Engine Classification Using Vibrations Measured by Laser Doppler Vibrometer on Background Surfaces[1, 2]

J. Wei*[a], Chi-Him Liu[a], Zhigang Zhu[a], Karmon Vongsy[b], Olga Mendoza-Schrock[b]
[a]Department of Computer Science, The City College of New York, New York, NY 10031
[b]Air Force Research Laboratory, Wright-Patternson Air Force Base, Fairborn, OH 45324

ABSTRACT

In our previous studies, vehicle surfaces' vibrations caused by operating engines measured by Laser Doppler Vibrometer (LDV) have been effectively exploited in order to classify vehicles of different types, e.g., vans, 2-door sedans, 4-door sedans, pick-ups and buses, and different types of engines, such as Inline-four engines, V-6 engines, 1-axle diesel engines and 2-axle diesel engines. The results are achieved by employing methods based on a great array of machine learning classifiers such as AdaBoost, random forests, neural network and support vector machines. However, to achieve effective Intelligence, Surveillance and Reconnaissance (ISR) in utmost interest to military applications, signals directly picked up from vehicle surfaces are problematic: in a contested environment, enemies can intentionally change or conceal vehicle surfaces and thus compromising the vibration signals taken by the LDV thus counter-measuring the efficacy of the LDV's remote sensing prowess. We hence need a more reliable approach to pick authentic vibrations of vehicle engines from a trustworthy surface, not the ones fully controlled by enemies. Toward this end, we propose to pre-plant a number of retro-reflective and well-vibrating small objects such as flat metal sheets/bars along the road, so that when the vehicles move close to these objects their engines will cause the objects' surface to vibrate, which in turn will be picked up by the LDV. Compared with vibrations directly taken from the uncooperative vehicle surfaces that are rigidly connected to the engines, these vibrations are much weaker in magnitudes. However, the fact that they are difficult to be contaminated by enemies makes it an exceedingly appealing approach. In this work we conducted a systematic study on different types of objects that could be pre-planted in the environment. We tested different types of engines ranging from shavers, electric fans, and coffee machines over different surfaces such as white board, cement wall, and steel cases to investigate the characteristics of the LDV signals on these surfaces, in both the time and spectral domains. Preliminary results in engine classification using several machine learning algorithms point to the right direction on the choice of type of object surfaces to be planted for LDV measurements. This method has great potential to be exploited in contested, uncooperative environment for a more effective ISR.

Keywords: Laser Doppler Vibrometry (LDV) sensors, Intelligence, Surveillance and Reconnaissance (ISR), classification algorithms, Neural network, random forests, cross-validation

1. INTRODUCTION

In Intelligence, Surveillance and Reconnaissance (ISR) applications that is of utmost importance for military institutions and law enforcements agencies, Laser Doppler Vibrometry (LDV) sensors have gained increasing popularity.

The benefits provided by use of LDV sensors can be summarized below.

1) *Non-contact and non-invasive measurement*: no mass or pressure is ever applied during LDV's measurement process, hence the long-range sensing is achieved without even being noticed. The laser beams employed by most LDVs are mostly eye safe.

2) *High spatial and spectral resolution in a long range*: the wide range of amplitudes and frequencies offered by LDV sensing data from a relatively long distance, to as far as 100 feet, confer researchers and developers

---

[1] This work is released by Air Force Research Laboratory with PA approval # 88ABW-2014-5297.
[2] This work is supported by NSF I/UCRC Center for Surveillance Research, the Wright State University Site.

valuable information to work on in both spatial and frequency domains for intensive analysis, classification and clustering.

Recent promising applications of LDV for explosive detection by Adams and colleagues [1], have been the cover-page news of national news for its non-invasive nature and precision attained, which could revolutionize the homeland security and international security against crimes and terrorism. LDV's applications go far more than this: it has been widely used in art museums to protect and detect possible cracks from valuable frescos [2], used by civil engineers for effective railway inspection and building structure inspections have been reported by engineering researchers across the world [3].

In this work being sponsored by NSF I/UCRC center for surveillance research, the main theme is to examine means to classify suspicious vehicles in a manner that is extremely hard, if not impossible, to counter-measure. Many previous efforts have been made to address the issues of classifying vehicles making use of LDV data with varying degrees of success. In [4] the auto-correlation function of LDV signals was employed as the workhorse to distinguish engine type, speed, and number of cylinders with impressive precision. Averbuch and colleagues developed a diffusion map based framework to detect moving vehicles based on wavelet packets within the dynamic programming framework [5]. In [6], a prototype automatic vehicle classification system was developed, where a grid of accelerometers are installed on roadways to characterize road vibrations and the number of axles is classified accordingly. The corresponding performance compared with the ground truth is exceedingly valuable at about 99%. Note here accelerometer data is used instead of LDV data in the classification. The relevance to our LDV sensory data based classification will be revealed later. In [7] the original Mel-frequency cepstral coefficients (MFCC) after a principal component analysis (PCA) dimension reduction was fed to a forward Neural network to classify different Arabic speakers. In [8], vehicle operating conditions were classified using 11 extracted features, including MFCC and others such as zero-crossings, dominant frequencies and Flux. In [9], engine classifications for parked vehicles of four different types of engines were achieved by our group, where some initial success has been evidence in the use of the new tone-pitch vibration index and neural network.

One major hurdle blocking the effective utility of existing ISR approaches is the ease of counter-measure from the adversaries: if they know their vehicle surfaces are to be measured by the LDV sensors, they may put on some special materials (plastics boards) or treatments (thick paints) over their vehicle surfaces which can significantly compromise the signals measured by the laser points of the LDV. As one crucial objective of our on-going ISR researches using LDV sensors for reliable target detection, we endeavor to find means for classification without worrying about the counter-measures by adversaries. One way to do it is to not directly measure the vehicle surfaces that are under full control of the adversaries, instead, the LDV sensors should only pick up signals from surfaces owned and pre-planted in the environment by the detectors with full knowledge of the surface physics and vibration phenomenology that may arise over the surface. However, it remain unknown what type of surfaces could be employed toward this end. In this paper, experiments and performances are reported and analyzed on the different choices of surfaces for engine classification purpose, which has a crucial role to play in our eventual choice of external objects to be put on the field for ISR utilities.

In the next section, the representational index of LDV signals, the spectral tone-pitch vibration index, and the rationale behind it are described. The methodology for data collection and machine learning strategies employed in this study are presented in Sec. 3. Sec. 4 concludes this paper with more remarks.

2. SPECTRAL TONE-PITCH VIBRATION INDEXING FOR LDV MEASUREMENTS

Because of the immense success of speech recognition and music encoding [10], techniques such as MFCC and short-term Fourier transform, which have been found to be exceedingly useful in speech representations, were used to encode and index LDV signals in most state-of-the-art analysis approaches, such as those briefly reviewed in Sec. 1. However, to effectively exploit the LDV signals for ISR purposes, the literal use of speech coding and recognition methods are problematic since they are carefully tailored to take advantage of human auditory systems (HAS). In vehicle classification applications, HAS has no role to play, and the careful exploitations achieved by approaches such as MFCC is irrelevant. Instead, the vibration data collected by LDV sensors should be treated as a sequence of physical data or time series: the running vehicle engine is the periodic vibrating source propagating/dissipating its energy over the rigid surface of vehicles as vibrational waves, what a LDV sensor picks up is the vibrations or waves on the vehicle surfaces. Since the measurement point subtended by the LDV sensor is extremely small, in the order of merely micro meters, the surface where the LDV collects the vibrations can be treated mathematically as perfect 2-D planes or sheets, thus in our following mathematical analysis, although the vehicle surfaces are 3-D surfaces, the geometry around the laser measurement points is in essence only of two dimension. The planar wave partial differential equation (PDE) is as below is thus the commanding equation of the nature of the function *u(x,t):*

$$\Delta u = a\, u_{tt} \tag{1}$$

Where Δ is the Laplacian operator, a is a certain constant related to the permeation speed of the surface vibrations on the vehicle surfaces. The default signals measured by LDV sensors are the velocities, i.e., u_t, according to Eq. (1), to correlate the LDV measurements to the vibrations on the vehicle surfaces, the second time derivatives, u_{tt}, the accelerometers, should be used as the proxy for the vibrations on the vehicle surfaces.

To achieve desirable classification results, it is necessary to develop new features by taking account of the special nature of vehicle engines and their interplays with surfaces. After intensive exploratory data analysis of available vehicle engine data, the following special signatures are observed:

1. *Time unit*: to encode the rich details of the operating conditions of vehicle engines with RPM ranging from several hundred up to several thousand, we observed that the RPM resolution with magnitude lower than 60 is hard to be noticed by LDV sensors, that is, the largest possible frequency resolution must be at least 1Hz (60 RPM). Hence a duration with about 1 second must be subtended by the representation to attain adequate resolution of the frequency details of the vehicle engines. We found that a duration of 1.25 seconds provides the optimal performance.
2. We don't find the logarithmic transformation suppression used almost by default in most speech encoding work to be necessarily useful. Conversely, Fourier magnitudes (phase information is discarded in this work) in different bands have different roles, as discussed below.
   a) Very low AC bands (<5 Hz) are mostly caused by factors other than engines, e.g., wind and turn of driving wheel, which should thus be discarded.
   b) Very high AC bands (>120 Hz) are mostly due to random noise: in theory some vehicle may attain 7200 RPM, but those signals are seriously compromised/corrupted by overwhelming noises with low signal to signal ratio due to variations in low AC bands. Fourier coefficients in these bands are thus also dropped.
   c) For relatively high AC bands (in-between 50-120 Hz): the signals are corrupted considerably by noises, yet generally the noises are of smaller magnitudes. Instead of using the original Fourier magnitudes, which over-stressed noises, after trying out suppressing transforms such as square root, cubic root, we found that the logarithmic transform over these relatively higher bands results in the most valuable performance in suppressing noises while keeping useful signals as much as possible.
   d) Some strong signals due to the fundamental frequencies/modes of the vehicle surfaces are standing frequencies, i.e., the peaking magnitudes present repeatedly all through the entire spectra with strong and compactly supported energy. Simply suppressing or even dropping these valuable signals in higher bands will lose valuable information about the vehicle engines and the surfaces. Consequently, in view of the compactness of energy and periodicity of these special phenomenology, a second Fourier transform applies to the *magnitudes* of the first Fourier transform, these compact and periodic signals corresponding to the fundamental frequencies of the vehicle surface are well preserved in the relatively lower bands of the double Fourier transformation domain. The objective of this second application of Fourier transform is similar to the efforts made in music encoding to capture the *pitch* information [11], this part of information saved in our vibration representation is thus called the pitch index. While those encoded in cases b) and c) are the ordinary spectral domain, which is the actual *tone* information.

Based on the preceding exploratory studies, the following **spectral tone-pitch vibration indexing scheme** is formulated below.

1) Basic representative unit is the vibration data d with duration s=1.25 second
2) Apply Fourier transform to d and only keep the magnitudes:

$$F_d = |FFT(d)| \tag{2}$$

3) High frequency detail suppression:

$$F_d(1 : H_{high}) = F_d(1 : H_{high}); \quad F_d(H_{high}+1 : end) = |log(F_d(H_{high}+1 : end))| \tag{3}$$

4) Band passing step:

$$S_d = F_d(H_{high} : 2*H_{high}), \tag{4}$$

5) Apply another Fourier transform on F_d use the band-passed P_d to represent the *pitch information of d:*

$$P_d = |FFT(F_d)(H_{low} : H_{high})\ | \tag{5}$$

6) The vector [S_d, P_d] is the spectral-pitch vibration index of the time series *d* of duration s

From our intensive tests, this spectral domain index carries sufficient information to classify different types of vehicle engines for all datasets in AFRL or CCNY available to us with increasingly challenging properties. We tried out more than ten different ready-made indexes such as MFCC and STFT, combined with different mature classifiers such as kNN (k-Nearest Neighbor), random forests, AdaBoost [12], however, none of them can deliver accuracies consistently higher than 60%. Almost all of them can only yield accuracies 30~40% in our moving vehicle dataset, which is not much better than the random guess 25%. Therefore in this work we only report classification results based on the preceding tone-pitch index as the representation of engine types.

3. EXPERIMENTS AND CLASSIFICATION REPORTS

3.1. The setup and dataset of experimental data for surfaces

The main objective of this study is to determine the type of surfaces that can be used to pick up the vibrations of external engines. In CCNY parking lot, we have tried to use traffic cones, the cement base of lighting poles, the steel base of lighting poles, the stop signs, and surfaces of another car as the external surface to measure the vibrations of the target external vehicles, but in vain: the signals picked up from these various surfaces cannot generate signals strong enough to distinguish the different operating conditions of all vehicles we tested. A more careful choice of surfaces other than those available in the parking lot is thus needed. To resolve this issue, in this lab test, three different types of surfaces inside our lab are used: the surface of a white board, the cement wall, and the surface of a steel cabin. As shown in Fig. 1. For each type of surface, we ran three types of engines: a shaver, a coffee machine and an electric fan from different distances: 1 feet, 3 feet and 5 feet away from the measuring position. And for each surface and each engine, we focus the laser vibe on the reflective tapes (the small red squares in Fig. 1) and record the vibrations impacted on these surfaces by the three engines, each recording lasts for 5 seconds. We would like to see among these three surfaces by use of the preceding tone-pitch indexes, which surface can distinguish the three engines using a classifier. For each experimental layout we recorded for at least multiple times to ensure we have adequate number of data points (>500) to work on. To further test the possible ramifications for signals of different durations, the 5-second recordings are digitally partitioned (with possible overlapping) into segments with 3-second and 2-second duration.

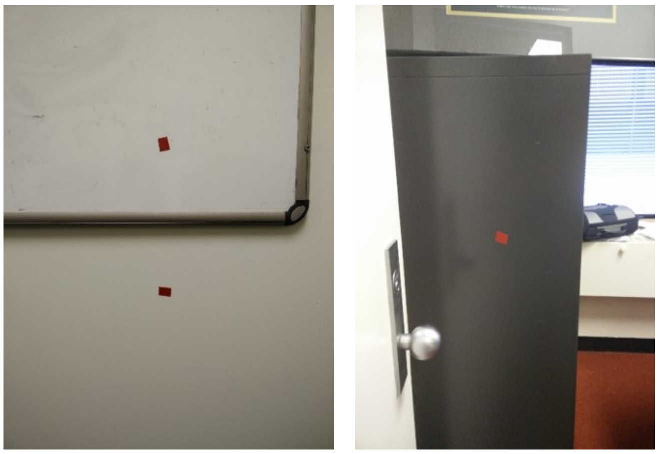

Figure 1. Three surfaces to pick up the external engine signals. Left: upper—white board; lower—cement wall. Right: a steel cabin. Red squares: reflective tapes for laser vibe signal measurements.

3.2. Procedures of classification performance evaluations

In our performance tests, as usual in machine learning [13], the dataset for the three engines and three different surfaces are partitioned into three separate sets for different purposes: the training data: to train the classifiers, cross-validation (CV) set: the data different from the training data, used to tune the parameters for different classifier and find the one with optimal performance, and test set: the data not part of the training and CV data, used to provide the performance of each classifier selected from the CV procedure. In this work, the choices of these three datasets are entirely randomized: 20%, 30% and 50% of the data points collected as described in 3.1 are randomly selected using random partitioning of all the data. We ran this simulation five times, the average performances for all classifiers are reported.

As described in Sec. 2, the duration of the basic unit of our vibration tone-pitch index is 1.25 seconds; however, each data recording segment **d** for all three afore-mentioned datasets is longer than that (2, 3 or 5 seconds). From the data **d** a large set of overlapping sequences of duration 1.25 seconds are first formulated. For a k-second (k=2,3,5 in our tests) measurement, from the i^{th} 1.25-second segment $\mathbf{s_i}$, if the next $(i+1)^{st}$ segment $\mathbf{s_{i+1}}$ is formed by shifting the 1.25-second window by 0.25 second, in total d can give rise to several 1.25-second segment (depending on the value of k), whose tone-pitch indexes are generated as an array of positive data points corresponding to the engine type label the original data **d** subtended. As can be seen, this so-called *framing* procedure in the training phase can generate more training data points to render the attained classifier more reliable. Whereas in the CV and test phase, to determine the label of a given data **d,** we just need to first frame it into an array of 1.25-second overlapping data sequence $\mathbf{s_i}$'s, the class label $\mathbf{l_i}$ of each $\mathbf{s_i}$ is next dictated by the trained classifier. The final label of **d** is arrived at by casting the majority vote of all the associated $\mathbf{l_i}$'s.

To seek out the best classifier(s) for our task, thanks to the MATLAB's various toolboxes such as statistics, image processing, computer vision, and neural network, we can easily call upon a wide array of different classifiers to produce evidential reports of our experimental datasets. There are more than ten different classifiers that are available, after intensive tests, the most competitive ones are used in this work to report the classification performances: they are Adaboost, kNN, random forest, LogitBoost, and Neural network. More mathematical and algorithmic details of these mature classifiers can be readily found in most machine learning textbooks, e.g., [13]. Except the neural network with 20 neurons in the hidden layer, all other classifiers yield accuracies less than 60% in a consistent manner. Since we are only interested in finding the optimal surfaces as a proxy of external engines, we here only tabulate performances produced by the best classifier on this dataset, i.e., the neural network using tone-pitch index.

The test accuracy rates over all three surfaces are summarized in Tab. I.

Table 1. Summary accuracy rates for the three surfaces in classifying the three engines using tone-pitch index and neural network with 20-hidden neurons.

Data	White board	Cement wall	Steel cabin
Test1: 2-second	46	78	90
Test 2: 3-second	51	78	98
Test 3: 5-second	59	82	96

From Table 1 it can be consistently observed that the surface of steel cabin gives rise to the best quality signal as a proxy for external engines due to the following two causes: 1) The surface is metal, similar to that of vehicles, hence the fundamental frequencies as handled by tone-pitch index differ not much from vehicle surfaces. 2) The steel cabin is firmly placed on the ground with no other sources of vibrations except the ones due to the close-by engines. Conversely, the reason why our tests of using a different car to pick up external engines does not work is mostly because of the tires that significantly dampens vibrations from external engines. Surfaces like white board and traffic cones are too sensitive to vibrations other than the targeted engines, where the signals of interest are too seriously compromised to be useful. The hard surfaces like cement wall, steel/cement bases of lighting poles are overly rigid making them less sensitive to external vibrations, in our lab tests, they are better than the white board. In the parking lot tests, they can only pick up very strong vehicle engine vibrations (>3000 RPM).

4. CONCLUDING REMARKS

In this work, after careful examinations of various state-of-the-art speech encoding and recognition techniques, we presented our new spectral tone-pitch vibration index as the content-based concise representation of engine vibrations. Unlike other successful indexing approaches such as MFCC by exploiting the human auditory systems, this new index attain its efficacy by dealing with the spectra of engine's vibrations and the modes of vehicle surfaces. Consequently it has yielded promising results on our LDV measurements for stationary vehicle test data. However, to conduct ISR in situations of great interest to military applications, LDV measurements that are hard to counter-measure are of utmost importance.

In this work, instead of picking up laser measurements directly from the target vehicles, which can be easily changed by putting on extra plastic covers and special paints, which makes our ensuing classification and identification difficult, we set out to take laser measurements from surfaces that are *outside* the adversary vehicles and are pre-planted in the environment and under our control. After trying out a great array of different surfaces, our preliminary tests suggested that the surface of steel cabin can serve as a reliable object planted in the field to pick up close-by engines for classification and identification purposed.

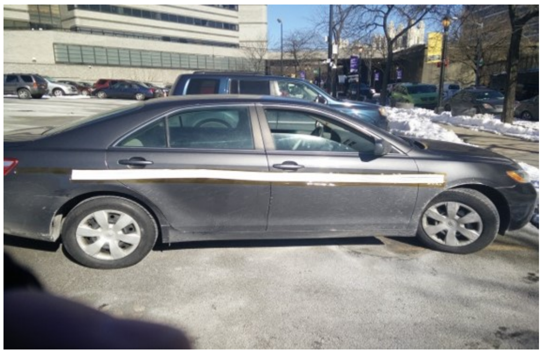

Figure 2 long reflective tapes on the car body to record moving data.

Furthermore, it remains a hard problem to use LDV to directly measure vehicles moving at normal speed since the duration of the measurements are too short and it is difficult to maintain the high quality of the resultant recording—the LDV must be 1) focused in order to ensure the data quality at a fixed distance, and 2) the surface focused by the LDV should have great reflectivity, the reflective tapes—those red squares shown in Fig. 1—are necessary, otherwise lots of noise will be present rendering the recorded signals hard to use. As shown in Fig. 2, in order to record data from moving vehicles, a very long and wide reflective tape has to be put on the car body from the front to the back bumper to ensure the data collection quality. Plus, to make sure the data collected can be up to several seconds, the car has to move extremely slowly, therefore the available moving vehicle data are still a long way from the scenarios of practical utility for military's ISR interest. However, the adversaries' vehicle certainly will not permit us to put these necessary tapes and give us a duration long enough for the tone-pitch index to represent. The use of external objects in the background with reliable surfaces have the potential to avoid the foregoing troubles: the object surface is not only under our full control that the adversary have no way or knowledge to compromise it; furthermore, this surface is not moving and we can freely put reflective tapes to ensure the data quality. In the imminent future we will conduct intensive tests along this line on real moving vehicles in real urban traffic to further inspect the performance of this line of attack.

REFERENCES

[1] C. Watson, J. Rhoads and D. Adams, " Structural Dynamic Imaging Through Interfaces Using Piezoelectric Actuation and Laser Vibrometry for Diagnosing the Mechanical Properties of Composite Materials," in *ASME Dynamic Systems and Control Conference*, Palo Alto, CA., 2013.

[2] P. Castellini, N. Paone and E. Tomasini, "The laser Doppler vibrometer as an instrument for non-intrusive diagnostic of works of art: application to fresco paintings," *Optics & Lasers in Engineering,* vol. 25, pp. 227-246, May 1996.

[3] K. Kubota, "Development of a remote non-contact measurement system combining laser doppler vibrometer and total station for monitoring of structures," in *The 3rd Int. Conf. on Structural Health monitoring of Intelligent Infrastructure*, 2007.

[4] K. Sigmund, S. Shelley, M. Bauer and F. Heitkamp, "Analysis of vehicle vibration sources for automatic differentiation between gas and diesel piston engines," in *SPIE 8391, Automatic Target Recognition XXII*, 2012.

[5] A. Averbuch and K. Hochman, "A Diffusion Framework for detection of moving vehicles," *Digital Signal Process,* vol. 20, no. 1, pp. 111-122, 2010.

[6] W. Ma, Xing, D., A. McKee, R. Bajwa, C. Flores, B. Fuller and P. Varaiya, "A Wireless Accelerometer-Based Automatic Vehicle Classification Prototype," *IEEE Trans. on Intelligent Transportation Systems,* vol. 15, no. 1, pp. 104-111, May 2013.

[7] A. Rashed, "Fast Algorith for Noisy Speaker Recognition Using ANN," *International Journal of Computer Engineering and Technology,* pp. 56-65, 2014.

[8] D. Petrovich, M. Snorrason and M. Stevens, "Identifying Vehicles using Vibrometry Signatures," in *Proceedings of the 16th International Conference on Pattern Recognition, Vol. 3*, 2002.

[9] J. Wei, K. Vongsy, O. Mendoza-Schrock and C. Liu, "Vehicle engine classification using spectdral tone-pitch vibration indexing and neural network," *Int. J. of Surveillance and Monitoring Research,* 2015 (in revision).

[10] X. Huang, A. Acero and H. Hon, Spoken Language Processing: A Guide to Theory, Algorithm, and System Development, Prentice-Hall, 2001.

[11] W. Hartmann, Principles of Music Acoustics, Springer, 2013.

[12] R. Szeliski, Computer Vision: Algorithms and Applications, Springer, 2011.

[13] K. P. Manning, Machine Learning: A Probabilistic Perspective, MIT Press, 2012.