Economical Dynamic Surface Sensing: Recognition of Affective Touch and Toucher

1st Author
1st author's affiliation
1st line of address
2nd line of address
1stauthor@place.com

2nd Author
2nd author's affiliation
1st line of address
2nd line of address
2ndauthor@place.com

3rd Author
3rd author's affiliation
1st line of address
2nd line of address
3rdauthor@place.com

ABSTRACT

Social touch is an essential non-verbal channel, whose interactive possibilities can be unlocked by the ability to recognize gestures directed at inviting surfaces. To assess impact on recognition of motion, substrate and coverings, we collected gesture data from a low-cost multi-touch fabric location/pressure sensor, and compared recognition performance. For six gestures most relevant in a haptic social robot context plus a no-touch control, we studied: sensor (1) stationary, varying *substrate* and *covering* (n=10); (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 arbitrary sensed objects.

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*

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

1. INTRODUCTION

A broad variety of interactive applications will be enabled by the capacity to recognize relevant nuances of social touch. In many cases, this requires a sensor with *high flexibility*, to support wearability on the body or embedding on curving, flexing surfaces of interactive objects; and sufficient resolution of *touch pressure* and *location* for a signal processing system to interpret its output in realtime as accurately as the application needs. Flexible pressure sensors have become affordable, and approaches have proliferated with the help of inexpensive electronics and Do-It-Yourself (DIY) guides [19]. However, simultaneously sensing location is more difficult, and wearable applications require that fidelity be balanced with cost. This balance will come through signal processing techniques as well as sensor technologies.

In this paper, we explore the potential of gesture recognition using a highly flexible, low-cost fabric pressure sensor that also localizes input on a grid with fingerpad-scale taxels. To assess the extent to which interference due to movement and other factors interfere with gesture recognition, we first collected touch data for a set of six validated social gestures plus one control [30] on a stationary sensor with 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, and examined the degree to which these factors impacted our ability to distinguish both touch gesture, and identity of the toucher.

Accuracy needs will vary by application. Our recognition rates were within 80-95% for all conditions we tested (chance 14.2%), which will suffice for many purposes and are enough to merit empirical comparison to human recognition ability in future work. At the same time, we found enough idiosyncrasy in individuals' touch signatures to permit identification of toucher within this sample, at an accuracy rate similar to that of the gestures themselves.

1.1 Detailed Requirements

Our goal is to add social-touch recognition capability to surfaces that flex and move in a biological manner, such as human-worn garments that might be touched by others, or interactive robots with malleable, active, touch-inviting bodies and skins.

Movement and elasticity: These surfaces are neither static nor uniform. Therefore, an effective sensor for these environments must be highly flexible, somewhat elastic, and perform well while mounted on non-rigid and/or actuated surfaces.

Pressure range: To read gestures as non-verbal communications, we must accurately interpret a wide range of pressures. While the gestures we measured in the experiments were not delivered with communicative intent, they were chosen because in less constrained circumstances, they could be. Based on a preliminary survey of these touch pressures, we determined that our sensor needed to register touches between 0.005 and 1 kg.

Multitouch: Reading multiple points of contact at once expands the space of intent that can be read [30]; e.g., differentiating *constant* and *rub* from *tickle* and *scratch*.

Resolution and computational cost: Taxel resolution, sampling rate and computational cost must be traded off to achieve usable recognition accuracy. For real-time sensing and recognition, our computational cost is dominated by sensor polling and scales to the number of taxels per grid edge. Our recognition tasks and feature selection explicitly analyze the differences between frames. In this case, recognition accuracy plateaus with fingerpad-scale taxels, when sampled fast enough to capture voluntary movement (peaking at 10 Hz [23]). We must be able to recognize changes in pressure and localized hand and finger motions up to this frequency.

Single-fingerpad resolution (~2 taxels per inch) could capture small fluctuations; however, our gestures (not including our control *no touch*) either involved the flat or palm of hand, (*constant*, *pat*, *rub*, *stroke*) or tended 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 1-inch taxels[8].

Versatility of sensor technology: The sensor must be able to access both the high-pressure surface gestures (as exhibited by a *pat*) as well as the lighter touches (*tickle*) without being lost in the noise of various deformations. But it also must be durable and easy to construct, and withstand stretch under dynamic motion.

1.2 Questions, Contributions and Applications

We were interested in learning:

Q1: how accurate our flexible fabric sensor could be in predicting gesture and differentiating between users,

Q2: determining how the sensor would perform under deformation due to curvature and motion such as that due to a zoomorphic social robot, and

Q3: the computational viability of real-time gesture recognition.

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

- Initial results of deployable accuracy in gesture recognition (6 key gestures and a control): 91.4% on a firm, flat surface, 90.3% on a foam, curved surface, and 88.4% on a foam, curved, moving surface

- And of differentiating toucher (with all 26 participants' data) at 88.8% accuracy.

- Investigation of factors underlying recognition performance including feature selection

- Discussion of feasibility of embedded real-time gesture recognition.

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

In the context of social robots, a flexible sensor could present a way for robots and humans to effectively use gestural touch as a method of communication. Close proximity and direct contact is a requirement for many tasks, such as a robot attendant that engages in affective interaction via a touch on the arm.

However, we foresee applications that go beyond touch-sensitive skin for robots. Accurate gesture recognition on fabric touch sensors opens up gesture-based controls on any electronic device that could incorporate a flexible, low-cost sensor. For example, patients with limited speech capacity could use a smart blanket's gesture recognition capabilities to perform a gesture that could be mapped to a set of requests for comfort or health-reporting purposes.

There are also applications outside of explicit gesture recognition. Inpatients who have been rendered effectively immobile often require nursing staff to regularly shift and rotate their body so as to prevent bed sores. Sheets fitted with pressure and location sensors could help to alert hospital staff to areas that are at high risk of bedsores and require pressure relief.

A robot capable of recognizing touch signatures may be able to predict or influence emotional state [12]. In a behavioural education context, a soft touch-sensing playmate may be able to aide students testing on the autism spectrum by responding to anxious or agitated strokes with slow, soothing, regulated breathing.

2. RELATED WORK

We situate our work in the context of social robots and affective tactile communication. Gestural touch has been identified as a key component of human-robot cooperation [2] wherein a communicative act emerges from nuances in touch data. Should one need to effectively communicate with a robot, naturally occurring non-verbal cues bring a depth to the interaction that are otherwise inarticulable. It would be difficult to teach a novice potter how to throw clay without guiding their hands directly. Each touch could either halt, contribute, or modify the behaviour execution [2]; which can be inferred through the emotional content of the touch [12].

2.1 Uses & Needs of Affective Touch Sensing

Human-Robot collaboration in an industrial or manufacturing setting presupposes a lexicon of social touch for operational interactions [9]. To ensure safe and effective communication, Gleeson et al indentifies the requirements of both a comprehensive gestural dictionary and lightweight sensing technology. Social touch could even extend beyond the factory floors to become collaborative household help as in a homecare assistant [1].

Correlations found between emotional and gestural touch [12] suggest that with sufficient gesture recognition accuracy and nuance, we can detect the toucher's emotional state by this route. With consistently higher accuracy results for within-subject classification than between-subject [8, 13], there is potential for touch behaviours to be used to identify individuals.

Much of the current work on affective touch recognition occurs on a sensor worn by either a human or robotic arm [12, 24, 13, 14]. The arm on an humanoid robot could be a primary source of communicative touch and offers a static surface onto which to affix a sensor. Sensors thus deployed must be flexible enough to wrap around an irregular, padded zoomorphic form.

Animals [7] and interactive robots in animal form (such as Sony's pet-dog AIBO [3, 26], the seal-shaped PARO [29, 11, 16, 22]) suggest potential benefits for mental health. Other touchable social robots include the teddy bear-like Huggable [25]; and Probo [21], which does not have a recognizable Earth animal analogue. However, while real pets respond to complex touch commands anywhere on the body, this has been difficult to achieve before now.

2.2 Flexible Pressure–Location Sensors

The requirement of representing pressure and location despite deformations including stretching, curving, masking etc., restricted construction materials. While many highly accurate pressure-location sensors exist, such as those developed for robot grippers used in dextrous manipulation (e.g., [20, 28]), these tend to be insufficiently flexible, overkill in terms of performance, and considerably too expensive for the objectives outlined here.

Stretch sensors designed for medical purposes by Vista Medical[1] inspired much of our implementation. Vista's sensors recognized only pressure, however, without localization. Further, gesture recognitions require multi-touch capacity.

Using Force-Sensing Resistors (FSRs) affixed to a hard shell restricts the need to account for recalibration of sensor movement; however, the trade-off is difficulty in detecting touch between sensors, limits in rendered motion, and non-aesthetic tactility [5, 2].

Several multitouch, flexible fabric sensors are available [14]. However, flexibility alone does not afford a full range of motion; it must be able to stretch and deform to approximate animal skin.

The design and sensing capabilities described by Flagg et al [8] informed many of our requirements and suggested that the bulk of the recognition accuracy could be achieved by the "below surface" sensor alone. However, this study did not consider the full design space of a robot in motion including a non-sensing cover 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.

Still more comprehensive is the collection of touch sensor projects developed by Perner-Wilson [19]. The demonstrations depicted on the Plusea site[2] boast a range of textile-based sensors with a myriad of purposes, including a stroke sensor featuring conductive threads and a flexible neoprene sensor that could recognize when the surface was being bent.

While many of these designs addressed recognizing touch contact and/or position, we wanted to evaluate the quality of the touch recognition on a sensor employing both position and pressure by testing how well the data held up to gesture classification.

3. STUDY

We hypothesized that:

H1: *gesture recognition rates will decrease with increase in 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. Between them are layers of a plastic standoff mesh separator and a sheet of Eeontex SLPA 20K Ω resistive fabric. The resistance value across a given taxel drops when pressure is applied to that taxel, compressing the mesh separator such that the conductive layers more closely approach each other.

A circuit is constructed using an Arduino Mega microprocessor. Each strip of the fabric is connected to a single I/O pin, where the top layer is connected to analog input pins, and the bottom layer is connected to digital output pins as seen in Figure 1.

The sensor is polled by sending a high voltage through each of the bottom layer's digital pins in a sequence. When the top layer contacts the bottom layer, a signal is sent to the analog pins which read the amount of current. Resistance (therefore current) varies with pressure. In this way, the sensor can read location and pressure, i.e. which of the top layer's rows are being activated per bottom layer column, and by how much.

[1]Stretch based sensors can be purchased commercially from Vista Medical http://www.vista-medical.com/subsite/stretch.php
[2]Many wearable sensors are displayed at http://www.plusea.at/
[3]Fabric was purchased through Eeonyx at www.eeonyx.com
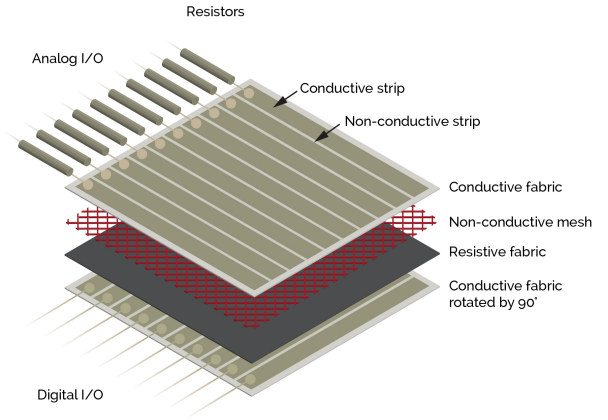
Figure 1: The fabric pressure sensor constructed out of Eeon-Tex conductive fabric and wired to an Arduino Mega microprocessor.

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

Dynamic range is modulated through choice of resistor value. We found that values $> 1K\Omega$ 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 aimed to assess the effects of different conditions interfering with the sensor; gestures and data collection method were held constant.

3.2.1 Gestures and Sampling

We selected gestures from Yohanan et al's social touch dictionary [30], choosing items most appropriate for human-animal interactions (Flagg et al [8]). The sensor was placed on a table in front of a seated participant, with a reference sheet with very general definitions for six selected gestures and one control (Table 1). Participants were asked to interpret each gesture as they saw fit; no further performance clarifications were provided.

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

Table 1: General suggested guideline for touch gestures shown as reference.

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

As a baseline, we first evaluated the impact on recognition accuracy of the sensor's substrate stiffness and curvature and covering thickness, under static conditions (unmoving surface). This produced a factorial design of $4 \times 3 \times 7$ ($covering \times substrate \times gesture$), using gestures listed in Table 1.

Covering: the cover fabric's pile or density varied from no cover at all (direct contact with the sensor) to a very long, very 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 $covering \times substrate$ conditions. Condition order was randomly generated for every participant, and gesture order was further randomized over each condition set. That is, we randomly generated a *masking* condition set (e.g., one set was fur with flat foam) and ran all ⌈**KM** *what does 'seen' mean here?*⌋ seen gestures in a randomly generated order before changing the masking condition and running another full gesture set, with a new random ordering of gestures. All participants completed all twelve ⌈**KM** *masking??*⌋ conditions, with each generating 48 2-second sample windows per gesture.

A study session took approximately 50 minutes to complete. 10 volunteers (4 female, 6 male) with cultural backgrounds from Canada, England, Southeast Asia, and the Middle East were compensated with $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 $motion$ = {breathing, not breathing}, $cover$ = {cover, no cover}, 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 (Figure 2(a)), a cat-sized robot designed for therapeutic use (Figure 2(b)). The robot's 'breathing' motion was created by extending and contracting the two ribs assemblies in a $14°$ arc from the spine at 0.5 Hz (Figure 2(c)). We draped and pinned fabric over the sensor, approximating a full fur jacket for condition randomization while limiting sensor disruption (Figure 2(d)),

Each session began by asking the participant to interact freely with the covered, moving robot for XX minutes, to reduce novelty. Each condition was then presented randomly twice, 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 2-second samples of each gesture for every condition.

3.3 Analysis and Results

We discarded the first and last second of each 10s, XXHz gesture capture and divided the remaining 8s into 4 2s windows. The 2s window (108 frames) was chosen to allow each gesture some periodicity; all fit completely within 1s (Flagg [8]). Further, 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 gestures that move across the sensor, we also calculated a centroid (average geometric centre) weighted based on the measured pressure reading for each frame. Each centroid was defined by 2 dimensions: a row-position C_x (1) and a column-position C_y (C_y is calculated 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 rows, multiplying by index, and dividing by the unweighted frame sum (the sum of the full sensor reading from a given time sample). We repeated the procedure for each column, obtaining a tuple of frame sum and centroid for each frame.

We calculated seven features across these three dimensions (frame value, C_x, C_y) for each 2s window (108 frames) 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.

We use Weka, an open-source application that applies common machine learning algorithms to our classification problem [10]. We ran k-fold cross validation on participant data from Study 1 for k= 5, 10, 20, 100, 200) and found that improvement between 20- and 100- folds was within 1% on average with some decreasing performance. A previous comparison between random forest and a number of other algorithms showed that random forest performed best in gesture recognition of this type [8]. All reported classification performance is therefore based on 20-fold cross validation of random forest models. Accuracy is defined as the percentage of data instances that are correctly classified.

To assess relative noise levels against the control condition, we report t-statistics and Cohen's d effect size (negligible $|d| < 0.2$, small $|d| < 0.5$, medium $|d| < 0.8$ or large $|d| \geq 0.8$) [6].

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 roughly ∼88%).

When we considered training by individual, Study 1 showed little change in gesture prediction rate compared to models trained on all data. Study 2 results are similar to other studies of social touch, which also report training on data drawn from the same condition. (??)! [14, 8, 18, 27].

Cover-substrate-motion: Fig. 3 shows overall gesture recognition accuracy by study and condition sets. A comparison of the relative noise contributions by $cover$ and $substrate$ conditions of Study 1 are given in Table 2. In Study 2, we compare $motion$ and $cover$ as binary conditions, with effect sizes in Table 3.

A confusion matrix sheds some light onto which gestures are being misclassified as seen in Figure 4(a). We see that in both studies, the most commonly misclassifed gesture is $tickle$.

Participant: We performed gesture classification with models trained by participant. In Study 1, mean accuracy was 89.1% with a maximum of 94.6%[4]. Training on all Study 1 data was ⌈**KM** *do you mean: Models trained on the full Study 1 dataset were ...?*⌋ within a percentage point of the mean accuracy of the models trained
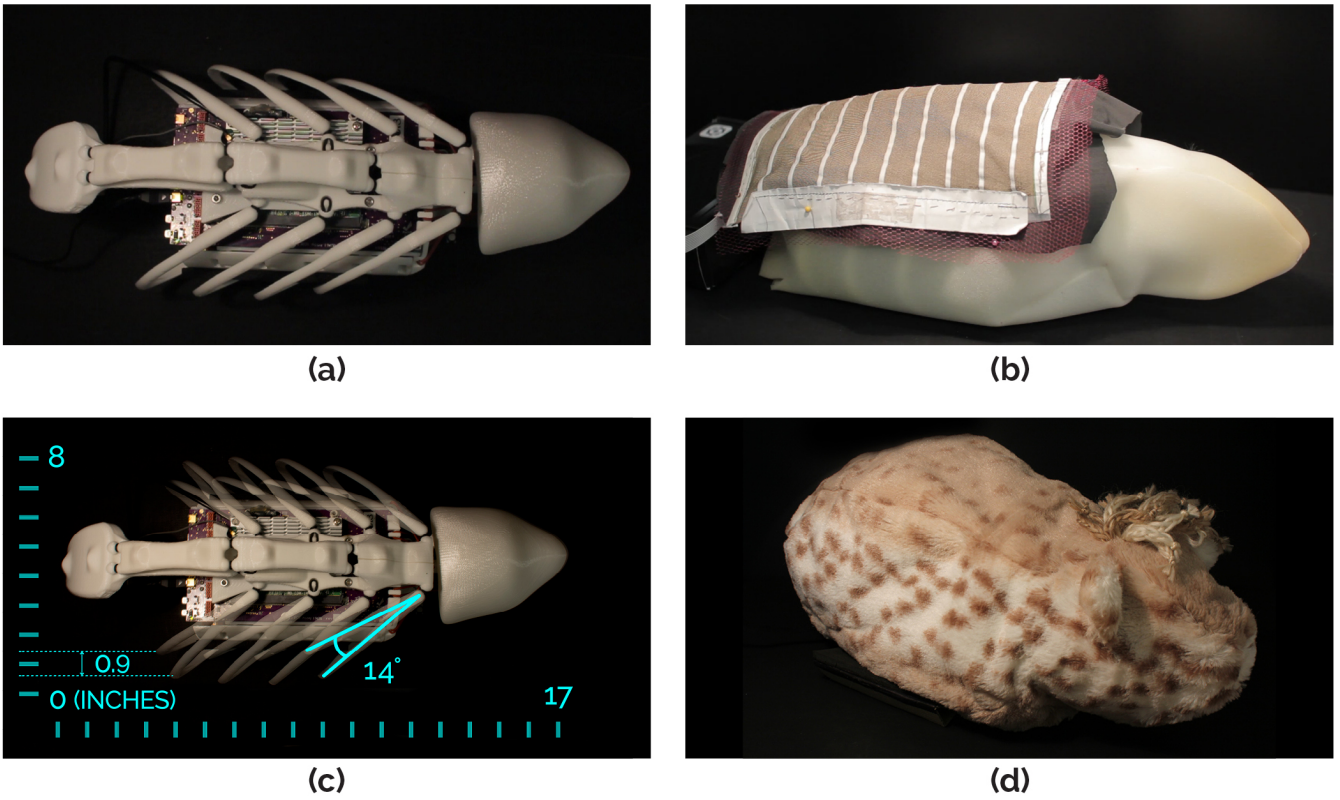
Figure 2: (a) Top view of the CuddleBot skeleton. (b) Touch sensor, pinned to foam substrate wrapped around the skeleton and corresponding to a *No touch* × *No motion* × *No cover* condition. (c) Full range of breathing motion used. (d) The fully-covered robot; a cover of nearly identical material was used in the study to facilitate quick condition changes.

Participant	Gesture prediction accuracy	Motion × cover
S1	90.2	80.9
S2	86.6	86.2
S3	86.6	86.2
S4	91.1	91.6
S5	81.3	80.9
S6	86.1	73.7
S7	84.8	88.0
S8	79.5	88.0
S9	95.5	95.1
S10	79.5	86.2
S11	90.2	89.8
S12	93.8	95.1
S13	97.3	93.4
S14	83.0	91.6
S15	79.5	64.8
S16	79.5	79.1
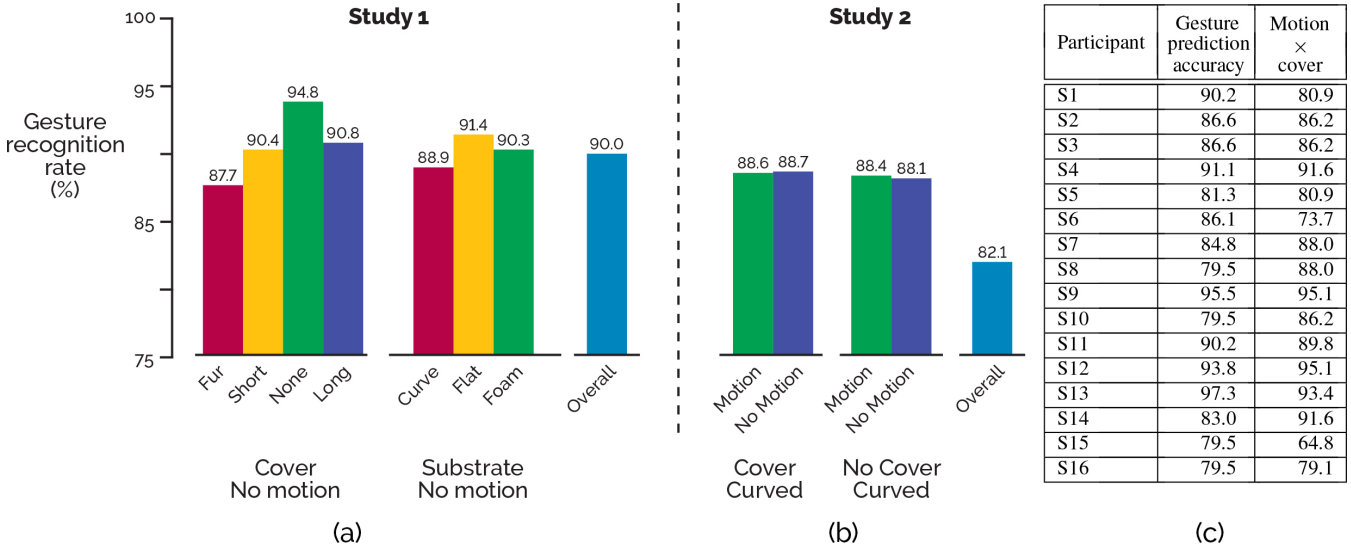
Figure 3: A comparison of gesture prediction accuracy rates with added pressure noise when varying *substrate* or *cover* in Study 1. Means are different at a significance level $p < 0.001$. Each bar represents an average over 10 trials; error bars are omitted as the accuracy rate change across the trials amounted to under 0.001% in each case.

on individuals (90.0%), indicating that training on participants did not improve recognition when data was not conditioned on noise-creating factors.

For Study 2, which had fewer noise factors, a greater effect was

---

[4] ⌈**KM** *better in a small table, in row with other small tables?*⌋

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%

	no substrate vs		no cover vs		
	foam	curve	fur	long	short
t-statistic	-48.0	-256.6	-89.0	-83.1	-60.1
cohen's d	-0.5	-2.6	-1.0	-0.9	-0.1

Table 2: A comparison of pressure readings varying *substrate* **or** *cover* **in Study 1 ($p < 0.001$). Each bar represents an average of 10 trials; error bars are omitted because accuracy rate change across trials was $< 0.001\%$ in each case.**

	A	B	C	D	E
t-statistic	-7.7	-123.4	-74.2	-6.9	-89.5
cohen's d	-0.06	-0.9	-0.8	-0.08	-1.0

Table 3: A comparison of pressure readings by motion and cover conditions. A: *motion vs. no motion*; **B:** *cover vs. no cover*; **C:** *cover vs no cover | no motion*; **D:** *motion vs no motion | no cover*; **E:** *no cover & no motion vs cover & motion*.

seen for models trained on participants. Training across all data gives 82.1% accuracy; models trained by participant averaged 86.5%, maximum 95.5% [5].

The $motion \times cover$ condition had an overall 79.4% recognition rate. Surprisingly, the subset of data with what we expected to be the most challenging conditions (in-motion, with-cover) produced the highest accuracy for individual-trained models with mean accuracy of 85.7% [6]. ⌈**KM** *verify*⌉

Comparisons of mean pressure of gesture behaviours by individual and that of the entire pool were made (that is, how P1 performed *scratch* versus how all participants performed *scratch*); all incidences were significant at $p < 0.05\%$. We report the Cohen's d effect sizes in Fig. 4(b).

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.

Being able to recognize the 'toucher' may have great impact on reading emotional state. We compared performance in subject 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, we see a mean of 87.9%. The highest contributing gesture is *constant* at 92.7%, followed by *pat* at 88.9%.

Training on all data from Study 2, we were able to recognize participants at a rate of 80.3%. Training by gesture again shows an improvement; the highest performer is *constant* at 93.8% followed by *pat* at 89.8%. Mean accuracy rate is 85.4%

Conditioning on only the *in-motion, with cover* factor in Study 2, average recognition rates of participants are 89.8%. Further splitting data to additionally train models by gesture does not provide a additional improvement in mean performance (85.2%); *constant* again is the highest performer (90.6%).

We can also cross-reference Fig. 4(b) to consider how pressure can play a role in recognition of participant; considerable variation is shown in pressures patterns exhibited by different individuals in making different gestures.
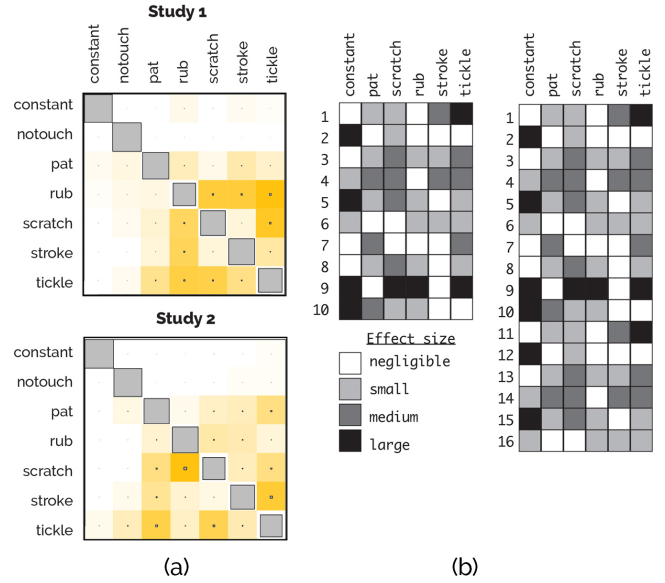
Figure 4: A Hinton diagram depicting the confusion matrix of gesture classification: the horizontal (row) gestures is classified as the vertical (column) gesture. The relative tint of the colored squares are representative of misclassification rates.
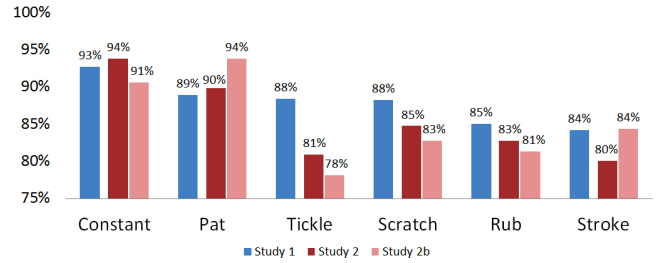
Figure 5: Subject recognition rates by gesture and study. Study 2b refers to the 'in-motion, with cover' condition.

4. DISCUSSION

We organize our discussion around the research questions which guided this investigation.

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. Our real question was the degree to which various factors that increase the realism of operating conditions impact this top performance.

The most realistic conditions we tested, and the ones we expected to perform the most poorly, were 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 on recognition of nonuniform motion, as well as unknown gesture segmentation boundaries in less well controlled conditions.

Q1b: Potential accuracy of sensor in user differentiation

Our studies show that user recognition in picking a particular 'toucher' out of a known group varies by gesture. A priori knowledge of a condition also improves the prediction accuracy, jumping from

constant	1.7	0.1	0.7	0.6	0.4	■
pat	1.5	-0.2	0.6	0.4	■	
rub	1.3	-0.4	0.4	■		
scratch	0.8	-0.7	■			
stroke	1.6	■				
tickle	■					

Table 4: MISSING CAPTION

80.3% trained over all data to 89.8% when trained on $in-motion$, $withcover$, the noisiest condition. Evaluating how this may change over the various gestures, we refer to Figure 5 which ranks the gestures $constant$ and pat as the highest performing. Figure 4(b), which compares the effect size of pressure reading by participant and gesture, reveals that there are many large effects seen in $constant$ gesture. This focus on pressure suggests that there may be revealing variations in 'heaviness of hand' by individuals.

Q2: 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. These factors are summarized in Figure 3, and below we discuss the impact of each factor individually.

Effect of cover: We found that the effect of having a cover on classification performance is significant; and more so than the underlying motion as in Table 3.

Figure 3 further illustrates the effect of the cover, which was a variable for both studies. Regardless of whether we partition our data by $cover$ on/off or $motion$ present/absent, we achieve a gesture recognition rate 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 cause a loss of information in touch pressure and/or location, thus confusing some gestures.

Another possible explanation could be from the variability in ecological validity 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 the 16 participants in Study 2. More research is needed to determine if the fur caused more naturalistic touches.

Effect of substrate: Compared to a flat, hard surface, a flat foam substrate decreased recognition accuracy by about 1 percentage point (Figure 3). It had slightly less impact than curvature or, comparing to Study 2, than motion. Given the sensor's piezoresistive construction, we anticipated this kind of small effect of springy backing up to some point, and this finding confirms that a somewhat compliant 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.

Effect of motion alone: The relatively small effect size of motion in raw frame data is unexpected, especially when considering the recognition improvement when conditioning over motion, and varying cover presence.

It may be that the periodic nature of robot breathing produced a similar effect to that of the cover. For example, an average over a noisy window could appear similar to a lower-pressure constant interference.

Interaction of motion and covering:

Interestingly enough, a large effect size is seen in the interaction between $cover$ and $motion$, however, we don't see this reflected in the gesture recognition performance conditioned on the added noise factors as per Figure 3. Based on this consistent improvement over training on all data (Overall at 82.1%),this suggests that these large effect sizes of noise interference has little effect on recognition so long as we train and test on the same condition.

Q3a: Gesture Recognizability
Comparing how often gestures are confused with others reveals that there is a much higher misclassification rate for certain gestures than others (represented by the more saturated cells in Figure 4(a). In Study 1, the most commonly misclassified gesture is rub as $tickle$ while in Study 2 is $scratch$ classified as rub. Both sets of gestures are quick, fast, back and forth motions. This may be related to relative gesture pressure by individual. Gestures like $constant$ which require less moving around are predicted much more consistently, and as Figure 4(b) shows, a larger effect size of pressure. This suggests that the quick motions are lost in the heavier covering conditions.

Q3b: Feature Utility
In a realistic setting we aim to be able to perform real-time recognition, requiring that computational economy. Prioritized feature selection allows us to focus on high-performing dimensions.

To help understand the relative utility of the features we used in our various recognition tasks, we used Weka's Attribute Evaluator function, to pick the highest-weighted features for the Random Forest model (Table presented in Table 5.

The feature set which offered the largest differentiation abilities in $gesture$ recognition tasks related to pressure variance; meanwhile, location variance dominated the task of recognizing the $toucher$. People's touch signatures may vary more in physical location range; however, a gesture may be better characterized using pressure when $toucher$ is known (as evidenced in Figure 4(b).

These results suggest that a subset of the features we used here could usefully be extracted to 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 [13].

Q4: Computational viability of real-time gesture recognition
The conditions evaluated here approached real-world use in some respects, specifically that of sensor covering, substrate and to some degree, underlying motion. While computation of the segmented gestures performed by our subjects was not performed in realtime, our post-hoc analysis indicated that reasonably high performance modern embedded microprocessor (such as XX) 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 our ongoing work). However, at the present, computational load is dominated by sampling rather than recognition, an overhead cost that will not necessarily change with realtime use (unless more selective sampling can be employed based on observed patterns of touching). We thus consider it quite likely that a more capable recognition engine will also be feasible with comparable computational resources.

5

5. CONCLUSIONS AND FUTURE WORK
The results described here represent a first feasibility assessment of the impact of mobile surfaces on gesture recognition performance. We found recognition ranges between 80-95% for conditions ranging from optimal to rather adverse, when distinguishing

		Study 1		Study 2	
		G	S	G	S
Variance	L	■	■	■	■
	P	■	■	■	■
Total Variance	L	■		■	■
	P	■		■	■
Min	L		■		■
	P			■	■
Max	L		■	■	■
	P		■	■	■
Median	L	■	■		■
	P			■	■
Mean	L		■	■	■
	P				

Table 5: Top features as selected by Weka for each study. Classification tasks are Gesture and Subject, by Location and Pressure features. Features selected at under 25% frequency in 20-fold cross validation are omitted.

between 7 social touch gestures found to be most important interacting with a small touch-centric robotic entity. We further found an ability to distinguish individual toucher at a rate of XX %. These results are highly encouraging and support taking the next step with a more comprehensive set of movement conditions.

The implication of a sensing system able to detect both individualization of toucher and the type of touching being done is considerable. For example, a sensor able to differentiate between users could provide a personalized set of experiences or controls, which has the potential to create a very powerful application for a very low-resolution, low-cost sensor hardware.

Further, with a sensor that can identify both the 'touch' and the 'toucher', differentiating affective intent is not a far leap [12]. For example, a sensor could build a personal touch profile, determine how far the individual is deviating from that profile on a particular day, and infer current emotional status from the difference. To properly build such a profile, it will be important to clarify the dimensions that constitute a 'touch signature'.

5.1 Future 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 be irregular motions, including sudden ones.

Continuous gestures: Our participants made single gestures, sparing us the need to segment our data in pre-processing. In reality, an algorithm will not know a priori of gesture boundaries or length, and will also need to handle the case of different gestures that are seamlessly connected.

We have implemented one possible limited real-time gesture recognition engine. However, our current model is built on 2-second windows of the same gesture. This has two drawbacks: in real conditions, we do not know when a gesture will start and end. Secondly, it potentially misses the nuances that might expose differentiating characteristics. One approach is to run several different sampling windows of different length, to search for touch activations of varying extents (but at the cost of increased computational load). Future work needs to explore this and other architectures, and optimize them for computational efficiency.

Pragmatic gestures: In this study, participants *were* instructed as to gesture, but not in 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.

In the present study, participants often varied in how they interpreted a given gesture, both between-participants for a given condition, and within-participant between conditions. For example of the latter, we suspect that users may have performed more authentic gestures on the zoomorphic robot covered with fur and moving compared to a surface that was flat, stationary and/or uncovered; but our sensing mechanisms are unable to distinguish between differences in the actual gesture, and our ability to accurately sense whatever gesture was made due to experiment condition.

Gesture stabilization and impact of system interactivity: Finally, with more efficient algorithms deployable in realistic conditions, we will need to conduct a longitudinal study where we collect data over long-term interactions with a sensor in a natural setting to investigate (for example) how individual gestures may change over time as the toucher learns and is encouraged to interact with the system being sensed.

6. REFERENCES

[1] J. Allen and K. E. MacLean. Personal space invaders: Exploring robot-inulated 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] M. Bhuyan, A. Kumar, K. MacDorman, and Y. Iwahori. A novel set of features for continuous hand gesture recognition. *J. Multimodal User Interfaces*, 8(4):333–343, 2014.

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

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

[7] 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.

[8] 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.

[9] 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.

[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] 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.

[12] 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.