Smartphone Cover-State Classification via Acoustic Sensing for Smartphone Search in Indoor Environments

Haruya Nishi*, Shigemi Ishida*, Tomoki Murakami† Shinya Otsuki†
*Graduate School/School of Systems Information Science, Future University Hakodate, Japan
†Access Network Service Systems Laboratories, Nippon Telegraph and Telephone Corporation, Japan
Email: *{g2123046, ish}@fun.ac.jp

***Abstract*—When we lose our smartphone, we often try to search the smartphone by listening to ringing sounds from the smartphone. This approach is inefficient because it requires individuals to rely on their sense of hearing and walk around their homes. We propose a smartphone searching method relying on the state of the smartphone's surroundings. The state of the smartphone's surroundings can be described by a cover state, contacting objects, and the room where the smartphone is located. In this paper, we show the feasibility of the cover-state classification via acoustic sensing. Reverberation data of a ringing sound was collected for different cover states with two different covering objects. We evaluated the cover-state classification performance using the collected data and confirmed that our method successfully classified the cover state with a mean accuracy of 0.66.**

***Index Terms*—acoustic sensing, smartphone, cover state,**

I. Introduction

In recent years, the number of smartphones lost indoors has increased. Many users search for their smartphones by making a call from another device and by identifying the source of the ringing sound. However, this method relies on the users' senses and requires time and effort to find the smartphone.

There are researches on indoor location estimation using smartphones' built-in sensors [1], [2] or external devices [3], [4], which do not assume the case where a smartphone is lost in a home. These methods face three problems in a house environment. The first problem is that objects near the smartphone can cause errors in estimation accuracy. The second is that the estimation results in terms of the relative distance between the reference station and the smartphone are not easy to understand for users looking for their smartphone. The third is that there is a risk of the smartphone's built-in sensors being exploited because of the sensors' remote use.

We therefore propose an in-home smartphone search-assistance system that uses a smart speaker to help users find a lost smartphone more quickly and accurately. In this paper, as a first step to realizing the proposed system, we present a smartphone cover-state classifier via acoustic sensing. We experimentally evaluate the proposed cover-state classifier and show that the mean classification accuracy was 0.66.

II. Smartphone Cover-State Classifier via Acoustic Sensing

The key idea of our cover-state classifier is to focus on how the sound is muffled by the cover state. Muffled sound affects harmonic structure for frequency components. We examine the harmonic structure of ringing sounds to identify the cover state.

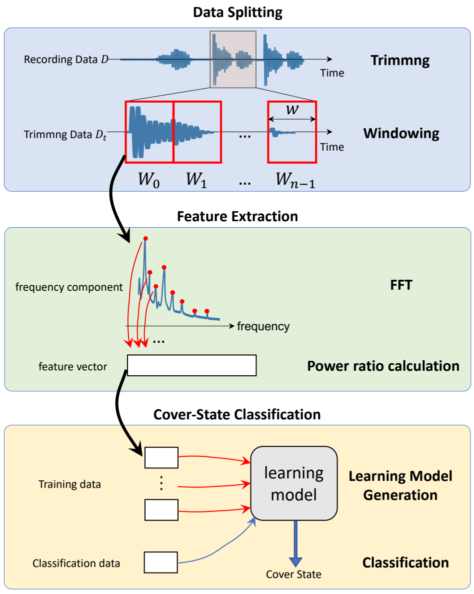

Fig. 1. System overview

In this study, we define three types of cover states: no-cover (covered by nothing), single-covered (covered by a single object), and multi-covered (covered by multiple objects).

Figure 1 shows the overview of the cover-state classifier. The classifier consists of data splitting, feature extraction, and cover-state classification blocks.

The data splitting block divides data into fixed data-length windows by performing trimming and windowing steps. The trimming step extracts single ring-tone data from the recorded sound data, which are divided into fixed-length windows in the windowing step.

The feature extraction block extract frequency-domain features for each window in two steps. The first step performs the fast Fourier transform (FFT), deriving frequency components of sound signals for each window. In the second step, the power ratio of fundamental to harmonic frequency components is calculated to construct a feature vector.

The cover-state classification block finally estimates cover state as a multi-class classification problem using a supervised learning model. In a training step, the learning model is

This is the author's version of the work.
© 2023 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
doi: 10.1109/GCCE59613.2023.10315431

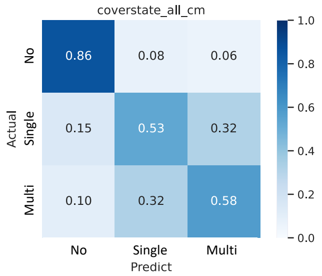

Fig. 2. Confusion matrix of cover-state classification results

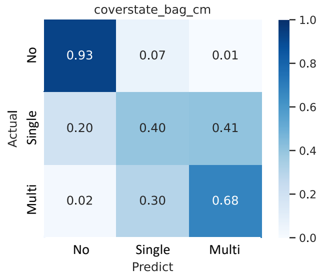

Fig. 3. Confusion matrix of classification results (covered by bag)

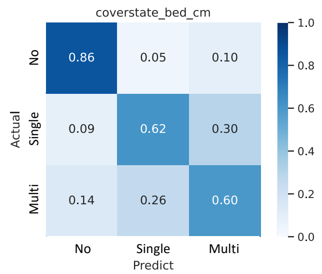

Fig. 4. Confusion matrix of classification results (covered by bedquilt)

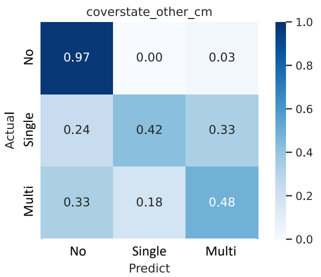

Fig. 5. Confusion matrix of classification results (covered by unknown)

trained with labeled data collected in various rooms with many objects. In a classification step, the cover state is estimated using the given data.

III. Evaluation

A. Experiment Environment

We collected ringing-sound recordings while changing the cover state and the covering objects of a smartphone in an indoor environment. An audio-technica AT2050 microphone, emulating a microphone on a smart speaker, was installed at the center of a room. An Apple iPhone XR smartphone was installed 0.8 meters away from the microphone on a bag or on a bedquilt. The bag and bedquilt were also used as covering objects. We executed the *Find My* app to make the smartphone ringing and recorded the ringing sound with a ZOOM H6 recorder for three calls of ringing. The sound was recorded at a sampling rate of 44.1 kHz and with a code length of 16 bits. The smartphone orientation was also changed within portrait, landscape, and upside down. For each combination of the covering object, cover state, and smartphone orientation, we collected 11 recordings, resulting in 198 recordings in total. We performed 10-fold cross-validation to evaluate the cover-state classification performance. The window size in the data splitting block was set to 4096 samples, which corresponds to approximately 93 milliseconds.

B. Cover-State Classification Performance

We first evaluated the classification performance. Figure 2 shows the confusion matrix of the classification results. The mean classification accuracy was 0.66. The accuracies for single- and multi-covered states were 0.53 and 0.58, respectively, which were lower than the accuracy of 0.86 for the no-cover state. The similarity of the sound muffling between single- and multi-covered states made it difficult to distinguish these cover states.

C. Effect of Covering Objects

To verify the effect of covering objects on classification accuracy, we evaluated the classification performance when we only used a single covering object. We trained and classified using the data of each of the covering objects, i.e., a bag and bedquilt, to calculate classification accuracy.

Figures 3 and 4 show the confusion matrices of the classification results for bag and bedquilt data, respectively. The mean classification accuracies for no-cover and multi-covered states were approximately 0.8 and 0.6, respectively. We can say that there was no significant effect of the covering object on classification performance. On the other hand, the mean accuracies for single-covered state for bag and bedquilt were 0.40 and 0.62, respectively, showing a discrepancy in classification accuracy. Single-covered states with a bag are often misclassified into multi-covered states. Bags are made of thin cloth, which made it difficult to distinguish single- and multi-covered states.

D. Classification with Unknown Covering Objects

To verify the classification performance when a smartphone is covered by unknown objects, we trained the model using a covering object and classified using the other covering object. Figure 5 shows the confusion matrix of the classification results for unknown covering objects. The mean classification accuracy was 0.63. The accuracies for no-cover, single-covered, and multi-covered states were 0.97, 0.42, and 0.48, respectively. Compared to Fig. 2, we can see that the accuracy was significantly lower than that for known objects except for the no-cover state. The sound muffled by an unknown object is unknown, which restricts the performance of supervised learning. In particular, the distinction between single- and multi-covered states was nearly random. We can still realize *covered or not* classification for unknown objects as the accuracy for the no-cover state was significantly high.

IV. Summary

In this paper, we present a smartphone cover-state classifier via acoustic sensing, helping people find their smartphone lost in a house. We conducted experimental evaluations and showed the feasibility of the proposed cover-state classification. In our future work, we plan to improve the accuracy of the cover-state classification and also plan to estimate the rooms where smartphones are located.

Acknowledgment

This work was supported in part by JSPS KAKENHI Grant Numbers JP20KK0258 and JP21K11847 as well as the Cooperative Research Project of RIEC, Tohoku University.

References

[1] Cho, J., Hwang, I. and Oh, S.: Vibration-Based Surface Recognition for Smartphones, *IEEE RTCSA*, pp. 459–464 (2012).
[2] He, S. and Shin, K. G.: Geomagnetism for Smartphone-Based Indoor Localization: Challenges, Advances, and Comparisons, *ACM CSUR*, Vol. 50, No. 6, pp. 1–37 (2017).
[3] Kriz, P., Maly, F. and Kozel, T.: Improving Indoor Localization Using Bluetooth Low Energy Beacons, *Mobile information systems*, Vol. 2016 (2016).
[4] Xiao, C., Yang, D., Chen, Z. and Tan, G.: 3-D BLE Indoor Localization Based on Denoising Autoencoder, *IEEE Access*, Vol. 5, pp. 12751–12760 (2017).