Poster Abstract: Tap it and You Know What It is: A Surface Identification System Based on Acoustic Dispersion

Baojie Yuan, Shicong Hong, Yongpan Zou, Kaishun Wu
{yuanbaojie2018}@email.szu.edu.cn;KyoDante@outlook.com;{yongpan,wu}@szu.edu.cn
College of Computer Science and Software engineering, Shenzhen University

ABSTRACT

Surface identification provides contextual services during human-computer interaction, which is important for target detection and scene understanding. A robust and ubiquitous surface recognition system has a wide range of applications such as context awareness and robot operation. Existing methods have shortcomings of requiring specialized devices and limited usage scenarios. In this paper, we introduce Surtify, a surface identification system based on acoustic dispersion with a smartphone. By combining the intrinsic physical phenomenon (*i.e.*, acoustic dispersion) with a deep learning model, Surtify can identify eleven kinds of surfaces with accuracies up to 96%, even in cross-person and cross-location scenarios.

CCS CONCEPTS

• **Human-centered computing → Ubiquitous and mobile computing systems and tools**.

KEYWORDS

material recognition, touch interface, acoustics

ACM Reference Format:

Baojie Yuan, Shicong Hong, Yongpan Zou, Kaishun Wu. 2020. Poster Abstract: Tap it and You Know What It is: A Surface Identification System Based on Acoustic Dispersion. 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/10.1145/3384419.3431187

1 INTRODUCTION

Surface identification provides rich physical information about the target object and has important applications in robot operation and context awareness. By identifying objects around smart devices, they can provide users help more conveniently. For example, in the context of smart home, by easily identifying the surface where a device is located, it can provide services with other smart devices on the surface. Aiming at this goal, researchers have come up with many different techniques to infer the material categories, which are mainly based on vision and radio-frequency (RF) signals. Nevertheless, they have disadvantages such as privacy risks, bulky specialized hardware and being easily affected by environmental changes. There are also methods based on sound signals or combining sound signals with other inertial sensors. Due to the difference in density and other physical properties of solid materials, there are great differences in the sound generated by tapping on its surface, which makes the use of sound recognition technology to recognize surface has a certain theoretical value and research. significance. However, Most of the existing sound-based surface recognition methods extract traditional acoustic features or combine them with other sensor data[1]. There are some researches have explore the tapping points on different solid surface[2] but they also need special hardware to carry.

To overcome the above shortcomings, we propose a convenient and robust solid material recognition system in this paper. As shown in Fig.1, the system only needs a commercial device equipped with a microphone to complete the solid material recognition instead of carrying other special devices. When tapping a solid material surface, the sound generated by the tapping propagates along the surface of the solid material and a physical phenomenon called *acoustic dispersion* occurs that is different frequency components of tapping sound propagate at different speeds on the surface. In this paper, We first convert the original audio signals into time-frequency matrices and build a deep learning model to preliminarily verify the feasibility of the dispersion feature to recognize different materials. After that, we construct a multi-channel datasets by combining data from different tapping points to explore the dual-dispersion characteristics on surface categories in-depth. The recognition accuracy of the system based on our multi-channel datasets can reach 96% in cross-person scenario.

2 BUILD MULTI-CHANNEL DATASET

For a certain surface, the propagation speed $V(f)$ of one certain frequency component f is constant determined by the physical properties of solid material. Additionaly, the high frequency components arrive more faster than lower frequency components. We can get a formula for the Time-Differenceof-Arrival (TDoA) between two different frequency components of a tapping sound which is linearly proportional varying with the propagation distance of the sound to the microphone (denoted as D) as follows:

$$TDOA(f_H, f_L) = D \cdot (\frac{1}{V(f_H)} - \frac{1}{V(f_L)}), \quad (1)$$

where $T^A(f_i)$ is the arrival time of the tapping sound at frequency f. This make tapping points at different distances to microphone show different dispersion patterns on their time-frequency distribution. Based on this observation, we combine tapping points at different distances from microphones to build multi-channel data to achieve data augmenting. An overview of our proposed system is shown in Fig.1.

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

ACM ISBN 978-1-4503-7590-0/20/11…$15.00
https://doi.org/10.1145/3384419.3431187

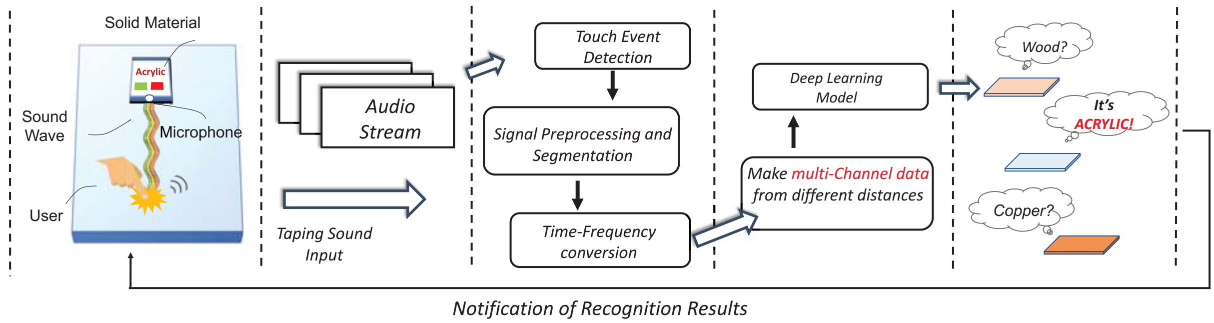

Figure 1: An overview of the system architecture of Surtify.

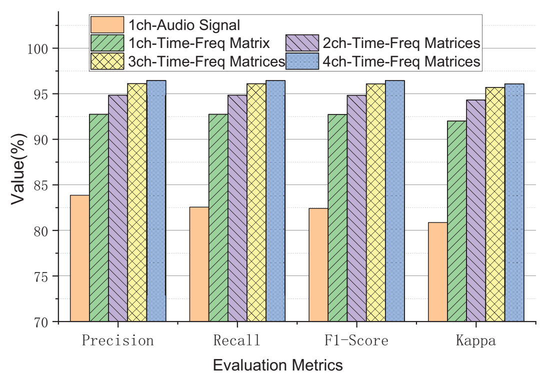

Figure 2: The comparison of recognition performance with different channel composition

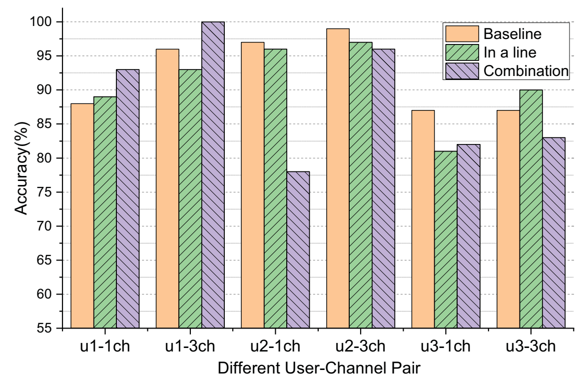

Figure 3: The comparison of recognition performance with different channel composition

3 EVALUATION

To evaluate the performance of Surtify, we implement a prototype on Huawei Mate 9 pro and conduct experiments with 25 participants. We test a total number of 11 kinds of common surfaces including There are nine baseline points. We set tapping points every 5 cm from 10 cm to 50 cm away from the microphone as baseline points. To capture data properly, the sampling rate of our system is 44100 Hz. In order to verify whether our dispersion features can be classified effectively, we train resnet50 on our multi-channel datasets for validation.

We mainly display the results of following two scenarios.

- **Leave-one-user-out scenario:** In order to test the performance of our system in this scenario, we train the model with data of 24 users and test it on that of the remaining one, evaluating our multi-channel methods on Resnet50.
  Fig. 2 indicates that our multi-channel methods can better deal with the diversity of users and produce a more general model. With the number of channel increasing, the recognition accuracy increases.
- **Random tapping-point scenario:** To verify that the system can be easily applied by users with little limitation, we also tested when tapping happened outside the baseline point. Fig. 3 shows that even the user isn't in the training dataset, our model can still recognize the type of target solid material.

4 CONCLUSION

In this paper, we prove the potential of identifying different surfaces based on an interesting physical phenomenon, *i.e.*, acoustic dispersion. To reduce the user training overhead without performance degradation, we extract intrinsic acoustic propagation patterns from received signals via time-frequency transformation. Moreover, we fully make use of the spatial characteristics of acoustic dispersion and construct multi-channel spectrogram data to train a deep learning model. The experiments show that our system can identify eleven kinds of surfaces with high accuracy in cross-person and -location scenarios.

ACKNOWLEDGMENTS

This research was supported in part by the China NSFC Grant 61802264, Shenzhen Science and Technology Foundation (No.JCYJ 20180305124807337), Natural Science Foundation of SZU (No.860-000002110537).

REFERENCES

[1] Semin Ryu and Seung-Chan Kim. 2020. Embedded Identification of Surface based on Multirate Sensor Fusion with Deep Neural Network. *IEEE Embedded Systems Letters* (2020).
[2] Semin Ryu and Seung-Chan Kim. 2020. Impact Sound-based Surface Identification using Smart Audio Sensors with Deep Neural Networks. *IEEE Sensors Journal* (2020).