Poster Abstract: Towards Recognizing Perceived Level of Understanding for Online Lectures using Earables

Dongwoo Kim
KOREATECH, Republick of Korea
dongwoo.kim@misl.koreatech.ac.kr

Chulhong Min
Nokia Bell Labs, UK
chulhong.min@nokia-bell-labs.com

Seungwoo Kang*
KOREATECH, Republic of Korea
swkang@koreatech.ac.kr

ABSTRACT

We envision that our earbuds recognize how much we understand learning materials while taking online lectures for effective learning and teaching, e.g., to pinpoint the part for which we need to put more effort to learn. To this end, we explore the feasibility of recognizing the perceived level of understanding of online learners based on IMU sensor data from earbuds. We present an exploratory study to identify head-related behaviors that can be detected by in-ear IMU data, which are associated with the perceived level of understanding for online lectures.

CCS CONCEPTS

• **Human-centered computing** → **Ubiquitous and mobile computing systems and tools**; • **Applied computing** → *E-learning*.

KEYWORDS

Online Learning, Understanding Level, Head Motions and Postures, Recognition

ACM Reference Format:

Dongwoo Kim, Chulhong Min, and Seungwoo Kang. 2020. Poster Abstract: Towards Recognizing Perceived Level of Understanding for Online Lectures using Earables. In *The 18th ACM Conference on Embedded Networked Sensor Systems (SenSys '20), November 16–19, 2020, Virtual Event, Japan.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3384419.3430428

1 INTRODUCTION

While online lecture provides attractive benefits, e.g., accessibility, comfort, learner's self-pacing, and low cost, it also has limitations compared to traditional, face-to-face learning. Since teachers cannot interact with students and catch nonverbal and behavioral cues from them during the lecture, it is extremely challenging to keep track of students' understanding and update the next lectures accordingly. One traditional method is to give quizzes after a lecture and analyze the answers, but such a manual process is limited in its scalability and granularity due to its significant costs.

To address the problem, we explore the feasibility of automatically recognizing the perceived level of understanding of online learners based on IMU sensor data from earbuds. Our research question is *whether some head motions and upper body movements of online learners can be a clue to estimate their understanding level for online lectures they are taking.* We note that students often use earbuds when they watch online lecture videos on their computers or smartphones. Utilizing earbud-integrated sensing modality is unobtrusive and efficient since it does not require other sensing devices that need to be worn by users or installed in an environment.

In this work, we examine the possibility of identifying behaviors associated with the perceived level of understanding which can be detected with in-ear IMU data. For the purpose, we collect in-ear IMU sensor data from 4 participants and their self-reported understanding levels while they watch online lecture videos. Based on the collected data, we identify key behaviors that naturally occur while taking online lectures and are relevant to the perceived understanding level. To recognize the understanding-relevant behaviors, we build a machine learning model with the IMU data and annotated behavior labels. Our result shows that the behavior detection model achieves 0.79 of F_1 score on average. Based on the results, we believe that we can build a model to estimate the perceived level of understanding from the detected behaviors.

2 PRELIMINARY STUDY

2.1 Data Collection

We recruited 4 participants in a university campus. We use eSense earbuds [2] to collect IMU data on the head/upper body movement of the participants taking an online lecture. The collected data are 3-axis accelerometer and 3-axis gyroscope data sampled in 32 Hz. We also record a video using a webcam to analyze the behaviors such as head and body movements of the participants exhibited during the lectures. We use online lectures provided by K-MOOC, a Korean MOOC established in 2015.

The participants watch the selected lecture videos and answer the questionnaire for their perceived level of understanding of the lectures. After taking the first lecture out of two, they take a break of about 5 minutes before the second lecture.

We consider the participant's self-reported understanding levels as the ground truth. In particular, we asked each study participant to answer a short survey for each lecture slide. We consider *a slide* as a unit of the *understanding* estimation. A slide often conveys a single topic and content, and thus it is naturally expected by teachers to be mapped the level of learner's understanding. For the survey, we made a questionnaire that consists of 5 statements. To make the questionnaire, we refer to the revised version of Bloom's Taxonomy regarding educational objectives [3].

*Corresponding author

SenSys '20, November 16–19, 2020, Virtual Event, Japan

ACM ISBN 978-1-4503-7590-0/20/11…$15.00
https://doi.org/10.1145/3384419.3430428

Table 1: Correlation coefficient between behavioral patterns and perceived level of understanding.

Behavior	GM			LD			PC			BD			NH			TH		
Stats	total	avg.	std.	total	avg.	std.	total	avg.	std.	total	avg.	std.	total	avg.	std.	total	avg.	std.
P1	-0.27	0.03	-0.18	0.03	0.07	0.13	-0.19	-0.19	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
P2	-0.26	-0.33	-0.48	0.13	0.12	0.22	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
P3	-0.32	-0.03	0.15	0.23	0.23	-0.02	-0.67	-0.67	N/A	-0.07	-0.07	N/A	N/A	N/A	N/A	N/A	N/A	N/A
P4	-0.58	-0.39	-0.35	0.51	0.13	0.36	-0.04	-0.04	N/A	N/A	N/A	N/A	0.35	0.35	N/A	-0.53	-0.53	N/A

2.2 Understanding-relevant Behaviors

We extract a set of behaviors that learners naturally make during online lectures. The researchers transcribe and code the recorded videos with observable, repetitive behaviors that the participants naturally performed while watching the online lectures.

We identify six observed behaviors from the analysis: 1) gazing at a monitor (GM), 2) looking down desk (LD) for note-taking or material-reading, 3) posture change (PC), 4) being distracted (BD) to smartphones, etc, 5) nodding the head (NH), and 6) tilting the head (TH). Interestingly, while learners' macroscopic, whole-body movements were quite limited during online lectures as expected, mostly to sitting and watching the video, a diversity of their microscopic movements were observed, especially around the head and upper body. The results validate our choice of earbuds as a sensing device to examine learners' behaviors on online lectures.

We identify key behaviors that can be used as clues to estimate the perceived level of understanding. To understand the impact of our observed behaviors, we assess statistical associations for every slide, between a) the statistic of observed behaviors and b) the reported levels of understanding in the questionnaire. For the statistic of the observed behaviors, we annotate the start and end timestamp of every behavior event, measure each event's duration, normalize the duration by dividing by the slide duration, and compute the total, average, and standard deviation duration of each behavior. For the analysis, we use Spearman's rank correlation.

Table 1 shows the correlation coefficients. The results show two important findings. First, some behavioral patterns show a moderate relationship to the perceived level of understanding, over 0.15 or under -0.15, showing the feasibility of estimating a learner's *understanding* from the observation of behavioral patterns. Interestingly, we can see that the statements of the participants in the interviews support our analysis. For example, P1 reported that he kept gazing vacantly at the monitor when he understood poorly. Such patterns are observed through the coefficient analysis, e.g., the negative relationship of the total duration of *gazing at a monitor*, -0.27. Second, the relationship between behavioral patterns and the perceived level of understanding is different depending on the individual, showing the need for personalized estimation.

Based on the analysis, we determine to detect two behaviors as primitive contexts, gazing at a monitor and looking down desk, as they generally show the meaningful relationship to the understanding level and are also commonly observed across the participants.

2.3 Behavior Detection on Earbuds

We develop a machine learning model to detect understanding-relevant behaviors using IMU signals on earbuds. One of challenges in distinguishing between *gazing at monitor* and *looking down desk* is that learners mostly remain stationary when they make these behavior. The relative direction of the earbuds to a person's head is mostly fixed. Thus, we extract features from X, Y, and Z streams separately without taking the magnitude and gather the features for the classification.

We first segment IMU streams into 1-second non-overlapping windows. Then, we extract time-domain and frequency-domain features reported in [1], from X, Y, and Z-axis streams. Then, we use PCA from the gathered features, and random forest as a classifier.

To investigate the detection accuracy of our model, we conduct 10-fold cross validation. To address the imbalanced data distribution, we use the synthetic minority over-sampling technique. The results show that our model reasonably detects understanding-relevant behaviors under the context of online lectures, 0.79 of average F_1 score (0.78, 0.75, and 0.86 for the the *gazing at monitor*, *looking down desk*, and others, respectively). We believe that we can further improve the performance by leveraging a larger amount of data and adopting advanced models.

3 CONCLUSION AND FUTURE WORK

We aim at estimating the perceived level of understanding of online learners based on IMU data from earbuds. We conduct an initial study with the IMU data from 4 participants and their self-reported understanding level scores while watching online lecture videos.

Based on a positive result in this initial study, we have several further study plans. First, we plan to develop a machine learning model to predict the perceived level of understanding based on detected behaviors. Second, we plan to collect data from more participants and lectures to build a more robust and accurate model. Lastly, we will extend and improve the current behavior detection model to support more diverse set of behaviors.

ACKNOWLEDGMENTS

This research was supported by National Research Foundation of Korea(NRF) and Institute of Information & Communications Technology Planning & Evaluation(IITP) funded by the Ministry of Science and ICT (2017M3C4A7066473, IITP-2020-0-01778)

REFERENCES

[1] Davide Figo, Pedro C Diniz, Diogo R Ferreira, and Joao MP Cardoso. 2010. Preprocessing techniques for context recognition from accelerometer data. *Personal and Ubiquitous Computing* 14, 7 (2010), 645–662.
[2] Fahim Kawsar, Chulhong Min, Akhil Mathur, and Alesandro Montanari. 2018. Earables for personal-scale behavior analytics. *IEEE Pervasive Computing* 17, 3 (2018), 83–89.
[3] David R Krathwohl. 2002. A revision of Bloom's taxonomy: An overview. *Theory into practice* 41, 4 (2002), 212–218.