AudioTouch: Minimally Invasive Sensing of Micro-Gestures via Active Bio-Acoustic Sensing

Yuki Kubo[*]
IPLAB, University of Tsukuba
Ibaraki, Japan
kubo@iplab.cs.tsukuba.ac.jp

Yuto Koguchi
IPLAB, University of Tsukuba
Ibaraki, Japan
koguchi@iplab.cs.tsukuba.ac.jp

Buntarou Shizuki
IPLAB, University of Tsukuba
Ibaraki, Japan
shizuki@cs.tsukuba.ac.jp

Shin Takahashi
IPLAB, University of Tsukuba
Ibaraki, Japan
shin@cs.tsukuba.ac.jp

Otmar Hilliges
AIT Lab, ETH Zurich
Zurich, Switzerland
otmar.hilliges@inf.ethz.ch

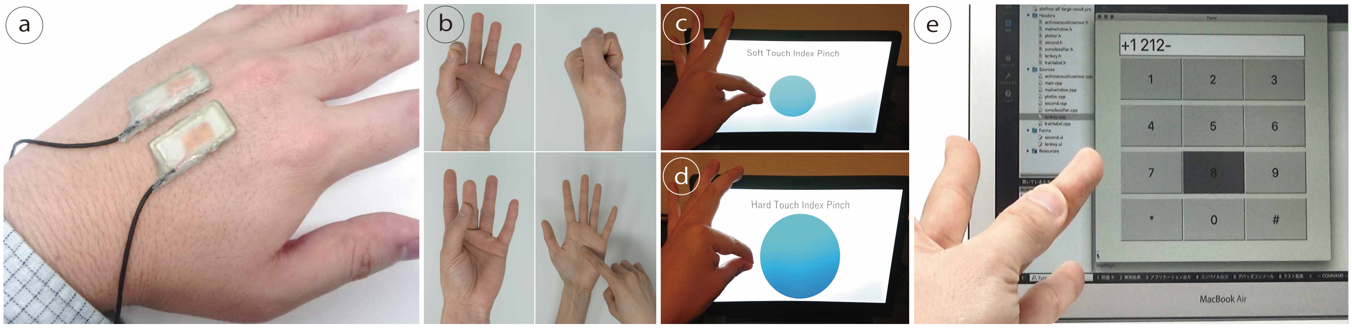


Figure 1: (a) AudioTouch is a micro-gesture recognition approach based on active bio-acoustic sensing without requiring any instrumentation on users' fingers or palm. (b) It recognizes micro-gestures with small differences among various finger gestures. (c+d) It also allows for discrimination of force, further expanding interaction vocabulary. (e) This approach enables several compelling application scenarios such as device-free input in mobile scenarios.

ABSTRACT

We present AudioTouch, a minimally invasive approach for sensing micro-gestures using active bio-acoustic sensing. It only requires attaching two piezo-electric elements, acting as a surface mounted speaker and microphone, on the back of the hand. It does not require any instrumentation on the palm or fingers; therefore, it does not encumber interactions with physical objects. The signal is rich enough to detect small differences in micro-gestures with standard machine-learning classifiers. This approach also allows for the discrimination of different levels of touch-force, further expanding the interaction vocabulary. We conducted four experiments to evaluate the performances of AudioTouch: a user study for measuring the gesture recognition accuracy, a follow-up study investigating the ability to discriminate different levels of touch-force, an experiment assessing the cross-session robustness, and, a systematic evaluation assessing the effect of sensor placement on the back of the hand.

CCS CONCEPTS

• **Human-centered computing → Gestural input**; *Ubiquitous and mobile devices*;

KEYWORDS

Bio-acoustics; Gesture Recognition; Finger Inputs; Touch-Force Recognition, One-handed Interaction

ACM Reference Format:
Yuki Kubo, Yuto Koguchi, Buntarou Shizuki, Shin Takahashi, and Otmar Hilliges. 2019. AudioTouch: Minimally Invasive Sensing of

---

[*] Also with Nippon Telegraph and Telephone (NTT) Corporation.

MobileHCI '19, October 1–4, 2019, Taipei, Taiwan

ACM ISBN 978-1-4503-6825-4/19/10…$15.00
https://doi.org/10.1145/3338286.3340147

Micro-Gestures via Active Bio-Acoustic Sensing. In *21st International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '19), October 1–4, 2019, Taipei, Taiwan.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3338286.3340147

1 INTRODUCTION

With computing becoming increasingly mobile and ubiquitous, interaction beyond the desktop increasingly highlights the need for complementary forms of input when no mouse or keyboard is available. While touchscreens are clearly intuitive and popular, they also come with drawbacks. In particular, as mobile devices continue to be miniaturized, touchscreen real-estate becomes increasingly limited, leading to smaller on-screen targets and occlusion by the fingers of displayed content. This issue becomes pronounced once we consider devices that do not even require a screen such as smart appliances and remote controllers, or devices that have a screen but with which we want to interact at a distance such as a phone in a pocket or a smart thermostat.

In response to these changing interaction needs, studies on human-computer interaction (HCI) have explored many approaches of recognizing gestural interaction with mobile devices, including sensing via the built-in camera [43], infrared (IR) proximity sensors [1, 31], magnetic tags [19], electromyography (EMG) [18, 33, 36, 41], electro-tomography [56] and ultrasound-imaging [32]. However, these often require significant user instrumentation or instrumentation of the environment, presenting significant barriers to adoption. Furthermore, many approaches require users to perform large hand or wrist motions to be able to discriminate various gestures.

In this paper, we present AudioTouch, a novel approach for sensing micro-gestures using active bio-acoustic sensing. AudioTouch recognizes micro-gestures with small differences among various finger gestures, for example, only articulating the thumb to parts of the other fingers such as the fingertip or individual segments of the finger. This approach is less invasive compared to those involving ring and glove-type sensors. Since it only requires two piezo-electric elements to the back of the hand (BoH) and does not require any instrumentation on the palm or fingers and does not encumber interaction with physical objects. Hence, we use the term 'minimally invasive' to denote this property.

We demonstrate that the acquired signal is rich enough to accurately discriminate a large number of micros-gestures via standard machine-learning methods. Furthermore, given the signal properties, a robust classifier can be trained from very few sample; hence, it is straightforward to personalize the gesture set. We also demonstrate that AudioTouch is capable of discriminating two levels of touch force in thumb-to-finger gesture sets of the middle and bottom phalanges with 85.0% and 85.5% accuracy, respectively. In addition, we investigated the effect of sensor location on the recognition accuracy using ten sensor locations. Except for one location, we did not find any difference in the recognition accuracy of AudioTouch. Finally, we discuss interaction scenarios in which AudioTouch may be leveraged.

2 RELATED WORK

With the explosive proliferation of mobile computing devices, the need for alternative input paradigms has grown significantly. We review work touching upon various aspects of input recognition in HCI research, focusing on suitable sensing modalities for mobile gesture recognition.

Input Recognition

Camera-based gesture recognition. Much work has been dedicated to estimating the full 3D hand-pose of the user or detect discrete gestures using RGB or depth cameras [12, 27, 43, 44, 46]. For example, LeapMotion [27] uses a short-baseline stereo pair in combination with active IR illumination to track 6-degrees-of-freedom hand poses. Taylor et al. [46] proposed a real-time hand tracking system using depth cameras that combines a discriminative machine-learning model, for initialization, with an energy minimization process for temporal tracking. Song et al. [43] proposed a data-driven technique to recognize static hand postures using only an RGB camera of commercial mobile devices. While providing rich means of input, all these methods rely on the line-of-sight between an environment-mounted or body-worn camera and the user's hands; hence, there are issues with (self-)occlusions, which may limit user mobility.

Non-visual sensing modalities. In the light of challenges faced by vision-based approaches, researchers have explored several alternative sensing modalities. Many gesture recognition approaches that leverage radio-frequency signals [28, 29, 47], capacitive sensing [42], and thermal sensing [9] have been proposed. WiFinger [28] is a gesture recognition system that exploits the channel state information of existing Wi-Fi signals to detect coarse user interaction. Soli [29] is a millimeter-wave radar that can be leveraged for gesture recognition. A follow-up work [47] has demonstrated that this signal can be leveraged to recognize 11 dynamic gestures based on a deep learning algorithm. Touché [42] recognizes the configuration of a hand touching a conductive object using capacitive profiles of the object. Pyro [9] recognizes pinching gestures using pyroelectric IR sensing. These approaches rely on external sensors or require significant user instrumentation and often can only detect coarse gestures (e.g., wave gestures). In contrast, AudioTouch recognizes a rich set of fine-grained hand postures based on light-weight on-body sensing only.

Finger and Glove-based Sensing. To avoid occlusion issues, ring-type sensors such as cameras [3], hall effect sensor arrays [19], and contact microphones [55] have been proposed. Cyber-gloves have been used provide user input in the context of augmented reality (AR) and virtual reality (VR) [4, 7, 35, 50, 51]. While such approaches can recognize fine-grained interactions, they do require instrumentation on the users' fingers and palm, which can be an issue. In contrast, AudioTouch does not require any instrumentation on the users' fingers and palm and solely relies on two piezo-electric sensors attached to the BoH.

On-body sensing. For overcoming the need for an environment-mounted sensing infrastructure, a number of approaches leverage body-worn sensors.

Various works have explored the use of optical sensors mounted on the user's wrist, arm, or shoulder [12, 24, 39, 44]. Similar to environment-mounted cameras, such approaches are susceptible to (self-)occlusion and varying lighting conditions and may be bulky due to the optical apparatus.

Much research has been dedicated to measuring signal changes induced by the configuration of the bones, muscles, wrist, or arm using sensors such as IR sensors [1, 31], capacitance/impedance sensors [40, 56, 58], pressure sensors [2, 5], IMUs [26, 54], EMG [18, 33, 36, 41], or ultrasound imaging [32]. Most closely related to AudioTouch are approaches that attempt to recognize subtle micro-gestures such as thumb-to-fingertip gestures. Tomo [56, 57] recognizes coarse hand gestures and pinch gestures using electrical impedance tomography on the user's wrist. Several approaches [18, 41] leverage forearm-mounted EMG electrodes to recognize a small number of pinch gestures. However, EMG-based sensing requires many electrodes. SensIR [31] is a gesture recognition method that can recognize 12 gestures, including pinch gestures, using near-IR sensing on the user's wrist with a wrist-worn bracelet composed of pairs of emitters and receivers. EchoFlex [32] can recognize ten discrete hand gestures quite accurately but requires the mounting of a large and expensive ultrasound probe on the user's arm. ThumbSlide [1] can detect the thumb sliding over a user's finger using an array of IR sensors mounted on the wrist. Finally, photo-reflective sensors [53] or strain gauges [30] mounted on the BoH have been explored for gesture recognition. AudioTouch also leverages BoH-mounted sensors but recognizes a rich set of thumb-to-finger gestures and, thumb gestures and detects touches to the user's palm as well as discriminates between different levels of touch-force across different pinch gestures, while only requiring minimal user instrumentation.

Acoustic Sensing

In our work, we leveraged bio-*active* acoustic sensing; hence, our work relates to those that leveraged acoustic sensing of some form for (input) recognition. Passive acoustic sensing has been used for gesture recognition [6, 11, 13], on-body sensing [14], and object identification [15]. Active acoustic sensing has been used in prototyping methods [25, 37, 38] for gesture recognition [23, 34, 52] and context recognition [48]. The Sound of Touch [34] tracks finger positions on the user's arm and recognizes arm-grasp gestures. Yokota et al. [52] proposed an on-skin touch sensing method using a transducer mounted on the user's wrist and receivers mounted on the user's wrist and tip of index finger. To the best of our knowledge we are the first to propose a method for the micro-gesture recognition of hand postures using only two piezo-electric sensors attached to the BoH. We investigated the effect of sensor locations on our method's gesture recognition accuracy.

Force and Pressure Sensing

Several approaches measure force on a touchscreen using the built-in sensors [8, 17, 45] or leveraging additional equipment [16, 21]. Others involve the sensing of stylus force in digital ink applications [20]. AudioTouch senses force as an additional modifier for gestural interaction and achieves this despite the lack of direct force sensing capabilities. In this sense, AudioTouch is similar to [50], which requires a force sensitive glove, but AudioTouch does not require any instrumentation on the fingers.

3 AUDIOTOUCH

Our goal was to recognize subtle, micro-gestures without requiring excessive user instrumentation to enable interaction in mobile and unconstrained settings. Prior work has shown that large arrays of strain gauges or photo reflectors can be leveraged to recognize finger postures [29, 30] and that users touching passive objects can change the standing wave patterns induced by surface mounted speakers [37]. We built upon and extended this previous work by leveraging piezo-electric sensors as a surface speaker and microphone mounted on the BoH to recognize micro-gestures based on differences in the resonant properties of the hand. We experimentally demonstrated that this approach, AudioTouch, can be used to discriminate between a rich set of one-handed gestures even if the changes in the hand are small. AudioTouch is so sensitive that it is even possible to discriminate between soft and hard presses across four different pinch gestures.

Sensing Principle

Our sensing principle is based on active bio-acoustic sensing. At the core of AudioTouch lies the observation that the shape of the hand and the configuration of bones and muscles in the hand, which serve as media for acoustic waves, change depending on finger postures. Generally, an object has its own resonant property, which depends on the shape, material, and boundary condition [37]. Prior work (e.g., [37]) focused on the boundary condition of an object to recognize how the object is touched. When an object is touched, the boundary condition changes, which changes the resonant property of the object. This change can be observed by vibrating an object and obtaining the frequency response.

AudioTouch, however, focuses on changes in shape, including the internal configuration, of the hand. When a user changes finger postures, the bones and muscles move; thus, the resonant property of the hand changes (Figure 2). These changes are observed as different resonant spectra. Using this basis, AudioTouch recognizes gestures by observing the resonant spectra. Specifically, it emits ultrasound from one piezo element attached to the BoH to both through the hand and along the surface and obtains the frequency response with another piezo element attached to the BoH. AudioTouch uses supervised machine-learning to recognize the gesture with the obtained resonant spectra.

Hardware

Our prototype system consists of a piezo-electric microphone and speaker, audio interface, and laptop running our custom software for signal processing and machine-learning-based micro-gesture recognition. Figure 3 shows our prototype implementation, illustrating the minimally invasive nature of AudioTouch. We use a piezo-electric microphone and speaker made of commercial bimorph type piezo elements (THRIVE K2512BP1, $25 \times 12 \times 0.23$ mm). The microphone induces a wave-pattern that travels through the user’s hand and surface and is received (after pose-determined attenuation) as vibration response by the surface microphone. The piezo elements are attached to an acrylic plate with hot glue to prevent breakage.

We mount the piezo elements on the user’s hand via simple double-sided tape. This mounting method does not result in any external forces being applied to the piezo elements. More specifically, we use double-sided medical tape (3M, 2477 Double-Coated TPE Silicone Acrylate Medical Tape) to ensure safe mounting on the skin. The piezo elements are amplified and connected to a laptop (Apple MacBook Pro, CPU: Intel Core i7 3.5 GHz, RAM: 16GB) through an audio interface (Steinberg UR44).

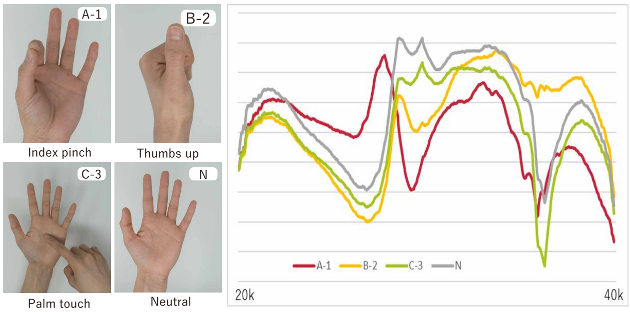


Figure 2: Examples of resonant spectra in hand postures. Differences in shape including internal configuration of hand are observed as different acoustic resonant spectra.

Software

The software consists of a sweep signal generator, vibration response analyzer, and machine-learning-based gesture recognition engine. The audio signals are handled via the BASS audio library.

Our sweep signal generator emits sinusoidal sweep signals from 20 – 40 kHz. The sweep signal increases linearly in 20 ms then repeated. That is, the duration of the sweep is 20 ms. The sampling rate is 96 kHz. We determined the frequency range in reference to a previous work [37]. This range is also inaudible to the human ear; hence, the approach appears to be entirely silent to the user.

The vibration response analyzer converts incoming audio signals from the time domain into the frequency domain. This module samples audio signals at 96 kHz. It first uses 4096 samples to calculate the frequency domain of the signals ranging from 0–48 kHz using the fast fourier transform (FFT), resulting in a set of 2048 frequency domain values. From these values, it extracts values in the range of 20–40 kHz. Finally, it constructs a 400-element feature vector using a peak detection algorithm, which obtains valid values (avoids using very small values) for the recognition engine. After the recognition, the system performs the above process again.

Using these features as input, the recognition engine classifies the different, discrete gestures in real-time using a standard support vector machine (SVM) classifier. The engine recognizes gestures every 20 ms. We used the SVM algorithm of the WEKA Machine Learning Toolkit [10] with its default parameters. The classifier was trained using the sequential minimal optimization algorithm.

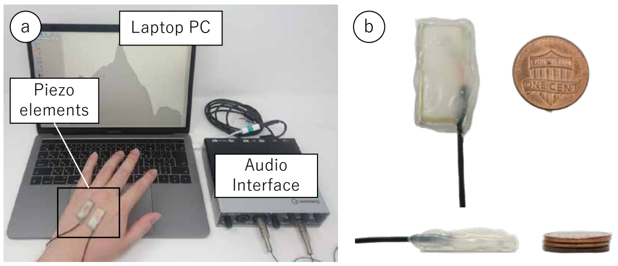


Figure 3: System overview and piezo elements. (a) AudioTouch consists of a piezo-electric microphone and speaker, audio interface, and laptop. (b) Piezo elements ($25 \times 12 \times 6$ mm) are attached to acrylic plates with hot glue.

4 EXPERIMENTAL EVALUATION

We conducted four experiments to evaluate the micro-gesture and touch-force recognition accuracies and feasibility of AudioTouch: a user study for measuring the micro-gesture recognition accuracy (Study 1), follow-up study investigating the ability to discriminate different levels of touch-force (Study 2), experiment assessing the cross-session robustness (Study 3), and systematic evaluation assessing the effect of sensor location on the BoH on the gesture recognition accuracy of AudioTouch (Study 4).

Participants

We recruited 11 participants (10 male and 1 female, P1 – P11) ranging in age from 21 to 24 (SD = 1.17) for Studies 1–3. These 11 participants took part in Study 1. In addition, we randomly assigned the 11 participants for Study 2 (8 participants: all male) and Study 3 (3 participants: 2 male and 1 female) to decrease the influence of fatigue. We also recruited another 10 participants (all male, R1 – R10) ranging in age from 21 to 25 (SD = 1.23) for Study 4.

Procedure

We first asked participants to attach the piezo elements to the back of their left hand. The locations were 10 mm from the metacarpal (MCP) joints (the knuckle between the hand and a finger). This distance was chosen according to the results reported in a previous work [30]. The distance between piezo elements was experimentally determined to be 10 mm. In this setting, the extensor tendon of the middle finger serves as the centerline between the two elements. During the experiment, the participants placed their left elbow and forearm on the desk and an armrest, respectively, to prevent fatigue (Figure 4a). The monitor in front of the participants displayed different hand poses, which were then performed by the participants. The order of presentation was randomized. When they performed the thumb gesture set, we asked them to make their palm perpendicular to the desk. When they performed the thumb-to-finger and palm gesture sets, we asked them to make their palm parallel to the desk. In a trial, we asked the participants to perform the displayed hand pose and to press a foot switch. After pressing the foot switch, data collection for each gesture started. The observed vibration response showed small variance over time. To suppress the effect of variance, we collected 20 samples (i.e., 20 × 400 element vectors) in each trial.

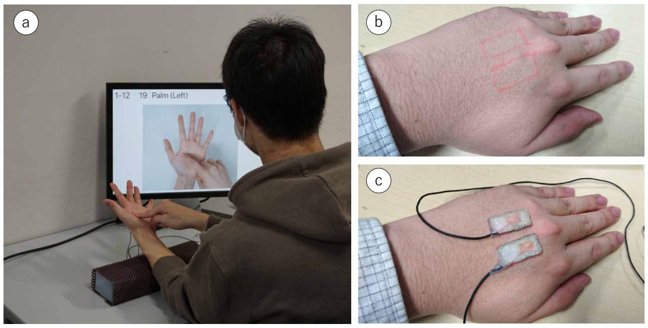


Figure 4: Experimental setup. (a) Participants placed their left elbow and left forearm on the desk and an armrest, respectively, during experiments. (b, c) We asked participants to mark the position of two piezo elements using a marker to investigate the reproducibility of AudioTouch (two red rectangles on BoH).

Study 1: Micro-Gesture Recognition

To evaluate the micro-gesture recognition accuracy of AudioTouch, we used three gesture sets: thumb-to-finger (Figure 5), thumb (Figure 6), and palm touch (Figure 7). We evaluated such recognition accuracies using these three individual gesture sets and a combined gesture set consisting of a total of 24 gestures (Figures 5 – 7). We also added a neutral state to simulate an idle (i.e., no interaction) state.

Procedure. The instruction order of the three gesture sets and neutral state was randomized, and participants were asked to conduct ten sessions. In each session, he/she performed the 24 gestures once. We asked them to take a break after every session for at least one minute. We collected the following number of samples: 11 participants × 10 sessions × 24 gestures × 20 samples = 52,800 samples.

Results. To attain meaningful micro-gesture recognition accuracies, we conducted a leave-one-session-out cross-validation for each participant and averaged the accuracies across participants. AudioTouch achieved average recognition accuracies of 87.1% (SD = 11.3), 82.9% (SD = 13.5), and 89.3% (SD = 13.8) for thumb-to-finger, thumb, and palm touch gestures, respectively. It also achieved an average recognition accuracy of 84.4% (SD = 10.9) when the classifier was trained on all gestures simultaneously.

Study 2: Touch-Force Recognition

In addition to recognizing several micro-gestures (as in Study 1), we observed that the tension on the BoH changes significantly when fingers touch each other lightly versus when pressed together tightly. To evaluate whether this difference can yield a meaningful gesture modifier (i.e., the same gesture can be mapped to different functionalities based on the exerted force), we conducted an additional user study to evaluate the touch-force recognition accuracy of AudioTouch, which would show its advanced potential and limitations. We again asked participants to perform gestures, but this time at two levels of touch-force (soft and hard), and evaluated AudioTouch's touch-force recognition accuracy using

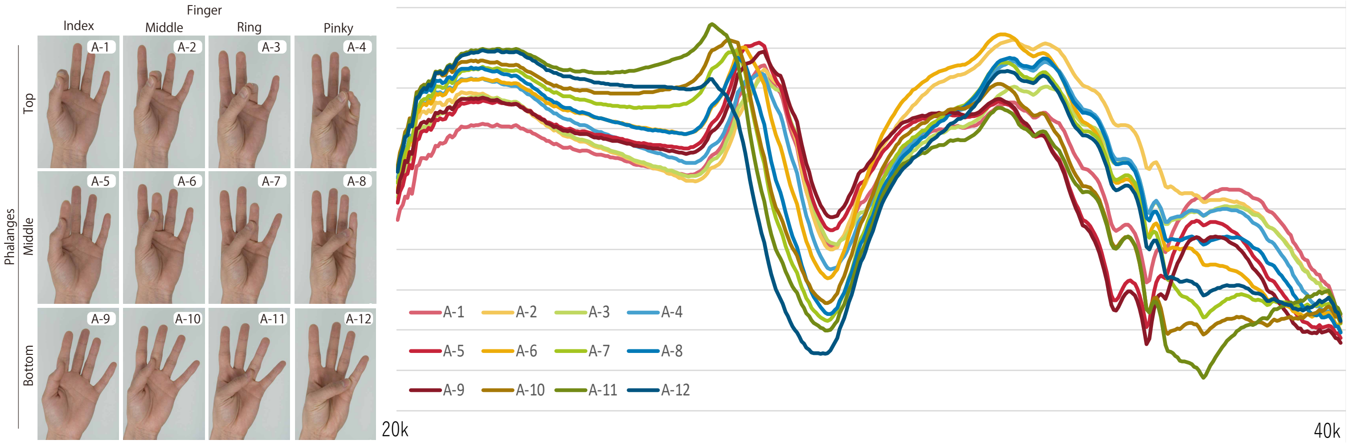


Figure 5: Thumb-to-finger gesture set and example bio-acoustic spectra of this set. Thumb touches each phalange on the hand.

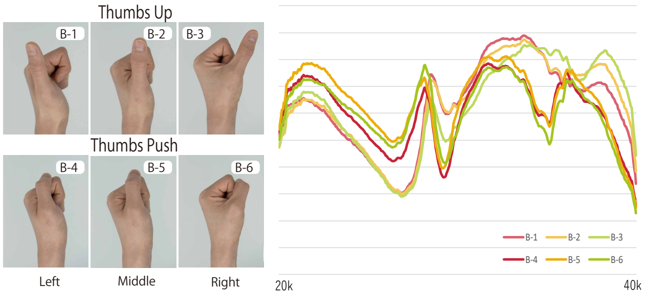


Figure 6: Thumb gesture set and example bio-acoustic spectra of this set.

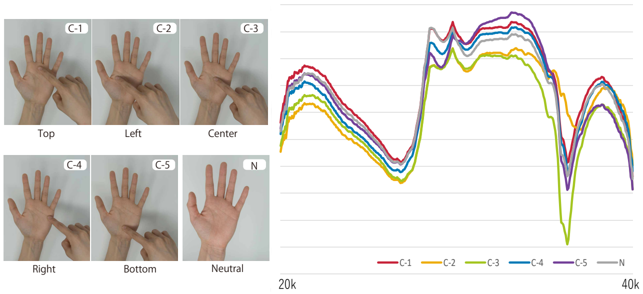


Figure 7: Palm touch gesture set and a neutral state, and example bio-acoustic spectra of this set and neutral state.

the thumb-to-finger gesture set. That is, we analyzed the recognition accuracy using 25 gestures (A-1 – A-12 gestures (Figure 5) × 2 levels + neutral state). We also used the collected data to evaluate such recognition accuracies from this gesture set when limited to a specific phalange (i.e., top, middle, or bottom) to examine the recognition accuracies in detail. That is, we analyzed the following three gestures sets: thumb-to-finger gesture set of the top phalanges (A-1–A-4 × 2 levels + neutral state), that of the middle phalanges (A-5–A-8 × 2 levels + neutral state), and that of the bottom phalanges (A-9–A-12 × 2 levels + neutral state).

Procedure. The presentation order of the gestures was randomized, and the participants were asked to conduct ten sessions. In each session, they conducted the 25 gestures once. We asked the participants to either only lightly touch fingers or press them tightly. We asked them to take a break after every session. We collected the following number of samples: 8 participants × 10 sessions × 25 gestures × 20 samples = 40,000 samples.

Results. We conducted a leave-one-session-out cross-validation for each participant. AudioTouch achieved average touch-force recognition accuracies of 74.8% (SD = 14.2), 78.2% (SD = 15.7), 85.0% (SD = 13.0), and 85.5% (SD = 13.2) for the thumb-to-finger gesture set, and those limited to the top, middle and bottom phalanges, respectively. The results indicate that AudioTouch can recognize two levels of touch-force in thumb-to-finger gestures of middle and bottom phalanges with 85.0% and 85.5% recognition accuracies, respectively. Note that the purpose with Study 2 was to explore the advanced potential and limitations of AudioTouch. As a result, it may be difficult to recognize two levels of touch-force in the thumb-to-finger gesture set (i.e., 25 gestures).

Study 3: Effect of Re-mounting of Sensors on Gesture Recognition Accuracy

An important limitation of many wearable gesture recognition approaches is the need to calibrate and retrain the classifier every time the device is taken off then placed on the user's body before it can be used. Since the operating principle of AudioTouch should not be overly sensitive to

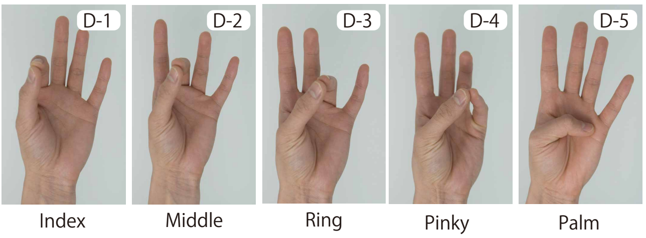


Figure 8: Pinch gesture set.

the re-mounting of the piezo elements, we conducted a study to assess the cross-session robustness of AudioTouch. We used the pinch gesture set shown in Figure 8.

Procedure. We first asked participants to coarsely mark the position of the two piezo elements (see Figure 4b, c). These elements were then taken off and re-attached after a short period (roughly 1–3 minutes). The presentation order of the gestures was randomized. We asked the participants to perform these gestures once per session. Each participant completed five rounds of ten sessions. We asked the participants to take a break for at least one minute and to remove then re-attach the piezo elements after every round. We collected the following number of samples: 3 participants × 5 rounds × 10 sessions × 6 gestures × 20 samples = 18,000 samples.

Results. AudioTouch yielded an average recognition accuracy of 76.2% (SD = 8.8) for the pinch gesture set in a leave-one-round-out cross-validation setting across all participants. This result indicates that AudioTouch has the potential of recognizing gestures even if the piezo elements accidentally come off the BoH by re-attaching them to almost the same locations. Precise sensor (re-)mounting was assumed in many previous work (and ours), and we leave an exhaustive study of the effect of misplaced sensors on reproducibility for future work.

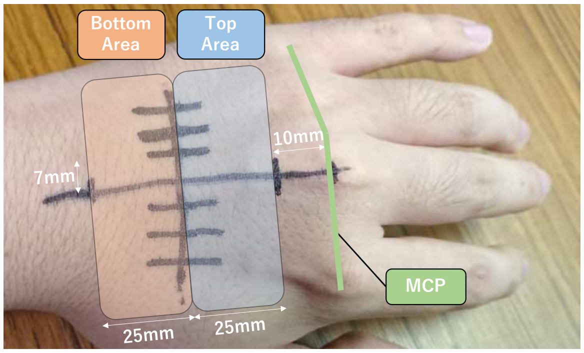


Figure 9: Marked lines on the BoH for Study 4.

Study 4: Effect of Sensor Location on Gesture Recognition Accuracy

In Studies 1–3, we used specific sensor locations (10 mm from the MCP joints with the two elements separated by 10 mm), which we determined with a trial-and-error approach. However, other sensor locations may yield better gesture recognition accuracies of AudioTouch.

To systematically investigate the effect of sensor locations on gesture recognition accuracy of AudioTouch, we used pinch gestures (Figure 8) at the ten locations (L1–L10) shown in Figures 9 and 10. At L1–L6, the extensor tendon of the middle finger serves as the centerline between the two elements. In this experiment, we used distances of 14, 28, and 42 mm (vertical lines in Figure 9 are drawn at intervals of 7 mm to avoid overlapping of piezo elements whose width is 12 mm). At L7–L10, the piezo elements were attached to only the left or right side of the BoH. We used two areas as vertical locations of the sensor: top and bottom. The top area is 10 mm from the MCP joint of the middle finger. The bottom area is 35 mm (i.e., 10 mm plus 25 mm, which is the height of a piezo element). We determined the two vertical areas to avoid the effect of the MCP joints and wrist.

Procedure. We first asked participants to mark the lines for ten sensor locations (see Figure 9) then conduct ten sessions at each sensor location. The presentation order of the pinch gestures (Figure 8) was randomized, and we asked participants to perform these gestures once per session. After finishing the ten sessions, the piezo elements were then taken off and re-attached to an indicated next location. We used a Latin square design (e.g., $O_{R1} = \{L_1, L_2, ..., L_{10}\}$, $O_{Rn} = \{L_n, L_{n+1}, ..., L_{10}, L_1, ..., L_{n-1}\}$) as the location order to remove the order effect because this study had many conditions. Each participant completed ten rounds of ten sessions each. We asked the participants to take a break for at least one minute and to remove then re-attach the piezo elements after every round. We collected the following number of samples:

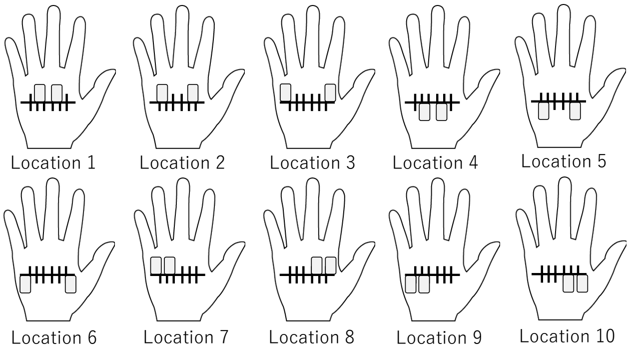


Figure 10: Sensor locations.

10 participants × 10 locations × 10 sessions × 6 gestures × 20 samples = 120,000 samples.

Results. We averaged the gesture recognition accuracies using a leave-one-session-out cross-validation for a location of a single participant and averaged the accuracies across participants. The results are shown in Figure 11. To compare the gesture recognition accuracies between locations, we conducted a one-way ANOVA. The test showed significant differences in these accuracies between locations ($F(9, 90) = 1.99$, $p = .049 < .05$). We also conducted a Tukey's HSD test, which showed that the recognition accuracy at L3 (96.6%) was significantly higher than that at L9 (90.6%).

The results indicate that AudioTouch's gesture recognition accuracies were approximately 95% at L1 – L8. However, those at L9 and L10 were 90.6% and 91.8% (approximately 90%), respectively. Therefore, L9 (bottom, only left or right side) may not be suitable as a sensor location on the BoH. In addition, L10 may not be suitable as a sensor location, though there were no significant differences in the recognition accuracy between L10 and L1–L9.

5 DISCUSSION

Recognition Robustness and Accuracy

In Study 1, we instructed a hand orientation (i.e., parallel or perpendicular) for each micro-gesture. That is, micro-gestures were performed with different hand orientations in Study 1. The hand orientation might affect the micro-gesture recognition accuracies across different gesture sets since the wrist rotation affects recognition. Therefore, further studies on the effect of hand orientation on micro-gesture recognition accuracy are necessary.

While AudioTouch achieved 89.3% micro-gesture recognition accuracy with a gesture set in Study 1, this accuracy with all gesture sets remained relatively low. To address this issue, we calculated confusion matrices of recognition for all gesture sets (Figure 12 left) and found that micro-gestures involving the pinky (i.e., A-4, A-8, A-12 in Figure 5 in the thumb-to-finger gesture set) tends to be confused with other gestures, whose average accuracies are the lowest among four fingers (index: 87.8%, middle: 89.4%, ring: 85.0%, pinky: 78.3%). To examine this effect, we calculated recognition accuracy by removing the pinky finger gestures. The resulting accuracy was 92.1% with 10 gestures of the thumb-to-finger gesture set. We also calculated the accuracy for all micro-gestures without these three micro-gestures. The result shows an accuracy of 86.6% with 21 micro-gestures (Figure 12 right).

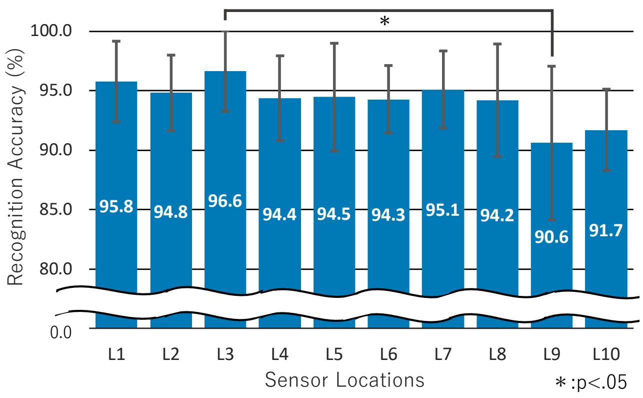


Figure 11: Gesture recognition accuracies at each location. Each bar means ± one SD between participants.

Compared to the gesture recognition accuracy of AudioTouch, that of Tomo [57], which uses 32 electrodes on a user's wrist, was 94.3% for five pinch gestures. In contrast, AudioTouch recognized ten thumb-to-finger gestures with 92.1% micro-gesture recognition accuracy.

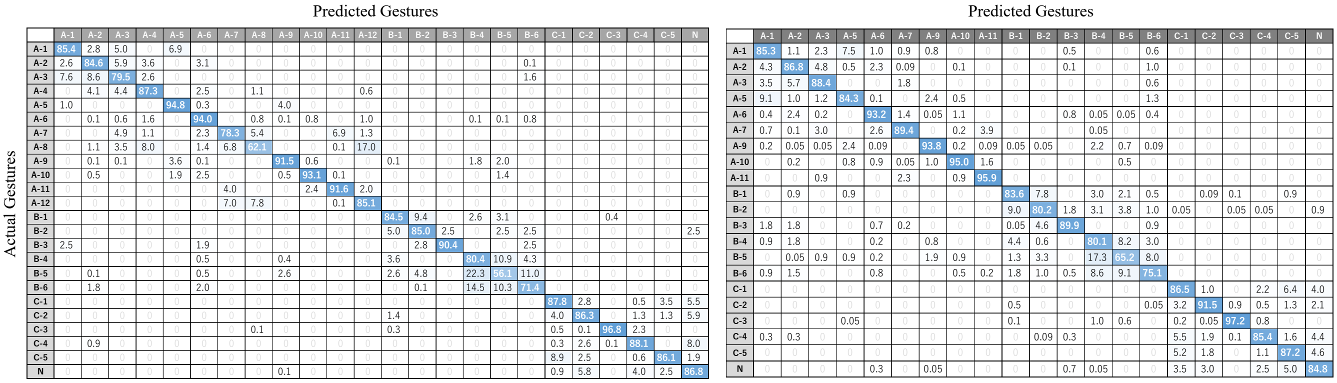
Other Sensor Location

We conducted another pilot study to compare the BoH to one location on a wrist (the outer wrist) with four participants and the six gestures (pinch gesture set and a neutral state) shown in Figures 7 and 8. AudioTouch yielded a gesture recognition accuracy of 93.1% for the BoH and 82.4% for the wrist. A dependent t-test also showed that this accuracy for the BoH was higher ($p = .003 < .05$). Although, this pilot study compared these accuracies at one location on the BoH and wrist, we felt that it might be difficult for AudioTouch to recognize small differences between various finger gestures when the sensor location is on the wrist and that the BoH would be a more suitable location.

Limitations

While we believe that AudioTouch is a promising and interesting direction for always-available, wearable interaction, it is not without limitations. Of course, further engineering would be required to further miniaturize the setup and improve robustness and accuracy. AudioTouch is currently not robust in the cross-user setting: it currently needs to be trained for each user. Considering the significant differences in body type, skin, and gesture performance, this is to be expected. While we believe that advanced machine-learning methods can improve this situation, addressing this fully is beyond this initial exploration.

External pressure changes the vibration patterns picked-up by the piezo-electric microphone. Hence, band-type mounting solutions would be problematic. However, retraining the classifier or custom mounts could alleviate these issues. We also assume that holding heavy or vibrating objects would affect gesture recognition, but we have not yet investigated this.



	A-1	A-2	A-3	A-4	A-5	A-6	A-7	A-8	A-9	A-10	A-11	A-12	B-1	B-2	B-3	B-4	B-5	B-6	C-1	C-2	C-3	C-4	C-5	N
A-1	85.4	2.8	5.0	0	6.9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A-2	2.6	84.6	5.9	3.6	0	3.1	0	0	0	0	0	0	0	0	0	0	0	0.1	0	0	0	0	0	0
A-3	7.6	8.6	79.5	2.6	0	0	0	0	0	0	0	0	0	0	0	0	0	1.6	0	0	0	0	0	0
A-4	0	4.1	4.4	87.3	0	2.5	0	1.1	0	0	0	0.6	0	0	0	0	0	0	0	0	0	0	0	0
A-5	1.0	0	0	0	94.8	0.3	0	0	4.0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
A-6	0	0.1	0.6	1.6	0	94.0	0	0.8	0.1	0.8	0	1.0	0	0	0	0.1	0.1	0.8	0	0	0	0	0	0
A-7	0	0	4.9	1.1	0	2.3	78.3	5.4	0	0	6.9	1.3	0	0	0	0	0	0	0	0	0	0	0	0
A-8	0	1.1	3.5	8.0	0	1.4	6.8	62.1	0	0	0.1	17.0	0	0	0	0	0	0	0	0	0	0	0	0
A-9	0	0.1	0.1	0	3.6	0.1	0	0	91.5	0.6	0	0	0.1	0	0	1.8	2.0	0	0	0	0	0	0	0
A-10	0	0.5	0	0	1.9	2.5	0	0	0.5	93.1	0.1	0	0	0	0	0	1.4	0	0	0	0	0	0	0
A-11	0	0	0	0	0	0	4.0	0	0	2.4	91.6	2.0	0	0	0	0	0	0	0	0	0	0	0	0
A-12	0	0	0	0	0	0	7.0	7.8	0	0	0.1	85.1	0	0	0	0	0	0	0	0	0	0	0	0
B-1	0	0	0	0	0	0	0	0	0	0	0	0	84.5	9.4	0	2.6	3.1	0	0	0	0.4	0	0	0
B-2	0	0	0	0	0	0	0	0	0	0	0	0	5.0	85.0	2.5	0	2.5	2.5	0	0	0	0	0	2.5
B-3	2.5	0	0	0	0	1.9	0	0	0	0	0	0	0	2.8	90.4	0	0	2.5	0	0	0	0	0	0
B-4	0	0	0	0	0	0.5	0	0	0.4	0	0	0	3.6	0	0	80.4	10.9	4.3	0	0	0	0	0	0
B-5	0	0.1	0	0	0	0.5	0	0	2.6	0	0	0	2.6	4.8	0	22.3	56.1	11.0	0	0	0	0	0	0
B-6	0	1.8	0	0	0	2.0	0	0	0	0	0	0	0	0.1	0	14.5	10.3	71.4	0	0	0	0	0	0
C-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	87.8	2.8	0	0.5	3.5	5.5
C-2	0	0	0	0	0	0	0	0	0	0	0	0	1.4	0	0	0	0	0	4.0	86.3	0	1.3	1.3	5.9
C-3	0	0	0	0	0	0	0	0.1	0	0	0	0	0.3	0	0	0	0	0	0.5	0.1	96.8	2.3	0	0
C-4	0	0.9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3	2.6	0.1	88.1	0	8.0
C-5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8.9	2.5	0	0.6	86.1	1.9
N	0	0	0	0	0	0	0	0	0.1	0	0	0	0	0	0	0	0	0	0.9	5.8	0	4.0	2.5	86.8

	A-1	A-2	A-3	A-5	A-6	A-7	A-9	A-10	A-11	B-1	B-2	B-3	B-4	B-5	B-6	C-1	C-2	C-3	C-4	C-5	N
A-1	85.3	1.1	2.3	7.5	1.0	0.9	0.8	0	0	0	0	0.5	0	0	0.6	0	0	0	0	0	0
A-2	4.3	86.8	4.8	0.5	2.3	0.09	0	0.1	0	0	0	0.1	0	0	1.0	0	0	0	0	0	0
A-3	3.5	5.7	88.4	0	0	1.8	0	0	0	0	0	0	0	0	0.6	0	0	0	0	0	0
A-5	9.1	1.0	1.2	84.3	0.1	0	2.4	0.5	0	0	0	0	0	0	1.3	0	0	0	0	0	0
A-6	0.4	2.4	0.2	0	93.2	1.4	0.05	1.1	0	0	0	0.8	0.05	0.05	0.4	0	0	0	0	0	0
A-7	0.7	0.1	3.0	0	2.6	89.4	0	0.2	3.9	0	0	0	0.05	0	0	0	0	0	0	0	0
A-9	0.2	0.05	0.05	2.4	0.09	0	93.8	0.2	0.09	0.05	0.05	0	2.2	0.7	0.09	0	0	0	0	0	0
A-10	0	0.2	0	0.8	0.9	0.05	1.0	95.0	1.6	0	0	0	0	0.5	0	0	0	0	0	0	0
A-11	0	0	0.9	0	0	2.3	0	0.9	95.9	0	0	0	0	0	0	0	0	0	0	0	0
B-1	0	0.9	0	0.9	0	0	0	0	0	83.6	7.8	0	3.0	2.1	0.5	0	0.09	0.1	0	0.9	0
B-2	0	0	0	0	0	0	0	0	0	9.0	80.2	1.8	3.1	3.8	1.0	0.05	0	0.05	0.05	0	0.9
B-3	1.8	1.8	0	0	0.7	0.2	0	0	0	0.05	4.6	89.9	0	0	0.9	0	0	0	0	0	0
B-4	0.9	1.8	0	0	0.2	0	0.8	0	0	4.4	0.6	0	80.1	8.2	3.0	0	0	0	0	0	0
B-5	0	0.05	0.9	0.9	0.2	0	1.9	0.9	0	1.3	3.3	0	17.3	65.2	8.0	0	0	0	0	0	0
B-6	0.9	1.5	0	0	0.8	0	0	0.5	0.2	1.8	1.0	0.5	8.6	9.1	75.1	0	0	0	0	0	0
C-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	86.5	1.0	0	2.2	6.4	4.0
C-2	0	0	0	0	0	0	0	0	0	0.5	0	0	0	0	0.05	3.2	91.5	0.9	0.5	1.3	2.1
C-3	0	0	0	0.05	0	0	0	0	0	0.1	0	0	1.0	0.6	0	0.2	0.05	97.2	0.8	0	0
C-4	0.3	0.3	0	0	0	0	0	0	0	0	0.09	0.3	0	0	0	5.5	1.9	0.1	85.4	1.6	4.4
C-5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5.2	1.8	0	1.1	87.2	4.6
N	0	0	0	0	0.3	0	0.05	0	0	0	0	0.7	0.05	0	0	3.5	3.0	0	2.5	5.0	84.8

Figure 12: Confusion matrices of micro-gesture recognition accuracy in Study 1. Left: all gestures, right: all gestures without pinky finger gestures in thumb-to-finger gesture set.

Environmental Factors Affecting Micro-Gesture Recognition Accuracy

Environmental resistance is generally important in gesture recognition of wearable devices. In a situation in which continuous vibration occurs, a sensing method for detecting such environmental factors (e.g., a method using an accelerometer, which is used in lifelog applications) could be used to reduce such effects by classifying the noise-floor level and subsequently adjusting the classifier. However, AudioTouch is free from problems such as the effect of external light and limitations of the viewing angle of camera-based methods and does not require full instrumentation of the hand as is the case with glove-type methods.

Acceptability

Our current proof-of-concept implementation uses large piezo elements and a home-made mounting mechanism, resulting in higher invasiveness than bracelet type sensors, which are not good at recognizing micro-gestures. However, we believe that there is a promising path towards miniaturization, potentially reducing the form factor to tattoo-like patches, such as DuoSkin [22], since the piezo element is extremely thin (0.23 mm). There are also flexible piezo elements as well as flexible and thin cables, which were used to implement SkinMarks [49]. If AudioTouch adopts such flexible components, acceptability would become higher.

Portability

One of the most appealing properties of AudioTouch is the possibility to use it as an input mechanism for mobile and wearable devices. Our experiments show that our proof-of-concept implementation achieved good gesture recognition accuracy in challenging settings. We also showed that AudioTouch can be miniaturized and made entirely self-contained, as shown in Figure 13. The module has a programmable wave generator to emit sinusoidal sweep signals from 20–40 kHz. The setup is connected to the host via Bluetooth and is entirely powered by a 3.7 V, 400 mAh battery. This unoptimized prototype can be continuously used for 1 hour, and the latency is about 0.1 seconds. We believe that miniaturization and runtime improvement are straightforward engineering tasks.

Machine-Learning Algorithm

We used a well established and simple machine-learning algorithm with default parameters for the proof-of-concept implementation, demonstrating that the sensing approach can be used with current methods. To further improve gesture recognition accuracies of AudioTouch, adopting domain specific recognition approaches or more powerful machine-learning methods would be a choice. Nonetheless, one aspect of the SVM is that it achieves high accuracy with a small set of data, allowing end-user retraining of the classifier.

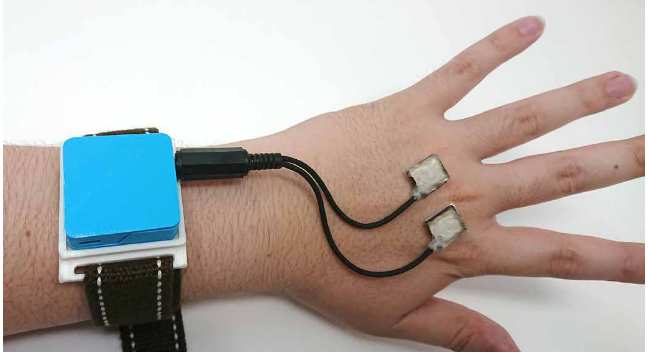


Figure 13: Wearable type of AudioTouch.

Future Work

In Studies 1 – 3, we used specific sensor locations (10 mm from MCP joints and elements between 10 mm) to conduct experiments focused on micro-gesture and touch-force recognition accuracies of AudioTouch. In Study 4, we evaluated the affect of ten sensor locations on gesture recognition accuracy; however, there are many conditions other than sensor locations that could not be evaluated. For future work, it is necessary to explore the effects of sensor conditions such as other locations (e.g., wrist and arm), angles, sizes, and shapes to evaluate gesture recognition accuracy. We also plan to explore how sensitive is AudioTouch is to different locations.

We conducted all experiments in controlled environments. Therefore, it is necessary to verify user factors (e.g., effects of user's fatigue, touching/removing elbow from the desk, wrist bending angle, wrist and arm movement, and body movement, e.g., walking) in subsequent studies. In addition, each user has different physical conditions such as skin, muscle, and age. Therefore, we need to conduct user studies with participants of diverse demographics (e.g., wider age ranges or better representation of gender).

6 INTERACTIVE SCENARIOS

To envision utility of AudioTouch, we briefly discuss potential interactive scenarios.

AudioTouch can be leveraged to robustly detect a rich repertoire of micro-gestures only involving actuation of individual fingers but no wrist or arm movement. This allows for discrimination of pressure, further expanding the interaction vocabulary. The lightweight and small form-factor of AudioTouch forms a straightforward path to real-world implementation. These unique properties enable compelling interactive scenarios for which we implemented proof-of-concept prototypes.

While the below proof-of-concept implementations are simple, they demonstrate the general purpose of AudioTouch, and we believe it can be used in many on-the-go situations that currently require a user to retrieve a smartphone or tablet from his/her pocket or use a special purpose input device such as a remote control. AudioTouch may also enable rich interactions in AR/VR, industrial, and automotive settings.

Number keypad

AudioTouch can be leveraged as an efficient means of numerical input. Mapping a set of numerical keys onto the phalanges of the user's in combination with thumb-to-finger gestures enabled with AudioTouch allows for efficient number input, providing a natural analogy to a physical or touch number keypad. In our implementation (see Figure 1e), we map each finger to a row of a number keypad, then each phalange is used as a key.

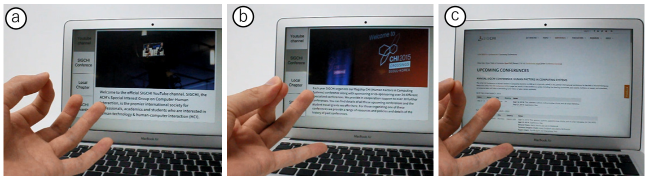


Figure 14: Menu selection application. (a) Soft-press pinch brings up menu and enables cycling through menu items. (b) Hard-press pinch selects menu item and triggers corresponding action. (c) Result of selecting menu item.

Menu Selection

Using menuing systems on remote devices such as smart TVs can be a challenge with conventional input methods. The capability of AudioTouch to discriminate pressure allows for the implementation of an efficient menu selection application. For example, a soft-press pinch may bring up a menu and enable cycling through a menu items, whereas a hard-press of the same gesture selects a menu item, as illustrated in Figure 14.

7 CONCLUSION

In this paper, we presented AudioTouch, a minimally invasive approach for sensing micro-gestures for always available input in on-the-go interaction settings. This approach leverages an active bio-acoustic sensing scheme and robustly recognizes a rich repertoire of micro-gestures involving only actuation of individual fingers. The sensing scheme is entirely self-contained and does not rely on any instrumentation of the environment or the user's fingers and palm. AudioTouch also allows for discrimination of force, further expanding the interaction vocabulary.

We conducted experiments to evaluate the micro-gesture recognition accuracy of AudioTouch. Such accuracies were 92.1% and 86.6% with 10 gestures of the thumb-to-finger gesture set and 21 gestures, respectively. We also demonstrated that AudioTouch is capable of discriminating two levels of touch-force in the middle and bottom phalanges of thumb-to-finger gestures with 85.0% and 85.5% recognition accuracies. Except for one location, we did not find any difference in the gesture recognition accuracy of AudioTouch. We also implemented proof-of-concept interactive scenarios and showed how AudioTouch can be made fully wearable and wireless.

8 ACKNOWLEDGEMENTS

This work was partially supported by Japan Science and Technology Agency (JST) ACT-I Grant Number JPMJPR16UA.

REFERENCES

[1] Shuhei Aoyama, Buntarou Shizuki, and Jiro Tanaka. 2016. ThumbSlide: An Interaction Technique for Smartwatches Using a Thumb Slide Movement. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16)*. ACM, New York, NY, USA, 2403–2409. https://doi.org/10.1145/2851581.2892435

[2] Riley Booth and Peter Goldsmith. 2017. Detecting Finger Gestures with a Wrist Worn Piezoelectric Sensor Array. In *Proceedings of the 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC '17)*. 3665–3670.

[3] Liwei Chan, Yi-Ling Chen, Chi-Hao Hsieh, Rong-Hao Liang, and Bing-Yu Chen. 2015. CyclopsRing: Enabling Whole-Hand and Context-aware Interactions through a Fisheye Ring. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology (UIST '15)*. ACM, New York, NY, USA, 549–556. https://doi.org/10.1145/2807442.2807450

[4] CyberGlove Systems Inc. 2009. CyberTouch Glove. http://www.cyberglovesystems.com/cybertouch. Last accessed: 18.03.2017.

[5] Artem Dementyev and Joseph A. Paradiso. 2014. WristFlex: Low-power Gesture Input with Wrist-worn Pressure Sensors. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. ACM, New York, NY, USA, 161–166. https://doi.org/10.1145/2642918.2647396

[6] Travis Deyle, Szabolcs Palinko, Erika Shehan Poole, and Thad Starner. 2007. Hambone: A Bio-Acoustic Gesture Interface. In *Proceedings of the 11th IEEE International Symposium on Wearable Computers*. IEEE, 3–10.

[7] Laura Dipietro, Angelo M. Sabatini, and Paolo Dario. 2008. A Survey of Glove-Based Systems and Their Applications. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 38, 4 (July 2008), 461–482. https://doi.org/10.1109/TSMCC.2008.923862

[8] Mayank Goel, Jacob Wobbrock, and Shwetak Patel. 2012. GripSense: Using Built-in Sensors to Detect Hand Posture and Pressure on Commodity Mobile Phones. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology (UIST '12)*. ACM, New York, NY, USA, 545–554. https://doi.org/10.1145/2380116.2380184

[9] Jun Gong, Yang Zhang, Xia Zhou, and Xing-Dong Yang. 2017. Pyro: Thumb-Tip Gesture Recognition Using Pyroelectric Infrared Sensing. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17)*. ACM, New York, NY, USA, 553–563. https://doi.org/10.1145/3126594.3126615

[10] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H. Witten. 2009. The WEKA Data Mining Software: An Update. *SIGKDD Explorations Newsletter* 11, 1 (2009), 10–18. https://doi.org/10.1145/1656274.1656278

[11] Teng Han, Khalad Hasan, Keisuke Nakamura, Randy Gomez, and Pourang Irani. 2017. SoundCraft: Enabling Spatial Interactions on Smartwatches using Hand Generated Acoustics. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17)*. ACM, New York, NY, USA, 579–591. https://doi.org/10.1145/3126594.3126612

[12] Chris Harrison, Hrvoje Benko, and Andrew D. Wilson. 2011. OmniTouch: Wearable Multitouch Interaction Everywhere. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST '11)*. ACM, New York, NY, USA, 441–450. https://doi.org/10.1145/2047196.2047255

[13] Chris Harrison and Scott E. Hudson. 2008. Scratch Input: Creating Large, Inexpensive, Unpowered and Mobile Finger Input Surfaces. In *Proceedings of the 21st Annual ACM Symposium on User Interface Software and Technology (UIST '08)*. ACM, New York, NY, USA, 205–208. https://doi.org/10.1145/1449715.1449747

[14] Chris Harrison, Desney Tan, and Dan Morris. 2010. Skinput: Appropriating the Body As an Input Surface. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 453–462. https://doi.org/10.1145/1753326.1753394

[15] Chris Harrison, Robert Xiao, and Scott Hudson. 2012. Acoustic Barcodes: Passive, Durable and Inexpensive Notched Identification Tags. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology (UIST '12)*. ACM, New York, NY, USA, 563–568. https://doi.org/10.1145/2380116.2380187

[16] Seongkook Heo and Geehyuk Lee. 2011. Force Gestures: Augmenting Touch Screen Gestures with Normal and Tangential Forces. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST '11)*. ACM, New York, NY, USA, 621–626. https://doi.org/10.1145/2047196.2047278

[17] Seongkook Heo and Geehyuk Lee. 2011. ForceTap: Extending the Input Vocabulary of Mobile Touch Screens by adding Tap Gestures. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '11)*. ACM, New York, NY, USA, 113–122. https://doi.org/10.1145/2037373.2037393

[18] Donny Huang, Xiaoyi Zhang, T. Scott Saponas, James Fogarty, and Shyamnath Gollakota. 2015. Leveraging Dual-Observable Input for Fine-Grained Thumb Interaction using Forearm EMG. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software and Technology (UIST '15)*. ACM, New York, NY, USA, 523–528. https://doi.org/10.1145/2807442.2807506

[19] Da-Yuan Huang, Liwei Chan, Shuo Yang, Fan Wang, Rong-Hao Liang, De-Nian Yang, Yi-Ping Hung, and Bing-Yu Chen. 2016. DigitSpace: Designing Thumb-to-Fingers Touch Interfaces for One-Handed and Eyes-Free Interactions. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1526–1537. https://doi.org/10.1145/2858036.2858483

[20] Sungjae Hwang, Andrea Bianchi, and Kwangyun Wohn. 2012. MicPen: Pressure-sensitive Pen Interaction using Microphone with Standard Touchscreen. In *Proceedings of the 30th International SIGCHI Conference Extended Abstracts on Human Factors in Computing System (CHI EA '12)*. ACM, New York, NY, USA, 1847–1852. https://doi.org/10.1145/2212776.2223717

[21] Sungjae Hwang and Kwang-yun Wohn. 2012. PseudoButton: Enabling Pressure-Sensitive Interaction by Repurposing Microphone on Mobile Device. In *Proceedings of the 30th International SIGCHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '12)*. ACM, New York, NY, USA, 1565–1570. https://doi.org/10.1145/2212776.2223673

[22] Hsin-Liu (Cindy) Kao, Christian Holz, Asta Roseway, Andres Calvo, and Chris Schmandt. 2016. DuoSkin: Rapidly Prototyping On-skin User Interfaces Using Skin-friendly Materials. In *Proceedings of the 2016 ACM International Symposium on Wearable Computers (ISWC '16)*. ACM, New York, NY, USA, 16–23. https://doi.org/10.1145/2971763.2971777

[23] Hiroyuki Kato and Kentaro Takemura. 2016. Hand Pose Estimation based on Active Bone-conducted Sound Sensing. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct (UbiComp '16)*. ACM, New York, NY, USA, 109–112. https://doi.org/10.1145/2968219.2971403

[24] David Kim, Otmar Hilliges, Shahram Izadi, Alex D. Butler, Jiawen Chen, Iason Oikonomidis, and Patrick Olivier. 2012. Digits: Freehand 3D Interactions Anywhere using a Wrist-worn Gloveless Sensor. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology (UIST '12)*. ACM, New York, NY, USA, 167–176. https://doi.org/10.1145/2380116.2380139

[25] Gierad Laput, Eric Brockmeyer, Scott E. Hudson, and Chris Harrison. 2015. Acoustruments: Passive, Acoustically-Driven, Interactive Controls for Handheld Devices. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 2161–2170. https://doi.org/10.1145/2702123.2702414

[26] Gierad Laput, Robert Xiao, and Chris Harrison. 2016. ViBand: High-Fidelity Bio-Acoustic Sensing Using Commodity Smartwatch Accelerometers. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)*. ACM, New York, NY, USA, 321–333. https://doi.org/10.1145/2984511.2984582

[27] LEAP MOTION, INC. 2012. LEAP MOTION. https://www.leapmotion.com. Last accessed: 01.02.2019.

[28] Hong Li, Wei Yang, Jianxin Wang, Yang Xu, and Liusheng Huang. 2016. WiFinger: Talk to Your Smart Devices with Finger-grained Gesture. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. ACM, New York, NY, USA, 250–261. https://doi.org/10.1145/2971648.2971738

[29] Jaime Lien, Nicholas Gillian, M. Emre Karagozler, Patrick Amihood, Carsten Schwesig, Erik Olson, Hakim Raja, and Ivan Poupyrev. 2016. Soli: Ubiquitous Gesture Sensing with Millimeter Wave Radar. *ACM Transactions on Graphics* 35, 4, Article 142 (July 2016), 19 pages. https://doi.org/10.1145/2897824.2925953

[30] Jhe-Wei Lin, Chiuan Wang, Yi Yao Huang, Kuan-Ting Chou, Hsuan-Yu Chen, Wei-Luan Tseng, and Mike Y. Chen. 2015. BackHand: Sensing Hand Gestures via Back of the Hand. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology (UIST '15)*. ACM, New York, NY, USA, 557–564. https://doi.org/10.1145/2807442.2807462

[31] Jess McIntosh, Asier Marzo, and Mike Fraser. 2017. SensIR: Detecting Hand Gestures with a Wearable Bracelet Using Infrared Transmission and Reflection. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology (UIST '17)*. ACM, New York, NY, USA, 593–597. https://doi.org/10.1145/3126594.3126604

[32] Jess McIntosh, Asier Marzo, Mike Fraser, and Carol Phillips. 2017. EchoFlex: Hand Gesture Recognition Using Ultrasound Imaging. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 1923–1934. https://doi.org/10.1145/3025453.3025807

[33] Jess McIntosh, Charlie McNeill, Mike Fraser, Frederic Kerber, Markus Löchtefeld, and Antonio Krüger. 2016. EMPress: Practical Hand Gesture Classification with Wrist-Mounted EMG and Pressure Sensing. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 2332–2342. https://doi.org/10.1145/2858036.2858093

[34] Adiyan Mujibiya, Xiang Cao, Desney S. Tan, Dan Morris, Shwetak N. Patel, and Jun Rekimoto. 2013. The Sound of Touch: On-body Touch and Gesture Sensing Based on Transdermal Ultrasound Propagation. In *Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces (ITS '13)*. ACM, New York, NY, USA, 189–198. https://doi.org/10.1145/2512349.2512821

[35] NeuroDigital Technologies. 2015. Gloveone Glove. https://www.neurodigital.es/gloveone/. Last accessed: 01.02.2019.

[36] North Inc. 2014. Myo. https://www.bynorth.com/about. Last accessed: 01.02.2019.

[37] Makoto Ono, Buntarou Shizuki, and Jiro Tanaka. 2013. Touch & Activate: Adding Interactivity to Existing Objects Using Active Acoustic Sensing. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology (UIST '13)*. ACM, New York, NY, USA, 31–40. https://doi.org/10.1145/2501988.2501989

[38] Makoto Ono, Buntarou Shizuki, and Jiro Tanaka. 2015. Sensing Touch Force using Active Acoustic Sensing. In *Proceedings of the 9th International Conference on Tangible, Embedded, and Embodied Interaction (TEI '15)*. ACM, New York, NY, USA, 355–358. https://doi.org/10.1145/2677199.2680585

[39] Manuel Prätorius, Dimitar Valkov, Ulrich Burgbacher, and Klaus Hinrichs. 2014. DigiTap: An Eyes-free VR/AR Symbolic Input Device. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology (VRST '14)*. ACM, New York, NY, USA, 9–18. https://doi.org/10.1145/2671015.2671029

[40] Jun Rekimoto. 2001. GestureWrist and GesturePad: Unobtrusive Wearable Interaction Devices. In *Proceedings of the International Symposium on Wearable Computers (ISWC' 01)*.

[41] T. Scott Saponas, Desney S. Tan, Dan Morris, Ravin Balakrishnan, Jim Turner, and James A. Landay. 2009. Enabling Always-available Input with Muscle-computer Interfaces. In *Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology (UIST '09)*. ACM, New York, NY, USA, 167–176. https://doi.org/10.1145/1622176.1622208

[42] Munehiko Sato, Ivan Poupyrev, and Chris Harrison. 2012. Touché: Enhancing Touch Interaction on Humans, Screens, Liquids, and Everyday Objects. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, New York, NY, USA, 483–492. https://doi.org/10.1145/2207676.2207743

[43] Jie Song, Gábor Sörös, Fabrizio Pece, Sean Ryan Fanello, Shahram Izadi, Cem Keskin, and Otmar Hilliges. 2014. In-air Gestures Around Unmodified Mobile Devices. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. ACM, New York, NY, USA, 319–329. https://doi.org/10.1145/2642918.2647373

[44] Srinath Sridhar, Anders Markussen, Antti Oulasvirta, Christian Theobalt, and Sebastian Boring. 2017. WatchSense: On- and Above-Skin Input Sensing Through a Wearable Depth Sensor. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 3891–3902. https://doi.org/10.1145/3025453.3026005

[45] Ryosuke Takada, Wei Lin, Toshiyuki Ando, Buntarou Shizuki, and Shin Takahashi. 2017. A Technique for Touch Force Sensing using a Waterproof Device's Built-in Barometer. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17)*. ACM, New York, NY, USA, 2140–2146. https://doi.org/10.1145/3027063.3053130

[46] Jonathan Taylor, Lucas Bordeaux, Thomas Cashman, Bob Corish, Cem Keskin, Toby Sharp, Eduardo Soto, David Sweeney, Julien Valentin, Benjamin Luff, Arran Topalian, Erroll Wood, Sameh Khamis, Pushmeet Kohli, Shahram Izadi, Richard Banks, Andrew Fitzgibbon, and Jamie Shotton. 2016. Efficient and Precise Interactive Hand Tracking Through Joint, Continuous Optimization of Pose and Correspondences. *ACM Transactions on Graphics* 35, 4, Article 143 (July 2016), 12 pages. https://doi.org/10.1145/2897824.2925965

[47] Qinglong Wang, Xiangshi Ren, and Xiaoying Sun. 2016. EV-Pen: an Electrovibration Haptic Feedback Pen for Touchscreens. In *SIGGRAPH ASIA 2016 Emerging Technologies (SA '16)*. ACM, New York, NY, USA, Article 8, 2 pages. https://doi.org/10.1145/2988240.2988241

[48] Hiroki Watanabe, Tsutomu Terada, and Masahiko Tsukamoto. 2017. Gesture Recognition Method utilizing Ultrasonic Active Acoustic Sensing. *Journal of Information Processing* 25 (2017), 331–340. https://doi.org/10.2197/ipsjjip.25.331

[49] Martin Weigel, Aditya Shekhar Nittala, Alex Olwal, and Jürgen Steimle. 2017. SkinMarks: Enabling Interactions on Body Landmarks Using Conformal Skin Electronics. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 3095–3105. https://doi.org/10.1145/3025453.3025704

[50] Eric Whitmire, Mohit Jain, Divye Jain, Greg Nelson, Ravi Karkar, Shwetak Patel, and Mayank Goel. 2017. DigiTouch: Reconfigurable Thumb-to-Finger Input and Text Entry on Head-mounted Displays. *Proceedings*

of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 1, 3, Article 113 (Sept. 2017), 21 pages. https://doi.org/10.1145/3130978

[51] Pui Chung Wong, Kening Zhu, and Hongbo Fu. 2018. FingerT9: Leveraging Thumb-to-finger Interaction for Same-side-hand Text Entry on Smartwatches. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 178, 10 pages. https://doi.org/10.1145/3173574.3173752

[52] Tomohiro Yokota and Tomoko Hashida. 2016. Hand Gesture and On-body Touch Recognition by Active Acoustic Sensing throughout the Human Body. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16 Adjunct)*. ACM, New York, NY, USA, 113–115. https://doi.org/10.1145/2984751.2985721

[53] Sugiura Yuta, Nakamura Fumihiko, Kawai Wataru, Kikuchi Takashi, and Sugimoto Maki. 2017. Behind The Palm: Hand Gesture Recognition through Measuring Skin Deformation on Back of Hand by using Optical Sensors. In *56th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE) (SICE '17)*. 1082–1087. https://doi.org/10.23919/SICE.2017.8105457

[54] Cheng Zhang, Xiaoxuan Wang, Anandghan Waghmare, Sumeet Jain, Thomas Ploetz, Omer T. Inan, Thad E. Starner, and Gregory D. Abowd. 2017. FingOrbits: Interaction with Wearables using Synchronized Thumb Movements. In *Proceedings of the 2017 ACM International Symposium on Wearable Computers (ISWC '17)*. ACM, New York, NY, USA, 62–65. https://doi.org/10.1145/3123021.3123041

[55] Cheng Zhang, Qiuyue Xue, Anandghan Waghmare, Ruichen Meng, Sumeet Jain, Yizeng Han, Xinyu Li, Kenneth Cunefare, Thomas Ploetz, Thad Starner, Omer Inan, and Gregory D. Abowd. 2018. FingerPing: Recognizing Fine-grained Hand Poses using Active Acoustic On-body Sensing. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, Article 437, 10 pages. https://doi.org/10.1145/3173574.3174011

[56] Yang Zhang and Chris Harrison. 2015. Tomo: Wearable, Low-Cost Electrical Impedance Tomography for Hand Gesture Recognition. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software and; Technology (UIST '15)*. ACM, New York, NY, USA, 167–173. https://doi.org/10.1145/2807442.2807480

[57] Yang Zhang, Robert Xiao, and Chris Harrison. 2016. Advancing Hand Gesture Recognition with High Resolution Electrical Impedance Tomography. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)*. ACM, New York, NY, USA, 843–850. https://doi.org/10.1145/2984511.2984574

[58] Yang Zhang, Junhan Zhou, Gierad Laput, and Chris Harrison. 2016. SkinTrack: Using the Body As an Electrical Waveguide for Continuous Finger Tracking on the Skin. In *Proceedings of the 34th SIGCHI Conference on Human Factors in Computing Systems (CHI '16)*. ACM, New York, NY, USA, 1491–1503. https://doi.org/10.1145/2858036.2858082