An Ear Canal Deformation based Head Gesture Recognition Using In-ear Wearables

Youngone Lee
Department of Computer Science
Trinity University
San Antonio, Texas, USA
email: ylee5@trinity.edu

Sheng Tan
Department of Computer Science
Trinity University
San Antonio, Texas, USA
email: stan@trinity.edu

***Abstract*—Hands-free interfaces have become increasingly popular due to the growing demands for convenient control/interaction with mobile and wearable devices. Among all of the hands-free interfaces, head gestures interaction has shown great potential in providing alternatives under various real-world scenarios such as interfaces for people with disabilities, Virtual/Augmented Reality (VR/AR), and vehicle driving. However, existing head gesture recognition systems require either Line-Of-Sight (LOS) or the user to wear specialized hardware. Additionally, those approaches could raise potential privacy concerns. In this work, we propose a novel in-ear wearable system that can achieve head gesture recognition by utilizing off-the-shelf earbuds with a built-in microphone. Specifically, we leverage the relationship between the deformation of the ear canal and the head motion to distinguish different head gestures. A preliminary study shows our system can achieve over 94% recognition accuracy for various head gestures.**

***Keywords*—wearable; human computer interaction; head gesture; Internet-of-Things (IoT).**

I. Introduction

Up until recently, Human Computer Interactions (HCIs) on mobile devices have been dominated by contact interactions including touching the screen or pressing physical buttons. Because of the technological advancement of hardware along with the booming development of ubiquitous computing, a growing number of mobile and wearable devices (e.g., smart glasses, Internet of things devices, virtual reality/augmented reality devices, in-ear wearable devices) have been developed. To better facilitate the control over those emerging devices, more and more hands-free interfaces have been proposed such as gaze tracking [1], voice/speech interaction [2], brain wave control [3], and head posture recognition [4]. Among those novel approaches, head gesture recognition has shown great potential in providing alternatives for various real-life applications. For example, people with certain disabilities and drivers can leverage head gestures to interact with mobile and wearable devices [4]. Additionally, such an approach can be used to control head-mounted VR/AR devices.

Much research effort has been dedicated to developing different techniques for head gesture recognition. Traditionally, Computer Vision (cv) based approaches utilize cameras that can capture the image of the user's head motion to achieve gesture recognition. But, such a solution cannot work under Non-Line-Of-Sight (NLOS) scenarios and suffers from performance degradation in poor lighting conditions. Moreover, CV based approach could raise serious privacy concerns if the image data of the users is not managed properly.

Another body of work leverages motion sensors or Radio Frequency (RF) devices worn by the users to achieve head gesture recognition [4]. The motion sensor-based approach mainly relies on sensors such as accelerometers, and gyroscopes to infer the user's head motion speed and direction. On the other hand, RF devices (e.g., Radio-frequency Identification (RFID) tags, WiFi transceivers) mounted on the users can be used to measure the relative distance to the access point for head gesture recognition. However, those approaches all require specialized or customized hardware, which incurs non-negligible deployment costs. Additionally, some users might be reluctant or feel uncomfortable wearing additional devices.

In this paper, we aim to resolve those issues by proposing an in-ear wearable based head gesture recognition system. This work takes advantage of the Commercial Off-The-Shelf (COTS) earbuds to infer various head movements. It is done by sensing the unique ear canal deformation that closely correlates with distinctive head motions. The proposed system does not require any specialized or additional hardware other than COTS earbuds. Furthermore, our system is unobtrusive to the user during the recognition process and can enhance system security by leveraging user biometrics.

In particular, our system exploits the acoustic sensing approach that can detect the unique ear canal deformation caused by the head motion. The proposed system utilizes a sonar-like technique that can be implemented using any COTS earbud with a built-in microphone. To achieve this, the earbud speaker continuously sends an inaudible acoustic signal through the ear canal when the user is performing a head gesture. The in-ear microphone will capture the signal reflections that encompass ear canal deformation information. Next, our system will analyze the captured signal reflections to detect and distinguish various head gestures. We evaluate our system with a preliminary study and achieve over 94% accuracy in recognizing different head gestures.

The rest of the work is structured as follows. In Section II, we briefly describe the system flow. In Section III, we present the results of the preliminary study. Finally, in Section IV, we conclude and discuss future work.

II. System Design

The underlying principle of our head gesture recognition system lies in the fact that, when a user is performing the head gesture, the motion generated would result in dynamic ear canal deformation with distinctive features. Our system leverages COTS in-ear wearable devices to sense the unique ear canal deformation associated with the head movement of the user for gesture recognition. As shown in Figure 1, our system consists of four major components: *HGI (Head Gesture Interface) Activation*, *Signal Collection*, *Signal Processing*, and *Head Gesture Recognition*. The system first requires HGI Activation to initialize the recognition process which can be triggered by a particular head gesture of the user's choice (e.g., nodding the head or shaking the head). After the activation, the earbud speaker continuously emits an inaudible chirp signal (e.g., over 16kHz) to probe the ear canal. The signal reflected from the ear canal when the user is performing a head gesture will be captured by the inward-facing microphone, which can be further used to extract ear canal deformation information. For the *Signal Processing* component, we first apply various denoising techniques on the collected signal to reduce the inference contained in the captured signal reflections. Next, the denoised signals will go through the segmentation process to find the starting and ending times of the corresponding head gesture. It is done by leveraging the fact that the captured signal reflection will remain relatively consistent when there is no or minimal head motion. Then, the system will move on to the feature extraction component. In particular, we utilize the time-frequency analysis to extract the acoustic characteristics that represent the dynamic ear canal deformation. Lastly, *Head Gesture Recognition* will identify the gesture through the Support Vector Machine (SVM) based classification module and compare it against the pre-built user profile. If the head gesture is identified as unknown, our system will prompt the user to either perform the head gesture again or enroll the unknown gesture into the user profile.

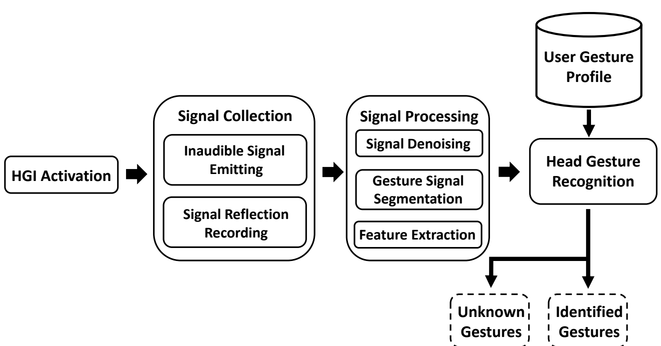

Fig. 1. Overview of system flow.

III. Feasibility Study

To demonstrate the feasibility of the proposed system, we built a prototype device utilizing a COTS in-ear earbud with an embedded microphone chip. The microphone is inward-facing and located in the center area of the speaker. We use Google Pixel 4a with Android 12 that connects to the prototype to control the inaudible probe signal emitting and the reflected signal recording. A chirp signal range from 16kHz to 23kHz is used for the probe signal. We designed four commonly used head gestures inspired by existing work [5]: down and up, up and down, clockwise rotation, and counter-clockwise rotation. Four participants were recruited - two females and two males for the feasibility study. We collected 100 samples from each participant by asking them to perform each head gesture 25 times in their manner. The environments involved in the study are the typical living room and bedroom area. The results are shown in Figure 2. We observe that the proposed system can achieve overall recognition accuracy of around 94%. This demonstrates that our system can effectively recognize various head gestures across different users.

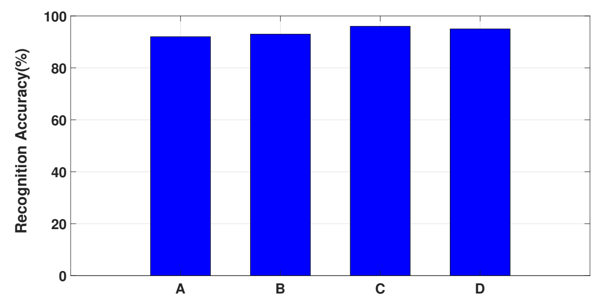

Fig. 2. Recognition accuracy of four different head gestures (A: down and up; B: up and down; C: clockwise rotation; D: counter-clockwise rotation).

IV. Conclusion and Future Work

In this work, we propose a head gesture recognition system utilizing COTS in-ear wearable devices which does not require LOS or any specialized sensor to work. The preliminary study shows that our system can recognize various head gestures with high accuracy. We plan to include more experiments under various scenarios/environments and use more sophisticated deep learning algorithm to further improve recognition accuracy in the future.

References

[1] C. H. Morimoto and M. R. Mimica, "Eye gaze tracking techniques for interactive applications," *Computer vision and image understanding*, vol. 98, no. 1, pp. 4–24, 2005.

[2] C. M. Rebman Jr, M. W. Aiken, and C. G. Cegielski, "Speech recognition in the human–computer interface," *Information & Management*, vol. 40, no. 6, pp. 509–519, 2003.

[3] C. K. Ho and M. Sasaki, "Brain-wave bio potentials based mobile robot control: wavelet-neural network pattern recognition approach," in *2001 IEEE International Conference on Systems, Man and Cybernetics. e-Systems and e-Man for Cybernetics in Cyberspace (Cat. No. 01CH37236)*, vol. 1, pp. 322–328, IEEE, 2001.

[4] K. Chen, F. Wang, M. Li, B. Liu, H. Chen, and F. Chen, "Headsee: Device-free head gesture recognition with commodity rfid," *Peer-to-Peer Networking and Applications*, vol. 15, no. 3, pp. 1357–1369, 2022.

[5] Y. Yan, C. Yu, X. Yi, and Y. Shi, "Headgesture: Hands-free input approach leveraging head movements for hmd devices," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 4, pp. 1–23, 2018.

 ISBN: ISBNFILL