AtaTouch: Robust Finger Pinch Detection for a VR Controller Using RF Return Loss

Daehwa Kim
daehwakim@kaist.ac.kr
HCI Lab, School of Computing, KAIST
Daejeon, Republic of Korea

Keunwoo Park
keunwoo@kaist.ac.kr
HCI Lab, School of Computing, KAIST
Daejeon, Republic of Korea

Geehyuk Lee
geehyuk@gmail.com
HCI Lab, School of Computing, KAIST
Daejeon, Republic of Korea

ABSTRACT

Handheld controllers are an essential part of VR systems. Modern sensing techniques enable them to track users' finger movements to support natural interaction using hands. The sensing techniques, however, often fail to precisely determine whether two fingertips touch each other, which is important for the robust detection of a pinch gesture. To address this problem, we propose AtaTouch, which is a novel, robust sensing technique for detecting the closure of a finger pinch. It utilizes a change in the coupled impedance of an antenna and human fingers when the thumb and finger form a loop. We implemented a prototype controller in which AtaTouch detects the finger pinch of the grabbing hand. A user test with the prototype showed a finger-touch detection accuracy of 96.4%. Another user test with the scenarios of moving virtual blocks demonstrated low object-drop rate (2.75%) and false-pinch rate (4.40%). The results and feedback from the participants support the robustness and sensitivity of AtaTouch.

CCS CONCEPTS

• **Human-centered computing** → **Gestural input**; **Interaction devices**.

KEYWORDS

Touch segmentation; Hand gesture; Virtual Reality (VR); Antenna; RF Sensing
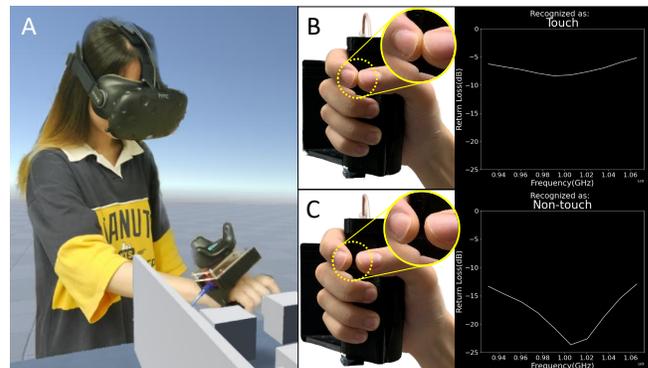
1 INTRODUCTION

Figure 1: AtaTouch precisely detects the touches of two fingertips and supports a robust pinch gesture input on VR interaction (A). AtaTouch utilizes the change in antenna return loss when human fingers form a loop (B, C).

Handheld controllers are an essential part of virtual reality (VR) systems because they enable precise tracking of the hand, immersive interaction through haptic feedback, and auxiliary inputs through buttons, thumbsticks, and triggers. While retaining these advantages of the controllers, modern sensing techniques have enabled the controllers to support natural interaction using gestures (e.g., Facebook's Oculus Touch [19], Valve Index Controller [10], and Tactual Labs [23]). Hand gestures such as a pinch gesture enable users to manipulate objects dexterously, grabbing and moving small virtual blocks, clicking a button, and controlling a slide bar [13]. Further, a simple thumb and finger pinch interaction was reported to be useful in comfortable virtual text entry methods [4, 12].

Many researchers have studied finger tracking and hand gesture detection using cameras. However, vision-based methods of finger tracking have a drawback in a VR environment because a VR controller occludes the finger movements while the controller is an essential part of VR interactions [3]. Human–computer interaction (HCI) studies have actively explored approaches for solving the finger-occlusion problem, ranging from leveraging the change in the back of the hand [15, 16, 22, 29, 48], wrist [11, 14, 20, 32, 50], or arm [21, 39, 40, 44], modifying the handheld controller with capacitive sensors [3], to harnessing bio-acoustic signals [2, 17, 24]. However, their objective was to detect the overall shapes of hand gestures, but often fail to precisely detect whether two fingertips touch each other or not, which is important for the robust detection of a pinch gesture. This sensing limitation causes users to perform

exaggerated trajectories [51]; they open their fingers wide apart and close them strongly.

Therefore, we propose AtaTouch (at a touch), which is a novel, precise, robust, and sensitive technique to detect the closure state of pinching fingers by using the fingers as an impedance-matching component of an antenna. Our approach leverages a change in the return loss of an antenna when human fingers make a loop (Figure 1 B, C). Therefore, AtaTouch enables users to effortlessly perform pinch gestures. We designed a prototype VR controller with AtaTouch inside and verified that it could detect the finger pinching state of the hand that held the controller.

The remainder of this paper is structured as follows. We review related work and describe AtaTouch's sensing principle. We then describe the AtaTouch prototype and quantify its accuracy, robustness, and sensitivity through an accuracy test and user studies with common VR application scenarios. Finally, we discuss the possibilities and limitations of our approach.

2 RELATED WORK

The objective of our work is to detect the pinch gesture leveraging RF signals on the human body. Therefore, we introduce previous sensing techniques for finger tracking and gesture detection. Subsequently, we describe sensing techniques that use the human body and RF signals.

2.1 Sensing Techniques for Finger Tracking and Gesture Detection

Many researchers have studied hand tracking using RGB and depth cameras [5, 34–36]. Oculus Quest supports bare hand tracking using cameras in the headset [13]. However, depending on the hand position and orientation, the fingertips are occluded by the back of the hand or other fingers, and the vision-based approach fails to reliably fingertip touches. In addition, while handheld controllers are essential devices for VR environments, the occlusion of the fingers by the controllers makes vision-based finger-tracking more challenging [3]. The use of depth information is one of the options, but the noise captured by a depth camera makes touch-detection problems difficult [46, 51]. Modifying the camera location is another option [7, 31], but this solution may suffer from problems of finger occlusion, and some light conditions may make finger-segmentation more difficult [51].

HCI studies have addressed the finger-occlusion problem. One approach utilizes the skin deformation of the back of the hand using wearable photo reflective sensors [22], strain gauge sensors [29], capacitive sensors [15, 16], and a camera [48]. Although skin deformation implies movements of the fingers, these sensing techniques have large mean Euclidean finger tracking errors of the index and thumb [22], often resulting in a pinch detection error.

Another approach utilized the internal change of the wrist or arm when a pinch gesture is performed. For example, tomography of electrical impedance [50, 52], and near-infrared diffusion [32] measured internal changes in the tissues of the wrist. Electrical signals, generated from movements of the arm and wrist muscles, were measured and analyzed via electromyography (EMG) to recognize hand gestures [20, 39, 40, 44]. Some studies measured the

changes in skin contour using force-sensitive resistors [11], air-pressure sensors [21], and photo reflectors [14]. However, these sensing approaches may not distinguish the flexion of the fingers from the finger pinch, which is important for the robust detection of a pinch gesture.

Direct modification of the handheld controller with capacitive sensors was explored. Oculus Touch controllers have capacitive sensors on buttons and provide thumbs-up gestures and pointing gestures with the index [19, 42]. Arimatsu et al. [3] implemented capacitance-sensing electrodes that cover the surface of the controller to track fingers grabbing the controller. However, Arimatsu et al. mentioned that their method made the gap between the actual and estimated hand poses and caused users to experience difficulty in performing a pinching action.

Bioacoustic signals generated when users perform pinching and tapping were measured by a piezoelectric microphone [2], a smartwatch accelerometer [2], and multiple-cantilevered piezo films [17]. However, users have to pinch their fingers strongly to create enough bioacoustic signals and users get fatigued easily. FingerPing [49] recognized thumb touches by analyzing acoustic resonance features that travel through the hand from a thumb-worn speaker. However, wearing a ring and a wrist band in addition to holding the VR controller could be cumbersome for VR interactions.

2.2 Sensing Techniques using Human Body and RF signals

Some prior studies utilized the human body as an electrical transmission medium [25, 41, 51, 53]. These studies have two common separate components: a signal emitter and a receiver attached to the wrist and head [51, 53], two people's wrists [41], or a wrist and an object [25]. For example, ActiTouch [51] enabled precise on-skin touch segmentation by emitting and receiving RF waves through the human body when the two arms form a loop. The transmitter and receiver were separated on a wristband and a headset, respectively. They tried integrating both the receiver and transmitter in one wristband, but the signal-to-noise ratio was poor in that configuration. Although they did not explore their method for the loop formation of fingers, their findings imply that their approach to emit and receive an RF signal through the hand requires two separate devices on a finger and an extra body part to detect touches of fingertips. Then, users have to wear additional equipment other than the controller. While these sensing techniques directly measure RF signals propagating through the human body injected from a body-attached device, our technique measures the change in the coupled impedance of the antenna and hand when fingers form a loop. Therefore, our approach is wireless, does not require wearing a gadget, and only requires a handheld controller.

Doppler motion sensing for gesture recognition is also popular, where a receiver captures reflected RF waves with phase and frequency shifts by moving fingers [18, 26, 28, 45]. However, these sensing techniques explore reflected RF waves in a temporal domain rather than a spatial domain, thereby detecting static gestures and spatial configurations such as maintaining pinch gestures is another challenging problem [28]. Our approach is different from them in utilizing the static shape (loop) of fingers rather than motion.
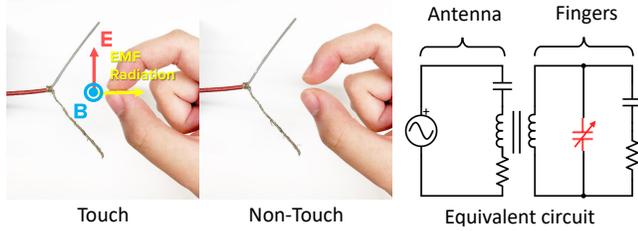
Figure 2: Sensing principle of AtaTouch. An electromagnetic field radiates from the antenna and induces an electric current in the human fingers (left). "E (red)" indicates electric field polarization and "B (blue)" indicates magnetic field polarization. The impedance of the antenna is electrically coupled with that of the fingers (right). The inductance and capacitance are changed when the fingers form a loop (left) or not (middle). The change in impedance causes the return-loss change of the antenna, which is measured by a VNA.

Cohn et al. [8, 9] measured the voltages of the human body, which acted as an antenna and captured electromagnetic noise at home, to detect whole-body gestures. Some studies, the most related to our sensing approach, leveraged the change in antenna return loss to classify hand gestures [1, 30, 47]. The hantenna [30] observed the change in return loss when people touched a coaxial cable's inner conductor using one finger and an outer conductor using the other finger. The cause of this observation was that the fingers were directly connected to the coaxial cable and the beginning section of the fingers behaved as dipole antennas. The hantenna [30] requires direct contact of the cable between two fingertips, thus limiting natural interaction. Other researchers used the perturbation of antenna impedance over time when fingers move near an antenna [1, 47]. However, their methods identified temporal finger-motion patterns, which are less sensitive to hand gestures with small motions, such as effortless pinching. On the contrary, our approach considers the characteristics of the antenna recognizing by the fingers forming a loop, thus can recognized small changes in hand posture. In addition, we explored the integration of an antenna into VR controllers.

3 SENSING PRINCIPLE

A radio-frequency wave is partially transmitted and reflected when the impedance of its transmission medium suddenly changes. The ratio of the reflected power to the incident power is called return loss. When a radio-frequency wave is fed to an antenna, we can obtain the return loss of the antenna due to the sudden impedance change between a signal source and the antenna. A vector network analyzer (VNA) is used to measure the antenna parameters, including return loss. The signal source of a VNA transmits high-frequency waves to an antenna through a device under test (DUT) port of the VNA.

We leverage the impedance change of an antenna when electromagnetically coupled with human fingers. When a high-frequency AC signal flows to an antenna, an electromagnetic field is induced (Figure 2 left). The human body is electrically conductive at radio

frequency. Therefore, the electromagnetic field induces a current in the fingers near the antenna when fingers are aligned to the electric field of the antenna (arrows in Figure 2 left). Owing to the inductance components in the antenna and the fingers (Figure 2 right), their impedance is mutually coupled. As the pinching fingers are closed or opened, the capacitor component in the hand changes. This impedance change by the finger touch state subsequently changes the antenna return loss.

We verified the sensing principle of AtaTouch through a pilot experiment. We fabricated a V-shaped antenna using a coaxial cable (Figure 2 left, middle) as it is one of the fundamental shapes of the antennas. We stripped off the coaxial cable sheath and divided it into a feedline and a ground-side line. From our experience, observing the change in return loss was not possible when the antenna was too short. Therefore, the length of each line was set 6 cm, and the angle between the two wires was set $110°$. Our hypothesis was that return loss would significantly change as fingers are closed or opened when fingers are aligned to the longitudinal direction of the antenna (i.e., when fingers are aligned in the direction of electric field oscillation).

The first author closed the thumb and index finger of the right hand (Figure 2 left) and then opened the two fingers (Figure 2 middle). The tips of the fingers were placed 4 cm from the feed point of the antenna. The hand rested on the desk, and the gap between the fingers when opened was maintained below 5 mm to minimize the effect of hand movements. Return loss data in the frequency range from 800 MHz to 1.5 GHz were collected from a miniVNA tiny [33]. Data collection was performed for four different combinations of hand orientations and locations (Figure 3). The hand orientations were transversal and longitudinal to the stranded wires of the antenna. The hand locations were in front of the antenna or lateral to the antenna. In each combination, we collected the return loss data 20 times for each of the closed (*Touch*) and opened (*Non-touch*) states of the pinching fingers.

The differences between the return loss data of consecutive *Touch* and *Non-touch* in the corresponding frequencies were calculated, and the difference of maximum values was subsequently chosen. In total, 20 differences in return loss were obtained for each of the four combinations. When the hand orientation was in the longitudinal direction of the wires, the average differences were -14.57 and -6.24 dB for the frontal and lateral locations of the hand, respectively. However, when hand orientation was in the transversal direction of the wires, the average differences were -0.55 and -0.59 dB for the frontal and lateral location of the hand, respectively. We concluded that the changes in coupled impedance and return loss are maximized when the pinching fingers align with the poles of the antenna. This pilot experiment result supports our hypothesis.

4 IMPLEMENTATION

In this section, we describe a prototype VR controller with an Ata-Touch inside. Our design goals were 1) to significantly revise return loss values when fingers are closed (*Significant Magnitude*) and 2) to make return loss always increase or decrease when fingers are closed regardless of the location of the fingers (*Consistent Trend*). We considered the finger location to deal with the case of re-grabbing the controller.
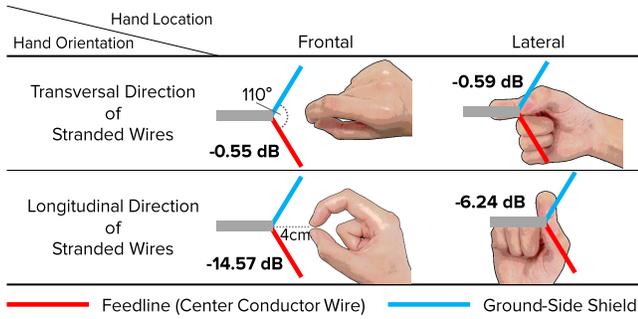
Figure 3: Average return loss difference between when the thumb and index finger are closed or opened. Data were collected for four different hand locations and orientations.
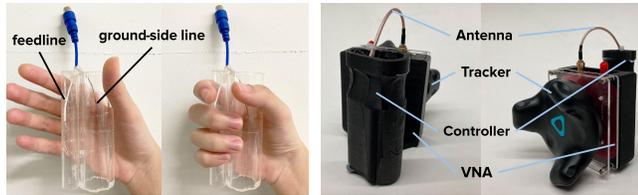

Figure 4: The first prototype of the AtaTouch (left) is shown to help understand the antenna location in the controller. An image on the right shows the final AtaTouch prototype.

4.1 Antenna Placement

As we verified in the Sensing Principle section, the orientation of the finger should be aligned with the electric field polarization to affect the impedance of an antenna. There were two options: one was the frontal space of a V-antenna and the other was the lateral space of a V-antenna, wherein fingers were in a longitudinal direction of stranded wires, as shown in Figure 3 bottom. Using the frontal space of the antenna makes it difficult to embed the antenna in a controller considering the common controller formfactor and grip posture. Therefore, we chose the placement of the fingers to the lateral space of the antenna. In this case, the antenna could be placed inside the controller that users grip (Figure 4 left).

4.2 Design Parameter Optimization through Simulation

To satisfy the aforementioned design goals, *Significant Magnitude* and *Consistent Trend*, we needed to know the influences of antenna length and finger position on return loss. Therefore, we conducted electromagnetic simulations which are commonly used in antenna design [6, 27, 43].

We made a virtual V-shaped antenna using two metal wires (feedline and ground line). The antenna was inverted and placed inside a virtual hand to simulate the situation where a hand grips the controller (similar to Figure 4 left). We used a virtual, one-side-open, metallic cylinder shell as a palm (Figure 5). The dimensions of the cylinder shell was 8.5 cm (width) × 6 cm (height) × 1 cm
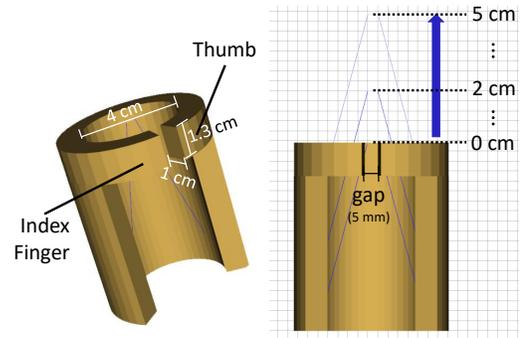

Figure 5: Metallic cylinder shells were used as a palm and fingers (left). The V-shaped antenna shifted from 0 cm to 5 cm along the blue arrow (right).

(thickness). The thumb and index finger were represented by a ring with an inner diameter of 4 cm and a size of 1.3 cm (height) × 1 cm (thickness). The ring was opened with a 5-mm gap to represent non-touched fingertips. The palm was placed below the fingers. We varied the length of the antenna from 2 cm to 8 cm in 2-cm steps. For each antenna length, the hand positions were varied from 0 cm to 5 cm in 1-cm steps. The end of the feedline and ground line is 2 mm apart from the hand, leading to different angles for each antenna length. We used OpenEMS [27] with MATLAB as simulation tool, which uses the finite-difference time-domain (FDTD) method. As a result of the simulation, we obtained a return loss value for each frequency. As the operating frequency of the VNA [33] of our prototype was from 100 MHz to 3 GHz, we used the return loss values in this range.

We analyzed the delta value of return loss from *Touch* to *Non-touch* (Figure 6). For an antenna length of 2 cm, the delta is positive for 0 cm of the finger position and negative for 3 cm of the finger position. This indicates that the trend of return loss change as fingertip touches is not consistent depending on the position of the fingers. For an antenna length of 4 cm, we observed a consistent trend of the delta at approximately 800 MHz, but the magnitudes of the delta are smaller than those of antenna lengths of 6 cm and 8 cm. We found a large delta and a consistent trend of delta at frequencies from 800 MHz to 1.2 GHz for the antenna lengths of 6 cm and 8 cm. Finally, we chose an antenna length of 6 cm because it showed a slightly larger magnitude of delta and will enable a more compact controller design.

In a supplemental simulation, we observed the effect of the antenna angle on return loss. The configuration of the palm and the fingers in the simulation were the same as the previous one. For antenna length of 6 cm, the antenna angles were set to 10, 20, and 30°. For each antenna angle, the finger positions were set to 0, 2, and 4 cm. Figure 7 shows the delta value of return loss from *Touch* to *Non-touch*. The return loss change was the largest for the antenna angle of 30° and the smallest for 10°.

4.3 Hardware Configuration

We 3D-printed the prototype controller in the shape of two attached cylinder shells. A larger cylinder (radius = 1.2 cm) was placed on the
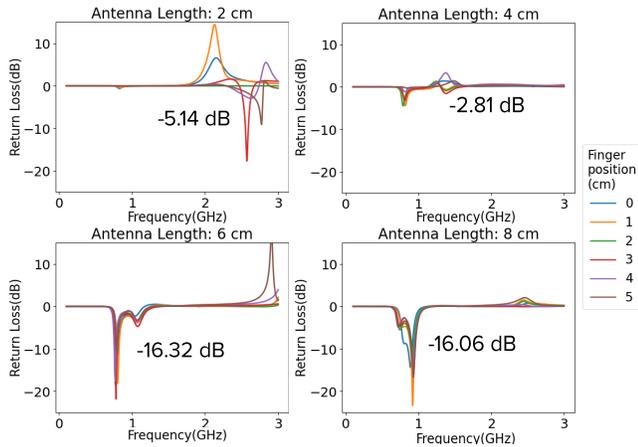
Figure 6: Simulated return-loss differences between *Touch* and *Non-touch* with varying antenna lengths and finger positions. The numbers in the graphs indicate the average of the maximum changes in return loss.
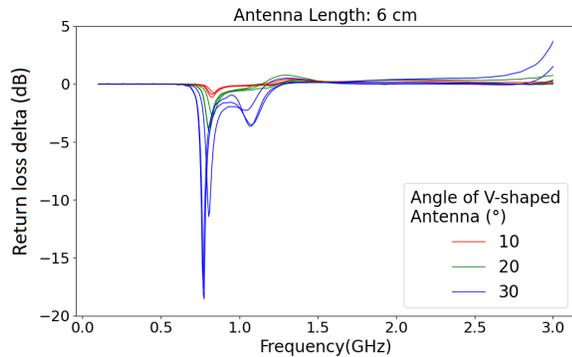

Figure 7: Simulated return-loss differences between *Touch* and *Non-touch* with varying antenna angles and finger positions. The same color indicates the same antenna angle with varying finger positions.

thumb side, and a smaller cylinder (radius = 0.6 cm) was placed on the side of the other fingers. The heights of the cylinder shells were 10 cm, and the thickness was 2 mm. For convenience of the grip, we reduced the radius of the cylinders of the part where the thumb and index finger meet. We fabricated a V-shaped antenna (length = 6 cm, angle = 35°) using a coaxial cable. We placed the antenna in the controller so that the front face of the controller and plane, comprising two wires of the antenna, are parallel. We attached the VNA and HTC VIVE tracker to the back of the controller prototype (Figure 4 right). The SMA (SubMiniature version A) connector of the antenna was connected to the DUT port of the VNA. We observed that the palm position holding the controller also affected the return loss. For robust sensing, we found that wrapping the palm side of the controller with an aluminum foil (i.e., a metal to shield RF) effectively reduced the influence of the palm position. Therefore,

we wrapped the palm side of the controller with 8.5 cm (width) × 10 cm (height) of aluminum foil.

4.4 Classifier

A laptop with a 4-core Intel i7 processor was used to run the software obtaining the return loss data and determining the touch states of pinching fingers. The return loss data were sent to the laptop from the VNA over a USB serial link. We collected 10 data points in the range of 930 MHz to 1.08 GHz, where the change in the return loss was most dynamic. Although there was a clear difference in the return losses between *Touch* and *Non-touch*, we could not use the same threshold criteria for all users because the finger length and hand size affect the return loss values. Therefore, we decided to calibrate a classifier for each person. This calibration process needs to be performed once. To simplify the calibration process, we collected only *Touch* data and used a one-class classification algorithm to classify *Touch* and *Non-touch*, which utilizes the nearest neighbors of a data point. We collected 80 *Touch* data for each user as calibration data. We compared new input with 30 nearest neighbors in the calibration data and obtained distances between the new input and its neighbors. We accepted the input as *Touch* if the mean of the distances was smaller than a threshold; otherwise, we accepted it as *Non-touch*. We designed the threshold optimization method as follows. We obtained the distances between each datum in the calibration data and 30 nearest neighbors, except for itself. A total of 2400 (80×30) distances were gathered, and we calculated the mean of the distances. Subsequently, we set the threshold (Θ) as the addition of the mean of the distances and a margin (α) (i.e., $\Theta = mean + \alpha$). Through a pilot test with six participants, we found that the classifier worked well with a margin of 2 (i.e., $\alpha = 2$).

5 FINGERTIP TOUCH DETECTION ACCURACY TEST

We evaluated how precisely AtaTouch can detect the touch state of pinching fingers. We recruited 12 participants (6 females and 6 males, all right handed) aged between 18 and 26 years from a recruiting board on a university campus. The accuracy test took approximately 20 min.

Obtaining reliable ground-truth systems for detecting fingertip touch states is difficult. We considered an optical tracking system using markers, but human fingertips have complex curved shapes and markers could not represent real fingertip positions. Moreover, we considered attaching a colored sticker to the thumb tip and checking the occlusion of the sticker when the fingertips touch each other. However, the touch positions on the thumb tip were not always in the same position, and an object between fingertips might affect the results of our system. On the contrary, a human observer was used as an accurate touch-event detector in a previous study [51]. Therefore, we decided to use a human observer as a ground-truth system for detecting the touch state in this test.

5.1 Procedure

We evaluated the touch detection accuracy of AtaTouch when participants performed pinch gestures using their thumb and index finger (*Index-pinch*) and thumb and middle finger (*Middle-pinch*).

We collected one training dataset and two test data sets. The training data were collected only for the case where two fingers were closed (*Touch* state). The test data were collected for both cases: for the *Touch* state as well as for the *Non-touch* state where fingers are open. To ensure a uniform gap between fingers in this case, we allowed participants to hold a 5-mm-thick transparent acrylic piece between fingers. The order of the *pinch types (Index pinch vs. Middle pinch)* was randomized. For every five pinch gestures, we asked participants to put down the controller on the desk and then hold it again to allow for grip variations.

The training data, *Non-touch* test data, and *Touch* test data were collected sequentially. The experimenter confirmed the fingertip touches of the participants and then pressed the keyboard button to let the participants know the next pinch type to use. There were only instructions on the screen about what pinch type to use, and any live-prediction result of AtaTouch or return loss data from VNA was not shown to both the experimenter and the participants.

The training data contained 40 data points for each pinch type. The process to gather training data was done in less than 5 min. For the test set, 2400 trials (12 participants × 50 datapoints × 2 fingers × 2 touch-states) were collected. We recorded both live prediction of touch states from AtaTouch and raw data of return loss from VNA.

5.2 Results

We evaluated the accuracy of touch detection using only the live-prediction results from AtaTouch. The overall touch detection accuracy of AtaTouch was 96.4% (SD=7.6). Figure 8 shows the accuracies for the touch states of the ground truth and fingers. In the touch states, AtaTouch predicted *Non-touch* and *Touch* with accuracies of 95.0% (SD=14.6) and 97.8% (SD=6.6), respectively. For *Index-pinch* and *Middle-pinch*, the accuracies of touch detection were 97.7% (SD=6.6) and 95.2% (SD=14.0), respectively.

While AtaTouch showed a decent pinch detection accuracy, the standard deviation of *Non-Touch* state detection accuracy was relatively high. This standard deviation came from one participant with a long middle finger (*Touch* accuracy: 100%, *Non-touch* accuracy: 49%). When there is a gap between the finger and the side of the controller owing to the long finger length, the difference in the return loss between *Touch* and *Non-touch* is smaller (~5dB) than for people with shorter fingers (~20dB). The threshold (Θ), which we chose through the pilot test (please see the Classifier subsection) worked well enough for most participants, but the margin (α) may need to be adjusted depending on the finger length of a user. We performed post-hoc analysis using the recorded data from the test. We reduced the margin (α) to 0.3 (i.e., Θ = *mean* + 0.3), which was originally 2. The pinch detection accuracy of this participant increased to 95% for *Touch* and 93% for *Non-touch*. With this accuracy result, the standard deviation of *Non-touch* accuracy for 12 participants was reduced from 14.6 to 2.0. We expect that this process of adjusting the threshold could be done by a user through a device personalization interface.
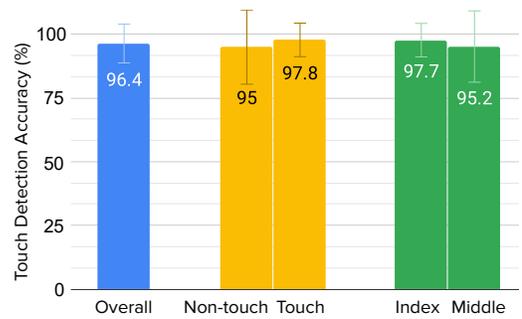

Figure 8: Fingertips touch detection accuracies of AtaTouch on 12 participants. Error bars indicate standard deviation.

6 USER STUDIES ON VR INTERACTION SCENARIOS

We evaluated AtaTouch with VR interaction scenarios wherein people move VR objects using a pinch gesture. This experiment had two sub-experiments; one was "Robustness of Touch Detection", and the other was "Sensitivity and Effortlessness of AtaTouch". We recruited 12 participants (4 females, all right handed) aged between 19 and 27 years from a recruiting board on our university campus. These two sub-experiments took approximately 40 min.

The VR environment was run on Unity with an HTC VIVE Cosmos headset with a tracker attached to the AtaTouch controller (Figure 4 right). The pinch states estimated by AtaTouch were sent to Unity. Subsequently, the participants could attain visual feedback on a hand model and grab an object. The hand model had a translucent green spherical cursor, and the block turned blue when the cursor was hovering over the block. After introducing both sub-experiments, we collected calibration data from each participant as we did in the accuracy test. We asked participants to wear the headset and hold the controller on their right hand.

6.1 Robustness of Touch Detection

The goal of this sub-experiment was to evaluate how AtaTouch robustly and seamlessly maintains touch states for various hand postures in a VR interaction scenario. The task was to move nine virtual blocks on the desk from the left side to the right side over a wall on the center of the desk. The task of moving a block from left to right was to evaluate the object drop rate; the object dropped when AtaTouch recognized *Touch* as *Non-Touch*. The task of moving the hand from right to left after putting down the block was to evaluate false pinch rate; false pinch denotes recognition of *Non-touch* as *Touch*.

Participants could grab the block by performing a pinch gesture. When only participants saw and heard a "grab" instruction, they could grab the block and then move it to the right side of the desk. They were not allowed to put down the block until they see or hear "release" instruction ("grab" trial). The number of object drops was counted in between "grab" and "release" instructions. Subsequently, participants moved their hands from right to left. They were not allowed to grab the block until they see or hear
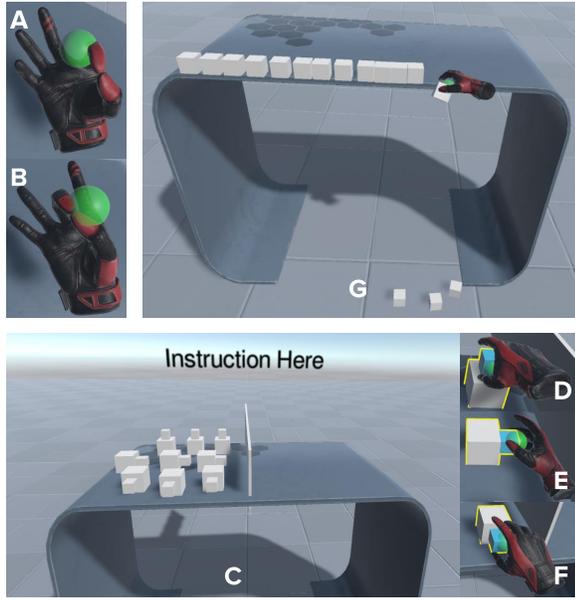
Figure 9: User studies on VR interaction scenario. When participants touch their fingertips, the hand model shows visual feedback (A, B). The first sub-experiment was to move virtual blocks to the left side of the desk (C). The participants moved the blocks in three hand postures (D, E, F). The second sub-experiment was to drop virtual blocks to the bottom as fast as possible (G).

"grab" instruction ("release" trial). The number of false pinches was counted in between "release" and "grab" instructions.

The blocks were placed in 3 rows × 3 columns, and the blocks in each row had a handle in the same position. The position of the handle was to induce different hand postures when participants grab the block. The purpose of different hand postures was to verify that AtaTouch can detect pinch gestures in various hand postures. As shown in Figure 9 (D–F), we requested participants to grab the blocks through three hand postures. Participants repeatedly moved nine blocks for six sets. In the first three sets, participants picked up the block with *Index-pinch*, and in the latter three sets, they used *Middle-pinch*. In total, 1296 trials (12 participants × 9 blocks × 6 sets × 2 ("grab" and "release" trials)) were recorded for the evaluation.

The overall object-drop rate ($\frac{\text{\# of object drop}}{\text{\# of grab trials + \# of object drop}}$) was 2.75% (SD=3.2). The overall false-pinch rate ($\frac{\text{\# of false pinch}}{\text{\# of release trials + \# of false pinch}}$) was 4.4% (SD=5.45). Figure 10 shows the error rates for the *Index-pinch* and *Middle-pinch*. As the tracker was sometimes not properly tracked, the block was stuck when the participants performed pinch gestures and they needed to re-pinch. These trials were counted as object drops. Some participants performed pinch gestures just for fun when the "release" instruction was provided; the false-pinching errors contained these human errors. Despite these measurement errors, AtaTouch showed reasonable accuracy in a VR interaction scenario.
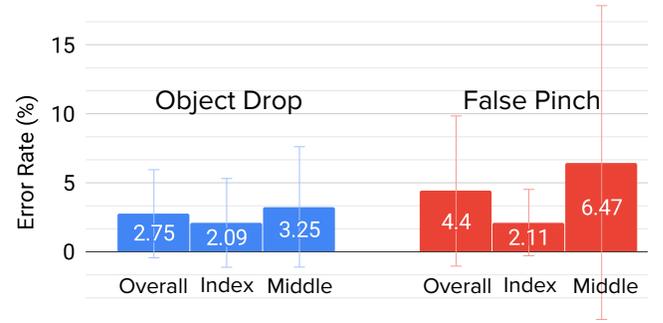

Figure 10: Error rates on object drop and false pinch of Ata-Touch on VR interaction scenario of moving virtual blocks.

6.2 Sensitivity and Effortlessness of AtaTouch

The objective of this sub-experiment was to evaluate how well Ata-Touch supports a light pinching interaction. The task was to grab the 16 blocks on the edge of the desk and quickly and successively drop them to the floor. We designed this consecutive quick pinching task to induce participants to put less effort into each pinch gesture. Participants could grab the block by performing a pinch gesture. Participants repeated the dropping of 16 blocks over six times. The participants used *Index-pinch* for the first three trials and used *Middle-pinch* for the latter three trials. After completing the tasks, participants scored the two statements on a 7-points Likert scale for each pinch type. The statements were as followed: S1) "Compared to the usual experience of dropping objects in the real world, I had to make my fingers more apart to drop the block.", S2) "Compared to the usual experience of picking up objects in the real world, I had to pinch or tap my fingers more strongly to pick up the block.". On the scale, 1-point indicated strong disagreement, and 7-point indicated strong agreement on each statement.

As shown in Figure 11, participants responded that they could drop the virtual block without making their fingers apart largely for both *Index-pinch* (S1 score = 2) and *Middle-pinch* (S1 score = 2.17) cases. In addition, participants reported that they could pick up the virtual block without strong pinching or tapping gestures for both *Index-pinch* (S2 score = 2) and *Middle-pinch* (S2 score = 2.17) cases. Feedback from the participant were "I was amazed that it worked so well than I expected. (p6)", "Sometimes, the middle finger and thumb were a little apart, but the block was still in my hand. Except for those cases, grabbing or dropping objects could be done with efforts similar to or even with less effort than the real experience. (p8)".

7 LIMITATION AND FUTURE WORK

The prototype controller appears cumbersome owing to the size of the VNA. Modern RF technologies, however, have enabled a VNA integrated on a chip [37, 38]. As AtaTouch requires only an antenna and a VNA in a controller, a future AtaTouch prototype could have a much lighter form factor than the current one. In addition, we simulated the antenna lengths and changes in return losses in the frequency range from 100 MHz to 3 GHz, which is the operating range of the VNA we used. It might be possible to
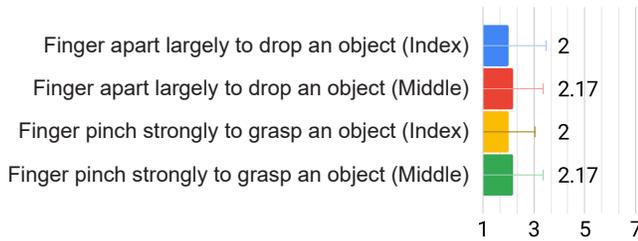
Figure 11: Feedback of the participants on light pinch. (1: Strongly disagree to 7: Strongly agree). The lower the score, the better.

find a higher frequency range that shows significant changes in the return loss for a shorter antenna, and the antenna will enable a smaller controller form factor.

It took 152 ms on average from a user's fingertip touch to touch state classification. Considering that the sub-millisecond time constants of the involved RF physics and the RF measurement circuitry and the low computational cost of the classifier's nearest neighbor search algorithm, we presume that the majority of the processing time must be due to the bottleneck of the serial link from VNA to the computer. As we were using an off-the-shelf VNA for prototyping, we did not have control over the serial link and its protocol. We expect that the processing time would be significantly reduced to allow for measuring more data samples per second if we build a prototype with customized electronics in the future.

The ambient RF signals may not affect the return loss measurement because AtaTouch leverages the electromagnetically coupled impedance of the antenna and human fingers when they are close to each other. Therefore, the fingertip touch segmentation accuracy of our system would be robust to changes in the environment.

AtaTouch may be useful for distinguishing between two fingers touching a conductive surface and one finger touching a conductive surface. Zhang et al. [51] proposed a precise on-skin touch segmentation method and discussed that their method cannot reliably distinguish single-finger touches and multi-finger touches. If one finger touches the opposite palm, no looping is achieved between the finger and thumb. However, when both thumb and a finger touch the palm, a loop is formed between a finger and thumb through the contacted palm. As our sensing principle detects loop formation between thumb and a finger, it may be possible to detect loop formation of a thumb-palm-finger, which would be useful for robust detection of two-finger interaction such as an on-skin zoom in/out gesture.

Finger identification techniques can be combined with the Ata-Touch controller. For instance, we attached capacitive sensors to the AtaTouch controller to obtain the thumb position. From the thumb position, we could identify whether the thumb touches the index or middle finger (please also see Video Figure).

AtaTouch may work with other electronic components, such as buttons and batteries, which are commonly used on controllers. We observed that the aluminum film, which was originally used to reduce the effect of the palm position, effectively shielded the capacitive sensors and electric wires (i.e., metallic components).

Further, we tested the AtaTouch controller with electric wires inside the controller. The return loss changed when the wires were moved, but remained unchanged when the wires were fixed. The effect of electronic devices on AtaTouch could be addressed with RF shielding or fixation of electronic devices.

For more robust sensing, further optimization of the antenna design and classifier would be needed. One possible solution is to use an antenna with high directivity. Directional antennas would achieve better SNRs by reducing the effects of palm noise and sensitivity to other electrical devices. Another possibility is to improve the classifier. We implemented a simple classifier that only uses *Touch* data as training data and uses a static threshold. Achieving higher accuracy through the exploration of another type of classifier such as a two-class classifier would be possible. To avoid personal calibration, classifiers that set the threshold dynamically or based on the first derivatives of the raw return-loss signal are possible.

Although we observed that finger length and hand size affect the return loss values, a deeper investigation is required to determine which and how user variance (e.g., BMI, hand size, skin moisture) affects our sensing principle in a future study.

8 CONCLUSION

While modern sensing techniques enable VR controllers to support natural interaction using hands, the techniques have a limitation for the precise segmentation of two fingertip touches, which is important for robustly supporting a pinch gesture. To address this problem, we proposed AtaTouch, which is a novel, precise, and robust fingertip-touch sensing technique. We evaluated the performance of AtaTouch on the prototype VR controller. Our system could precisely detect fingertip touch with an accuracy of 96.4%. Another user test conducted on VR environments demonstrated that participants could robustly move virtual blocks with a low object-drop rate (2.75%) and a low false-pinch rate (4.40%). Feedback from participants showed non-exaggerated finger movements in the test, and the result supported the sensitivity of AtaTouch.

ACKNOWLEDGMENTS

REFERENCES

[1] Ibrahim Alnujaim, Hashim Alali, Faisal Khan, and Youngwook Kim. 2018. Hand gesture recognition using input impedance variation of two antennas with transfer learning. *IEEE Sensors Journal* 18, 10 (2018), 4129–4135.

[2] Brian Amento, Will Hill, and Loren Terveen. 2002. The sound of one hand: a wrist-mounted bio-acoustic fingertip gesture interface. In *CHI'02 Extended Abstracts on Human Factors in Computing Systems*. 724–725.

[3] Kazuyuki Arimatsu and Hideki Mori. 2020. Evaluation of Machine Learning Techniques for Hand Pose Estimation on Handheld Device with Proximity Sensor. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.

[4] Doug A Bowman, Vinh Q Ly, and Joshua M Campbell. 2001. Pinch keyboard: Natural text input for immersive virtual environments. (2001).

[5] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, and Yaser Sheikh. 2018. OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields. *arXiv preprint arXiv:1812.08008* (2018).

[6] Shibaji Chakraborty, Uddipan Mukherjee, and Khushal Anchalia. 2014. Circular micro-strip (Coax feed) antenna modelling using FDTD method and design using genetic algorithms: A comparative study on different types of design techniques. In *2014 International Conference on Computer and Communication Technology (ICCCT)*. IEEE, 329–334.

[7] 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 & Technology*. 549–556.

[8] Gabe Cohn, Daniel Morris, Shwetak Patel, and Desney Tan. 2012. Humantenna: using the body as an antenna for real-time whole-body interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1901–1910.

[9] Gabe Cohn, Daniel Morris, Shwetak N Patel, and Desney S Tan. 2011. Your noise is my command: sensing gestures using the body as an antenna. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 791–800.

[10] Valve Corporation. 2019. Valve Index Controller. Retrieved August 24, 2020 from https://store.steampowered.com/valveindex.

[11] 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*. 161–166.

[12] Jacqui Fashimpaur, Kenrick Kin, and Matt Longest. 2020. PinchType: Text Entry for Virtual and Augmented Reality Using Comfortable Thumb to Fingertip Pinches. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–7.

[13] Oculus from Facebook. 2020. Oculus Picks: 5 Hand Tracking Experiences on Quest. Retrieved September 14, 2020 from https://www.oculus.com/blog/oculus-picks-5-hand-tracking-experiences-on-quest/.

[14] Rui Fukui, Masahiko Watanabe, Tomoaki Gyota, Masamichi Shimosaka, and Tomomasa Sato. 2011. Hand shape classification with a wrist contour sensor: development of a prototype device. In *Proceedings of the 13th international conference on Ubiquitous computing*. 311–314.

[15] Oliver Glauser, Daniele Panozzo, Otmar Hilliges, and Olga Sorkine-Hornung. 2019. Deformation capture via soft and stretchable sensor arrays. *ACM Transactions on Graphics (TOG)* 38, 2 (2019), 1–16.

[16] Oliver Glauser, Shihao Wu, Daniele Panozzo, Otmar Hilliges, and Olga Sorkine-Hornung. 2019. Interactive hand pose estimation using a stretch-sensing soft glove. *ACM Transactions on Graphics (TOG)* 38, 4 (2019), 1–15.

[17] 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*. 453–462.

[18] Souvik Hazra and Avik Santra. 2018. Robust gesture recognition using millimetric-wave radar system. *IEEE sensors letters* 2, 4 (2018), 1–4.

[19] Facebook Inc. 2016. Oculus Touch. Retrieved August 24, 2020 from https://www.oculus.com/rift/accessories/.

[20] Peter Ju, Leslie Pack Kaelbling, and Yoram Singer. 2000. State-based classification of finger gestures from electromyographic signals. In *ICML*, Vol. 8. Citeseer, 439–446.

[21] Pyeong-Gook Jung, Gukchan Lim, Seonghyok Kim, and Kyoungchul Kong. 2015. A wearable gesture recognition device for detecting muscular activities based on air-pressure sensors. *IEEE Transactions on Industrial Informatics* 11, 2 (2015), 485–494.

[22] Wakaba Kuno, Yuta Sugiura, Nao Asano, Wataru Kawai, and Maki Sugimoto. 2017. 3D Reconstruction of Hand Postures by Measuring Skin Deformation on Back Hand.. In *ICAT-EGVE*. 221–228.

[23] Tactual Labs. 2019. Controller. Retrieved August 24, 2020 from https://www.tactuallabs.com/.

[24] 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*. 321–333.

[25] Gierad Laput, Chouchang Yang, Robert Xiao, Alanson Sample, and Chris Harrison. 2015. Em-sense: Touch recognition of uninstrumented, electrical and electromechanical objects. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. 157–166.

[26] Ziheng Li, Zhenyuan Lei, An Yan, Erin Solovey, and Kaveh Pahlavan. 2020. Thu-Mouse: A micro-gesture cursor input through mmWave radar-based interaction. In *2020 IEEE International Conference on Consumer Electronics (ICCE)*. IEEE, 1–9.

[27] Thorsten Liebig. [n.d.]. openEMS - Open Electromagnetic Field Solver. https://www.openEMS.de

[28] 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 (TOG)* 35, 4 (2016), 1–19.

[29] 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*. 557–564.

[30] Sergio Rueda Linares, David Martin Garcia, Sen Yan, Vladimir Volski, and Guy AE Vandenbosch. 2017. The hantenna: experimental assessment of the human hand as an antenna. *IET Microwaves, Antennas & Propagation* 12, 5 (2017), 773–778.

[31] Christian Loclair, Sean Gustafson, and Patrick Baudisch. 2010. PinchWatch: a wearable device for one-handed microinteractions. In *Proc. MobileHCI*, Vol. 10. Citeseer.

[32] 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*. 593–597.

[33] mini Radio Solutions. 2019. miniVNA Tiny. Retrieved September 7, 2020 from http://miniradiosolutions.com/54-2/.

[34] Franziska Mueller, Florian Bernard, Oleksandr Sotnychenko, Dushyant Mehta, Srinath Sridhar, Dan Casas, and Christian Theobalt. 2018. Ganerated hands for real-time 3d hand tracking from monocular rgb. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 49–59.

[35] Franziska Mueller, Micah Davis, Florian Bernard, Oleksandr Sotnychenko, Mickeal Verschoor, Miguel A. Otaduy, Dan Casas, and Christian Theobalt. 2019. Real-time Pose and Shape Reconstruction of Two Interacting Hands With a Single Depth Camera. *ACM Transactions on Graphics (TOG)* 38, 4 (2019).

[36] Franziska Mueller, Dushyant Mehta, Oleksandr Sotnychenko, Srinath Sridhar, Dan Casas, and Christian Theobalt. 2017. Real-time Hand Tracking under Occlusion from an Egocentric RGB-D Sensor. In *Proceedings of International Conference on Computer Vision (ICCV)*. 10. http://handtracker.mpi-inf.mpg.de/projects/OccludedHands/

[37] Johannes Nehring. 2017. Highly Integrated Microwave Vector Network Analysis Circuits and Systems for Instrumentation and Sensing Applications. (2017).

[38] Johannes Nehring, Martin Schütz, Marco Dietz, Ismail Nasr, Klaus Aufinger, Robert Weigel, and Dietmar Kissinger. 2016. Highly integrated 4–32-GHz two-port vector network analyzers for instrumentation and biomedical applications. *IEEE Transactions on Microwave Theory and Techniques* 65, 1 (2016), 229–244.

[39] T Scott Saponas, Desney S Tan, Dan Morris, and Ravin Balakrishnan. 2008. Demonstrating the feasibility of using forearm electromyography for muscle-computer interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 515–524.

[40] 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*. 167–176.

[41] Kenji Suzuki, Taku Hachisu, and Kazuki Iida. 2016. Enhancedtouch: A smart bracelet for enhancing human-human physical touch. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 1282–1293.

[42] Facebook Technologies. 2020. Oculus Controller Data. Retrieved September 14, 2020 from https://developer.oculus.com/documentation/native/pc/dg-input-touch/.

[43] Y Wang, AJ Pretorius, and AM Abbosh. 2016. Low-profile antenna with elevated toroid-shaped radiation for on-road reader of RFID-enabled vehicle registration plate. *IEEE Transactions on Antennas and Propagation* 64, 4 (2016), 1520–1525.

[44] Kevin R Wheeler, Mindy H Chang, and Kevin H Knuth. 2006. Gesture-based control and EMG decomposition. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 36, 4 (2006), 503–514.

[45] Te-Yen Wu, Shutong Qi, Junchi Chen, MuJie Shang, Jun Gong, Teddy Seyed, and Xing-Dong Yang. 2020. Fabriccio: Touchless Gestural Input on Interactive Fabrics. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.

[46] Robert Xiao, Scott Hudson, and Chris Harrison. 2016. DIRECT: Making Touch Tracking on Ordinary Surfaces Practical with Hybrid Depth-Infrared Sensing. In *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces*. 85–94.

[47] Bin Xu, Yang Li, and Youngwook Kim. 2017. Classification of finger movements based on reflection coefficient variations of a body-worn electrically small antenna. *IEEE Antennas and Wireless Propagation Letters* 16 (2017), 1812–1815.

[48] Hui-Shyong Yeo, Erwin Wu, Juyoung Lee, Aaron Quigley, and Hideki Koike. 2019. Opisthenar: Hand Poses and Finger Tapping Recognition by Observing Back of Hand Using Embedded Wrist Camera. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*. 963–971.

[49] Cheng Zhang, Qiuyue Xue, Anandghan Waghmare, Ruichen Meng, Sumeet Jain, Yizeng Han, Xinyu Li, Kenneth Cunefare, Thomas Ploetz, Thad Starner, et al. 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*. 1–10.

[50] 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 & Technology*. 167–173.

[51] Yang Zhang, Wolf Kienzle, Yanjun Ma, Shiu S Ng, Hrvoje Benko, and Chris Harrison. 2019. ActiTouch: Robust Touch Detection for On-Skin AR/VR Interfaces. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and*

Technology. 1151–1159.

[52] 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.* 843–850.

[53] 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 2016 CHI Conference on Human Factors in Computing Systems.* 1491–1503.