11th ITS European Congress, Glasgow, Scotland, 6-9 June 2016

Paper number EU-TP0042

Design of simple vehicle counter using sidewalk microphones

Shigemi Ishida[1*], Kohei Mimura[1], Song Liu[1], Shigeaki Tagashira[2], Akira Fukuda[1]

1. ISEE, Kyushu University, Japan

{ishida,mimura,ryu,fukuda}@f.ait.kyushu-u.ac.jp

2. Faculty of Informatics, Kansai University, Japan

shige@res.kutc.kansai-u.ac.jp

Abstract

Vehicle counting is one of the fundamental tasks in the intelligent transportation system (ITS). Although automatic vehicle counters have been proposed to retrieve realtime traffic data, current automatic vehicle counters are suffered from high deployment costs, resulting in limited number of deployments. In this paper, we present a vehicle counter using sidewalk microphones that are easily deployed at a low cost. Our vehicle counter only relies on two microphones and counts vehicles using a sound map, which is a time-difference map of vehicle sound on the two microphones. We developed a vehicle count algorithm using a sound map. To minimize the effect of practical issues on vehicle counting, we also developed a simple image processing technique. We herein describe the design and implementation of our vehicle count system. Experimental evaluations reveal that our vehicle count system successfully counted vehicles with an F-measure of 0.92.

Keywords:

Vehicle count, acoustic sensing, sound map.

1. Introduction

Increasing attention has been focused on the intelligent transportation system (ITS) due to the change of road transportation strategy. The main purpose of the ITS is to improve the safety, efficiency, dependability, and cost effectiveness of transportation systems. In the past decade, products such as car navigation systems have brought the ITS to daily life.

Vehicle counting is one of the fundamental tasks in the ITS. In Japan, vehicle counting has been mainly conducted as a road traffic census almost every five years since 1928 by the Ministry of Land, Infrastructure, Transportation and Tourism (MLIT). The traffic census investigates temporal traffic volume, which restricts usage of the traffic data to non-realtime applications.

To retrieve realtime traffic data, automatic vehicle count systems have been deployed. However, the deployment of the automatic vehicle count system is limited to high traffic roads because of its high installation and maintenance costs. The automatic vehicle count systems are also suffered from a motorbike counting problem. Although camera-based vehicle counters that are capable of motorbike counting are proposed, restrictions on camera location and angle make difficulties on practical deployment.

Therefore, this paper proposes a simple vehicle counter that can be easily deployed at a low cost. Our vehicle counter only relies on two microphones at a sidewalk. Because sound waves are diffracted over obstacles, we can deploy microphones in a low height configuration, which drastically reduces roadwork costs in terms of road closure as well as safety installation. Our vehicle counter can detect all types of vehicles including motorbikes as long as the vehicle generates sound.

There are several studies reporting a vehicle monitoring system using acoustic sensors [1–4]. These studies used a microphone array to draw a *sound map*, i.e., a map of time difference of vehicle sound on different microphones. The studies manually analyzed the sound map and demonstrated that the sound map can be used for vehicle counting.

We extend these sound map studies to realize an automatic vehicle count system. We present a simple vehicle count algorithm using a sound map. In a practical environment, a sound map sometimes becomes noisy due to distributed sound sources on a vehicle. We therefore develop a simple image processing technique to minimize the effect of the sound map noise. By conducting experiments in our university, we demonstrate that our vehicle count system accurately counted vehicles with an F-measure of 0.92.

The remainder of this paper is structured as follows. Section 2 briefly looks through related works on vehicle counting. We present our vehicle count system in Section 3 and extend the vehicle count system in Section 4 to address practical issues. In Section 5, we implement our vehicle count system and conduct experiments to reveal the basic performance. Section 6 concludes the paper.

2. Related works

Current vehicle counters are categorized into two types: intrusive and non-intrusive.

Loop coils and photoelectric tubes are categorized into the intrusive vehicle counters. These vehicle counters are required to be installed under the road surface. The installation and maintenance therefore require roadwork closing a target road section, which is suffered from high costs. Loop coils and photoelectric tubes also have difficulties in motorbike detection due to their small coverage.

The non-intrusive vehicle counter is based on sensors such as laser, infrared, ultrasound, radar, or camera. The non-intrusive vehicle counter needs to be installed above or by a road for better performance. Deployment above a road is suffered from high installation and maintenance costs in terms of roadwork. Roadside non-intrusive vehicle counters are capable of single lane detection and only works on small roads. Most of non-intrusive counters are based on laser, infrared, or ultrasound. These counters have small coverage and are suffered from the motorbike detection problem.

To reduce installation and maintenance costs, camera-based vehicle counters using CCTVs installed in the environment have been proposed [5,6]. CCTVs, however, are only available in limited areas, especially in city areas. Performance of vehicle counters using CCTVs is also affected by weather condition because camera location and angle are not suitable for vehicle counting but for security surveillance.

On the contrary, acoustic approach is a promising candidate for vehicle counting at a low installation and maintenance costs. Using roadside microphone array, we can locate a sound source, i.e., a vehicle on a road. Acoustic approach is capable of counting vehicles on multiple lanes at a sidewalk because sound waves are diffracted over obstacles.

Several studies have reported on a vehicle monitoring system using acoustic sensors. Forren et al. and Chen et al. proposed traffic monitoring schemes using a microphone array [1–3]. The monitoring schemes draw a sound map, i.e., a map of time difference of vehicle sound on different microphones and analyze the sound map to monitor vehicles. The monitoring schemes are missing design details of vehicle counting. The monitoring scheme also installs a microphone array in a high height configuration at a roadside and monitors vehicles on multiple lanes. The high height configuration fails to reduce installation cost in terms of safety installation.

Barbagli et al. reported an acoustic sensor network for traffic monitoring [4]. The acoustic sensor network installs sensor nodes at road sides. Each sensor node draws a sound map and combines the sound map with an energy detection result to monitor traffic flow distribution. The sensor network requires many sensor nodes at both sides of the road to monitor realtime traffic flow, which results in high deployment and maintenance costs. The paper also lacks an evaluation of accuracy on vehicle counting because the sensor network focuses on monitoring traffic flow with small energy consumption.

3. Sidewalk vehicle count system

A. Overview

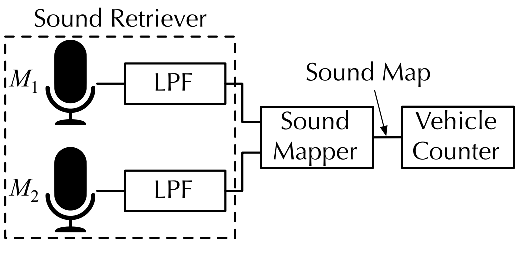

Figure 1 - Overview of sidewalk vehicle count system

Figure 1 depicts an overview of our sidewalk vehicle count system. Our vehicle count system consists of three components: a sound retriever, sound mapper, and vehicle counter. The sound retriever consists of two microphones followed by a low-pass filter (LPF). We install two microphones at a sidewalk of a road and record vehicle sound. The LPF removes high frequency environmental noise and the sound mapper calculates cross-correlation function between sounds on the two microphones to draw a sound map. Finally, we apply a simple vehicle count algorithm to the sound map, retrieving the number of vehicle passed in front of the microphones.

In the following subsections, we present design details of the sound retriever, sound mapper, and vehicle counter.

B. Sound retriever

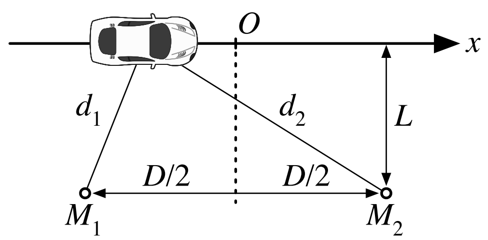

Figure 2 - Microphone setup

Figure 2 depicts a microphone setup. Two microphones M_1 and M_2 are installed at a sidewalk of a road parallel to the road.

Distance D between the microphones and distance L between the road and microphones affect vehicle count performance. Sound arrival time difference on the two microphones is maximum when a vehicle is at $x = \pm\infty$. The maximum sound arrival time difference Δt_{max} is calculated to be

$$\pm\Delta t_{max} = \pm\frac{D}{c}, \quad (1)$$

where c is the speed of sound in air. As D increases, a curve on a sound map increases or decreases more quickly. We therefore expect count accuracy improvement as D increases. Increasing D, however, degrades count accuracy because of increase in coverage, which tends to be affected by environmental noise. Distance L is in similar situation. Distance D and L might be determined by preliminary experiments as well as physical restrictions.

To minimize the effect of environmental noise, our vehicle count system applies a low-pass filter (LPF) to vehicle sound signals. Majority of frequency component of sound signals generated by vehicle tires is less than 2.0 kHz [7]. The cut-off frequency of the LPF is therefore 2.5 kHz including a margin. Because the tire sound is generated by all types of vehicles, our vehicle count system can detect every type of vehicles including buses, trucks, motorbikes, and electric vehicles.

C. Sound mapper

A sound map is time series data keeping track of sound arrival time difference between two microphones. As shown in Figure 2, we install two microphones M_1 and M_2 separated by D by a road at the distance of L. Sound signals generated by a vehicle on the road reach the two microphones with different traveling distance. Let x be the location of a vehicle. The sound traveling distance d_1 and d_2 are calculated to be

$$d_1 = \sqrt{\left(x + \frac{D}{2}\right)^2 + L^2}, \quad (2)$$

$$d_2 = \sqrt{\left(x - \frac{D}{2}\right)^2 + L^2}. \tag{3}$$

We derive the time difference Δt of sound arrival between the two microphones using the speed c of sound in air:

$$\begin{aligned}\Delta t &= \frac{d_1 - d_2}{c} \\ &= \frac{1}{c}\left\{\sqrt{\left(x + \frac{D}{2}\right)^2 + L^2} - \sqrt{\left(x - \frac{D}{2}\right)^2 + L^2}\right\}.\end{aligned} \tag{4}$$

Using Equation (4), we can locate a vehicle from sound delay. Sound delay can be derived using a cross-correlation function $R(t)$. Let $s_1(t)$ and $s_2(t)$ be sound signals on the two microphones. The cross-correlation function $R(t)$ is defined as

$$R(t) = \int s_1(\tau)\, s_2(\tau + t) d\tau. \tag{5}$$

We assume that the two microphones receive same sound signals with time shifted by Δt, i.e., $s_1(t) = s_2(t + \Delta t)$. Because $R(t)$ is maximum at $t = \Delta t$, we can estimate the sound delay Δt by finding a peak of $R(t)$.

In our vehicle count system, the generalized cross-correlation (GCC) function [8], which is commonly used in the field of acoustic source localization, is used to estimate the sound delay. The count system divides sound signals with small window and applies the GCC to each windowed data to estimate the sound delay. The sound map is derived by plotting the sound delay of each windowed data.

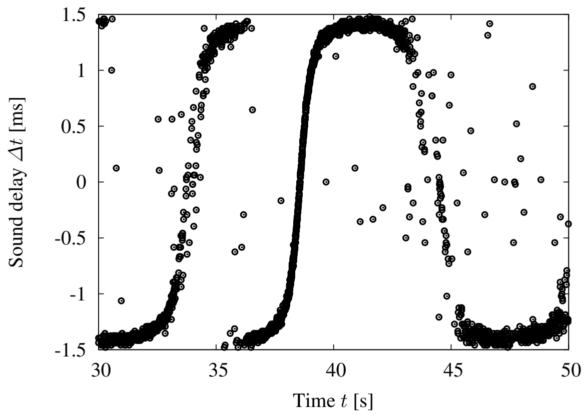

Figure 3 - Example of sound map

A typical sound map, i.e., sound delay Δt as a function of time, is shown in Figure 3. As a vehicle passes in front of microphones, sound delay increases or decreases drawing an S-curve; direction of the vehicle decides the direction of the S-curve.

D. Vehicle counter

Figure 4 illustrates an S-curve on a sound map when a vehicle passes in front of microphones from left to right. We divide an S-curve into three sub-curves as shown in Figure 4; sub-curves 1, 2, and 3 are observed when a vehicle coming toward microphones, the vehicle passing in front of microphones, and the vehicle going away from microphones, respectively. The sub-curves 1 and 3 are close to asymptotes $\Delta t = \pm \Delta t_{max}$ given by Equation (1).

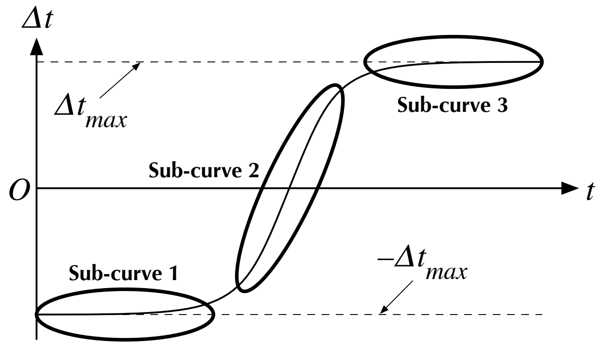

Figure 4 - Vehicle passing drawn on sound map

Order of sub-curves depends on vehicle direction because the direction of an S-curve depends on vehicle direction. We separately apply a vehicle count algorithm for each vehicle direction to the sound map.

Here we describe our vehicle count algorithm using a sound map curve of a vehicle passing from left to right as an example. The count algorithm is a state machine that keeps track of sound delay to detect the sub-curves.

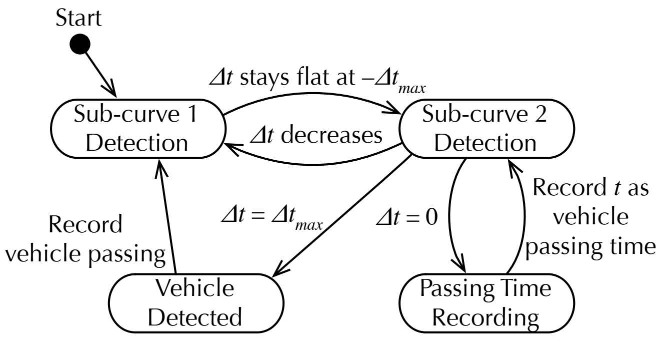

Figure 5 - State machine diagram of vehicle count process (for left to right vehicles)

Figure 5 illustrates a state machine diagram of our vehicle count process. The count process starts from a *sub-curve 1 detection* state. Sub-curve 1 is close to an asymptote $\Delta t = -\Delta t_{max}$. If the sound delay Δt remains flat at $\Delta t = -\Delta t_{max}$ for a specific duration, the count process switches its state to a *sub-curve 2 detection* state.

In the sub-curve 2 detection state, the count process keeps track of the sub-curve increasing as time increases. When the sound delay $\Delta t = 0$, the count process temporarily switches its state to a *passing time recording* state to make a tentative record of vehicle passing time.

When Δt reaches Δt_{max}, the count process switches its state to a *vehicle detected* state and finalizes a vehicle passage. The count process switches to the sub-curve 1 detection state to start detection of next vehicle passage. In the sub-curve 2 detection state, the count process state returns to the sub-curve 1 detection state when Δt decreases to restart the count process.

4. Extend to practical environment

A. Big vehicle detection problem

Sound delay draws an S-curve for a vehicle, as shown in Figure 3. For big vehicles such as buses and trucks, however, sound delay partially splits into two curves, which results in failure of vehicle count.

Figure 6 shows an example of sound map when a big vehicle passes in front of microphones. A curve partially splits into two curves around $\Delta t = 6$ s. The maximum separation between the two curves is approximately 1 millisecond. A curve on a sound map indicates a vehicle; two curves seem to indicate two vehicles, resulting in false positive vehicle counting.

The two curves on a sound map are generated by front and rear tires. Consider the case of a bus passage right in front of microphones as shown in Figure 7. Let l be a wheelbase of the bus. Sound delays Δt_F and Δt_R of front and rear tires are calculated by Equation (4). We derive difference δ between the sound delays of front and rear tires as:

$$\begin{aligned} \delta &= \Delta t_R - \Delta t_F, \\ \Delta t_F &= \frac{1}{c}\left\{\sqrt{\left(-\frac{l}{2}+\frac{D}{2}\right)^2 + L^2} - \sqrt{\left(-\frac{l}{2}-\frac{D}{2}\right)^2 + L^2}\right\}, \\ \Delta t_R &= \frac{1}{c}\left\{\sqrt{\left(\frac{l}{2}+\frac{D}{2}\right)^2 + L^2} - \sqrt{\left(\frac{l}{2}-\frac{D}{2}\right)^2 + L^2}\right\}. \end{aligned} \tag{6}$$

When $l = 4.5$ m, $D = 0.5$ m, and $L = 4$ m, the sound-delay difference δ is 1.4 milliseconds. The sound-delay difference is considerable compared to the maximum sound delay $\Delta t_{max} \approx 1.5$ milliseconds in Figure 3.

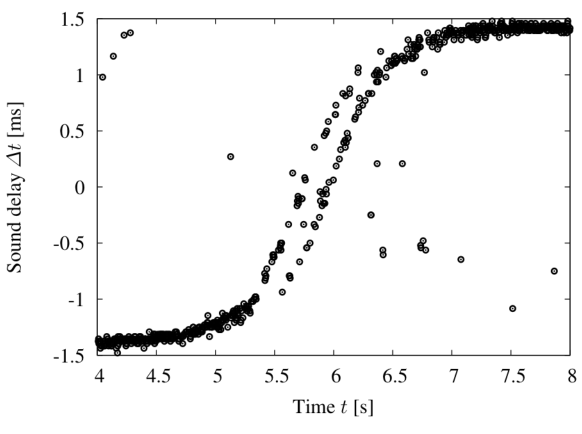

Figure 6 - Example of sound map when big vehicle passes

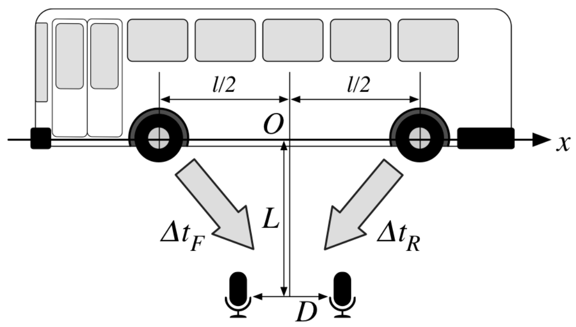

Figure 7 - Bus passing right in front of microphones

B. Image processing for sound map

To address the big vehicle detection problem, we developed a simple image processing technique for a sound map. The image processing technique combines two curves generated by a big vehicle into a single bold curve. No modification on the count algorithm is required. The image processing technique also reduces the effect of noise and improves count accuracy.

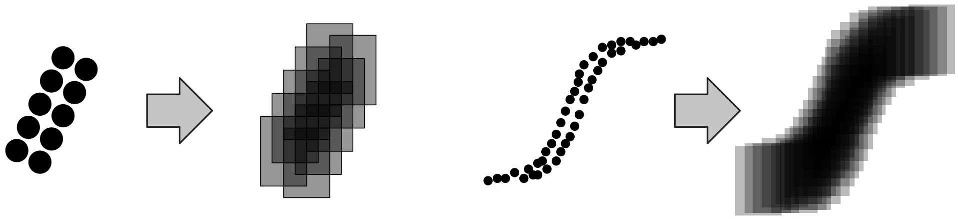

Figure 8 - Overview of image processing

Figure 8 illustrates an overview of image processing. The image processing is replacement of each point on a sound map with translucent rectangle. Rectangles, having height in sound-delay dimension and width in time dimension, overlap each other and reproduce a bold curve with a vague outline.

The height and width of the rectangles is determined from the maximum length of a vehicle and a speed limit of a road. The height of the rectangles is set to the maximum sound-delay separation. Maximum sound-delay separation is calculated by substituting the maximum length of a vehicle for l in Equation (6). The width of the rectangles should be set to a time duration such that a vehicle at the speed limit moves negligible distance compared to the vehicle length.

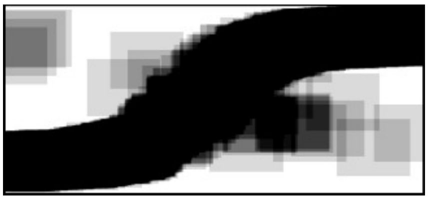

Figure 9 - Sound map after image processing

We apply our image processing technique to Figure 6 and derive Figure 9. Comparing the figures reveals the effect of the image processing. We can observe a single bold curve on a sound map for a big vehicle.

5. Evaluation

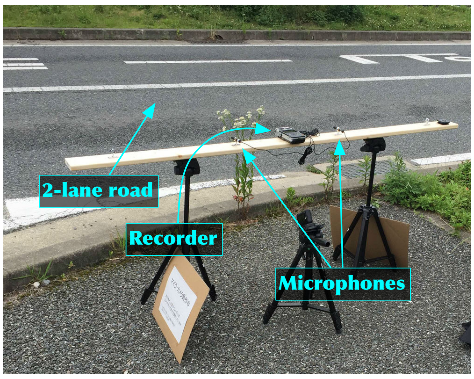

Figure 10 - Experiment setup

As an initial evaluation, we conducted experiments in our university evaluating the basic performance of our vehicle count system. Figure 10 shows an experiment setup. A target road has two lanes, one lane for each direction. Two microphones were installed approximately two meters away from the road center. Distance between the two microphones was 50 centimeters, which is determined based on preliminary experiment results. We recorded vehicle sound for about 30 minutes using a Sony PCM-D100 recorder with OLYMPUS ME30W microphones. The sound was recorded with a sampling frequency of 48 kHz and word length of 16 bits. We also recorded video monitoring the road, which was used as ground truth data.

Comparing the results derived by our vehicle count system with the video, we evaluated the numbers of true positives (TPs), false negatives (FNs), and false positives (FPs). TP, FN, and 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. We excluded true negatives (TNs), which is defined as the case that no vehicle detected when no vehicle passing, because TNs were not countable in our experiments. Using the numbers of TPs, FNs, and FPs, we also evaluated an accuracy, precision, recall, and F-measure. Note that we used the number zero of TNs in these calculations.

Table 1 - Experiment results

	TP	FN	FP
Left to right	63	11	0
Right to left	87	15	0
Total	150	26	0

Table 2 - Experiment results vs. vehicle types

	Accuracy	Precision	Recall	F-measure
Normal cars	0.78	1.00	0.78	0.88
Buses, trucks	0.91	1.00	0.91	0.95
Small cars	0.97	1.00	0.97	0.98
Motorbikes	0.81	1.00	0.81	0.89
All	0.85	1.00	0.85	0.92

Table 1 summarizes the numbers of TPs, FNs, and FPs. As described in Section 3-C, our vehicle count system separately counts vehicles for each vehicle direction. We therefore derived the numbers of TPs, FNs, and FPs for each vehicle direction and summed the results to retrieve a total result.

Using Table 1, we calculated an accuracy, precision, recall, and F-measure of 0.85, 1.00, 0.85, and 0.92, respectively. Our vehicle count system successfully counted vehicles with an F-measure of 0.92. Table 2 shows the experiment results for each type of vehicles. We confirmed that all types of vehicles were accurately counted with a minimum F-measure of 0.88.

The number of FPs was zero. Our vehicle count system exhibited high tolerance to environmental noise such as wind and people chattering. There was no mistake on detection of vehicle direction.

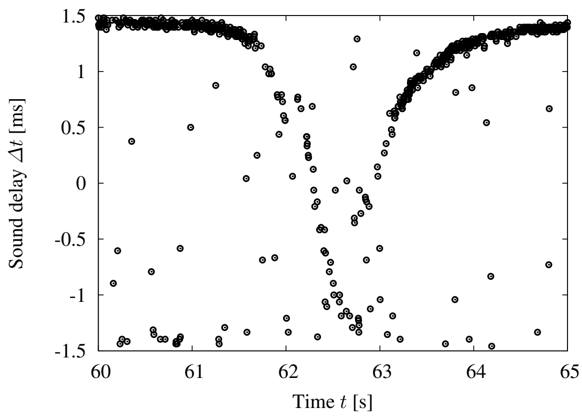

Figure 11 - Example sound map when two vehicles simultaneously coming from opposite directions

We analyzed the false negatives as performance improvement discussion. 22 false negatives in the 26 were the case that two vehicles were simultaneous coming from opposite directions and consecutively coming from same direction. Figure 11 depicts an example sound map when two vehicles were simultaneously coming from opposite directions. The vehicle count algorithm has less considerations on simultaneous vehicle detection, resulting in light colored curves that are hardly ever detected.

6. Conclusion

In this paper, we presented an acoustic vehicle count system using sidewalk microphones that are easily deployed at a low cost. Our vehicle counter only relies on two microphones and counts vehicles using a sound map, which is a time-difference map of vehicle sound on the two microphones. To minimize the effect of practical issues on vehicle counting, we developed a simple image processing technique. We conducted experimental evaluations and demonstrated that our vehicle count system accurately counted vehicles with an F-measure of 0.92.

Acknowledgment

This work was supported in part by JSPS KAKENHI Grant Numbers 15H05708 and 15K12021 as well as the Cooperative Research Project of the Research Institute of Electrical Communication, Tohoku University.

References

1. J. Forren, and D. Jaarsma (1997), "Traffic monitoring by tire noise", in Proceedings *IEEE Conf. Intelligent Transportation Systems (ITSC)*, pp.177–182.
2. S. Chen, Z.Sun, and B. Bridge (1997), "Automatic traffic monitoring by intelligent sound detection", in Proceedings *IEEE Conf. Intelligent Transportation Systems (ITSC)*, pp.171–176.
3. S. Chen, Z. Sun, and B. Bridge (2001). "Traffic monitoring using digital sound field mapping", *IEEE Transactions on Vehicular Technology*, vol.50, no.6, pp.1582–1589.
4. B. Barbagli, G. Manes, R. Facchini, et al. (2012). "Acoustic sensor network for vehicle traffic monitoring", in Proceedings *1st Int'l Conf. Advances in Vehicular Systems, Technologies and Applications*, pp.1–6.
5. N. Buch, M. Cracknell, J. Orwell, et al. (2009). "Vehicle localisation and classification in urban CCTV streams" in Proceedings *ITS World Congress*, pp.1–8.
6. A. Nurhadiyatna, B. Harjono, A. Wibisono, et al. (2012), "ITS information source: Vehicle speed measurement using camera as sensor", in Proceedings *Int'l Conf. Advanced Computer Science and Information Systems (ICACSIS)*, pp.179–184.
7. H. Wu, M. Siegel, and P. Khosla (1998), "Vehicle sound signature recognition by frequency vector principal component analysis", in Proceedings *IEEE Instrumentation and Measurement Technology Conf. (IMTC)*, vol.1, pp.429–434.
8. C. Knapp, and G. Carter (1976), "The generalized correlation method for estimation of time delay", *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol.29, no.4, pp.320–327.