Enabling Low-Cost Full Surface Tactile Skin for Human Robot Interaction

Xiaoran Fan, Daewon Lee, Larry Jackel,
Richard Howard, Daniel Lee and Volkan Isler

***Abstract*— Realizing full coverage, low-maintenance, and low-cost tactile skin is one of *de facto* design dreams since the invention of robots. It ensures safety and enables collaborative work protocols for human robot interactions (HRI). The on-robot tactile capability is realized by deploying an array of external sensors or inferring from proprioceptive information that comes with the robot, such as motor torque. However, these methods may be cumbersome, introduce extra management cost, expensive, lack real-world robustness, or require special robot designs. In this paper, we present *SonicSkin*, a low-cost ($< \$2$) and easy to deploy system that localizes the on-robot human touch and estimates the touch pressure without actually attaching sensors at potential touch locations. The system requires only a single pair of piezoelectric transducers (*i.e.* one transmitter and one receiver) attached on the target robot and turns the robot itself into a versatile sensor. We present a set of novel algorithms to progressively address the unique challenges posed by our system design. We put together an end-to-end *SonicSkin* system on a Jaco robot arm that runs in real-time, and conducted an extensive real-world study including 57019 actual evaluation datapoints under various challenging conditions from 12 human subjects. *SonicSkin* achieves less than 2 cm localization error for 96.4% of touches, with more than 96.7% cross-correlation similarity between the predicted touch pressure and the ground truth touch pressure.**

I. Introduction

Humans can explore the unknown physical world by feel, weigh and grasp when identifying the location of the contact. Also, interpersonal touch is a fundamental but undervalued aspect of human nature. Similarly, tactile sensing plays a key role in human robot interactions (HRI), these include safe robot operation around humans, providing emotional support in HRI, and human guidance of robot behaviors. Yet, a low-cost full surface tactile skin is still unavailable despite numerous efforts.

Existing approaches for realizing tactile skins can be divided into two groups depending on the sensing method. The first group relies on exteroceptive sensors such as capacitive, magnetic and IR based sensors [1]–[6]. These methods directly attach tens to thousands of sensor elements to the robot to create a tactile skin that covers an adequate area. However, they can be bulky, difficult to deploy, and may introduce extra overhead for robot operation. Furthermore, they can be expensive. The second group of methods employ proprioception, *e.g.* motor torque, position, velocity and momentum readings coupled with inverse kinematics and dynamics, to create an on-robot tactile skin [7]–[9]. The goal is to infer the contact location and contact force by learning or building models that use measured values internal to the system. However, in practice, it is challenging to obtain dynamic parameters due to noisy and time varying properties of motors, and not every robot is equipped with dedicated proprioceptive sensors. As a result, it is difficult to implement robust real-world systems based on this group of methods.

All authors are with Samsung AI Center NY, 837 Washington Street New York, New York 10014

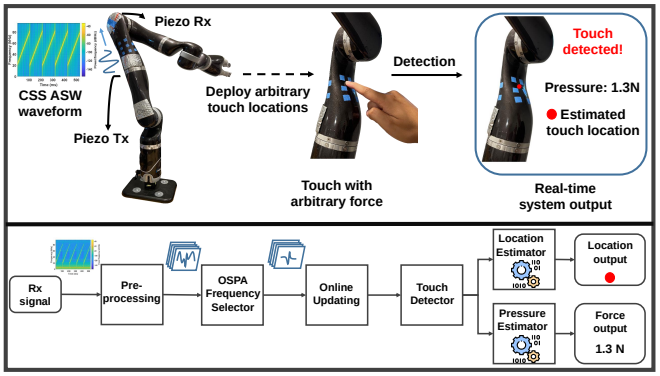


Fig. 1. *SonicSkin* system configuration and overview.

In this paper, we present *SonicSkin*, a virtual tactile skin covering the robot surface that can localize the touch and estimate the touch pressure without deploying sensors at the touch location. The system has three desirable properties: 1) It is low-cost ($< \$2$). 2) It can cover the entire target robot linkage without any dead spots. And 3) it can be deployed on an off-the-shelf robotic manipulator or mobile robot body with minimal modifications. To do so, *SonicSkin* leverages the Acoustic Surface Wave (ASW) introduced in our previous work [10] to create a sensing area within the entire target robot linkage. A single pair of piezoelectric elements is attached to the robot surface for launching and receiving the ASW signal. The core physical intuition is that ASW transforms the entire robot surface into an acoustical sensor, and an external contact at different locations with varying force will introduce different changes in damping, mode, and acoustic impedance for the sensing system. This change can be used to accurately localize the touch and estimate the touch pressure even for a soft contact.

However, to realize this concept and put together a real-world system that robustly works in a dynamic environment is a daunting task. First, both contact location and force will alter the ASW signal. It's challenging to localize the finger touch to centimeter accuracy with uncertain contact force. We address this challenge by signaling a unique wide-band ASW carrier wave. Second, a moving robot introduces significant internal and external acoustical noise. The noise is a challenge for *SonicSkin* because of the acoustical nature of our sensing principle. To overcome the influence of noise,

we created an optimal spectrum prorating algorithm (*OSPA*) to select the frequencies that are sensitive to external touch but insensitive to robot noise. Third, changes in temperature can modify acoustic properties of the robot. We address this issue by introducing an online updating mechanism that enables the system to work robustly during a three months of test time using an initial one-shot training data. Fourth, determining the contact pressure given the above challenges is also a hard problem. *SonicSkin* leverages the wide-band nature of our customized ASW waveform to overcome this challenge.

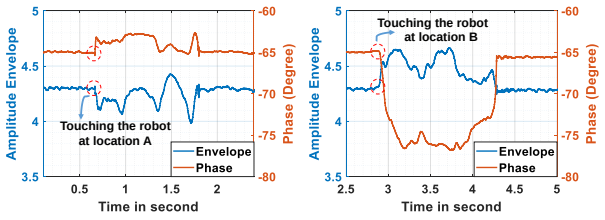
We built a real-time system with a pair of piezoelectric elements attached to an off-the-shelf commercial robot arm, and demonstrated its capability to accurately localize human touch and predict the touch pressure under several difficult but realistic settings. This paper contributes as follows:

- *SonicSkin* is the first real-world system that enables full surface tactile sensing with a single pair of low-cost ($< \$2$) piezoelectric sensors.
- We identified realistic real-world challenges and customized several signal processing mechanisms to effectively address these issues. The system requires only one-shot training, runs in real-time, and corrects for temperature changes.
- We implemented a *SonicSkin* prototype on a 7 DOF manipulator, and conducted comprehensive evaluations with 57019 actual received datapoints from 12 human subjects. Our multi-scenario and multi-person real world experiments show *SonicSkin* localizes 96.4% of the touches with less than 2 cm error even when the robot is moving, and attains root mean square error (RMSE) of 0.59 N when the robot is stationary and 1.86 N for the contact pressure estimation when the robot is moving. Additionally, we also validated the feasibility of deploying *SonicSkin* on robots made of various materials by implementing the system on 15 separate items with different acoustic properties.

II. Related Work and Background

A. Touch Localization and Force Estimation

As briefly discussed in Section I, in the past years, many efforts have been made to create on-robot tactile skins using exteroceptive or proprioceptive sensors. Representative work on tactile skin using exteroceptive sensors includes e-skin [1], artificial sensor skin network [2], flexible tactile sensing array [3], soft magnetic tactile skin [4], and triboelectric nanogenerators skin [5]. These methods usually provide an accurate tactile reading. However, because the sensor elements have to be deployed at every possible touch location, they introduce extra overhead in hardware and software management, and can be costly. On the other hand, tactile sensing using proprioceptive sensors tries to address these inconveniences by intelligently leverage the readings from motor torque, position, velocity and momentum readings. However, representative work in this category [7]–[9] heavily focuses on simulation or requires a special hardware



(a) Touch at location A. (b) Touch at location B.

Fig. 2. The amplitude envelope and phase of the recorded signal from the first proof-of-concept experiment, showing responses to contact at different locations with varying force. (a) The finger touches at location A on the robot. (b) The finger touches at location B that is 2 cm away from location A.

platform (specific proprioceptive sensors). The noisy nature of proprioceptive reading limits the real-world feasibility for these methods.

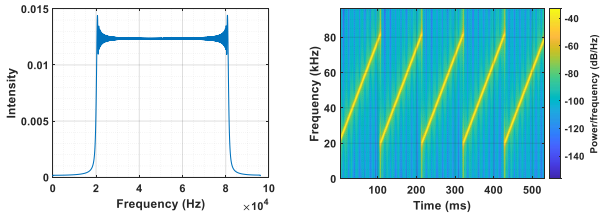
SonicSkin differs from above methods by transforming the target robot linkage into a versatile sensor. The most related works are Panotti [11] and ForcePhone [12]. Panotti leverages multiple microphones deployed on the robot to localize collisions by passively listening to the sound of the contact. But Panotti can not localize soft contact or estimate the force of the contact. ForcePhone transmits ultrasound to probe the contact force applied on cellphones. However, ForcePhone does not use sound for localization and there is no motor noise on cellphones.

B. Background on ASW

We use the tiny "leaks" acoustic energy generated by ASW for collision avoidance (detecting obstacles before the contact) in our previous work [10]. However, the major acoustic energy still resides within the surface. The acoustic impedance changes are significantly stronger when an obstacle contacts the robot. Different materials are characterized by different loss factors and sound velocities; the adaptive nature of the ASW allows us to get useful results with a wide range of materials. The acoustic impedance between the transmitter and receiver changes significantly when an obstacle contacts the robot, and it shows different characteristics when the obstacles with different acoustic properties contact the robot surface at different locations and/or with different force. As a result, anything, regardless of hardness and size, that contacts the robot surface will change the modes and resonances of the entire acoustic system. Using this principle, we transform the entire robot surface into a tactile skin that can accurately localize human touch and estimate the touch pressure. In this paper, we present protocols and algorithms that enable this full surface tactile skin, demonstrate how we address a series of unique challenges when we implement this design on real-world manipulators, and describe an end-to-end system *SonicSkin* that runs in real-time.

III. System Design

SonicSkin employs the ASW generated by a piezoelectric element deployed on the robot to both localize human contact and estimate touch pressure. To provide insight on how *SonicSkin* works and we look at the behavior of a single tone



(a) Chirp in frequency domain. (b) Spectrogram of chirps.

Fig. 3. We customize a CSS ASW waveform that sweeps from 20 kHz to 80 kHz. (a) The frequency domain of our designed chirp shows the signal bandwidth and constant amplitude. (b) The spectrogram calculated by Short-time Fourier transform shows 5 consecutive chirps. The period T of our designed chirp is 106.7 ms, which determines the lower bound of the system response time.

19 kHz ASW signal launched on a Jaco manipulator [13]. Figure 2 shows the amplitude and phase of the received ASW signal when a human finger touches the robot at two different locations (2 cm apart) with varying force. As shown in the figure, the amplitude envelope and phase of the received ASW signal respond differently to the touches at different locations. It indicates ASW could differentiate touches at two very close locations. However, the ASW signal is also sensitive to the touch pressure, which confuses the localization if we directly use the envelope or phase as a fingerprint to do localization.

A. ASW Waveform Design

Challenge 1: localize touches with varying contact force.

The key to reliably localize the touch is to find a combination of physical parameters that is independent of contact force at a given location. The more physical parameters we measure, the better chance we can find this parameter combination. This intuition motivates us to revisit the design of ASW waveform itself. In *SonicSkin*, we customize a signal that changes its frequency linearly between two frequencies f_1 and f_2, namely Chirp Spread Spectrum (CSS) [14]. A linear acoustic chirp is defined as:

$$s(t) = \sin(at^2 + \Omega_1 t + \phi) \tag{1}$$

Where a is the slope of the frequency change in spectrum, ϕ is the initial phase of the chirp, t is the time vector, $0 < t < T$, T is the period of the chirp, $2\pi f_1 = \Omega_1$, $2\pi f_2 = \Omega_2$, and $\Omega_1 + aT = \Omega_2$. Figure 3 shows our designed chirp waveform. *SonicSkin* sends consecutive short chirps with period $T = 106.7$ ms. The localization and force estimation are implemented within one chirp. which determines the lower bound of the system response time.

The received chirp can be written as $r(t) = h(t) * s(t) + n(t)$, where $*$ is convolution operator, n is the noise, and $h(t)$ is the wireless channel between the transmitter and receiver piezoelectric elements on the robot. Note the channel $h(t)$ is frequency selective and the frequency domain of the received signal can be written as $R(f) = H(f)S(f)+N(f)$. When an external contact is made with the robot, the channel response changes accordingly [15]. This process can be written as:

$$\bar{R}(f) = [H(f) + \Delta H(f)]S(f) + N(f) \tag{2}$$

Where $\Delta H(f)$ is the channel alteration introduced by the external contact. Figure 4(a) shows example channel responses when there is no touch and when there are touches at two close locations (2 cm apart). It shows that the channels are mutually different across the frequency band for these three cases. Note for the same location, the $\Delta H(f)$ is also correlated with the contact force. However, as we purposely introduce a wide range of frequencies in ASW signaling, it is now much more likely to find an unique combination of frequency dependent channel variations $\Delta H(\bar{f})$ that can reliably recognize the contact location under varying force. On the other hand, the frequency domain response is unique to the touch at each location. This is because of the inhomogeneity of the interference pattern of the waves refelecting off the boundaries of the robot skin. The pattern changes in a unique way when an objects touches the robot. The wavelength of the sound is a few centimeters which means there is a pattern of this size scale over the entire surface that can be learned by a classfier. We empirically validated this in our large scale field study with 57019 real-world evaluation datapoints.

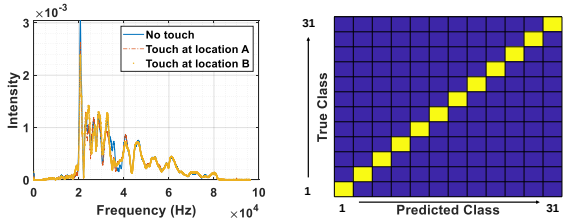
In *SonicSkin* design, the localization framework works in two phases. In the first phase, the system collects the ASW signal when the user touch numerous locations only one time with varying force. The system calculates $\Delta H(f)$ for each touch and perform a one-shot training using a Support Vector Machine (SVM) with a Radial Basis Function (RBF) Kernel. The trained SVM model is essentially the linear combination of frequency dependent channel variations $\Delta H(\bar{f})$. We transform the localization into a classification problem. In the second phase, the system adopts the pre-trained SVM model to localize the user touches. To validate the effectiveness of this design, we performed an experiment with 13 touch locations on a *stationary* Jaco manipulator. In this experiment, adjacent predefined locations are 2 cm apart, and the robot arm is encircled by these 13 locations. During testing, we collected 46 chirps while varying forces were applied at each predefined touch location. The experiment results are shown in Figure 4(b) using a confusion matrix. *SonicSkin* accurately localizes all touch locations on the robot arm. In our evaluation section, we demonstrate that *SonicSkin* can differentiate touch locations within one centimeter accuracy with even more predefined touch locations.

B. Optimal Spectrum Prorating Algorithm

Challenge 2: localize touches when the robot is moving.

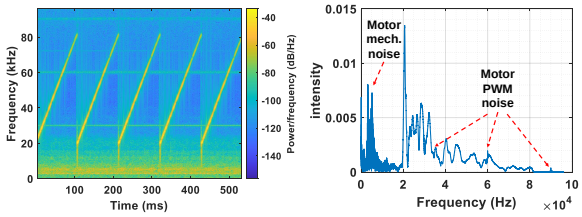
Figure 5 shows an example received chirp signal when the robot is moving. The robot mechanical noise and motor Pulse Width Modulation (PWM) noise significantly distort the feature $\Delta H(f)$ that we used for localization. *SonicSkin* is likely to fail the touch localization if we use $\Delta H(f)$ directly.

We created an Optimal Spectrum Prorating Algorithm (*OSPA*) to address this daunting task. The first step of *OSPA* is to send the spectrogram of the received signal $\mathcal{R}(t, f)$ to a matched filter that removes the lower register (less than 20 kHz) and three PWM harmonic noise bands (30 kHz, 60 kHz, and 90 kHz). We can write the spectrogram of the



(a) Chirp channel responses. (b) Localization confusion matrix.

Fig. 4. We employ the chirp channel response differences $\Delta H(f)$ caused by touches to localize the contact. (a) The chirp channel responds differently when there is no touch, touch at location A, and touch at location B (2 cm away from location A). (b) *SonicSkin* can accurately localize every touch in all locations that surround the robot arm in a proof-of-concept experiment.



(a) Spectrogram for mobile arms. (b) Frequency domain.

Fig. 5. The Spectrogram $\mathcal{R}(t,f)$ and frequency domain $R(f)$ of received ASW chirp when the robot is moving. The major mechanical noise resides in the lower frequency while there are three distinct noise peaks from the motor PWM at 30 kHz, 60 kHz, and 90 kHz.

filtered received ASW chirp as $\mathcal{SR}(t,f)$, where t is the time index and f is the frequency bin index in spectrogram.

The second step in *OSPA* is to find the frequencies that are more sensitive to touches from the remaining frequency bins. Specifically, during the one-shot training, we collect the $\mathcal{SR}(t,f)$ when we touch all predefined touch locations. We next calculate a touch variance vector:

$$TV(f) = \frac{\sum_{t=0}^{n-1} (\mathcal{SR}(t,f) - \frac{\sum_{t=0}^{n-1} \mathcal{SR}(t,f)}{n})^2}{n} \tag{3}$$

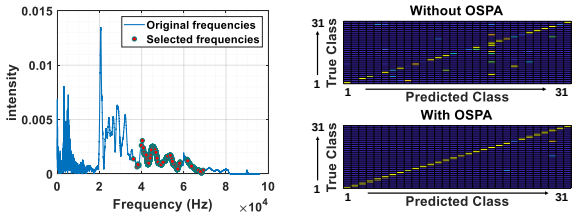
Where n is the length of the signal. $TV(f)$ is the energy variance of each remaining frequency bin caused by touching all predefined locations on the robot. A higher $TV(f)$ value means the frequency f is more sensitive to the touch.

The third step in *OSPA* is to find the frequencies that are less sensitive to the robot movement. Similarly, we collect another $\mathcal{SR}^\star(t,f)$ without touching the robot when the robot is moving. The moving variance vector $MV(f)$ can be calculated similarly as Equation 3. A smaller $MV(f)$ value means the frequency f is less sensitive to the robot moving.

The last step of *OSPA* employs $TV(f)$ and $MV(f)$ work in an adversarial manner. We define the variance contesting vector $VC(f)$ as follow:

$$VC(f) = \frac{MV(f)}{TV(f)} \tag{4}$$

A smaller $VC(f)$ value means the frequency f works as a better option for the localization feature because it translates to a smaller $MV(f)$ and/or larger $TV(f)$ value. From the physics point of view, a smaller $VC(f)$ translates to the frequency f is sensitive to the finger touch (a larger $TV(f)$) but insensitive to robot motion (a smaller $MV(f)$). *OSPA* gets the final selected optimal frequency bins by sorting and thresholding the number of selected frequencies. An implementation of *OSPA* is illustrated in Algorithm 1.



(a) Frequency selection in *OSPA*. (b) Performance wo/w *OSPA*.

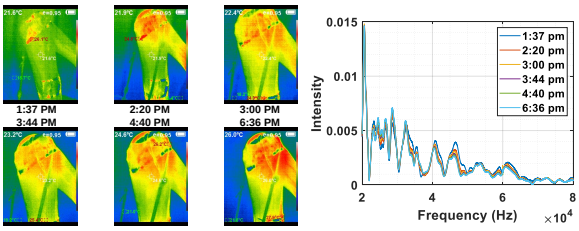
Fig. 6. *OSPA* selects frequencies that are sensitive to the touch but insensitive to the robot movement. (a) The selected frequencies from *OSPA* are mostly from 40 kHz to 70 kHz. (b) An example experimental result. The top and bottom confusion matrix shows the localization performance without and with *OSPA*, respectively. *OSPA* greatly improves the localization accuracy for mobile robot.

Algorithm 1 Optimal Spectrum Prorating Algorithm

function OSPA($\mathcal{R}(t,f)$, $\mathcal{R}^\star(t,f)$)
  $F_{selected} = NULL$
  $\mathcal{SR}(t,f)$ and $\mathcal{SR}^\star(t,f)$ $\leftarrow$ Matched Filtering $\mathcal{R}(t,f)$ and $\mathcal{R}^\star(t,f)$
  $TV(f) \leftarrow$ Variance $\mathcal{SR}(t,f)$
  $MV(f) \leftarrow$ Variance $\mathcal{SR}^\star(t,f)$
  $VC(f) = \frac{MV(f)}{TV(f)}$
  $L(f) \leftarrow$ Ascending Sort $VC(f)$
  for `i` $\leftarrow$ `1 to` γ **do**
    $F_{Selected}(i) = L(i)$
  end for
  return $F_{Selected}$
end function

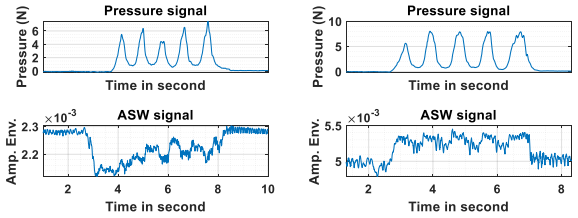
Where $\mathcal{R}(t,f)$ and $\mathcal{R}^\star(t,f)$ are the spectrogram of received signals for touching the robot and for not touching the robot during the one-shot training. γ is a preset threshold. It's an integer number less than the total number of frequency bins when we calculate the spectrogram. We evaluate this parameter setting in Section V-A. $F_{Selected}$ includes the finally selected frequencies from *OSPA*.

Figure 6(a) visualizes the frequency domain of the received ASW chirp with selected frequencies from *OSPA*. Surprisingly, the selected frequencies are not from 20 kHz to 30 kHz where the coupled piezoelectric transceiver has the highest efficiency, but the selected frequencies mostly range from 40 kHz to 70 kHz. This result indeed shows the elegance of *OSPA* design. The physical intuition behind is that *OSPA* takes many design considerations into account, such as sensor efficiency, wavelength, localization resolution, motor noise profile, channel inhomogeneity, etc. The algorithm balances them internally and selects the optimal frequencies for moving robots.



(a) Heat images of the Jaco arm. (b) Frequency response vs time.

Fig. 7. The temperature of the robot shifts gradually after powering on. The features from the frequency domain change accordingly, which makes the localization framework fail over time.



(a) 19 kHz ASW vs. pressure. (b) 24 kHz ASW vs. pressure.

Fig. 8. The ASW signal responds to the touch force differently for different frequencies. The 24 kHz ASW signal in this example projects the pressure better than the 19 kHz ASW signal.

We did an example experiment with 31 predefined touch locations on a mobile robot to show the effectiveness of *OSPA*. Adjacent touch locations this time were 1.2 cm apart, which is similar to the spacing of human fingers, hence one can slide on the robot surface and localize the touch continuously if the classification is 100% correct. The robot is programmed to reach random end effector locations during the whole experiment. Figure 6(b) shows the localization accuracy increases from 54.3% to 96.7% by applying *OSPA* for the collected data. We will present a thorough evaluation of the *OSPA* performance in Section V-B. Note the *OSPA* steps are the same for robots made of different materials. Additionally, *OSPA* selects a collection of optimal frequencies, in the future designs, the system can transmit only those frequencies instead of the entire chirp to provide faster contact detection response time.

C. Online Feature Updating

Challenge 3: compensate the robot temperature drift.

Temperature can drastically change the acoustic properties of a material, such as wavelength and damping [16], thus affecting *SonicSkin* results. We observed the robot continuously warming up several hours after powering on. The robot surface temperature also drifts with changes in the environment temperature. Figure 7(a) shows a series of heat images after powering on the Jaco arm (collected by a HTI HT19 Thermal Imager [17]). As can be seen, the temperature increases 5.4°C after 5 hours. Correspondingly, the frequency domain $R(f)$ of the received chirp also changes as the temperature drifts. This temperature drift causes the *SonicSkin* localization to fail after hours of running.

We overcame this challenge by introducing an online feature updating mechanism. After training, *SonicSkin* localizes touch by sending the $\mathcal{R}(t, F_{Selected})$ selected by *OSPA* to the trained classifier. The classifier yields a score vector S during each detection (*i.e.* every chirp). S is a N by 1 vector where N is the number of predefined touch locations. The final touch location is determined by finding the highest score from S. In order to let the classifier to adapt the temperature drift, *SonicSkin* collects 3 seconds of chirp data for calibration without touching the robot every time the system bootstraps. Then we calculate the average prediction score for each predefined touch location during these 3 seconds, S_c. After this short calibration process, *SonicSkin* predicts the touch location by using a new feature S_{Final}, $S_{Final} = S - S_c$. Additionally, *SonicSkin* updates S_c every 20 minutes. We illustrate this mechanism in Algorithm 2.

Algorithm 2 Prediction By Online Feature Updating

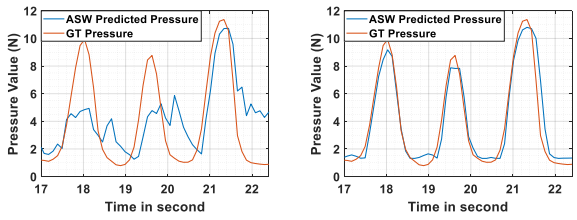
function UPDATING($\mathcal{R}(t, f)$, $\mathcal{R}_c(t, f)$)
  $S_{Final} = NULL$
  for `Every` α `minutes` **do**
    for `t` $\leftarrow$ `1 to` β **do**
      $\mathcal{R}_c(t, F_{Selected}) \leftarrow$ *OSPA* $\mathcal{R}_c(t, f)$
      $S_c(t) \leftarrow$ Predict $\mathcal{R}_c(t, F_{Selected})$
    end for
    $S_c \leftarrow$ Average $S_c(t)$
  end for
  $\mathcal{R}(t, F_{Selected}) \leftarrow$ *OSPA* $\mathcal{R}(t, f)$
  $S(t) \leftarrow$ Predict $\mathcal{R}(t, F_{Selected})$
  $S_{Final}(t) = S(t) - S_c$
  return $S_{Final}(t)$
end function

Where $\mathcal{R}_c(t, f)$ is the spectrogram of the received signal during calibration, $\mathcal{R}(t, f)$ is the spectrogram of received signal during testing, α is the updating rate, and β is length of calibration (α = 20 minutes and β = 3 seconds in *SonicSkin*). Note that after the online updating algorithm, the prediction scores are close to zero when there is no touch, but the prediction scores are high when there is a touch. As such, touch detection is done by thresholding the prediction score. Upon using our online feature updating mechanism, the classifier is able to robustly work even on data collected two months ago. Detailed evaluations of this mechanism and the study of parameter α setting are presented in Section V-A.

D. Touch Pressure Estimation

Challenge 4: how to estimate the touch pressure.

As indicated in Figure 2, the amplitude of received ASW signal responds sensitively to the touch pressure. We next measure and demonstrate the touch pressure using a Singletact miniature force sensor [18] as the ground truth, $G(t)$. Figure 8 shows the envelope of the received ASW signal at 19 kHz and 24 kHz when applying a varying touch force. As can be seen, the 19 kHz signal is not ideal for the touch pressure sensing, and 24 kHz is better but not optimal for this purpose either. To better estimate the touch pressure, *SonicSkin* again takes advantage of the wide-band nature of



(a) Force prediction using $E(t)$. (b) Force prediction using $E^O(t)$.

Fig. 9. *SonicSkin* uses frequency dependent features from the wide-band chirp signal to do force prediction. (a) The force estimation performance by using the total energy across all frequencies from the chirp. (b) The force estimation performance by using the total energy from the selected frequencies from the chirp.

the ASW waveform design. We detail the pressure estimation framework next.

Instead of using the energy from a single frequency, we investigate the relationship between the touch pressure and the total energy across the entire chirp frequency band, $E(t) = \sum_{f=f_1}^{f_2} \mathcal{R}(t,f)$, where f_1 and f_2 are the lower and upper frequency bond of the designed chirp. We define the pressure estimator $P(t)$ by the first order Fourier model:

$$P(t) = a\cos(\omega E(t)) + b\sin(\omega E(t)) + c \tag{5}$$

The parameter a, ω, b, and c are calculated by $\arg\min_{a,b,\omega,c} \sum_{t=0}^{n-1} (P(t) - G(t))^2$. Figure 9(a) shows a snapshot of the predicted pressure and ground truth pressure using this model. We can see an improvement over using the single frequency energy. However, the prediction is still not ideal and the curve fitting R^2 score is 0.46[1]. A detailed evaluation for the curve fitting model is presented in Section V-A.

The *OSPA* presented in Section III-B selects an optimal collection of frequencies that are sensitive to the touch but insensitive to the robot movement. As such, we use the total energy across the selected frequencies as the force prediction feature, $E^O(t) = \sum_{f \subseteq F_{Selected}} \mathcal{R}(t,f)$. Then we calculate the optimal set of parameters in Equation 5 with $E^O(t)$. For the same experiment shows in Figure 9(a), the R^2 score increases to 0.92 during curve fitting by using $E^O(t)$. Figure 9(b) shows the prediction result. As expected, it is apparent that the force estimation accuracy is greatly improved. Finally, we noticed that the optimal parameter set is different for different touch locations. As a result, *SonicSkin* calibrates each touch location separately during the one shot-training. The system uses the complete parameter set $\mathcal{S} = [A, B, \Omega, C]$ for the touch estimation in the prediction phase, where $\mathcal{S} \in \mathbb{R}^{N\times 4}$, N is the number of predefined touch locations.

IV. Putting together a *SonicSkin* system

The *SonicSkin* system is lightweight and can be deployed on the robot with minimum modifications. We deploy two CPT-2065-L100 piezo elements [20] 20 cm apart on a Kinova Jaco Gen 2 7-DoF manipulator. Both piezo elements are attached to the surface of the robot with thermoplastic adhesive (e.g., with hot glue) and covered by a sound deadening material. We use a RME Fireface UFX+ USB Audio Interface [21] as the analog-digital/digital-analog converter (AD/DA). The sampling rate Fs for the AD/DA is 192 kHz and the bit depth is 24 bit. Signal processing is done on an Intel NUC7i7BNH computer [22]. The ground truth for the contact pressure is collected by a SingleTact miniature force sensor [18]. We deploy training contact locations that fully surround one linkage of the robot arm. Note only one pair of piezoelectric elements is required at each linkage for multi-body manipulators because the robot joint decouples the acoustic signal between two adjacent robot linkages. We can adopt time duplex or frequency duplex to avoid interference among sensors.

V. Evaluation

We present the evaluation results in this section. We first describe a set of micro-benchmark experiments to study the parameter settings in *SonicSkin*, followed by an extensive field study to demonstrate the localization accuracy and force estimation accuracy of *SonicSkin* under various conditions. In total, 57019 chirp datapoints were collected for evaluation with 12 human subjects.

A. Micro-benchmark

We performed a collection of micro-benchmark experiments to investigate the optimal parameter settings in *SonicSkin* system design. All experiments were conducted on the Jaco manipulator. As shown in Figure 10, we deployed 31 target touch locations on the robot. In order to demonstrate that *SonicSkin* can accurately differentiate two very close touch locations, the predefined touch locations were densely deployed: the distance between the center of two adjacent touch locations were 1.2 cm. Note in this section, we use the classification accuracy to show the performance of localization under various system parameters. We use the R^2 to show the performance of the curve fitting for the pressure estimation under various models.

1) The numbers of frequencies selected in OSPA: We first study the classification accuracy for localizing the touch on a continuously moving robot when we vary the number of selected frequencies γ in Algorithm 1. In the *SonicSkin* implementation, we have 20480 FFT points when we calculate the spectrum $R(f)$ for the received chirp. We have a max number of 5013 frequency bins after the first step of *OSPA*. Using the same data from one test experiment, we investigate the classification accuracy for the 31 predefined touch locations every 100 frequency bins. Figure 11 shows the results. We can see that the trend of classification accuracy drops when we choose more frequency bins. Note for each γ, the classification performance also varies with the optimization process during training the SVM model. Hence we choose three γ values and run the SVM model training 10 times each to demonstrate the variance of the classification accuracy. We

[1] R^2 a statistical measure that represents the proportion of the variance for a dependent variable that's explained by an independent variable or variables in a regression model [19].

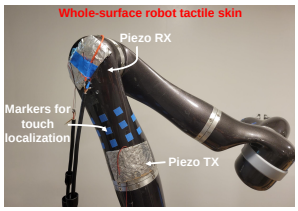


Fig. 10. Predefined touch area. We also touch between markers.

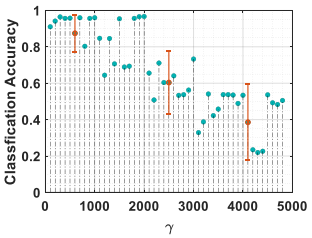


Fig. 11. Classification accuracy vs. number of frequency bins γ.

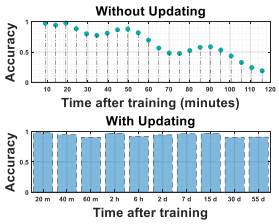


Fig. 12. Classification accuracy vs. time elapsed.

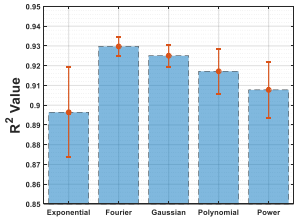


Fig. 13. Curve fitting performance vs. various fitting models.

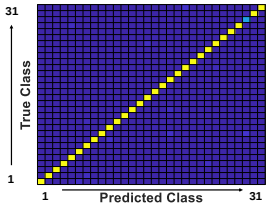
can see that the classification accuracy increases to more than 95% with only a few hundred frequency bins. As a result, we choose $\gamma = 300$ in the *SonicSkin* implementation. Note that the computation load is also reduced when fewer frequency bins are used for the classification task.

2) Online updating rate: We next investigate how often we should update the prediction feature in Algorithm 2. For the experiment results shown in the top of Figure 12, we performed the one-shot training right after powering on the robot, and we tested the accuracy of this model by touching all 31 predefined touch locations every 5 minutes in the next two hours. We can see that the classification accuracy drops significantly after 30 minutes if we don't update the prediction feature. Therefore we set the parameter $\alpha = 20$ minutes in Algorithm 2. The bottom of Figure 12 shows the classification accuracy when we update the prediction features. We can see that the classification accuracy stays relatively stable after nearly two months of operation. This result indicates that the feature updating mechanism presented in Section III-C is more than enough to support the system in the long run.

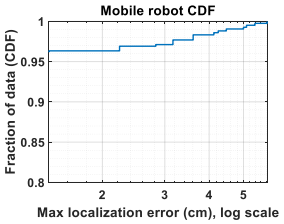
3) Curve fitting models for pressure estimation: We study the curve fitting model selection in the last part of micro-benchmark experiments. We simply put a first order Fourier model in Section III-D, however we also evaluated other models when we design the force estimator. We collected the data in this experiment by touching all 31 predefined locations with the SingleTact force sensor in these experiments. Figure 13 shows the R^2 values when we least square fit the data. Note that we also vary the order for each model when we implement the curve fitting (hence we see a variation in the R^2 for each model). As a result, the first order Fourier model is our optimal choice considering both fitting accuracy and computation overhead. Please note that a simple model such as a first order polynomial fitting (linear fitting) also yields convincing R^2 value ($R^2 = 0.9$), which validates the effectiveness of the frequency selection mechanism in our force estimator design.

B. Field Study

We employed the optimal parameters selected from the micro-benchmark experiments and conducted an extensive field study to evaluate the real-world robustness of *SonicSkin*. Again we deployed 31 dense touch locations on the Jaco arm as described in the micro-benchmark (the distance between the center of two adjacent touch locations are 1.2 cm). We first studied the localization and force estimation accuracy for



(a) Localization confusion matrix.



(b) Localization error CDF.

Fig. 14. The localization (a) confusion matrix and (b) error CDF in *SonicSkin* for mobile robot. We collected 16492 real-world datapoints in this experiment.

both stationary and moving robots. Next we demonstrated the localization accuracy for multiple persons. Finally we investigated the feasibility of deploying *SonicSkin* on robots made of other materials than the Jaco arm.

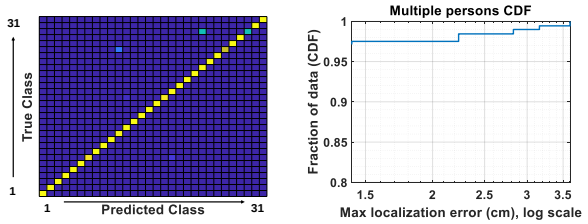
1) Localization accuracy: In this study, during a three month period, we collected 6944 testing datapoints when the robot was stationary, and 16492 testing datapoints when the robot was moving. Each datapoint was collected by recording one received chirp signal when touching one target location with varying force. The classification accuracy for the stationary robot was 99.4%, which indicates it is a trivial task for *SonicSkin* in this scenario. For the moving robot, Figure 14(a) shows the confusion matrix for this 31 classes. Note the mis-classifications are rare and they can be seen when we zoom-in this chart. The overall classification accuracy is 96.2%. To better understand the localization error in distance, we demonstrate the (Cumulative distribution function) CDF in Figure 14(b). it shows that 97.0% of the touch localization have less than 3 cm of error.

2) Force estimation accuracy: In this section, we report the RMSE, standard deviation σ, and the cross-correlation between the predicted force and ground truth force. In total 9175 datapoints for a stationary robot and 3912 datapoints for a moving robot were collected. Table I shows the force estimation performances under these two scenarios. As expected, the estimation accuracy drops when the robot is moving. However, the predicted force is still very similar as the ground truth (more than 95% of cross-correlation). Note the ground truth reading might not be accurate when the robot is moving, because the SingleTact sensor might be affected by the robot motion.

3) Localization accuracy for multiple persons: In order to study the robustness of *SonicSkin* in more realistic scenarios, we invited 12 volunteers to touch all 31 predefined locations at will (varying force, fingers, etc) and collected 10416

	RMSE (N)	σ (N)	Cross-correlation
Stationary robot	0.59	0.58	99.4%
Mobile robot	1.86	1.73	96.7%

TABLE I. **Performances of the force estimator in** *SonicSkin***.**



(a) Localization confusion matrix. (b) Localization error CDF.

Fig. 15. Localization performances for multi-person experiments.

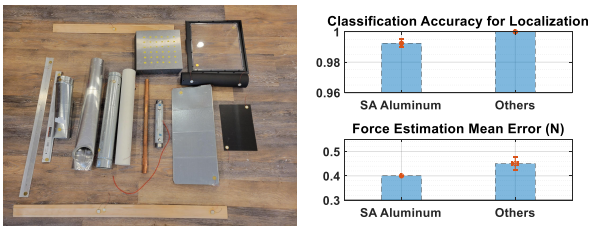
testing datapoints when the robot is stationary. Figure 15 shows the confusion matrix and localization error CDF of this study. The overall classification accuracy is 97.2%, and 98.5% of the touch location estimations have less than 3 cm of localization error.

4) Various robot surface materials: Last but not least, not every robot is made of carbon fiber composite found on the Jaco arm. To investigate the feasibility of *SonicSkin* for robots made of other materials, we deployed the piezoelectric element on various materials as shown in Figure 16(a). The same as on-robot experiments, the piezoelectric elements are simply attached to various target surfaces (e.g., with hot glue). Note it's easy to reproduce our system on a variety of substrates and geometries. But the one-shot calibration for *OSPA*, online feature updating, and classifier retraining need to be implemented for any new surface materials.

We conducted this study in two parts. Since many robots are made of aluminum, we first studied the *SonicSkin* on a standalone aluminum plate with 36 predefined touch locations (3 cm spacing between adjacent touch locations). We collected 5040 datapoints in this experiment. Then we studied *SonicSkin* on 14 other materials shown in Figure 16(a) with two very close predefined touch locations (1.2 cm spacing). We collected 360 datapoints for each material. We report the classification accuracy and RMSE for force estimation in Figure 16(b). Note the classification accuracy is 100% for the materials other than the standalone plate. The force estimation error is lower than 0.5 N for all materials we tested. This study indicates it is likely to effectively implement *SonicSkin* on other robots made of different materials than the Jaco arm.

VI. Conclusion

In this paper, we presented *SonicSkin*, a low-cost on-robot full surface tactile skin that can be deployed on commercial off-the-shelf robots with minimum modifications. *SonicSkin* realized less than 2 cm touch localization error for 96.4% of tests, with more than 96.7% cross-correlation similarity between the predicted touch pressure and the ground truth touch pressure. We also validated the feasibility of implementing *SonicSkin* on robots made of various materials.



(a) Various dummy robot materials we tested. (b) Localization and pressure estimation performances.

Fig. 16. We study the feasibility of implementing *SonicSkin* on 15 different materials used as "dummy" robots.

We have also demonstrated that *SonicSkin* works with contact by objects other than human fingers. In the future we will validate its accuracy on a wide range of living and non-living contact objects. We also plan to further investigate the possibility and challenges to deploy *SonicSkin* on mobile robots (*e.g.* vacuum cleaners).

References

[1] R. Dahiya, "E-skin: from humanoids to humans [point of view]," *Proceedings of the IEEE*, 2019.

[2] P. Mittendorfer, E. Yoshida, and G. Cheng, "Realizing whole-body tactile interactions with a self-organizing, multi-modal artificial skin on a humanoid robot," *Advanced Robotics*, 2015.

[3] Z. Ji, H. Zhu, H. Liu, N. Liu, T. Chen, Z. Yang, and L. Sun, "The design and characterization of a flexible tactile sensing array for robot skin," *Sensors*, 2016.

[4] T. Hellebrekers, N. Chang, K. Chin, M. J. Ford, O. Kroemer, and C. Majidi, "Soft magnetic tactile skin for continuous force and location estimation using neural networks," *IEEE Robotics and Automation Letters*, 2020.

[5] G. Yao, L. Xu, X. Cheng, Y. Li, X. Huang, W. Guo, S. Liu, Z. L. Wang, and H. Wu, "Bioinspired triboelectric nanogenerators as self-powered electronic skin for robotic tactile sensing," *Advanced Functional Materials*, 2020.

[6] S. Sundaram, P. Kellnhofer, Y. Li, J.-Y. Zhu, A. Torralba, and W. Matusik, "Learning the signatures of the human grasp using a scalable tactile glove," *Nature*, 2019.

[7] L. Manuelli and R. Tedrake, "Localizing external contact using proprioceptive sensors: The contact particle filter," in *IROS*. IEEE, 2016.

[8] D. Popov and A. Klimchik, "Real-time external contact force estimation and localization for collaborative robot," in *ICM*. IEEE, 2019.

[9] J. Zhou, Y. Chen, X. Chen, Z. Wang, Y. Li, and Y. Liu, "A proprioceptive bellows (pb) actuator with position feedback and force estimation," *IEEE Robotics and Automation Letters*, 2020.

[10] X. Fan, R. Simmons-Edler, D. Lee, L. Jackel, R. Howard, and D. Lee, "Aurasense: Robot collision avoidance by full surface proximity detection," in *IROS*. IEEE, 2021.

[11] X. Fan, D. Lee, Y. Chen, C. Prepscius, V. Isler, L. Jackel, H. S. Seung, and D. Lee, "Acoustic collision detection and localization for robot manipulators," in *IROS*. IEEE, 2020.

[12] Y.-C. Tung and K. G. Shin, "Expansion of human-phone interface by sensing structure-borne sound propagation," in *MobiSys*. ACM, 2016.

[13] "Kinova Jaco Prosthetic robotic arm," Webpage, 2018.

[14] H. Cho and S. W. Kim, "Mobile robot localization using biased chirp-spread-spectrum ranging," *IEEE transactions on industrial electronics*, 2009.

[15] A. V. Oppenheim, A. S. Willsky, S. H. Nawab, G. M. Hernández, *et al.*, *Signals & systems*. Pearson Educación, 1997.

[16] J. H. Kuypers, C.-M. Lin, G. Vigevani, and A. P. Pisano, "Intrinsic temperature compensation of aluminum nitride lamb wave resonators for multiple-frequency references," in *IFCS*. IEEE, 2008.

[17] "HT 19 Thermal Imager," Webpage, 2021.

[18] "8 mm Diameter, 10N/2.2LB Force (USB)," Webpage, 2021.

[19] "Coefficient of determination," Webpage, 2021.

[20] "CPT-2065-L100 Piezo Element," Webpage, 2021.

[21] "Fireface UFX+," Webpage, 2021.

[22] "Intel NUC7i7BNH," Webpage, 2019.