A vehicular noise surveillance system integrated with vehicle type classification

Chuang Shi¹, Woon-Seng Gan², Yong-Kim Chong³, Agha Apoorv⁴, and Kin-San Song⁵ ¹⁻⁴School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore

¹E-mail: shichuang@ntu.edu.sg Tel: +65-67906901 ²E-mail: ewsgan@ntu.edu.sg Tel: +65-67904538

³E-mail: eykchong@ntu.edu.sg Tel: +65-67904535

⁴E-mail: agha.apoorv@ntu.edu.sg Tel: +65-67906901

⁵Pollution Control Department, National Environment Agency, Singapore

E-mail: song kin san@nea.gov.sg Tel: +65- 67319935

Abstract— This paper introduces an ongoing project on the surveillance of noisy vehicles on the road. Noise pollution created by vehicles on urban roads is becoming more severe. To enforce current measures, we developed a vehicular noise surveillance system including a vehicle type classification method. Samples of vehicular noise were recorded on-site using this system. Harmonic features were extracted from each sample based on an average harmonic structure. The k-nearest neighbor (KNN) algorithm was applied to achieve classification accuracies for the passenger car, the van, the lorry, the bus, and the motorbike of 60.66%, 65.38%, 52.99%, 62.02%, and 80%, respectively. This study was motivated by the demand of monitoring noise levels generated by different types of vehicles. The classification method using audio features is independent of lighting condition, thus providing a replacement to machine vision based techniques in vehicle type classification.

I. INTRODUCTION

In this day and age, nations are facing increasingly serious pollution problems caused by a variety of sources. Noise is one of the most widely-cited pollution sources. In particular, traffic noise is steadily gaining attention because of the rapid growth in the number of vehicles on the road [1]. In urban areas, residential buildings have been built close to highways. Exposure to traffic noise at night causes significant reduction in sleep quality. A past study shows that the quality of sleep can be significantly affected when more than 30 noise events happen, even though the maximum noise level of these events are merely around 60 dB [2]. Unlike noise issues in confined areas, whereby the three elements of noise, namely vibrating source, transmission path, and receiver, can be fine controlled [3-5], traffic noise can only be moderated through measures promoted by the government. Systems to enforce traffic noise measures are of increasing importance. For instance, current vehicle inspection measures the engine noise when the vehicle is stationary in a workshop. The result is generally different from the actual noise emission when the vehicle is moving.

To address this problem, a noise source localization method was proposed by Houzu et al. [6-7] to detect noisy vehicles in a traffic flow. A hybrid method of beamforming and sound intensity mapping was used in a microphone array consisting

of 31 microphones with an overall length of 1.2 meters. Realtime implementation was carried out in a field-programmable gate array (FPGA), which was able to handle a block size of 40 milliseconds. In their field test, the microphone array was setup by the roadside and placed 4 meters above the ground. Moving vehicles were localized with good accuracies in real road conditions.

One possible extension to Houzu's noise source localization method is to integrate vehicle type classification. In this way, noise threshold levels can be set for different types of vehicles in accordance with lawful noise regulations. Machine vision based classification methods commonly fail when the contour of vehicles is unable to be extracted due to its shadow and blurriness [8]. Alternatively, in 1999, Succi et al. [9] observed that vehicular noise was resulted from rotating components, which was periodic in nature. Then, Tung and Yao [10] used a triangular microphone array and implemented a nonlinear dynamic method to classify vehicle types. A wavelet-based algorithm was investigated by Averbuch et al. [11] to process the audio information captured at a large distance, which may be inaudible to humans. The wavelet-based method showed a better result when the ratio of vehicular noise to background noise was low. Moreover, Srour [12] discussed the feasibility of classifying military vehicles using audio features. Recently, a support vector machine was developed to classify vehicle types using harmonic features and a Bayesian decision fusion at overall accuracies of 73.4% and 82.4%, respectively [13]. Watts [14] combined a speed camera, a video camera, and a microphone concurrently to capture the noise emitted from a moving vehicle, and the relation between noise level and speed were studied for different types of vehicles. They showed that the vehicle speed can be a good complement to audio features in vehicle type classification.

A local project in Singapore is being carried out to identify noisy vehicles on the road. The purpose of this project besides lawful enforcement is to reduce the number of noisy vehicles on the roads. Moreover, this work will help to improve drivers' driving manners, especially to avoid sudden acceleration that leads to annoying engine noises. A prototype of the project has been deployed at a road junction, where the noise level is recorded daily during peak hours. The samples of vehicular noise are recorded and attempted to be classified into five vehicle types, *i.e.* the passenger sedan car, the bus, the van, the lorry, and the motorbike.

This paper begins with an introduction to the design of the vehicular noise surveillance system. In Section 3, an audio feature extraction method is derived from the average harmonic structure. The KNN algorithm to perform the classification is described in Section 4. Section 5 presents preliminary results. Sections 6 concludes this paper.

II. DESIGN OF THE NOISE SURVEILLANCE SYSTEM

One challenge of designing the noise surveillance system is choosing a correct equipment to pick out the noise from a fast moving vehicle. Because of the requirements of this project, the noise level is only captured from a small spot which is as large as the size of one vehicle. A microphone array forms a sharper polar pattern by applying phased array techniques. However, the cost of the system greatly increases due to the usage of more microphones, more channels in data acquisition, and a more powerful processing unit for real-time processing. Hence, a directional microphone could be a balanced solution between the performance and the cost.

A conventional type of directional microphone is known as the shotgun microphone [15]. It is initially designed for use in professional broadcasting and high-quality sound recording in film production. The highly directional pattern of a shotgun microphone is a result of the interference tube placed in front of the condenser element. A longer tube generally results in a sharper polar pattern. Alternatively, a novel design of a highly directional microphone was proposed to adopt the parametric array effect in air [16]. Although the opposite application of the parametric array effect in air has been thoroughly studied [17-21], the feasibility of a parametric microphone has yet to be figured out.

The development of a noise surveillance system is planned in two phases. Firstly, an overhead mounted design is carried out. The system diagram is shown in Figure 1(a). One shotgun microphone is used to record the noise emitted from vehicles passing beneath it. The processing unit processes the recorded samples and determines whether the noise level exceeds the threshold. The camera will be activated by the processing unit to capture the vehicle's number plate after detecting the noise violation. In the second phase, a road-side deployed design will be carried out. The system diagram is shown in Figure 1(b). The single shotgun microphone in the overhead mounted design is replaced by an array of shotgun microphones, which serves dual roles of noise detection and noise tracking. In the road-side deployed design, the monitoring area is divided into three zones for (i) alarming, (ii) tracking, and (iii) capturing. The microphone array will continuously monitor in-coming traffic in the alarm zone, and switches to the tracking mode once the noise level limit has been exceeded. It will track the offending vehicle and determines its lane in the measurement zone. The lane number is sent to the camera to capture the number plate of the offending vehicle in the correct lane.



Figure 1 System diagrams of the vehicular noise surveillance system: (a) overhead mounted design; (b) road-side deployed design.



Figure 2 An experimental setup of vehicular noise surveillance system.

Figure 2 presents the experimental setup of vehicular noise surveillance system for the overhead mounted usage. A laptop is used as the processing unit, and an audio interface is used as the data acquisition hardware. They are both connected via the universal serial bus (USB). The camera and the shotgun microphone are pointed to the same spot. Daily measurement has been carried out using this experimental setup.



Figure 3 Extraction of harmonic features in one frame of a noise sample.

III. HARMONIC FEATURE EXTRACTION USING THE AVERAGE HARMONIC STRUCTURE

The average harmonic structure was originally proposed by Duan *et al.* [22] for unsupervised musical source separation. Their motivation was to model the timbre characteristics of different musical instruments and subsequently solve a blind source separation problem. The average harmonic structure is



Figure 4 Harmonic features for different types of vehicles using the average harmonic structure and data recorded from our prototype setup of the noise surveillance camera.

more robust than mel-frequency cepstral coefficients (MFCC) in the presence of additive noise. In our project, the vehicular noise can be similarly treated a harmonic source. It is excited from rotating components and reshaped by the vehicle body, which is a resonating system. Automobile engines are typically operated below 3000 revolutions per minute, equivalent to 50 Hz.

The harmonic features are extracted based on the average harmonic structure for each sample of the vehicular noise in the following steps: (a) Power spectrum is plotted for each frame of the sample, and the frame size is 100 milliseconds in this paper, as shown in Figure 3(a); (b) Locations of peaks in the power spectrum are extracted to decide the fundamental frequency by a linear least squares method, as illustrated in Figure 3(b); (c) A triangular filter bank with equal bandwidth is designed to have center frequencies given by harmonics of the fundamental frequency in each frame, and the number of filters in the filter bank is fixed at 20, as shown in Figure 3(c); (d) After applying the filter bank in the power spectrum, the maximum value in the bandwidth of each filter is taken as the harmonic feature of one frame at one order of harmonic. They are then averaged in time domain across all the frames to result in 20 harmonic features of one vehicular noise sample, corresponding to the fundamental frequency and from the 1st harmonic up to the 19th harmonic.

We have recorded vehicular noise from 122 sedan cars, 258 buses, 104 vans, 268 lorries, and 100 motor bikes. Each sample is with a length of at least 1 second (10 audio frames). Half of the recorded samples from each type of vehicles are used as reference data. Their harmonic features are plotted in Figure 4 with respect to the harmonic order from 1 (corresponding to the fundamental frequency) to 20 (corresponding to the 19th harmonic). Motor bikes as being two-wheeled vehicles show distinguishable harmonic features as compared to four-wheeled vehicles.

IV. KNN CLASSIFICATION ALGORITHM

A basic KNN algorithm is carried out for classification of vehicle types in this paper. The KNN algorithm computes the Euclidean distances from one testing sample to each reference sample in the harmonic feature space. A shorter distance shows a closer match between the testing sample and one reference sample. When k = 1, the testing sample is straightforwardly classified into the type of its nearest neighbor. This algorithm is one of the oldest methods yet widely applied in data mining [23].

Accuracies of classifying vehicle types are plotted in Figure 5 using 1NN, 2NN, 3NN, 4NN, and 5NN algorithms. It shows that sedan cars are the most difficult to be distinguished as their noise levels are relatively low as compared to the other types of vehicles. The accuracy of classifying lorries is almost independent from the number of neighbors, while the 1NN algorithm gives the most accurate classification for vans and motorbikes.

Instead of using a fixed number of neighbors across all the vehicle types, we investigated to use a fixed portion of the reference samples to be the number of neighbors in the KNN



Figure 5 Classification accuracies when different numbers of neighbors are selected in the KNN algorithm.



Figure 6 Classification accuracies when the numbers of neighbors selected in the KNN algorithm are given by a ratio of the number of samples to be classified.

algorithm for each vehicle type. The resulted accuracies of classifying vehicle types are plotted in Figure 6. It shows that the average classification accuracy achieves its highest level when one eightieth of the number of the reference samples of each vehicle type is used in the KNN algorithm. Note that the number of neighbors are always rounded upwards towards the next highest integer. In this case, classification accuracies of the passenger car, the van, the lorry, the bus, and the motorbike are obtained at 60.66%, 65.38%, 52.99%, 62.02%, and 80%, respectively.

V. CONCLUSIONS

In this paper, we introduced two designs of vehicular noise surveillance system, which could be mounted on an overhead bridge or placed along the roadside, respectively. Differing from most of the past studies, we propose to use one shotgun microphone or an array of shotgun microphones to sharpen the polar pattern and reduce the overall cost of the system. Daily measurements were carried out using a prototype setup to get on-site vehicular noise samples. The average harmonic structure was used to extract harmonic features of each noise sample. Subsequently, the KNN algorithm was employed to classify the vehicle type of each noise sample. Based on the preliminary results of this paper, a fixed ratio of the correct classification samples is preferred, probably because the total number of the correct classification data is limited. This paper shows the feasibility of classifying vehicle type in a vehicular noise surveillance system on the roads.

VI. ACKNOWLEDGMENT

This work is supported by the National Environment Agency of Singapore.

REFERENCES

- [1] D. Ouis, "Annoyance from road traffic noise: A review," J. Environmental Psychology, vol. 21, no. 1, pp. 101-120, 2001.
- [2] E. Öhrström, "Effects of low levels of road traffic noise during the night: A laboratory study on number of events maximum noise levels and noise sensitivity," *J. Sound Vib.*, vol. 179, no. 4, pp. 603-615, 1995.
- [3] C. M. Harris, *Handbook of Acoustical Measurements and Noise Control*, 3rd ed., NY: McGraw-Hill, 1991.
- [4] C. Shi and W. S. Gan, "Product directivity models for parametric loudspeakers," *J. Acoust. Soc. Am.*, vol. 131, no. 3, pp. 1938-1945, 2012.
- [5] C. Shi and W. S. Gan, "Using length-limited parametric source in active noise control applications," in *Proc. 20th Int. Congr. Sound Vib.*, Bangkok, Thailand, 2013.
- [6] H. Houzu, T. Tanaka, I. Sakamoto, T. Murakami, M. Ishihama, and K. Sawatari, "Hybrid source localization method to detect noisy vehicles traveling on the road," in *Proc. InterNoise 2009*, Ottawa, Canada, 2009.
- [7] H. Houzu, I. Sakamoto, T. Tanaka, T. Murakami, M. Ishihama, and K. Sawatari, "Development of sound source localization system to detect noisy vehicles from traffic flow," in *Proc. InterNoise 2011*, Osaka, Japan, 2011.
- [8] J. W. Hsieh, S. H. Yu, Y. S. Chen, and W. F. Hu, "Automatic traffic surveillance system for vehicle tracking and classification," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 2, pp.175-187, 2006.
- [9] G. P. Succi, T. K. Pedersen, R. Gampert, and G. Prado, "Acoustic target tracking and target identification: recent results," in *Proc. SPIE 3713*, Orlando, FL, 1999.
- [10] T. L. Tung, Kung Yao, "Classification of vehicles using nonlinear dynamics and array processing," in *Proc. SPIE 3807*, Denver, CO, 1999.
- [11] A. Z. Averbuch, V. A. Zheludev, I. Kozlov, "Wavelet based algorithm for acoustic detection of moving ground and airborne targets," in *Proc. SPIE 4419*, San Diego, CA, 2000.
- [12] N. Srour, "Back propagation of acoustic signature for robust target identification," in *Proc. SPIE 4232*, Boston, MA, 2000.
- [13] B. Guo, M. S. Nixon, and T. R. Damarla, "Acoustic information fusion for ground vehicle classification," in *Proc. 11th Int. Conf. Information Fusion*, Cologne, Germany, 2008.
- [14] G. Watts, "Identifying the characteristics of noisy vehicles at the roadside," in *Proc. InterNoise 2011*, Osaka, Japan, 2011.
- [15] J. Eargle, The Microphone Handbook. 1st ed., NY: Elar, 1981.

- [16] T. Merkel, "How to receive sound through ultrasound: the interaction between ultrasonic and audio waves in the air," in *Proc. 16th Int. Congr. Sound Vib.*, Cracow, Poland, 2009.
- [17] C. Shi and W. S. Gan, " Modeling the directivity of parametric loudspeaker," in *Proc. of 19th Int. Symp. Nonlinear Acoust.*, Tokyo, Japan, 2012.
- [18] C. Shi and W. S. Gan, "Analysis and calibration of system errors in steerable parametric loudspeakers," *Appl. Acoust.*, vol. 73, no. 12, pp. 1263–1270, 2012.
- [19] C. Shi, W. S. Gan, Y. K. Chong, "Calibration of parametric acoustic array," in *Proc. 8th Int. Conf. Inform. Commun. Signal Process.*, Singapore, 2011.
- [20] C. Shi and W. S. Gan, "Grating lobe elimination in steerable parametric loudspeaker," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 58, no. 2, pp. 437-450, 2011.
- [21] C. Shi and W. S. Gan, "On grating lobe elimination of difference frequency in parametric loudspeaker," in *Proc. 2nd APSIPA Annu. Summit Conf.*, Singapore, 2010.
- [22] Z. Duan, Y. Zhang, C. Zhang, and Z. Shi, "Unsupervised singlechannel music source separation by average harmonic structure modeling," *IEEE Trans. Audio Speech Language Proc.*, vol. 16, no. 4, pp. 766–778, 2008.
- [23] D.T. Larose, "k-Nearest neighbor algorithm," in *Discovering Knowledge in Data: An Introduction to Data Mining*, NJ: John Wiley & Sons, 2005.