Turk J Elec Eng & Comp Sci
() : –

doi:10.3906/elk-1807-49

A new spectral estimation based feature extraction method for vehicle classification in distributed sensor networks

Erdem KÖSE[1] **, Ali Köksal HOCAOĞLU**[1*]
[1]Department of Electronics Engineering, Faculty of Engineering, Gebze Technical University, Gebze/Kocaeli, Turkey

Received: .201 • **Accepted/Published Online:** .201 • **Final Version:** ..201

Abstract: Ground vehicle detection and classification with distributed sensor networks is of growing interest for border security. Different sensing modalities including electro-optical, seismic and acoustic were evaluated individually and in combination to develop a more efficient system. Despite previous works which mostly studied frequency-domain features and acoustic sensors, in this work we analyzed the classification performance for both frequency and time-domain features and seismic and acoustic modalities. Despite their infrequent use, we show that when fused with frequency domain features, time-domain features improve the classification performance and reduce false positive rate especially for seismic signals. We investigated the performance of seismic sensors and showed that the classification performance varies with the type of road due to distinct spectral characteristics of the medium. Our proposed classifier fuses time and frequency domain features and acoustic and seismic modalities to achieve the highest classification rate of 98.6% using a relatively small number of features.

Key words: Spectral estimation, feature extraction, distributed sensor networks, vehicle classification, border security

1. Introduction

Detecting and classifying ground vehicles in the battlefield is an important task. Acoustic, seismic and magnetic sensors [1] are commonly employed to detect and classify ground vehicles due to their fewer restrictions for scenarios where optical/radar-based sensor systems are inhibitive. Research on the classification of ground vehicles has been recently accelerated by the advances in wireless sensors and sensor networks. Also, the increase in sensitivity and signal-to-noise ratio of these sensors has opened up new opportunities and challenges for battlefield awareness and other surveillance applications along with the advances in wireless sensor networks.

Sensor networks are common to use for human/animal classification [2], human footstep discrimination [3], condition monitoring in railway industry [4], vehicle detection and classification [5, 6], urban traffic management [7], vehicle speed estimation [8], supporting environments to multimedia surveillances [9] or discriminating human, animal and vehicles [10]. Tracked and trackless vehicle detection and classification with distributed sensor networks as a counter camouflage technique is also one of the popular application areas [11]. In this paper, we consider a wireless distributed sensor network that is equipped with a microphone and geophone at each node to discriminate military targets. We mainly focus on developing a methodology to extract features from frequency-domain and time-domain signals yielding a high classification performance to classify military targets.

*Correspondence: khocaoglu@gtu.edu.tr

An increasing interest can be observed in the vehicle classification problem in distributed networks in the last two decades. Many methods have been proposed to improve the classification performance. These studies focus mostly either on extracting new features or adapting the well-known classification algorithms to different situations in pattern classification tasks. For feature extraction, the both frequency domain [11] and time domain [12] methods are proposed. However, frequency domain features are dominant in the literature. Duarte et. al. used the coefficients of Discrete Fourier Transform (DFT) of the acoustic and seismic signals. Kornaropoulos et. al. used one dimensional Discrete Wavelet Transform (DWT) [13]. Wang et. al. proposed the use of Mel-Frequency Cepstral Coefficients (MFCC) to extract multi-dimensional frequency spectrum features of target vehicles from acoustic sensors [14]. Kangyan Wang et al. used sparse representation computed via l1-minimization on MFCC [15] and they demonstrated superiority of this approach. Taheri et al. estimated the acoustic data of the dataset as a time-varying autoregressive stochastic model [16] and gave performance assessments in a qualitative manner.

The vehicle classification studies focus also on the classifiers. In [14], Wang et. al. proposed a discrimination dictionary learning framework called Fisher Discrimination Dictionary Learning (FDDL). This method is based on the Fisher discrimination criterion and it was shown that especially for a small size of training samples this approach outperformed Support Vector Machine (SVM). Guo et. al. proposed a Hybrid Dictionary Learning (HDL) method[17] based on learning a hybrid dictionary which has an analysis dictionary to generate discriminative codes and a synthesis dictionary to achieve class-specific discriminative reconstruction. They showed that their method outperformed FDDL [14] in both time consumption and classification terms. Ntalampiras proposed Echo State Network (ESN) based classifier [18], which is based on echo state property of Reservoir Network (RN), and he showed that his method outperformed Hidden Markov Model (HMM) based classifier.

Wireless sensor networks require optimization of the resources such as the use of battery and the bandwidth. Sending the raw data or the feature vectors increases the bandwidth and is not feasible. A common approach to achieve this is to make a local decision at each sensor node and sending the decisions to a local fusion center to make a final decision. In this work, we also followed this strategy and evaluated the feature vectors locally. However, our final decision is node-based. Fusion of the features extracted from acoustic and seismic sensors were performed at individual nodes.

This work uses the real data collected at the third Sensor Information Technology (SensIT) situational experiment of Defense Advanced Research Projects Agency (DARPA) to verify the effectiveness of the proposed methodology. The dataset is called the third SensIT situational EXperiment (SITEX02) dataset and it has received a lot of attention from researchers [11]. The dataset contains acoustic and seismic signals from two types of armored vehicles, which are Assault Amphibious Vehicle (AAV) and Dragon Wagon (DW). AAV is a fully tracked amphibious landing vehicle and DW is a fully wheeled tank recovery truck-trailer. The DARPA program is based on the concept of detecting and identifying targets at the sensor node level and then combining these findings across the sensor field to support remote situation awareness capabilities. For this level of fusion, the accuracy of the decision extracted from each sensor field is extremely crucial. The main purpose of this work is, therefore, to improve the classification performance at individual nodes.

In this paper, we propose a framework for vehicle classification in a wireless sensor network setting. We consider the use of acoustic and seismic sensors together. We make a local decision by fusing features extracted from these two sensors. The contribution of our work is three-fold. First, we propose an efficient approach to reduce the frequency domain features. Second, we conduct an extensive analysis for the contribution of time-domain and frequency domain features on the performance of the vehicle classification problem. Third, the effect of fusing different sensor modalities on reducing the false alarm rate and increasing the classification rate are thoroughly analyzed.

The rest of the paper is organized as follows: In Section 2, the proposed framework for vehicle classification using acoustic and seismic sensors is described. Section 3 describes the dataset and introduces the performance criteria and test results are given. In Section 4, comparison with previous works is given. Section 5 gives summary and conclusion of the work.

2. Feature extraction for vehicle classification

There is growing interest for detecting and classifying military vehicles to identify friend or foe and counter camouflage techniques. In this section, we present a new approach for feature extraction to classify military vehicles more effectively and efficiently. We use a real dataset to show the effectiveness of these features and compare the proposed approach to the existing methods which use the same dataset. Table 1 summarizes the main features of these methods. All methods except [12] use frequency-domain features to discriminate the vehicle classes. The proposed method in this paper uses the both, i.e. fuses the time-domain and frequency domain features. Also, all methods except [11, 12] use only the acoustic sensor for classification. We use both modalities as in [11, 12] and thoroughly investigate the effect of fusing these two sensors.

Table 1: Features and methods.

Methods	Feature extraction	Feature count	Classification	Domain	Sensors
M. F. Duarte et al.[11]	DFT	50	kNN	Frequency	Both
G. P. Mazarakis et al.[12]	TESPAR	45	ANN	Time	Both
E. M. Kornaropoulos et al.[13]	DWT-SαS	8	kNN	Frequency	Acoustic
K. Wang et al.[15]	MFCC	12	SRC	Frequency	Acoustic
S. Ntalampiras[18]	DFT	50	ESN	Frequency	Acoustic
Proposed	PSD-Time Domain	23	ANN	Both	Both

For feature extraction, Duarte et. al. [11] calculated 512-points DFT of the sensor data and used only the first 50 points for each sensor corresponding to 0-968.75 Hz with step size of 19.375 Hz for acoustic modality and, 0-484.375 Hz with step size of 9.6875 Hz for seismic modality. Then, they evaluated some classifiers such as SVM, Maximum Likelihood (ML) and k-Nearest Neighborhood (kNN) and achived the best performance by the kNN classifier. Mazarakis in [12] used a customized Time Encoded Signal Processing and Recognition (TESPAR) alphabet as feature extractor and got a total of 45 features for acoustic and seismic data. They used a customized Artificial Neural Network (ANN) classifier which was named as Archetype C1. Kornaropoulos et. al. in [13] used one-dimensional DWT to get a 4-level decomposition for feature extraction. This method assumes that decomposition outputs are characteristic functions in frequency domain and estimates parameters for Statistical Alpha-stable Distribution (SαS) model and gets a total of 8 features. A specific distance measure is used for kNN classifier. Wang et. al. in [15] used MFCC to extract features from frequency domain for acoustic data and classified the vehicles by Sparse Representation Classification (SRC) method. Similar to the work in [11], Ntalampiras [18] also used 512-points DFT and used the first 50-points as the features for acoustic data. However, he used an ESN-based classifier.

The proposed method in this paper differs from the above studies mainly at feature extraction level.

Unlike from these works, we used both time-domain and frequency-domain features as shown in Figure 1. As frequency-domain features, we estimated spectral densities. However, instead of using the DFT of the signal as features directly, we analyzed the spectral density of the training data and then determined the discriminative frequency bands. As we show in the following subsections, this approach increases the effectiveness of the frequency-domain features, i.e. causes an increase in the classification performance and a reduction in the false alarm rates. It is also worth to mention that this approach yields a very small number of features.

2.1. The proposed method

Feature extraction is one of the critical steps of classification. In this work we used frequency and time domain features to classify vehicles. In literature, several frequency domain feature extraction methods were proposed for vehicle classification. However, time domain features are not commonly employed. We used five different descriptive features of the time domain signals. We fused them with frequency domain features and analyzed their individual contribution to the overall performance. To extract frequency domain features, we first estimated the spectral component frequencies. We used the MUltiple SIgnal Classification (MUSIC) algorithm to find the roots to locate where the peaks occur in the estimated spectrum. We then used the Welch method to extract the frequency domain features. The whole process is summarized in Figure 1.

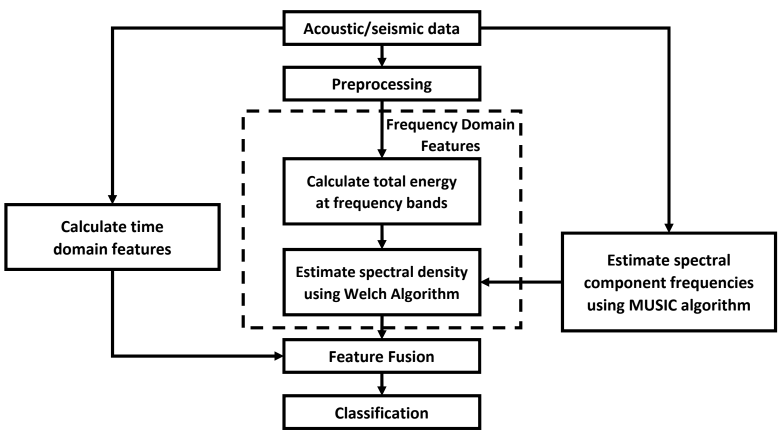


Figure 1: Flow graph for vehicle classification

2.1.1. Frequency domain feature extraction

Frequency-domain features have been used commonly in earlier studies. As mentioned earlier, the coefficients of the discrete Fourier transform of the acoustic and seismic signals have been utilized as frequency-domain features. We use the same approach. However, instead of using the coefficients directly, we estimate the energy of specific frequency bands. In order to determine these bands, we first model the underlying signal as a sum of sinusoids. We then estimate the energy of the bands centered at these frequencies. We apply the MUltiple SIgnal Classification (MUSIC) algorithm to estimate the spectral component frequencies as this algorithm provides us the spectral component frequencies with an accuracy higher than that of autoregressive spectral estimation techniques and other classical spectral estimation techniques such as periodogram and correlogram. We assume a signal model as follows:

$$x[n] = \sum_{k=1}^{K} A_k exp(j2\pi f_k n) + e[n] \quad (1)$$

where $x[n]$ denotes the noise-free complex-valued sinusoidal signal, A_i and f_i are its amplitudes and frequencies, respectively, and $e[n]$ is an additive observation noise. Although the signals acquired from the sensors are real-valued, we use this model due to the convenience from a mathematical standpoint. If we assume as usual that $e[n]$ is white noise with variance σ^2, the covariance of the signal x has the form

$$R_x = \sum_{k=1}^{K} |A_k|^2 s_k s_k^H + \sigma^2 I \tag{2}$$

where

$$s_k = \begin{bmatrix} 1 \\ exp(j2\pi f_k) \\ \vdots \\ exp(j2\pi f_k(M-1)) \end{bmatrix} \tag{3}$$

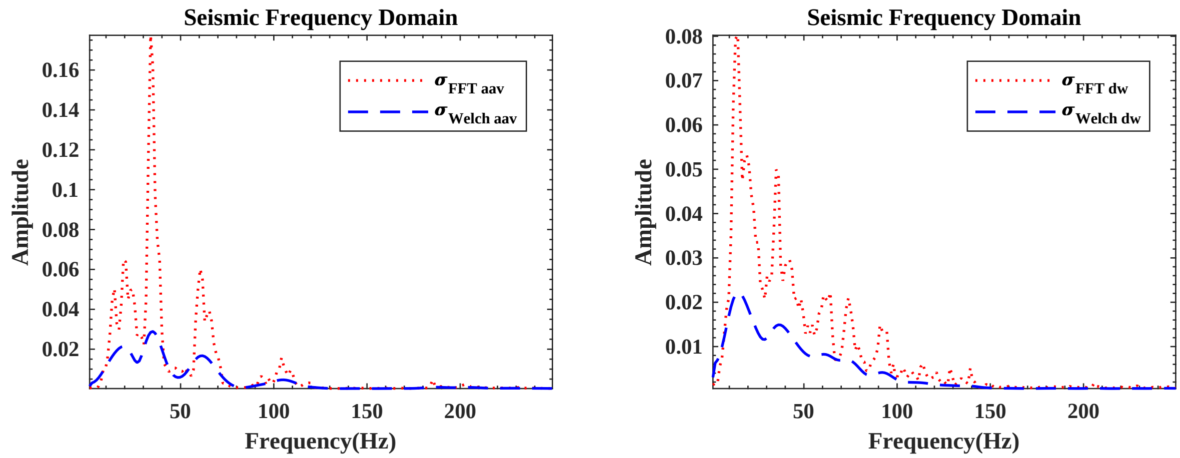
and $M > K$.

The MUSIC algorithm uses the eigen-decomposition of the covariance matrix to determine the frequency estimates as the locations of the K highest peaks of the function

$$\frac{1}{s_k^H (\sum_{k=K+1}^{M} v_k v_k^H) s_k} \tag{4}$$

where v_k, $k = K+1, ..., M$, are signal eigenvectors corresponding to the smallest eigenvalues of the covariance matrix.

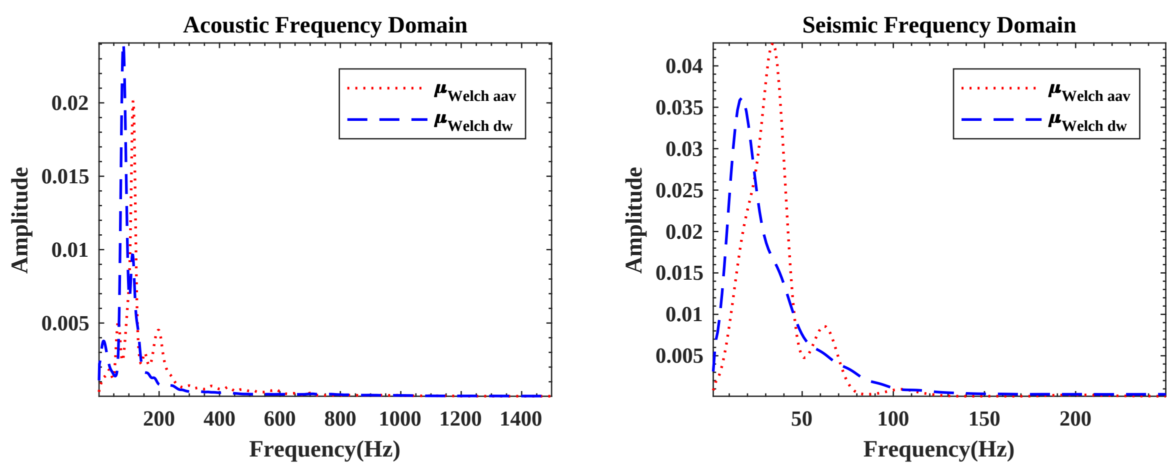
This function is called pseudo-spectrum since it indicates the presence of sinusoidal components in the studied signal; but it fails to provide true Power Spectral Density (PSD). Therefore, we use the MUSIC algorithm only to determine frequencies of the component, and don't use the pseudo-spectrum provided by the MUSIC algorithm as a discriminative feature. Instead, we use the Welch algorithm to estimate the spectrum and then calculate the energy of the frequency bands centered at the frequencies determined by the MUSIC algorithm. Since the Welch algorithm decreases the variance of the estimated PSD by allowing an overlap between data segments, we think that it provides more robust features than the Fourier transform, where the frequency and amplitude of the frequency components vary significantly due to measurement noise. This can be easily seen from Figure 2. The figure on the left shows the magnitude of DFT and the PSD estimated by the Welch algorithm for seismic signal of AAV type vehicle and the one on the right is for DW type vehicle. Although it is presented only for seismic data in this figure, it is also true for acoustic signals that the standard deviation of the Fourier transform is large and varies significantly over the frequency for the training data of the two vehicles being considered in this work. The peaks in the variance occur around the frequencies of the sinusoidal components. Therefore, high variance at these frequencies imply less-dependable features. This analysis suggests that the PSD provided by the Welch algorithm is more dependable than the features obtained by DFT.



(a) AAV Seismic. (b) DW Seismic.

Figure 2: Seismic data's standart deviation of power spectral density spectrums for AAV and DW.

As a preprocessing step, the Direct Current (DC) component was removed from signal. The Hamming window is applied to the blocks of time series data when estimating the spectrum to avoid spectral leakage. Then, the signals are normalized by their energy at each window. For the vehicle classification problem, the acoustic signals have significant energy in the interval 0-250 Hz and the seismic signals have significant energy in the interval 0-100 Hz as shown in Figure 3. The rest of the frequencies can be ignored due to very low energy.



(a) Acoustic. (b) Seismic.

Figure 3: Average Welch graphics for acoustic and seismic data of AAV and DW.

2.1.2. Time domain feature extraction

Time-domain features that we extract in this work are shape based and commonly employed for acoustic and seismic signals. Before extracting time-domain features, we remove the DC component from time series data. We then extract the energy of the signal and the zero crossing density $\frac{\text{\#of Zero Crossing Points}}{\text{\#of Total Sample Points}}$ [19] for acoustic feature extraction. We also extract the skewness and the kurtosis[20] in addition to energy and zero crossing density for seismic feature extraction. In addition to these features, we extract peak-to-peak value as a feature. Thus, we extract 5 features for acoustic signals and 5 features for seismic signals. Thus, we use a total of 10 time-domain features.

2.2. Classification

We designed a two-class classifier to distinguish AAV and DW type vehicles. As shown in Table 1, earlier works used kNN, ANN, SRC and ESN methods to classify the vehicles in the SITEX02 dataset. In this study, we used Multilayer Perceptron (MLP) and Stochastic Gradient Descent (SGD) with cosine annealing learning rate[21].

Our classifier has adaptive learning algorithm, minimum error-restart procedure and holds the best validation performance. It also has adjustable Exponential Linear Unit (ELU) activation function for hidden layers and adjustable hyperbolic tangent activation function for output layers. Choosing an appropriate activation function is an important problem for neural networks due to vanishing gradient problem. We used hyperbolic tangent for output layer. But, it was not used for hidden layers due to vanishing gradient problem. ELU[22] is one of the best solutions for vanishing gradient problem in literature.

Initial weights are another problem because of the speed of convergence. If the initial weights are not chosen wisely, MLP converges in a long period. Instead of assigning random numbers from uniform distribution, we assigned weights from a normal distribution with zero mean and a standard deviation based on number of nodes [23, 24].

Another problem is choosing the right learning rate. We used the SGD with warm restarts[21], which uses cosine annealing with iterations of epochs. This method prevents sticking in local minimum and random initialization point problems.

We used the classification performance on the validation set to determine best result and hold the best until a better result is obtained. We used Boosting and Bagging to verify the result.

3. Results and Dicussion

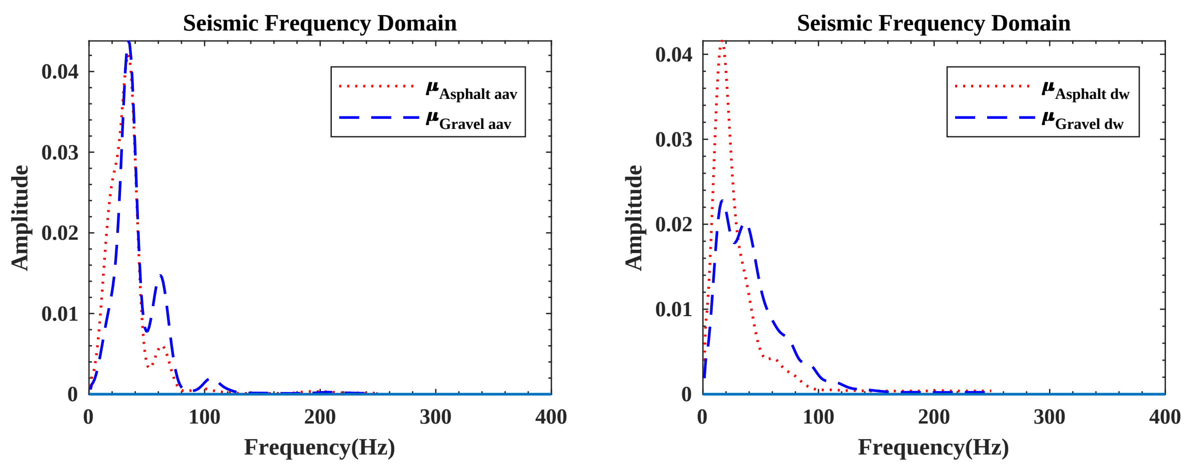
We used the publicly available SITEX02 dataset in our experiments. The dataset contains acoustic and seismic signals from AAV and DW type of armored vehicles. We used the collections labeled as AAV3, AAV6 and AAV9; DW3, DW6, DW9 and DW12 from SITEX02 dataset and excluded *no vehicle* states using Constant False Alarm Rate (CFAR) detection algorithm [11], which extracts the actual event from the raw data. We used 70% of the available data for training, 15% of the data for validation and the rest for the test. We report the performance on the test set and use True Positive Ratio (TPR) and False Positive Ratio (TPR) as performance criteria. True positive ratio is calculated by $\frac{TP}{TP+FN}$, where TP is number of true positives and FN is number of false negatives. False positive ratio is calculated by $\frac{FP}{FP+TN}$, where FP is number of false positives and TN is number of true negatives. We performed 10 trials for each pairs of different sensor modalities and domains, and reported the best achieved performance for each case.

We mainly used ANNs as classifier. However, we also used Bagging and Gentle Boosting methods to verify the results that we obtained using the ANN classifier. Each sensor modality and feature domain combinations are tested to see the contribution of each feature type and sensor modality. Test details are given in this section and the comparison to earlier works are given in the next section.

3.1. Performance of seismic sensors

Most of the earlier works report the classification performance only on acoustic modality. Duarte et.al. [11] and Mazarakis et al.[12] are the only works that report the classification performance on the seismic modality for the SITEX02 dataset. When used individually, seismic sensors provide a poor classification performance of 64% at about 50% false alarm rate. One of our main contributions in this paper is to investigate the reason behind this poor performance and to provide a solution based on this analysis.

Duarte et.al. [11] used the frequency spectrum of the seismic signals of the event. The DFT of these signals was calculated and the first 50 points, containing frequency information of up to 484 Hz, were used to classify the vehicles. Our investigation on different road types shows that seismic signals have distinct characteristics on different medium. Figure 4 depicts the spectral densities of the signals for two types of vehicles on asphalt and gravel roads. As can be seen from the figure clearly, the seismic signal has higher energy at low frequencies for asphalt roads and higher energy at high frequencies for gravel roads. Another important observation is that spectral density of AAV on asphalt road is very similar to that of DW on gravel road in contrast to dissimilarity of them on gravel roads. This indicates that classifying these two type of vehicles on asphalt roads is more difficult.



(a) AAV Seismic. (b) DW Seismic.

Figure 4: Spectral densities for AAV (left) and DW (right) on asphalt and gravel roads.

This analysis suggests that building a classifier for each type of road can yield an improvement in the classification performance. We trained one classifier for asphalt roads and another classifier for gravel roads. The results are presented in Table 2. The classifier trained using for gravel roads provides 88.7% classification performance at 11.4% false alarm rate. The performance on the asphalt roads is slightly lower and drops to 85.5%. The classifier trained for unified dataset which include both asphalt and gravel roads has 85.5%. Since this approach increase the classification performance on certain type of roads, then it may be worth to train a separate classifier for each road type. This is clearly the case for the gravel road for this dataset.

Even a single classifier trained for all road types gives much higher classification performance than the earlier studies. For seismic data, we used the total energy in certain frequency bands as we discussed in Section 2.1. The features obtained in our approach are different than the frequency features in [11]. One difference is the number of features and the other is the variance of the features. We used only 5 features obtained by the spectral energy in certain frequency bands centered at those frequencies provided by the MUSIC algorithm. The spectral energy is calculated using the Welch algorithm. The Welch algorithm provides a smoother spectral information than the Fourier transform. Therefore, the features obtained by the Welch algorithm has lower variance for the same class. This, in combination with the advantage provided by the MUSIC algorithm, which is to determine the major frequency contents precisely, result in higher classification rate.

Table 2: Performance of two classifiers for two different types of road.

	True positive ratio (%)			False positive ratio (%)		
	Classifier for Asphalt	Classifier for Gravel	Classifier for unified dataset	Classifier for Asphalt	Classifier for Gravel	Classifier for unified dataset
AAV	85.5	90.9	82.4	14.7	13.6	11.4
DW	85.4	86.4	88.6	14.5	9.1	17.7
Overall	85.5	88.7	85.5	14.6	11.35	14.5

3.2. Fusion results

In this section, we not only analyze the effect of fusing frequency-domain and time-domain features, we also analyze fusing acoustic and seismic modalities. For each case, the features were extracted as described in Section 2.1. We then report the fusion results to see if there is a positive interaction or synergy between acoustic and seismic modalities and, frequency-domain and time-domain features.

The MUSIC algorithm generates a pseudo-spectrum with seven peaks for the acoustic training data and five peaks for the seismic training data. The rest of the peaks are too weak to recognize. The energies of the bands with a width of 8 Hz. and centered at these frequencies were used as frequency-domain features.

Table 3: Effects of fusion to performance with MLP classifier.

MLP		True positive ratio %			False positive ratio %		
		Frequency	Time	Together	Frequency	Time	Together
AAV	Acoustic	95.9	92.9	96.5	3.7	12.2	4.7
	Seismic	82.4	80.6	89.4	11.4	27.2	10.2
	Together	97.1	94.7	97.7	2.0	9.8	0.4
DW	Acoustic	95.9	87.8	95.9	4.1	6.5	3.5
	Seismic	88.6	72.8	89.8	17.7	19.4	10.6
	Together	97.6	90.2	99.6	2.9	5.3	2.4
Overall	Acoustic	95.9	90.4	96.2	3.9	9.3	3.8
	Seismic	85.5	76.7	89.6	14.5	23.3	10.4
	Together	97.3	92.5	98.6	2.5	7.5	1.4

We used three classifiers to test the quality of the features and we report the results of only one of them in this section. In the next section, we also provide the performances of the other two classifiers. For MLP, we used two hidden layers with 3 and 2 nodes. Number of nodes was kept small in order to avoid over learning and to take advantage of generalization properties of MLPs. As the hidden layer activation function, ELU[22] was used. Activation function of the output layer was chosen to be tangent hyperbolic. Training was performed on the training dataset and ended when the performance started to decrease on the validation set. The classification performance is given for all cases in Table 3.

The effect of sensor fusion and using the time and frequency domain features can be seen in Table 3. The time-domain features are especially useful when only seismic sensor are employed. Frequency domain features are able to classify AAV type vehicles at a rate of 82.4%. Fusing it with the time-domain features causes 7% improvement. The frequency domain features for the acoustic sensor achieve a TPR of 95.9% at 3.9% FPR. An improvement of 1.4% is achieved both on TPR and FPR when they are fused with frequency domain features of seismic data. If only acoustic modality is used, the best performance becomes 96.2% for TPR and 3.8% for FPR when frequency and time domain features are fused. The best overall performance is 98.6% TPR and 1.4% FPR and it is achieved by fusing both acoustic-seismic modalities and frequency and time domain features. This shows that using only acoustic sensors may be a good choice for some cases. However, for cases where a low FPR is important, using the two modalities becomes a necessity.

4. Comparison with previous works

In the last two decades, several methods were proposed to classify vehicles in distributed networks. As mentioned earlier, these methods are based on either time domain or frequency domain features and, most of them use only acoustic modality. In this section, we compare our results with the ones report on the SITEX02 dataset. Table 4 shows the performance for seismic modality and Table 5 shows for fusion of acoustic and seismic modalities for the methods given in Table 1. The proposed method in this work achieves a significant performance improvement by the proposed features. The improvement in TPR is around 10%.

Table 4: Comparison with studies in the literature using only seismic sensor.

	True positive ratio %			False positive ratio %		
	AAV	DW	Together	AAV	DW	Together
Duarte et al.[11]	58.0	56.8	57.4	48.6	47.6	48.1
Mazarakis et al.[12]	87.0	69.0	78.0	-	-	-
Proposed Method w. Gentle Boosting	84.1	87.8	86.0	12.2	15.9	14.0
Proposed Method w. Bagging	84.7	89.0	86.9	11.0	15.3	13.1
Proposed Method w. MLP	89.4	89.8	89.7	10.2	10.6	10.4

Ntalampiras [18] recently achieved 96.3% TPR using acoustic modality for 50 features. We achieve about the same performance using only acoustic data. However, we use only 12 features. On the other hand, by fusing both modalities and fusing time and frequency domain features, the proposed method in this work improves the overall performance by 2% against the best performance in earlier works and reduces FPR to a low figure of 1.4%.

Table 5: Comparison with studies in the literature.

	True positive ratio %			False positive ratio %		
	AAV	DW	Overall	AAV	DW	Overall
Duarte et al.[11]	85.9	81.8	83.6	3.0	10.7	6.9
Mazarakis et al.[12]	100.0	77.0	88.5	-	-	-
Kornaropoulos et al.[13]	94.3	88.9	91.6	11.1	5.7	8.4
Wang et al.[15]	100.0	90.0	95.0	-	-	-
Ntalampiras[18]	95.8	96.7	96.3	-	-	-
Proposed Method w. Gentle Boosting	97.7	98.4	98.0	1.6	2.4	2.0
Proposed Method w. Bagging	99.4	98.4	98.9	1.6	0.6	1.1
Proposed Method w. MLP	97.1	99.6	98.6	0.4	2.4	1.4

5. Conclusion

In this work, we propose a framework for vehicle classification in a wireless sensor network setting. The proposed approach reduces the frequency domain features efficiently while still achieving a high classification rate. Specifically, we used the MUSIC algorithm to determine the major frequency components and then apply the Welch algorithm to estimate the PSD with a low variance. We investigated the performance of seismic sensors and showed that the classification performance varies with the type of roads due to distinct spectral characteristics of the medium. We also conducted an extensive analysis for the contribution of time-domain and frequency domain features on the performance of vehicle classification problem. We showed that when fused with frequency domain features, time-domain features improve the classification performance and reduce false positive rate especially for seismic signals and have insignificant effect on acoustic signals. The effect of fusing different sensor modalities on reducing the false alarm rate and increasing the classification rate are thoroughly analyzed. The proposed approach achieved a performance of 98.6% for TPR and 1.4% for FPR when both acoustic-seismic modalities and frequency and time domain features were fused. This performance was obtained by using only 23 features.

References

[1] Padmavathi G, Shanmugapriya D, and Kalaivani M. A study on vehicle detection and tracking using wireless sensor networks. Wireless Sensor Network 2010; (2):173-185. doi: 10.4236/wsn.2010.22023

[2] Narayanaswami R, Gandhe A, Tyurina A, and Mehra RK. Sensor fusion and feature-based human/animal classification for unattended ground sensors. In: IEEE International Conference on Technologies for Homeland Security (HST); Waltham, MA, USA; 2010. pp. 344–350.

[3] Faghfouri AE and Frish MB. Robust discrimination of human footsteps using seismic signals. In: Proceedings of SPIE 8046, Unattended Ground, Sea, and Air Sensor Technologies and Applications XIII; Orlando, FL, USA; 2011. p. 80460D. doi: 10.1117/12.882726

[4] Hodge VJ, O'Keefe S, Weeks M, and Moulds A. Wireless sensor networks for condition monitoring in the railway industry: A survey. IEEE Transactions on Intelligent Transportation Systems 2015; (16):1088–1106. doi: 10.1109/TITS.2014.2366512

[5] Mayvan AD, Beheshti SA, and Masoom MH. Classification of vehicles based on audio signals using quadratic discriminant analysis and high energy feature vectors. International Journal of Soft Computing 2015; (6):53–64. doi: 10.5121/ijsc.2015.6105

[6] Yang B and Lei Y. Vehicle detection and classification for low-speed congested traffic with anisotropic magnetoresistive sensor. IEEE Sensors Journal 2015; (15):1132–1138. doi: 10.1109/JSEN.2014.2359014

[7] Nellore K and Hancke GP. A survey on urban traffic management system using wireless sensor networks. Sensors 2016; (16):157. doi: 10.3390/s16020157

[8] Odat E, Shamma JS, and Claudel C. Vehicle classification and speed estimation using combined passive infrared/ultrasonic sensors. IEEE Transactions on Intelligent Transportation Systems 2018; (19):1593-1606. doi: 10.1109/TITS.2017.2727224

[9] Atrey PK, Maddage NC, and Kankanhalli MS. Audio based event detection for multimedia surveillance. In: IEEE International Conference on Acoustics, Speech and Signal Processing; Toulouse, France; 2006. pp. 1-5.

[10] Küçükbay SE, Sert M, and Yazici A. Use of acoustic and vibration sensor data to detect objects in surveillance wireless sensor networks. In: International Conference on Control Systems and Computer Science; Bucharest, Romania; 2017. pp. 207–212.

[11] Duarte MF and Hu YH. Vehicle classification in distributed sensor networks. Journal of Parallel and Distributed Computing 2004; (64):826–838. doi: 10.1016/j.jpdc.2004.03.020

[12] Mazarakis GP and Avaritsiotis JN. Vehicle classification in sensor networks using time-domain signal processing and neural networks. Microprocessors and Microsystems 2007; (31):381–392. doi: 10.1016/j.micpro.2007.02.005

[13] Kornaropoulos EM and Tsakalides P. A novel KNN classifer for acoustic vehicle classification based on alpha-stable statistical modeling. In: IEEE/SP Workshop on Statistical Signal Processing; Cardiff, UK; 2009. pp. 1–4.

[14] Wang R, Guo S, Li Y, and Zhang Y. Fisher discriminative dictionary learning for vehicle classification in acoustic sensor networks. Journal of Signal Processing Systems 2017; (86):99–107. doi: 10.1007/s11265-016-1105-x

[15] Wang K, Wang R, Feng Y, Zhang H, Huang Q, Jin Y, and Zhang Y. Vehicle recognition in acoustic sensor networks via sparse representation. In: IEEE International Conference on Multimedia and Expo Workshops; Chengdu, China; 2014. pp. 1–4.

[16] Taheri SM and Nosrati H. Acoustic signature identification using distributed diffusion adaptive networks. In: International Symposium on Communication Systems, Networks & Digital Signal Processing; Manchester, UK; 2014. pp. 943–948.

[17] Guo S, Wang R, Liu B, Wei Q, and Li Y. Vehicle classification in acoustic sensor networks based on hybrid dictionary learning. In: IEEE International Conference on Big Data Intelligence and Computing; Auckland, New Zealand; 2016. pp. 861–865. doi: 10.1109/DASC-PICom-DataCom-CyberSciTec.2016.147

[18] Ntalampiras S. Moving vehicle classification using wireless acoustic sensor networks. IEEE Transactions on Emerging Topics in Computational Intelligence 2018; (2):129–138. doi: 10.1109/TETCI.2017.2783340

[19] Kakar VK and Kandpal M. Techniques of acoustic feature extraction for detection and classification of ground vehicles. International Journal of Emerging Technology and Advanced Engineering 2013; (3):419–426.

[20] Lara-Cueva R, Bernal P, Saltos MG, Benítez DS, and Rojo-Álvarez JL. Time and frequency feature selection for seismic events from Cotopaxi Volcano. In: Asia-Pacifc Conference on Computer Aided System Engineering; Quito, Ecuador; 2015. pp. 129–134.

[21] Loshchilov I and Hutter F. Sgdr: Stochastic gradient descent with warm restarts. In: International Conference on Learning Representations; Toulon, France; 2017. pp. 1-16

[22] Trottier L, Giguere P, and Chaib-draa B. Parametric exponential linear unit for deep convolutional neural networks. In: IEEE International Conference On Machine Learning And Applications; Cancun, Mexico; 2017. pp. 207-214.

[23] Fernández-Redondo M and Hernandez-Espinosa C. Weight initialization methods for multilayer feedforward. In: European Symposium on Artificial Neural Networks; Bruges, Belgium; 2001. pp. 119–124.

[24] Glorot X and Bengio Y. Understanding the difficulty of training deep feedforward neural networks. In: International Conference on Artificial Intelligence and Statistics; Sardinia, Italy; 2010. pp. 249–256.