and that the range and extrema of $\mathcal{G}_x^2(\omega)$ occur when all pairwise $\mathcal{C}_{x_i, x_j}^2(\omega)$ values are at the same extremes (0 or 1).

4 Slip Experiments

Three sets of experiments, each utilizing a different setup and tactile array sensor technology, were performed in order to generate examples of slip occurring at both interfaces. While elements of the interaction were held constant across these experiments for comparison purposes, a breadth of conditions were tested to evaluate the ability to work with a range of textures from smooth to rough and at low and high grasp

forces and high and low sliding speeds typically used in manipulating objects.

In all experiments, the *held object* consisted of a specific texture sample taken from the set described in Table 1 and Figure 5. The set includes smooth surfaces, various grits of sandpaper, regular patterns of holes or slats, and even a silicone mold of the skin on the outer surface of the capacitive sensor (Section 4.2). This last texture was chosen specifically to present a difficult classification problem, especially for the capacitive sensor, in that the same surface is involved in both object-world and object-hand slip.

In this section we describe the sensors and the experiments performed. Any sensor specific considerations taken when extracting the features described in the previous section are discussed with the results in Section 5.2



Figure 5: The 8 textures plates used in experiments (described in Table 1).

4.1 BioTacsTM and MC Hand

The first set of experiments utilize the Syntouch Bio-Tac¹ sensor mounted on a manually controlled MotionControl MC² prosthetic hand.

The BioTac sensor is based on the the work of Wettels and Loeb (Wettels, 2012) and consists of a rigid central core surrounded by an elastic, fluid-filled skin with fingerprint ridges like those of the human fin-

¹www.syntouchllc.com

²<http://www.utaharm.com>

Table 1: Textures used in controlled manipulation experiments (labels match samples in Figure 5)

	description
A	smooth high-density fiberboard (HDF)
B	HDF with 1.59mm holes on 6.35mm grid
C	HDF with 3.18mm holes on 6.35mm grid
D	sensor skin (textured silicone rubber)
E	HDF with 6.35mm wide rectangular slats
F	60 grit sandpaper
G	150 grit sandpaper
H	400 grit sandpaper

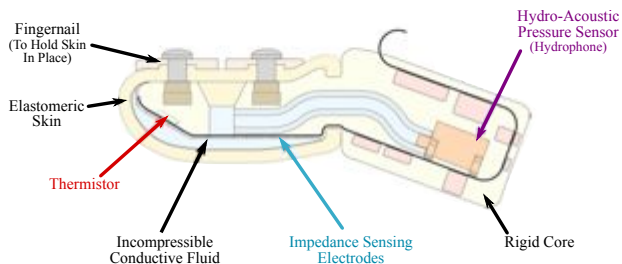


Figure 6: Crosssection of a BioTac[®] sensor demonstrates its construction and the location of various sensing elements. The pressure sensor and impedance electrodes are useful in slip discrimination, mimicking the FA-II and FA-I human mechanoreceptors, respectively. *Reproduced with permission of SynTouch LLC*

gertip (Figure 6). The sensor includes thermal, vibration and skin deformation sensing capabilities, of which the latter two are useful for slip and texture classification.

Skin deformation is sensed via an array of 19 passive sensing electrodes distributed across the surface of the inner core of the sensor. These electrodes measure the impedance through the fluid to 4 active excitation electrodes at 100Hz. In addition, as the finger rubs across surfaces, vibrations are detected by a hydro-acoustic pressure transducer with a filtered bandwidth of 1040Hz. The signals from this dynamic sensor, roughly akin to the human FA-II mechanoreceptors, have been used to characterize surface textures (Fishel and Loeb, 2012).

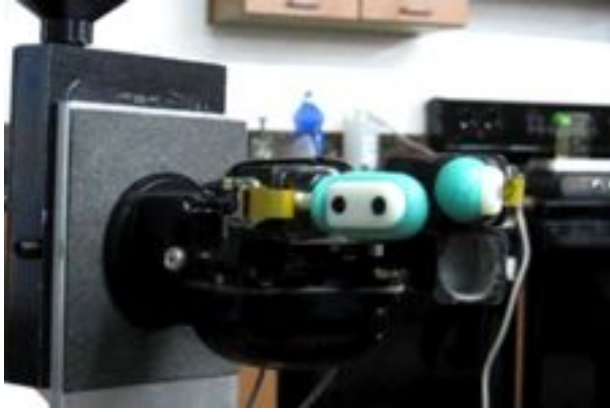


Figure 7: (Experiments with BioTac[®]sensors used a passively loaded prosthetic MC Hand[®] equipped with two opposing sensors.

A single embedded microcontroller performs sequential analog-to-digital conversions of the signals described above, in addition to temperature information and low frequency pressure measurements. The different measurements are interleaved, such that every high frequency pressure reading is followed by a different electrode, low-frequency pressure, or temperature measurement. This results in a 4400Hz sample rate for the pressure sensor and a 100Hz sample rate for each electrode. Up to three BioTacs can be synchronized such that each pressure or electrode measurement is simultaneous *interfinger*, while the electrodes are sampled sequentially *intrafinger*.

A MotionControl MC prosthetic hand was instrumented with two BioTacs in the thumb and forefinger positions as in Figure 7. The hand, statically mounted on a stand, was closed upon objects passively with a grip force of $\sim 2.5\text{N}$ using a pre-load spring. Motion during the experiments was produced manually by the experimenter.

Four separate motions were performed in order to produce slip at the two interfaces (Table 2). In the first, the MC hand was closed upon one of the texture plates which was pulled through the grasp via external weights, inducing slip between the fingers and the texture plate. The other three motions produced slip

Table 2: Types of motions used to generate different slip conditions on the BioTac[®] experimental system

No.	Description	Slip
1	Held object drops with additional weights.	object/hand
2	Stylus drags across surface of held object.	object/world
3	Soft secondary gripper grasps object and pulls.	object/world
4	Hard secondary gripper grasps object and pulls.	object/world

between the held texture plate and an external object. The second scenario consisted of bracing the held texture plate against a vertical rod to prevent large oscillations while a blunt aluminum stylus, approximately 15 cm long with a tip radius of 3 mm, was dragged across it with a normal force of approximately 2 N. The third and fourth motions were identical except for the contact surfaces involved. Both utilized an additional passive gripper mounted on a vertical shaft to pinch the texture plate. This additional gripper was manually drawn along the shaft, slipping against the texture plate. In one case the gripper’s surface was hard plastic while in the other the gripping surface was covered with the same silicone skin used for texture D.

In these experiments slip was produced at either a low speed, of approximately 2 cm/s or a high speed of approximately 5 cm/s. Each motion, texture, and speed combination was performed 10 times for a total of 640 trials. Data were recorded using custom software provided by Syntouch from both BioTacs for a period of 3 seconds, including the slip events of interest. The data were manually cropped after collection by visually observing the pressure signal to determine when the trial occurred in that 3 s window.

4.2 Capacitive Sensors and Adaptive-Gripper

For the second set of experiments, capacitive sensors were affixed to a Robotiq Adaptive Gripper[®]. The suite, shown fully installed on the hand in Figure 8,

consists of 132 taxels arrayed across the grasping surfaces. The taxels have an average area of $\sim 1\text{cm}^2$ and are covered with a silicone rubber skin textured with 1.6 mm tall posts with 2 mm diameter hemispherical ends with a 77% packing ratio (same as texture D in Figure 5). This configuration yields a sensor and skin with stiffness of 300N/m and low force sensitivity of 6mN . In addition to the capacitive taxels, each distal phalanx is further equipped with a three-axis accelerometer for measuring vibrations of the sensor package on that phalanx. The sensor design is covered in (Aukes et al, 2014; Heyneman and Cutkosky, 2012). The sensors have a physical bandwidth of $> 200\text{Hz}$ and are sampled by an Analog Devices AD7147 capacitive to digital converter (CDC). Each AD7147 sequentially samples up to 12 taxels at a rate of 300Hz .

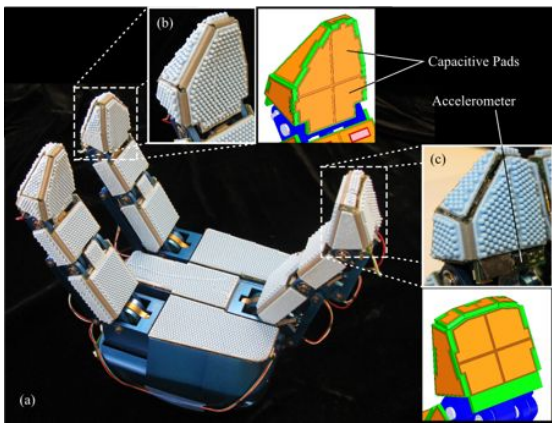


Figure 8: The Robotiq AdaptiveGripper[®] with the capacitive sensing suite installed (a). The full suite includes 132 capacitive taxels and 3, 3-axis accelerometers mounted on the back of each distal phalanx. Insets show closeups of the outer skin texture and the configuration of the capacitive pads on the inner surface (b) and back (c) of the distal phalanx, as well as the location of the 3-axis accelerometer.

The Robotiq AdaptiveGripper equipped with the capacitive sensor suite was mounted on a position-controlled AdeptOne 5-axis SCARA robot arm, (Figure 9) to move the gripper into position and produce manipulation primitives to generate both types

of slip conditions. The gripper was used at its lowest strength setting producing grip forces of $\sim 15\text{N}$ during all trials.

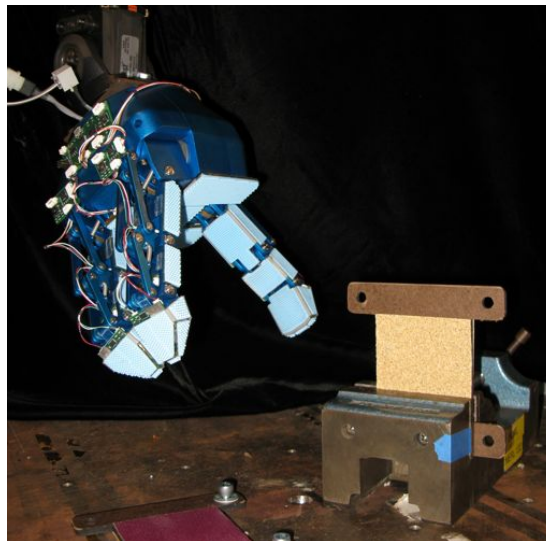


Figure 9: Experiments using the capacitive sensor suite for the AdaptiveGripper[®] were performed with the gripper mounted on an AdeptOne robot arm. Manipulation primitives were scripted via simple coordinated motion of the arm and gripper, while data were recorded from the tactile sensor suite.

Five motions were executed, as outlined in Table 3. Motion 4 was repeated twice; once with a rounded hemispherical stylus and once with a sharp pointed stylus. Motions 1 and 2 both produce slip between the finger surface and the texture plate, while motions 3-5 produce slip between the held texture plate and the environment. Motion 5 was specifically chosen to be difficult. In this case the surfaces involved at both interfaces are the same: texture plate and silicone molded skin. The speed was held constant in all trials and each motion/texture combination was repeated 10 times for a total of 480 trials. Markers were automatically inserted into the captured data after the Adept had finished accelerating into its motion and after the motion had completed to allow for automatic extraction of relevant slip data.

Table 3: Manipulation primitives used in experiments with the AdeptOne

No.	Description	Slip
1	Pinch stationary object and pull away	object/hand
2	Rub single finger across a stationary object	object/hand
3	Remove object wedged (but not clamped) in a vise opening	object/world
4	Drag stylus across stationary object	object/world
5	Drag object across patch of silicone with skin texture	object/world

4.3 PVDF Sensors and Benchtop Platform

The final set of experiments utilized another custom sensor design. These sensors incorporate embedded strips of polyvinylidene-fluoride, or PVDF, a piezoelectric polymer often used for dynamic tactile sensing (Cutkosky and Ulmen, 2014). The embedded strips are thin and flexible, having little effect on the elastomers in which they are embedded.

The fingertips are designed to be interchangeable with those of the Robotiq Adaptive Gripper and contain up to 8 strips of PVDF, along with an AnalogDevices AD7608 digital acquisition chip. The chip provides variable sampling rates, matched analog anti-aliasing prefilters, and simultaneous sampling on up to 8 channels. Due to communication throughput considerations in the rest of the acquisition system, the sensors were sampled at 3.125kHz, with an anti-alias filter with a cutoff at 1.5kHz.

Although they can be mounted to the Adaptive-Gripper, the experiments with the PVDF sensors were performed on a dedicated benchtop test platform designed for controlled slip experiments. As illustrated in Figure 11, the platform consists of three engagement mechanisms arranged around a centrally mounted texture sample. Each fingertip is brought into contact with the sample via a spring loaded 4-bar mechanism, and can then be moved vertically using

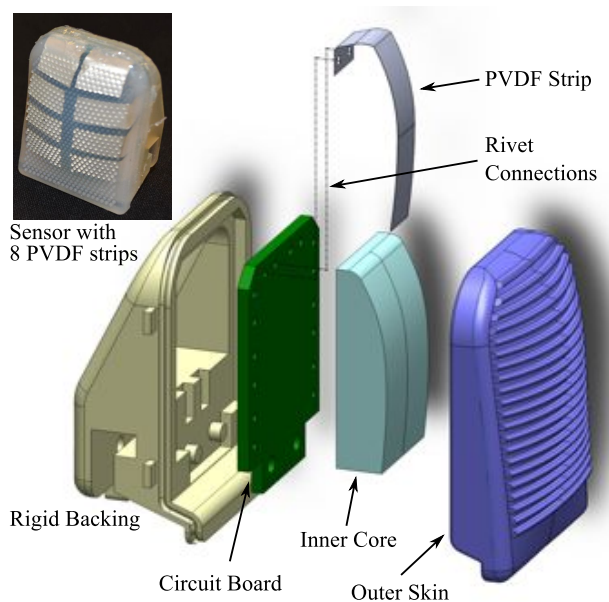


Figure 10: An exploded view of a PVDF sensor fingertip shows the basic construction. A two step molding process with both an inner core and outer skin allows the position of the PVDF strip(s) to be tightly controlled. Rivets connect the strip(s) to the circuit board, which is bonded to a 3D printed finger backing. The inset shows a completed sensor.

a linear slider. The texture sample is mounted on a commercial 6-axis JR3 force/torque sensor with 1kHz bandwidth. This sensor provides an initial contact signal for calibration as well as total shear force during experiments and an independent measure of the vibrations experienced by the sample. An infrared position sensor is mounted inside each engagement mechanism to measure motion of the samples before, during, and after slip.

As in the BioTac experiments, slip was produced manually by the experimenter. While the system allowed for much more sensing of the entire process, the breadth of motions was much more limited. Two fingertips with sensors, and a third dummy fingertip with no sensors, were brought into contact with the sample. Then either the dummy or the two fingertip sensors were manually slid along the vertical

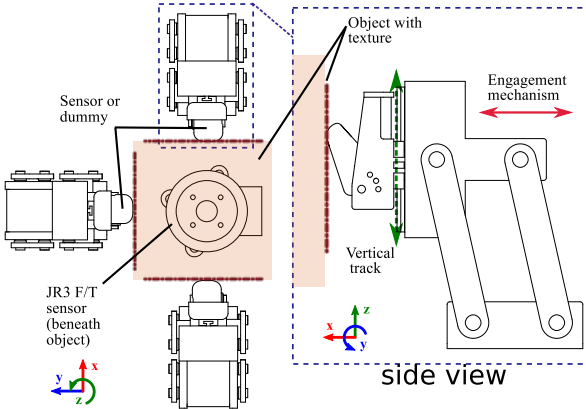


Figure 11: (Left) A top view of the test platform shows the arrangement of engagement mechanism around the JR3 force/torque sensor where the texture samples are mounted. (Right) A side view of the engagement mechanism highlights the parallel four-bar mechanism which is used to bring sensors or inert objects into contact with the texture samples and the vertical slider track which allows them to slip against it.

tracks. Sliding the passive dummy fingertip produces object/world vibrations that are transmitted through the object to the sensor fingertips; sliding the sensor fingertips produces object/finger slippage.

For these experiments the texture samples were limited to the three sandpapers used in the other tests. Contact forces of $\sim 5.5\text{N}$ and $\sim 10\text{N}$ were produced by varying the preload springs in the engagement mechanisms. To reduce abrasive wear, the object samples were covered with a thin sprayed-on rubber film. Each experiment was repeated 10 times, resulting in a total of 120 trials.

5 Classification Results

Once relevant features are extracted from raw sensor data there are many algorithms that can be applied in order to make classifications. Since the focus of this research is on the design of physically motivated and extendable features we evaluate their per-

formance using a simple naive-Bayes maximum likelihood estimator (MLE). The features are extracted from each trial and grouped based on their slip type: object/hand vs. object/world. Using a 10-fold validation scheme each group is split in to 10 sets and each used as the validation set in turn. The remaining 9 sets not used for validation are instead used to fit the features from each class to a normal distribution. The features from the validation trials are then assigned slip classes to maximize the likelihood of the measured features, given the probability distributions inferred from the test sets.

The performance of a feature is simply the number of trials that were correctly identified using this procedure. While this is certainly not the most sophisticated method for using these features, it does provide an equal footing for comparison.

5.1 General Feature Performance

Either of the features can be computed over any frequency range, including the entire signal spectrum. While the results from the naive-Bayes MLE serve as a quantitative comparison, performance can be qualitatively demonstrated by graphing the classifiers for a range of frequencies. In Figure 12a, the feature \mathcal{G}^2 (eq. 34) is computed for the BioTac sensor experiments in frequency bands of 10Hz. The 10 trials for a given experiment are averaged to produce each data point in the figure. In this plot, as well as Figures 15, 17, and 18, unfilled data points indicate object/hand slip while filled indicate object/world slip trials. To simulate the action of the naive-Bayes MLE, Figure 12b shows the regions bounded by one standard deviation for the sets of *all trials* for both classes of slip. In this example one can see very good separation of these two groups, suggesting that the MLE will adequately classify examples.

In order to investigate the general performance of these features the naive-Bayes MLE is used in two ways. First, the features are computed using the entire spectrum. Second, the frequency band which gives the maximum accuracy is determined by comparing the performance of the features when computed across all possible bands. The latter approach corresponds to the case where a sufficient body of

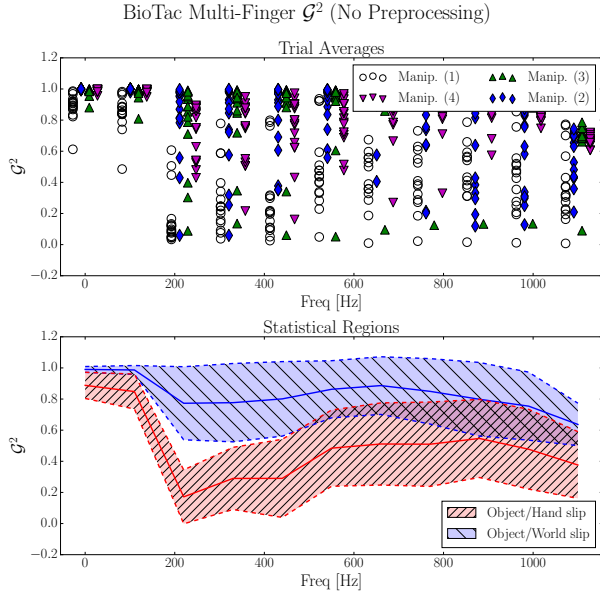


Figure 12: A visual comparison of feature performance for the average \mathcal{G}^2 value across the 10 trials for each BioTac experiment at a variety of frequencies (a). The Object/Hand slip trials (Manip. 1 from Table 2) are shown as unfilled circles; Manip. 2,3,4 denote Object/World slips. The regions describing the examples of each type of slip can then be defined by the average (solid line) and standard deviation (dotted line) at each frequency (b). The \mathcal{G}^2 value is generated using the two pressure sensors in the BioTac fingers, an example of *interfinger* feature analysis.

data can be collected such that the features are trained based on that body.

Tables 4-6 show the quantitative performance using the naive-Bayes MLE as described above. The feature \mathcal{G}^2 has been calculated from a single finger, as well as for multiple fingers. In the case of the BioTac, the single-finger \mathcal{G}^2 utilizes the array of impedance electrodes while the multi-finger \mathcal{G}^2 uses the high-frequency pressure signals from the two fingers.

While some combinations of sensors and features work across the entire spectrum, others do not; for example computing the single-finger \mathcal{G}^2 with the Bio-

Table 4: Power ratio Γ discrimination performance

Sensor	Full spectrum	Best band	(f, w) [Hz]
BioTac	50.39%	76.33%	(23.5, 0.5)
Capacitive	83.37%	89.55%	(0.5, 0.5)
PVDF	86.55%	87.39%	(122.6, 2.5)

Table 5: Single Finger Group Coherence \mathcal{G}^2 Performance

Sensor	Full spectrum	Best band	(f, w) [Hz]
BioTac	49.38%	79.84%	(48.0, 2.0)
Capacitive	66.31%	67.8%	(135.0, 2.0)
PVDF	92.86%	98.74%	(1540, 7.5)

Tac electrode data is only marginally better than randomly guessing. For combinations with poor performance it is important to consider how knowledge about the sensor may help in preprocessing the data or choosing a frequency band to use, without resorting to the generation of a large learning set.

5.2 Augmenting with Sensor Specific Knowledge

Although the two features are meant to be generalizable across sensors, they are not required to be *sensor agnostic*. In this section we examine properties of each of the sensors that can be used to improved performance.

There are two general categories of sensor-specific preprocessing that were discussed in Section 3; *sequential-sampling correction* and *contact estimation*. For the BioTac and capacitive sensors, which sample arrays sequentially, the taxel data are transformed into the frequency domain and a pure time-delay applied to each signal individually with the appropriate sampling delay. This yields a new set of signals which estimate a simultaneously sampled system. Note that this correction only affects the calculation of Γ , since signal coherence, including \mathcal{G}^2 , is insensitive to constant phase shifts due to time delays. Contact estimation is covered in the sections

Table 6: Multi Finger Group Coherence \mathcal{G}^2 Performance

Sensor	Full spectrum	Best band	(f, w) [Hz]
BioTac	75.78%	86.72%	(5.0, 5.0)
Capacitive	59.49%	75.9%	(146.5, 2.5)
PVDF	99.16%	100%	(22.55, 2.5)

below for each type of sensor.

5.2.1 BioTacs

The fluid-filled skin construction of the BioTacs introduces physical coupling between measurements across the electrode array. Because the fluid is incompressible, as the skin compresses in one area it bulges away from the core in others (Figure 13), leading to increased impedance at the contact location and decreased impedance away from it. A simple method for contact estimation is to look for electrodes which have an impedance above the nominal by some threshold amount. In this case, that threshold is determined by the noise on the sensor estimated from the standard deviation of the signals seen from the sensor at rest with no contact.

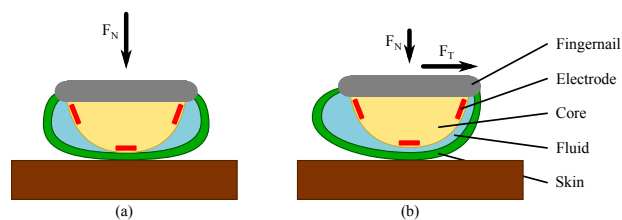


Figure 13: The fluid-filled elastomeric skin results in strong coupling between impedance measurements. A contact force with a significant shear component causes the skin to compress on one side and bulge on the other (b), while a purely normal contact force causes both sides to bulge as the contact area compresses (a). *Modified from (Wettels and Loeb, 2011)*

After accounting for both sequential sampling and contact estimation, the new performance values for the features on the BioTac data are shown in Fig-

ure 14. In all cases the corrected data either shows improvement in classification performance or is indistinguishable from the uncorrected versions.

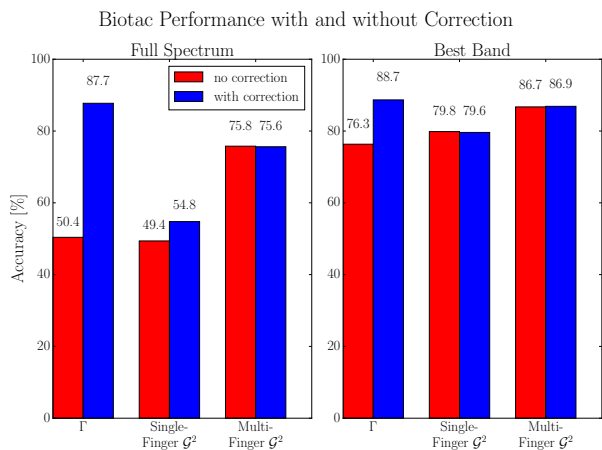


Figure 14: The classification performance of the various features are shown for the BioTac sensors using both original, uncorrected data as well as data that has been corrected for sequential sampling and uses the estimated contact location. These sensor specific processing steps result in significant improvements for some features and negligible changes for others, while none are negatively affected.

5.2.2 Capacitive Sensors

The capacitive sensors have little coupling between taxels. However, they are affected by being mounted on a hard, flat surface in contact with flat objects. Upon examining the qualitative results of the feature Γ (Figure 15) it is clear that the examples of object/hand slip have generally higher values than object/world slip, in contrast to what one might expect based on human mechanoreception and what is seen in the other two sensors.

This result can be explained by considering the flat-on-flat nature of the AdaptiveGripper/object contacts. Figure 16 demonstrates the different mechanisms at work in this scenario. When an object slips against an external object (Figure 16a) it predominantly rotates in the fingertip grasp of the Adaptive-

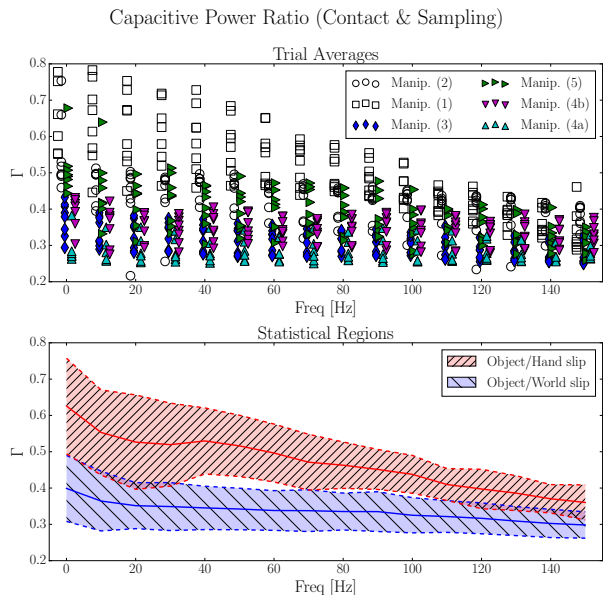
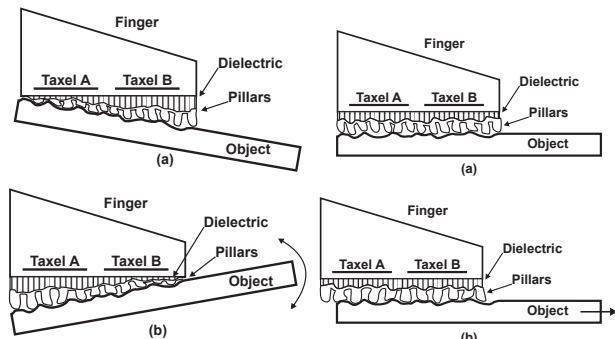


Figure 15: The visual comparison for the capacitive sensors demonstrates that Γ for object/world slip (Manip. 1, 2 from Table 3) is generally smaller than for object/hand (Manip. 3-5), unlike what is predicted from human mechanoreception. This inversion is likely due to the fact that the capacitive sensor experiments consisted primarily of flat-on-flat contact conditions.

Gripper. This rotation causes relatively large, out-of-phase signals on the taxels near the tip of the finger relative to those at the base. These signals significantly cancel when calculating $\mathcal{P}_\epsilon(\omega)$ resulting in small values of Γ . In contrast, when the object slips in the grasp this cancellation is not nearly as dominant, meaning the magnitudes of $\mathcal{P}_\epsilon(\omega)$ and $L(\omega)$ are of a similar scale, and Γ is closer to 1.

5.2.3 PVDF Sensors

For the PVDF sensor, the vibration frequencies of the surface texture elements are of importance when objects slide over the surface of the finger. Using a finite element model and sliding tests, it was confirmed that the fundamental vibration frequency of individ-



(a) Effect of slip between an object and the environment when held by the capacitive sensor fingertips (b) Effect of slip between an object and the capacitive sensor fingertips

Figure 16: Because the surfaces of both the capacitive sensor fingertips and the texture plates used in the experiments are globally flat, they exhibit pathological responses to the vibrations caused by object/world (left) and object/hand (right) slip. This is compounded by the fact that the taxels are relatively large compared to the surface texture features.

ual texture elements was at approximately 1360Hz. The classification features also show the largest separation for frequencies in this approximate range (Figure 17).

Slip between the object and hand will excite vibrations at this frequency as individual surface texture elements experience stick-slip and resonate at this fundamental frequency, generating significant incoherent signal content. In contrast, slip between the object and environment will not generally exhibit significant content at this frequency. By focusing on frequencies that are indicated by properties of the sensor itself, like the fundamental frequency of vibration of surface features, the signal processing and classification will operate on data which best captures the inherent differences between slip at the two different interfaces.

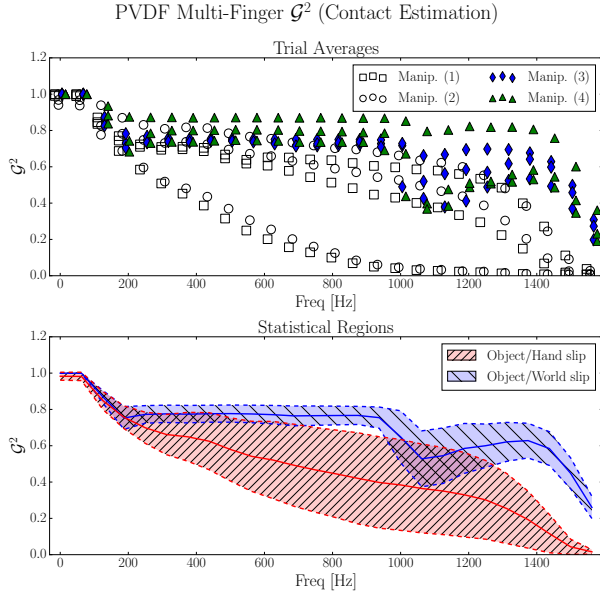


Figure 17: The visual comparison \mathcal{G}^2 between the two PVDF sensors highlights the utility of characterizing important frequencies in a sensor design. The largest separation between object/world and object/hand slip examples occurs above 1300Hz. This corresponds very closely to the first mode frequencies of the texture elements on the sensor surface, which are primarily excited during object/hand slip.

5.3 Biomimetic Sensors vs. Bio-inspired Processing

The power-ratio classification feature, Γ , presented in Section 3.1 can be computed for any tactile array sensor. While inspired by the function of the two relevant human mechanoreceptors it does not require, and may not produce comparable results with, a heterogeneous bio-inspired sensing suite.

Sensing suites described as *biomimetic* often use two or more types of sensors, each measuring a different quantity through different transduction methods and/or placement, in order to more directly mimic the FA-I and FA-II mechanoreceptors in the human hand. In the classification feature Γ presented here we have relied upon the ratio taken in Equation (6)

to cancel common-mode amplitude effects on both the individual and ensemble measures due to different slip conditions. While this is likely to be true when both measures are calculated from a single array, it may not be true when the signals are generated by diverse sensors. Similarly, for the coherence feature using two different sets of sensors invalidates the modeling used in its derivation and makes a relative comparison of group coherence less meaningful. For example, one sensor may generally have higher coherence than the other due to the mechanics of the transduction, regardless of the type of slip involved.

As an example, if we compute the Γ feature for a combination of electrode and pressure sensors in the BioTac sensor, Figure 18, it is readily apparent that there is no region in which the two types of slip demonstrate reasonable separation. Indeed, when the classifier is qualitatively evaluated, the classification accuracy drops by 8.9%.

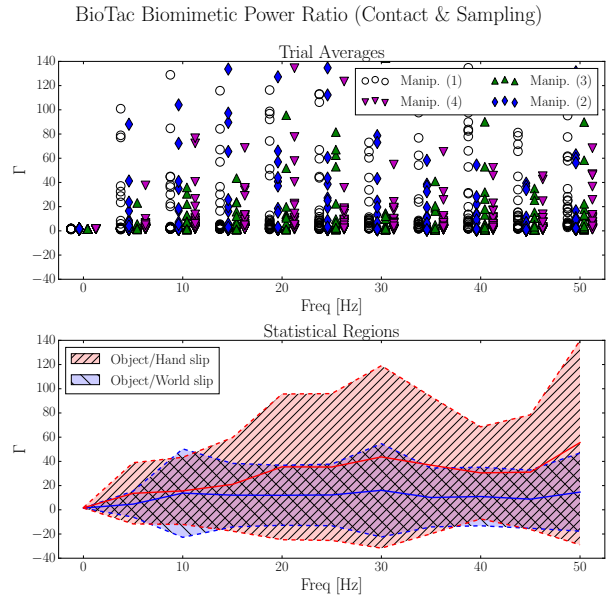


Figure 18: When using two separate sensors to more directly mimic FA-I and FA-II mechanoreceptors, the effectiveness of the classification feature Γ decreases, as indicated by the significant overlap between trials of different types of slip.

6 Conclusions and Future Work

This work focused on a new tactile classification problem relevant to advanced grasping and manipulation in real world environment: determining whether slippage has occurred between the hand and the held object or between the held object and the environment. By considering the function of human mechanoreceptors in the same scenarios and the mechanics of tactile sensors we proposed two simple, frequency-based features that can be extracted from tactile array signals.

These features have been evaluated using three different tactile array sensors in three different experimental setups. With appropriate sensor-specific processing steps, such as contact estimation and sequential sampling correction, a simple naive-Bayes maximum likelihood estimator using 10-fold cross validation was able to achieve $> 85\%$ classification accuracy for the majority of the feature and sensor combinations.

Some of the experiments included manually generated motions, producing only moderately controlled and repeatable trajectories, as typically experienced in manipulating objects. We believe the variations in slip speed, contact force, or other contact conditions caused by this manual process are beneficial in demonstrating the insensitivity of these classification features to those factors. As an example, when sensors with regular surface textures slip they generate frequency content based on the texture’s spatial frequency and the slip speed. If slip speed were controlled tightly from trial to trial the result could be a misapprehension that a certain frequency range is particularly useful.

We discussed the utility of sensor characterization for improving the generalizability of these features for use with other sensors. For example, understanding how to perform contact estimation, and which frequencies will be preferentially excited will help generate the most discriminating features. Our results demonstrate that these features can be used to distinguish between object/world and object/hand slip even when using broad-spectrum data. However, as discussed for the PVDF sensors, we expect they will

be most effective when the sensor mechanics, sensor electronics, and signal processing methods are working together: the surface texture should naturally generate signals at the frequency of peak sensitivity for the electronics, and the methods described here should use the signal content around that frequency.

The data presented here were post-processed. In implementing a real-time system the processing required to generate these features is not prohibitive. For example, the process of computing the coherence feature for an array of 20 sensors using 1024 sample points and Welch’s algorithm (256 sample windows, 128 sample overlap) takes 23.9ms, dominated by the time required to compute the largest eigenvalue. This benchmark is taken using unoptimized python code running on a 2.5GHz i7 processor. Instead, the main trade-off will be between the time required for an accurate spectral estimate and the acceptable latency in the system. Considering a goal similar to human reaction times on the order of 200ms, sensors with a lower sample-rate like the BioTac impedance array at 100Hz will have trouble making this trade-off, as there will not be enough new samples (20 in this example) from which to make a robust spectral estimate.

The classification features presented here are not intended to be used in isolation. To reduce the incidence of false positive and false negative errors in slip detection it will be important to take advantage of additional information from joint-torque sensors, contact sensors, low-frequency tactile arrays, and even vision systems. These sensors can be used to estimate the contact state of the hand/arm/object system in order to inform the use of any event classification algorithms. For example, when manipulating an object, knowledge about the proximity of other items in the environment may generate a prior probability distribution for use with classification methods like the naive-Bayes MLE used here; if there are no external objects in the vicinity of the hand it is quite unlikely that sensed vibrations in the tactile arrays are due to object/world slip.

Although not explicitly tested in this work, there are two situations in which a constant slip classification will generate false-positives when neither slip is occurring and noise or other vibrations in the sys-

tem or environment are sensed by the tactile arrays. Because noise sources are typically common to all channels (e.g. power supply noise, EMF coupling, etc.) and therefore coherent, we would expect false object/world slip detection. In order to alleviate this issue, characterizing the noise floor of the array is necessary so that classification is not performed when the sensor signals lie below it. The second situation could occur in a variety of scenarios such as a task being performed on a mobile base (vibrations of the moving base are picked up by all the sensors) or if another object slips along a nearby, but unsensored surface of the hand. In these cases we expect the classification to identify object/world slips more often. Correctly identifying periods of no slip in these conditions will likely require more general system knowledge, as discussed above.

Finally, this work has explored the use of data from multiple fingers in order to classify slip by computing the \mathcal{G}^2 feature across fingers. In real manipulations slip may occur at any of the many interfaces between the object and the multiple fingers *and* the environment. A further extension to coherence, following the same principles as the \mathcal{G}^2 , may be helpful in determining the coherence between populations of sensors relative to the coherence within those groups in order to identify all interfaces at which slip is occurring at any given moment.

7 Acknowledgments

This work has been supported in part by the DARPA ARM-H project with additional support from ARL MAST.

References

- Ascari L, Bertocchi U, Corradi P, Laschi C, Dario P (2009) Bio-inspired grasp control in a robotic hand with massive sensorial input. *Biological Cybernetics* 100(2):109–128
- Aukes DM, Cutkosky MR, Kim S, Ulmen J, Garcia P, Stuart H, Edsinger A (2014) Design and Testing of a Selectively Compliant Underactuated Hand. *International Journal of Robotics Research* 33(5):721–735
- Brisben AJ, Hsiao SS, Johnson KO (1999) Detection of Vibration Transmitted Through an Object Grasped in the Hand. *Journal of Neurophysiology* 81:1548–1558
- Cutkosky MR, Ulmen J (2014) Dynamic Tactile Sensing. In: Balasubramanian R, Santos VJ (eds) *The Human Hand as an Inspiration for Robot Hand Development SE - 18*, Springer Tracts in Advanced Robotics, vol 95, Springer International Publishing, pp 389–403
- Dahiya RS, Metta G, Valle M, Sandini G (2010) Tactile Sensing From Humans to Humanoids. *IEEE Transactions on Robotics* 26(1):1–20
- Decherchi S, Gastaldo P, Dahiya RS, Valle M, Zunino R (2011) Tactile-Data Classification of Contact Materials Using Computational Intelligence. *IEEE Transactions on Robotics* 27(3):635–639
- Fearing RS, Hollerbach JM (1985) Basic Solid Mechanics for Tactile Sensing. *The International Journal of Robotics Research* 4(3):40–54
- Fishel JA, Loeb GE (2012) Bayesian Exploration for Intelligent Identification of Textures. *Frontiers in Neurorobotics* 6(June):1–20
- Fox J, Daniel T (2008) A neural basis for gyroscopic force measurement in the halteres of *Holorusia*. *Journal of Comparative Physiology A* 194(10):887–97
- Fox JL, Fairhall AL, Daniel TL (2010) Encoding properties of haltere neurons enable motion feature detection in a biological gyroscope. *Proceedings of the National Academy of Sciences of the United States of America* 107(8):3840–5
- Fujimoto I, Yamada Y, Morizono T, Umetani Y, Maeno T (2003) Development of artificial finger skin to detect incipient slip for realization of static friction sensation. *Proceedings of IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems, MFI2003* pp 15–20

- Greenstein J, Kavanagh P, Rowe MJ (1987) Phase coherence in vibration-induced responses of tactile fibres associated with pacinian corpuscle receptors in the cat. *The Journal of Physiology* 386:263–275
- Gunji D, Mizoguchi Y, Teshigawara S (2008) Grasping Force Control of Multi-fingered Robot Hand based on Slip. In: *IEEE International Conference on Robotics and Automation*, pp 2605–2610
- Heyneman B, Cutkosky MR (2012) Biologically inspired tactile classification of object-hand and object-world interactions. In: *IEEE International Conference on Robotics and Biomimetics*
- Ho AV, Hirai S (2014) *Mechanics of Localized Slippage in Tactile Sensing*, Springer Tracts in Advanced Robotics, vol 99. Springer International Publishing, Cham
- Holweg E, Hoeve H, Jongkind W, Marconi L, Melchiorri C, Bonivento C (1996) Slip detection by tactile sensors: algorithms and experimental results. In: *IEEE International Conference on Robotics and Automation*, April, pp 3234–3239
- Hosoda K, Tada Y, Asada M (2006) Anthropomorphic robotic soft fingertip with randomly distributed receptors. *Robotics and Autonomous Systems*
- Howe RD, Cutkosky MR (1989) Sensing skin acceleration for slip and texture perception. In: *IEEE International Conference on Robotics and Automation*, IEEE Comput. Soc. Press, pp 145–150
- Howe RD, Cutkosky MR (1993) Dynamic tactile sensing: perception of fine surface features with stress rate sensing. *IEEE Transactions on Robotics and Automation* 9(2):140–151
- Hunt C (1961) On the nature of vibration receptors in the hind limb of the cat. *Journal of Physiology* pp 175–186
- Jamali N, Sammut C (2011) Majority Voting : Material Classification by Tactile Sensing Using Surface Texture. *IEEE Transactions on Robotics* 27(3):508–521
- Jamali N, Sammut C (2012) Slip prediction using Hidden Markov models: Multidimensional sensor data to symbolic temporal pattern learning. In: *2012 IEEE International Conference on Robotics and Automation*, IEEE, pp 215–222
- Jentoft LP, Tenzer Y, Vogt D, Wood RJ, Howe RD (2013) Flexible, stretchable tactile arrays from MEMS barometers. In: *2013 16th International Conference on Advanced Robotics (ICAR)*, IEEE, pp 1–6
- Johansson RS, Birznieks I (2004) First spikes in ensembles of human tactile afferents code complex spatial fingertip events. *Nature Neuroscience* 7(2):170–177
- Johansson RS, Flanagan JR (2009) Coding and use of tactile signals from the fingertips in object manipulation tasks. *Nature Reviews: Neuroscience* 10(5):345–359
- Johansson RS, Westling G (1984) Roles of glabrous skin receptors and sensorimotor memory in automatic control of precision grip when lifting rougher or more slippery objects. *Experimental Brain Research* pp 550–564
- Johnson M, Adelson E (2009) Retrographic sensing for the measurement of surface texture and shape. In: *2009 IEEE Conference on Computer Vision and Pattern Recognition*, Ieee, pp 1070–1077
- Kim SH, Engel J, Liu C, Jones DL (2005a) Texture classification using a polymer-based MEMS tactile sensor. *Journal of Micromechanics and Microengineering* 15(5):912–920
- Kim SH, Engel J, Liu C, Jones DL (2005b) Texture classification using a polymer-based MEMS tactile sensor. *Journal of Micromechanics and Microengineering* 15(5):912–920
- Knibestöl M (1973) Stimulusresponse functions of rapidly adapting mechanoreceptors in the human glabrous skin area. *The Journal of physiology* 232:427–452

- Lepora NF, Evans M, Fox CW, Diamond ME, Gurney K, Prescott TJ (2010) Naive Bayes texture classification applied to whisker data from a moving robot. In: The 2010 International Joint Conference on Neural Networks (IJCNN), IEEE, pp 1–8
- Lin CH, Erickson TW, Fishel JA, Wettels N, Loeb GE (2009) Signal Processing and Fabrication of a Biomimetic Tactile Sensor Array with Thermal, Force and Microvibration Modalities. In: IEEE International Conference on Robotics and Biomimetics
- Lucarotti C, Oddo CM, Vitiello N, Carrozza MC (2013) Synthetic and bio-artificial tactile sensing: a review. *Sensors (Basel, Switzerland)* 13(2):1435–66
- Maheshwari V, Saraf RF (2006) High-resolution thin-film device to sense texture by touch. *Science (New York, NY)* 312(5779):1501–1504
- Marple SLJ (1987) *Digital Spectral Analysis: with applications*. Prentice-Hall, Englewood Cliffs, NJ
- Mazid a, Russell R (2006) A Robotic Opto-tactile Sensor for Assessing Object Surface Texture. 2006 IEEE Conference on Robotics, Automation and Mechatronics pp 1–5
- Melchiorri C (2000) Slip Detection and Control Using Tactile and Force Sensors. *IEEE/ASME Transactions on Mechatronics* 5(3):235–243
- Mittendorfer P, Cheng G (2011) Humanoid Multimodal Tactile-Sensing Modules. *IEEE Transactions on Robotics* 27(3):401–410
- Oddo CM, Controzzi M, Beccai L, Cipriani C, Carrozza MC (2011) Roughness Encoding for Discrimination of Surfaces in Artificial Active-Touch. *IEEE Transactions on Robotics* 27(3):522–533
- Ramírez D, Vía J, Santamaría I (2008) A generalization of the magnitude squared coherence spectrum for more than two signals: definition, properties and estimation. In: IEEE International Conference on Acoustics, Speech and Signal Processing, pp 3769–3772
- Romano JM, Hsiao K, Niemayer G, Chitta S, Kuchenbecker KJ (2011) Human-Inspired Robotic Grasp Control with Tactile Sensing. *IEEE Transactions on Robotics* (99):1–13
- Schmitz A, Maiolino P, Maggiali M, Natale L, Cannata G, Metta G (2011) Methods and Technologies for the Implementation of Large-Scale Robot Tactile Sensors. *IEEE Transactions on Robotics* 27(3):389–400
- Sinapov J, Sukhoy V, Sahai R, Stoytchev A (2011) Vibrotactile Recognition and Categorization of Surfaces by a Humanoid Robot. *IEEE Transactions on Robotics* 27(3):488–497
- Teshigawara S, Tsutsumi T, Shimizu S, Suzuki Y, Ming A, Ishikawa M, Shimojo M (2011) Highly sensitive sensor for detection of initial slip and its application in a multi-fingered robot hand. In: 2011 IEEE International Conference on Robotics and Automation, IEEE, pp 1097–1102
- Tiwana MI, Redmond SJ, Lovell NH (2012) A review of tactile sensing technologies with applications in biomedical engineering. *Sensors and Actuators A: Physical* 179:17–31
- Wettels N (2012) Biomimetic tactile sensor for object identification and grasp control. Ph.d. thesis, University of Southern California
- Wettels N, Loeb GE (2011) Haptic Feature Extraction from a Biomimetic Tactile Sensor: Force, Contact Location and Curvature. In: IEEE International Conference on Robotics and Biomimetics
- Witney AG, Wing A, Thonnard JL, Smith AM (2004) The cutaneous contribution to adaptive precision grip. *Trends in Neurosciences* 27(10):637–643
- Yahud S, Dokos S, Morley JW, Lovell NH (2010) Experimental validation of a polyvinylidene fluoride sensing element in a tactile sensor. In: IEEE International Conference of the Engineering in Medicine and Biology Society, vol 2010, pp 5760–5763

Yousef H, Boukallel M, Althoefer K (2011) Tactile sensing for dexterous in-hand manipulation in robotics: a review. *Sensors and Actuators A: Physical* 167(2):171–187

Zhang X, Liu R (2012) Slip detection by array-type pressure sensor for a grasp task. In: *2012 IEEE International Conference on Mechatronics and Automation*, IEEE, pp 2198–2202

A Properties of \mathcal{G}^2

This definition of $\mathcal{G}_{\mathbf{x}}^2$ can be checked for four basic properties which indicate it is a useful extension of classical coherence:

- it has the same range as \mathcal{C}^2 ; $\mathcal{G}_{\mathbf{x}}^2 \in [0, 1]$
- it is equivalent to \mathcal{C}^2 when $N = 2$
- it takes on the extreme $\mathcal{G}_{\mathbf{x}}^2 = 0$ iff $\mathcal{C}_{x_i, x_j}^2 = 0$ for $i \neq j$
- it takes on the extreme $\mathcal{G}_{\mathbf{x}}^2 = 1$ iff $\mathcal{C}_{x_i, x_j}^2 = 0$ for $i \neq j$

The first property follows from the fact that $\mathcal{C}_{\mathbf{x}}$ is a normalized *hermitian* matrix. As a hermitian matrix, all its eigenvalues are non-negative real values. Because it is normalized, its diagonal entries are all 1's and therefore its eigenvalues sum to N . These together bound $\lambda_1 \in [1, N]$, which leads to $\mathcal{G}_{\mathbf{x}}^2 \in [0, 1]$.

The second property can be verified by examining $\mathcal{C}_{\mathbf{x}}$ for $N = 2$:

$$\mathcal{C}_{\mathbf{x}} = \begin{bmatrix} 1 & \mathcal{C}_{x_1, x_2} \\ \mathcal{C}_{x_1, x_2}^* & 1 \end{bmatrix}$$

and the eigenvalues of this matrix are $\lambda = 1 \pm |\mathcal{C}_{x_1, x_2}|$. Plugging into Equation (34) yields $|\mathcal{C}_{x_1, x_2}|^2$, as expected.

To prove that $\mathcal{G}_{\mathbf{x}}^2$ only attains the extreme value of 0 when all individual signals are pair-wise incoherent we simply note that the only normalized hermitian matrix with eigenvalue $\lambda_1 = 1$ is the identity matrix. If $\mathcal{C}_{\mathbf{x}} = \mathbf{I}$, then all the off-diagonal terms, the pair-wise coherence between signals, are 0 as expected.

The proof for the conditions under which $\mathcal{G}_{\mathbf{x}}^2 = 1$ is slightly more involved. First we note that $\lambda_1 = N$ from Equation (34). This implies that $\lambda_i = 0$ for $i \neq 1$ and therefore that $\text{rank}(\mathcal{C}_{\mathbf{x}}) = 1$. From this rank constraint we know that we can decompose the complex coherence matrix as

$$\begin{aligned} \mathcal{C}_{\mathbf{x}} &= \mathbf{v}^* \mathbf{v}^T \\ &= \begin{bmatrix} v_1^* v_1^T & v_1^* v_2^T & \cdots & v_1^* v_N^T \\ v_2^* v_1^T & v_2^* v_2^T & \cdots & v_2^* v_N^T \\ \vdots & \vdots & \ddots & \vdots \\ v_N^* v_1^T & v_N^* v_2^T & \cdots & v_N^* v_N^T \end{bmatrix} \end{aligned}$$

for some vector \mathbf{v} . From the constraint that all diagonal elements of $\mathcal{C}_{\mathbf{x}}$ must be 1, we have $|v_i| = 1$.

Therefore all elements of $\mathcal{C}_{\mathbf{x}}$ have magnitude 1, and we have

$$\mathcal{C}_{\mathbf{x}}^2 = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

as desired.

To prove the reverse direction, first note that for a set of completely coherent signals, $\mathbf{x}(t)$, each signal can be defined as a filtered version of the first signal:

$$\begin{aligned} x_i(t) &= (h_i * x_1)(t), \text{ or} \\ \mathbf{x}(t) &= (\mathbf{h} * x_1)(t), \text{ with} \\ \mathbf{h}(t) &= [1, h_2, \dots, h_N]^T \end{aligned}$$

This set of signals has the complex coherence matrix

$$\mathcal{C}_{\mathbf{x}} = \mathbf{h}^* \mathbf{h}^T$$

This shows that $\text{rank}(\mathcal{C}_{\mathbf{x}}) = 1$, and therefore $\lambda_1 = N$, $\lambda_i = 0$ for $i \neq 1$. Finally, from Equation (34) this yields $\mathcal{G}_{\mathbf{x}}^2 = 1$, as desired.