Initial Evaluation of Acoustic Train Detection System

Koji Sato[1], Shigemi Ishida[1], Jumpei Kajimura[1], Masato Uchino[1],
Shigeaki Tagashira[2], and Akira Fukuda[1]

[1] ISEE, Kyushu University Japan,
{k_sato,ishida,kajimura,uchino,fukuda}@f.ait.kyushu-u.ac.jp
[2] Faculty of Informatics, Kansai University, Japan,
shige@res.kutc.kansai-u.ac.jp

Abstract. Accidents between trains and vehicles at railroad crossing is one of the most common train accidents in Japan. To prevent railroad crossing accidents, we need to detect both trains and vehicles. In this paper, we propose an acoustic train and vehicle detection system that shares a microphone array to detect trains and vehicles. In our previous work, we have developed a vehicle detection system using stereo microphones. Here we present a train detection system using a microphone. The train detection system analyzes frequency components of the sound signals acquired by a railside microphone. We calculate probability of train passing using logistic regression model and apply a hysteresis thresholding with two thresholds to detect train passing. Simple filtering based on train length is also applied to increase robustness to noise including vehicle passing sounds. We conducted initial evaluations and confirmed that our train detection system successfully detected trains with an F-measure of 0.987 and Recall of 1.0.

Keywords: Train detection, Acoustic sensor, Logistic regression.

1 Introduction

Train is fast and large-capacity transportation, which plays an important role in modern society. Railroad accident cause not only human damages but also social problems such as train delay. It is essential to prevent railroad accidents. An accident between a train and vehicle at railroad crossing is one of the most common train accidents in Japan. There were a total of 242 railroad crossing accidents in Japan in 2015 [1].

In order to prevent the railroad crossing accidents, we need to detect both trains and vehicles. For train detection, we can categorize the train detection system into two types: on-board installation and railway installation types. Typical examples of the on-board installation type are based on GPS (Global Positioning System) or tachograph that records moving distance from a reference point. Typical example of the railway installation type is based on a track circuit, which

is used for railway signaling control. For vehicle detection, ultrasonic sensors, infrared sensors, and loop coils have widely installed. These sensors might be used on a railroad crossing.

These detection systems, however, are separately installed, which implies that we need multiple systems to simultaneously detect trains and vehicles on a railroad crossing. We need a new low-cost approach that detects both trains and vehicles using a single sensor to apply the detection system in rural areas.

We are developing a train and vehicle acoustic detection system that shares a microphone array to detect both trains and vehicles. The system analyzes sound signals of both trains and vehicles derived by a microphone array near a railroad crossing to detect trains and vehicles. Microphone is a cost-effective device, which places few physical restrictions on installation location of the detection system because sound signals are diffracted over obstacles.

In this paper, we present a train detection system using a microphone because we have developed a vehicle detection system using stereo microphone in our previous work [3, 2]. Our train detection system analyzes frequency components of the sound signals acquired by a railside microphone. We calculate probability of train passing using a logistic regression model on the frequency components. Regression coefficients are trained prior to system use. Finally, the system applies hysteresis thresholding with two thresholds to detect train passing.

To demonstrate the basic performance of our train detection system, we conducted experimental evaluation. We installed a microphone at a house nearby railway and recorded train passing sound for approximately 7.5 hours. The sound signals were then analyzed using our train detection system to detect passing trains. We confirmed that the train detection system successfully detected trains with an F-measure of 0.99 and a recall of 1.0.

The rest of this paper is organized as follows. Section 2 describes existing methods of train detection with their problems. Section 3 briefly introduces train and vehicle detection system and Section 4 describes our proposed train detection system. In Section 5, we conduct experimental evaluations to demonstrate the basic performance of our system, and Section 6 concludes the paper.

2 Related Works

To the best of our knowledge, train and vehicle simultaneous detection system is novel in both fields of train and vehicle detections. In this section, we briefly look through related works on train detection and vehicle detection.

2.1 Train detection

Train detection system is categorized into two types: on-board and railway installation types. Both types suffer from high installation and operation costs.

The methods using a tachograph and satellite positioning system are categorized an on-board installation type. Distance from a reference point is calculated using tachograph output. The reference points are measured using IC

tags installed along the railway [5]. Satellite positioning systems, including GPS (Global Positioning System) and GNSS (Global Navigation Satellite System), provide current train location calculated from radio signals from satellites [7, 4]. The on-board detection systems suffer from a communication problem: reliable communication between train and ground is mandatory. The satellite-based on-board detection systems are affected by wireless communication: performance degrades in a tunnel, for example, because we receive no satellite signals in a tunnel.

The train detection methods using track circuits or axle counters are categorized into a railway installation type. A track circuit is an electrical circuit using wheels and axle of a train, and tracks separated by insulators forming track sections. When a train passes through a section, the circuit is electrically shorted out, which indicates the presence of passing train in the section. An axle counter detects train passing when a train passes over an axle counter installed on a railway. The railway installation type requires high cost for installation because of railroad work, which sometimes disturbs train service.

2.2 Vehicle detection

Existing roadway vehicle detection systems are based on an infrared sensor, laser radar, loop coil, and stereo camera.

Vehicle detection system using an infrared sensor at a railroad crossing installs infrared transmitter and receiver across a railway to avoid device installation in a crossing. The receiver output explicitly indicates passing trains as trains shut out the infrared light. The infrared-based approach highly affected by weather condition and dust, degrading detection accuracy or resulting in failures.

Loop coils utilize magnetic field change caused by vehicle passing to detect vehicles. Loop coils come with high installation cost due to roadwork to embed loop coils under a road surface. In addition, loop coils have small sensor coverage, so that fails in accurate detection of small vehicles.

A laser-radar-based system transmits laser light into a railroad crossing and analyze the reflected light to detect vehicles. To detect vehicles in multiple lanes, we need multiple laser radar sensors installed above a railroad crossing, which results in high installation cost for safety installation as to prevent falling.

A stereo-camera-based system analyzes stereo camera image to detect obstacle in a railroad crossing [6]. However, installation is restricted due to power supply problem, i.e., two high-power cameras consumes much power as they are always turned on. In addition, detection accuracy degrades for high-speed vehicles at night.

3 Train and Vehicle Detection System

Figure 1 illustrates an overview of train and vehicle detection system. The detection system consists of two microphones followed by a low pass filter (LPF), vehicle detection system, and train detection system. Two microphones are installed
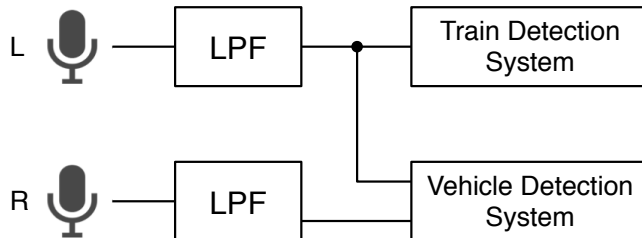
Fig. 1. Overview of train and vehicle detection system

beside a railway crossing and collect audio data of train and vehicle passing. After applying a LPF to reduce high frequency environmental noise, the vehicle and train detection systems share and analyze the audio data to detect vehicles and trains, respectively. The vehicle detection system uses two microphones and utilize time difference of arrival of vehicle sound to estimate where vehicle is running. Train detection system uses one microphone and analyzes frequency components of the sound to detect train passing.

We have developed an acoustic vehicle detection system in our previous work [3, 2]. The system draws a "sound map", which is a map of time difference of sound arrival on two microphones. The time difference is calculated from a cross correlation function of the sound data acquired on the two microphones. Although the system is not tested at a railroad crossing, we confirmed that the system detected vehicles with an F-measure of 0.92. We also confirmed that the types of vehicles have small influence on detection accuracy because the detection is based on vehicle running sound.

We present a train detection system in the following section.

4   Train Detection System

Figure 2 illustrates an overview of the train detection system. The train detection system consists of predict and detect blocks to analyze the sound signals acquired from a microphone installed by a railway crossing. The LPF (low pass filter) is applied prior to the analysis to reduce the influence of high frequency environmental noise. The predict block calculates probability of train existence based on frequency components of sound signals acquired from the microphone. Then the thresholding is applied on train existence probability to predict the presence of a passing train. The detect block calculates MA (moving average) over the output of predict block and applies hysteresis thresholding with two thresholds to detect train passing.

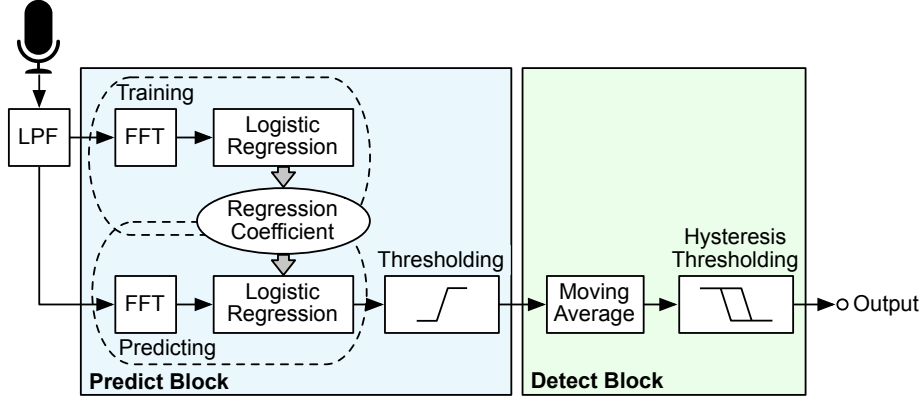Details of each block is described in the following subsections.

Fig. 2. Overview of train detection system

4.1 Predict block

Predict block consists of training and predicting phases because the block uses logistic regression, which is one of machine learning methods.

In a training phase, regression coefficients of logistic regression are trained using frequency components of sound signal acquired in advance. In a predicting phase, logistic regression analysis is performed on the frequency components of sound signal at each time point using the regression coefficients obtained in the training phase to calculate the existence probability of passing trains. Thresholding is performed to the existence probability to finalize the existence of passing trains.

Each phase is described in detail below.

Training phase In a training phase, the system trains regression coefficients of logistic regression using the training audio data. The ground truth of train passing is manually derived from video.

As shown in Fig. 2, we use frequency components calculated from fast Fourier transform (FFT) as feature values. The sound data is divided into fixed time-width data. FFT is applied to the each divided data to calculate amplitude of frequency components. Using amplitude of the frequency components, the system trains the regression coefficients of logistic regression.

Figure 3 shows an example of the frequency components. As shown in Fig. 3, the frequency components of the sound of passing train is concentrated on less than 1000 Hz. Therefore, we use frequency components less than 1000 Hz as training features.

In a logistic regression analysis for train detection, the system calculates probability of train existence using the frequency components derived from FFT as feature value. Probability of train existence is given by

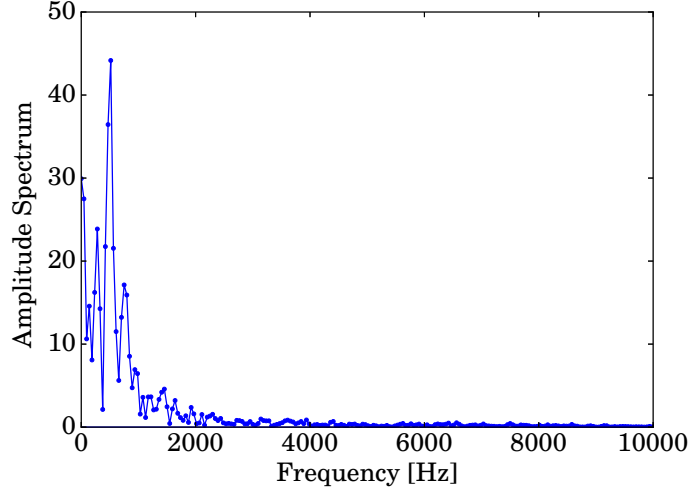$$P(Y = 1|\mathbf{X}) = \frac{1}{1 + e^{-\mathbf{A}\mathbf{X}}}, \tag{1}$$

Fig. 3. Frequency components of train passing sound

where $\mathbf{X} = {}^{t}[1, x_1, x_2, \ldots, x_n]$ is an input vector and $\mathbf{A} = [a_0, a_1, a_2, \ldots, a_n]$ is a regression coefficient vector.

In a training process, regression coefficients are calculated by minimizing a cost function $C(\mathbf{A})$:

$$C(\mathbf{A}) = \frac{1}{N} \sum_{i=1}^{N} \log P(Y = Y_i | \mathbf{X}_i), \qquad (2)$$

where $\{X_i, Y_i | i = 1, 2, \ldots, N\}$ is a training data set derived from the FFT.

Predicting phase In a predicting phase, the system calculates the existence probability of passing train using regression coefficients obtained in the training phase.

Logistic regression analysis is performed using frequency components calculated from FFT of sound data. The sound data is divided into fixed time-width data. FFT is again applied to the each divided data to calculate amplitude of the frequency component $\mathbf{X}$. Then, the probability of train passing on each time is calculated substituting the regression coefficients $\mathbf{A}$ obtained in the training phase and the frequency component $\mathbf{X}$ for regression model Eq. (1). Passing train is detected if the probability is higher than a threshold. The threshold value is determined in a preliminary experiment, as described in Section 5.2.

Figure 4 shows an example of the result of thresholding. A train passed in front of a microphone between 22 and 32 seconds. As shown in Fig. 4, the system successfully detected a passing train.
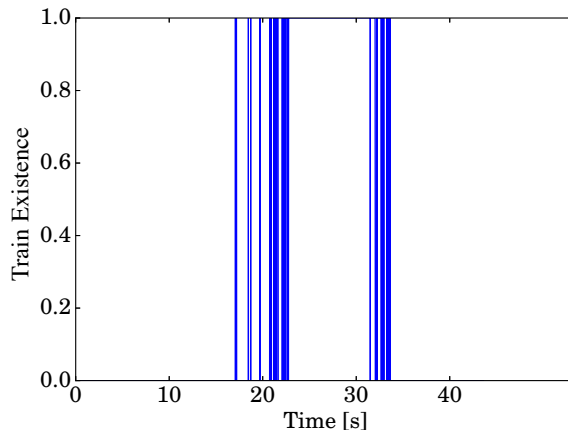
Fig. 4. Output example of predict block. `0` and `1` indicate that train exists and no train exists, respectively.

4.2 Detect block

As shown in Fig. 4, result of logistic regression analysis chatters when train is approaching or going away. Passing of large vehicle such as a truck sometimes leads faulty detection.

To reduce such faulty detections, moving average is applied to the output of predict block prior to apply thresholding. The length of the moving average is set to approximately 5 seconds based on the time length of train passing. Hysteresis thresholding is then applied to the output of moving average to finalize the existence of passing trains. Figure 5 briefly explains a train detection algorithm. A The blue curve in Fig. 5 represents the averaged probability of train passing. To detect a passing train, the system detects the train head and tail by hysteresis thresholding. When the averaged probability exceeds a higher threshold, the system detects the head of train. And when the averaged probability falls below a lower threshold, the system detects the tail of train, which implies that a train have passed.

5 Initial Evaluation

To demonstrate the basic performance of our train detection system, we conducted initial experiment.

5.1 Experiment setup

Figure 6 shows the experiment setup. A microphone was installed in a backyard of a house near a railway in Itoshima city, Fukuoka, Japan. We collected audio
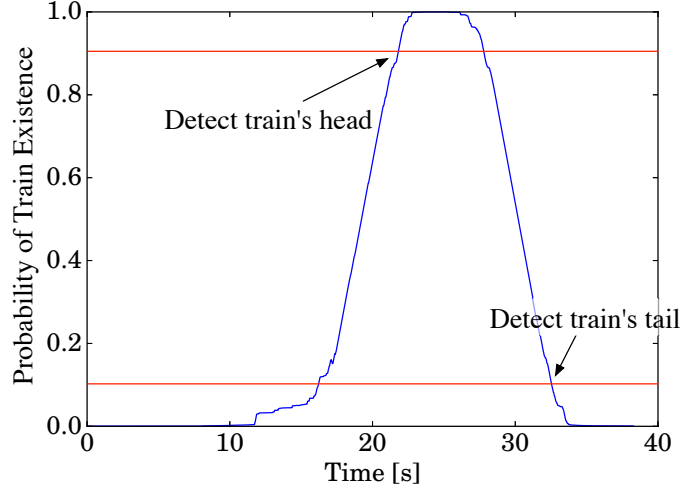
Fig. 5. Overview of hysteresis thresholding

data for approximately 7.5 hours. We also collected audio data at a different location for approximately 3 hours, which was used only for training to evaluate the influence of difference between locations of training and testing. A total of 39 trains were passed in the 7.5-hour testing data, while a total of 17 trains were passed in the 3-hour training data. Although two microphones were installed as shown in Fig. 6, we only used one microphone in this paper; the other microphone would be used in our future work.

The target railway has a single track, implying multiple trains never pass at the same time. The sound was recorded using a SONY HDR-MV1 recorder with an AZDEN SGM-990 microphone at a sampling frequency of 48 kHz and word length of 16 bits. Video monitoring the target railway was also recorded using a SONY HDR-MV1 video recorder, which was used as ground truth.

We manually labeled training data referring to ground truth, i.e., recorded video: `1` for train passing and `0` for no train passing. The label `1` is used only when a train was passing right in front of the microphone. The audio data when train was approaching and was going away was excluded for evaluation to improve training accuracy because these cases might include much noise other than train sound.

We define train passing time t_p as the time when the train is passing right in front of the microphone. Train passing sound is captured from $t_p = 0$ to $t_p = 5$ [seconds]. In a training phase, audio data from $t_p = -5$ [seconds] to $t_p = 0$ and from $t_p = 5$ [seconds] to $t_p = 20$ [seconds] was excluded from training data because the data includes ambiguous sound signals.
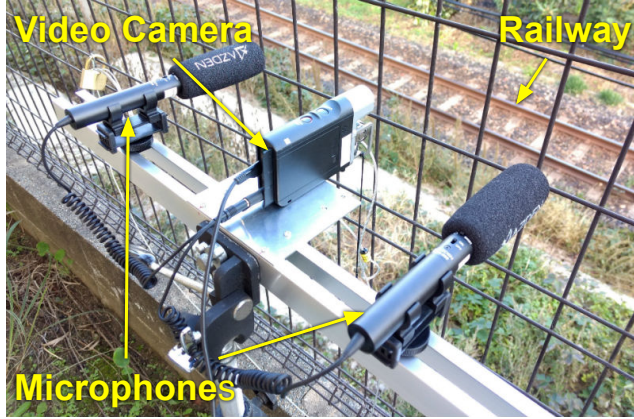
Fig. 6. Experiment setup

We extracted frequency components less than 1,000 Hz that are passed to a logistic regression module. Fast Fourier transform (FFT) was performed on 1024 samples at a sampling rate of 48 kHz, we picked 21 points from the FFT results.

Comparing the results derived by our train detection system with video, we evaluated the number of true positives (TPs), false negatives (FNs), and false negatives (FNs). TP, FN, and FP are defined as the case that a train was detected when a train passed, no train was detected when a train passed, and a train was detected when no train passed, respectively. We excluded true negatives (TNs), which is defined as the case that no train was detected when no train passed, because TNs were not countable in our experiment.

Using the numbers of TPs, FNs, and FPs, we also evaluated a precision, a recall, and F-measure defined as:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{3}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{4}$$

$$\text{F}_{\text{measure}} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{5}$$

5.2 Preliminary experiment

To determine a threshold for thresholding of logistic regression described in Section 4.1, preliminary experiment was conducted. The threshold was determined using a ROC (Receiver Operating Characteristic) curve, which is a plot of the true positive (TP) rate against the false positive (FP) rate at various threshold. The point that has the shortest distance from upper left corner (0,1) on the ROC curve gives an optimal threshold.
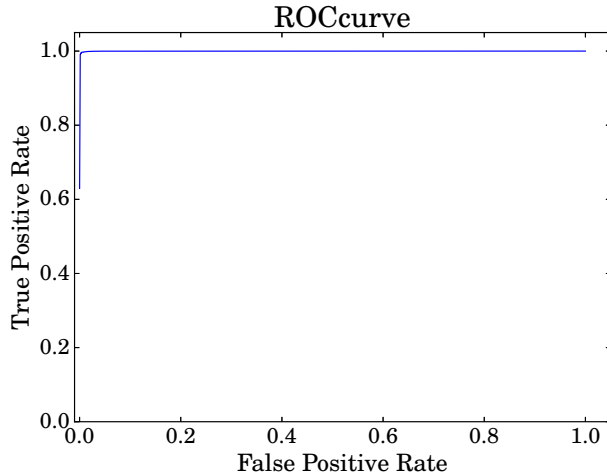
ROCcurve

True Positive Rate / False Positive Rate

Fig. 7. ROC curve with thresholds changed from 0 to 1 in steps 0.01

Table 1. Experiment results

TPs	FNs	FPs
39	0	1
Precision		0.98
Recall		1.0
F-measure		0.99

We drew a ROC curve using the audio data derived for testing. We performed a 10-fold cross-validation to calculate the TP and FP rates with various threshold. The sizes of train passing sound data and no train passing sound data were equalized in testing not to include the influence of training data size.

Figure 7 shows a ROC curve with thresholds changed from 0 to 1 in steps of 0.01. A threshold corresponding to the point that has the shortest distance from (0,1) was 0.37. We therefore used a threshold of 0.37 in the rest of evaluations.

5.3 Detection accuracy

We evaluated detection accuracy of our train detection system. Logistic regression coefficients were derived using 3-hour training sound data and 7.5-hour sound data was used for evaluation to include the influence of recorded location difference between training and testing. The threshold value of logistic regression was set to 0.37, as described in Section 5.2.

Table 1 shows evaluation results, i.e., the numbers of true positives (TPs), false positives (FPs), false negatives (FNs), precision, recall, and F-measure. Table 1 indicates the following:
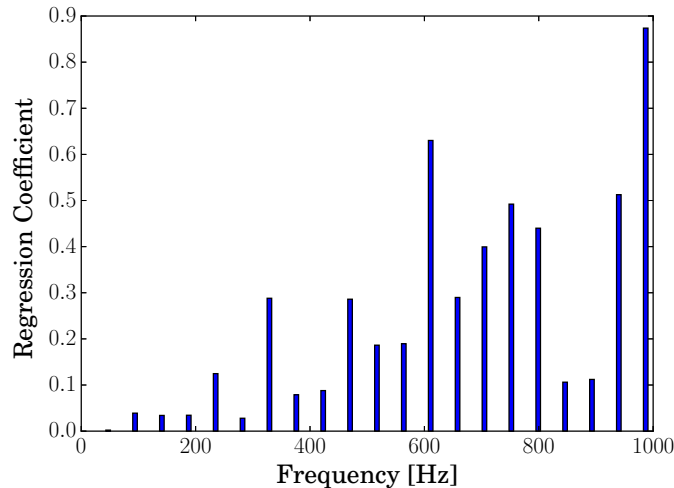
Fig. 8. Regression coefficients of logistic regression

1. Recall of 1.0 indicates that the train detection system detected all passing trains without any missed trains. Main purpose of our train detection system is to prevent accidents. No FN detection is a quite important feature because FN detection may cause a serious accident.

2. Precision of 0.98 indicates that the train detection system suffered from the small number of false positive detections. FP occurred when three motorbikes were successively passing near a microphone. Loud sound signals that partially include frequency components of train sound caused the FP detection. We believe that this type of noise could be excluded when we use sufficient amount of noise data for training.

3. F-measure of 0.99 indicates that the train detection system exhibited extremely high detection performance. One cause of this high performance might be the experiment environment. There was a single railway track in front of a microphone so that multiple trains never simultaneously passed. We need to perform an extended experiment to confirm the detection performance in a city area scenario.

The above result reveals that our acoustic train detection system successfully and effectively detected trains.

Figure 8 shows absolute values of regression coefficients used in logistic regression corresponding to each frequency component derived by FFT. The absolute value of regression coefficient implies the degree of each frequency component contributing to train detection. We can confirm that higher frequency signals tend to have higher regression coefficients; high frequency signals were dominant

for train detection. Logistic regression successfully extracted train sound using the regression coefficients, which provides train detection robust to noise.

6 Conclusion

In this paper, we propose an acoustic train and vehicle detection system. In our previous work, we developed a vehicle detection system using two microphones. We therefore focus on an acoustic train detection system in this paper. In our train detection system, frequency components of train sound are analyzed using a logistic regression model to calculate the probability of train passing. We apply a threshold to the probability to detect passing trains. We conducted experiment evaluations to demonstrate the detection performance of our train detection system. Experimental evaluations revealed that our train detection system successfully detected trains with an F-measure of 0.99 and recall of 1.0.

Acknowledgment

References

1. Cabinet Office, Government of Japan: White paper on traffic safety in japan 2016. available online (Jul 2016), `http://www8.cao.go.jp/koutu/taisaku/kou-wp.html`
2. Ishida, S., Liu, S., Mimura, K., Tagashira, S., Fukuda, A.: Design of acoustic vehicle count system using DTW. In: Proc. ITS World Congress. pp. 1–10. AP-TP0678 (Oct 2016)
3. Ishida, S., Mimura, K., Liu, S., Tagashira, S., Fukuda, A.: Design of simple vehicle counter using sidewalk microphones. In: Proc. ITS EU Congress. pp. 1–10. EU-TP0042 (Jun 2016)
4. Lu, D., Schnieder, E.: Performance evaluation of GNSS for train localization. IEEE Trans. Intell. Transp. Syst. 16(2), 1054–1059 (Apr 2015)
5. Matsumoto, M.: ATACS delivers moving block. Railway Gazette International 167(3), 43–45 (Mar 2011)
6. Ohta, M.: Level crossings obstacle detection system using stereo cameras. Quarterly Report of RTRI 46(2), 110–117 (Jun 2005)
7. Yordanov, R., Iontchev, E., Miletiev, R., Mladenov, V.: GPS/GPRS/INS system for real-time monitoring of the urban railway. Int. J. Transp. Syst. 1, 48–52 (Aug 2016)