Proposal for a Compressive Measurement-Based Acoustic Vehicle Detection and Identification System

Billy Dawton*, Shigemi Ishida*, Yuki Hori*, Masato Uchino*, Yutaka Arakawa*
*Graduate School/Faculty of Information Science and Electrical Engineering, Kyushu University, Japan
Email: {bdawton, hori, uchino}@f.ait.kyushu-u.ac.jp, {ishida, arakawa}@ait.kyushu-u.ac.jp

***Abstract*—As society becomes increasingly interconnected, the need for sophisticated signal processing and data analysis techniques becomes increasingly apparent, particularly in the field of Intelligent Transportation Systems (ITS) where various sensing applications generate data at an exponential rate. In this paper, we put a forward a compressive sensing-based system to extract information from passing vehicle sounds sampled at sub-Nyquist rates for Acoustic Vehicle Detection and Identification (AVDI) applications. The obtained compressive measurements are used to detect and identify passing vehicles. Initial evaluation performed using data obtained from roads on a university campus presents an accuracy of 86.2 % with a back-end ADC sample rate of 3 kHz.**

***Index Terms*—Intelligent Transportation Systems (ITS), Acoustic Vehicle Detection, Compressive Sensing (CS), Feature Extraction.**

I. Introduction

The past years have seen a marked increase in the development of Intelligent Transportation System (ITS) technologies. A growing number of novel applications such as smart navigation, traffic monitoring, and road safety have been accompanied by a corresponding improvement in overall system performance and efficiency.

This increase in performance comes with an increase in computational cost and complexity, requiring more data and processing power than ever before. This is particularly apparent in traffic monitoring applications, where the methods used for the detection and identification of vehicles often come with high computational and installation costs. In an effort to mitigate this, low-cost, low-complexity vehicle detection systems based on acoustic sensors have been proposed. Most recently, the authors have presented a stereo microphone-based vehicle detection and identification system in [1] and a sound map-based sequential vehicle detection system in [2].

Despite the low installation costs associated with acoustic sensing, the subsequent analysis and leveraging of the acquired data is often costly in terms of computational requirements, reducing overall system efficiency and restricting the deployment of such systems in power-critical applications. Indeed, current Acoustic Vehicle Detection and Identification (AVDI) systems usually contain a stage presenting relatively high computational complexity: this occurs either prior to the initial detection or classification stage like in [1] or [3] where the use of successive DWTs or DFTs are used to analyze and process the data, or during the classification stage itself where complex supervised learning methods such as deep neural networks (DNN) [4] or Multilayer Perceptrons (MLPs) [5] are employed. The computational cost associated with these stages somewhat mitigates the reduction in complexity entailed by the use of acoustic sensing methods.

In this paper, we look to improve upon existing AVDI methods by exploiting the nature of the signals under consideration to sample them at sub-Nyquist rates, reducing the amount of data and computational complexity involved at each stage of the detection and identification processes. To achieve this, we use a technique called compressive sensing (CS), first presented in [6], that enables the reconstruction of sparse or compressible signals from a reduced set of linear, non-adaptive measurements.

We propose an acoustic sensing system taking advantage of the dimensionality reduction properties of CS to acquire vehicle signals at sub-Nyquist rates and uses them in conjunction with a range of machine learning techniques for use in AVDI applications; to the best of our knowledge, our proposed system is the first of its kind. Our research aims to lower the overall computational cost, complexity, and power consumption when compared to existing setups whilst maintaining high classification accuracy.

II. Related Work

Vehicle detection and identification using features extracted from vehicle audio in tandem with supervised learning has been widely explored. Methods using Support Vector Machine (SVM) [7] classifiers, k-Nearest Neighbor (KNN) classifiers [8], Gaussian Mixture Models, and Hidden Markov Models [9] along with the frequency domain information of vehicle signals have been proposed. Whilst these systems share a similar goal and basic approach, they differ in their applications, performance and features.

A method for identifying passing vehicles based on the frequency-domain shape of their sound signature is put forward in [10]. The unique shape of each passing vehicle's envelope enables the system to accurately distinguish individual vehicles. However, this same uniqueness makes it impossible for the system to identify the type (i.e. the class label) of a passing vehicle.

In [5], a system capable of analyzing the acoustic signature of vehicles independently of any changes in engine speed is presented. By using wavelet packet analysis instead of more traditional time or frequency domain-based techniques and a MLP classifier, the system is able to extract engine

This is the author's version of the work.
© 2020 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.
doi: 10.1109/VTC2020-Fall49728.2020.9348569

speed-independent features from the sounds of passing vehicles. Whilst this improves system accuracy performance in a range of sensing environments, the computational and hardware requirements entailed by the use of a neural network make it difficult to deploy the system in low-power low-cost situations.

The authors have, in previous works, proposed several acoustic vehicle detection systems. SAVeD, the sequential acoustic vehicle detector put forward in [2] works by fitting S-curve models to points on a sound map using a random sample consensus (RANSAC) estimation method. The obtained system F-measure is 83 %. In [1], the authors designed a stereo microphone-based system which identifies passing vehicles based on frequency-domain features extracted from their sound signature. By time-shifting and combining the two signals to produce an emphasized sound signal, vehicle type estimation accuracy is improved, particularly when faced with simultaneously and successively passing vehicles. The obtained system accuracy is 95 %. Whilst both the above systems perform well when compared to existing microphone-based detection methods, the computational cost associated with the two methods is high and makes low-power, embedded applications difficult.

The authors also propose in [11] an ultra low-power vehicle detector (ULP-VD) capable of detecting passing vehicles with minimal computational cost; this system however is only able to detect the presence of a vehicle and must be used in combination with other techniques to identify them.

Traditionally, digital signal processing techniques are performed on a full set of samples acquired by sampling an analog signal at the Nyquist rate. In [12] the concept of using CS as a tool for signal processing on samples acquired at sub-Nyquist sample rates is explored. The authors find that it is possible to successfully perform a variety of processes including filtering, detection and classification directly on a reduced set of linear samples, without reconstructing the signal beforehand.

In [13], it is shown that it is viable to use the linear measurements as features in machine learning applications with only minimal pre-processing: by carefully selecting the sensing parameters, the authors demonstrate that it is possible to obtain enough relevant signal information to perform fault detection on industrial solenoids.

In [14] a license plate recognition system which uses an SVM classifier to identify plate numbers by sub-sampling the sparse, flattened 1-D representations of images obtained from traffic monitoring cameras is put forward.

The above methods serve to illustrate the viability of using CS measurement-based sensing systems for a wide variety of applications, whilst simultaneously highlighting the flexibility of acoustic detection methods.

III. Proposed System

A. System Overview

Figure 1 shows the average sound signals obtained from passing cars and scooters, and from periods without a passing vehicle: we can see that the overwhelming majority of the frequency content is contained below 6 kHz, and that the different signal classes can be distinguished by the power

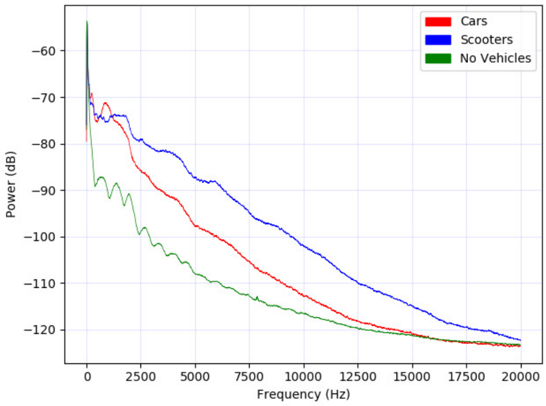

Fig. 1. Average audio signals for three vehicle classes

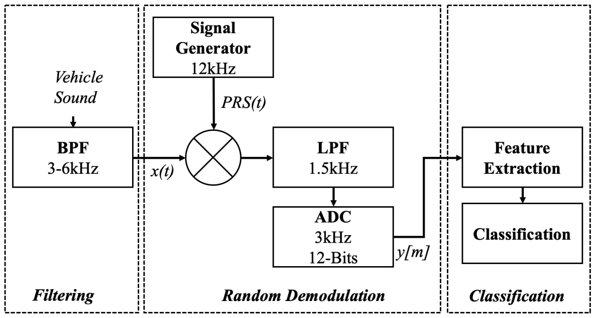

Fig. 2. Proposed system overview

contained in their respective frequency components from 3 kHz onwards. Rather than sample the signals at the Nyquist rate, our proposed system uses a combination of filtering and compressive sampling techniques to acquire the information located in the 3–6 kHz frequency band directly at sub-Nyquist rates. The system then detects and identifies passing vehicles using features extracted from the samples acquired in this manner.

The system can be seen in Figure 2 and is made up of three stages: filtering, random demodulation, and classification. The filtering stage consists of a single band-pass filter (BPF) operating over the 3–6 kHz band and removes unwanted frequency content. The input signal is then combined with a random chipping sequence before being sampled at a sub-Nyquist rate in the random demodulation stage. Finally, in the classification stage, features are extracted from the samples obtained in the previous stage and are used as inputs to a classifier for vehicle type detection and identification. The workings of these stages are explained further in Sections III-C and IV-B.

It is important to stress that in our proposed system the front-end band-pass and low-pass filtering and mixing are performed in the analog domain, and that the sampling operation, and thus the digitization of the signal, only occurs at the end of the acquisition process.

B. Compressive Sensing

Traditionally, analog signals are acquired according to the Shannon-Nyquist theorem: uniformly sampled at a rate at least twice as high as the highest frequency present in the signal of interest. Thus, a signal whose highest frequency component is $W/2$ Hz, has a minimum required rate of W Hz. However, we are often only interested in a small part of the information contained within a signal's bandwidth, in which case sampling at the Nyquist rate is inefficient. This is especially true when dealing with signals that have a sparse representation in a given domain (for example, significantly fewer relevant frequency components in a signal than what is permitted by its bandlimit). When dealing with this type of signal, rather than acquiring a full set of samples at the Nyquist rate and subsequently discarding unwanted information, we would like to access the relevant information directly.

In our paper this targeted sampling is achieved using CS, a method for efficiently acquiring sparse or compressible signals (a signal can be called compressible if only a small amount of its non-zero components have significant magnitude) first put forward in [6] and [15].

The procedure is described as follows: a signal $x \in \mathbb{R}^N$ of length T_s can be expressed as a combination of discrete coefficients $\alpha \in \mathbb{C}^N$ and vectors ψ_n which form the columns of an orthonormal basis matrix $\mathbf{\Psi} \in \mathbb{C}^{N\times N}$ for a given time window:

$$x(t) = \sum_{n=1}^{N} \alpha_n \psi_n(t)\ ,\ t \in [0, T_s) \tag{1}$$

Where our coefficients are computed as $\alpha_n =< x, \psi_n >$.

$x(t)$ is sparse in the domain defined by the matrix $\mathbf{\Psi}$, which in our case is a DFT matrix making x sparse in the frequency domain. Given that $\mathbf{\Psi}$ is known to us in advance, we are interested in obtaining the information in the sparse coefficient vector α. If we are able to obtain α, it is straightforward to reconstruct x through the appropriate DFT transform.

We define $y \in \mathbb{R}^M$ as the set of linear measurements obtained by performing a sequence of sampling operations represented by $\mathbf{\Phi} \in \mathbb{R}^{M\times N}$, such that $y = \mathbf{\Phi}x$ and crucially, $N > M$ (N and M are positive integers).

$$y = \mathbf{\Phi}x \tag{2}$$

We also define the product of the measurement and basis matrices as the sensing matrix $\mathbf{\Theta} \in \mathbb{C}^{M\times N} = \mathbf{\Phi\Psi}$.

From (1) and (2):

$$\begin{aligned} y &= \mathbf{\Phi\Psi}\alpha && (3)\\ &= \mathbf{\Theta}\alpha && (4) \end{aligned}$$

CS establishes that if $\mathbf{\Theta}$ satisfies the incoherence and RIP (restricted isometry property) conditions outlined in [16], it is possible to recover α, and thus x, from y with much fewer samples than would be required in traditional Nyquist sampling. The recovery process is typically performed using l_1 minimization.

$\mathbf{\Theta}$: $\mathbb{C}^N \rightarrow \mathbb{C}^M$ can be considered as a dimensionality reduction operator: it enables us to obtain a lower-dimension representation of α, the sparse coefficient vector of interest. If we wish to consider the compressive measurements as low-dimensional representations, or features extracted from an input signal, then certain conditions must be met. In particular, we need to ensure that for any two distinct signals x_1 and x_2, $\mathbf{\Phi}x_1 \neq \mathbf{\Phi}x_2$ in both noiseless and noisy conditions (in the presence of quantization noise for example). For this to be the case our matrix $\mathbf{\Phi}$ must again follow the set of restrictions defined in [16].

C. Random Demodulator

Initial theoretical work on CS only considers discrete signals, however our proposed system looks to obtain continuous-time audio signals which have a sparse or compressible representation in the frequency domain. In addition, to fully benefit from the advantages of CS, we are constrained to acquiring the signal of interest using exclusively analog components. To that end, our system takes inspiration from an architecture developed by Tropp et al. in [17] called the Random Demodulator (RD), which allows for analog signals to be used in CS applications and can be designed using exclusively analog components. The intuition behind the system is as follows: instead of sampling a continuous-time input signal at the Nyquist rate, the RD modulates the signal with a random chipping sequence, spreading the sparse input signal across the entirety of the frequency spectrum. This smeared signal is then low-passed before being sampled at a sub-Nyquist rate, and the original signal is obtained from these samples via a recovery algorithm.

Our analog signal described in (1) is combined with a pseudo-random chipping sequence defined as:

$$PRS(t) = \sum_{n=0}^{W-1} \epsilon_n(t)\ ,\ t \in \left[\frac{n}{W}, \frac{n}{W}+1\right) \tag{5}$$

ϵ_n is a random sequence which switches between ± 1 with equal probability (Rademacher sequence) at or above $x(t)$'s Nyquist rate.

The combined signal $x(t).PRS(t)$ is passed through an LPF $h(t)$ and sampled at a rate R below the Nyquist rate W with $R < W$ to obtain linear compressive samples $y[m]$. In the time domain, this corresponds to a multiplication followed by a convolution:

$$\begin{aligned} y[m] &= \int_{-\infty}^{\infty} x(\tau)PRS(\tau)h(t-\tau)d\tau \Big|_{t=mR} && (6)\\ &= \sum_{n=1}^{N} \alpha_n \int_{-\infty}^{\infty} \psi_n(\tau)PRS(\tau)h(mR-\tau)d\tau && (7) \end{aligned}$$

From which we can obtain the expression for the sensing matrix $\mathbf{\Theta}$, where each entry is defined as $\theta_{m,n}$ for row m and column n.

$$\theta_{m,n} = \int_{-\infty}^{\infty} \psi_n(\tau)PRS(\tau)h(mR-\tau)d\tau \tag{8}$$

The random demodulation process is represented by $\mathbf{\Phi}$. It is shown in [17] that if $\mathbf{\Psi}$ is a DFT matrix, then $\mathbf{\Theta}$ satisfies the RIP conditions as long as the amount of measurements M follows:

$$M = O\left(K \log \frac{W}{S}\right) \tag{9}$$

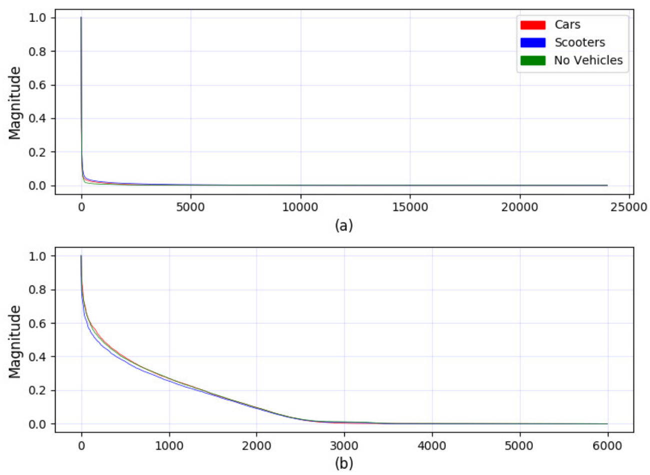

Fig. 3. Sorted coefficients α_n *(rescaled)* for: (a) original input signals, and (b) filtered input signals

Where S is the sparsity level of the input signal.

The RD in our proposed system is subject to two modifications. First, the presence of the BPF operating over the 3–6 kHz band at its input, whose role is to remove redundant sub-3 kHz information from the signals, improving classification performance, and low-pass the signal to 6 kHz reducing the signal's bandwidth and lowering the required chipping sequence frequency to 12 kHz.

Second, the absence of a reconstruction stage: we are not looking to reconstruct x and will instead extract features directly from y for use in classification, reducing the computational load of the system by bypassing the computationally expensive reconstruction phase.

D. Signal Model

An important consideration in CS applications is the input signal model. As previously mentioned, CS is applicable when the signal to be acquired is sparse or compressible. In reality, signals are only very rarely completely sparse, and are much more often compressible or approximately sparse. More formally, this means that the magnitude of their non-zero coefficients decay following a power law distribution of the type:

$$\alpha_n = Cn^{-p} \text{ for } n \in \{1; N\} \tag{10}$$

Where C and p are constants.

Figure 3 shows the coefficient distributions of (a) the original input signals, and (b) the filtered input signals. The original signals are compressible as their spectra are clearly dominated by a relatively small amount S of high magnitude coefficients. The filtered signals, due to the reduced size of their bandlimit, do not present the same degree of compressibility. This is due to the number of S high magnitude components representing a higher proportion of the total amount of components in these filtered signals. Whilst this is a departure from the signal model defined in CS literature, we are not in our case constrained by the usual recovery process requirements on signal sparsity and compressibility, and so assume that the filtered input signal model is fit for purpose in subsequent analysis.

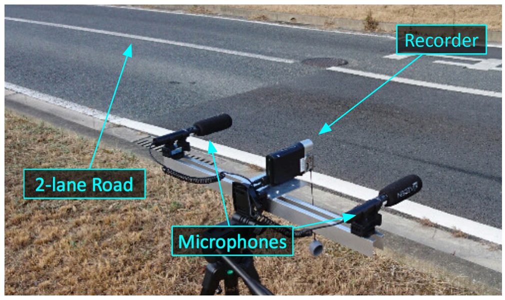

Fig. 4. Experimental setup

IV. Evaluation

An initial software-based evaluation of our proposed system is performed using audio data collected from the roads on a university campus.

A. Data Acquisition

The data acquisition setup can be seen in Figure 4. Two AZDEN SGM-990s microphones set to "cardioid" pickup pattern recording at a sample rate of 48 kHz and bit depth of 16 bits are installed at the side of a two-lane two-way road at a height of 1m from the ground, parallel to the road and connected to a SONY HDR-MV1 video camera. The microphones record the sound of passing vehicles for a duration of 20 minutes and the video camera records the ground truth video footage. The intra-microphone distance is 50 cm, the distance between the microphones and the center of the front lane is 3 m, and the distance between the microphones and the center of the back lane is 6 m. The signals at the two microphones are averaged to obtain a single-channel mono signal for use in subsequent analysis. Vehicle sounds were recorded on two separate occasions under the same conditions, with one set of data being used as a training set, and the other as a testing set. Classification was performed for 3 classes: cars, scooters/motorbikes, and no passing vehicle; referred to as: "Car", "Scooter", and "NoVeh" respectively. During this initial evaluation, we are only looking to perform classification on individual, non-overlapping vehicle sounds.

We set the time at which a vehicle passes in front of the mid-point between the microphones is t_p, and define the range $T_r = \left[t_p - \frac{T_s}{2} ; t_p + \frac{T_s}{2}\right]$ where $T_s = 2s$. From the t_p information obtained from the ground truth data, we extract the "Car" and "Scooter" signals whose T_r do not overlap with that of preceding or following vehicles. We obtain 40 "Car" and 57 "Scooter" signals from the first set, and 52 "Car" and 50 "Scooter" signals for the second set. Splitting the parts of the signal who do not correspond to the T_r of any vehicle into sections of length T_s gives us 115 "NoVeh" signals for the first set and 112 for the second for a total of 212 and 214 signals across the three classes under consideration for the first and second sets respectively.

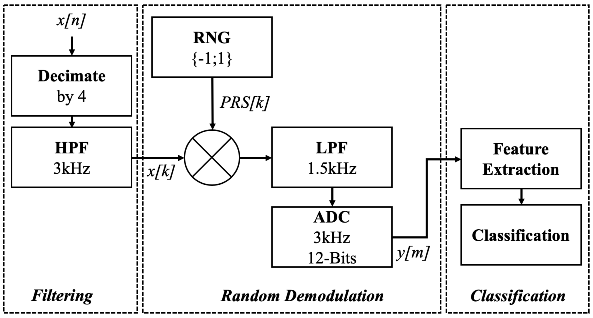

Fig. 5. System software implementation overview

TABLE I
SYSTEM PARAMETERS.

W	R	B	T_s	N	K	M
48 kHz	3 kHz	12 bits	$2s$	96000	24000	6000

B. System Simulation

1) Overview: We described in Section III-A the real-world analog implementation of the system. In this section, we perform an initial evaluation of the system by simulating its operation using a software-based digital-domain representation. Figure 5 shows the software implementation of the system proposed in Figure 2.

$x[n] \in \mathbb{R}^N$ is a discrete version of the analog input signal $x(t)$ sampled at 48 kHz. The input signal is decimated by a factor of 4, bringing the Nyquist rate down to 12 kHz and high-pass filtered at 3 kHz through a Type II Chebyshev filter. The resulting signal $x[k] \in \mathbb{R}^K$ is shown in Figure 6 for each average vehicle type. $PRS[k] \in \mathbb{R}^K$ is a vector containing an equiprobable random distribution from the set $\{-1.1\}$. The pre-ADC LPF is a 2nd order Butterworth filter and the measurements $y[m] \in \mathbb{R}^M$ are obtained by uniformly sampling and quantizing every $\frac{N}{R}$th entry from the combined $x[k].PRS[k]$ signal at rate $R = \frac{12kHz}{4}$ and bit-depth B. Compared to the experimental sample rate of 48 kHz, the reduction in sampling rate is $(\frac{48kHz}{12kHz})(\frac{12kHz}{3kHz}) = \frac{N}{M} = 16$.

This simultaneous filtering operation has the effect of further sparsifying the input signal by suppressing unwanted information contained in the signals' lower frequency range, whilst also performing anti-aliasing by attenuating the frequencies above the signal's Nyquist frequency of 6 kHz. As a result, in the unattenuated 3-6 kHz band, the frequency information of each different signal class is clearly distinct.

Figure 7 shows the linear samples $y[m]$ and their frequency domain representations. The action of the proposed system causes a distinct separation between the signal types, observable in both the frequency-domain representation of the $y[m]$ samples, as well as the $y[m]$ samples themselves.

2) Feature Extraction: We perform classification on a set of features extracted from the $y[m]$ measurements. Selecting a reduced set of relevant features for classification improves system performance by removing redundant information, mitigating the effects of overfitting, and reducing the system's computational cost.

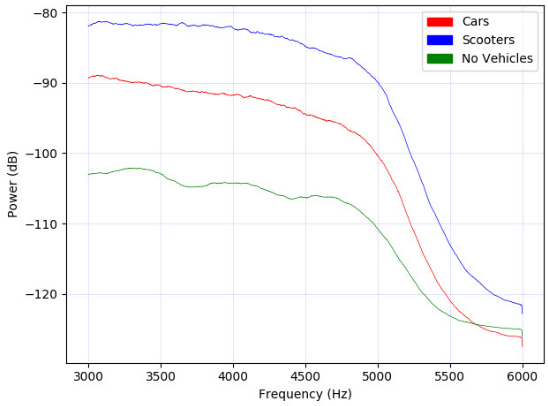

Fig. 6. Filtered average audio signals for three vehicle classes

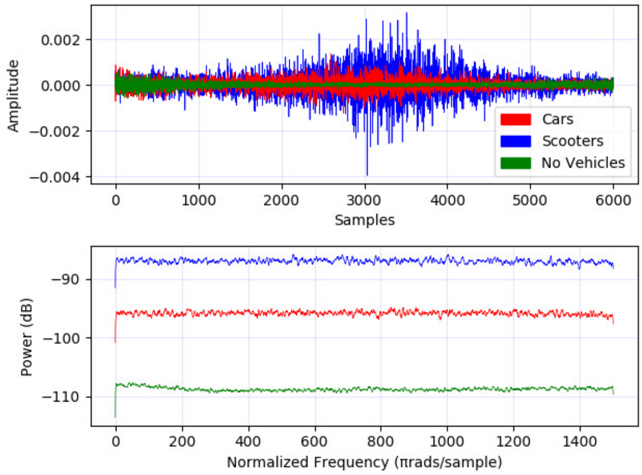

Fig. 7. Linear measurements of average audio signals for three vehicle classes

The 9 following features are extracted from $y[m]$: *mean*, *standard deviation*, *median*, *absolute max value*, *peak-to-peak range*, *interquartile range*, *percentage of data in 1st standard deviation*, *percentage of data between 2nd & 1st standard deviation*, *percentage of data between 3rd & 2nd standard deviation*

Prior to classification, the extracted feature data is randomly undersampled to obtain classes with equal amounts of entries.

C. Classification Results

A random forest classifier, chosen for its outlier robustness and low pre-processing requirements (no rescaling of input data), is trained on the first dataset and tested on the second. Results are averaged over 100 runs to obtain an average system accuracy of 86.2 %. In this initial evaluation, the detection and identification processes are performed simultaneously: for an unknown signal of length T_s, our system determines whether or not a vehicle is passing, and if so, its type.

Figure 8 shows that system identifies "Scooter" signals with high accuracy, but is less effective at identifying "Car" and "NoVeh" signals. We can see on Figure 1 that over the 3–6kHz band the signal power of "Scooter" signals is

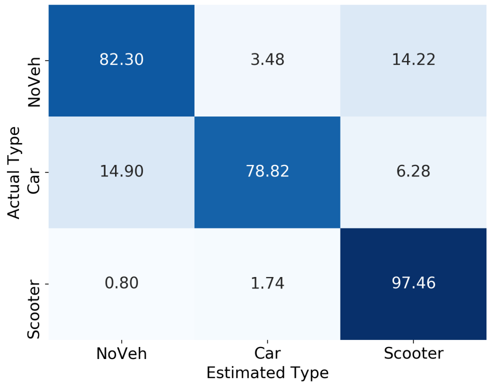

Fig. 8. System confusion matrix

higher than that of the other two, which translates to higher amplitude and crucially, higher variance in the corresponding linear samples. These $y[m]$ differences combined with the choice of feature set contribute to the model's bias towards the "Scooter" class. Additionally, the balanced datasets used during the classification process are relatively small in size; repeating the experiment with more data could help the system differentiate between the two classes, as well as giving us more certainty regarding its accuracy performance.

V. Discussion

We gauge the performance of our system by comparing it to those offering similar functionality. The system in [2] shows a vehicle detection accuracy of 71 %, obtained by comparing the signals obtained at both microphones of the stereo pair. For a $T_s = 2s$ time interval, sampling at 48 kHz results in $N = 192000$ points. This process is followed by the use of RANSAC, an optimization algorithm whose complexity increases which each successive iteration. The system in [1] also initially obtains $N = 192000$ points on which are performed the subsequent detection and identification processes through a succession of overlapping 4096-point FFTs. Obtained accuracy is 95 %. Initial results show a proposed system accuracy of 86.2 % with $N = 6000$ points processed over the same T_s interval. Operations are limited to simple mathematical processes for feature extraction, with detection and identification being done using random forest classification. This comparably high accuracy score, coupled with a significant reduction in computational requirements serves to highlight the viability of our proposed system in AVDI applications.

VI. Conclusion

This paper serves as an initial evaluation of a compressive measurement-based method for detecting and identifying vehicles based on their sound signature. We sought to improve upon existing AVDI methods by exploiting the nature of the signals under consideration to sample them at sub-Nyquist rates, reducing the amount of data and computational complexity involved at each stage of the detection and identification process. To this end, we designed and created a CS-based system with a pre-processing stage consisting only of successive filtering and mixing, and that performs classification on easy-to-extract features using a simple machine learning classifier. A software implementation of the proposed system capable of classifying different vehicle sounds with an accuracy of 86.2 % and a back-end ADC sample rate of 3 kHz. The proposed system shows an accuracy comparable to that of existing systems (71 % and 95 % in [2] and [1] respectively) whilst minimizing computational requirements. Future work includes fine-tuning the existing system, in particular removing the bias towards the "Scooter" class, adding additional functionality to the system, such as a separate vehicle presence detection or a steady-state noise reduction stage, and finally working towards a hardware implementation of the system.

Acknowledgments

This work was supported in part by JSPS KAKENHI Grant Number JP18K18041 and the Cooperative Research Project Program of RIEC, Tohoku University.

References

[1] B. Dawton, S. Ishida, Y. Hori *et al.*, "Initial evaluation of vehicle type identification using roadside stereo microphones," in *Proc. IEEE Sens. Appl. Symp. (SAS)*, Mar. 2020, pp. 1–6.

[2] S. Ishida, J. Kajimura, M. Uchino *et al.*, "SAVeD: Acoustic vehicle detector with speed estimation capable of sequential vehicle detection," in *Proc. IEEE Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2018, pp. 906–912.

[3] H. Göksu, "Vehicle speed measurement by on-board acoustic signal processing," *Meas. Control*, vol. 51, no. 5–6, pp. 138–149, May 2018.

[4] X. Liu, E. Gönültaş, and C. Studer, "Analog-to-Feature (A2F) conversion for audio-event classification," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2018, pp. 2275–2279.

[5] H. Göksu, "Engine speed-independent acoustic signature for vehicles," *Meas. Control*, vol. 51, no. 3–4, pp. 94–103, Apr. 2018.

[6] E. J. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[7] A. Aljaafreh and L. Dong, "An evaluation of feature extraction methods for vehicle classification based on acoustic signals," in *Proc. IEEE Int. Conf. Netw. Sens. Control*, Apr. 2010, pp. 570–575.

[8] Z. Changjun and C. Yuzong, "The research of vehicle classification using SVM and KNN in a ramp," in *Proc. Int. Forum Comput. Sci.-Technol. Appl.*, Dec. 2009, pp. 391–394.

[9] M. E. Munich, "Bayesian subspace methods for acoustic signature recognition of vehicles," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, Sep. 2004, pp. 2107–2110.

[10] S. S. Yang, Y. G. Kim, and H. Choi, "Vehicle identification using wireless sensor networks," in *Proc. IEEE SoutheastCon*, Mar. 2007, pp. 41–46.

[11] K. Kubo, C. Li, S. Ishida *et al.*, "Design of ultra low power vehicle detector utilizing discrete wavelet transform," in *Proc. ITS AP Forum*, May 2018, pp. 1052–1063.

[12] M. A. Davenport, P. T. Boufounos, M. B. Wakin *et al.*, "Signal processing with compressive measurements," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 445–460, Apr. 2010.

[13] C. Knoebel, H. Wenzl, J. Reuter *et al.*, "A compressed sensing feature extraction approach for diagnostics and prognostics in electromagnetic solenoids," in *Proc. Annu. Conf. Progn. Health Manage. Soc. (PHM)*, Oct. 2017, pp. 16:1–16:6.

[14] A. Jokić and N. Vuković, "License plate recognition with compressive sensing based feature extraction," *arXiv preprint*, pp. 1–4, Feb. 2019, arXiv:1902.05386 [cs.CV].

[15] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[16] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[17] J. A. Tropp, J. N. Laska, M. F. Duarte *et al.*, "Beyond Nyquist: Efficient sampling of sparse bandlimited signals," *IEEE Trans. Inf. Theory*, vol. 56, no. 1, pp. 520–544, Jan. 2010.