Poster Abstract: Toward Continuous Finger Positioning on Ear Using Bone Conduction Speaker

Yuke Zhang
zhangyuke@akg.t.u-tokyo.ac.jp
The University of Tokyo, Japan

Ken Takaki
takaki@akg.t.u-tokyo.ac.jp
The University of Tokyo, Japan

Hiroaki Murakami
murakami@akg.t.u-tokyo.ac.jp
The University of Tokyo, Japan

Takuya Sasatani
sasatani@akg.t.u-tokyo.ac.jp
The University of Tokyo, Japan

Yoshihiro Kawahara
kawahara@akg.t.u-tokyo.ac.jp
The University of Tokyo, Japan

ABSTRACT

The advancement of semiconductor and battery technologies popularized tiny acoustic wearable devices such as bone conduction wireless headsets. However, this small form factor poses inconvenience when controlling these devices, as they cannot equip large footprint intuitive interfaces such as volume sliders and touch screens. This paper presents a technique using acoustic responses measured by a bone conduction speaker and a microphone to utilize the ear as a touch input interface. We discovered that a finger placed on different parts of the ear affects the acoustic radiation characteristic of the ear, modulating the leaked sound, and by leveraging this effect, the touch position can be estimated. Experimental results show that five distinct frequency responses with five different finger positions can be obtained, which indicates that our method could allow bone conduction headsets to capture continuous finger positions without additional hardware.

CCS CONCEPTS

• **Human-centered computing → Interaction devices**; *Ubiquitous and mobile computing*.

KEYWORDS

acoustic sensing, bone conduction, positioning, human interface

ACM Reference Format:
Yuke Zhang, Ken Takaki, Hiroaki Murakami, Takuya Sasatani, and Yoshihiro Kawahara. 2022. Poster Abstract: Toward Continuous Finger Positioning on Ear Using Bone Conduction Speaker. In *The 20th ACM Conference on Embedded Networked Sensor Systems (SenSys '22), November 6–9, 2022, Boston, MA, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3560905.3568075

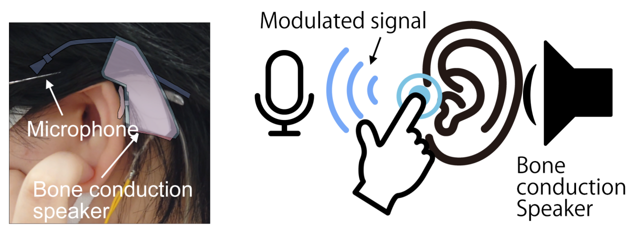

Figure 1: The setup for the finger positioning device using a bone conduction speaker and a microphone. These components are fixed on the ear with a 3D printed case.

1 INTRODUCTION

Wearable audio devices have become essential in various scenes of our lives and work, such as making phone calls and listening to music. The development of small speakers/microphones and the fundamental requirement for comfort and portability made these devices smaller and lightweight. However, miniaturization also results in a limited area for input interfaces, making it challenging for users to interact with these devices intuitively (*e.g.,* slider that turns the volume up/down).

Prior studies attempted to offer larger input areas by turning the ear into an input controller using sensing hardware to detect input events such as touch and deformation on the ear. EarTouch [1] added an optical proximity sensor on the headset to detect the shape deformation of the ear. However, it is hard to deploy on headsets because it needs additional sensor hardware, and the requirement to deform the ear limits the available input gestures. EarBuddy [3] detected eight tapping and sliding gestures by recording and classifying the sound generated by these gestures using machine learning. This approach allows users to interact with wearables using a variety of commands, yet it is hard to accurately sense analog inputs such as distance of the sliding, which can help manipulate sound-related features (*e.g.,* volume, playback speed), involving many analog parameters.

This paper presents an approach using a speaker/microphone pair to perform active acoustic sensing in the inaudible frequency range to achieve continuous finger positioning on the ear. Our setup uses (i) a bone conduction speaker behind the ear that presents sound and slightly vibrates the ear, resulting in a sound leakage which is modulated by the finger placement, and (ii) a microphone in front of the ear that records the leaked modulated sound. The remaining of this paper describes the experimental setup, evaluates the acoustic response changes according to finger position to demonstrate our approach's effectiveness, and finally discusses the results.

SenSys '22, November 6–9, 2022, Boston, MA, USA

ACM ISBN 978-1-4503-9886-2/22/11.
https://doi.org/10.1145/3560905.3568075

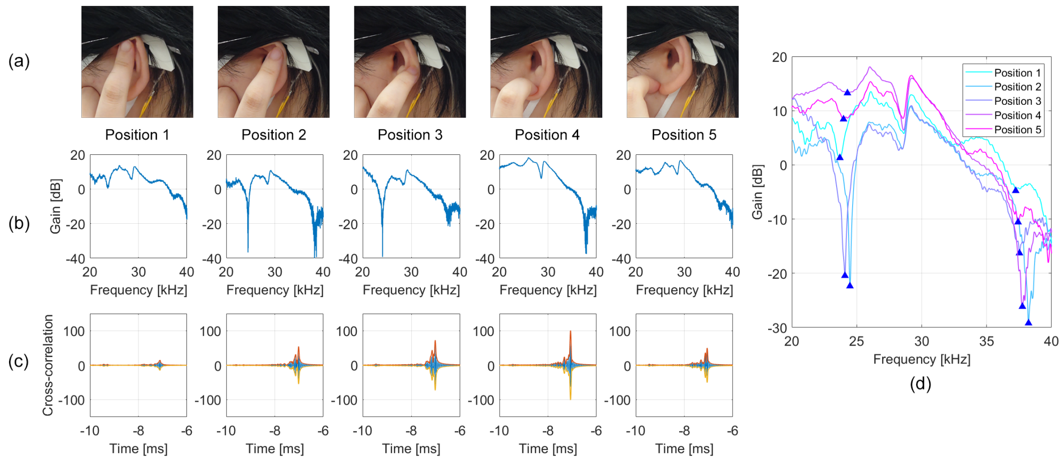

Figure 2: The finger was placed in five positions as shown in (a), and (b) shows the frequency response of the recorded sound, and (c) shows the cross-correlation between the recorded sound and the sound source. (d) overlays all frequency responses shown in (b), and the notches are highlighted by triangle marks.

2 PROPOSED FINGER POSITION ESTIMATION

The proposed approach uses a bone conduction speaker in contact behind the ear and a microphone in front of the ear to estimate the finger position, as shown in Figure 1. We selected this bone conduction speaker placement based on a finding that sound can be efficiently presented with a small pressure here [2]. The microphone was placed in front of the ear to efficiently pick up the sound leaking from the ear, modulated by finger placement. This way, the sound transmission path changes according to the finger position, changing the recorded signal.

3 EXPERIMENTAL EVALUATION

All the experiments in this section were conducted with the following setup: the bone conduction speaker placed behind the ear plays an inaudible linear chirp signal (20–44 kHz with a sampling frequency of 96 kHz, 8192 samples) while the user slowly swipes the ear with a finger (Figure 2 (a)). We ensured that the finger swiped the ear without pressing to observe the effect of finger position and suppress the effects caused by pressure or ear deformation. Then, the microphone (Classic Pro CEM1AK), placed in front of the ear, picks up each acoustic response.

As the acoustic wave propagation around the ears is generally complex and cannot be described with a straightforward model, we considered applying two approaches widely used for active acoustic sensing to track the finger's position. The first approach views the setup as an acoustic system with frequency characteristics dependent on finger position. The frequency spectrum of the signals picked up at the microphone captures the change in the modal frequency around the pinna caused by finger placement. Figure 2 (b) shows the frequency spectrum of the signal transmitted from the speaker to the microphone corresponding to each finger position shown in Figure 2 (a), and Figure 2 (d) shows an overlayed graph of the frequency spectrums.

The second approach focuses on the effect in the time domain of the recorded signals caused by the finger placement to estimate the position. This effect is usually captured through cross-correlation between the transmitted chirp signal and recorded signal, which is a common method for measuring the time of flight, and Figure 2 (c) shows the results. The figures show that distinct frequency responses characterized by different notch frequencies can be captured through the first approach, while the second approach does not show a significant change in the peak position. These results suggest that accurate finger positioning could be achieved by viewing the ear as an acoustic system affected by finger position and calculating the changes of notches in the frequency response.

4 CONCLUSION AND FUTURE WORK

This paper showed that our system consisting of a bone conduction speaker and a microphone could obtain distinct frequency responses according to five different finger positions. These results indicate that our method could enable bone conduction headsets to capture continuous finger positions without additional hardware. Developing a signal processing algorithm to estimate the finger position from the frequency response will be future work. Furthermore, the robustness of the proposed method against different ear shapes between individuals and re-wearing the bone conduction speaker has to be examined.

Acknowledgements. This work was supported by JST ACT-X Grant Number JPMJAX200H, Japan.

REFERENCES

[1] Takashi Kikuchi, Yuta Sugiura, Katsutoshi Masai, Maki Sugimoto, and Bruce H Thomas. 2017. EarTouch: turning the ear into an input surface. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '17, Article 27)*. ACM, New York, NY, USA, 1–6.

[2] Ryota Shimokura, Hiroshi Hosoi, Tadashi Nishimura, Toshiaki Yamanaka, and Harry Levitt. 2014. Cartilage conduction hearing. *The Journal of the Acoustical Society of America* 135, 4 (April 2014), 1959–1966.

[3] Xuhai Xu, Haitian Shi, Xin Yi, Wenjia Liu, Yukang Yan, Yuanchun Shi, Alex Mariakakis, Jennifer Mankoff, and Anind K Dey. 2020. EarBuddy: Enabling On-Face Interaction via Wireless Earbuds. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. ACM, New York, NY, USA, 1–14.