Roadway Water-splash Detection Method Using Acoustic Sensing

Taku Shiroto*, Remi Kanazawa*, Shigemi Ishida*, Yoh Shiraishi*

* Graduate School / School of Systems Information Science, Future University Hakodate, Japan
Email: {g2123030, g2122018, ish, siraisi}@fun.ac.jp

***Abstract*—The Japanese Road Traffic Law dictates that drivers must prevent making 'water splashes', to avoid pedestrians' clothing and belongings getting wet. However, drivers sometimes mistakenly make water splashes at night or in rain. It is effective to share information about water splashes between drivers to help drivers avoid water splashes easily. To collect the water-splash information, it is insufficient just to detect puddles or their causes, such as cracks, potholes, or ruts, because water splashes occurs not only on puddles but also on flat road surfaces where water flows. In this paper, we present a method to detect actual water splashes caused by vehicles. We use an acoustic sensing method that classifies sound recorded inside the vehicle into two classes, which are 'splashing' and 'non-splashing', with supervised learning. We achieved an F measure of approx. 90% and confirmed the effectiveness. Additionally, we confirmed that acoustic features focused on a low-frequency range are effective.**

***Index Terms*—water-splash detection, acoustic sensing, machine learning, microphone**

I. Introduction

Drivers have a responsibility to avoid making water splashes that can soak the clothes and belongings of pedestrians. Although some drivers recognize the road conditions and avoid puddles, they have still difficulties preventing water splashes.

To ensure that drivers avoid water splashes, it is effective to collect information about water splashes and share the information between drivers. Recent studies have proposed methods to detect puddles and their causes, such as cracks, potholes, and ruts where water can accumulate [1]–[4]. However, detecting puddles and their causes is insufficient to avoid water splashes because water splashes can also occur when a vehicle passes through an area where water is flowing. Small puddles might cause no water splashes, which results in false positive detection. Additionally, the accuracy of puddle detection methods with cameras could be adversely affected at night or in rain. To help drivers avoid water splashes, accurate water-splash detection is required.

In this paper, we present an acoustic sensing method to detect water splashes. Our method utilizes an embedded microphone installed in car equipment, such as a car navigation system or a dashcam, to detect the sound of water splashes. Acoustic features that reflect the sound of water splashes are extracted from the acoustic data. Using the features as training data, a supervised learning model is constructed to detect whether a water splash has occurred.

Specifically, our main contributions are as follows:

- We propose a water-splash detection method, which is important to enable drivers to fulfill their responsibility. To the best of our knowledge, water-splash detection is novel in the field of intelligent transportation systems.
- We show experimental evaluation to demonstrate that our detection method successfully detects water splashes with high accuracy of an F measure of approx. 90%.
- We show that the Mel-Frequency Cepstral Coefficient (MFCC) is an effective acoustic feature for water-splash detection. The frequency range below 6.3 kHz is particularly helpful in detecting water splashes.

The remainder of this paper is organized as follows. In Sec. II, we present the related works about puddle detection, road anomaly detection, and acoustic sensing. Sec. III analyzes the characteristics of water-splash sound collected in an actual environment. Sec. IV presents our water-splash detection method including acoustic features, followed by experimental evaluations and discussions in Sec. V. Finally, conclusions are presented in Sec. VI.

II. Related Works

Water splashes mainly occur in puddles. Some methods to detect puddles and their causes such as cracks, potholes, and ruts, have been reported. X. Han et al. and Kim et al. proposed image-based puddle detectors [1], [2]. They used the fact that the surfaces of puddles on an image reflects the surrounding scenery. Basavaraju et al. detected cracks and potholes on a road surface [3]. J. Han et al. estimated the depth of ruts on a road surface [4].

In the field of acoustic sensing, studies have been conducted on road condition classification. Bahrami et al. classified dry and wet road surfaces using two streams of Convolutional Neural Networks (CNNs), with an accuracy of 92.3% [5]. Abdić et al. also classified dry and wet road surfaces using supervised learning, and achieved an accuracy of 93.2% [6]. They collected acoustic data with a microphone installed on the outside of a vehicle, near the tires. They also extracted features including those based on the Mel scale, which is a feature of human hearing.

There are various acoustic sources outside the vehicle, such as the sound of other vehicles running and rainfall. The method of installing microphones outside the vehicle, as in the study by Bahrami et al. may adversely affect the accuracy of the machine learning model due to the inclusion of these noises.

978-4-907626-52-5/23/$31.00 ©2023 IPSJ

TABLE I
DATA COLLECTION ENVIRONMENT

Arguments	Environment
Wheather	Rain
Time of Day	Night
Road Condition	Wet, Asphalt surface
Recording Device	ZOOM Q2N-4K
Microphone	audio technica AT9944
Sampling Rate	44.1 kHz
Bit Length	16 bit

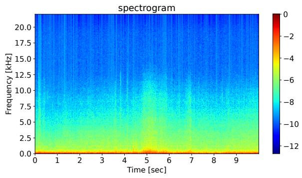


Fig. 1. Spectrogram of water-splashing sound

III. PRELIMINARY ANALYSIS

In this section, we analyze acoustic data collected from an actual vehicle during the occurrence of water splashes. The aim of this analysis is to clarify the characteristics of the water splashes in both frequency and time domains.

A. Data Collection Environment

Table I shows the data collection environment. A single microphone was installed above the rear seat of a vehicle. The collected recordings include 48 water splashes.

B. Calculation of Spectral Envelope Differences

To determine the acoustic characteristics in the frequency domain, we compared the spectral envelope of vehicle running sound with and without water splashes. We calculated the average difference in the spectral envelope in the following steps.

1) The recordings were divided into 10-second windows with 50% overlap. Then, windows including water-splash sound are used, discarding those with no water-splash sound. The windows including the water-splash sound consist of both parts in which the water-splash sound occurs and does not occur.
2) For the 10-second windows extracted in the first step, we extracted the first 0.37-second segments for each of the splashing parts and non-splashing parts. The 0.37-second segment corresponds to $16384 (= 2^{14})$ samples at 44.1-kHz sampling, which is determined based on our observation that the length of water-splash sounds is longer than approximately 0.5 seconds.
3) We performed a fast Fourier transform for each segment and calculated the sound log power, i.e., the logarithm of the square of sound intensity, in the segment as a function of frequency, deriving a spectral envelope for each segment.
4) For spectral envelopes of splashing and non-splashing segments in the same window, we calculated the difference and averaged over all the windows.

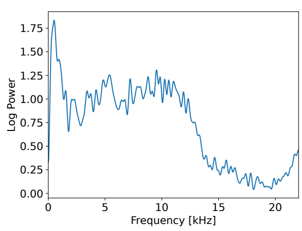


Fig. 2. Average difference in spectral envelopes of splashing and non-splashing sounds

C. Data Analysis Result

Figure 1 shows an example of the spectrogram of a 10-second window including a water-splash sound. In Fig. 1, the water splash occurs from 4.8 to 5.8 seconds. We can confirm that the main frequency components of the water-splash sound are below 12.5 kHz.

Figure 2 shows the average difference in the spectral envelope. In Fig. 2, we can see a large difference between splashing and non-splashing sounds at frequencies below 12.5 kHz. The difference is particularly large at frequencies below 2.5 kHz. From this analysis, we can confirm that the acoustic characteristics of water splashes appear in the lower frequency range of the human audible range, which is up to around 20 kHz.

IV. WATER-SPLASH DETECTION METHOD

A. Overview of Our Method

The proposed method consists of a learning pahse and detection phase. In the learning phase, acoustic data from each vehicle is collected along with the driving video. The acoustic data is divided into splashing and non-splashing parts based on the video images. The procedure of the data segmentation is described below:

1) The recordings are divided into 10-second windows including water-splash sounds, by the same process in the step 1) in Sec. III-B.
2) For each of the remaining windows, we extract segments, using a 0.37-second window with 50% overlap, from each of the splashing and non-splashing parts.
3) We extract acoustic features from each segment and derive sets of acoustic feature vectors for each of the splashing and non-splashing sounds.

Finally, the sets of feature vectors are used as training data to construct a machine learning model that detects whether a water splash has occurred. In the detection phase, the machine learning model detects water splashes with acoustic features extracted from the acoustic data newly acquired during driving.

In this paper, we define a splashing part as the moment when water droplets are observed on the window of a car in the driving video. Figure 3 shows the camera view from the inside of a vehicle while passing through splashing/non-splashing parts.

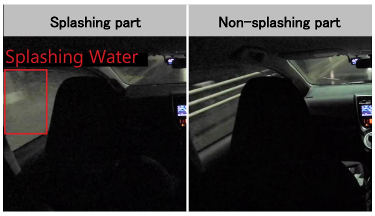


Fig. 3. Camera view from inside of vehicle while passing through splashing and non-splashing parts

B. Acoustic Features

In this paper, we use three types of acoustic features and compare the accuracy of water-splash detection for each. The details of each acoustic feature are described below. Each type of feature is created by multiplying the acoustic spectrum by the corresponding filter bank shown in Fig. 4 and performing a discrete cosine transform.

1) MFCC
2) Linear-Frequency Cepstral Coefficient (LFCC)
3) Splash-Frequency Cepstral Coefficient (SFCC)

1) MFCC is an acoustic feature that expresses high resolution in the low-frequency range and low resolution in the high-frequency range, mimicking the Mel scale. MFCC or Mel scale-based features have been used in related works [5], [6].

2) LFCC is an acoustic feature with uniform resolution over the entire frequency range. LFCC is used as a comparison target for MFCC and SFCC.

3) SFCC is an acoustic feature with varying frequency resolution, based on the analysis in Sec. III. The splash filter bank has a high resolution in the region below 2.5 kHz, a medium resolution in the frequency range between 2.5 kHz and 12.5 kHz, and low resolution above 12.5 kHz. SFCC is used to evaluate the acoustic features created through the data analysis.

V. Evaluation

This section reports the results of evaluating the machine learning models described in Sec. IV. The evaluation aims to clarify the effectiveness of the proposed water-splash detection method, the acoustic features effective for water-splash detection, and the frequency range effective for water-splash detection.

A. Conditions

The evaluation experiment was conducted by driving on public roads in Hakodate city, in two separate trips. The details of the data collection environment are shown in Table I. The collected acoustic data were segmented according to the method described in Sec. IV-A. Correct labels were assigned to splashing and non-splashing parts respectively. The window size was set to 0.37 seconds with 50% overlap. In order to prevent duplication and deterioration in the quality of acoustic features, acoustic data smaller than the window size were excluded at the end of each splashing/non-splashing part. 19 dimensional acoustic features were extracted for each window for each acoustic feature. The acoustic amplitude values were normalized in accordance with EBU-R128. We constructed the machine learning models with Support Vector Machine (SVM).

TABLE II
THE RESULTS OF ACCURACY EVALUATION (MFCC, LFCC AND SFCC)

Feature Type	F measure
MFCC	90.90%
LFCC	88.16%
SFCC	89.61%

B. Evaluation Method

1) Adjustment of the volume of acoustic data: The machine learning models were created and evaluated using acoustic data from 70 water-splash events. A total of 596 acoustic features were used as training data, 298 each for splashing/non-splashing. With the data segmentation described in Sec. IV-A, more acoustic features were extracted from the non-splashing part than from the splashing part. To prevent splashing/non-splashing bias in the machine learning model, the same number of data as that in the splashing part was randomly selected from the acoustic features in the non-splashing part.

2) Separation of training data and validation data: In this paper, the 70 water-splash events were divided into 5 groups of 14 water-splash events each, and a 5-fold cross-validation was performed with 1 group as the validation data and the remaining 4 groups as the training data. If the acoustic features are divided without considering which water-splash event they are extracted from, the accuracy of the machine learning model may be higher than in actuality because similar acoustic features are included in both the training and the validation data.

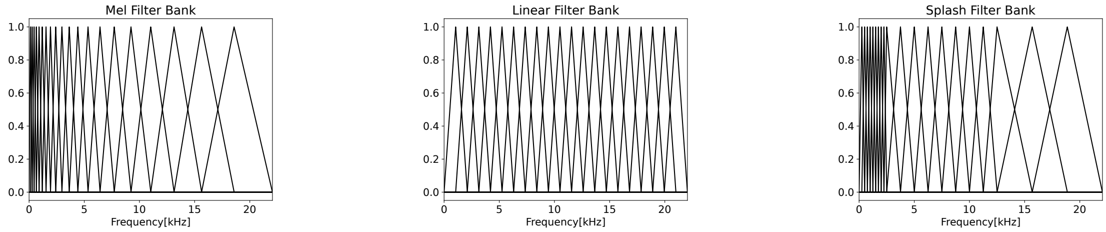
C. Evaluation Results and Analysis

In this section, we present the results of the evaluation of our method. We also discuss the effectiveness of acoustic sensing, effective acoustic feature types, and the frequency range effective for water-splash detection. Table II shows the average F measure obtained from 5-fold cross-validation for the machine learning models created for each acoustic feature.

As a result, it was confirmed that water splashes were detected with a high accuracy of approx. 90% for all features.

1) Analysis of Acoustic Features: The accuracy exhibited by the machine learning models was highest for MFCC, followed by SFCC, then LFCC. Acoustic features focused on the low-frequency region, such as MFCC and SFCC, showed higher accuracy than LFCC. This fact is consistent with the characteristic noted in Sec. III, i.e., the difference between splashing and non-splashing parts tends to appear in the low-frequency region. Therefore, it is suggested that acoustic features focused on the low-frequency region are effective for water-splash detection.

However, the accuracy of SFCC, which reflects the acoustic characteristics of water splashes in its resolution, was lower than that of MFCC. This suggests that we need to improve the



(a) Mel filter bank (MFCC) (b) Linear filter bank (LFCC) (c) Splash filter bank (SFCC)

Fig. 4. Filter banks used in acoustic feature extraction

TABLE III
TOP 4 FREQUENCY RANGES MOST STRONGLY CORRELATED WITH SPLASHING

Rank	Frequency Range	Correlation
1	4,200Hz-6,300Hz	0.35
2	1,050Hz-3,150Hz	0.33
3	11,500Hz-13,650Hz	0.30
4	10,500Hz-12,600Hz	0.28

criteria to determine the resolutions. In this paper, the criteria to determine the resolutions of SFCC were qualitative. To address this issue, it could be effective to reflect the correlation between frequency ranges and the correct answer data of splashing/non-splashing in the resolution for each frequency.

2) Analysis of Effective Frequency Range: The correlation coefficients between the correct labels for splashing/non-splashing and each dimension of the features were calculated for LFCC, where each dimension of the features has the same bandwidth of interest. Table III shows the top four frequency ranges that are highly correlated with splashing/non-splashing. The results indicate that the frequency range below 6.3 kHz, where the MFCC resolution is high and the SFCC resolution is medium, has a high correlation with the correct labels of splashing/non-splashing. This suggests that this frequency range is effective for the detection of water splashes.

On the other hand, the correlation coefficient with the correct labels of splashing/non-splashing was high in the mid-frequency region around 10 kHz, which was not focused on in MFCC and SFCC. These results suggest that in addition to the low-frequency region focused on by MFCC and SFCC, focusing on the mid-frequency region around 10 kHz could result in even higher water-splash detection accuracy.

3) Review of Analysis: The accuracy of our method reached an F measure of approx. 90%, indicating the effectiveness of our acoustic sensing method for water-splash detection. The results of Sec. V-C1 also show that the acoustic features focused on low frequencies, such as MFCC, are more effective for water-splash detection. Additionally, the frequency range below 6.3 kHz has a high correlation between the correct labels of splashing/non-splashing, indicating that the frequency range is effective for water-splash detection.

VI. CONCLUSION

The purpose of this study is to establish a method for detecting water splashes from vehicles, in order to prevent damage to pedestrians caused by water splashes. Toward this end, we proposed a method using acoustic sensing. We confirmed that our method is effective in detecting water splashes with an F measure of approx. 90%. Additionally, we confirmed that acoustic features focused on low-frequency regions, such as MFCC, are more effective. Particularly, we confirmed that the frequency range below 6.3 kHz has a high correlation with the correct answer of splashing/non-splashing.

Future work will include the development of acoustic features based on correlation coefficients between the correct labels of splashing/non-splashing and each frequency. We will also continue to collect acoustic data and construct machine learning models that are robust to various road conditions, weather, and internal and external noises.

ACKNOWLEDGEMENT

Part of the research presented in this paper was supported in part by JSPS Grant-in-Aid for Scientific Research JP20K11793.

REFERENCES

[1] X. Han, C. Nguyen, S. You, and J. Lu, “Single image water hazard detection using FCN with reflection attention units,” in *European Conference on Computer Vision (ECCV) 2018*, 2018, pp. 105–121.
[2] J. Kim, J. Baek, H. Choi, and E. Kim, “Wet area and puddle detection for advanced driver assistance systems (ADAS) using a stereo camera,” *International Journal of Control, Automation and Systems*, vol. 14, no. 1, pp. 263–271, 2016.
[3] A. Basavaraju, J. Du, F. Zhou, and J. Ji, “A machine learning approach to road surface anomaly assessment using smartphone sensors,” *IEEE Sensors Journal*, vol. 20, no. 5, pp. 2635–2647, 2020.
[4] J. Han, L. Cui, and S. Shi, “Road rut detection system with embedded multi-channel laser sensor,” in *The International Journal of Advanced Manufacturing Technology*, 2021, pp. 1–10.
[5] S. Bahrami, S. Doraisamy, A. Azman, N. A. Nasharuddin, and S. Yue, “Acoustic feature analysis for wet and dry road surface classification using two-stream cnn,” in *CSAI 2020: 2020 4th International Conference on Computer Science and Artificial Intelligence*, 2021, pp. 194–200.
[6] I. Abdić, L. Fridman, D. E. Brown, W. Angell, B. Reimer, E. Marchi, and B. Schuller, “Detecting road surface wetness from audio: A deep learning approach,” in *International Conference on Pattern Recognition (ICPR)*, 2016, pp. 3458–3463.