Different Strokes and Different Folks: Economical Dynamic Surface Sensing and Affect-Related Touch Recognition

Xi Laura Cang, Paul Bucci, Andrew Strang, Jeff Allen, Karon MacLean, H. Y. Sean Liu
Department of Computer Science
University of British Columbia
Vancouver, BC, Canada
{cang, pbucci, agstrang, jephalun, maclean, hliu02}@cs.ubc.ca

ABSTRACT

Social touch is an essential non-verbal channel whose great interactive potential can be realized by the ability to recognize gestures performed on inviting surfaces. To assess impact on recognition performance of sensor motion, substrate and coverings, we collected gesture data from a low-cost multitouch fabric pressure-location sensor while varying these factors. For six gestures most relevant in a haptic social robot context plus a no-touch control, we conducted two studies, with the sensor (1) stationary, varying *substrate* and *cover* (n=10); and (2) attached to a robot under a fur covering, *flexing* or *stationary* (n=16).

For a stationary sensor, a random forest model achieved 90.0% recognition accuracy (chance 14.2%) when trained on all data, but as high as 94.6% (mean 89.1%) when trained on the same individual. A curved, flexing surface achieved 79.4% overall but averaged 85.7% when trained and tested on the same individual. These results suggest that under realistic conditions, recognition with this type of flexible sensor is sufficient for many applications of interactive social touch. We further found evidence that users exhibit an idiosyncratic 'touch signature', with potential to identify the toucher. Both findings enable varied contexts of affective or functional touch communication, from physically interactive robots to any touch-sensitive object.

Categories and Subject Descriptors

H.5.2 [**INFORMATION INTERFACES AND PRESENTATION**]: User Interfaces—*Haptic I/O*; I.5.2 [**PATTERN RECOGNITION**]: Design Methodology—*Classifier design and evaluation, Feature evaluation and selection*; I.5.4 [**PATTERN RECOGNITION**]: Applications—*Signal Processing*

General Terms

Gesture, Touch and Haptics; Affective Computing and Interaction; Human-Robot Interaction; Non-verbal behaviors

Keywords

Haptics, tangible interaction, social touch, affective touch, flexible touch sensor, pressure-location sensing, recognition techniques

ICMI 2015, November 9–13, 2015, Seattle, WA, USA.

ACM 978-1-4503-3912-4/15/11 ...$15.00.
DOI: http://dx.doi.org/10.1145/2818346.2820756.

1. INTRODUCTION

Words can sometimes be inefficient for communicating instructions or affective content. In many contexts, touch may be the best modality for conveying directive and emotion: imagine how informing someone to get out of the way quickly and clearly with one simple touch. To harness this communication channel, social robots working in tandem with humans must recognize the same haptic language that we use, of which gestures and affect are key components.

We focus on exploring the range of touch gestures detectable by a custom-built flexible fabric pressure sensor and evaluating the added noise from curvature, motion and material cover. Using common machine learning techniques, we highlight salient features of touch for recognizing both the type of gesture being performed, and the person performing the gesture—both the 'touch' and the 'toucher'.

Reliable gesture recognition is an important step towards further research in the field of affective touch. A strong foundation of research on gesture may allow us to detect the toucher's emotional state [14]. Until recently, this kind of research was difficult, as touch sensors were not easily deformable nor cheap in both price and computational resources. Our 10×10 sensor has 100 fingerpad-scale taxels recording pressure and 2-D location data, and we use a random forest classification method to approximate in situ recognition rates.

We first collected touch data for a set of six validated touch gestures [32] plus one control on a stationary sensor under a variety of substrate stiffnesses and coverings. We then mounted the same sensor on an actuated robot skeleton and collected similar data while varying the sensor's covering and motion (Fig. 1). Recognition rates were within 80–95% for all conditions we tested (chance 14.2%), a level of accuracy which will suffice for many purposes and is enough to merit empirical comparison to human recognition ability in future work. At the same time, we found individuals' touch signatures were idiosyncratic enough to permit identification of toucher within this sample, at an accuracy rate similar to that of the gestures themselves.

1.1 Questions and Contributions

We wished to learn:

Q1: How accurate is our flexible fabric sensor in predicting gesture and differentiating between users?;

Q2: How does sensor performance hold up under deformation due to curvature and motion, such as that produced by a zoomorphic social robot?; and

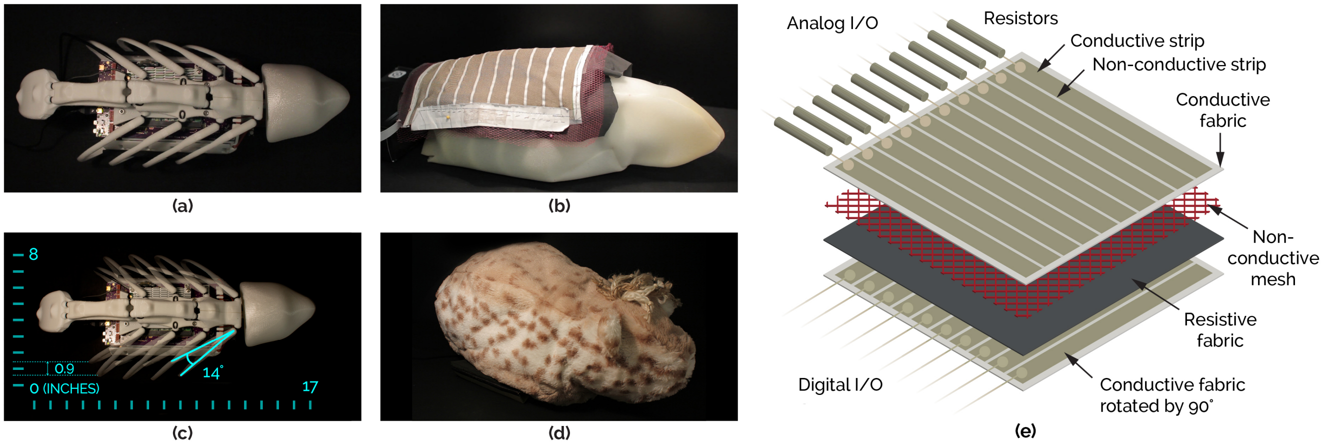

Figure 1: (a) Top view of the CuddleBot skeleton. (b) Touch sensor, pinned to foam substrate wrapped around the skeleton and corresponding to a *No touch* × *No motion* × *No cover* condition. (c) Full range of breathing motion used. (d) The fully-covered robot; a covering of nearly identical material was used in the study to facilitate quick condition changes. (e) The fabric pressure sensor constructed out of EeonTex conductive fabric <www.eeonyx.com>, wired to an Arduino microprocessor.

Q3: Is real-time gesture recognition computationally viable?

With 20-fold cross validation on random forest models, we contribute initial results of:

- deployable accuracy in gesture recognition (6 gestures + control): 91.4% on a firm, flat surface, 90.3% on a foam, curved surface, and 88.4% on a foam, curved, moving surface;
- differentiating toucher at 88.8% accuracy (n=26);
- factors underlying recognition performance;
- feasibility of real-time gesture recognition.

We also make our data and analysis publicly available[1].

Our study compares gesture recognition performance across a variety of conditions that approach real-time dynamic gesture recognition. Toucher recognition accuracy shows promise for incorporating personalized responses to an individual touch signature.

1.2 Applications

Accurate gesture recognition on a fabric touch sensor opens up gesture-based controls on any electronic device. For example, patients with limited speech could use a smart blanket with gesture recognition capabilities for comfort or health-reporting purposes. In the context of social robots, a sensor that can wrap around any irregular form could be used as a touch-sensitive skin. Outside of explicit gesture recognition, pressure-sensitive hospital sheets could alert caregivers of bedsore risk.

In a behavioural education context, a soft touch-sensing playmate capable of recognizing touch signatures may use this data to interpret and influence emotional state [14]. Such a robot could aid students testing on the autism spectrum by responding to anxious or agitated strokes with slow, soothing, regulated breathing—a behaviour shown to have calming benefits [23].

1.3 Detailed Requirements

Our sensing requirements are dictated by a zoomorphic robot, affectionately dubbed the CuddleBot, that invites touch with a soft furry body. Since a user will expect to interact with the CuddleBot via touch, having a full-body sensor that deforms with robot motion is required.

[1]All collected data and select analysis can be found at <www.cs.ubc.ca/labs/spin/data>

Movement and elasticity: The sensor must be highly flexible, somewhat elastic, and perform well while mounted on non-rigid and/or actuated surfaces.

Pressure range: Based on a preliminary survey of these touch pressures, we determined that our sensor needed to register touches between 0.005 and 1 kg. This range is appropriate for light tickles to heavy pats.

Multitouch: Multitouch capability allows us to compute varying pressure over an area, differentiating touches like *constant* and *pat* from *tickle* and *scratch*.

Resolution and computational cost: Taxel resolution, sampling rate, and computational cost must be balanced to achieve usable recognition accuracy. For real-time, our computational cost is dominated by sensor polling and grows with the number of taxels per grid edge. Our recognition tasks and feature selection explicitly analyze the differences between frames. In this case, accuracy plateaus with fingerpad-scale taxels, when sampled fast enough to capture voluntary movement (peaking at 10Hz [25]). We must be able to recognize changes in pressure and localized hand motions up to this frequency.

Single-fingerpad resolution ($\approx$2 taxels per inch) could capture small fluctuations; however, our gestures (not including our control *no touch*) either involve the flat or palm of hand (*constant*, *pat*, *rub*, *stroke*), or tend towards quickly crossing many taxels (*tickle*, *scratch*). This suggests that using statistical features that emphasized the changes from frame to frame could be used to achieve reasonable classification rates even at $\approx$1 inch taxels [7].

2. RELATED WORK

We situate our work in the context of social robotics and affect-encoding social touch. Gestural touch has been identified as a key component of human-robot cooperation [2]. However, the semantics of that touch is conveyed through nuance. For example, the same gesture could halt, contribute or modify another person's behaviour [2] depending on the emotional content inferred from pressure dynamics [14].

2.1 Social & Affective Touch Communication

In collaboration with human workers, robots employed in a laboratory or workshop setting presupposes a lexicon of social touch for operational interactions [8]. To ensure safe and effective communication, Gleeson et al identify the requirements of both a comprehensive gestural dictionary and lightweight sensing technology. The intimate nature of collaborative robotic household help emphasizes the importance of affect detection for social robots in this context [20, 1].

Previous work revealed correlations between gestural social touch and emotional communication [11, 14]. Humans recognize the affect encoded in gestural touch [12, 11], suggesting that machine recognition of emotional state can be achieved with sufficient sensing technology and clever feature extraction.

Much of the current work on social touch recognition uses a sensor worn on a static human or robotic arm [14, 26, 15, 16]. The collected data and signal processing procedures may not account for the added deformation noise of a soft-tissue zoomorphic robotic form in motion.

The use of animals [6] and interactive robots in animal form (such as Sony's pet-dog AIBO [3, 28], the seal-shaped PARO [31, 13, 18, 24]) suggest potential benefits in therapeutic use. Other touchable social robots include the teddy bear-like Huggable [27]; and Probo [22], which does not have a recognizable animal analogue. However, while real pets respond to complex touch commands anywhere on the body, this has been difficult to achieve without a generalized touch-sensitive skin.

In trying to establish zoomorphic robots as an emotional agent [7, 32], touch sensing strategies have included fur-level conductive threads, extensive biometric data, gyroscopes and accelerometers, to name a few. While this cavalcade of sensing produces encouraging results for social gesture classification [7, 32], it is far from the light-weight system required for automatic, real-time recognition.

An unexpected result emerging from social touch recognition is the demonstrably higher accuracy results for within-subject classification over between-subject [7, 15]. Leveraging this result may allow us to use touch behaviours to identify individuals and thus, recognize the nuances of an individual's "touch signature" to better predict touch gestures and, eventually, basic emotional content.

2.2 Flexible Pressure-Location Sensors

Real-time classification of social touch gestures on a flexing, noisy surface requires that we have manageable signal processing while retaining the ability to represent pressure and location.

Here we examine the suitability of existing sensing technology and recognize their influence on our custom build. We do not present our sensor as a contribution.

While many highly accurate pressure-location sensors exist, such as those developed for robot grippers used in dexterous manipulation [21, 30], these tend to be insufficiently flexible, overkill in terms of resolution, and considerably too expensive for the objectives outlined here.

Other work has used Force-Sensing Resistors (FSRs) affixed to a hard shell [32]. This reduces the need to calibrate for sensor drift over continued use, however, the trade-off non-aesthetic tactility, and difficulty in detecting touches between sensors—limiting rendered motion [2, 4].

Stretch sensors designed for medical purposes by Vista Medical[2] is the foremost inspiration for our custom sensor. However, Vista's sensors recognized only pressure without localization and did not have multitouch capability.

Several multitouch, flexible fabric sensors are available [16]. However, flexibility alone does not afford a full range of motion; it must be able to stretch and deform to approximate animal skin.

[2]Stretchable sensors can be purchased commercially from Vista Medical <www.vista-medical.com/subsite/stretch.php>

The design and sensing capabilities described by Flagg et al [7] informed many of our requirements and suggested that the bulk of the recognition accuracy could be achieved by the "below surface" sensor alone. However, Flagg's study did not consider the full design space of a robot in motion including a non-sensing fur and a variety of configurations. To evaluate how much information is compromised under these conditions, we applied a variety of realistic use noise sources to the sensor, both directly and indirectly.

3. STUDIES

We hypothesized that:

H1: *gesture recognition rates will decrease with noise-creating factors*—allowing us to rank these factors' impact on recognition performance, and their interactions therein.

H2: *variability in gesture execution will be higher between subjects than within subjects*—giving rise to the potential of differentiating individuals based on personal touch signatures.

3.1 Apparatus

We constructed a sensor by layering two squares of conductive EeonTex[3] Zebra fabric, aligned at 90 degrees, with a plastic standoff mesh separator and a sheet of EeonTex SLPA 20kΩ resistive fabric. Resistance value across a given taxel drops when pressure is applied, compressing the mesh separator so the conductive layers more closely approach each other. A circuit is constructed using an Arduino Mega microprocessor. Each fabric stripe is connected to a single I/O pin: the top layer is connected to analog input pins, and the bottom layer is connected to digital output pins (Fig. 1(e)).

The sensor is polled by sequentially sending a voltage through the bottom layer's digital pins. The analog pins read current; resistance (and hence current) varies with pressure.

Preliminary testing of our sensor using stationary weights showed that under ideal conditions, we were able to achieve a touch weight range of 0.005–1kg using 1kΩ resistors. Under the most severe conditions, lighter touches were lost in the dense fur; at the heavier end, touches were equalized by the yielding foam substrate. For Study 1, the curved-foam substrate with thick fur cover was the most obscuring condition; for Study 2, this was the cover condition with bot in motion.

Dynamic range is modulated through choice of resistor value. We found that values greater than 1kΩ allowed our sensor to register greater forces, but lost resolution; conversely, lower values gave greater granularity in recognizing very fine touches, but were too vulnerable to saturation at commonly applied force levels. The same sensor and microprocessor set were used in all studies described here.

3.2 Methods

Our two studies assessed how realistic conditions impacted sensor data and hence recognition accuracy; gestures and data collection procedures were unchanged.

3.2.1 Gestures and Sampling

We selected gestures from Yohanan et al's touch dictionary [32], choosing items most appropriate for human-animal interactions [7]. The sensor was placed on a table in front of a seated participant, a reference sheet with very general definitions for six selected gestures and one control was provided (Table 1). Participants were instructed to interpret each gesture as they saw fit; no further performance clarifications were provided.

[3]Sensor fabric purchased from <www.eeonyx.com>

Table 1: Touch gesture instructions as provided to participants.

Gesture	Suggested Definition
no touch	no contact with the sensor (control)
constant	touch contact without movement
pat	quick & gentle touches with the flat of the hand
rub	moving the hand to and fro with firm pressure
scratch	rubbing with the fingertips
stroke	moving hand repeatedly
tickle	touching with light finger movements

A frame consisted of pressure data from all 100 taxels in the 10×10 grid. We collected 10 seconds of continuous hand touch data at 54 frames per second for each combination of gesture and condition, randomizing gestures and conditions wherever possible.

3.2.2 Study 1: Cover and Substrate on Static Robot

We first measured gesture recognition for the static (unmoving) case, to assess impact of the sensor's substrate stiffness, curvature and covering thickness in absence of movement noise. This produced a factorial design of $4 \times 3 \times 7$ ($cover \times substrate \times gesture$), using gestures listed in Table 1.

$Cover$: The fabric's pile or density varied from no cover (participant touched sensor directly) to a very long, thick synthetic fur. Minky (a short furry fabric generally used for baby blankets), and a longer-furred fabric comprised intermediate variations.

$Substrate$: The material underneath the sensor consisted of a firm, flat surface (sensor affixed by velcro to a table); a spongy foam, flat surface; and a spongy foam, curved surface. In cases with foam, the sensor was pinned directly to the foam substrate.

To minimize sensor reading disturbances due to transitions (i.e., unwrapping and replacing the sensor on/off the robot body), we blocked our design on the $cover \times substrate$ conditions. Condition order was randomly generated for every participant, and gesture order was further randomized over each condition set. All participants completed all twelve masking conditions, with each generating 48 2s sample windows per gesture. A study session took approximately 50 minutes to complete. 10 volunteers (4 female, 6 male) were compensated $10 for their time.

3.2.3 Study 2: Stationary vs Moving Robot

Our second study focused on the impact of the robot's breathing movement. We varied $cover \times motion \times gesture$, for a $2 \times 2 \times 7$ factorial design. Factors consisted of $cover$ = {cover, no cover}, $motion$ = {breathing, not breathing}, and $gesture$ = {set of seven gestures}. Each participant performed each condition combination twice in a randomly generated order.

In the breathing condition, the sensor was attached to the CuddleBot, a cat-sized robot designed for therapeutic use. Fig. 1(a-b) shows the naked skeleton and the sensor pinned to the foam intermediary. The robot's 'breathing' motion was created by extending and contracting the paired rib assemblies in a 14° arc from the spine at 0.5Hz (Fig. 1(c)). We draped and pinned fabric over the sensor, approximating a full fur jacket for condition randomization while limiting sensor disruption (Fig. 1(d)),

Each session began by asking the participant to interact freely with the covered, moving robot for 1 minute to reduce novelty. Each condition was then presented twice, in random order, for a total of (($2 \times 2 \times 7$) + 1) = 57 trials. 16 participants (10 female, 6 male) were compensated $5 for the 30 minute session, each providing 32 2s samples of each gesture for every condition.

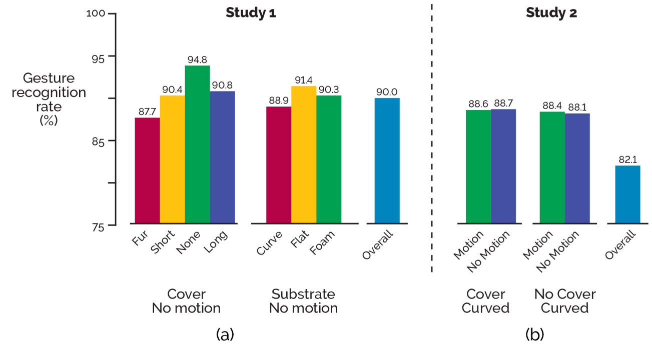

Figure 2: Mean gesture prediction accuracy rates with added pressure noise when (a) varying $substrate$ or $cover$ in Study 1 and (b) varying $motion$ and $cover$ on the same curved structure as in Study 2. Each bar represents an average accuracy rate over 10 trials; error bars are omitted as Δ across trials $< 0.001\%$ in each case.

3.3 Analysis and Results

We discarded the first and last second of each 10s gesture capture and divided the remaining 8s into four 2s windows. The 2s window (at 54Hz) was chosen to allow each gesture some periodicity; all gestures fit completely within 1s (Flagg [7]). Given the challenge of determining gesture boundaries in a realistic, real-time setting when a motion is steadily repeated, a 2s window allows capture of *at least* 1 complete gesture cycle.

To account for translatory gestures, we also calculated a centroid (average geometric centre) weighted by the pressure reading for each frame. Centroids were defined by row C_x (Eq. 1) and column position C_y (Eq. 1 with *i* and *j* indices reversed):

$$C_x = \frac{\sum_{i=1}^{10} \sum_{j=1}^{10} i * pressure(i,j)}{\sum_{i=1}^{10} \sum_{j=1}^{10} pressure(i,j)} \quad (1)$$

We calculated weighted *pressure* by summing readings across each row, multiplying by index, and dividing by the unweighted frame sum (the sum of the full frame sensor reading). Repeated for each column, this provided a tuple of frame sum and centroid per frame.

As a "baseline" for both studies, we sampled sensor frames in the absence of gestures. In Study 1, each of the 12 (4 $cover$ $\times$ 3 $substrate$) condition sets contributed 4320 frames; in Study 2, each of the 4 (2 $motion \times$ 2 $cover$) condition sets contributed 6912 frames. To establish the effect of noise under each condition, we ran MANOVA over three frame-level dependent variables: pressure, C_x-coordinate, and C_y-coordinate. In all cases except one[4], all three variables showed significant differences at the $p < 0.001$ level. This indicates that the sensor is sensitive to changes in these conditions.

The data fails the Shapiro-Wilks test of normality; however, visual inspections of residual Q-Q plots did not reveal any systematic patterns. Together with our large sample size (n>4000 frames / condition), we proceeded with the normality assumption, alert to risk of inflated Type I error.

The six gestures (omitting no-touch data) were then compared with each other under the conditions of each study. MANOVA over the same three metrics (pressure, C_x, C_y) showed that gesture and participant combinations were statistically significant ($p < 0.001$). Differences in participant touch were detectable at frame level.

[4]For Study 2, the condition of *with-cover*$\times$*with-motion* under *no touch* did not show statistically significant differences in C_x data.

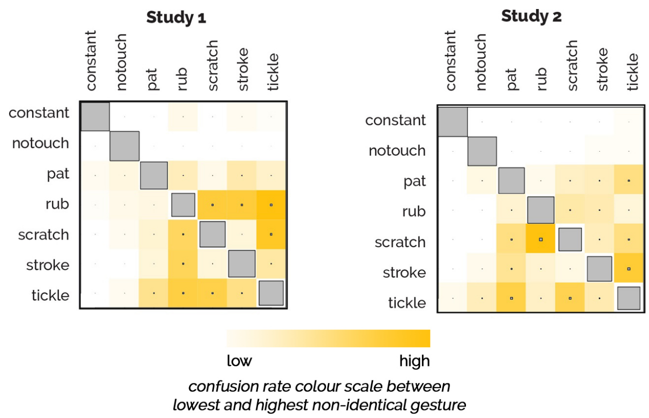

Figure 3: A modified Hinton confusion matrix for gesture classification. Horizontal (row) gestures are classified as the vertical (column) gesture. Saturation in non-diagonal squares represents number of misclassifications.

We calculated seven features across these three dimensions (frame value, C_x, C_y) for each 2s window for a total of 21 features. For each dimension, features are {maximum, minimum, mean, median, variance across all frames, total variance within the 2s window, area under the curve}. Condition variables ($curvature$, fur) or ($cover$, $motion$) make up the other features. $Participant$ labels were included for $gesture$ predictions and vice versa.

Each capture produced four 2s windows, providing repetition for training. Pairwise comparisons of all within- and between- capture windows generated two binomial distributions for statistically significant pairs using two-sample Kolmogorov-Smirnov (KS) tests. Permutation testing [9] using the KS test statistic did not detect a statistically significant difference ($p = 0.214$) between the distributions. This is consistent with our observations of participants varying touch behaviour both between and within captures.

We used Weka, an open-source machine learning application to classify gestures [10]. Flagg's comparison of random forest and a number of other algorithms showed that random forest performed best in gesture recognition of this kind [7]. We ran k-fold Cross Validation (CV) on Study 1 participant data for k = {5, 10, 20, 100} and found less than 1% improvement between 20- and 100-folds. While this CV technique does ensure that any one instance is included in the test or training set and not both, it cannot promise subject-independent classification. Running Leave-One-Out classification yielded slightly improved results but we were cautious to the inflated bias [17]. All reported classification performance is therefore based on the slightly conservative 20-fold cross validation of random forest models. Accuracy is defined as the percentage of data instances that are correctly classified.

3.3.1 Gesture Classification by Condition

H1: Gesture recognition rates will decrease with increase in noise-creating factors—*accepted*.
Comparing classification under Study 1 conditions (static surface), we found highest recognition accuracy with no cover on the firm, flat substrate case. Lowest performers were dense fur and curved, foam substrate. In Study 2 (dynamic surface, heavy versus no cover), conditioning across each of surface and motion factors had minor effect recognition rates (all ≈88%).

With models trained on individual, Study 1 showed little change in gesture prediction rate compared to all-data models. Study 2 individually-trained results are more similar to other studies, which also report training on single-condition data [7, 16, 19, 29].

Cover-substrate-motion: Fig. 2 shows overall gesture recognition accuracy by study and condition set.

We assessed relative noise levels by calculating effect sizes of significant conditions. Cohen's d reveals a large effect ($|d| \geq 0.8$ [5]) with the introduction of curvature (vs *no substrate*) and fur and short minky (vs *no cover*) in Study 1. Large effects ($|d| \geq 0.8$ [5]) from Study 2 were from introducing the cover (regardless of motion), and from the combination of having motion and cover. Interestingly enough, adding motion by itself produced a very low effect ($d \leq 0.08$). Further investigation into the interaction between cover and motion on pressure readings included Tukey's HSD of adjusted p-values to clarify the significance of stratified factors. While all other combinations remained significant at $p < 0.05$, the case of varying motion in the presence of a cover was alone insignificant at $p.adj = 0.7$.

A confusion matrix (Fig. 3) indicates how gestures were misclassified. In both studies, the most-misclassified was *tickle*.

Participant: We classified gestures with models trained by participant. In Study 1, mean accuracy was 89.1% (max=94.6%)[5]. Models trained on all Study 1 data were accurate at 90.0%, i.e. within 1% of the mean accuracy of the individual-trained models. This indicated that training on participants did not improve recognition when data was not conditioned on noise-creating factors.

For Study 2 (fewer noise factors) we found a greater effect for models trained on participants (mean=86.5%, max=97.3%)[6]. Training across all data gave 82.1% accuracy.

The $motion \times cover$ condition had an overall 79.4% recognition rate. Training on the subset of data with the most challenging conditions (in-motion, with-cover) still produced a higher recognition rate when using individual-trained models (mean=85.7%, min=73.7%, max=95.1%).

We compared mean pressure of gesture behaviours by individual versus that of the entire pool (i.e. how P1 performed $scratch$ versus how all participants performed $scratch$). All incidences were significant at $p < 0.05$ (Cohen's d effect sizes reported in Fig. 4).

3.3.2 Toucher Recognition

H2: Variability in gesture execution will be higher between subjects than within subjects—*partially accepted*, for the case of data compared within the same noise conditions.
The ability to recognize toucher may have great impact on reading emotional state. We compared performance in participant classification for models trained across the entire dataset, with those trained on the 6 meaningful gestures of our $gesture$ set (omitting *no touch*). We also look at accuracy rates on data collected in the most realistic condition (in-motion, with-cover).

Recognition rate by study: We compare recognition rate by study and gesture in Fig. 5. Study 1 achieves an overall accuracy rate of 78.5% (chance 10%), but for models trained by gesture, a mean of 87.9%. The highest contributing gesture is *constant* at 92.7%, followed by *pat* at 88.9%.

Using all Study 2 data, participant recognition was 80.3%. Training by gesture again showed recognition improvement; *constant* was best at 93.8%, followed by *pat* at 89.8% (mean, all 6 gestures: 85.4%).

Conditioning on only the *in-motion, with cover* factor, referred to in Fig. 5 as Study 2b, average recognition rates of participants

[5] Study 1 gesture recognition accuracy by participant: P1-93.0%, P2-83.8%, P3-85.0%, P4-92.6%, P5-93.2%, P6-88.0%, P7-94.6%, P8-91.7%, P9-86.0%, P10-83.4%

[6] Study 2 gesture recognition accuracy by participant: P1-90.2%, P2-86.6%, P3-86.6%, P4-91.1%, P5-81.3%, P6-86.1%, P7-84.8%, P8-79.5%, P9-95.5%, P10-79.5%, P11-90.2%, P12-93.8%, P13-97.3%, P14-83.0%, P15-79.5%, P16-79.5%

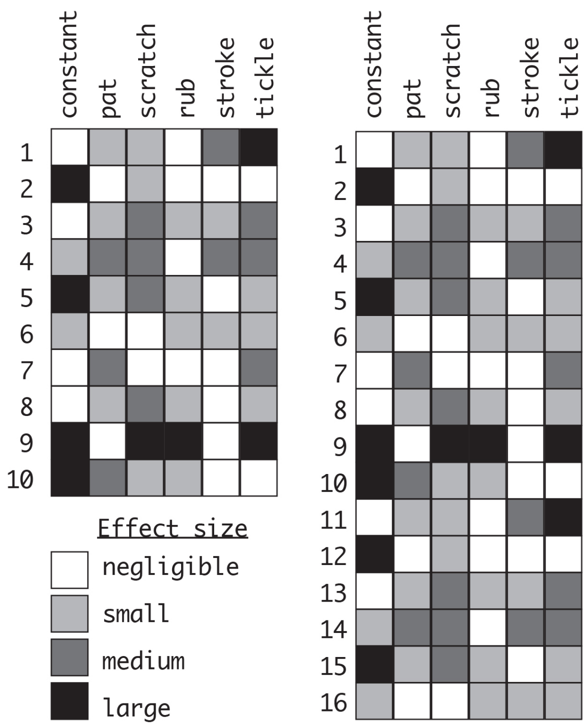

Figure 4: Cohen's d effect sizes of participant by gesture for each study.

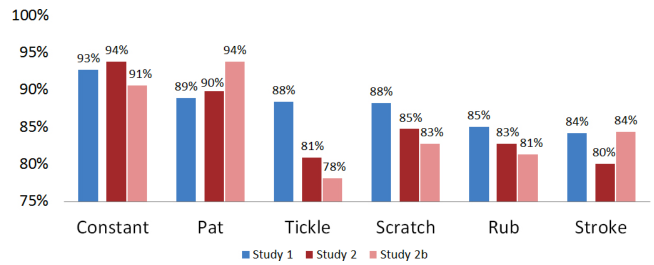

Figure 5: Mean subject recognition rates by gesture and study over 10 trials; error bars omitted as Δ across trials $< 0.01\%$ in each case. Study 2b refers to the 'in-motion, with cover' condition.

are 89.8%. Further splitting data to additionally train models by gesture does not provide additional improvement in mean performance (85.2%); this time *pat* is the highest performer (93.8%) and *constant* a close second (90.6%).

We again refer to effect sizes (Fig. 4) to consider the role of pressure in participant recognition; individuals making different gestures exhibit considerable variation in pressure patterns.

4. DISCUSSION

We discuss our findings in direct response to questions posed in Section 1.1.

Q1a: Potential accuracy of sensor in gesture recognition

Unsurprisingly, we found the highest recognition rate (94.8%) for the case of no covering and a flat, stiff, stationary surface (Study 1); these are the least demanding conditions and the ones we expected to perform the best.

In evaluating the degree to which noise factors degraded performance, we expected the noisiest conditions to be in Study 2: moving, curved, springy surface under a heavy fur cover. This achieved 88.6% recognition rate of our 6 gestures and 'no touch', among the lowest we observed. However, at just under 90%, this value is still usably high. Further work is required to assess the impact of nonuniform motion, as well as unknown gesture segmentation boundaries in lesser controlled conditions.

Q1b: Potential accuracy of sensor in user differentiation

Our studies show that the ability to pick a particular 'toucher' out of a known group varies by gesture. A priori knowledge of a condition also improves prediction accuracy, jumping from 80.3% trained over all data to 89.8% when trained on *in-motion*, *with-cover*, the noisiest condition. To see how this may change over the various gestures, we refer to Fig. 5 which ranks *constant* and *pat* as most identifiable. Fig. 4, which compares the effect size of pressure reading by participant and gesture, reveals that there are many large effects for *constant* gesture. This focus on pressure suggests that there may be revealing variations in individual 'heaviness of hand'.

Q2a: Impact on accuracy due to cover, substrate and motion

Over our two studies, we examined variations in cover thickness, substrate stiffness and curvature, and motion. Summarized in Fig. 2, we now discuss the impact of these factors individually.

Cover: The effect of a cover on classification performance is significant; more so than the underlying motion (as noted by Section 3.3.1). Fig. 2 further illustrates this. Regardless of whether we partition our data by $cover$ on/off or $motion$ present/absent, we achieve gesture recognition of at least 88.1%, 6% higher than training overall (82.1%).

The pressure applied over a denser, heavier fur cover may muffle some of the lighter touches and degrade transmission of touch pressure and/or location, thus confusing some gestures.

Another possible explanation could be from added familiarity that the cover affords. For example, according to one subject, *"When it had the fur on, I had a more pleasant experience...Without the fur, I found it difficult to touch it." (S7)* This opinion was expressed in some form by 10 of 16 Study 2 participants. More research is needed to determine if the fur invited more naturalistic touching.

Substrate: Compared to a flat, hard surface, a flat foam substrate decreased recognition accuracy by about 1% (Fig. 2a). It had slightly less impact than curvature or, comparing to Study 2, than motion. Given the sensor's piezoresistive construction, we anticipated the effect of firmly compliant backing to be small; this finding confirms that a somewhat springy underlying surface (helpful for conveying the sense of an animal body as well as a pleasant tactility) is feasible under a large-body touch sensor.

Motion: The relatively small effect size of motion in raw frame data is unexpected. However, in the context of Tukey's HSD results (with a cover, the motion effect is insignificant), we gain some further insight into just how small the effect of regular periodic motion is, and we confidently rank motion noise behind that of a cover.

This is very promising for the larger premise of reliable touch sensing on a flexing surface.

Interaction of motion and cover: There is a large effect size for the interaction between $cover$ and $motion$, which is absent in recognition performance conditioned on added noise factors (Fig. 2). This consistent improvement over training on all data (overall at 82.1%) suggests that these large effect sizes of noise interference have little effect on recognition as long as we train and test on the same condition.

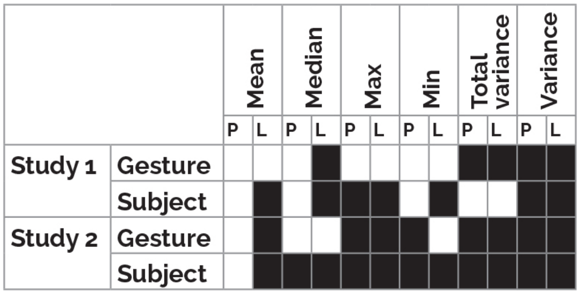

Figure 6: Top features as selected by Weka for each study. Classification tasks are Gesture and Subject, by Location (C_x,C_y) and Pressure features. Features selected at under 25% frequency in 20-fold cross validation are omitted.

Q2b: Gesture Recognizability:
Gesture confusion patterns reveal a considerable range of misclassification (the more saturated cells in Fig. 3). In Study 1, the most commonly misclassified gesture is rub as $tickle$; in Study 2 $scratch$ is most misclassified as rub. Both pairs are commonly executed as quick back-and-forth motions. This may be related to relative gesture pressure by individual: gestures like $constant$, generally more stationary, are predicted consistently and also indicate a larger effect size by pressure (Fig. 4). Quick motions being lost in the heavier covering may also contribute to these errors.

Q3a: Feature Utility
Making computational economy is the key to real-time recognition. Prioritized feature selection allows us to focus on high-performing dimensions. To help understand relative feature utility in our recognition tasks, we used Weka's Attribute Evaluator function to find the highest-weighted features for the random forest model (Fig. 6).

The feature set with the greatest ability to differentiate *gestures* related to pressure variance; meanwhile, location variance facilitated *toucher* recognition. People's touch signatures may vary more in physical location range, but, a gesture may be better characterized using pressure when *toucher* is known (Fig. 4).

These results suggest that using a subset of the features described here could increase computational efficiency, depending on the priority of recognition task needed and the variance exhibited by an actual data pool. Meanwhile, evaluating the performance of a reduced feature set is difficult due to the lack of a benchmark for comparing accuracy rates [15].

Q3b: Computational viability of real-time gesture recognition
The conditions evaluated here approached realism in some respects, specifically that of sensor covering, substrate, and underlying motion. Our post-hoc analysis indicated that a modern microprocessor could keep up with both sampling and recognition.

Our setup fell short of realism in at least one important factor: people are unlikely to perform distinct, discrete gestures with well-defined boundaries. A different computational architecture will be required to handle this problem (a topic of ongoing work). However at present, computational load is dominated by sampling rather than recognition, an overhead cost that will not necessarily change with real-time use (unless more selective sampling can be employed based on observed patterns of touching). It is thus quite likely that a more capable recognition engine will also be feasible with comparable computational resources. In situ real-time recognition may be better approximated by speaker-independent Leave-One-Out (LOO) sliding window. Our work uses k-fold CV as the more conservative accuracy rate (as compared to LOO), as we do expect a calibration process in which speaker behaviour is learned. Until we optimize, the best window size is unknown.

5. CONCLUSIONS

The results described here represent an initial feasibility assessment of the impact of flexing surfaces on gesture recognition performance. We found recognition rates from 80–95% for optimal to noisy conditions when distinguishing between social touch gestures relevant to interacting with a small touch-centric robotic entity. We further found an ability to distinguish individual toucher at a rate of 78.5% and 80.3% in Study 1 and Study 2 respectively. In the noisiest case (also the most realistic), training by condition increased participant recognition accuracy to 89.8%. The next step is evaluating more comprehensive sets of movement conditions.

The implication of a sensing system able to detect both individuality touch and toucher is considerable. For example, a sensor able to differentiate between users could provide a personalized set of experiences or controls.

Further, identifying the touch brings us closer to differentiating affective intent [14]; identifying toucher may allow us to qualify their touch behaviour. A sensor loaded with a personal touch profile could determine how far an individual deviates from that profile on a given day, and infer emotional status. To build such a profile, it will be important to establish the dimensions of a touch signature.

6. FUTURE WORK

We foresee many ways in which to extend this work.

More extensive movement conditions: The present study employed steady periodic motion of an underlying surface for a flexible sensor. A more general, and potentially challenging, environment will include irregular and unexpected motions.

Continuous gestures: The single-gesture samples of this study removed the need to segment data in pre-processing. In future, an algorithm will not know of gesture boundaries or length a priori, and will need to handle the case of seamlessly transitioning gestures.

One approach is to run several sampling windows of different length to search for varying touch activations at the cost of increased computational load. Future work needs to explore this and other architecture to determine a strategy to optimize for computational efficiency.

Pragmatic gestures: In this study, participants were instructed to perform a particular named touch gesture, but not with communicative intent or emotion context. The semantics of a "natural" touch will be dependent on context of situation and the user's own state; to determine communicative intent, it may be necessary to observe other factors as well.

Our participants often varied in how they interpreted a given gesture, both between participants, and individually between and within conditions. For the latter, we suspect users may have performed more authentic gestures on the moving, fur-covered robot than when it was flat, stationary and/or uncovered. We also observed differing touch behaviour from the beginning through the end of one capture, but our sensing mechanisms are unable to distinguish these cases.

Gesture stabilization and system interactivity: Finally, with more efficient algorithms deployable in realistic conditions, we plan a longitudinal study of long-term interactions in natural settings to investigate how individual gestures change over time as a toucher learns to interact with the sensing system.

7. ACKNOWLEDGMENTS

We would like to thank the Natural Sciences and Engineering Research Council of Canada (NSERC), SurfNet and the UBC SURE program for providing partial support of this work. We also thank Michael Phan-Ba for engineering the firmware and robot control, Philippe Kruchten for his mentorship, and finally, Oliver Schneider for the excellent title.

8. REFERENCES

[1] J. Allen and K. E. MacLean. Personal space invaders: Exploring robot-initated touch-based gestures for collaborative robotics. In *Proc ACM/IEEE Int'l Conf on Human-Robot Interaction - Extended Abstracts*, pages 185–186, 2015.

[2] B. D. Argall and A. G. Billard. A survey of tactile human–robot interactions. *Robotics and Autonomous Systems*, 58(10):1159–1176, 2010.

[3] M. R. Banks, L. M. Willoughby, and W. A. Banks. Animal-assisted therapy and loneliness in nursing homes: use of robotic versus living dogs. *J American Medical Directors Assoc*, 9(3):173–177, 2008.

[4] J. Chang, K. MacLean, and S. Yohanan. Gesture recognition in the haptic creature. In *Eurohaptics*, pages 385–391, 2010.

[5] J. Cohen. A power primer. *Psychological bulletin*, 112(1):155, 1992.

[6] K. M. Cole, A. Gawlinski, N. Steers, and J. Kotlerman. Animal-assisted therapy in patients hospitalized with heart failure. *American J of Critical Care*, 16(6):575–585, 2007.

[7] A. Flagg and K. MacLean. Affective touch gesture recognition for a furry zoomorphic machine. In *Proc Intl Conf on Tangible, Embedded and Embodied Interaction*, pages 25–32, 2013.

[8] B. Gleeson, K. MacLean, A. Haddadi, E. Croft, and J. Alcazar. Gestures for industry: intuitive human-robot communication from human observation. In *Proc. ACM/IEEE Int'l Conf on Human-robot interaction*, pages 349–356, 2013.

[9] P. Good. *Permutation tests: a practical guide to resampling methods for testing hypotheses*. Springer Science & Business Media, 2013.

[10] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. The weka data mining software: an update. *ACM SIGKDD Explorations Newsletter*, 11(1):10–18, 2009.

[11] M. J. Hertenstein, R. Holmes, M. McCullough, and D. Keltner. The communication of emotion via touch. *Emotion*, 9(4):566, 2009.

[12] M. J. Hertenstein, D. Keltner, B. App, B. A. Bulleit, and A. R. Jaskolka. Touch communicates distinct emotions. *Emotion*, 6(3):528, 2006.

[13] K. Inoue, K. Wada, and R. Uehara. How effective is robot therapy?: Paro and people with dementia. In *Intl Fed for Medical and Biological Engineering*, pages 784–787, 2012.

[14] H. Y. Joung and E. Y.-L. Do. Tactile hand gesture recognition through haptic feedback for affective online communication. In *Universal Access in Human-Computer Interaction*, pages 555–563. Springer, 2011.

[15] M. M. Jung. Towards social touch intelligence: developing a robust system for automatic touch recognition. In *Proc Intl Conf on Multimodal Interaction*, pages 344–348, 2014.

[16] M. M. Jung, R. Poppe, M. Poel, and D. K. Heylen. Touching the void–introducing cost: Corpus of social touch. In *Proc Intl Conf on Multimodal Interaction*, pages 120–127, 2014.

[17] R. Kohavi et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai*, volume 14, pages 1137–1145, 1995.

[18] P. Marti, A. Pollini, A. Rullo, and T. Shibata. Engaging with artificial pets. In *Proc. Annual Conf European Assoc Cognitive Ergonomics*, pages 99–106, 2005.

[19] K. Nakajima, Y. Itoh, Y. Hayashi, K. Ikeda, K. Fujita, and T. Onoye. Emoballoon. In *Advances in Computer Entertainment*, pages 182–197. Springer, 2013.

[20] P. Rani, C. Liu, N. Sarkar, and E. Vanman. An empirical study of machine learning techniques for affect recognition in human–robot interaction. *Pattern Analysis and Applications*, 9(1):58–69, 2006.

[21] B. Robins, K. Dautenhahn, F. Amirabdollahian, F. Mastrogiovanni, and G. Cannata. Developing skin-based technologies for interactive robots-challenges in design, development and the possible integration in therapeutic environments. In *IEEE ROMAN*, pages xxii–xxiii, 2011.

[22] J. Saldien, K. Goris, S. Yilmazyildiz, W. Verhelst, and D. Lefeber. On the design of the huggable robot probo. *J. Physical Agents*, 2(2):3–12, 2008.

[23] Y. Sefidgar, K. MacLean, S. Yohanan, H. Van der Loos, E. Croft, et al. Design and evaluation of a touch-centered calming interaction with a social robot *in press*.

[24] T. Shibata and K. Wada. Robot therapy: a new approach for mental healthcare of the elderly-a mini-review. *Gerontology*, 57(4):378–386, 2010.

[25] K. B. Shimoga. Finger force and touch feedback issues in dexterous telemanipulation. In *Proc Intelligent Robotic Systems for Space Exploration*, pages 159–178, 1992.

[26] D. Silvera-Tawil, D. Rye, and M. Velonaki. Interpretation of social touch on an artificial arm covered with an eit-based sensitive skin. *Intl J of Social Robotics*, 6(4):489–505, 2014.

[27] W. D. Stiehl, J. Lieberman, C. Breazeal, L. Basel, L. Lalla, and M. Wolf. Design of a therapeutic robotic companion for relational, affective touch. In *ROMAN Workshop on Robot & Human Interactive Communication*, pages 408–415, 2005.

[28] T. Tamura, S. Yonemitsu, A. Itoh, D. Oikawa, A. Kawakami, Y. Higashi, T. Fujimooto, and K. Nakajima. Is an entertainment robot useful in the care of elderly people with severe dementia? *J Gerontology Series A: Biological Sciences and Medical Sciences*, 59(1):M83–M85, 2004.

[29] D. S. Tawil, D. Rye, and M. Velonaki. Touch modality interpretation for an eit-based sensitive skin. In *IEEE Conf on Robotics and Automation (ICRA)*, pages 3770–3776, 2011.

[30] J. Ulmen and M. Cutkosky. A robust, low-cost and low-noise artificial skin for human-friendly robots. In *Proc. IEEE Robotics and Automation (ICRA)*, pages 4836–4841, 2010.

[31] K. Wada, Y. Ikeda, K. Inoue, and R. Uehara. Development and preliminary evaluation of a caregiver's manual for robot therapy using the therapeutic seal robot paro. In *ROMAN*, pages 533–538, 2010.

[32] S. Yohanan and K. E. MacLean. The role of affective touch in human-robot interaction: Human intent and expectations in touching the haptic creature. *International Journal of Social Robotics*, 4(2):163–180, 2012.