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

4th Author
4th author's affiliation
1st line of address
2nd line of address
4thauthor@place.com

5th Author
5th author's affiliation
1st line of address
2nd line of address
5thauthor@place.com

6th Author
6th author's affiliation
1st line of address
2nd line of address
6thauthor@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 performance of sensor motion, substrate and coverings, we collected gesture data from a low-cost multi-touch fabric location/pressure 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 *covering* (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 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 real-time 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 [17]. 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 [28] 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 (Fig. 1), and examined how these factors impacted 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.

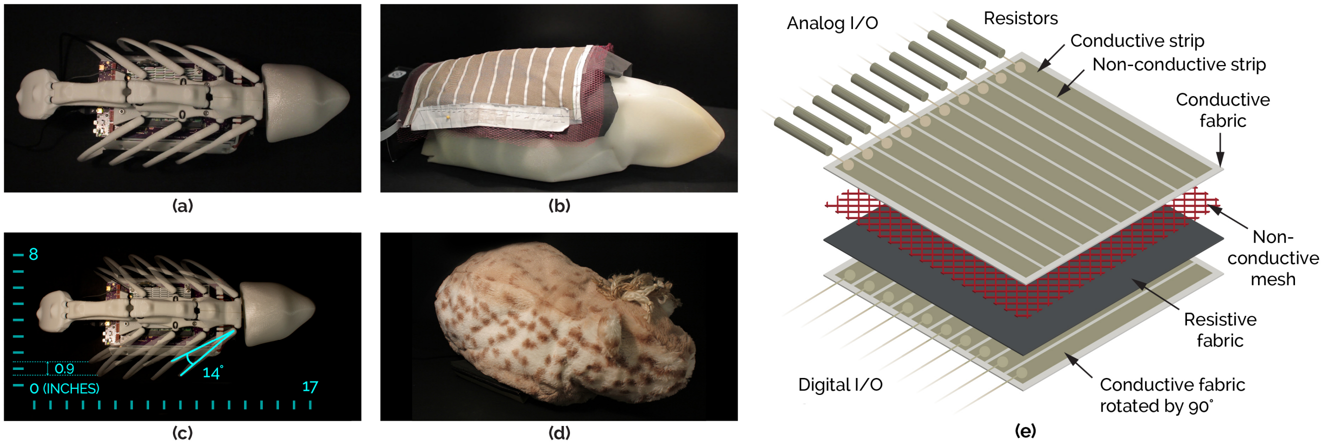


Figure 1: (a) Top view of the CuddleBot skeleton. (b) Touch sensor, pinned to foam substrate wrapped around the skeleton and corresponding to a *No touch* $\times$ *No motion* $\times$ *No cover* condition. (c) Full range of breathing motion used. (d) The fully-covered robot; a cover of nearly identical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

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 [28]; 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 [21]. We must be able to recognize changes in pressure and localized hand and finger motions up to this frequency.

Single-fingerpad resolution ($\sim$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 $\sim$1-inch taxels[7].

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 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.

Our study compares gesture recognition performance across a variety of conditions which approach real-time dynamic gesture recognition. Toucher recognition accuracy shows promise for incorppo-rating 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. E.g., bedridden individuals often require a nurse to shift and rotate their body so as to prevent bed sores. Sheets fitted with pressure and location sensors could alert caregivers of bedsore risk.

A robot capable of recognizing touch signatures may be able to predict or influence emotional state [11]. 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 [11].

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 [8]. 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 [11] 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 [7, 12], 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 [11, 22, 12, 13]. 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 [6] and interactive robots in animal form (such as Sony's pet-dog AIBO [3, 24], the seal-shaped PARO [27, 10, 14, 20]) suggest potential benefits for mental health. Other touchable social robots include the teddy bear-like Huggable [23]; and Probo [19], 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., [18, 26]), 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 [2, 4].

Several multitouch, flexible fabric sensors are available [13]. 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 [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, 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 [17]. The Plusea site[2] boast a range of textile-based sensors with myriad purposes, including a stroke sensor featuring conductive threads and a flexible neoprene bend sensor.

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, with a plastic standoff mesh separator and a sheet of Eeontex SLPA 20K Ω 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 easily able to achieve a touch weight range of 0.005 to 1 kg using $1K\Omega$ resistors. Under the most severe conditions, 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 assessed how more realistic conditions impacted sensor data and hence recognition accuracy; gestures and data collection were held constant.

3.2.1 Gestures and Sampling

We selected gestures from Yohanan et al's social touch dictionary [28], choosing items most appropriate for human-animal interactions (Flagg et al [7]). 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.

[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/

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: Touch gesture instructions.

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.

As a "baseline" for both studies, we sampled sensor frames in the absence of gestures. In Study 1, each of the 12 ($4cover \times 3substrate$) condition sets contributed 4320 frames; in study 2, each of the 4 ($2motion \times 2cover$) condition sets contributed 6912 frames. To establish the effect of noise under each condition, we ran MANOVA over 3 dependent variables: pressure, x-coordinate, and y-coordinate. The 6 gestures (omitting no-touch data) can then be compared with each other under the conditions of each study.

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 $4 \times 3 \times 7$ ($covering \times substrate \times gesture$), using gestures listed in Table 1.

$Covering$: The cover 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 $covering \times substrate$ 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 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 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 (Fig. 1(a)), a cat-sized robot designed for therapeutic use (Fig. 1(b)); [1]. 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 (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 randomly twice, for a total of ((2×2×7) + 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, 54Hz 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 [7]). 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 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):

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

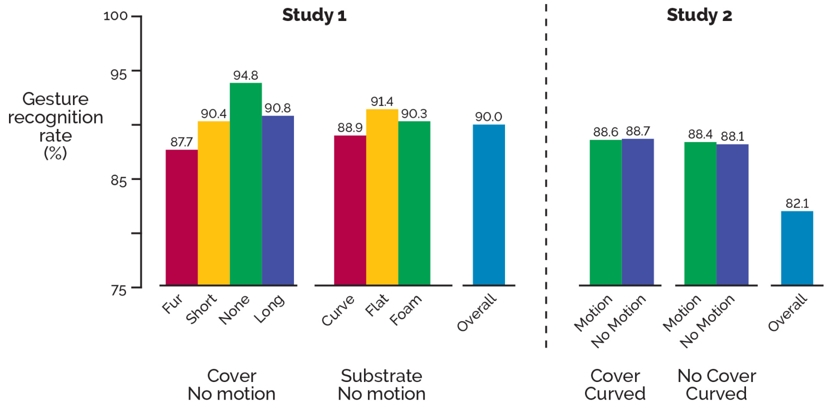
We calculated weighted *pressure* by summing readings across each row, multiplying by index, and dividing by the unweighted frame sum (the sum of the full frame sensor reading). Repeated for each column, this provided a tuple of frame sum and centroid per frame.

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 used Weka, an open-source application that applies common machine learning algorithms to our classification problem [9]. We ran k-fold cross validation on Study 1 participant data for k= 5, 10, 20, 100, 200) and found improvement between 20- and 100- folds 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 [7]. 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.

Results of MANOVA tests are shown in Table 2. We report Pillai's trace in place of the F-statistic [16]. 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.

To verify significance of various factors for frame-level dependent variables (prior to computing features, these were simply pressure and x,y taxel coordinates), we assessed relative noise levels against the baseline condition (Tables 3 and 4). The largely significant comparisons were identified with pairwise t-tests with a Bonferroni adjustment. We also report Cohen's d effect sizes, interpreted as: negligible to large, $|d|$<0.2 to $|d| \geq 0.8$ respectively) [5].



(a) (b)

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

(c)

Figure 2: (a) Mean gesture prediction accuracy rates with added pressure noise when varying $substrate$ or $cover$ in Study 1. Each bar represents an average accuracy rate over 10 trials; error bars omitted (Δ accuracy rate change across trials $< 0.001\%$ in each case). (b) Combined $motion$ and $cover$ impact on pressure noise (Study 2, $p < 0.01$.) (c) Gesture prediction accuracy when trained by participant under the $breathing, with fur$ condition (Study 2).

	factors	Pillai	df	Sig. metrics
Study 1 baseline†	substrate	0.976	2	p,x,y
	cover	0.376	3	p,x,y
	substrate*cover	0.285	6	p,x,y
Study 1‡	gesture	0.418	5	p,x,y
	subject	0.549	9	p,x,y
	gesture*subject	0.533	45	p,x,y
Study 2 baseline†	motion	0.007	1	p,x,y
	covered	0.185	1	p,x,y
	motion*covered	0.006	1	p, y
Study‡	gesture	0.387	5	p,x,y
	subject	0.443	15	p,x,y
	gesture*subject	0.450	75	p,x,y

Table 2: Significant multivariate effects (p<0.001). Metrics are p: pressure, x: row coordinate, and y: column coordinate. † Baseline tests for significance of noise under no touch control gesture across all conditions. ‡ Tests for significance of subject and gesture across all conditions.

	no substrate vs		no cover vs		
	foam	*curve*	*fur*	*short*	*long*
cohen's d	-0.5	-2.6	-1.0	-0.9	-0.1

Table 3: Effect size of Study 1 pairwise comparisons of pressure when varying $substrate$ or $cover$ factors (p<0.001).

factor 1: *versus*	motion+	cover+	cover+	motion+	cover+; motion+
factor 2:	motion-	cover-	cover-	motion-	cover-; motion-
data set:	cover (+,-)	motion (+,-)	motion (-)	cover (-)	cover (+,-); motion (+,-)
cohen's d	-0.06	-0.9	-0.8	-0.08	-1.0

Table 4: Effect size of Study 2 pairwise comparisons of pressure, with motion and cover conditions. +/- indicates factor presence / absence (e.g., motion+ & cover- represents *with motion, without cover*). All comparisons significant at p<0.001.

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 $\sim$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, 13, 15, 25].

Cover-substrate-motion: Fig. 2 shows overall gesture recognition accuracy by study and condition set. Table 3 shows noise effect sizes of $cover$ and $substrate$ for Study 1; Table 4 shows effect sizes of Study 2 $motion$ and $cover$ (both now binary). 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(a)) indicates how gestures were misclassified. In both studies, the most-misclassifed was $tickle$.

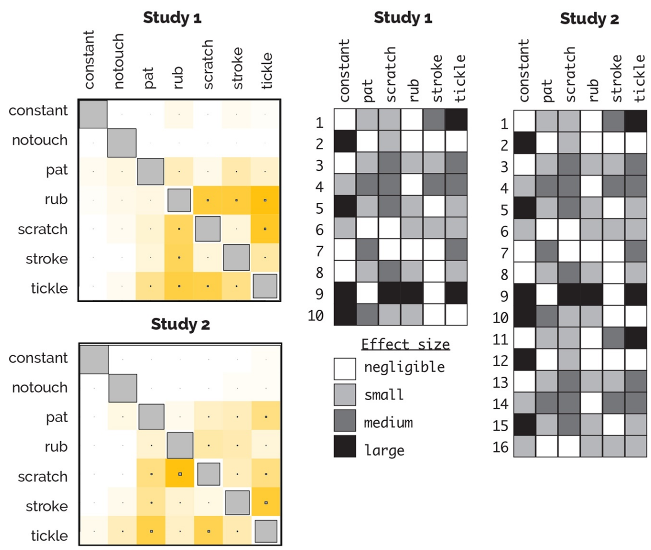
Participant: We classified gestures with models trained by participant. In Study 1, mean accuracy was 89.1% (max 94.6%[4]). Models trained on all Study 1 data were accurate at 90.0%, i.e. within 1% of the mean accuracy of the individual-trained models. This indicated that training on participants did 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 (86.5% (max 95.5%)[5]. 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%).

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% (see Cohen's d effect sizes in Fig. 3(b)).

[4]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%



(a) (b)

Figure 3: (a) 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. (b) Cohen's d effect sizes of participant by gesture for each study.

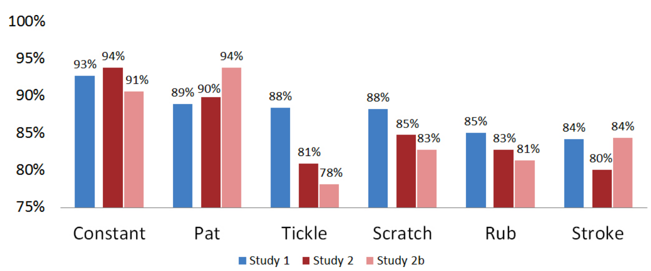


Figure 4: Mean subject recognition rates by gesture and study over 10 trials; error bars omitted (Δ accuracy rate change across trials $< 0.01\%$ in each case). Study 2b refers to the 'in-motion, with cover' condition.

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. 4. 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 improved things; *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 in Study 2,

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

(a)

	constant	pat	rub	scratch	stroke	tickle
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	■					

(b)

Figure 5: (a) 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. (b) Cohen's d effect sizes, pairwise comparisons between gesture pressures.

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 consider effect sizes (Fig. 3(b)) to consider the role of pressure in participant recognition; individuals making different gestures exhibit considerable variation in pressure patterns.

4. DISCUSSION

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 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. 4 which ranks $constant$ and pat as most identifiable. Fig. 3(b), 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 (Table 4). Fig. 2 further illustrates the cover effect, which was a variable for both studies. 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 / 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. 2). 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, and 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 motion is, and we confidently rank motion noise behind that of a cover.

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

Interaction of motion and covering: 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.

Q2b: Gesture Recognizability:

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

Q3a: Feature Utility

In a realistic setting we aim to perform real-time recognition, making computational economy key. 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. 5(a)).

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. 3(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 [12].

Q3b: Computational viability of real-time gesture recognition

The conditions evaluated here approached realism in some respects, specifically that of sensor covering, substrate and to a degree, underlying motion. While computation of the segmented gestures performed by our subjects was not performed in real-time, our post-hoc analysis indicated that a moderately high-performance modern embedded 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.

5. CONCLUSIONS AND FUTURE WORK

The results described here represent a first feasibility assessment of the impact of flexing surfaces on gesture recognition performance. We found recognition ranges between 80-95% for conditions ranging from optimal to quite adverse, when distinguishing between social touch gestures found to be most important interacting with a small touch-centric robotic entity (six plus no-touch). 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 most adverse conditions (also the most realistic case), knowing the conditions increased participant recognition accuracy to 89.8%. 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, identifying both the 'touch' and the 'toucher' is not far off from differentiating affective intent [11]. A sensor could build a personal touch profile, 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.

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

Herein we describe one possible limited real-time gesture recognition engine. However, this model is built on 2s 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.

Our participants often varied in how they interpreted a given gesture, both between participants, and individually between 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; 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 (for example) how individual gestures change over time as a toucher learns to interact with the system being sensed.

6. 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] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

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

[16] C. L. Olson. On choosing a test statistic in multivariate analysis of variance. *Psychological Bulletin*, 83(4):579, 1976.

[17] H. Perner-Wilson and M. Satomi. Diy wearable technology. In *ISEA 15th Intl Symp on Electronic Art*, 2009.

[18] 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.

[19] 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.

[20] 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.

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

[22] 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.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] 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.

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