Vol.2021-MBL-100 No.21
Vol.2021-UBI-71 No.21
Vol.2021-CDS-32 No.21
Vol.2021-ASD-21 No.21
2021/9/3

Preliminary investigation of obstacle recognition via smartphone active sound sensing for pedestrian safety

Thilina Dissanayake[1]　前川 卓也[1]　原 隆浩[1]

Abstract: To ensure the safety of the pedestrians that use a smartphone while walking, numerous smartphone applications have been developed. The most common method is to use the back camera of the smartphone to record the video in real-time to recognize obstacles such as approaching vehicles and people. However, this imposes problems such as limited detection ranges, and poor performance in the dark. To overcome these problems, we design a method that can recognize common obstacles such as vehicles, trees, signposts, etc. using smartphone active sound sensing. We mimic the echolocation of bats and emit sine waves and sweep signals from the smartphone and record the reflected waves. We exploit the spectral and spatial characteristics of the stationary and non-stationary obstacles and recognize them, making our method adaptable to numerous applications such as assisting visually impaired people, managing smartphone alert levels, and providing information regarding how crowded the sidewalk is..

Keywords: Mobile sensing, Pedestrian safety

1. Introduction

With the popularity of the smartphone on the rise, the safety of pedestrians using smartphones while walking has become a main concern to pedestrian safety as well as a popular theme for studies in the ubiquitous computing research community. Distracted pedestrians often bump into various roadside obstacles and even into other pedestrians causing severe injuries and even death. As an example, Nasar et al. [1] mention that the estimated number of injuries due to mobile phone use among pedestrians in the United States was 1506 in 2010, based on data from National Electronic Injury Surveillance System for 2004 - 2010. This issue has now become a socio-economic problem, increasing the burden on the health care and rehabilitation services sector.

To address the above problem, smartphone applications have been developed employing sensor data from different smartphone sensors. As an example, the rear camera of the smartphone is used to record the video in real-time and that footage is used to recognize obstacles such as approaching vehicles [2]. However, this solution has drawbacks such as limited detection range, poor performance in the dark, and high computational cost. Furthermore, recording videos in the public may sometimes be privacy-invasive and hence be illegal on some occasions.

Therefore, methods that are based on acoustic approaches have been exploited. As an example, Wen et al. [3] proposed a method that attaches an ultrasonic sensor to the back of the smartphone to detect abrupt changes in the ground in front of the user, such as stairs and potholes. Li et al. [4] proposed a method that employs passive noise from the vehicles to detect and count approaching vehicles. Furthermore, Wang et al. [5] proposed a method based on active sound sensing to detect upcoming obstacles using two microphones of off-the-shelf smartphones. However, these studies only concentrate on either (I) detecting the presence of an obstacle or (II) detecting a certain class of obstacles. In contrast, our method is designed not only to detect an obstacle prior to a collision but also to classify it into a class that provides rich information about it that can be adapted into

[1] Graduate School of Information Science and Technology, Osaka University

applications such as crowdsensing. As an example, by recognizing humans walking or riding bicycles on a sidewalk, crowdsensing applications can alert the user regarding the upcoming traffic. Furthermore, automation and managing alert levels in smart devices can also be achieved using obstacle recognition. For example, a person riding a bicycle towards the distracted smartphone user is much more dangerous than a garbage can placed on the sidewalk.

The contributions of our method can be listed as follows.

- To the best of our knowledge, this is the first study that recognizes classes of the obstacles in front of a smartphone user using active acoustic sensing.
- We propose a novel feature extraction method that can combine inherent acoustic features of an object such as the sound absorption using multiple microphones of the smartphone.
- We conducted a preliminary investigation on our methodology using data under several environmental conditions and under real-life situations.

2. Related works

Here, we discuss the studies related to improving pedestrian safety using smartphones.

2.1 Non-acoustic obstacle detection systems

In this section, we focus on pedestrian safety systems that employ sensors such as accelerometers, cameras, and infrared. Jain et al. [6] employ the camera of the smartphone to recognize the material on the pedestrian's walking surface. They identify when the distracted smartphone user is about to enter the street, which can be used to warn the user to be cautious. Wang et al. [2] propose a smartphone application that employs the rear camera of the smartphone to detect vehicles approaching the user and to alert the user prior to a collision. In contrast to the above approaches, we employ an acoustic sensing-based approach that minimizes the constraints to the holding orientation of the smartphone.

2.2 Acoustic-based obstacle detection systems

In this section, we present pedestrian safety systems that employ acoustic sensors such as microphones. Tange et al. [7] installed ultrasonic, gyro, and acceleration sensors in a white walking stick and create a system that can detect static obstacles such as walls and stairs to assist visually impaired persons. In contrast to this study, our proposed method is designed to work with smartphone sensors which enables it to be adapted to more versatile applications. Wang et al. [5] propose a method that can detect obstacles with smartphone active sound sensing. They emit inaudible sweep signals from the smartphone and use the reflections of the signals to calculate the distance from the user to the obstacle. In contrast, our method is not only designed to detect the obstacles but also to classify them into a specific obstacle class. Furthermore, we employ a combination of sweep signals and sine waves to recognize dynamic obstacles such as walking people.

Auto++, proposed by Li et al. [4] employ the passive noise from the approaching motor vehicles to detect them and estimate their driving direction. Furthermore, they extend their study and propose a method to count the number of vehicles around the user. In contrast to their method, our proposed method is designed to detect and recognize several obstacle classes using active sound sensing.

3. Preliminary investigation on obstacle detection using acoustic signals

In this section, we investigate the acoustic sensing methods to detect and classify stationary and locomotive obstacles. Furthermore, we design a probing signal which can be used for this task.

3.1 Detection of locomotive and static obstacles with Doppler shift

Here, we investigate the feasibility of employing a well-known phenomenon known as the "Doppler effect" or the "Doppler shift" to detect locomotive and static obstacles on the sidewalk. Doppler shift is the phenomenon of observing an increment or a decrement in the frequency of a wave when the observer is moving towards or away from the sound source. In this study, as we use the same smartphone to emit an 18 kHz sine wave and record the reflected waves, the smartphone becomes both the sound source and the observer. The Doppler shift is created on the reflected sine wave when the user moves with the smartphone relative to the obstacle, or the obstacle moves relative to the user by which the sound wave gets reflected.

For ease of analysis, we separate locomotive obstacle detection into two main scenarios.

(1) The user is stationary: Here, we assume that the user is holding the smartphone while staying stationary and the obstacle (a walking person) approaches the user. Figure 1 (a) and (b) shows the FFT spectro-

Vol.2021-MBL-100 No.21
Vol.2021-UBI-71 No.21
Vol.2021-CDS-32 No.21
Vol.2021-ASD-21 No.21
2021/9/3

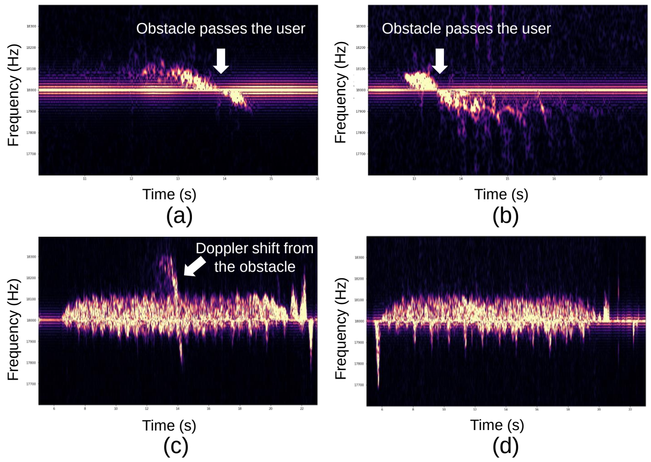

図 1 Doppler shifts created during different scenarios. (a) The user is stationary and a person is walking towards the user from the front. (b) The user is stationary and a person is walking towards the user from behind. (c) The user is walking towards a person while the person is also walking towards the user. (d) The user is walking towards a stationary person.

gram of the Doppler shift created on the sine wave when the walking person approaches the user from the front and from behind respectively. When the person walks towards the user, it creates a Doppler shift towards the positive direction of the spectrogram. When the person passes the user and moves away from the user, it creates a Doppler shift towards the negative direction of the spectrogram. Note that, when the obstacle is approaching from the front, it creates a longer positive Doppler shift followed by a shorter negative Doppler shift after the obstacle passes the user. Similarly, when the obstacle is approaching from behind, it creates a shorter positive Doppler shift followed by a longer negative Doppler shift after the obstacle passes the user. When the user is holding the smartphone in front of him, his body blocks the wave traveling to his rear, hence, the shorter Doppler shift when the obstacle is at the user's back.

(2) The user is mobile: Here, we assume that the user is holding the smartphone while moving forward and the obstacle (a person) is either moving towards the user from the front or remain stationary. Figure 1 (c) and (d) shows the FFT spectrogram of the Dopper shift created when the obstacle is moving towards the user and remains stationary respectively. Both (a) and (b) contain periodic positive and negative Doppler shifts created by moving arms, and legs of the user. However, Figure 1 (c) contains a characteristically large Doppler shift created by the person moving towards the user. Note that this Doppler shift is larger than the Doppler shifts created in Figure 1 (a) and (b) because in this scenario the user is also moving towards the person resulting in an increment of the relative velocity component of the person towards the user. However, when the person is stationary (Figure 1 (d)), the Doppler shift created by him cannot be distinguished clearly amongst the Doppler shifts created by the user. This is because now the relative velocity of the obstacle is roughly equal to the velocity of the user, hence having a relative velocity smaller than that of the moving body parts of the user. Therefore, it is necessary to implement a method to detect stationary obstacles that are not based on Doppler shifts.

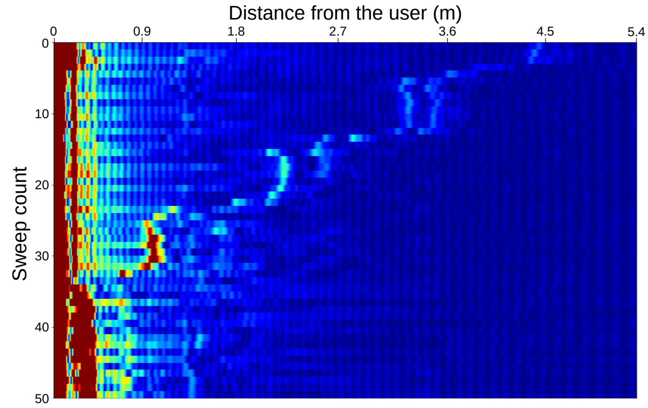

図 2 Reflections of the obstacle, recorded on the impulse response

3.2 Detecting stationary obstacles with sine sweeps

Here, we describe a method of detecting stationary obstacles with sine sweeps. The basic idea is to emit a short excitement signal to capture the reflections off the obstacle. Hence, we select a short sweep signal that sweeps the frequencies from 18 kHz to 21 kHz which lies in the inaudible range of human beings. We periodically emit (0.1 sec long sweeps every 0.5 sec) these sweeps from the smartphone and record the reflected waves. Next, we calculate the correlation between the recorded signal and the emitted sweep signal (See Section 4.3.2). Then, we extract the upper envelope of the resulted correlation.

Figure 2 shows the envelopes of 50 sweeps stacked according to their recorded time when the user is moving towards an obstacle (wall). As can be seen in the figure, the reflection from the wall gets closer to the user when

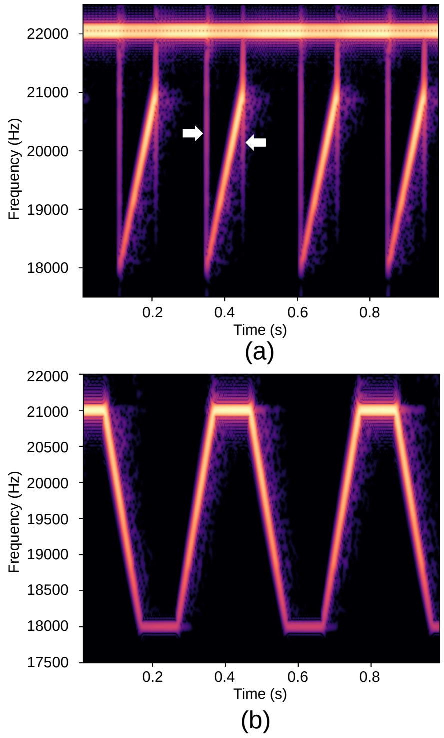

図 3 Two designs for the probing signal

he moves closer to the obstacle. Therefore, we will employ this method to detect stationary obstacles that are in front of the user.

3.3 Designing the signal

Since sweep signals can only be emitted discretely, by using only sweep signals, it is impossible to continuously probe for obstacles. Especially, when it comes to locomotive obstacles with high velocities, it is necessary to detect them from as far from the user as possible. Therefore, a probing signal that is compatible with both the sine sweep component and the sine wave component is necessary. Here, we explain two wave designs that can be used for this purpose.

Figure 3 shows the FFT spectrograms of the two wave designs. Figure 3 (a) shows a composite wave that is designed to emit periodic sweeps while emitting the sine wave uninterruptedly. This design allows detecting locomotive obstacles using the 22 kHz sine wave and stationary obstacles with 18 kHz - 21 kHz sweeps. The sine wave is separated from the upper bound frequency of the sweep with a 1 kHz bandwidth so that the Doppler shifts that occur on the sine wave will not get mixed with the sweep signal which will alter their frequency response. However, this design has two main flaws. The first one is that, as we are emitting two types of waves at the same time using the speakers, the sweep and the sine wave will be emitted with half the maximum power of the speakers, each. This adversely affects the maximum detection and recognition range of our method. The second flaw is caused by the signal power leakage between two consecutive sweep signals. Strong signal power leakages can be observed at the start and the end of each sweep due to frequency hopping. Unfortunately, these leaks make the sweep signals audible to the human ear. Xie et al. [8] proposed a method to reduce the influence of the signal power leakage by adding a tapered cosine window on the emitting signals. This method gradually increases the power of the sweep signal at the beginning from zero and decreases at the end back to zero. This reduces the power leakage at the start and the end of the sweep. However, this also changes the amplitudes of the sweep and alters the frequency response of the sweeps making it harder to be employed for obstacle classification. Therefore, it is necessary to design a novel signal to overcome these problems.

Since the signal power leakage is caused by frequency hopping, we design a wave combining sweeps and sine waves in a continuous manner. Figure 3 (b) shows the FFT spectrogram of the novel signal design. Here, every 18 kHz - 21kHz sweep is followed by a 21 kHz sine wave, which is then followed by a 21 kHz - 18 kHz inverse sine sweep and an 18 kHz sine wave. As can be seen in the figure, the signal power leakage has been diminished.

Using this method, we can probe for stationary obstacles using the sweep waves and locomotive obstacles using the sine waves. Since the locomotive obstacles can be considered to be high-priority obstacles, the length of the sine wave can be increased compared to the sweep signal. As an example, each 0.1 sec sweep can be followed up with a 0.5 sec sine wave. This way, it is possible to prioritize the locomotive obstacles

3.4 Classification of stationary obstacles

Here we explain how we classify the stationary obstacles. As mentioned in the Section 3.2, the frequency range of the sweep signal is from 18 kHz to 21 kHz. As the impulse responses of the sweeps contain reflections from the obstacle, we believe that the frequency responses of these impulse responses contain valuable information that can be used to distinguish between the obstacle classes. Therefore, we consider the amplitudes of the frequency bands in between the aforementioned frequencies.

Figure 4 shows the FFT spectra of six different obstacle classes between 18 kHz and 21 kHz, 50 different instances each. As can be seen in the figure, there are distinguish-

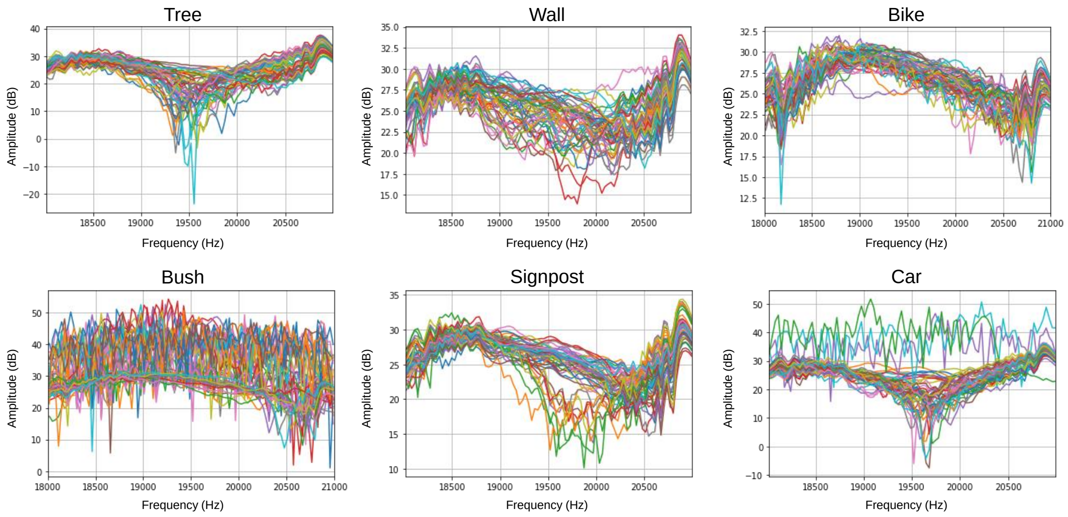

図 4 FFT spectra of different stationary obstacles

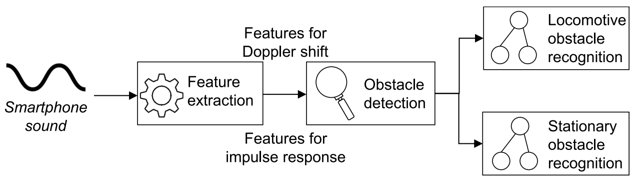

図 5 Overview of the proposed method

able features in spectra in different obstacle classes. As an example, the minimum value for the Tree class mostly occurs between 19 kHz and 19.5 kHz. In contrast, the minimum value for the Car class mostly occurs between 19.5 kHz and 20 kHz. We believe that we can employ these characteristic spectra to classify different obstacles into obstacle classes.

4. Method

4.1 Preliminaries

For this study, we assume that the user is holding a smartphone in front of his body. We assume the smartphone's position and the holding orientation, such that the user can view the content on the smartphone screen. A combination of inaudible short sweep signals and a sine wave will be emitted from the top and the bottom speakers of the smartphone with maximum volume. At the same time, the top and the bottom microphones of the smartphone record the reflected signals at a sampling frequency of 192 kHz.

4.2 Overview

Figure 5 shows the overview of the proposed method. After recording the reflections from the sine wave and the periodic sweeps, we extract two types of features. From the sine wave, we extract features for the Doppler shift. From the sweeps, we extract features for impulse response. During the next step, we use the features for Doppler shift to detect mobile obstacles and impulse responses to detect stationary obstacles. Furthermore, we employ these features and train Random Forest classifiers to recognize locomotive obstacle classes and stationary obstacle classes.

4.3 Feature extraction

4.3.1 Features for Doppler shift

Here we explain the process of extracting features for the Doppler shift. We apply 0.1 sec sliding window with 90% overlap to calculate the FFT spectrogram. We use the Hannig window as the window function. As we are only interested in the neighborhood of 18 kHz, we select 4000 frequency bins from either side of the bandwidths of the 18 kHz sine wave [9]. This produces an 8000-dimensional frequency vector sequence.

4.3.2 Features for impulse response

An impulse is a signal that is equal to one at time zero and is zero otherwise. It contains all the frequencies in the frequency domain. Since it is hard to generate such a strong pulse in the real world, we employ a sine sweep signal that sweeps the frequencies between 18 kHz and 21 kHz within 0.1 sec. At the same time, we record the reverberation of the reflected wave using the top and the bottom microphones (m_T and m_B) at the same time.

Next we calculate the impulse response of the recorded signal. Here, we calculate the convolusion of the recorded signal with the time-reverced trancemitted sweep signal. Let the received sweep signal at the top microphone m_T be $R_T(t)$. The impulse reponse $IR_k(t)$ of the signal $R_T(t)$ is given by

$$IR_T(t) = \int_{t_1}^{t_2} R_T(\tau)S(t-\tau)d\tau. \quad (1)$$

Here, $S(t-\tau)$ is the time-reversed signal of the emitted sweep.

Next, we extract the upper envelope of the resulted impulse response. The peaks of the envelop correspond to reflections from obstacles (Figure 2). By analyzing the time-series envelopes, that is, stacking a sequence of envelopes when the user is moving forward, we can detect the stationary obstacles in front of the user.

Next, we calculate the FFT spectrum of each impulse response. Furthermore, we extract 12th order Mel-Frequency Cepstral Coefficients (MFCC) of each sweep that give us the temporal variations of different frequency bands. Since the MFCC algorithm is more discriminative at lower frequencies and less discriminative at higher frequencies, we optimize it to our preferred 18 kHz - 20 kHz frequency range by adjusting the Mel-filter banks. Next, we flatten the MFCCs of each sweep along the time axis and concatenate them with the FFT spectrum vector of the sweep. Concatenated frequency and time-domain vectors are then used to recognize the stationary obstacles.

4.4 Detecting obstacles

The extracted envelops of the impulse responses from the sweeps will be employed to detect the stationary obstacles. We employ a peak detection algorithm that detects peaks that are higher than a given threshold. If a peak that exceeds the threshold is detected, we consider that an obstacle is nearby. Furthermore, by the position of the peak in the impulse response, the distance to the obstacle from the phone can be calculated. When looking at the first strongest peak in the impulse response, it can be considered as a result of direct propagation of the sweep wave from the speaker to the microphone. As the distance between the microphone and the speaker is not more than few centimeters, we can use this peak as the origin of a distance measuring coordinate system. After we recognize this peak, we calculate the number of samples until the next significant peak which represents an obstacle. Note that, there is a peak caused by the reflections from the user's body which occurs very close to the first peak. As we are recording the audio at a sampling rate of 192 kHz, by considering the speed of sound as 345 m/sec, we can calculate that each sample in the impulse response corresponds to 0.18 cm. In other words, the distance resolution of this method is 0.18 cm. As can be seen in Figure 2, for a considerable large obstacle such as a wall, we can detect reflections when the wall is 4.5 m away from the smartphone.

To detect locomotive obstacles, we leverage the Doppler shifts created by the moving obstacles.Here, we employ the frequency vectors we tailored in Section 4.3.1. We train a classifier with these vectors and recognize the Doppler shift created by a target obstacle. Furthermore, we use the characteristics of the Doppler shifts (see Section 3.1) to recognize if the obstacle is approaching from behind or from the front of the user.

4.5 Obstacle recognition

4.5.1 Locomotive obstacle recognition

Here, we recognize locomotive obstacles with the features extracted in Section 4.3.1. Since we are recognizing the obstacles that are on the sidewalk, we consider two main classes of locomotive obstacles, walking personals and moving bicycles. Implementing this section and testing the recognition range is an important part of our future works.

4.5.2 Stationary obstacle recognition

Here, we recognize the stationary obstacles with the features described in Section 4.3.2. We assume six stationary obstacle classes, namely; Tree, Wall, Bike, Bush, Signpost, and Car. We train a classifier using the feature vectors extracted from the impulse responses. Next, we use the classifier to recognize the class of the target obstacle.

5. Preliminary evaluation

5.1 Dataset

Here, we explain the process with which we collected the data for evaluation. A participant was asked to carry a Samsung Galaxy S20 Ultra 5G smartphone in his hand. He was asked to hold it in front of him in a way that he can read the content on the smartphone screen. Next, he was instructed to walk towards different obstacles in different environments at a normal speed. We collect eights sweeps when the user is within 1 m from each obstacle. As we are assuming six different obstacle classes, this can be considered as a six-class classification problem. We collected data from 20 different obstacles from each class, each situated in a different environment. As an example, the dataset of the Car class consists of different body types such as sedan-style cars, SUVs, and convertibles.

Furthermore, we collected additional nine sessions of data to evaluate the proposed locomotive obstacle detection. We separated the data into three different scenarios

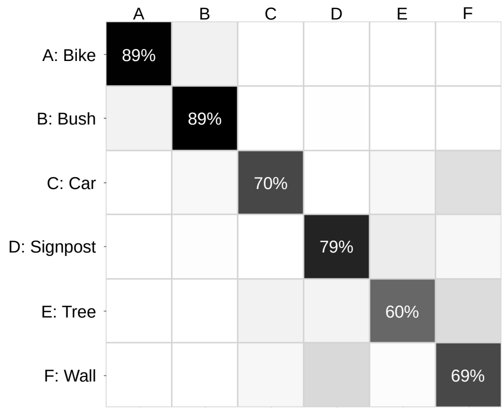

図 6 Visual confusion matrix of the overall results of stationary obstacle detection

表 1 Macro-averaged precision, recall, and F-measure of the stationary obstacle recognition

	Precision	Recall	F-measure
Overall	0.77	0.76	0.76
Majority vote	0.85	0.83	0.83

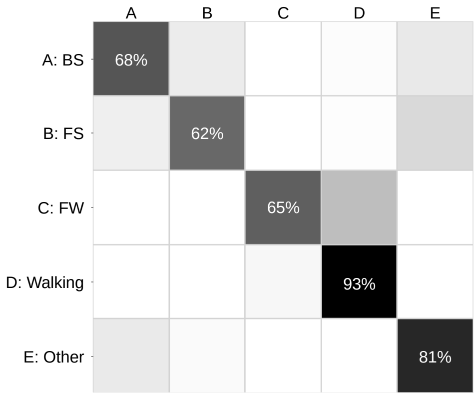

図 7 Visual confusion matrix of the overall results of locomotive obstacle detection

表 2 Macro-averaged precision, recall, and F-measure of the locomotive obstacle detection.

	Precision	Recall	F-measure
BS	0.63	0.68	0.66
FS	0.70	0.62	0.66
FW	0.77	0.65	0.71
Walking	0.84	0.93	0.88
Other	0.82	0.81	0.81

according to the locomotive patterns of the user and the obstacle.

- An obstacle is approaching from behind, the user is stationary (BS).
- An obstacle (a person) is approaching from the front, the user is stationary (FS).
- An obstacle is approaching from the front, the user is walking towards the obstacle (FW).

In addition, we also try to recognize the two additional classes, the Doppler shifts created during the user's walking ("Walking" class) and all the other data ("Other" class), including the Doppler shifts that occurred by moving the smartphone while standing still. Therefore, this problem can be considered as a five-class classification problem.

5.2 Evaluation methodology

We evaluate the proposed locomotive obstacle detection method, using "leave-one-session-out" cross-validation, where the Doppler shifts during one session were used as test data, and the Doppler shifts created from the remaining sessions were used as the training data.

We evaluate the proposed stationary obstacle recognition method, using "leave-one-obstacle-out" cross-validation, where sweeps collected from one obstacle were used as test data, and the sweeps collected from the remaining obstacles were used as the training data.

5.3 Stationary obstacle recognition

Figure 6 shows the visual confusion matrix of the overall results of the stationary obstacle recognition. As can be seen in the figure, all the classes have achieved good results. Table 1 shows the overall results of the stationary obstacle recognition. The overall macro-averaged precision, recall, and F-measure were 77%, 76%, and 76% respectively. The results of the Tree class is somewhat poor. Some of the Tree obstacles we collected the data from were situated close to a boundary wall. This could be the reason why some of the Tree instances have been misclassified into the Wall class.Designing a method to distinguish between multiple obstacles is a main part of our future work.

Since we have collected data by emitting eight sweeps per obstacle, we determine the obstacle class using the majority vote results of the sweeps. Table 1 also shows the results of the majority vote method. The macro-averaged precision, recall, and F-measure of the majority vote method were 85%, 83%, and 83% respectively.

5.4 Locomotive obstacle detection

Table 2 shows the overall results of the locomotive obstacle detection. As can be seen in the table, our method can differentiate between the Doppler shifts created by the user and the moving obstacles. Furthermore, it can also differentiate the Doppler shifts created by the moving obstacle under three different approaching scenarios.

6. Discussion

6.1 Device heterogeneity

To test the device heterogeneity of our method, we additionally collected sweep data from three instances from each class using the Galaxy Note20 Ultra 5G smartphone. We trained a classifier using the data collected from S20 and test the data collected from Note20. The macro-averaged precision, recall, F-measure was 15%, 14%, and 10% respectively. We believe that there are two main reasons why cross-device recognition failed. The first reason is that the frequency response of the microphone of Note20 is different from S20 at high frequencies, which made FFTs of the impulse responses to have different characters across the devices. Furthermore, we found out that there are random delays in the timestamps related to sweep emission times in Note20. Therefore, it was hard for our algorithm to correctly and accurately determine the start and the end of each sweep which made the sweeps from the Note20 to have different MFCCs from the S20 sweeps. Since the consistency of both the frequency and time-domain features in Note20 have failed, our method could not accurately recognize the classes of the obstacles. As a part of our future work, we will design a method to dynamically detect sweeps recorded in time-series audio data.

6.2 Limitations

We employ the speakers of the smartphone to emit sine sweeps and sine waves. Since the sweeps have to be emitted frequently to recognize the obstacles, it affects the battery life of the smartphone. However, it is not necessary to emit sweeps that frequently if there is no obstacle is in close proximity. During our experiments, we found out that to accurately recognize stationary obstacles, the obstacle should be within 1 m from the smartphone. Therefore, we believe that we can increase the battery life of the smartphone by dynamically selecting how frequently the sweeps should be emitted. By emitting sweeps less frequently (2 sweeps per sec), we can probe for obstacles in close proximity. If a stationary obstacle is detected and it is within 1 m of the phone, the sweeps can be emitted more frequently (10 sweeps per sec) to recognize it.

7. Conclusion

We presented a preliminary investigation on a method for detecting and recognizing obstacles using smartphone active sound sensing for pedestrian safety. We discussed the theory of operation with which we can detect stationary and locomotive obstacles and recognize them. We also proposed two novel designs for probing signals and discussed a method to reduce the effect of the signal power leakage from frequency hopping. We evaluated our method on recognizing stationary obstacles and detecting and recognizing locomotive obstacles and achieved state-of-the-art performance. As a part of our future works, we plan to evaluate the sensing range of our method.

謝辞 This work is partially supported by JSPS KAKENHI Grant Number JP16H06539, JP17H04679, JP21H03428, and JP21K19769.

参考文献

[1] Nasar, J. L. and Troyer, D.: Pedestrian injuries due to mobile phone use in public places, *Accident Analysis & Prevention*, Vol. 57, pp. 91–95 (2013).

[2] Wang, T., Cardone, G., Corradi, A., Torresani, L. and Campbell, A. T.: Walksafe: a pedestrian safety app for mobile phone users who walk and talk while crossing roads, *Proceedings of the twelfth workshop on mobile computing systems & applications*, pp. 1–6 (2012).

[3] Wen, J., Cao, J. and Liu, X.: We help you watch your steps: Unobtrusive alertness system for pedestrian mobile phone users, *Proceedings of the 2015 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, IEEE, pp. 105–113 (2015).

[4] Li, S., Fan, X., Zhang, Y., Trappe, W., Lindqvist, J. and Howard, R. E.: Auto++ detecting cars using embedded microphones in real-time, *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 1, No. 3, pp. 1–20 (2017).

[5] Wang, Z., Tan, S., Zhang, L. and Yang, J.: ObstacleWatch: Acoustic-based obstacle collision detection for pedestrian using smartphone, *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 2, No. 4, pp. 1–22 (2018).

[6] Jain, S. and Gruteser, M.: Recognizing textures with mobile cameras for pedestrian safety applications, *IEEE Transactions on Mobile Computing*, Vol. 18, No. 8, pp. 1911–1923 (2018).

[7] Tange, Y., Konishi, T. and Katayama, H.: Development of vertical obstacle detection system for visually impaired individuals, *Proceedings of the 7th ACIS International Conference on Applied Computing and Information Technology*, pp. 1–6 (2019).

[8] Xie, Y., Li, F., Wu, Y., Yang, S. and Wang, Y.: HearSmoking: Smoking Detection in Driving Environment via Acoustic Sensing on Smartphones, *IEEE Transactions on Mobile Computing* (2021).

[9] Dissanayake, T., Maekawa, T., Amagata, D. and Hara, T.: Detecting Door Events Using a Smartphone via Active Sound Sensing, *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, Vol. 2, No. 4, pp. 1–26 (2018).