Design of Ultra Low Power Vehicle Detector utilizing Discrete Wavelet Transform

Kazuo Kubo[1], Chengyu Li[1], Shigemi Ishida[1],
Shigeaki Tagashira[2], and Akira Fukuda[1]

[1] ISEE, Kyushu University Japan,
{kubo,licy0012,ishida,fukuda}@f.ait.kyushu-u.ac.jp
[2] Faculty of Informatics, Kansai University, Japan,
shige@res.kutc.kansai-u.ac.jp

Abstract. Vehicle traffic is important information in the intelligent transport system (ITS). We have developed an acoustic vehicle detection system that relies on two microphones at a sidewalk. The system has already been confirmed that it successfully detected the vehicle and direction of travel with an F-measure of 0.92. However, the power consumption of the system is high, which puts power restrictions on deployment. We therefore propose a low power vehicle detection system utilizing a wake-up mechanism. The acoustic vehicle detection system, presented in our previous work, is activated when a newly developed ultra low power vehicle detector (ULP-VD) detects vehicles. Initial experimental evaluations reveal that the ULP-VD successfully detected vehicles with a precision of 0.94 and recall of 0.95.

Keywords: Vehicle detection, discrete Wavelet transform, logistic regression.

1 Introduction

Vehicle traffic is important information in the intelligent transport system (ITS). Vehicles on a road are generally detected using a vehicle detection system. The deployment of the vehicle detection system is currently limited to high traffic roads because of its high installation and maintenance costs, which restricts ITS applications.

We are developing an acoustic vehicle detection system using microphones installed at a sidewalk. Our system detects vehicles using the sound generated from tires during vehicle running. Because sound waves are diffracted over obstacles, we can deploy microphones in a low height configuration, which drastically reduces roadwork cost in terms of road closure as well as safety installation.

To realize the vehicle detection system, we have developed an acoustic vehicle detection method using a sound map. The sound map is a map of sound arrival time difference on two microphones. Our vehicle detection method analyzes the sound map using a state-machine-based algorithm [9] or template matching [8] to detect passing vehicles. We use deterministic approaches to detect vehicles

© 2018 Shigemi Ishida. This is an accepted version of the paper. Personal use of this material is permitted. Not for redistribution.

that require no training data prior to system use. Experimental evaluations revealed that our vehicle detection method successfully detected vehicles with an F-measure of 0.92.

The acoustic vehicle detection system, however, consumes much power requiring a power cable or a big battery for 24/7 operation. The system relies on a sound map that is generated with much computation including image processing and convolutional integral, putting difficulties on power reduction. Power saving is one of the most important tasks to apply the vehicle detection system to an environment with limited power resources.

We therefore propose a low power vehicle detection system, which is a combination of a high performance vehicle detector (HP-VD) and a newly developed ultra low power vehicle detector (ULP-VD). The HP-VD is a vehicle detector presented in our previous work described above. The ULP-VD is a simple vehicle detector and is implemented on an ultra low power MCU such as MSP430. The ULP-VD is always activated and detects the presence of vehicles in front of microphones. When passing vehicle is detected, the ULP-VD wakes the HP-VD to retrieve more detailed vehicle information including direction of the vehicle.

In this paper, we describe the design of the ULP-VD. Our key idea is to use discrete wavelet transform (DWT) to analyze frequency components of vehicle sound instead of power hungry fast Fourier transform (FFT) to reduce power consumption. Wavelet coefficients derived from DWT include features enough to detect vehicles [1]. The ULP-VD extracts the features and applies a logistic regression analysis to detect vehicles. The ground truth for construction of a logistic regression model would be derived from a HP-VD after installation. By conducting experiments in our university, we demonstrated that the ULP-VD successfully detected vehicles with a precision, recall of 0.94, 0.95, respctively.

The remainder of this paper is structured as follows. Section 2 briefly looks through related works on vehicle detection. In Section 3, we present overview of a low power vehicle detection system and the ULP-VD. Section 4 conducts experiments to demonstrate the effectiveness of the ULP-VD. Finally, Section 5 concludes the paper.

2 Related Works

2.1 Vehicle detection systems

The typical vehicle detection systems are divided into three types of detectors: presence detector, speed detector, and density detector. The presence detector detects passing vehicles. The speed detector measures speed of passing vehicles in addition to the presence detection. The density detector measures traffic volume in addition to the speed detection.

In this paper, we focus on presence detectors because speed and density are generally calculated based on the results of presence detection. Presence detectors are categorized into invasive and non-invasive detectors.

Vehicle detection systems using a loop coil or a geomagnetic sensor are categorized into invasive detectors. Vehicle sensors are installed under a road surface, which results in high resilience to physical damages such as pressure or dust resulting in low maintenance cost. The invasive detectors, however, has limited sensor coverage and require multiple sensors to cover multiple lanes. Moreover, roadwork for installation is an issue that drastically increases initial cost.

Vehicle detection systems using ultrasound or an infreared sensor are categorized into non-invasive detectors. The non-invasive detectors also shows high resilience to physical damages as the detectors detect vehicles from a distance. The coverage of the non-invasive detectors is wider than that of the invasive detectors. However, the non-invasive detectors suffer from high installation cost because the detectors are required to be installed above or by a road for better detection performance. Performance degradation by a bad weather condition is another issue of non-invasive detectors.

To reduce installation and maintenance costs, camera-based vehicle detectors using CCTVs, which are already installed in the environment, have been proposed [3, 10]. However, CCTVs are only available in limited areas, especially in city areas.

On the contrary, vehicle detection using acoustic signals is another approach that comes with low installation and maintenance costs. Several studies have reported on a vehicle monitoring system using acoustic sensors [7, 4, 5, 2]. These studies used a microphone array to draw a sound map, i.e., a map of time difference of sound arrival on two microphones. The studies manually analyzed the sound map to show the feasibility of vehicle detection using the sound map.

We also have developed an acoustic vehicle detection system using a sound map. We applied a state-machine-based algorithm [9] or template matching [8] to a sound map to detect vehicles as well as their direction. Our vehicle detector, however, suffers from high power consumption because the sound map is derived via peak detection in correlation, which requires much computation.

2.2 Vehicle detection system using DWT

Averbuch et al. proposed a robust vehicle detection system including car model [1]. Features of vehicle sound is first extracted using discrete Wavelet transform (DWT) of vehicle sound. The features are then supplied to a machine learning system to detect vehicles and vehicle types. However, this method requires training of the machine learning system prior to system use. Vehicle sound is dependent on vehicle types, which implies the requirement of much training data for many vehicle types.

Choe et al. proposed an acoustic signal analysis method to remotely recognize military vehicles [6]. Type of vehicle is estimated using pattern matching on vehicle sound signatures, which is derived using DWT. However, this method has difficulties in collecting sufficient data for signature creation. For this reason, features are almost manually extracted.

To the best of knowledge, no paper reported an acoustic vehicle detection method employing machine learning with training after deployment. Vehicle passing sound varies from place to place depending on vehicle speed, distance between vehicle and microphones, and road condition. In order to use vehicle

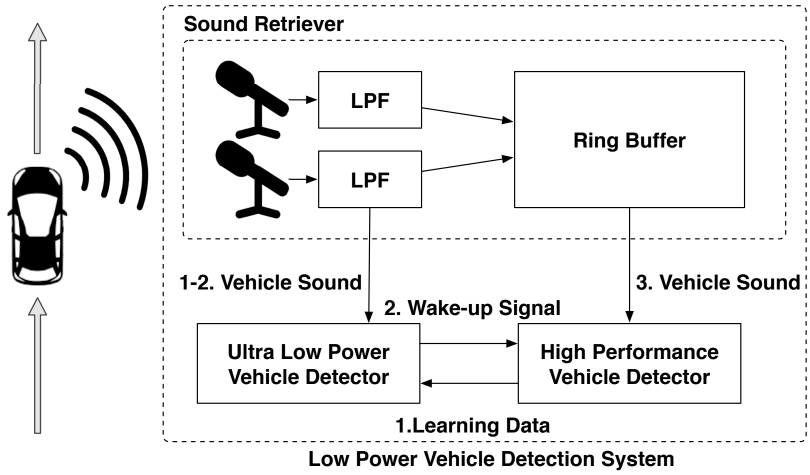

Fig. 1. Overview of low power vehicle detection system

detection system in many conditions, we need training data in each condition, which is a high barrier for deployment.

3 Low Power Vehicle Detection System

3.1 Overview

Fig. 1 shows an overview of a low power vehicle detection system. The low power vehicle detection system consists of three blocks: a sound retriever, high performance vehicle detector (HP-VD), and ultra low power vehicle detector (ULP-VD). Average power consumption is successfully reduced by employing wake-up scheme: the HP-VD is activated only when vehicles are detected by the ULP-VD. We implement the ULP-VD on an ultra low power MCU such as MSP430 to reduce power consumption because the ULP-VD is always activated.

The sound retriever consists of two microphones followed by low-pass filters (LPFs) and a ring buffer. The two microphones are installed at a sidewalk to retrieve vehicle passing sound. The LPF reduces influence of environmental noise. Main component of vehicle passing sound is no more than 2.0 kHz [11]. The cut-off frequency of the LPF is therefore set to 2.5 kHz with a 0.5-kHz margin. Low-pass-filtered sound signals are then stored in the ring buffer.

The ULP-VD retrieves sound signals from one microphone of the sound retriever. The sound data is divided with a time-window. The windowed data is analyzed using discrete Wavelet transform (DWT) to detect vehicles. The ULP-VD wakes HP-VD when a vehicle is detected.

The HP-VD reads sound data from a ring buffer in the sound retriever and draws a sound map to detect vehicles, as presented in our previous work [9, 8]. Vehicle sound is captured at a low frequency of 8 kHz with 16-bit data length.

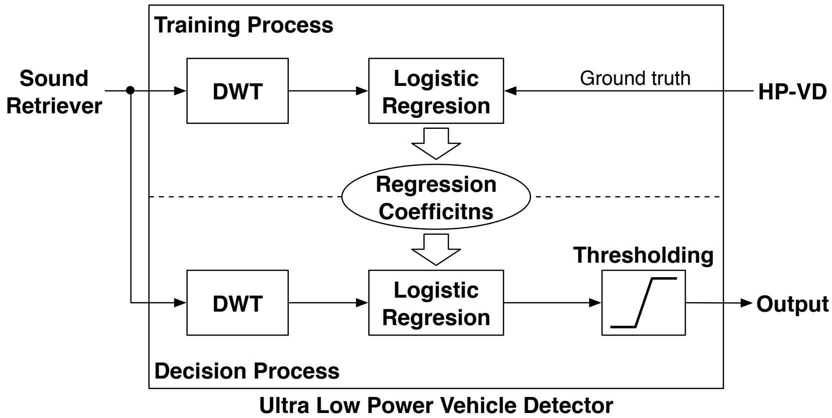

Fig. 2. Overview of ultra low power vehicle detector

The size of ring buffer would be several tens of kilobytes to store sound data for few seconds.

In the following subsections, we describe the ULP-VD because the HP-VD has been presented in our previous papers.

3.2 Overview of Ultra Low Power Vehicle Detector

Fig. 2 shows an overview of the ultra low power vehicle detector (ULP-VD). The ULP-VD is composed of three components: a discrete Wavelet transform (DWT) block, logistic regression block, and thresholding block. We employ machine learning using logistic regression, which consists of training and decision processes.

In a training process, the DWT block analyzes time-frequency components of sound signals and extract feature values. The feature values are then passed to logistic regression as training data to create a regression model; regression coefficients are calculated. Ground truth, i.e., labels for training is derived a high performance vehicle detector (HP-VD), as shown in Fig. 1.

In a decision process, feature values are again extracted from DWT analysis results. Logistic regression block calculates probability of existence of vehicles with the regression coefficients derived in the training process. We finally apply a threshold to the probability to determine if the vehicle is passing in front of microphones.

The following subsections describe the each process in more details.

3.3 Training process

In a training process, logistic regression model is created using vehicle sound and ground truth derived from the HP-VD.

Table 1. Relationship among decomposition level, frequency range, and number of samples (with 512 samples at sampling frequency f_s of 8 kHz)

Decomposition level	Frequency range	Number of samples
6	0 – 0.125 kHz	16
5	0.125 – 0.25 kHz	16
4	0.25 – 0.5 kHz	32
3	0.5 – 1.0 kHz	64
2	1.0 – 2.0 kHz	128
1	2.0 – 4.0 kHz	256

As shown in Fig. 2, time-frequency components derived by DWT is used as feature values. A DWT block analyzes time-windowed sound data to retrieve time-frequency components of the data. We apply Haar wavelet transform to the each windowed data with n decomposition levels.

We extract the maximum value in the each decomposition level, which is used as feature values in a logistic regression block. The numbrer of samples in the each level is different. Table 1 shows an example of relationship among decomposition level, frequency range, and number of samples in discrete Wavelet transform with 512 samples at a sampling frequency f_s of 8 kHz. Higher decomposition levels describe lower frequency bands and have small number of samples, which is one of typical characteristics of discrete Wavelet transform. In our vehicle detector, DWT is used instead of fast Fourier transform (FFT) for frequency component analysis. We choose the maximum value as a representative frequency component in each frequency band.

Fig. 3 shows an example of feature values extracted from motorbike sound, compared with original spectrum. Although absolute value is different, we can confirm the feature values roughly represent the original spectrum.

A logistic regression block calculates probability of vehicle passing based on a logistic regression model. Let $x_1, x_2, \ldots, x_n$ be extracted feature values and $Y = \{0, 1\}$ be a random variable describing vehicle passing. Probability of vehicle passing is derived by

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

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 $C(\mathbf{A})$:

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

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

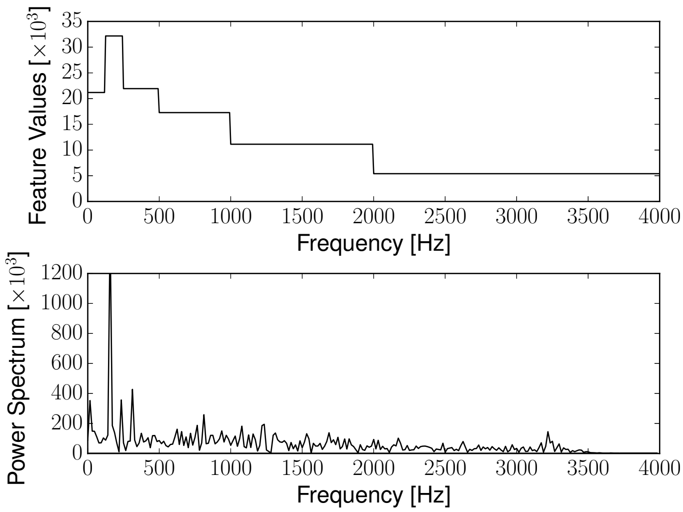

Fig. 3. Example of feature values with six decomposition levels plotted as a function of frequency (above) compared with original spectrum (below)

3.4 Decision process

In a decision process, vehicle passing probability is calculated with the regression coefficients calculated in a training process.

We first perform feature extraction in the same manner as in the training process. Sound signals are divided with a small time window, which are passed to a DWT block extracting feature values, i.e., representative frequency components. A logistic regression block then calculates probability of vehicle passing using Eq. (1).

We finally apply a threshold to the probability to detect vehicles. The threshold is determined from a receiver operating characteristics (ROC) curve in Section 4.

4 Evaluation

As an initial evaluation, we conducted experiments in our university campus.

4.1 Experiment setup

Figure 4 shows experiment instruments. A target road has two lanes, one lane in each direction. Two microphones were installed approximately two meters away from the road center separated by 50 centimeters at a height of one meter. We recorded vehicle sound for approximately 30 minutes using a Sony HDR-MV1 recorder with AZDEN SGM-990 microphones. The sound was recorded

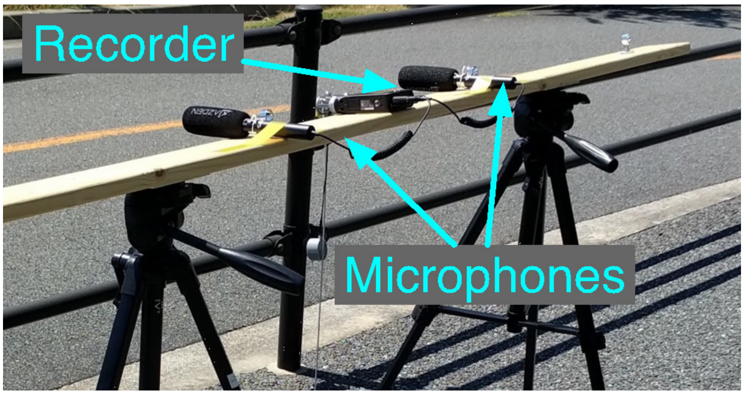

Fig. 4. Experiment instruments

with a sampling frequency of 48 kHz and with a word length of 16 bits. 151 vehicles were passed in front of the microphones during our experiment including buses, motorbikes, trucks, and small cars. Envorinmental noise such as wind and insects chirping was recorded as well as the vehicle sound. Video monitoring the target road was also recorded as ground truth data. Although we installed the two microphones for HP-VD, one of the microphones was used for ULP-VD evaluations in this paper.

4.2 Evaluation procedure

In a first step, the sound data was divided with a 512-sample window. We labeled each windowed data as `1` for vehicle passing and `0` for no vehicle passing. There was a large difference between the number of the labeled data samples with vehicles and no vehicle because time length of vehicle passing was relatively small compared to time length of no vehicle. We randomly picked up no-vehicle data samples to equalize the numbrer of samples in a training process.

In a second step, we determined the threshold described in Section 3.4. The threshold was determined using a receiver operating characteristic (ROC) curve, which is a plot of the true positive (TP) rate against the false positive (FP) rate at various threshold. We can determine an optimum threshold by finding the point shortest to the upper left corner of a ROC curve.

In a final step, comparing the output of decision process with ground truth data, we evaluated the numbers of true positives (TPs), false negatives (FNs), and false positives (FPs). TP, FN, FP are defined as the case that a vehicle detected when a vehicle passing, no vehicle detected when a vehicle passing, and a vehicle detected when no vehicle passing, respectively. Using the numbers of TPs, TNs, FPs, and FNs, we also evaluated precision, recall, and F-measure

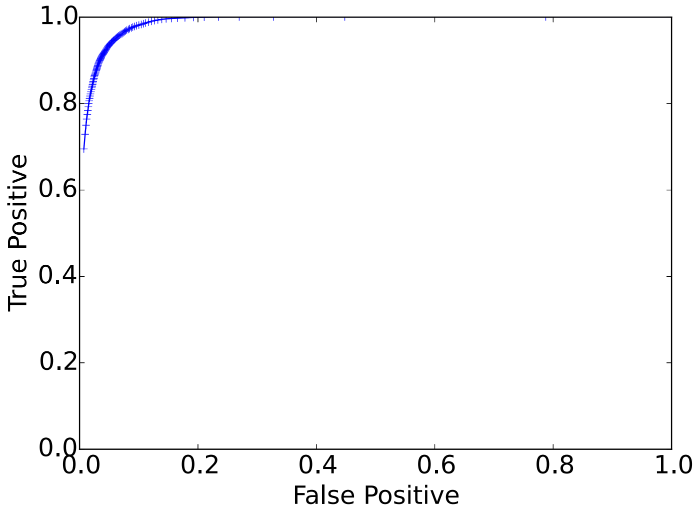

Fig. 5. ROC curve using regression coefficients obtained from the average of 50 times of 10 divided cross-validations. a threshold is changed from 0 to 1 and every 0.01.

defined as:

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

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

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

In this evaluation, we performed 10-fold cross-validations of 50 times. We calculated precision, recall, F-measure for each cross-validation. 50-time cross-validation results were averaged out to derive final results.

4.3 Evaluation results

Figure 5 shows a ROC curve with thresholds changed from 0 to 1 in steps of 0.01. Threshold of 0.37 corresponds to the point shortest to the upper left corner of the ROC curve. The area under the curve (AUC) was 0.99. We threfore used a threshold of 0.37 in the rest of evaluations.

Table 2 shows the average numbers of TPs, FNs, and FPs over 50-time cross-validations. Precision, recall, and F-measure calculated from the numbers of TPs, FNs, and FPs were also shown in Table 2.

Table 2 and Fig. 5 indicate the following:

Table 2. Evaluation results

TPs	FNs	FPs
735	37	45
Precision		0.94
Recall		0.95
F-measure		0.95

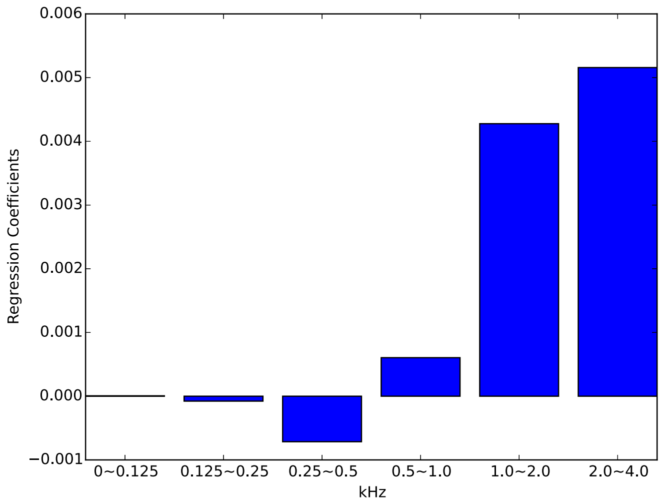

Fig. 6. Regression coefficients as a function of the corresponding frequency band

1. F-measure of 0.95 indicates that the ULP-VD achieved high performance compared to counting by hand that suffers from counting error of few percent. The high f-measure also indicates that the performance of ULP-VD was not very affected by vehicle types such as buses, motorbikes, trucks, and small cars.
2. Recall of 0.95 indicates the small number of FNs. The ULP-VD is used to activate high performance vehicle detector (HP-VD). Small number of FNs is an important feature because FNs result in wake-up failures of the HP-VD.
3. Precision of 0.94 indicates the small number of FPs. FPs are allowed in our low power vehicle detection system, though FPs increase average power consumption. We can see that our ULP-VD successfully prevents unnecessary wake-ups.
4. The AUC of 0.99 implies high performance of the ULP-VD. The ULP-VD exhibits high detection performance with a wide range of thresholds.

The above result reveals that the ULP-VD successfully detected vehicles with accuracy comparable with that of counting by hand.

Figure 6 shows the average regression coefficients derived in a training process as a function of the corresponding frequency band. High regression coefficients indicate a high contribution to vehicle detection. We can confirm that the frequency components between 1.0 kHz and 4.0 kHz were dominant for vehicle detection. Many types of vehicles commonly generate sound signals within this frequency band, resulting in the high regression coefficients.

5 Conclusion

This paper presents a low power vehicle detection system that relies on a high performance vehicle detector (HP-VD) and ultra low power vehicle detector (ULP-VD) employing a wake-up mechanism. Because we have developed the HP-VD in our previous work, an ULP-VD using the discrete Wavelet transform (DWT) and logistic regression was newly developed, which can be implemented on a low power MCU. We conducted experiments in our university to evaluate the basic performance of the ULP-VD. Initial experimental evaluations reveal that the ULP-VD successfully detected vehicles with a precision and recall more than 0.9.

Acknowledgments

This work was supported in part by JSPS KAKENHI Grant Numbers JP15H05708, JP17K19983, and JP17H01741 as well as the Cooperative Research Project of the Research Institute of Electrical Communication, Tohoku University.

References

1. Averbuch, A., Zheludev, V.A., Rabin, N., Schclar, A.: Wavelet-based acoustic detection of moving vehicles. Multidimentional Systems and Signal Processing 20(1), 55–80 (Mar 2009), article no.3
2. Barbagli, B., Manes, G., Facchini, R., Manes, A.: Acoustic sensor network for vehicle traffic monitoring. In: Proc. IEEE Int. Conf. on Advances in Vehicular Systems (VEHICULAR). pp. 1–6 (Jun 2012)
3. Buch, N., Cracknell, M., Orwell, J., Velastin, S.A.: Vehicle localisation and classification in urban CCTV streams. In: Proc. ITS World Congress. pp. 1–8 (Sep 2009)
4. Chen, S., Sun, Z.P., Bridge, B.: Automatic traffic monitoring by intelligent sound detection. In: Proc. IEEE Conf. Intelligent Transportation Systems (ITSC). pp. 171–176 (Nov 1997)
5. Chen, S., Sun, Z., Bridge, B.: Traffic monitoring using digital sound field mapping. IEEE Trans. Veh. Technol. 50(6), 1582–1589 (Nov 2001)
6. Choe, H.C., Karlsen, R.E., Gerhart, G.R., Meitzler, T.J.: Wavelet-based ground vehicle recognition using acoustic signals. In: Proc. SPIE. Wavelet Applications III, vol. 2762, pp. 434–445 (Mar 1996)
7. Forren, J.F., Jaarsma, D.: Traffic monitoring by tire noise. In: Proc. IEEE Conf. Intelligent Transportation Systems (ITSC). pp. 177–182 (Nov 1997)

8. 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)
9. 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)
10. Nurhadiyatna, A., Hardjono, B., Wibisono, A., Jatmiko, W., Mursanto, P.: ITS information source: Vehicle speed measurement using camera as sensor. In: Proc. Int. Conf. on Advanced Computer Science and Information Systems (ICACSIS). pp. 179–184 (Dec 2012)
11. Wu, H., Siegel, M., Khosla, P.: Vehicle sound signature recognition by frequency vector principal component analysis. In: Proc. IEEE Instrumentation and Measurement Technology Conf. (IMTC). vol. 1, pp. 429–434 (May 1998)