Classification of Uncooperative Vehicles with Sparse Laser Doppler Vibrometry Measurements[1, 2]

J. Wei*[a], Chi-Him Liu[a], Zhigang Zhu[a], Olga Mendoza-Schrock[b], Karmon Vongsy[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

Recently Laser Doppler Vibrometry (LDV) has been widely employed to achieve long-range sensing for the purpose of Intelligence, Surveillance and Reconnaissance (ISR) in military applications, due to its high spatial and spectral resolutions in vibration measurements that facilitates effective analysis using signal processing and machine learning techniques. Based on the collaboration of The City College of New York and the Air Force Research Laboratory in the last several years, we have developed a bank of algorithms to classify different types of vehicles, such as sedans, vans, pickups, motor-cycles and buses, and identify various kinds of engines, such as Inline-4, V6, 1- and 2-axle truck engines. Thanks to the similarities of the LDV signals to acoustic and other time-series signals, a large of body of existing approaches in literature has been employed, such as speech coding, time series representation, Fourier analysis, pyramid analysis, support vector machine, random forest, neural network, and deep learning algorithms. We have found that the classification results based on some of these methods are extremely promising. For instance, our vehicle engine classification algorithm based on the pyramid Fourier analysis of the engine vibration and fundamental frequencies of vehicle surfaces over the data collected by our LDV in the summer of 2014 have consistently attained 96% precision. In laboratory studies or well-controlled environments, a great array of high quality LDV measured points all over the vehicles are permitted by the vehicle owners, therefore extensive classifier training can be conducted to effectively capture the innate properties of surfaces in the space and spectral domains. However, in real contested environments, which are of utmost interest and practical importance to military applications, the uncooperative vehicles are either fast moving or purposively concealed and thus not many high quality LDV measurements can be made. In this work an intensive study is performed to compare the performance in vehicle classifications under the cooperative and uncooperative environments via LDV measurements based on a content-based indexing approach. The method uses an iterative Fourier analysis and an artificial feed-forward neural network. As our empirical studies have suggested, even in uncooperative and contested environments, with adequate training dataset for similar vehicles, our classification approach can still yield promising recognition rates.

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

1. INTRODUCTION

For the purpose of Intelligence, Surveillance and Reconnaissance (ISR) in military applications, Laser Doppler Vibrometry (LDV) sensors, as illustrated in Fig. 1, have been extensively employed to achieve long-range sensing. A LDV works by sending out a laser beam to a reflective surface, which is then reflected back to and received by the LDV, therefore the amplitude and frequency of the surface vibrations are extracted from the Doppler shifts between the outgoing laser beams and the reflected ones. Many advantages are provided by use of LDV sensors. Two most outstanding ones are given below.

1) *Non-contact and non-invasive measurement*: no mass or pressure is ever applied during LDV's measurement process, thus achieving long-range sensing in a non-invasive manner for ISR purposes, and causing no extra

[1] This work is released by Air Force Research Laboratory with PA approval # 88ABW-2014-5269.
[2] This work is supported by NSF I/UCRC Center for Surveillance Research, The Wright State University Site.

damages in non-contact civil engineering applications. The laser beams employed by most LDVs are mostly eye safe (only some high power types may cause damage to human eyes if stared directly for a long duration). By contrast, a small dose of radiation from CT or X-ray in medical applications can have enormous side effects to human cells; while the water penetration and/or corrosion are always the annoying side effects in traditional ultrasonic structural tests.

2) *High spatial and spectral resolution in a long range*: the expansive and wide range of amplitudes and frequencies offered by LDV sensors from a relatively long distance give researchers and developers valuable information in both spatial and frequency domains for intensive analysis, classification and clustering. Based on our experiments, LDV signals with a COTS sensor shown in Fig. 1 from as far as 100 feet are of excellent qualities.

The use of LDV in many research and development subjects has become increasingly more popular due precisely to the foregoing advantages. Recent promising applications of LDV for bomb detection, artwork (fresco) scanning in museums [1], railway inspection and building structure inspection [2] [3] have been reported by engineering researchers across the world.

Many previous efforts have been made to address the issues of classifying vehicles making use of LDV data with varying degrees of success. Wang and collaborators at the City College of New York (CCNY) and the Air Force Research Laboratory (AFRL) [4] developed a method to detect and classify civilian vehicles based on multimodal audio-visual features, including visual tokens using global geometric features and local structure features (Histogram of Oriented Gradients), as well as various LDV features, such as short term energy. Radial-based support vector machine was employed to classify vehicles with promising performances. In [5] Smith and colleagues at AFRL put forward a hierarchical vehicle classification approached using laser-vibrometry data was developed, where a wide array of time and frequency domain features such as spectral flux, Mel-frequency cepstral coefficients (MFCCs), and number of zero-crossings are tested and automatically selected by a parameter selection procedure to generate a tree of different types of vehicles such as vans and sedans.

Figure 1. The LDV sensor used in our tests.

In the following sections, signatures of LDV signals of utmost importance for classifications will be developed.

2. VEHICLE ENGINE INDEXING AND CLASSIFICATION USING LDV MEASUREMENTS

LDV data are 1-dimensional signals similar to other sound or music data. Thanks to the immense success of speech recognition and music encoding [6], techniques of widely use with great success such as MFCC and short-term Fourier transform, were employed to encode and index LDV signals in most state-of-the-art LDV analysis approaches. However, to effectively exploit the LDV signals for ISR purposes, the direct use of speech coding methods are problematic since they have been carefully tailored to take advantage of human auditory systems (HAS). Two most outstanding methodologies behind the success of speech coding are the following two data suppression approaches:

1) *Logarithmic transformation for magnitude suppression*: HAS's sensitivity is not linear to the magnitude I of sound, but in direct proportion to its logarithmic transformation, namely, the essential *loudness* L is determined by log(I):

$$L = 10 * \log\left(\frac{I}{I_0}\right),$$

Where I_0 is the HAS's hearing threshold. Because of this, in MFCC the logarithm transformation of the Fourier magnitudes applies, which attained great success to exploit/cheat HAS.

2) *Octave band collection for signal length suppression*: HAS's sensitivity decreases as the frequency of the sound signal increases. This gives rise to a new quantization methodology referred to as octave band quantization: when formulating the digital representation of sound signals, a binning or histogramming procedure applies, that is, the values within the same bin or bracket are summed as one value; the size of the

bins doubles in the sequence. As an illustration, for a signal of length 15 of constant value 1s: [1 1 1 1 1 1 1 1 1 1 1 1 1], the resultant octave band collection produces a length-4 vector [1 2 4 8].

However, in vehicle classification applications, HAS is no longer the judge for the recognition quality, these great efforts to successfully exploit the HAS for optimal encoding and recognition purposes are inappropriate or even irrelevant any more. 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 surface. 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 of 3-D surfaces, the geometry around the laser measurement points is in essence only of two dimensional. The focal points should thus be the phenomenology of periodic vibrations on rigid surfaces and the corresponding mathematics, e.g., partial different equations (PDE) for waves, and physics, e.g., fundamental frequencies of sheets, modeling and analysis. Most of those concepts and ideas targeted for the HAS should be carefully re-examined in this new direction of classification endeavors.

Based on the preceding arguments, using features directly from prior successful efforts of human speech recognition is problematic. To achieve desirable classification results, it is necessary to develop new features by taking account of the special nature of vehicle engines. After intensive exploratory data analysis of vehicle engine data available to us from the measurement experiments to be detailed in the next section, in order to develop the representations of different vehicle engines for effective classification, the following special signatures are observed:

1. *Time unit*: unlike MFCC where the duration is usually around 40 milliseconds, we will have 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 of about 60 is hard to be picked up by LDV sensors, that is, the largest possible frequency resolution must be at least 1Hz (60 RPM). Hence a duration with at least 1 second must be subtended by the representation to attain adequate resolution of the frequency details of the vehicle engines. By trials and errors we found a duration 1.25 seconds provides the optimal performance.
2. *Analysis domain*: a great array of time-domain features such as zero-crossings, short-term moving averages have been tested, but none of them presented noticeable usefulness. Even as part of the machine learning committee together with other spectral domain (after Fourier transformation) features, no considerable performance increase was evidenced by using time-domain features. Consequently in this work we only focus on analysis in the spectral domain.
3. We don't find the logarithmic transformation suppression to be beneficial. We have also found that still Fourier coefficients in different bands play different roles:
   a) Very low AC bands (<5 Hz) are mostly caused by other factors, e.g., wind and turn of driving wheel, which are irrelevant to vehicle engines and 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 buried by overwhelming noises 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 (not always). Instead of using the original Fourier magnitudes, which will 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 can yield the best result in suppressing noises while keeping useful signals as much as possible.
   d) In Cases b) and c), the higher AC bands are either

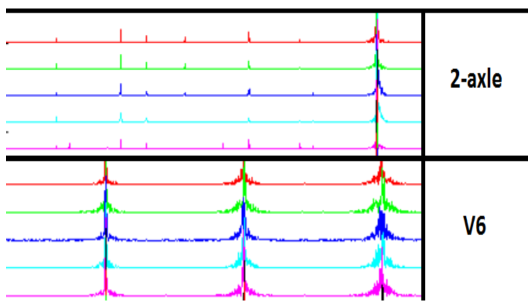

Figure 2 periodic peaks due to fundamental frequency.

suppressed or discarded altogether to reduce the impacts of noise in those bands. However, As depicted in Fig. 2, some strong signals due to the fundamental frequency 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 signal, 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 second application of Fourier transform over the magnitudes of the results of the first Fourier transform is similar to the *pitch* information in music encoding [7], this part of information kept 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. If we view the vehicle as a large music instrument blew by the engine, the signals picked up by the LDV contain the tone (first application of Fourier transform) and pitch information as a complete picture of the signature of vehicle engines.

e) Last but not the least, the default measurements made by LDV are velocity, a first derivative u_t of vibrating distance u. However, the partial differential equation (PDE) for a 2-D wave function on a plane is

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

Where Δ is the Laplacian operator: the sum $u_{xx}+u_{yy}$ of the second spatial derivatives of u. As afore-mentioned, the place for the small laser point is an actual 2-D space, according to the foregoing wave PDE, to measure the innate vibration u_{xx} of the vehicle surface as a proxy of the engine's vibrations, the second time derivation u_{tt} should be employed. Consequently in our work the recorded LDV data are further differentiated to change it to the correct representation of the vehicle surface vibration u_{xx}.

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

1) Basic representative unit is the vibration data d with duration s = 1.25 seconds
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 test, this spectral domain index carries sufficient information to classify different types of vehicle engines for all three sets of experimental data 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 and variant LogitBoost [8], however, none of them can deliver accuracies higher than 60%. Almost all of them can only yield accuracies 30~40%, not much better than the random guess 25%. Therefore in this work we only report classification results based on our tone-pitch index as the representation of engine types.

3. EXPERIMENTS AND CLASSIFICATION REPORTS

3.1. Three sets of experimental data

To test the viability of our new vibration indext, well-controlled LDV measurements of vehicles are needed. As a team sponsored by the Air Force Summer Faculty Fellowship Program (SFFP) in the summer 2014 and NSF I/UCRC Center for Surveillance Research since September 2015 to now, we conducted three sets of experiments with different targets in mind.

Dataset 1—the WSU 2014 dataset (WSU-14)

Figure 3. Two I4 cars with reflective tapes. Left: the passenger side; right: the front side.

During the summer of 2014, from June 30 to August 5, together with several other summer interns and AFRL scientists, the first two authors (Wei and Liu) participated in an AFRL project headed by the last two authors (Mendoza-Schrock and Vongsy). In this project our team collected LDV measurements in the parking lot of the Nutter Center of Wright State University (WSU) for 12 civilian vehicles with four different types of engines, namely, Inline-4 (I4) and V-6 (V6) for 4-door sedans, 1-axle (1A) and 2-axle (2A) diesel engines for semi-trucks. All of them are collected multiple times under different weather conditions. During measurement time for each vehicle, the vehicle was put in parking; the driver was asked to vary the pushing of the gas pedal to let engine run at the condition *Idle* (no pedal pushing), *2000 RPM* (rotations per minute), *sweep* (the RPM indicator moves slowly from idle to 3000 rpm and backwards), and *Idle with fan/radio on.* As shown in Fig. 3[3], 24 small squares of reflective tapes were put around the vehicle surfaces: three on the front and back bumpers each, six on the passenger side, and 12 on the driver side. The LDV sensor was wheeled around the vehicle recording four run conditions for each of the 24 locations for a duration of 30 seconds—during the 30-second recording the LDV is stationary and the driver was asked to perform and literally stick to the 4 designated operating conditions. One complete measurement of a vehicle hence generated around 100 30-second LDV measurements. The same vehicle was measured at least twice at different times to see if the variations in temperatures and other weather conditions will significantly change the performance of classification.

Dataset 2—the CCNY Nov. 2014 dataset (CCNY-14)

In November 2014, three months after the first dataset collection in WSU, sponsored by the NSF CSR grant, the first two authors conducted data collections on two more I4 vehicles in CCNY parking lot, shown in Fig. 2, using exactly the same methodologies used in the WSU dataset: The operating conditions and placements of reflective tapes as just described in dataset 1 are the same. These two cars are not the used in the dataset 1 collection, and the weather condition is even more different from the ones in the summer. This set of data should serve as yet another even more challenging test data to inspect the performance of our vehicle engine classification algorithms.

Dataset 3—the CCNY Jan. 2015 dataset (CCNY-15a and b)

Figure 4 Tests of moving sensors or vehicles. The long reflective tapes placed along the passenger side.

In Jan. 2015, in CCNY parking lot the first two authors performed another data collection: this time instead of always taking LDV measurements over parking vehicles where no movements are involved, which is unrealistic in real world scenarios, two subsets of experimental data were collected on the long reflective tape placed on the passenger side of an I4 car, as demonstrated in Fig. 4.

a) The car was parked and not moving, while the driver was performing the three operating conditions (excluding the 4th condition—radio on, which was found to be entirely undistinguishable from Idle from the two previous datasets) as before, instead of letting the LDV sensor fixating on one small square of reflective tapes for a certain duration a the measurement of the surface vibration, the sensor now panned from the front to the back end or from back-forth while recoding the vibrations from the tape. The purpose of this mobile test is to see when the sensor is moving during data collection, if the signals recorded on the vehicle surface are still of

[3] The two vehicles shown are those taken in the second dataset—the CCNY dataset. The AFRL vehicles' dataset is not shown here for security reasons. The placements of reflective tapes on the vehicle surfaces and LDV measurement procedures are the same: the first two authors participated in both dataset collections using the same LDV sensor shown in Fig. 1.

acceptable qualities. We conducted this test several times with different recording durations to gain deeper insights into the mobile sensor case.

b) This time the LDV sensor is stationary, while the car was moving back and forth when the LDV recorded the vibration, some minor re-adjustments may be applied just to make sure the laser point stayed on the reflective tapes to guarantee the data quality[4]. To allow for at least a data duration of at least several seconds, the car has to move very slowly.

The third dataset is one step closer to the practical use of the LDV sensor in real world ISR applications, but still a small step since the movements involved are admittedly much slower than real urban traffics.

3.2 Procedures of classification performance evaluations

In our performance tests, as usual in machine learning [9], all the datasets available to use are partitioned into three separate sets for different purposes:

1. *Training* and *cross-validation (CV) data*: one complete (4 operating conditions for all 24 measuring point) measurement of one randomly chosen vehicle for each of the four types of engines serves as the training and CV data. For each type, four randomly chosen points, each with all its four operating conditions, are opted as the training data; while the remainder data serve as the CV data points for each engine. The training data were used to train classifiers; whereas the CV data were used to tune the optimal parameters that give rise to the best possible classification accuracy for different classifiers, for instance, the number k in kNN, and number of neurons in the hidden layer of a neural network.
2. *Test data:* All the measurements that were not used to train and validate the classifiers are kept as the objective data to produce the hard evidence of the classifiers' performance.

As described in the preceding section, 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. From the data **d** a large set of overlapping sequences of duration 1.25 seconds are first formulated. For a 30-second 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 more than 100 1.25-second segment, 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., [9].

The CV and test accuracy rates over all four test datasets, i.e., the WSU 2014 summer set, the CCNY 2014 Nov. set, and the CCNY 2015 a) and b) sets, are summarized in Tab. I.

Table 1. Summary accuracy rates using five classifiers for the four datasets

Data	**AdaBoost**	**Random Forest**	**kNN**	**logitBoost**	**Neural network**
CV	96	95	97	96	97
Test1: WSU-14	75	74	68	75	96
Test 2: CCNY-14	37	29	53	34	96

[4] So far without reflective tapes, the quality of the LDV measurements is still unacceptable.

Test 3: CCNY-15a	23	16	60	20	97
Test 4: CCNY-15b	46	32	68	39	99

From Table 1 it can be consistently observed that the neural network, the one chosen by the CV procedure to have the optimal performances has 20 neurons in the hidden layer, attained the best performances for all four test phases. Since the classifiers were trained using a small subset of WSU summer data, while the test data in Tests 2-4 were half a year later in entirely different weather (Fall and winter vs. summer), and different scenarios (moving sensor in test 3 and moving cars in test 4 to simulate the uncooperative scenarios), it is understandable that the four general classifiers failed miserably to as low as 16%, the accuracies of the NN are exceedingly outstanding: the linear combinations and sigmoid thresholding involved in this shallow network must have encoded the RPM and surface fundamental frequencies in the optimized manner.

4. CONCLUDING REMARKS

In this paper, after analyzing the properties of the state-of-the-art speech recognition methodologies, a new indexing scheme to better represent the contents of the LDV measurements for the purpose of engine type classification in the uncooperative environments, the vibration tone-pitch index is put forward as a result. Based on this index we run tests on four datasets we collected with the sponsorship from AF SFFP, AFRL, and NSF CSR from WSU, CCNY for parked or moving vehicles using stable or moving LDV sensors. We discovered that although most classifiers, such as adaboost, random forest, and kNN, cannot achieve desirable classification accuracies, the neural network with 20 neurons in the hidden layer is able to deliver consistently outstanding performance with accuracies consistently better than 96% in all CV and test cases. As reported in another paper by our group [10], even the deep net cannot attain this high level of performance in a consistent manner.

Currently we are actively working on several fronts to make our work on LDV sensor based classification of more utility to ISR efforts. These include:

1) Analyze the inner workings of the neural network with the hope of optimal spectrum feature selection, so that instead of using more than 100 numerical values (due to the tone-pitch index), we can use data of much reduced dimension to achieve similar classification performances [11].

2) Right now the allowable movements of vehicles and sensors are still in very slow motion, which is far from the authentic uncooperative vehicles that will move much faster. Readjustments have to be made to the index to reduce the duration of coding vectors, namely, from 1.25 seconds to the order of 0.1 second, then the practical uncooperative vehicles can be realistically handled.

3) Given the immense success of deep net [12] [13], we need to investigate more about the reason why it cannot achieve performances similar to the shallow neural network yet. One of the reasons could be the small size of the training data and lack of variety of training objects to sufficiently train the deep nets. In the current settings it may not be realistic for us to produce training data with size, say, more than millions of data points, that the deep net can have an edge. However, as more datasets are available, we will see if the deep net can improve its performances as abundantly evidenced in many other disciplines.

REFERENCES

[1] 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.

[2] D. Willemann, P. Castellini, G. Revel and E. Tomasini, "Structural Damage Assessment in Composite Material using Laser Doppler Vibrometry," in *Int. Conf. on Vibration measurements by Laser Techniques*, Ancona, Italy, 2004.

[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] T. Wang, Z. Zhu and C. Taylor, "A multimodal temporal panorama approach for moving vehicle detection, reconstruction and classification," *Computer Vision and Image Understanding,* vol. 117, pp. 1724-1735, 2013.

[5] A. S. A. Smith, S. Kungas, M. Derking and O. Mendoza-Schrock, "Vechicle Classification using Laser-Vibrometry," in *SPIE DSS 111 Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR V*, 2014.

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

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

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

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

[10] 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).

[11] J. Wei, "On Markov Earth Mover's Distance," *Int. J. on Image and Graphics, 14(4),* 2014.

[12] G. Hinton and R. Salakhutdinov, "Reducing the Dimensionality of Data with Neural Networks," *Science,* vol. 313, no. 5786, pp. 504-507, 2006.

[13] J. Wei, "Small moving object detection from video sequences," *Int. J. of Image and Graphics,* vol. 14, no. 3, 2013.