AUTOMATED ACOUSTIC IDENTIFICATION OF VEHICLES

N Evans Intelligent Systems Research Group, University of York, Heslington, York, UK
D Chesmore Intelligent Systems Research Group, University of York, Heslington, York, UK

1 INTRODUCTION

A number of vehicle recognition researches have already been published. Although monitoring passing vehicles is important for areas including traffic control or road planning[1-3] as well as in recent development of driver assistant systems,[4-6] needless to say monitoring unauthorised intruding vehicles for security has become one of the most critical issues for society, governments and industries.[7-9]

This paper describes a currently ongoing research aiming to develop a system to perform automated vehicle category identification, possibly in real-time, thus will be applicable to security purposes. The organisation of the rest of paper is: Section 2 explains a variety of sensors that have been employed in automated vehicle detection as well as the present approach, Section 3 describes automated identification, of which the current focuses are on time domain signal processing techniques such as Time Domain Signal Coding (TDSC) and Co-Occurrence Matrix, then the primary results are presented and discussed in Section 4 followed by some conclusions and prospect for the near future studies.

2 AUTOMATED VEHICLE DETECTION

Sensors are categorised as active or passive, depending on whether they provide measurement outputs with an external circuit or not.[10] The following are types of sensors that have been utilised in automated vehicle detection.

2.1 Active Sensing and Passive Sensing

In active sensing, radar[11,12] and lasers[13] have been applied mainly because of their capabilities of monitoring relatively wide ranges, as well as their lesser susceptibility to meteorological conditions, allowing detection through day and night, although generally the installation and maintenance costs can be high.[1] On the other hand, inductive loops[14] and magnetic sensors[15,16] may be operated at lower costs, although their complexity and relatively poor accuracy due to noise sensitivity are drawbacks.

The main advantages of using passive sensors in comparison to any active sensing technologies are that data collection activities can be carried out: firstly without interrupting the movements of the target objects hence potentially useful for traffic monitoring applications,[1,5] secondly without betraying the existence of sensors and their implementations therefore would be suitable for military applications for example.[17] Passive sensors in vehicle detection can be divided into two groups: techniques using image processing or acoustic signal processing. The former, such as video camera,[4,5,14] satellite[3] and thermal infrared technology,[18] outperform the latter in stationary target detection[19]. Nevertheless they are more vulnerable to tall vegetation and terrain conditions than the latter. Moreover video cameras and satellite are less effective under dark or foggy situations.[1,11,13]

In addition to the advantages of passive sensors that have been mentioned above, acoustic and seismic sensors can be managed at relatively lower costs[1,17,20-22] and distributed easily according to requirements of practical applications because of the flexibilities supported, for example, due to lower power consumption as well as the physically smaller and lighter devices that are available.[20,22]

2.2 Acoustic Signature of Vehicles

Acoustic signals of vehicles are mostly generated by engine, propulsion, exhaust system, vibration of the body and also friction noise between tyres and the ground surface.[23-30] It was suggested that the dominant components of acoustic signals of a vehicle moving faster than 30 miles per hour (approximately 48 km per hour) are due to tyre friction noise.[31,32] The major elements of moving vehicle signals are observed in the relatively lower frequency part of human perception range, e.g. up to a few kilo hertz.[1,21,26,33] On the contrary, it is also claimed in some papers that often higher frequency components of the sound source, up to several kilo hertz, can be identified in spectrum analyses due to their higher harmonic components.[28] It was suggested that dominant frequency components of seismic signals, generated by a moving vehicle, are below 500 Hz.[33]

2.3 Current Approach

A more accurate observation of unauthorised intruding vehicles can be achieved when utilising multiple methods together rather than relying on one type of sensor. Therefore sensor fusion is a desired near future goal for vehicle detection. Nevertheless development and/or operation of such systems would demand a huge amount of resources. Consequently this research concentrates on acoustic and seismic vehicle detection.

The data presented in this paper were obtained during a recording session at outskirts of York on a cloudy afternoon in January 2008: the temperature and maximum wind speed were measured at 10 centigrade and 1.4 m/h respectively. The apparatus is listed below and Figure 1 shows its set up.

- Two Rode NT5 condenser cardioid microphones (placed vertically to the road)
- Two Sensor SM-4 vertical basic geophone units
- Two AKG CK93 hyper cardioid microphones (placed in parallel to the road)
- Edirol portable 4-channel recorder
- Marantz Portable 2-channel recorder

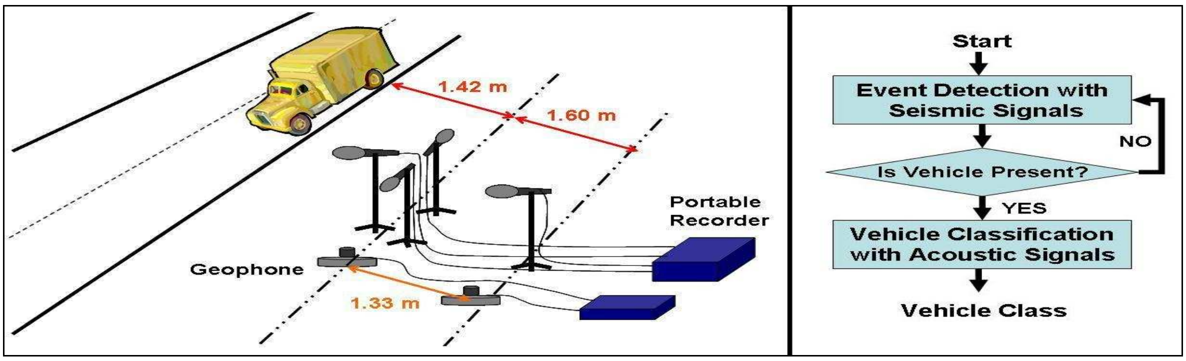

Figure 1: Recording Set Up (left) and System Structure Flow Chart (right)
(Picture of a truck was obtained from Microsoft Office 2003 Clip Art.)

3 AUTOMATED VEHICLE IDENTIFICATION

As presented in Figure 2, an automated identification system consists of input signal pre-processing, feature extraction and decision making stages,[1,34,35] although there are variations in ways of modelling general recognition systems.[36] First of all, data of the target source are collected by suitably placed sensors, and then fed into the system. Secondly unwanted signals in the collected data are removed at the pre-processing stage, while signals of interest may be boosted.

Next, the processed signals are treated further to separate selected features of signals that represent source characteristics well, so that discrimination of signals of different classes can be improved. Finally, a classifier identifies to which category the source of input signal should belong through decision making processes.

Figure 2: General structure of an automated identification system

3.1 Time Domain Feature Extraction

The time domain approach is primarily the examination of signal shape, which can be observed by an oscilloscope for example. Behaviours of input signals are studied in terms of magnitude and the timing of amplitude variation. Now, commonly frequency components of acoustic signals generated by vehicles are dominant in reasonably narrow bands[37] therefore features obtained in the frequency domain may not lead to a clear distinction of vehicle types. Furthermore, time domain signal processing tends to be less computationally expensive and/or capable of achieving faster processing speed than frequency domain or time-frequency domain signal processing. Hence time domain signal analyses may be more appropriate for high speed vehicle recognition.

3.1.1 Time Domain Signal Coding (TDSC)

Time Domain Signal Coding (TDSC) has been developed for diverse applications such as diagnosis of heart condition,[38] machinery maintenance[39] and animal species identifications.[40-45] A standard TDSC is a time domain signal processing technique that focuses on waveform descriptors generated purely in the time domain: such as duration between two consecutive zero crossings and shape information, represented by the number of minima per segment.[40] Since some potential in improving TDSC algorithms further has been recognised, research has continuously been carried out by testing variations of shape descriptors, constructing automated codebook generation algorithms[38,40] and building algorithms that does not require a codebook.[43] More on the current TDSC version is described in 4.1.2.

3.1.2 Co-Occurrence Matrix

In general, it is desirable to achieve efficient descriptions of relatively complex waveforms that maintain the majority of characteristics whilst data are reduced. The Co-Occurrence Matrix, which had been used actively in image processing, specifically in texture analysis,[45] was first introduced to speech signal processing by Terzopoulos suggesting that the Co-Occurrence Matrix could be exploited to advance what had already been known as ‘voiced/unvoiced/silence classification and pitch detection’ in speech signal processing researches.[46,47] Adaptation of Co-Occurrence Matrix to the present research is discussed in 4.2.2.

3.2 Dimension Reduction and Identification

It was realised that increasing the number of features to depict input data attributes beyond a certain point would actually lead poorer classification results: hence executing appropriate dimension reduction algorithms on extracted features before processing the final decision making algorithms could improve overall recognition performance.[36] In general, converging acquired input signals (or features) according to pre-determined rules of assessing similarities and dissimilarities between each other is called clustering: whereas separating input signals one after another into pre-defined finite groups is called classification. Typically clustering algorithms perform unsupervised learning: on the other hand both supervised and unsupervised learning can be seen in classification algorithms. For the actual real-time recognition system, as a finished product of the current research, classification may be more appropriate although clustering techniques can certainly benefit verification of feature extraction algorithms for instance.

4 RESULTS TO DATE

4.1 Event Detection with Seismic Signals

4.1.1 Pre-processing Seismic signals

So far, most data processing has been performed with MATLAB R2007a. Seismic data, collected at a sampling frequency of 44.1k Hz and then normalised with Cool Edit Pro 2.0, were treated by a Blackman window FIR filtering function with the cut-off frequency set to meet the upper band limit of the geophone measurement, i.e. 180Hz. The filtered data were then decimated by a factor of 44 to make the new data sampling rate to be approximately 1k Hz. It is realised that acquiring seismic data at a much lower initial sampling rate would lead more efficient processing.

4.1.2 A New Version of TDSC Feature Extraction

Figure 3 shows features collected by the TDSC algorithm. Firstly input signals are segmented at each zero-crossing point so that each signal frame (named ‘epoch’) can be specified. Secondly the duration (D) and either the number of positive minima or negative maxima (S), depending on the polarity of the signal amplitude, are found about each epoch. Thirdly in the original TDSC, a codebook is generated with training data by manually studying the clustering behaviour of the above D and S combinations about each class, and then finally D and S combinations of the test data are categorised by using the codebook.

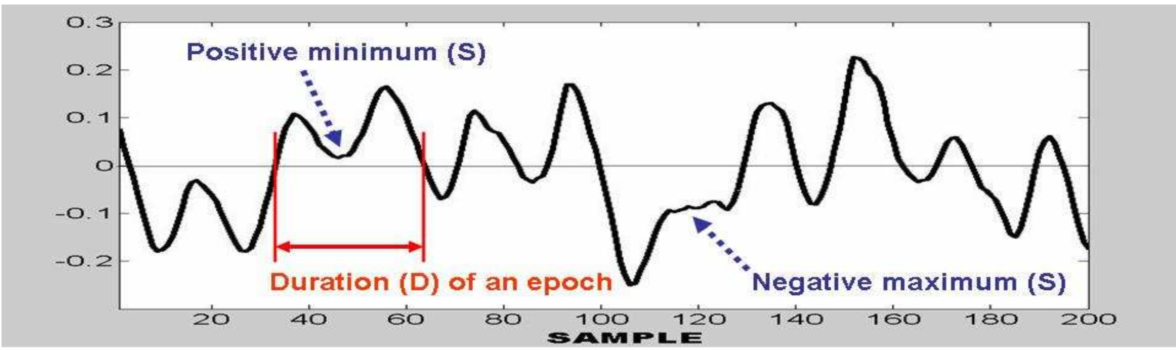

Figure 3: Features Extracted by TDSC from Example Sample Set

Nevertheless, the issues regarding the construction of a suitable codebook are concerning. Thus a proposed new algorithm, which is similar to using D-matrix[43] but involving less computation, has been considered as a novel algorithm that may improve identification. The new procedures are explained below with a set of example samples, of which the waveform is indicated in Figure 3 and the values of collected features are listed in Table 1 and Table 2 .

1. Samples of a certain length (a frame) are collected.
2. Zero crossing points (border of each epoch) within the frame are found.
3. About each epoch, *D* and *S* values are gathered (first and second row of Table 1).
4. For each *D-S* pair per an epoch, $Q=kS+D$ is calculated, where an appropriate k is found according to the maximum value of *D*. For this example, $k=100$ was used (third row of Table 1).
5. The *Q* values are then sorted in ascending order to obtain *Q’* (Table 2).

D	30	30	7	12	6	11	30	11	7	14	7	8	13	5
S	1	2	0	0	0	0	3	0	0	0	0	0	0	0
Q	130	230	7	12	6	11	330	11	7	14	7	8	13	5

Table 1: Sample values for TDSC (D, S, Q)

Q'	5	6	7	7	7	8	1	1	12	13	14	130	230	330

Table 2: Sample values for TDSC (Q')

4.1.3 Vehicle / Non Vehicle Classification with Seismic Signals

By following the above procedures, 9 sets of vehicle samples and 11 sets of background noise samples, all over 7 seconds, were processed. The collected *Q'* values were then plotted on a graph as in Figure 4 so as to examine their characteristics.

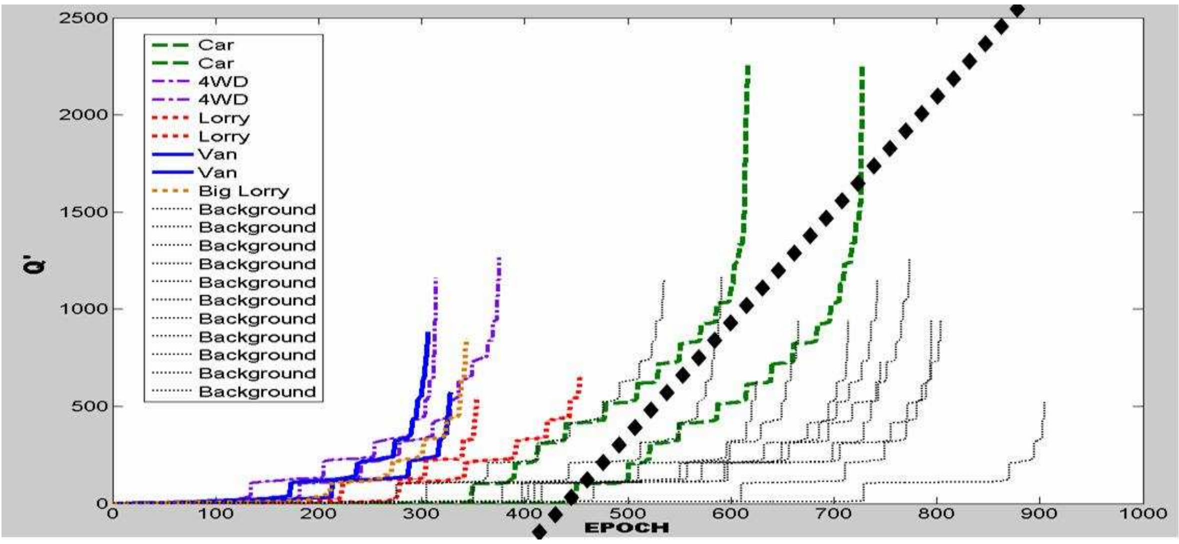

Figure 4: Plotted Graph for Visual Comparison of Q' Values.

By visually inspecting the plotted graph in Figure 4, it can be realised that within a set of samples collected over the same length of time, longer durations (less number of zero crossings i.e. fewer epochs) and greater shape counts (more number of minima) are found in samples of vehicles comparing to that of background noise (i.e. non vehicle). It may be possible to discriminate sample sets between vehicle and non vehicle sounds by the obtained *Q'* values. Then a simple method, a linear line drawn as shown in Figure 4, was adopted to obtain the classification results in Table 3.

	Vehicle Class	Non Vehicle Class
Vehicle Samples	0.89 (8 out of 9)	0.11 (1 out of 9)
Non Vehicle Samples	0.09 (1 out of 11)	0.91 (10 out of 11)

Table 3: Classification Results by Visual Inspection

4.2 Vehicle Category Identification with Acoustic Signals

4.2.1 Pre-Processing of Acoustic Signals

Normalised acoustic signals were firstly processed with a Hamming window FIR band pass filter with cut-off frequencies at 180 Hz and 18k Hz. Nonetheless, other noise cancellation methods should also be studied in order to improve speed and efficiency.

4.2.2 Feature Extraction with Co-Occurrence Matrix Algorithms

Sets of feature vectors were acquired by using the Co-Occurrence Matrix algorithms[47] from acoustic signal samples corresponding to the 8 of the 9 seismic signals processed above. The Co-Occurrence Matrix consists of a two-dimension matrix that embodies frequencies of occurrence of particular amplitude pairs at a '*temporal lag' k* over a period of time, which is framed by the length *M* of window function *W[n]*.

$$w(n)\begin{cases}=1, & 0\le n\le M-1\\ =0, & \text{otherwise}\end{cases} \quad \textbf{(1)}^{47} \qquad S_n(m)=s(m)w(n-m) \quad \textbf{(2)}^{47}$$

$$\Psi ij(S_n,k)=\#\{(m_1,m_2)\mid n-M+1\le m_1,m_2\le n;\ m_2-m_1=k;S_n(m_1)=i,S_n(m_2)=j\} \quad \textbf{(3)}^{47}$$

where # denotes set cardinality.[47]

Some descriptors have revealed distinguishing attributes of acoustic signals generated by various types of vehicles: particularly the lorry was apparent when some parts of the features (0.6 seconds long) were plotted on graphs. Within those, Figure 5 shows 'entropy' and 'variation' descriptors, for which equations to calculate these values are shown in (4) and (5) below.

$$f_{ENTROPY}[\Phi(S_n,k)]=\sum_{i=0}^{Q-1}\sum_{j=0}^{Q-1}-\phi_{ij}(S_n,k)\bullet\log\phi_{ij}(S_n,k)\ ; \quad \textbf{(4)}^{47}$$

$$f_{VARIATION}[\Phi(S_n,k)]=\sum_{i=0}^{Q-1}\sum_{j=0}^{Q-1}(i-j)^2\phi_{ij}(S_n,k)\ ; \quad \textbf{(5)}^{47}$$

where $\phi_{ij}(S_n,k)$ are cells of a co-occurrence matrix.[47]

These findings and effects of variations in signal frame size, window length, lag *k*, and quantisation level will be continuously examined.

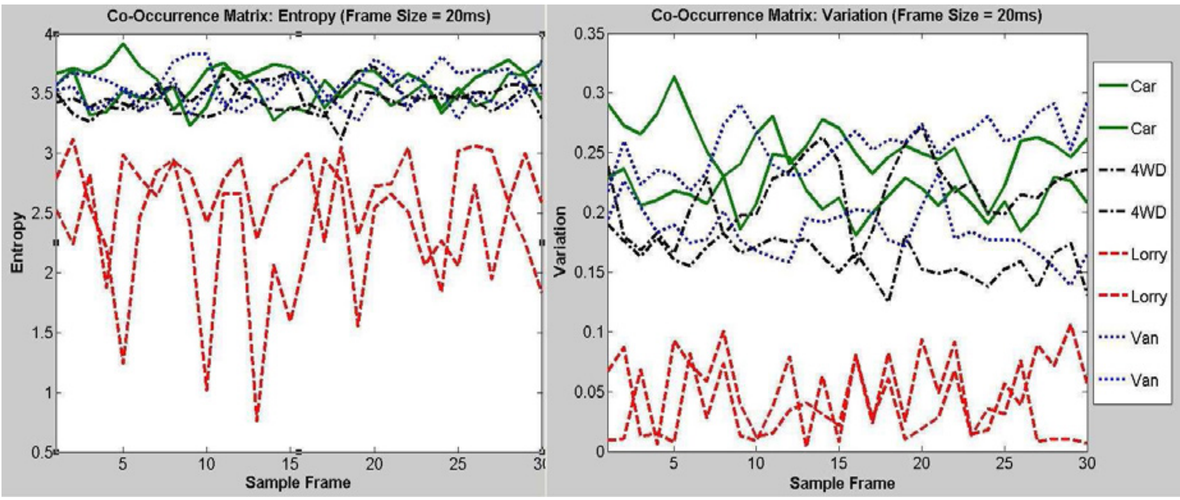

Figure 5: Examples of Collected Descriptor, Entropy (left) and Variation (right)

5 CONCLUSIONS AND FUTURE WORK

Efficient high speed automated vehicle identification is required in a range of applications, for example surveillance and traffic monitoring. Hence development of such a system is the aim of ongoing research at University of York. Real data have been collected and analysed using diverse methods. Because their potential of providing low power, low cost, as well as less intensive computation: acoustic and seismic sensors combined with time domain signal processing techniques have been explored and the primary results have shown some positive outcomes. Nevertheless, the research project is investigating further to improve signal processing efficiency and accuracy. Moreover, carrying out some comparative studies into techniques in other signals processing domains might well be interesting. Additionally, the study of decision making algorithms should be the next milestone of the research.

6 REFERENCES

1. A. Y. Nooralahiyan, M. Dougherty, D. McKeown and H.R. Kirby, A field trial of acoustic signature analysis for vehicle classification, Transportation Research Part C: Emerging Technology, 5(3-4) 165-177. (1997).
2. C. Toth, A. Bars and T. Lovas, Vehicle recognition from LiDAR data, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 23(Part 3/W13), 163–166. (2003).
3. F. Meyer, S. Hinz, A. Laika, D. Weihing and R. Bamler, Performance analysis of the TerraSAR-X Traffic monitoring concept, ISPRS Journal of Photogrammetry and Remote Sensing, 61(3-4), 225-242. (2006).
4. M. Bertozzi and A. Broggi, GOLD: a parallel real-time stereo vision system for generic obstacle and lane detection, IEEE Trans. Image Process., 7(1), 62-81. (1998).
5. Z. Sun, G. Bebis and R. Miller, On-Road Vehicle Detection: A Review, IEEE Trans. Pattern Analysis and Machine Intelligence. 28(5), 694-711. (2006).
6. L. Tsai, J. Hsieh and K. Fan, Vehicle Detection Using Normalized Color and Edge Map, IEEE Trans. Image Process., 16(3), 850-864. (2007).
7. H. C. Choe, R. E. Karlsen, G.R. Gerhart and T.J. Meitzler, Wavelet-based ground vehicle recognition using acoustic signals, Proc. of the SPIE, 2762, 434-445. (1996).
8. J. Altmann, S. Linev and A. WeiB, Acoustic-seismic detection and classification of military vehicles-developing tools for disarmament and peace-keeping, Applied Acoustics 63(10), 1085-1107. (2002).
9. H. G. Xiao, C. Z. Cai and Y. Z. Chen, Military vehicle classification via acoustic and seismic signals using statistical learning methods, International Journal Of Modern Physics. C 17(2), 197-212. (2006).
10. J. S. Wilson, ed., Sensor Technology Handbook, Elsevier Inc. (2005).
11. W. Koch, J. Koller and M. Ulmke, Ground target tracking and road map extraction, ISPRS Journal of Photogrammetry and Remote Sensing .61(3-4), 197-208. (2006).
12. S. Suchandt, M. Eineder, H. Breit and H. Runge, Analysis of ground moving objects using SRTM/X-SAR data', ISPRS Journal of Photogrammetry and Remote Sensing. 61(3-4), 209-224. (2006).
13. H. Cheng, B. Shaw, J. Palen, B. Lin, B. Chen and Z. Wang, Development and field test of a laser-based nonintrusive detection system for identification of vehicles on the highway, IEEE Trans. Intelligent Transportation Systems. 6(2), 147-155. (2005).
14. C. Sun, G. S. Arr, R. P. Ramachandran and S. G. Ritchie, Vehicle Reidentification Using Multidetector Fusion, IEEE Trans. Intelligent Transportation Systems, 5(3), 155-164. (2004).
15. M. J. Caruso and L. S. Withanawasam, Vehicle Detection and Compass Applications using AMR Magnetic Sensors, Technical report, Honeywell International Inc. (1999).
16. J. Ploetner and M. M. Trivedi, A multimodal framework for vehicle and traffic flow analysis, IEEE Intelligent Transportation Systems Conference (2006).
17. G. Becker and A. Gudesen, Passive sensing with acoustics on the battlefield, Applied Acoustics 59(2), 149-178. (2000).
18. B. Nelson, Automatic vehicle detection in infrared imagery using a fuzzy inference-based classification system, IEEE Trans. Fuzzy Systems, 9, 53-61 (1999).
19. S. Hinz and U. Stilla, Car detection in aerial thermal images by local and global evidence accumulation, Pattern Recognition Letters, 27(4), 308-315. (2006).
20. K. B. Eom, Analysis of acoustic signatures from moving vehicles using time-varying autoregressive models, Multidimensional Systems and Signal Processing, 10(4), 357-378. (1999).
21. J. Altmann, Acoustic and seismic signals of heavy military vehicles for co-operative verification, J. Sound Vibration, 273(4-5), 713-740. (2004).
22. J. Lan, S. Nahavandi, T. Lan and Y. Yin, Recognition of moving ground targets by measuring and processing seismic signal, Measurement, 37, 189-199. (2005).
23. D. W. Thomas and B. R. Wilkins, Determination Of Engine Firing Rate From Acoustic Waveform, Electronics Letters. 6(7), 193-194. (1970).

24. D. W. Thomas and B. R. Wilkins, Analysis Of Vehicle Sounds For Recognition, Pattern Recognition. 4(4), 379-389. (1972).
25. R. E. Karlsen, G. R. Gerhart, T. J. Meitzler, R. Goetz and H. C. Choe, Wavelet analysis of ground vehicle acoustic signatures, Proc. SPIE-The International Society for Optical Engineering. (1995).
26. A. Y. Nooralahiyan, H. R. Kirby and D. McKeown, Vehicle classification by acoustic signature, Mathematical and Computer Modelling. 27(9-11), 205-214. (1998).
27. H. Wu, M. Siegel and P. Khosla, Vehicle sound signature recognition by frequency vector principal component analysis, IEEE Trans. Instrumentation and Measurement. 48(5), 1005-1009. (1999).
28. X. Wang and H. Qi, Acoustic target classification using distributed sensor arrays, Proc. IEEE ICASSP, 4, IV-4186. (2002).
29. M. E. Munich, BAYESIAN SUBSPACE METHODS FOR ACOUSTIC SIGNATURE RECOGNITION OF VEHICLES, Proc. 12th EUSIPCO, (2004).
30. B. F. Necioglu, C. T. Christou, E. B. George and G. M. Jacyna, Vehicle Acoustic Classification in Netted Sensor Systems Using Gaussian Mixture Models, Proc. SPIE: Signal Processing, Sensor Fusion, and Target Recognition XIV, 409-419. (2005).
31. C. T. Christou and G. M. Jacyna, Simulation of Vehicle Acoustics In Support of Netted Sensor Research and Development, SPIE Defense & Security Symposium, (2005).
32. G. M. Jacyna, C.T. Christou, B. George and B. F. Necioglu, Netted sensors-based vehicle acoustic classification at Tier 1 nodes, Proc. SPIE: The International Society for Optical Engineering, Vol.5796. (2005).
33. J. Chen, K. Yao and R. Hudson, Source localization and beamforming, IEEE Signal Processing Magazine. 19(2), 30-39. (2002).
34. C. G. Looney, Pattern Recognition Using Neural Networks: Theory and Algorithms for Engineers and Scientists, Oxford University Press, Inc. New York. (1997).
35. R. O Duda, P. E. Hart and D. G. Stork, Pattern Classifiaction, John Wiley & Sons. (2001).
36. C. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, (1995).
37. A. Averbuch, E. Hulata, V. Zheludev and I. Kozlov, A wavelet packet algorithm for classification and detection of moving vehicles, Multidimensional Systems And Signal Processing 12(1), 9-31. (2001).
38. M. D. Swarbrick, Acoustic Diagnosis of Heart Defects using Time Domain Signal Processing and Artificial Neural Networks, PhD thesis, University of Hull. (2001).
39. W. Lucking, The Application of Time Encoded Signals to Automated Machine Condition Classification using Neural Networks, PhD thesis, University of Hull. (1997).
40. E. D. Chesmore, Application of time domain signal coding and artificial neural networks to passive acoustical identification of animals, Applied Acoustics. 62, 1359-1374. (2001).
41. E. D. Chesmore, Automated bioacoustic identification of species, Anais da Academia Brasileira de Ciências, 76 (2), 435-440. (2004).
42. D. Chesmore, Robust In-field Automated Bioacoustic Identification of Species for Rapid Biodiversity Studies, Proc. IOA 4th International Conference on Bio-Acoustic. (2007).
43. I. J. Farr, Automated Bioacoustic Identificaiton of Statutory Quarantined Insect Pests, PhD thesis, Univeristy of York. (2007).
44. I. J. Farr and D. Chesmore, Automated bioacoustic detection and identification of wood-boring insects for quarantine screening and insect ecology, Proc. IOA 4th International Conference on Bio-Acoustics, (2007).
45. K. D. I. Haralick, Textural Features for Image Classification, IEEE Trans. Systems, Man and Cybernetics. 3 (6), 610 - 621. (1973).
46. B. S. Atal and L. R. Rabiner, A pattern recognition approach to voiced-unvoiced-silence classification with applications to speech recognition, IEEE Trans. Speech and Audio Processing 24 (3), 201 - 212. (1976).
47. D. Terzopoulos, Co-occurrence Analysis of Speech Waveforms, IEEE Trans. Acoustics, Speech, and Signal Processing. ASSP-33(1), 5-30. (1985).