VibAware: Context-Aware Tap and Swipe Gestures Using Bio-Acoustic Sensing

Jina Kim[*] KAIST Daejeon, Republic of Korea

MinYung Kim[†] KAIST Daejeon, Republic of Korea

Woo Suk Lee[‡] Applied Sciences Group, Microsoft, Redmond, WA, USA

Sang Ho Yoon[§] KAIST Daejeon, Republic of Korea

Abstract

We present VibAware, context-aware tap and swipe gesture detection using bio-acoustic sensing. We employ both active and passive sensing methods to recognize microgestures and classify inherent interaction contexts. Here, the interaction contexts refer to interaction spaces, graspable interfaces, and surface materials. With a context-aware approach, we could support adaptive input controls to enable rich and affordable interactions in Augmented Reality and Mixed Reality for graspable, material-based interfaces. Through an investigation and preliminary study, we confirmed the feasibility of tap and swipe gesture recognition while classifying associated contexts.

Index Terms: Human-centered computing—Interaction techniques; Human computer interaction(HCI)—Interaction paradigms—Mixed / augmented reality;

1 Introduction

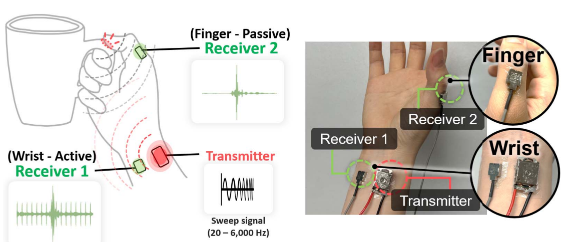

Figure 1: Vibaware employs sensing nodes at the wrist and finger which capture active and passive signals accordingly. Microgesture creates vibration captured by the passive sensor and distorts acoustic signals from the transmitter captured by the active sensor.

With the advancement in wearable sensing techniques, researchers have explored a wide category of hand interactions and diverse sensing approaches of recognizing microgestures [2,6]. Particularly, micro hand gestures like tap and swipe using a single hand have been adopted for contemporary augmented and virtual reality (AR/VR) [3]. To this extent, on-body sensing approaches promoting user comfort in a socially acceptable form factor have been suggested for microgesture recognition [1,4]. However, these works only focused on detecting gestures in a discrete manner. Extra care still must be taken to robustly recognize micro hand gestures under different interaction spaces such as graspable interfaces or surface materials. Therefore, our aim is to enable context-aware microgesture recognition through bio-acoustic sensing. This allows us to cover broader interaction scenarios. In our work, we further advance the sensing capability by understanding deeper semantics gestures (*e.g.*, tap while grasping a pen).

*e-mail: jina1190@kaist.ac.kr
†e-mail: min.kim@kaist.ac.kr
‡e-mail: woolee@microsoft.com
§e-mail: sangho@kaist.ac.kr

2 Design Space for Context-Aware Microgestures

We explored the representative microgestures and related contexts to provide and enrich the microgestures interface. Previously, researchers emphasized the importance of microgestures including Tap, Swipe, Draw, and Press which could provide direct and subtle interactions [5]. Among them, tap and swipe have been reported as the most preferred gestures while the thumb and index finger have shown high flexibility and comfort. To this end, we select **tap** and **swipe** performed by **thumb** and **index finger** as our representative microgestures as shown in Figure 2. As we mentioned, we categorize the interaction space into *within-hand, graspable interfaces, and everyday surfaces*. This adaptive approach would broaden the interaction scope without adding extra gestures or hardware.

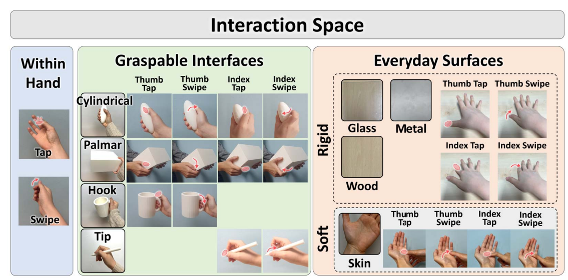

Figure 2: Design space - Gesture set categorized as *Within-Hand, Graspable Interfaces*, and *Everyday Surfaces*

3 System Overview and Investigation

3.1 Proof-of-concept system

Our Proof-of-concept system is composed of 2 accelerometers (1-axis accelerometer: VS–BV203–B, KEMET) as our active & passive receiver and a surface transducer (COM-10917, Sparkfun) as our active acoustic source for sinusoidal sweep signals from 20 Hz to 6,000 Hz. The components are small enough to be placed on the finger and the wrist as shown in Figure 1(Right). We capture motion-induced low bandwidth (passive) and surface transducer-generated broad bandwidth (active) bio-acoustic signals for our proposed work.

3.2 Signal pattern observation

We observed the signal pattern of gestures in a different context. Figure 3 shows that the frequency response of active signal under 6,000 Hz and passive signal under 500 Hz exhibits distinctive changes upon gestures. These acoustic signals during different conditions illustrate that discernible and rich sensing signals are generated to be used for context-aware gesture recognition. Based on our investigations, we expect that the change in the frequency response from microgestures would create unique features for recognition.

2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) | 979-8-3503-4839-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/VRW58643.2023.00147

979-8-3503-4839-2/23/$31.00 ©2023 IEEE
DOI 10.1109/VRW58643.2023.00147

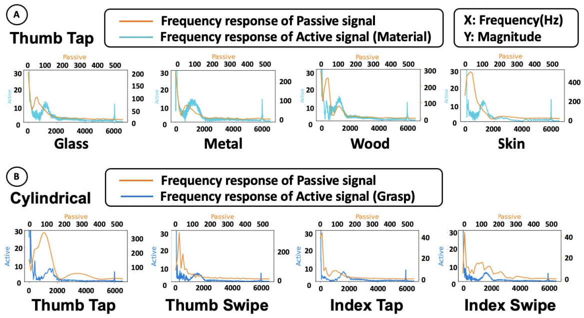

Figure 3: The observed pattern of the signal is different according to each condition: (A) ThumbTap on different materials and (B) Different gestures on cylindrical grasp

3.3 Preliminary Study

We conducted a preliminary study to verify the basic performance of gesture recognition in different interaction spaces (5 participants, mean age of 24). Here, we also compared the performance among different pre-processing methods while each processing method has the same time segment & feature extraction step. 1) **PP1**: Without applying bandpass filter, 2) **PP2**: Applying single bandpass filter (finger-attached passive sensor: 10∼500 Hz, wrist-attached active sensor: 10∼6,000 Hz), 3) **PP3**: Adding more bandpass filters to wrist-attached active sensor (10∼500 Hz, 100∼3,000 Hz). For context-aware gesture recognition, the processing pipeline is divided into 3 steps as shown in Figure 4(B).

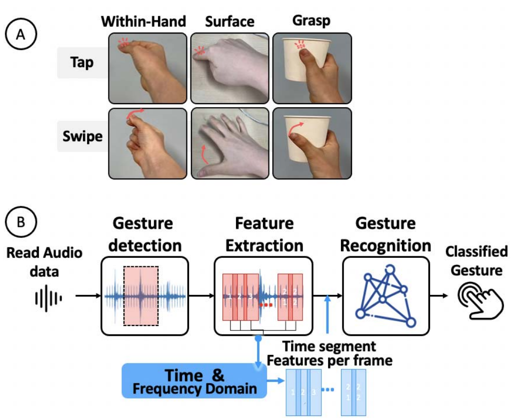

Figure 4: (A) Gesture set for preliminary study, (B) Gesture recognition processing workflow

We asked participants to perform provided gestures as shown in Figure 4(A) in random order. A total of 3,000 samples from each receiver (2 receivers × 5 participants × 5 sessions × 6 gestures × 10 trials). The average accuracy across the participants was improved by 96.8%, 97.2%, and 97.5% when employing multiple bandpass filters. This tells us it is crucial to focus on the effective range of bandwidth to extract meaningful acoustic features.

4 Envisioned Application

VibAware could be used for a context-aware gesture. **For finger interaction for AR Input**, sending materials to a partner by a swipe gesture with both hands or replying to a message by a tap gesture within a single hand without getting interrupted by bringing their hands to the camera to recognize hand gesture based on camera-vision. **For grasp-based** example, the user gets a phone call while

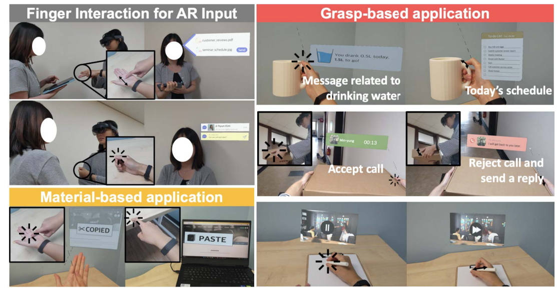

Figure 5: Example Applications

moving the box by tapping it or rejects it with a reply message by swiping it. Also, playing or pausing a video by tapping the pen and skipping forward by swiping it while writing. **For material-based** example, distinguishing between soft hands and solid materials could be used as a copy&paste function between AR glasses and other digital devices. The user can tap her hands to copy the image on her AR glasses and by tapping the desk, the captured page opens on the computer monitor on the desk.

5 Discussion and Future Work

Our study shows the potential that active and passive acoustic sensing recognizes the gestures while understanding associated contexts. In the future, we would like to further understand the frequency response of other physical properties in different interaction spaces for a more rigorous study and expand the types of hand grasps and materials. And we need to design hardware to make the system a wearable form factor. Our research would contribute to bringing expanded interactions for future input metaphors and adaptive context-aware microgesture-based interactions which provide distinctive controls using according to different interaction contexts.

Acknowledgments

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2022R1F1A1065277) and (No. 2022R1A4A5033689).

References

[1] Y. Iravantchi, Y. Zhang, E. Bernitsas, M. Goel, and C. Harrison. Interferi: Gesture sensing using on-body acoustic interferometry. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2019.

[2] Y. Kubo, Y. Koguchi, B. Shizuki, S. Takahashi, and O. Hilliges. Audiotouch: Minimally invasive sensing of micro-gestures via active bio-acoustic sensing. Association for Computing Machinery, Inc, 10 2019. doi: 10.1145/3338286.3340147

[3] G. Li, D. Rempel, Y. Liu, W. Song, and C. H. Adamson. Design of 3d microgestures for commands in virtual reality or augmented reality. *Applied Sciences (Switzerland)*, 11, 7 2021. doi: 10.3390/app11146375

[4] A. Mujibiya, X. Cao, D. S. Tan, D. Morris, S. N. Patel, and J. Rekimoto. 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*, pp. 189–198, 2013.

[5] A. Sharma, J. S. Roo, and J. Steimle. Grasping microgestures: Eliciting single-hand microgestures for handheld objects. Association for Computing Machinery, 5 2019. doi: 10.1145/3290605.3300632

[6] C. Zhang, Q. Xue, A. Waghmare, R. Meng, S. Jain, Y. Han, X. Li, K. Cunefare, T. Ploetz, T. Starner, O. Inan, and G. D. Abowd. Fingerping: Recognizing fine-grained hand poses using active acoustic on-body sensing. vol. 2018-April. Association for Computing Machinery, 4 2018. doi: 10.1145/3173574.3174011