Hover Detection using Active Acoustic Sensing

Masaya Tsuruta, Shuhei Aoyama, Arika Yoshida,
Buntarou Shizuki, and Jiro Tanaka

University of Tsukuba, Japan
{tsuruta,aoyama,yoshida,shizuki,jiro}@iplab.cs.tsukuba.ac.jp

Abstract. In this paper, we present a technique for hover and touch detection using Active Acoustic Sensing. This sensing technique analyzes the resonant property of the target object and the air around it. To verify whether the detection technique works, we conduct an experiment to discriminate between hovering the hand over the piezoelectric elements placed on a target object and touching the same object. As a result of our experiment, hovering was detected with 96.7% accuracy and touching was detected with 100% accuracy.

1 Introduction

Hover is the gesture that is placing a hand or a finger closely above an object. It can be detected easily by simply adding proximity sensors to the object. This method is useful for improving usability since it can add another interaction modality (i.e., hover) to the object. For example, a hover-sensitive device can automatically wake up from standby mode for an instant interaction.

In this paper, rather than adding proximity sensors to an object, we show a lightweight technique that can detect hover over an object by attaching a pair of piezoelectric elements to the object: a vibration speaker and a contact microphone. This technique is based on our Active Acoustic Sensing [1], which is a technique to make existing objects touch-sensitive using ultrasonic waves and frequency analysis. The Active Acoustic Sensing utilizes the fact that the resonant property of a solid object is sensitive to how they are touched. In other words, the resonant property of an object changes with respect to the manner in which the object is touched. Since these changes can be observed as different resonant frequency spectra, we can estimate how the object is touched by analyzing the spectra.

The contribution of this paper, in addition to determining how an object is touched, is identifying that hover can also be observed as different resonant frequency spectra. Thereby, this technique will further enrich the vocabulary of interaction for prototyping objects.

2 Related Work

2.1 Passive sensing

Some studies detect tactile gestures by passively capturing the sound or vibrations created by these gestures using microphones [2–4]. For example, Toffee [3] detects around-device taps by capturing mechanical vibrations generated by a finger tap; it uses four microphones attached to each corner at the bottom of a device and uses time differences between the arrivals of acoustic vibrations to estimate the tapped position. Braun et al. [4] sense interaction on everyday surfaces with some microphones attached to these surfaces. Their system detects taps and swipe gestures using machine learning. In contrast to these studies which need a user-activated interaction that emits sounds, our study does not require users to emit any sound as our device emits the sounds actively.

2.2 Active sensing

In addition to the aforementioned techniques, there are others that can detect user interaction using active acoustic signals [5–8]. For example, Acoustruments [5] is a sensing technique used for tangible interaction on a smartphone. This technique connects the speaker and the microphone of the smartphone by a plastic tube. The sensing system detects changes of acoustic waves in or around the plastic tube to recognize a variety of interaction. While this study uses changes of acoustic waves in or around the tube with a smartphone, our study focuses on using changes of acoustic waves diffused in the air to detect hover over arbitrary solid objects. SoundWave [6] measures Doppler shifts with the speakers and microphones already embedded in commodity devices to detect in-air gestures. In contrast, our study is a lightweight technique, which uses a pair of piezoelectric elements attached to an existing object to detect hover over the object, in addition to touch gestures that can be detected by [1]. EchoTag [7] enables a smartphone to tag and remember indoor locations by transmitting an acoustic signal with the smartphone's speaker and sensing its environmental reflection with the smartphone's microphone. This study makes use of reflection of acoustic signals around the device. On the other hand, our study uses acoustic signals to sense hover. Wang et al. [9] use a swept frequency audio signal to detect tapped positions on a paper keyboard placed on a desk. Our study can detect hover in addition to how the object is touched.

2.3 Hover detection

Various methods to detect hover have been researched. Wilson et al. [10] use a depth camera. Withana et al. [11] use IR sensors. Rekimoto [12] uses capacitive sensing. By contrast, our study uses Active Acoustic Sensing [1] to detect hover over an existing object.

3 Detection Mechanism

We will now present our mechanism to detect hover and touch using ultrasonic waves and their frequency analysis.

3.1 Principles of touch and hover detection

The principle of our Active Acoustic Sensing [1], which is a technique to detect how an existing solid object is touched using its acoustic property and also serves as a basis of our hover detection, is simple. Every object has a resonant property. This causes a vibrational reaction, which is unique to that object. When the object is touched, its resonant property changes with respect to the nature of the touch. As a result, the vibrational reaction also changes. Our technique uses this phenomenon to estimate how the object is touched (i.e., touch gestures) in accordance with the following procedure. Initially, an actuator attached to the object causes the object to vibrate at a wide range of frequency. Next, the frequency response from the object is acquired using a sensor attached to the object. Finally, our technique uses machine learning with frequency response labeled as touch gestures.

In this work, to realize hover detection, we focus on the waves leaked into the air. If there is nothing in the air around the sensor, the waves observed by the sensor mainly consist of the waves emitted from the actuator and propagated through the body of the object where the sensor is attached. In contrast, if a hand or a finger is above the sensor (i.e., hovering), the waves emitted from the actuator into the air are reflected by the hand or finger and propagated to the sensor (Fig. 1). Consequently, the waves observed by the sensor change (i.e., now contain such waves) before hovering. This change can also be detected as hover using machine learning with frequency response. As described above, our technique is a lightweight method to detect hover over an object, in addition to the manner in which the object is touched, by attaching a pair of actuator and sensor to the object.
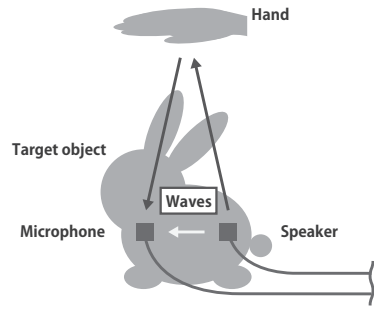

Fig. 1. Principle of hover detection.

3.2 System

The overview of our system is shown in Fig. 2. This system consists of the following: software to generate waves and analyze the frequency response, a pair of actuator and sensor that are attached to the object, and an amplifier.

The actuator and sensor used in this system are our own piezoelectric vibration speaker and microphone respectively. The two have the same composition, which consists of a piezoelectric film and an acrylic plate (Fig. 3). This acrylic plate is pasted to the piezoelectric film by an adhesive to enhance durability of elements and stability of the wave propagation. The piezoelectric vibration speaker and microphone should be attached near the location where we want to detect hover. In our system, we used a double-sided tape to attach these elements to the surface. Note that these elements should be exposed to air to make the ultrasonic waves propagate adequately into the air.
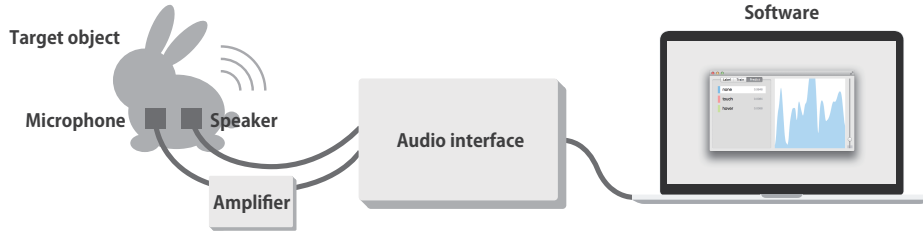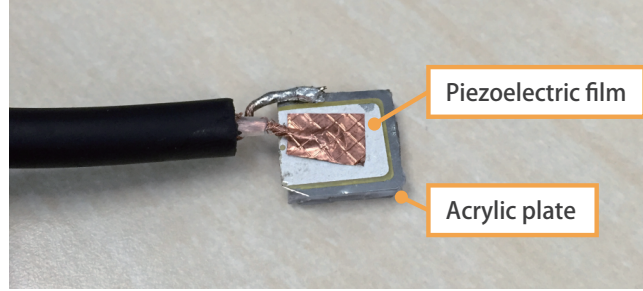

Fig. 2. System overview.

Fig. 3. Microphone/speaker.

The software generates a sinusoidal sweep signal from 20 kHz to 40 kHz as the wave emitted from the speaker. To detect touch and hover, the software uses Fast Fourier Transform (FFT) to obtain the spectrum of the wave acquired by the

microphone and a Support Vector Machine (SVM) as the classifier similar to that used in [1]. The user interface of the software is shown in Fig. 4. On the left side of the window is an area for labels representing the different types of interaction. In the "Label" tab, the user initializes the different types of interaction he wants to detect. In the "Train" tab, he iterates the actions to train and build an SVM model. In the "Predict" tab, each label shows its likelihood correspondence with an action. The right side of the window is an area for visualized sound spectrum of the frequency response, which is the result of FFT. It is used to understand how our system works and to confirm whether the system is runs normally. When user interactions such as hover or touch induce changes, the shape of the spectrum change as well, and the corresponding label in the "Predict" tab is highlighted.
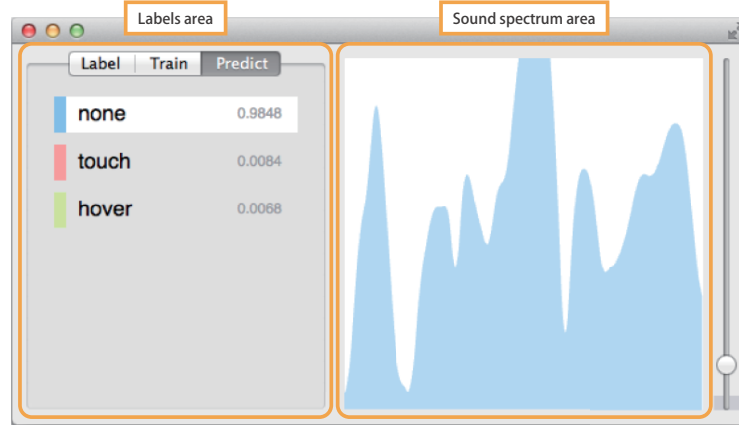

Fig. 4. Software.

4 Experiment

To verify whether the detection technique works, we conducted an experiment to discriminate hover and touch. For this purpose, we attached a piezoelectric vibration speaker/microphone pair to an acrylic table as shown in Fig. 5. We also used a doorknob and a portable safe as target objects. We call the three scenes of experiment "Table", "Doorknob", and "Safe" according to the target objects. The distance between the two piezoelectric elements was 20 mm.

A participant covered the piezoelectric elements with his hand over 5 mm above them (Hover). He touched the point that was 11.18 mm away from the elements (Touch). To measure the accuracy of our technique, the participant first touched the target object 10 times and then performed hover 10 times. We counted the number of times that the gesture were detected correctly under both conditions.

Following are the results. For "Table", Hover was correctly detected 9 times implying an accuracy of 90% (The failed trial was detected as Touch); Touch was correctly detected 10 times implying an accuracy of 100%. For "Doorknob", Hover was correctly detected 10 times implying an accuracy of 100%; Touch was correctly detected 10 times implying an accuracy of 100%. For "Safe", Hover was correctly detected 10 times implying an accuracy of 100%; Touch was correctly detected 10 times implying an accuracy of 100%. In total, Hover was detected with 96.7% accuracy; Touch was detected with 100% accuracy. The observed sound spectra supported these results. Fig. 6–8 show examples of sound spectra observed in the following three conditions: Default, Touch, and Hover. These sound spectra are different from each other as shown in Fig. 9 for example. Thus, our system could detect hover and touch accurately.
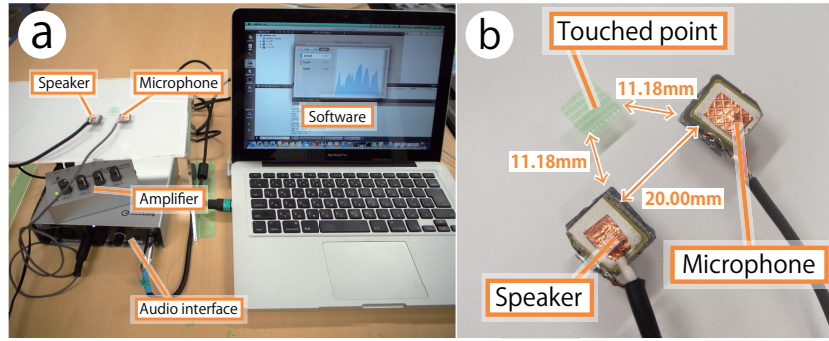

Fig. 5. Setup of the experiment: a) Overview, b) the attached piezoelectric vibration speaker and piezoelectric microphone.
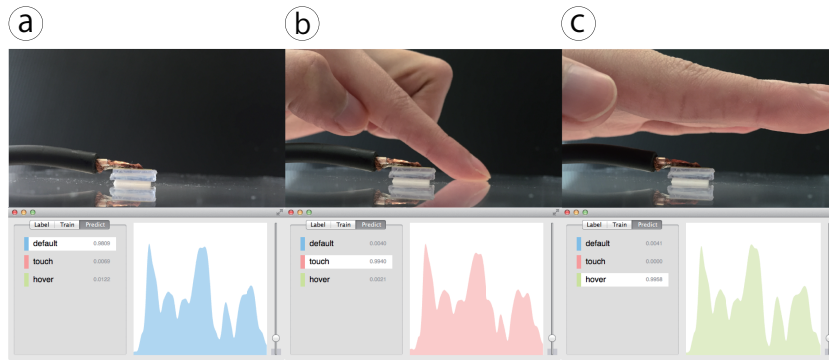

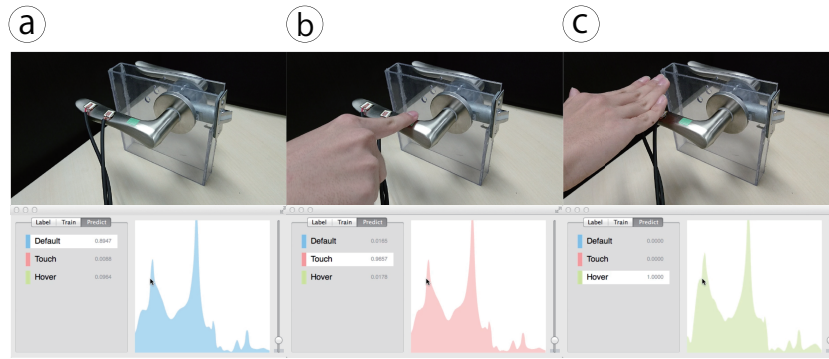**Fig. 6.** Observed sound spectra of "Table": a) Default, b) Touch, and c) Hover.

Fig. 7. Observed sound spectra of "Doorknob": a) Default, b) Touch, and c) Hover.
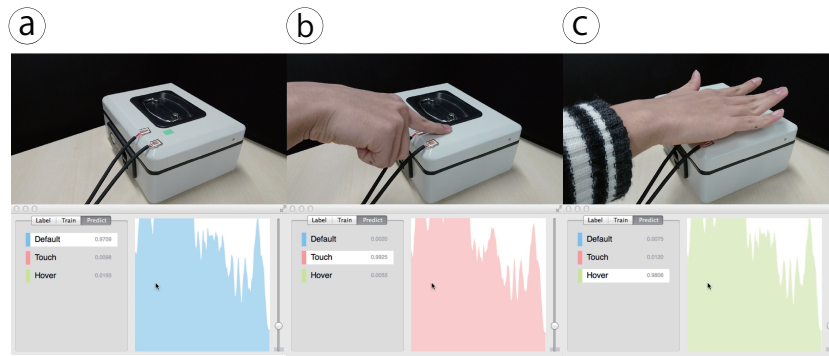

Fig. 8. Observed sound spectra of "Safe": a) Default, b) Touch, and c) Hover.
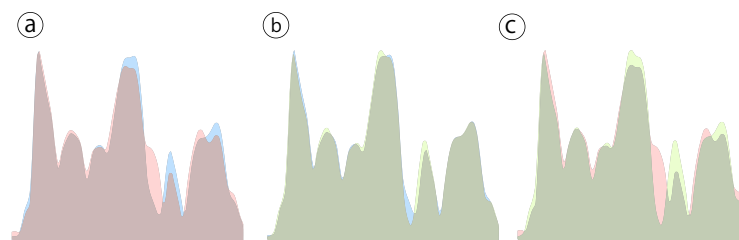

Fig. 9. Differences of sound spectra of "Table" between the three conditions: a) Default – Touch, b) Default – Hover, and c) Touch – Hover.

5 Applications

We present three applications as examples of our hover detection technique.

5.1 Shortcuts on computers

As shown in Fig. 10a, we attached the piezoelectric elements to the lower right corner of a computer. Using this setup, we were able to recognize four touch gestures and a hover gesture on the computer (Fig. 11). We assigned a shortcut to each of them. For example, we assigned the finger top gesture to Ctrl+z, which is a widely used shortcut for undoing, and the hover gesture to Ctrl+s, which is a frequently used shortcut for saving.

5.2 Around-device interaction on smartphones

By using our technique, hover can be made available even outside of the screen area without using dedicated smartphones. As shown in Fig. 10b, we pasted a piezoelectric speaker and a microphone at the back of a smartphone. In this case, back-of-device interaction [13] was realized by assigning the undo shortcut to it. As shown in this application, our technique is useful for prototyping around-device interaction including back-of-device interaction and will possibly promote researchers to conduct such research as [14, 15].

5.3 Prototyping objects with hover detection

As shown in Fig. 12, our technique can be applied to various everyday objects. This shows that our technique is useful for prototyping objects with hover detection. Fig. 12a and 12b are prototypes of crockery sensitive to grasp/hover. Our technique can also detect whether or not the crockery is filled with water. When there are contents in it, hover triggers a voice indication such as "The vessel has content." Fig. 12c is a touch/stack/hover sensitive toy car model made out of blocks. When hover is detected, it emits the sound of an engine, which the user has applied in advance. Fig. 12d and 12e are proximity sensitive objects with locking mechanisms. When these devices are approached by the hands of an unwelcome person, they automatically lock themselves even if they are carelessly left unlocked. For example, in the application described in Fig. 12d, if a user performs hover over the doorknob three seconds before grasping it, the system detects him as an owner. In contrast, if he grasps it immediately, the system detects him as an unwelcome person. As this application shows, our technique is also useful to detect whether the person is the owner or not. Fig. 12f is an example of adding function to a ready-made product. While conducting an iterative development of a prototype, it is possible to add another function quickly even though it does not exist as a button.
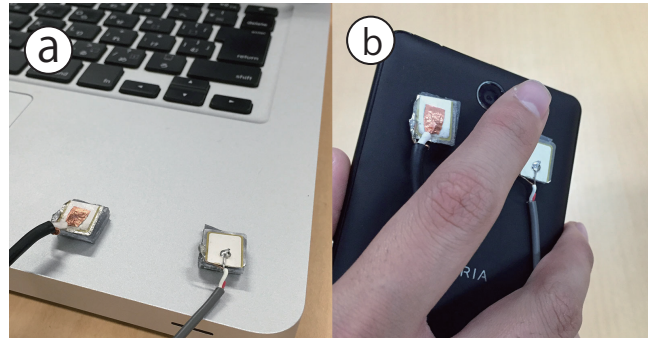
Fig. 10. Applications: a) Shortcuts on computers and b) back-of-device interaction on smartphones.
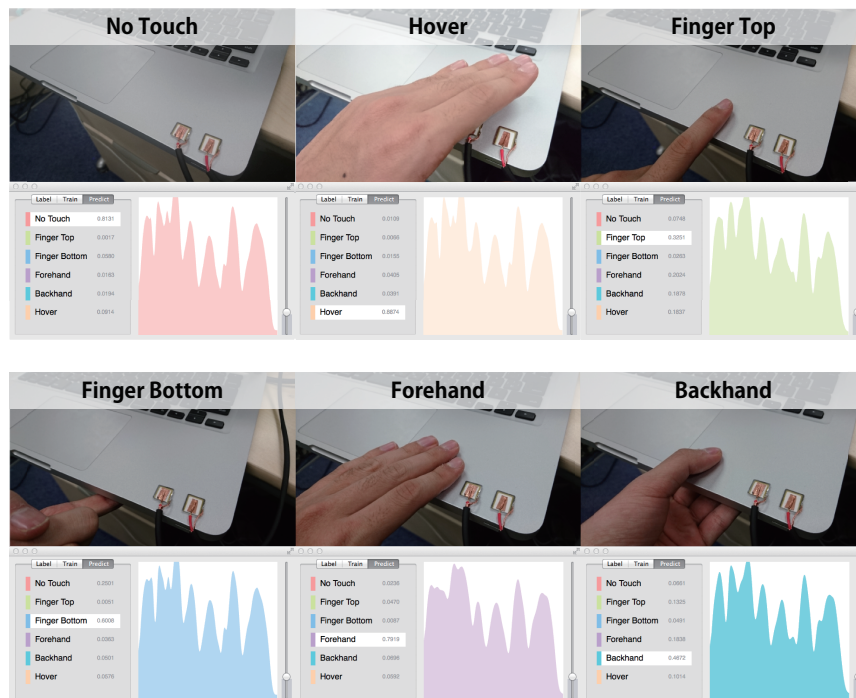

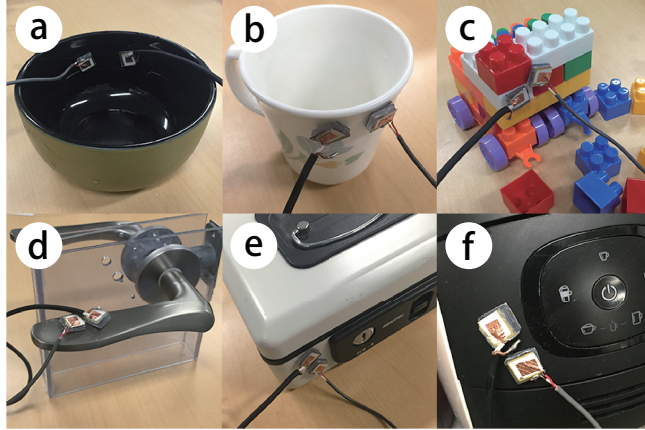**Fig. 11.** Four touch gestures and hover on the computer.

Fig. 12. Prototyping objects with hover detection.

6 Discussion

Our experiments show that, our technique can detect hover near the piezoelectric speaker and microphone. Moreover, we tested the detectable height of hover; the height was 13.8 mm in our current implementation. This implies that it is only possible to detect whether a hand or a finger hovers or not. However, this also implies that a high-output speaker with an amplifier and a more sensitive microphone will increase the detectable height, thus making it possible for more complex gestures to be detected.

Further experiments can also be conducted. In this paper, we discriminated between a single type of hover and touch gesture. While this experiment shows a possibility of application of our sensing technique, we should investigate the discrimination accuracy of various types of hover and touch gestures in our next experiment.

7 Conclusions and Future Work

In this paper, we presented a technique to detect hover by our Active Acoustic Sensing and acquiring ultrasonic waves leaked into and reflected from the air. As a result of our experiment, hovering was detected with 96.7% accuracy and touching was detected with 100% accuracy. However, the detectable height limits the detection to a hover gesture made by a hand or a finger and at present more complex gestures are not detectable. We believe dedicated implementation may accomplish detection of more complex gestures. For example, we already confirmed that when we use an ultrasonic speaker designed for emitting waves into the air, the detection height is extended to around 15 cm, and multiple steps of hover can be detected. Moreover, this also can be used with touch gestures. We

are now dedicated to implement a new method with Support Vector Regression (SVR), which can detect hover gestures as continuous values and we plan to complete it as a part of our immediate future work.

References

1. Ono, M., Shizuki, B., Tanaka, J.: 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, New York, NY, USA, ACM (2013) 31–40
2. Murray-Smith, R., Williamson, J., Hughes, S., Quaade, T.: Stane: Synthesized surfaces for tactile input. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '08, New York, NY, USA, ACM (2008) 1299–1302
3. Xiao, R., Lew, G., Marsanico, J., Hariharan, D., Hudson, S., Harrison, C.: Toffee: Enabling ad hoc, around-device interaction with acoustic time-of-arrival correlation. In: Proceedings of the 16th International Conference on Human-computer Interaction with Mobile Devices & Services. MobileHCI '14, New York, NY, USA, ACM (2014) 67–76
4. Braun, A., Krepp, S., Kuijper, A.: Acoustic tracking of hand activities on surfaces. In: Proceedings of the 2nd International Workshop on Sensor-based Activity Recognition and Interaction. WOAR '15, New York, NY, USA, ACM (2015) 9:1–9:5
5. Laput, G., Brockmeyer, E., Hudson, S.E., Harrison, C.: 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, New York, NY, USA, ACM (2015) 2161–2170
6. Gupta, S., Morris, D., Patel, S., Tan, D.: SoundWave: Using the doppler effect to sense gestures. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '12, New York, NY, USA, ACM (2012) 1911–1914
7. Tung, Y.C., Shin, K.G.: EchoTag: Accurate infrastructure-free indoor location tagging with smartphones. In: Proceedings of the 21st Annual International Conference on Mobile Computing and Networking. MobiCom '15, New York, NY, USA, ACM (2015) 525–536
8. Harrison, C., Schwarz, J., Hudson, S.E.: TapSense: Enhancing finger interaction on touch surfaces. In: Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology. UIST '11, New York, NY, USA, ACM (2011) 627–636
9. Wang, J., Zhao, K., Zhang, X., Peng, C.: Ubiquitous keyboard for small mobile devices: Harnessing multipath fading for fine-grained keystroke localization. In: Proceedings of the 12th Annual International Conference on Mobile Systems, Applications, and Services. MobiSys '14, New York, NY, USA, ACM (2014) 14–27
10. Wilson, A.D.: Using a depth camera as a touch sensor. In: ACM International Conference on Interactive Tabletops and Surfaces. ITS '10, New York, NY, USA, ACM (2010) 69–72
11. Withana, A., Peiris, R., Samarasekara, N., Nanayakkara, S.: zSense: Enabling shallow depth gesture recognition for greater input expressivity on smart wearables. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. CHI '15, New York, NY, USA, ACM (2015) 3661–3670

12. Rekimoto, J.: SmartSkin: An infrastructure for freehand manipulation on inter-
active surfaces. In: Proceedings of the SIGCHI Conference on Human Factors in
Computing Systems. CHI '02, New York, NY, USA, ACM (2002) 113–120
13. Fukatsu, Y., Hiroyuki, H., Shizuki, B., Tanaka, J.: Back-of-device interaction using
halls on mobile devices. In: Proceedings of Interaction 2015, Information Processing
Society of Japan (2015) 412–415 (In Japanese).
14. Kratz, S., Rohs, M.: Hoverflow: Expanding the design space of around-device inter-
action. In: Proceedings of the 11th International Conference on Human-Computer
Interaction with Mobile Devices and Services. MobileHCI '09, New York, NY, USA,
ACM (2009) 4:1–4:8
15. Hakoda, H., Kuribara, T., Shima, K., Shizuki, B., Tanaka, J.: AirFlip: A double
crossing in-air gesture using boundary surfaces of hover zone for mobile devices. In
Kurosu, M., ed.: Human-Computer Interaction: Interaction Technologies. Volume
9170 of Lecture Notes in Computer Science. Springer International Publishing
(2015) 44–53