Fall 9-14-2011

LOW COMPLEXITY FEATURE EXTRACTION FOR CLASSIFICATION OF HARMONIC SIGNALS

Peter William
University of Nebraska-Lincoln, peter@huskers.unl.edu

LOW COMPLEXITY FEATURE EXTRACTION FOR CLASSIFICATION OF HARMONIC SIGNALS

by

Peter E. William

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Philosophy

Major: Engineering

Under the Supervision of Professor Michael W. Hoffman

Lincoln, Nebraska

September, 2011

LOW COMPLEXITY FEATURE EXTRACTION FOR CLASSIFICATION OF HARMONIC SIGNALS

Peter E. William, Ph.D.

University of Nebraska, 2011

Advisor: Michael W. Hoffman

In this dissertation, feature extraction algorithms have been developed for extraction of characteristic features from harmonic signals. The common theme for all developed algorithms is the simplicity in generating a significant set of features directly from the time domain harmonic signal. The features are a time domain representation of the composite, yet sparse, harmonic signature in the spectral domain. The algorithms are adequate for low-power unattended sensors which perform sensing, feature extraction, and classification in a standalone scenario. The first algorithm generates the characteristic features using only the duration between successive zero-crossing intervals. The second algorithm estimates the harmonics' amplitudes of the harmonic structure employing a simplified least squares method without the need to estimate the true harmonic parameters of the source signal. The third algorithm, resulting from a collaborative effort with Daniel White at the DSP Lab, University of Nebraska-Lincoln, presents an analog front end approach that utilizes a multichannel analog projection and integration to extract the sparse spectral features from the analog time domain signal. Classification is performed using a multilayer feedforward neural network. Evaluation of the proposed feature extraction algorithms for

classification through the processing of several acoustic and vibration data sets (including military vehicles and rotating electric machines) with comparison to spectral features shows that, for harmonic signals, time domain features are simpler to extract and provide equivalent or improved reliability over the spectral features in both the detection probabilities and false alarm rate.

ACKNOWLEDGMENTS

I would like to express my gratitude to those who contributed in bringing today a happy and a very satisfied graduate student. First of all, I would like to thank my advisor Prof. Michael W. Hoffman for his guidance, motivation, patience, valuable advise, and for providing an excellent research environment. I have been awarded a scholarship from the College of Engineering to obtain my PhD and during my first few weeks at UNL, I was hoping to work under the supervision of a good professor, a talented instructor, and a kind person. I remember a graduate student at UNL told me "Dr. Hoffman takes good care of his students", and I believe many graduate students are missing this in their advisors, but I don't.

I have been a teacher assistant with the Electrical Engineering Department since August 2006, and I would like to thank the EE Department for the support and special thanks for Prof. Jerry Hudgins and Prof. Sohrab Asgarpoor for their care, consideration, great recommendations, and valuable feedback. I believe a significant part of the happy package comes from the appreciation I am experiencing as a teacher assistant in my department. I would like to thank all the EE faculty, staff, and colleagues for giving me a remarkable experience at UNL, especially Prof. Sina Balkır and my lab-mate Daniel White. I really enjoyed the technical discussions and the valuable experiences I earned working with Daniel. Thanks in advance for the committee readers Prof. Byrav Ramamurthy and Prof. Mustafa Gursoy for their efforts in reviewing this dissertation.

When I started writing the acknowledgment, I was hoping that I find the words that express my deepest appreciation to all my family. I would like to thank my lovely

wife and best friend, Nivin, for her love, care, support, and for being there whenever I need her. I would like to thank my parents for their support, prayers, and endless encouragement. I know they have been waiting for this day more than me. I am pretty sure this means a lot for them as their happiness and satisfaction means a lot to me.

Contents

List of Figures

List of Tables

List of Symbols

μ	Mean
Σ	Covariance
ϵ	Noise vector
θ	Ball contact angle (with the races)
φ_k	Phase of the k^{th} harmonic
ΔF	Spectral resolution
$\mathcal{I}$	Identity matrix
κ	Normalized energy
ψ_I	In-phase basis function
ψ_Q	Quadrature basis function
ξ	Assumed fundamental frequency $= \frac{1}{\tau}$
$alpha_k$	Amplitude of the k^{th} harmonic
C	Number of classes
d_B	Ball diameter
d_i	Desired output of the i^{th} neuron

d_p	Pitch diameter
$\widetilde{FF}$	Assumed fundamental frequency
F_s	Sampling frequency in Hz
f'	Derivation of the function f
f_B	Ball defect frequency
f_C	Cage defect frequency
f_I	Inner race defect frequency
f_O	Outer race defect frequency
f_R	Motor's rotational frequency
h_i^g	Net input to the i^{th} neuron at the g^{th} layer
L	Feature vector length
M	Number of Harmonics
n	number of observations
N	Number of samples
N_B	Number of rolling elements
$P(.)$	Probability density
P_D	Probability of correct detection
P_{FA}	Probability of false alarm
P_M	Probability of target miss
T_L	Length of observation window
T_Q	Longest expected time duration between successive ZC

W_{ji}^{g}	Weight from the i^{th} neuron to the j^{th} neuron at the g^{th} layer
y_{i}^{g}	Net output from the i^{th} neuron at the g^{th} layer

List of Abbreviations

A/D	Analog-to-Digital
AAV	Amphibious Assault Vehicle
ANN	Artificial Neural Network
ARL	Army Research Laboratory
ASIC	Application Specific Integrated Circuit
ASIP	Application Specific Integrated Processor
ATI	Analog-to-Information
BPSK	Binary Phase Shift Keying
BVP	Bochum Verification project
BW	Bandwidth
CFAR	Constant False Alarm Rate
CCA	Canonical Correlation Analysis
CISC	Complex Instruction Set Computer
CNAPS	Connected Network of Adaptive Processors
CPA	Closest Point of Approach

CS	Compressed Sensing
DAT	Digital Audio Tape
DARPA	Defense Advanced Research Projects Agency
DFT	Discrete Fourier Transform
DPKO	Department of Peace-keeping Operations
DSP	Digital Signal Processing
DTFT	Discrete Time Fourier Transform
DW	Dragon Wagon
DWT	Discrete Wavelet Transform
ECG	Electro-Cardio Graph
EDM	Electro-Discharge Machining
ESHC	Estimated Strongest Harmonic Component
FF	Fundamental Frequency
FFT	Fast Fourier Transform
FOV	Field Of View
GAO	Government Accountability Office
GPP	General Purpose processors
HLA	Harmonic Line Association
HMM	Hidden Markov Model
IR	Infrared
kNN	k-Nearest Neighbor

LBG	Linde, Buzo, and Gray
LOS	Line Of Sight
LS	Lease Square
LVQ	Linear Vector Quantization
MAP	Maximum a Posterior
MGC	Multivariate Gaussian Classifier
MEMS	Micro Electro Mechanical Systems
MIPS	Million Instructions Per Second
ML	Maximum Likelihood
MLE	Maximum Likelihood Estimator
MSE	Mean Square Error
NLOS	Non-Line Of Sight
NN	neural Network
PCA	Principal Component Analysis
PARC	Palo Alto Research Center
PSE	Power Spectrum Estimates
PTSVQ	Parallel Tree Structured Vector Quantization
RBF	Radial Basis Functions
RISC	Reduced Instruction Set Computer
Rprop	Resilient back-propagation
SensIT	SENSor Information Technology

SNR	Signal-to-Noise Ratio
SUV	Sport Utility Vehicle
SVM	Support Vector Machines
TDHA	Time Domain Harmonics Amplitude
TESPAR	Time Encoding Signal Processing And Recognition
TI	Texas Instruments
UGS	Unattended Ground Sensors
VLSI	Very Large Scale Integrated Circuit
VQ	Vector Quantization
WAM	Wide Area Munitions
WDSN	Wireless Distributed Sensor Network
WINS	Wireless Integrated Network Sensors
ZC	Zero-Crossings

Chapter 1

Introduction

1.1 Wireless Sensor Modules

The development of small and reliable wireless sensor modules and the need to utilize these modules in performing complex tasks have driven the focus of much recent research to develop communication and signal processing algorithms for sensor modules with limited resources. In spite of the applications' complexity, two factors should always be fulfilled, reliability and low power consumption. Reliability is achieved by having individual sensor modules capable of performing the assigned tasks with high quality of service and with minimum failure. The low power consumption will allow those deployed battery operated sensor modules to have a long lifetime relying only on their internal batteries or on energy harvesting from their environments.

Wireless sensor modules' key role is to enable the access of information located at a remote location or distributed over a large geographical area. We can denote that a

large number of the applications that utilize wireless sensor modules are monitoring applications. Some of these applications are monitoring the environment to record and identify the occurrence of a certain phenomena or incidents, while others obtain periodic measurements to identify the status of a defined object. Although in monitoring applications a large amount of data may be acquired, only a certain aggregate function of the acquired data is required.

The main constraint on wireless sensor modules is their limited resources. Sensor modules' requisite low cost, small size and weight, limited power and low data transmission rate may limit their ability to realize a highly efficient monitoring system. Recent advances in wireless communications, signal processing techniques, and Micro-Electro Mechanical Systems (MEMS), however, have enabled the development of low cost sensor modules that are capable of performing low power sensing, detection and classification of targets [1].

Although the problem is highly dependable on the monitoring application, almost all monitoring techniques share the common interest in reducing the acquired data into a smaller set of representative features to proceed further in the monitoring task. This step represents a bottleneck for low power monitoring approaches. The reason for extracting the features on the sensor module, is to minimize the energy intensive use of communication resources. The effectiveness of the system depends on its ability to maintain a low power processing profile and to generate a robust set of features.

1.2 Harmonic Signals

Harmonic signals are very prevalent and are found in many applications. Beside their simple structure, they represent a good example of sparse signals that can be totally defined by a small number of parameters. Although these harmonic parameters highly characterize the source [2–5], the harmonic parameters are not constant over time and the harmonic signal is either embedded in noise or other non-harmonic signals requiring care when developing a feature extraction algorithm.

The developed algorithms are evaluated on two different applications. The first, detection and classification of ground vehicles for peacekeeping operations, is an event driven application that requires high detection and classification rates. In addition, the generated features must be robust to many challenges (i.e., the background noise, different environments, and the wide selection of ground vehicles). The second application is the early detection of bearing faults in rotating machinery. Although detection of bearing faults might seem to be a simpler problem, the major challenge lies in the difficulty in defining a set of features that can precisely identify a defect type that might occur anywhere in the bearings and the signature varies with increasing defect size until a complete failure is reached.

When developing a feature extraction algorithm, attention needs to placed on the classifier as well, since both the set of generated features and the classifier share the responsibility for the resulting decision. The classifier should be selected such that it can effectively separate between candidate classes based on the information embedded in the generated features.

1.3 Contributions of this Dissertation

The primary goal of this dissertation is developing simple signal processing algorithms for extracting characteristic features from signals acquired by passive sensors. The extracted features will be used in signal/source identification. These algorithms are proposed for low power wireless sensor modules.

Although our focus is on the feature extraction stage, we investigated the detection and classification techniques that should be combined with the proposed feature extraction algorithms to maintain a low power profile and high reliability. Detection is an essential step for event driven applications and its outcome is very important for determining the system performance. The performance of each of the event detector, the feature extraction algorithm, and the classifier significantly affects the overall system performance.

A summary of specific contributions of this dissertation are listed below:

- Development of two feature extraction algorithms and the contribution in developing a third algorithm that targets an analog hardware platform.
- The common theme regarding all three algorithms, is extracting the features from the time domain signal without the need to transfer the acquired signal into the spectral domain.
- The evaluation process includes comparing the detection and false alarm rate of the proposed algorithms with other published algorithms on the same real world data sets. Our goal includes achieving at least the same classification rate with substantially fewer computations.

- The first developed algorithm "Zero-Crossings (ZC)":
  - Represents the simplest algorithm and it generates the features from the binary acquired signal.
  - The dimensionality of the generated feature vector is small compared to other time domain and spectral domain techniques.
  - The algorithm is more appropriate for applications with a small number of candidate classes.
- The second developed algorithm "Time Domain Harmonics' Amplitudes (TDHA)":
  - Employing simple techniques to estimate harmonic parameters that are crucial for the feature extraction algorithm.
  - The selection of a simplified harmonic model to approximate the acoustic and seismic harmonic structure.
  - Precisely defining the harmonic model parameters and employing a simplified Least Square (LS) method on a template of the acquired data to achieve a representative set of features with fewer computations compared to state of the art spectral features.
- The third developed algorithm "Analog-to-Information (ATI)":
  - The ATI approach replaces a high resolution Analog to Digital (A/D) converter followed by a Fast Fourier Transform (FFT) with a multichannel analog projection and integration to extract the sparse spectral features for signal classification.

  - The proposed approach does not require knowing the Fundamental Frequency (FF) of the harmonic series or the number of harmonics.

- The generated set of features can be applied to a variety of harmonic signal classifiers.

- The implementation of a multi-modal fusion algorithm that is suitable for multi-modal low power signal processing schemes.

- The generation of simulation results for single event decisions, multi-modal fusion and decision fusion between modalities, and the application of the proposed algorithms on three different real world data sets representing acoustic and seismic signals of military vehicles acquired at two different sites and vibration signals acquired using an accelerometer mounted over the housing of an induction motor. The results establish the validity of the selections made with respect to the proposed applications/acquired signals.

We presented our work in a number of conferences/symposiums and peer refereed journals:

- Peter E. William and Michael W. Hoffman, "Identification of Bearing faults using time domain zero-crossings", *Journal of Mechanical Systems and Signal Processing*, Vol. 25(8), pp. 3078-3088, November 2011.

- Peter E. William and Michael W. Hoffman, "Classification of military ground vehicles using time domain harmonics' amplitudes", *IEEE Transaction on Instrumentation and Measurement*, in press (DOI:10.1109/TIM.2011.2135110).

- Peter E. William and Michael W. Hoffman, "Acoustic classification of battlefield vehicles based on their seismic detection", 156^{th} *Meeting of the Acoustical Society of America (ASA)*, Vol. 124(4), pp. 2508-2508, November 2008.

- Peter E. William and Michael W. Hoffman, "Identification of Battlefield Vehicles using the zero-crossings and the strongest harmonic component of the time domain acoustic signal", *Military Sensing Symposia - Battlespace Acoustic and Magnetic Sensors (BAMS)*, August 2008.

- Peter E. William and Michael W. Hoffman, "Efficient sensor network vehicle classification using peak harmonics of acoustic emissions", *Proceedings of SPIE in Unattended Ground, Sea, and Air Sensor Technologies and Applications X (Defense and Security Symposium)*, Vol. 6963, pp. 1-12, March 2008.

- Daniel White, Peter E. William, Michael W. Hoffman, and Sina Balkır, "Analog sensing front end for harmonic signal classification applications", *to be submitted to the IEEE Sensors Journal*.

1.4 Outlines of this Dissertation

Chapter 2 reviews the existing open work for feature extraction algorithms proposed in the literature. The emphasis will be on the signal processing and classification schemes designed for low power sensor modules to enable signal detection, feature extraction and classification on the sensor node. In Chapter 3, we describe the first developed feature extraction algorithm and the experimental results for applying the proposed

algorithm for identification of bearing faults in induction motors. In Chapter 4, the second developed algorithm is described. The second algorithm represents the major contribution of the Dissertation and the experimental results for detection and classification of military vehicles over two different data sets are included. In Chapter 5, we describe the third developed algorithm and the experimental results obtained for applying the feature extraction algorithm on two different applications (classification of military vehicles and identification of bearing faults). The summary of the dissertation and the potential future research topics can be found in Chapter 6.

Chapter 2

Literature Review

This chapter reviews the algorithms proposed in the literature for extraction of characteristic features from harmonic signals for signal/source identification. The focus will be on features extracted for early identification of bearing faults and those extracted for detection and classification of moving military vehicles. Section 2.1 reviews an important operational factor, which involves the power profile of wireless modules and the selection of the hardware technology that appears to be viable for our proposed monitoring system based on the consumed power of individual sensors. Section 2.2 gives a brief overview on the utilization of low power sensor modules in unmanned monitoring. Section 2.3 summarizes the fundamental feature extraction algorithms proposed in the literature that utilize vibration/acoustic signals for identification of bearing faults and classification of military ground vehicles. Section 2.4 describes the classifiers that were proposed in the literature for integration over low power sensor modules and their robustness with the generated features. Section 2.5 reviews

data/decision fusion algorithms as a possible approach for increasing the network reliability.

2.1 Low Power Wireless Sensor Modules

The majority of existing hardware platforms relies on batteries, which dominates their size and weight. Non re-chargeable batteries are often chosen, such as AA, AAA, and coin-type batteries [6]. Chargeable batteries are not preferable, since in many applications charging is not practically available. Using renewable energy and scavenging/harvesting techniques is an interesting alternative and has been actively researched as a possible solution to this problem [7]. There are several system functions that comprise the energy profile for individual sensor modules. These functions, shown in Figure 2.1, include data acquisition, processing, power mode change, sleep/idle mode, and communication. Modern hardware implementations allow for several models that can be discussed for each function. For example, the Mica2 [8] mote has multiple sleeping modes that depend on the duration which the module will spend in sleep mode. Although our primary concern is the development of feature extraction algorithms for classification of harmonic signals, recognizing the limitations of available hardware platforms for sensor modules increases the credibility of the proposed algorithms.

Microcontrollers are now the primary choice for processing in wireless sensor networks [6]. The key metric in the selection of a microcontroller is power consumption. Four different types of microcontrollers, with different characteristics, might

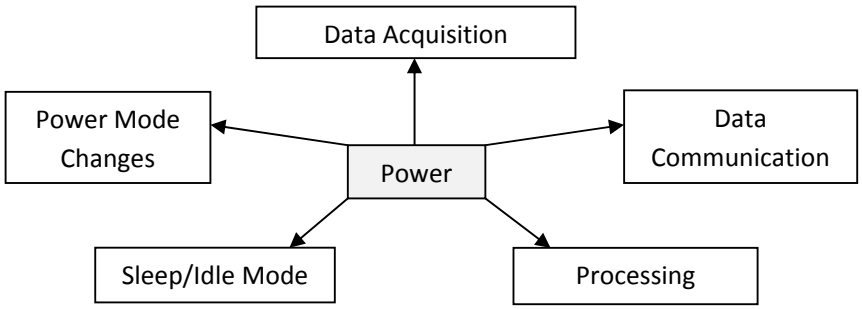

Figure 2.1: Sensor module Power Profile

be considered for these sensor modules: Application Specific Integrated Processor (ASIP), Reduced Instruction Set Computer (RISC), Complex Instruction Set Computer (CISC), and Application Specific Integrated Circuit (ASIC). Among these microcontrollers, the ASIC might seem to be the most convenient microcontroller, since it will be designed specifically to handle functions required by the application. On the other hand, there are high costs and limited flexibility in the ASIC approach compared to the other available general purpose microcontrollers. The MSP430 [9] is an example of a RISC controller, the 8051 [10] is an example of a CISC controller, and the CoolFlux [11] is an example of the ASIP. The CoolFlux is an ASIP specialized for Digital Signal Processing (DSP). The following sections address the different components of the sensor node power profile depicted in Figure 2.1.

Communication

The communication power depends mainly on the rate, separation, protocol, channel, and SNR. Most sensor modules use commercial-off-the-shelf radio modules with transmission power between -25 dBm and 10 dBm, while the receiving sensitivity can

be as good as -110 dBm [6]. Radios with low bit rates ($\leq$ 100Kbps) are very advantageous in terms of power consumption, which motivates performing the processing on the sensor module instead of forwarding the acquired data to a central location, which minimizes the transmission data rate. It must be noted that although radio modules with low bit rates have a low average power consumption, the effective energy/bit consumption is lower than radio modules with higher bit rates. Some monitoring applications require sampling and transmission at a high rate, which will exhaust the modules' power source. In [12], a comparison between the communication power versus the processing power for low power wireless modules show that instead of transmitting the acquired data to a central module, a low complexity processing algorithm implemented on the sensor module might result in a longer lifetime of the sensor module. In [12], the energy cost of transmitting 1 Kb a distance of 100 meters was found to be approximately 3 joules (assuming a 1 GHz carrier frequency, Binary Phase Shift Keying (BPSK) modulation, 10^{-6} error probability and Rayleigh fading), by contrast, a typical general purpose processor (e.g., Intel Atom [13]) can execute more than 3 million instruction for the same amount of energy. Given modern advances in low power microcontrollers and focusing on energy efficient signal processing schemes, performing the majority of the processing on the sensor module is highly recommended. This results in transmitting only the important findings from the module so that the radio is turned off most of the time.

Low power sensor modules commonly employ a RISC controller as the MSP430 [9] or the ATmega 128L [14] as their central processing unit. Table 2.1 lists the most popular battery operated low power sensor modules designed using a RISC controller

and a low power radio chip. These controllers are characterized by: high speed operation (8 MHz); high precision (16 bit-MSP430 and 8 bit-ATmega 128L); fast sleep and fast wakeup ($\leq 6\ \mu sec$); multiple sleep modes with ultra low power in deep sleep mode; and large memory with an optional external memory. Most of the wireless modules employ a Texas Instruments (TI) Chipcon RF transceiver designed specifically for low power and low voltage wireless applications. The current consumption of the Chipcon CC2420 is < 20 mA at a maximum data rate of 250 Kbps, the operating voltage is between 2.1 V and 3.6 V, with programmable transmission power between -25 dBm and 0 dBm.

Table 2.1: Wireless Sensor Modules

Module	Controller Unit	Radio	Year
BTnode [15]	ATmega 128L	Chipcon CC1000	2006
Epic mote [16]	MSP430	Chipcon CC1000	2008
FireFly [17]	ATmega 128L	Chipcon CC2420	2007
NeoMote [18]	ATmega 128L	Chipcon CC2420	2007
Mica2 [8]	ATmega 128L	Chipcon CC1000	2002
MicaZ [8]	ATmega 128L	Chipcon CC2420	2004
T-Mote sky [19]	MSP430	Chipcon CC2420	2006
Tinynode [20]	MSP430	XE1205	2005
TelosB [8]	MSP430	Chipcon CC2420	2004
Sensenode [21]	MSP430	Chipcon CC2420	2006

Data Acquisition

The data acquisition function represents the most frequent operation performed by the sensor module. The power consumed during this phase can be divided into sensing power and signal conditioning power. The sensing power depends on the transducer

and for most passive transducers (acoustic, seismic, infrared, etc.), the amount of consumed power is very reasonable. The signal conditioning power includes the power consumed during amplification, filtering and sampling (A/D) of the acquired signal. The resolution of the A/D has a direct impact on energy consumption [6]. General Purpose Processors (GPPs) are very power efficient since GPPs have integrated peripheral interfaces which reduces the time and power required to communicate with the different transducers on the module [22]. RISC and CISC processors are both GPPs. The DSP ASIP is very inefficient during data acquisition and requires a large amount of power compared to GPPs [22].

Processing

The processing power is the overall power consumed during performing all the computations that follow data acquisition and prior to communication or switching to a sleep mode. Million Instructions Per Second (MIPS) is one common approach to determine the amount of computations required to perform specific tasks. Some controllers such as DSPs can perform multiple instructions in one clock cycle, while other controllers perform a single instruction per cycle and others require hundred of cycles to perform a single instruction. The microcontroller precision represents another important factor in determining its performance, since performing 16 bit operations using an 8 bit microcontroller requires additional cycles and more time for execution. In general, DSPs are more efficient than GPPs with respect to both processing power consumed and run time. At higher sampling rates RISC controllers outperform CISC ones with respect to consumed power during processing and run time [22].

Power Mode Change and Sleep mode

Microcontrollers switch to lower power modes (sleep/idle) in order to reduce the overall power consumption. The overall power consumed, during a specific period, in low power modes in addition to the power consumed during switching must be smaller than keeping the microcontroller at the higher mode for the same period. DSPs (in general ASIPs) are not generally designed to stay for a long time at low power modes. ASIPs switch between modes very quickly compared to GPPs but on the other hand the power consumed in their low power modes is orders of magnitude higher than the equivalent power consumed by GPPs. Some GPPs like the MSP430 supports five low power modes with the current drawn at the lowest mode ≤ 100 nA, where the CPU and all clocks are disabled [9]. The ASIP consumes about 25% of its active power in sleep mode [22]. By spending a long time in sleep mode, the overall power spent in sleep mode exceeds the power spent in all the other functions in the given application.

Even with the assumption that the application will require minimum data transmission and all the processing will be performed on the sensor module, the overall power consumption will depend on the duty cycle, operating frequency, complexity of computations, how many transducers will be considered over the sensor module, etc. A comparison, inspired by the experimental study in [22], between the three different microcontroller types (DSP, RISC, and CISC) using commercial off the shelf microcontrollers with regard to the average consumed power with respect to each function is shown in Table 2.2.

We make the following assumptions about our monitoring system:

Table 2.2: Average power comparison

Function	ASIP	RISC	CISC
Data acquisition	high	low	moderate
Processing	very low	low	high
Power mode changes	very low	low	moderate
Sleep/Idle mode	high	moderate	low

- Spends more than 90% of its time in sleep mode;
- Acquires a significant amount of data using multiple transducers on the sensor module;
- Performs signal detection, feature extraction and classification on the sensor module; and
- Requires sampling of the data at high rates ($\geq$ 500 Hz).

Using the simple comparison between ASIP and GPPs, we conclude that a GPP with RISC technology, such as the MSP430, is the preferred microcontroller for a sensor module that uses off the shelf components for deployment of an automated monitoring system. The comparison in table 2.2 is roughly addressing off the shelf hardware microcontrollers. The indicated average power consumption during different functions are not intrinsic and one could make an ASIP with good peripheral and sensor interfaces, or low sleep power consumption, but this will result in more time to switch to low power mode and will reduce the processing performance.

When the application requires a complicated set of computations in addition to the data acquisition phase, a multi-processor approach might be a useful solution. A combined GPP (efficient for data acquisition) and an ASIP (efficient for high precision

computations) can be embedded on the sensor module for better performance. It was shown in [22] that the overall power consumed by a wireless sensor module performing heart rate recognition from the Electro-Cardio Graph (ECG) data using a real time algorithm is reduced by a factor of 28 % compared to a single processor platform. The efficiency of the multi-processor solution will be threatened by the complexity and cost of the overall system which was the primary reason for not considering an ASIC microcontroller among other candidates, yet Chapter 5 addresses an approach suitable for an ASIC microcontroller.

Table 2.3 shows an overview for the key features of the two most popular commercial low power microcontrollers in existing hardware platforms. Many of the sensor modules are currently employed for a number of monitoring applications including habitat monitoring [23], structure health monitoring [24], etc. Most of these applications require operation at low speed and the amount of computations are limited. Evaluating the performance of an MSP430 for a high speed application was tested in [22]. The microcontroller was responsible for acquiring the ECG data and performing real time processing at 500 Hz and detecting the QRS signal as in [25] to determine the heart rate. The total power consumed by the MSP430 (RISC) was almost half the power consumed by the CISC (8051) for the same procedures. The largest amount of power consumed was during the processing mode followed by data acquisition and finally communication since only the detected heart rate was transmitted without transmitting any of the acquired ECG signals.

Table 2.3: Microcontrollers used in low power sensor modules

Feature	ATmega 128L [14]	MSP430 [10]
Company	Atmel	TI
Data size (bits)	8	16
Flash (kB)	128	256
RAM (kB)	4	16
ADC (bits)	10	12
Operating voltage (V)	2.7 - 5.5	1.8 - 3.6

2.2 Monitoring using Wireless Sensor Networks

Some monitoring have been performed solely by human observers to efficiently detect the occurrence of a certain phenomena. With the technological advances especially in wireless communications and sensor systems, it has become more effective to implement an automated monitoring system that performs complex tasks, covers a large geographic area, and is able to take certain actions in response to changes in the environment. Typically, the monitoring system may be required to perform either condition monitoring or surveillance.

The vital tasks performed by electric machines (wind turbines, generators, etc.) and the financial commitment of operators require efficient monitoring and fault diagnosis schemes that are capable of early detection and identification of defects. The major electric machine faults include bearing defects, stator faults, broken rotor bar and end-ring, and eccentricity-related faults [26]. These faults may lead to increased vibration and noise levels. Fault detection can be realized by monitoring machine vibrations, acoustic emissions, or motor current. Almost 50% of all motor faults are bearing related [26, 27]. Bearing faults can be categorized into outer race defects, inner race defects, ball defects, and cage defects. Condition monitoring using vibration

data has been successfully used for detection and identification of bearing faults [28]. The challenges in the early detection and identification process lie in the complexity of the generated harmonic structure, the need to estimate the variable rotational frequency, and prior knowledge of the bearing geometry.

Target detection and classification is required in several civilian as well as military (peacekeeping) applications. The cost of monitoring civilian construction sites such as oil pipelines [29] using human observers is very high. These construction sites cover a large region and the deployment of a reliable wireless sensor network over the same region to perform continuous monitoring would be more efficient and reliable than a human network. In cease-fire and peacekeeping territories, human observation is not sufficient to detect ground vehicles that perform illegal actions such as arms smuggling. Human observation (human eye aided by binoculars) is limited by the Line of Sight (LOS). When violence escalates, it becomes more dangerous to maintain human presence and in many situations a lot of effort and resources are provided to protect peacekeepers and human observers [30]. Historically, classification of ground vehicles found its way into the open literature first through civilian applications as in traffic monitoring [31, 32] and automatic classification of vehicles passing through expressway toll stations [33]. Acoustic detectors were proposed as early as 1970 [34] for vehicle actuated traffic signal light control and other control services. Regarding military applications, great demand for ground vehicle detection and classification evolved since technological advances in DSP, low power Very large Scale Integrated circuit (VLSI) design and wireless communications and the need to detect threats covertly approaching military assets. In the early 70's, a system called

REMote Battlefield Acoustic and Seismic System (REMBASS) was developed in the US [35]. Later on, an Improved REMBASS system was brought to the battlefield. The REMBASS system used three different sensors onboard (acoustic, seismic and magnetic) for detection and classification of target vehicles. In the late 80's, the Wide Area Munitions (WAM) project [35] was developed. Sensors onboard include an acoustic array of microphones and a seismic sensor. The WAM detects approaching vehicles, estimates vehicle's bearing and range, and deploys munitions when the target is at the closest point. An important mission for Unattended Ground Sensors (UGS) developed by the US army was to build an early-warning remote target recognition and surveillance system. During the early 90's, the US army showed more interest in utilizing acoustic sensors in identifying vehicles in the battlefield. Their goal was to develop low cost disposable sensor modules that are capable of discriminating between multiple targets [4]. They started developing low power signal processing schemes that enable individual sensors to perform detection, classification, localization and tracking functions with a guaranteed long operating life under typical conditions [5, 36].

In this section, we will present two important aspects with respect to monitoring using wireless sensor modules. The first addresses sensing modalities appropriate for low power sensor modules and the capability of those modalities in preserving the signature of underlying phenomena. The second, illustrates *in-network* computations, where individual sensor modules become responsible for performing different functions (data acquisition, detection, feature extraction and classification).

2.2.1 Sensing Modalities

Choosing the sensing modality for early detection and identification of bearing faults and vehicle detection and classification for peacekeeping operations represents a fundamental concern, since the performance of the whole system depends mainly on how much information is embedded in the extracted features that can be used in discrimination among all candidate classes.

To efficiently identify major faults in electric machines, many diagnostic methods has been proposed [26] that utilize one or more of the following signatures: noise and vibration signals, acoustic noise, electromagnetic field, radio-frequency emissions, and motor-current signals. Among all these measurements, monitoring the machine vibrations and performing current signature analysis has been widely used in performing early detection of machine failure, diagnosis, and prognosis. It was proven in the literature that the harmonic signature is evident in both the vibration and the current signals [27, 28, 37, 38].

Heavyweight and lightweight military ground vehicles can be sensed using seismic, acoustic, thermal, electric, magnetic, chemical or optical sensors. Among these sensors, seismic and acoustic sensors are characterized by their passive nature, Non-Line of Sight (NLOS) sensing, long range, and low power requirements and they can be used either during day or at night [1, 39]. Night reduces the utility of visual sensors, when the threat of surprise aggression is greater. Under these conditions, passive detection using acoustic and seismic sensors operating at their maximum level of performance are likely to yield substantial benefits. Acoustic-Seismic sensors are now found in various battlefield ground sensors, generally known as UGS [35].

The main challenge is defining the set of acoustic or seismic features that can accurately and efficiently represent the vehicle signature and are less sensitive to the vehicle speed, engine RPM, direction of motion, sensor-target separation, and background noise. The selected set of features must be capable of representing different types of vehicles. Acoustic features have proven to be quite useful in vehicle classification, equivalent seismic features less so [40,41]. The most significant limitation that affects the performance of acoustic sensors is wind noise [5]. While acoustic propagation is dependent on atmospheric conditions, the attenuation of seismic waves varies relatively more significantly from site to site and is very difficult to predict, making seismic sensors less able to discriminate between vehicles [29]. Seismically propagating frequencies above 100 Hz can decrease by orders of magnitude over tens of meters. Thus, reliable discrimination of ground vehicles depending exclusively on seismic signals is not promising for long range applications [42]. Certain applications require detecting other kinds of threats beside ground vehicles, such as moving personnel. In this situation, seismic signature can be more reliable in the discrimination of personnel among other threats. In [43–47] a number of algorithms were proposed for personnel and vehicle detection using seismic signals. Several approaches perform multi-modal processing combining seismic and acoustic features for increasing the robustness of the classification process [42, 48–52].

In many peacekeeping operations, civilian vehicles are used by competing parties in smuggling weapons which threatens established peacekeeping agreements. It is more challenging to be able to differentiate not only among large armored tanks but also between small Sport Utility Vehicles (SUVs) and pickups which produce lower acoustic emissions and their gasoline engines have more effective mufflers [53].

2.2.2 In-Network Computations

Primary monitoring scenarios divide sensor modules into sensing modules and central modules, where all data sensed by distributed sensing modules is collected at a central module for further processing (centralized approach) [54]. Considering the fact that most of the sensor module energy is consumed during communication (radio module) [55], an alternate solution was proposed such that computations performed by the central modules are performed *in-network* (distributed approach) [54]. In this case individual sensor modules are responsible of performing operations on the collected data. Different scenarios have been proposed that vary in the amount of computation carried out by the sensor module. Recent approaches [55] propose an energy efficient single point scenario. Single point systems are individual sensors that are completely self-contained; they are capable of performing the complete set of functions required including sensing, processing and communication in a completely standalone manner. Self-contained modules transmit decisions and vital information to a remote location [56]. These standalone sensor modules can benefit from the distributed deployment of the network where individual decisions can be combined at higher level modules (cluster head modules) to increase the reliability of the collective decision. Sensor array processing algorithms allow some computations on the sensor module while a higher level module performs a certain aggregate function on the combined data. Sensor array processing is generally required for spatially distributed target localization and tracking [57].

2.3 Characteristic Features

Feature extraction is a mapping of the original higher dimensional signal measurements into a lower dimensional feature space. Four feature spaces (time, statistical, spectral, and principal component analysis and combinations thereof) have been proposed in the literature for extracting the characteristic features for identification of bearing faults and classification of military vehicles. The optimum measure of the effectiveness of all extracted features is Bayes error [36]. In order to determine Bayes error, one would need to obtain the posterior probabilities, which requires performing a time consuming estimation of the nonparametric densities. We will determine the effectiveness according to the experimental classification simulations employing the same classification algorithm between the same set of detected events using the same data sets. In this section, we will discuss these four feature spaces and their popularity and general effectiveness with respect to each monitoring application.

2.3.1 Time domain features

Previous experiments on the effects of amplitude distortion on speech have shown that intelligibility is affected little by the type of amplitude distortion known as peak clipping [59]. It has been found that speech recognition is possible when the speech signal is reduced to a series of rectangular waves in which the discontinuities correspond to the crossings of the time axis in the original speech signal. A 97.9 % identification of speech waveforms has been achieved based only on the ZC (binary) information. The effect of amplitude clipping on the intelligibility of speech waveforms helped in

observing that a significant amount of speech information lies in the ZC. The number of samples between the ZC of the waveform contains the information necessary to discriminate clipped speech samples. This observation provided the motivation for the development of the Time Encoded Signal Processing and Recognition (TESPAR) method [60] that was subsequently applied to vehicle classification. Time domain encoding and feature extraction for classification of ground vehicles was proposed in [41, 61]. Both acoustic and seismic signatures were used for vehicle identification along with a feedforward neural network. The proposed algorithm generated a fixed size feature matrix for identification of the target independent on the duration of the time domain signal. This feature matrix represents the number of points of inflections (minima/maxima) between each two successive ZC. The acoustic and seismic feature vectors were represented through a histogram like structure. The authors in [41, 61] compared the acoustic and seismic features extracted at each individual sensor. Features extracted from the acoustic signal have been proven to be more effective in identifying military vehicles when compared to their equivalent seismic features. Although TESPAR is characterized by generating a fixed length feature vector regardless the duration of the time domain acquired signal, it requires defining a coding scheme that codes each pair of discriminators that characterize the duration and shape between successive ZC into a single alphabet. The coding table is highly dependent on the application and the modality, raising the complexity of generating the characteristic features [60].

We previously proposed extracting features from the time domain acoustic signal [62, 63] for vehicle detection and classification. The algorithm extracts features

from the harmonic model approximating the deterministic signature of the acoustic signal in the upper frequency band ($\geq$ Strongest harmonics component). Classification was performed using a feedforward neural network and was tested on a dataset with two heavyweight military vehicles.

By extracting the features directly from the time domain signal, we no longer benefit from the energy compaction of either the acoustic or seismic signal in the lower frequency band of their spectral representation. However, extracting the features from the time domain signal eliminates the need for an FFT or Discrete Wavelet Transform (DWT) before feature extraction. Time and frequency domain features can be evaluated according to the amount of computation (energy) required just before feeding the features to the classifier, length of the feature vector, and how effective the features are for discrimination.

2.3.2 Statistical measurements

Statistical measurements require relatively simple processing algorithms compared to the other feature spaces unless they are extracted from the spectral domain which requires computing the FFT coefficients prior to feature extraction [64]. Authors proposed statistical measurements to provide a fast and reliable feature space to discriminate between classes that are easily separable as in the separation between healthy and defective bearings [65]. Statistical measurements by themselves are more popular with identification of bearing faults [65, 66] and are usually combined with other time/spectral features for classification of military vehicles [62, 67].

Statistical measurements include signal energy, mean, variance, skewness, kurtosis, etc. [65]. Most of these measurements are normalized and/or weighted prior to classification. The number of measurements are determined according to the number of candidate classes and how close they are to each other in the feature space. Most of the statistical measurements are obtained from the acoustic, vibration, or current signals [64].

2.3.3 Spectral features

Most of the proposed vehicle classification techniques [2, 40, 55] transform the acoustic signal into another favorable space, such as the frequency domain, in order to preserve the characteristic features with fewer components. Usually the time domain acoustic signal is processed in short windows (0.25 - 1 second), such that the harmonic structure remains fixed during this interval. An important advantage for representing the acoustic signal using its spectral components is energy compaction in the frequency domain, where it was reported in [2, 5, 48] that more than 80% of the energy is typically observed in the frequency band $20 - 400$ Hz for ground vehicles.

Vehicle detection and classification algorithms that utilize spectral features begin by either obtaining the FFT components [2, 3, 36, 40, 68–73] or the DWT components [74] after signal detection. The differences among these published algorithms are in the feature selection phase that creates the characteristic feature vector and/or the use of a particular classification scheme. In [40, 73] the FFT is computed for events of 0.75 second duration. A 512 FFT characteristic vector is generated for each

detected acoustic event sampled at 4960 Hz. The first 100 components are selected with resolution of 9.6875 Hz carrying the information for frequencies up to ~ 1 kHz. Each 2 successive components are averaged into a single component to generate a 50-dimensional feature vector. The same procedure is performed for the seismic signal, where the first 50 components are selected to generate a 50-dimensional feature vector with resolution of 9.6875 Hz without averaging, carrying the information for frequencies up to ~ 0.5 kHz. They compared performance based on either processing acoustic or seismic signals using three different classification schemes (Maximum Likelihood (ML), k-Nearest Neighbor (kNN), and Support Vector Machines (SVM)). The classifier produces decisions between two candidate vehicles. The simplicity and the low computational power and storage requirements in applying the relatively simple ML technique to decide between two vehicles as compared to kNN and SVM, makes ML the preferred classification scheme in this scenario [40].

A common misconception is that increasing the dimensionality of the feature vector will result in improved classifier performance [2]. In an attempt to reduce the dimensionality of the feature vector while still making use of the harmonic model of the acoustic emission and preserving the vehicle signature, the energy of selected harmonics' was proposed for vehicle classification. In [75] the Harmonic Line Association (HLA) technique was developed by Robertson and Weber to create harmonic line sets for each second of data samples. The HLA technique organizes local spectral peaks that exceed the noise level into families of harmonically related narrow band lines. The presence of a harmonic set indicates the presence of a coupled harmonic signal source which is likely to be a moving vehicle. Using the HLA technique, only

harmonically related peaks were selected for vehicle discrimination. In [2], only the three most powerful harmonic components were selected for vehicle discrimination using a Bayesian Maximum a Posteriori (MAP) classifier. The drawbacks of this algorithm include: first, the HLA technique requires the effective resolution of the computed FFT components to be 1 Hz or less, which raises the overall number of computations to obtain the FFT coefficients; second, the fact that estimating specific harmonics' energy must be preceded by an exact estimate of the FF and the number of harmonics. Estimating the FF consumes a considerable amount of energy since this harmonic component for ground vehicles is not distinct. The great interest in the exact values of the FF for determining the harmonic series motivated the need for reliable and energy efficient approaches for estimating the FF. In [69] the performance of a variety of methods for tracking the non-stationary FF of military vehicles has been evaluated and it was shown that tracking such changes is not granted.

In [42, 52], spectral features representing the first 15 harmonically related peaks were selected for military vehicle classification. They showed that there is a relation between the order of the strongest harmonic component and the number of cylinders in the vehicle's engine. The features were extracted using an approach similar to the HLA algorithm. FFT components were obtained followed by line series analysis to define the harmonic structure/peaks, followed by estimation of the FF. Generated features were classified using Linear Vector quantization (LVQ). The processed events were of 0.82 second duration such that during each interval the FF is not expected to change by more than 10%.

In [3, 68], an ML estimate of the FF was proposed given that an initial estimate within 5 Hz of the exact value was obtained. By correctly estimating the FF and

fixing the number of harmonics, the remaining parameters in the coupled harmonic model are the amplitude and phase of each component. The coupling between the harmonics' amplitudes and phases estimated using the LS method was proposed to allow discrimination among military vehicles.

With regard to dimensionality reduction, other algorithms suggested that the selection of specific spectral components reduces the ability to classify the candidate vehicle. In [71], a simple vehicle detection and classification scheme was developed that uses the whole shape of the spectrum instead of a limited number of spectral components. They compared normalized spectral components between the testing vehicle and a reference vehicle. A single value is obtained by summing up the absolute value of the Euclidean distance between each two relative components. If this single value is less than a pre-defined threshold, the testing vehicle and the reference vehicle are considered the same vehicle.

Discriminating between different kinds of ground vehicles is a tough problem because, in many situations, the features extracted from the simple harmonic model are not sufficient to classify vehicles that have similar features. Since there are multiple sources that are responsible for the acoustic emissions of a vehicle, the simple coupled harmonic model in [3,68,69] does not model spectral lines that appear at non-integral multiples of the FF. In [76], two sets of characteristic features were extracted from the frequency response of the same acoustic signal. Each feature vector was fed to a different classifier. Two classifiers were selected, a Multivariate Gaussian Classifier (MGC) and a SVM classifier. A Bayesian-based decision fusion algorithm was applied to the output of the two classifiers (MGC and SVM). The Bayesian-based decision

fusion output achieved $7 - 11\%$ better classification rate than using a single characteristic feature. In [77], two different harmonic sets have been defined for generation of characteristic feature vectors, one set for verifying the engine and the other decides between tires and tracks. They utilized separate neural networks for classification.

In the continuing effort to define a robust set of acoustic features that allows reliable classification of civilian and military vehicles, two transformation techniques, the DWT and the Discrete Fourier Transform (DFT) have been used for feature extraction. Classification of these types of characteristic features extracted from two different data sets representing civilian and military vehicles was performed using a backpropagation neural network [74]. Identification of civilian vehicles was a more challenging problem, since the harmonic structure is less distinct in their acoustic signatures.

Several schemes that have been used successfully in speech recognition have been adapted for vehicle classification. Most of these algorithms extract features from either the frequency or the time-frequency representation of the acoustic signal. In [72], a biologically motivated feature extraction method based on a human hearing model was adopted for battlefield vehicle detection and classification. Different Vector Quantization (VQ) clustering algorithms were implemented and tested on actual recorded vehicles' sounds. The performance of different VQ clustering algorithms was explored. The results showed that the Parallel Tree Structured Vector Quantization (PTSVQ) and LVQ achieved the highest classification rates. In [78], Hidden Markov Models (HMM) based classifiers that use time-frequency analysis of the acoustic signal have shown superior performance in classifying events embedded in noise when compared

to other spectrum based classifiers. There are two challenges in applying speech recognition schemes: first, there is background noise and the effect of the surrounding environment on the acquired acoustic signal; second, most speech recognition algorithms require large amounts of memory, a high sampling rate and a significant amount of computational power which may not be available at our sensor modules.

2.3.4 Principal Components' Features

Principal Component Analysis (PCA) represents a favored feature space since PCA features show high degree of energy compaction. In the literature, PCA was utilized in [70] to project the frequency and time-frequency acoustic features into a low-dimensional subspace. Vehicle detection and classification benefits from the lower-dimensional principal components' features in two situations: first, consuming less energy in transmitting these features to a remote module for feature fusion; second, constructing a simpler classifier for discriminating between characteristic features. The penalty paid is more computation and more energy consumption at the feature extraction stage [55].

Principal components' features have been compared to two other feature extraction techniques [36]: the Power Spectrum Estimates (PSE) technique, and the HLA technique for classification of military vehicles using their acoustic signatures. PSE was generated using short time Fourier transforms of the sampled acoustic signal followed by selection of the low frequency bins to approximate up to the 200 Hz frequency range. HLA generates a characteristic harmonic line set for small windows,

subject to a fixed harmonic structure (number of harmonics equals 11) and an FF in the range $8 - 20$ Hz. PCA requires the computation of autocorrelation values for eigen-analysis using small data blocks, then the generation of the principal components using the original data, and projection onto the subspace containing the 11 largest eigenvalues. With an Artificial Neural Network (ANN) utilized for classification, PCA extracted features outperform both PSE and HLA extracted features but they require an increase in computational costs [36].

2.4 Classification

The performance of classification systems is highly affected by the diversity of candidate classes, the set of selected features, and the classifier itself. It was shown in [36, 58] that a key problem lies in the selection of proper feature vectors that will be stable and class specific. Several approaches focus on how much information is revealed in the feature vector without considering reducing the feature vector which subsequently reduces the computational requirements for classification, decreases the time required for training the classifier, and reduces the classification bias. Several classification algorithms that have shown significant performance with respect to classification of acoustic or seismic or speech signals have been used for identification of bearing faults and verification of military vehicles, such as MAP [2,3,5,39,40,57,79], k-nearest neighbor (k-NN) [40,70,73,79], LVQ [40,48], ANNs [36,41,50,61,74], fuzzy logic [80, 81] and SVM [40, 55, 76]. The most frequently used classifiers for vehicle classification are MAP and ANN. MAP is characterized by its minimal storage and

computation requirements if the number of candidate classes is small. ANNs are capable of efficiently discriminating between a wide variety of classes while reducing the effect of noise within the generated features. The hardware implementation of an ANN using VLSI benefits from the parallel structure of neural networks and performs the required computations (training and testing) in a shorter duration when compared to the software implementation of the same network [36]. In this section, three different classification algorithms will be discussed: LVQ, MAP, and ANN and their performance with respect to the extracted features will be evaluated in Chapter 4.

Linear Vector Quantization

A representative L-dimensional code vector is selected to represent a group of the L-dimensional generated FVs [82]. The set of all code vectors form a codebook. Training of the classifier takes place by generation of this codebook from a selected training set FVs. Classification of each new FV is done by comparing this FV with every code vector in the codebook to find the closest code vector to this new FV. LVQ has been widely used for performing compression and pattern recognition tasks. The performance of LVQ is very sensitive to the size of the codebook and the training set. Assigning more code vectors for each class allows more confidence in generating a wide selection of representative vectors. One of the main concerns in LVQ is generation of the codebook. In our simulations, we selected the splitting technique, introduced by Linde, Buzo, and Gray (LBG) in their paper [83], for generation of the codebook.

LVQ searches for the wining code vector by finding the closest vector to the characteristic FV in the L-dimensional space. If the L-dimensional characteristic FV is formed from different extracted features (time/spectral/statistical) concatenated in

one large vector, we must perform some sort of weighting because each characteristic feature has a different dynamic range and our cost function is the square Euclidean distance.

LVQ is a special case of one of the most accurate, yet time consuming, classification methods called kNN [40]. kNN classifies each FV according to the majority class of its k ($k \geq 1$) nearest neighbor in the training set. The time required to classify a FV is large, especially when k is large. LVQ is obtained by setting k to 1 and generating a codebook with a limited number of representative code vectors.

Maximum a Posteriori

Several applications that utilize wireless sensor networks for monitoring have considered MAP classifiers as the most feasible classifier for resource constrained modules, due to its minimal storage and computation requirements [84]. Assuming that the generated features for each of the C classes have been drawn independently according to the probability $P(x|w_i)$, $i = 1, 2, ...C$. Assuming that the $P(x|w_i)$ is a multivariate normal with the density [40]

$$P(x|w_i) = \frac{1}{(2\pi)^{\frac{L}{2}} |\Sigma_i|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(x-\mu_i)^T \Sigma_i^{-1}(x-\mu_i)\right] \tag{2.1}$$

where L is the length of the characteristic feature. The probability distribution is totally defined by the parameters (the mean μ_i, and the covariance matrix Σ_i). The set of parameters of this model will be denoted by $\Psi_i = (\mu_i, \Sigma_i)$. The training set of FVs is used in generating an estimate for Ψ_i for every class $i, i = 1, 2, ...C$. We assume that for each distinct class, there are n observations, drawn independently from the

probability density $P(x|\Psi_i)$. Then the log likelihood function can be represented as:

$$l(\Psi_i) = \sum_{k=1}^{n} \ln P(x_{ik}|\Psi_i), \tag{2.2}$$

The Maximum Likelihood Estimate (MLE) of Ψ_i, is the value $\hat{\Psi}_i$ that maximizes $l(\Psi_i)$. This estimate can be found by differentiating (2.2) with respect to μ and Σ which gives,

$$\hat{\mu_i} = \frac{1}{n}\sum_{k=1}^{n} x_{ik}, \tag{2.3}$$

$$\hat{\Sigma_i} = \frac{1}{n}\sum_{k=1}^{n}(x_{ik} - \hat{\mu_i})(x_{ik} - \hat{\mu_i})^T, \tag{2.4}$$

Using a set of discriminant functions $g_i(x)$, the classifier will assign the characteristic vector x to class w_i if $g(w_i|x) > g(w_j|x)$, for all $j = 1, 2, ...C$ and $j \neq i$. We set the discriminant function equal to the posterior probability $P(w_i|x)$ for minimum error classification. Therefore,

$$g_i(x) = P(w_i|x) = \frac{P(x|w_i)P(w_i)}{\sum_{j=1}^{C} P(x|w_j)P(w_j)} \tag{2.5}$$

Based on our assumption that the density $P(x|w_i)$ is a multivariate normal with known (or estimated) mean and variance, the vector x is assigned to class i if either $P(w_i|x)$ (or the log of the numerator of (2.5) which can be simplified as $lnP(x|w_i) +$

$lnP(w_i)$) is maximum. The log of the discriminant function can be written as:

$$ln(g_i(x)) = -\frac{1}{2}(x-\mu_i)^T\Sigma_i^{-1}(x-\mu_i) - \frac{d}{2}\ln(2\pi) - \frac{1}{2}\ln|\Sigma_i| + \ln P(w_i), \tag{2.6}$$

By further assuming that all classes are equal probable and by removing constant terms, the computation of MAP can be simplified such that the classified class is the one that minimizes $(x-\mu_i)^T\Sigma_i^{-1}(x-\mu_i)$, i.e., ML.

MAP classification is simple and takes advantage of assuming a normal distribution for the components generating the feature vector. Assuming equal distribution among all targets, the problem reduces to finding the minimum or maximum of a simple discriminant function among all possible targets. However, one problem is that the assumption of a normally distributed feature vector is frequently not true and only a limited number of applications have normal distribution features. One more advantage is the ability to easily add or remove targets from the set of available targets. However, the large number of required computations and the amount of energy consumed increases exponentially with the size of the feature vector. A system with a feature vector of length 50 and 3 or more classes, requires a large number of real multiplications during the training phase just to estimate the distribution parameters.

Artificial Neural Networks

ANNs derive their computational power from their ability to learn. Figure 2.2 illustrates the structure of a simple neuron. ANN have been proven to be very reliable in identification of bearing faults using the acquired vibration signature [67, 85] and

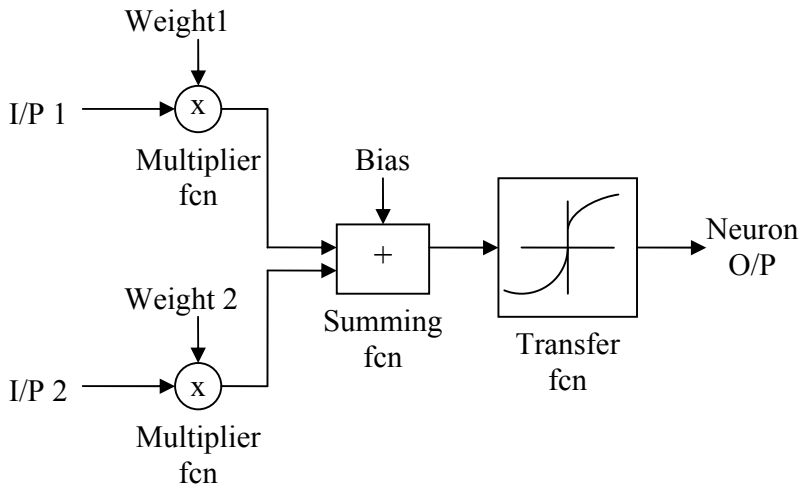

Figure 2.2: Structure of a Simple neuron

in classification of military vehicles using their acoustic signature [36, 53, 62]. Since extracted features might include some sort of corruption or redundancy because of propagation effects, background noise, etc., the feature extraction procedure and the classification technique must be tolerant to some extent to these problems. Because of the multilayer topology of ANNs, they explicitly carry out an additional feature extraction step prior to classification, where ANNs perform a non-linear transformation of the input data into a subspace such that classes become easily separable [86]. ANNs are formed from at least two layers, an input layer and an output layer. If we add one or more hidden layers, ANNs become capable of forming the internal subspace representation of the input data which allows an advanced discrimination compared to all other classification schemes [62].

There are two different Neural Network (NN) architectures, feedforward and recurrent (or feedback) networks. Feedforward NNs are simpler and require fewer computations and lower memory requirements because feedforward NNs are static (i.e.,

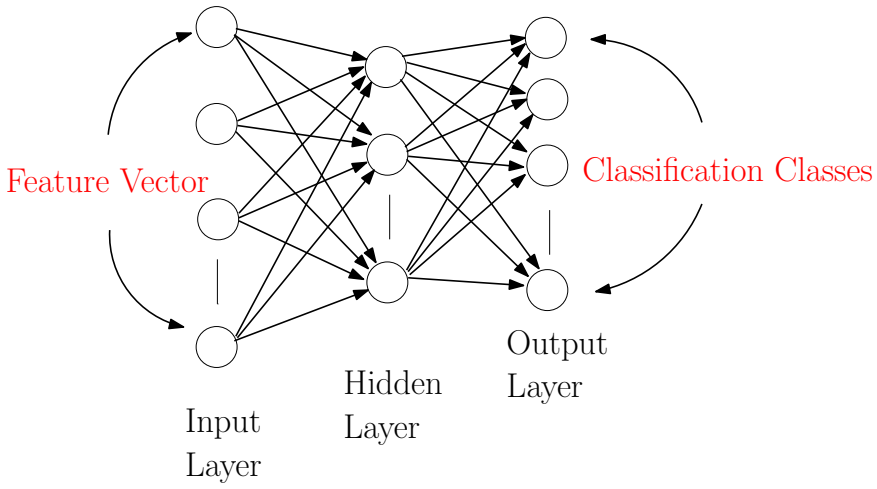

Figure 2.3: A three layer feedforward perceptron neural network

produce one set of output values for each given input) and memoryless with respect to their response. All these features are appropriate for our event driven application with the assumption that detected events are independent. The most common family of feedforward networks is called multilayer perceptron [87], where neurons are arranged into vertical layers that have unidirectional connections between them. Figure 2.3 shows the common structure of a three layer feedforward perceptron.

The number of neurons in the input layer is equal to the input FV length, and those at the output layer are equal to the number of candidate classes. Selecting the number of neurons in the hidden layer is very important. Increasing the number of neurons in the hidden layer will allow the input data to be more fitted. If the input data is noise free, the number of neurons in the hidden layer might be large and the only disadvantage may be a waste of the computational power. In the case of noisy data, if the number of neurons is too large, the noise components also become fitted [53].

The learning process in NNs is responsible for adjusting networks' architecture, weights, and biases such that the network becomes more capable of performing the required task. The network learns from the available training data (supervised learning), where the network is supplied with an output for every input. Neural networks are very sensitive to the training data and their performance degrades in the presence of new data outside their training set. In [77] a modification in the neural network architecture was proposed. Instead of using a single network with multiple outputs that are equivalent to the number of classes; they used multiple networks with a single output. The reason for that is to reduce the computational effort needed to re-train the network when new classes are added to the network. It has been shown that it is easier to train one neural network at a time [77].

There are four basic types of learning rules: error-correction, Boltzman, Hebbian, and competitive learning [87]. Error correction is the most popular supervised learning rule, where the error signal (difference between network output and the desired output) is used in modifying the network weights and biases. The back-propagation learning algorithm (Figure 2.4) represents one of the most widely used supervised error correction learning algorithms in pattern classification and it is known for its fast convergence and improved performance [36]. The following steps illustrate the error back-propagation learning algorithm (for a network of G layers) [87]:

1. Initialize the weights to small random values.

2. Choose the input training set and their associated targets (desired outputs).

3. Propagate the signal forward through the network.

4. Compute δ_i^G at the output layer

$$\delta_i^G = f'(h_i^G)[d_i - y_i^G], \tag{2.7}$$

where h_i^G represents the net input to the i^{th} neuron at the G^{th} layer, f' is the derivative of the activation function f, d_i is the desired output at neuron i, and y_i^G is the output from neuron i at the G^{th} layer.

5. Compute the deltas at each neuron for each of the preceding layer by propagating the error backwards

$$\delta_i^g = f'(h_i^g) \sum_j W_{ij}^{g+1} \delta_j^{g+1}, \tag{2.8}$$

for $g = G-1, G-2, ..., 1$, where W_{ji}^g is the weight from the i^{th} neuron at layer $g-1$ to the j^{th} neuron at layer g.

6. Update the weights using

$$\Delta W_{ji}^g = k \delta_i^g y_j^{g-1}, \tag{2.9}$$

where k is a constant value.

7. Repeat step 2 to 6 until the error (usually the Mean Square Error (MSE)) at the output layer is lower than a pre-specified threshold or a maximum number of iterations (epochs) is reached.

There are three more advantages for selecting an ANN for classification; First, neural networks are capable of delivering some sort of confidence to the final decision, which is very useful when fusion between individual decisions is considered. The

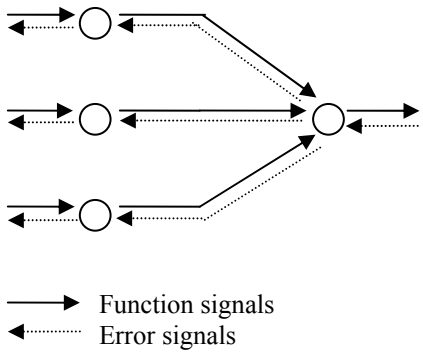

Figure 2.4: Error back-propagation learning algorithm

neural network is trained for pattern recognition problems such that a single neuron or more fires at the output layer for each significant class. In many situations, many neurons at the output layer have output values very close to the winning neuron, which might raise the question regarding the confidence in the output decision. Second, neural networks adjust their weights and biases for each neuron based on the propagated error, independent on the dynamic range of the remaining $L-1$ dimensions in the same FV. This helps in generating characteristic vectors from multiple features (spectral, statistical, etc.) without worrying about weighting the input vector. Third, the hardware implementation of an ANN using VLSI benefits from the parallel structure of neural networks and perform the required computations (training and testing) in a shorter time when compared to the software implementation of the same network. For example, a feedforward neural network employing error back-propagation learning for a network with 11 hidden layers for military vehicle verification, performed by the Army Research Laboratory (ARL) [36] was software

trained in 13 minutes, while the same training was performed using CNAPS [1] in 90 seconds.

In [50], three different neural networks paradigms (Back-propagation, Probabilistic, Radial Basis Functions (RBF)) suited for classification tasks were evaluated for military vehicle identification. Although all networks have shown similar performance with respect to the classification rate, the probabilistic neural networks, which implement Bayesian classification, consumed less time for training while the Back-propagation neural networks, consumed less time during testing. The key performance metrics for classification include energy and time consumed during training and testing, the probability of correct detection, or P_D, the probability of false alarm, or P_{FA}, and the probability of target miss, or $P_M = (1 - P_D)$.

2.5 Fusion

In-Network computations enable low power sensor modules to sense, process and identify events. One of the major advantages of wireless sensor modules is the dense deployment of the modules, which allows combining correlated captured data, features or decisions, increasing the reliability of the network and decreasing the effect of background noise [1]. Fusion can be performed over time, space or modality and it represents the key to fidelity and accuracy for unattended ground sensors. Data

[1] CNAPS (Connected Network of Adaptive Processors) is a hardware realization of neural networks using a group of analog parallel processors, based on VLSI technology and manufactured by Adaptive Solutions.

and decision fusion were used in [40] as a possible enhancement for the overall performance of the classifier, versus reporting decisions solely based on either acoustic or seismic events. In [76] fusion between two sets of acoustic features using two different classifiers was proposed. Three different fusion algorithms were examined for classification of military vehicles, feature level fusion, decision level fusion and modified Bayesian decision fusion. Collaborative processing between different sensors has also been used for tasks other than classification such as target counting, speed estimation, localization, and tracking [56, 73, 88].

Data and decision fusion have been successfully applied to several wireless sensor network applications. A principle concern centers around whether battery operated sensor modules are capable of handling the large amount of computations and memory requirements to perform fusion. The limited power of these modules reduces the number of computations that can be performed and the amount of data that can be transmitted to remote modules. Usually the confidence in the network's decision is increased as a result of applying fusion at one processing level or more [51, 76]. The main benefits obtained from applying data fusion for refining network decisions are:

- Increasing system reliability: gathering information from multiple sensors reduces the effect of random noise, hence producing more reliable decisions.
- Wide coverage area: Sensors at multiple modules are capable of delivering decisions based on spatial measurements.
- Forecast: Combining current and previous data from several sensors enhances the ability of the system to predict future events, which plays a significant role in the design of the network management protocol.

WSNs consider fusion at one or more of three levels (data level, feature level and decision level). Data level fusion is a low level model and is more appropriate for a decentralized processing scenario. Feature level fusion is a medium level fusion model and is utilized in distributed processing to reduce the computational burden on the modules by having multiple modules share in the generation of the characteristic features of any single observation, and then the combined FV is passed to a pattern classifier to identify the observation. Decision level fusion represents a high level fusion model and is the mostly used fusion technique in low power WSNs. At this level, each sensor determines its decision with respect to the underlying observation and a single decision is produced based on all individual decisions. Examples of decision level fusion methods include weighted decision methods (voting techniques), classical inference, Bayesian inference, and Dempster-Shafer's method [89].

Several examples have shown that multimodal data fusion results in reduced error as in determining the identity/location of a moving object. In [89], the error in determining the location of an aircraft using both a pulsed radar and an infrared imaging sensor is less than the error obtained by either of the independent sensors. The improved localization was obtained due to the fact that the radar has a higher ability in determining the aircraft range, while the infrared imagery sensor can accurately determines the aircraft's angular direction better than the radar. A number of approaches have used Canonical Correlation Analysis (CCA) for multimodal signal processing. In [90], Iyengar showed that the performance of a CCA detector (processing both seismic and acoustic data) is better than the detector exploiting individual modalities.

Single Modality Fusion

Beside the ability to report individual incidents, our model combines temporal decisions for verification of the vehicle's identity with no extra communications between modules. Temporal decision fusion concerns successive events that are detected by the same module within a certain period (indicating a single target). Spatial decision fusion requires determination of a cluster head, where the cluster contains 2 or more modules, to combine decisions. Spatial decision fusion requires intra-module synchronization and sending more data over the network. Spatial decision fusion will raise network traffic and is not preferred in low power wireless networks [89].

Multimodal Decision Fusion

The availability of multiple transducers at the same module, and the ability to measure different signals representing different physical representations of the same source enable better verification of the source. Although multimodal monitoring systems have shown enhanced performance applying multimodal data/decision fusion [91], finding an energy efficient methodology for computing the correlation and the dependence between different modalities for the same event represents the main challenge in both data and feature fusion scenarios [89, 91]. Decision fusion does not require determining the correlation between modalities, but features are extracted from each modality and the decisions of the classifiers are combined and a single decision is produced per module.

Multimodal Relay Fusion

A relay fusion model is proposed for target detection and classification using multimodal sensor data. The model does not require defining a correlation or a coupling

scheme for fusion between modalities. Instead of estimating the correlation between modalities to combine data obtained from multiple sensors or obtaining multiple decisions and combining these decisions using a specific fusion algorithm, relay fusion assign specific processing duties to certain modalities and pass on decisions to other modalities. Duties are assigned depending on the energy profile and performance with respect to the underlying observation. Relay fusion can be applied to any application that considers multimodal data obtained from multiple sensors.

In the proposed seismic-acoustic relay fusion model for vehicle detection and classification [92], acoustic features extracted from events that have seismic activity higher than the background noise are selected for target identification. The role of seismic data ends by defining the events and the role of the acoustic data starts with extracting the features. Acoustic data does not need to be processed for the event detection and there is no need to produce classification decisions based on the seismic data. Figure 2.5 illustrates the difference between multimodal decision fusion and the proposed multimodal relay fusion in addition to the combined relay-decision fusion model. The only requirement for applying relay fusion is some level of synchronization between different sensors/transducers at the same sensor module. Since the proposed multimodal fusion model is a decentralized fusion scheme, therefore synchronizing sensors over the same module can be simply achieved.

A distributed cluster-based algorithm for detection and classification of battlefield vehicles has proposed to benefit from the distributed deployment of sensor modules [79]. Dividing the monitored territory into clusters, where a single sensor module is nominated as the cluster head in each cluster. Correlated data is fused at the

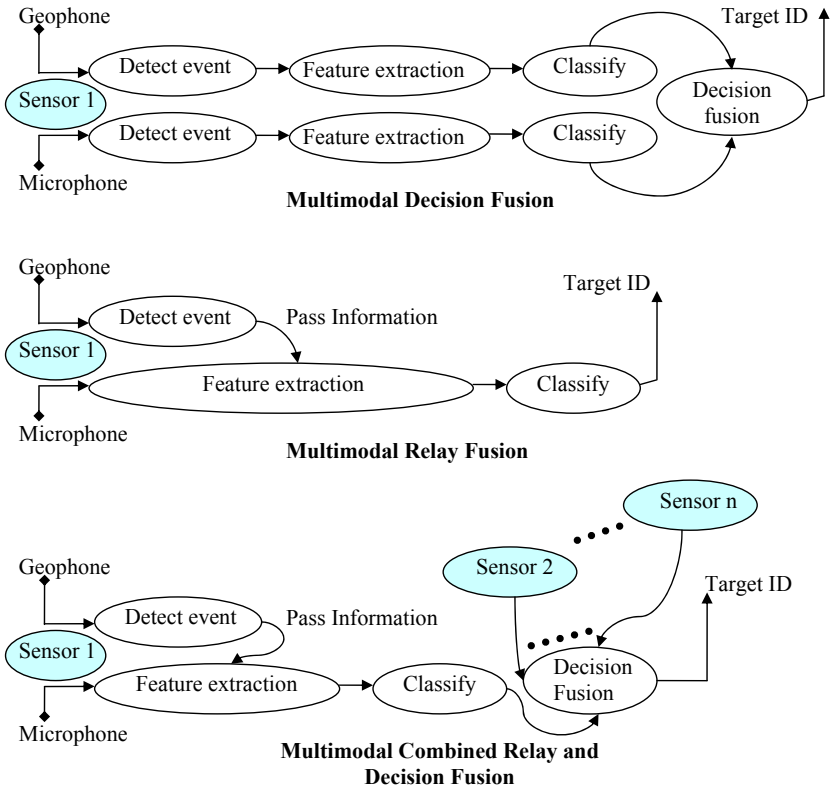

Figure 2.5: Seismic-Acoustic fusion models

cluster head for decision generation. Two approaches were proposed: the first combines extracted features; and the second combines individual decisions. Classification results using decision fusion and a ML classifier led to the best results.

In [91] different forms of collaborative signal processing (fusion techniques) were illustrated. The advantages and disadvantages of different combinations of data and decision fusion over both modalities (acoustic and seismic) and modules were discussed. They did not address temporal fusion, where individual decisions produced along successive events are combined to produce a single decision. Table 2.4 illustrates the most common collaborative signal processing schemes across both time, modality and space with emphasis on required network resources. Energy efficiency is determined according to the density of data transmitted across the network and the complexity of combining data/decisions. Usually data fusion between different modalities requires computation of a correlation function prior to fusion. Since data fusion can potentially yield the best performance at the cost of higher communication and computational burdens, data/feature fusion across modalities at the same module (no communication) and decision fusion across modules is attractive.

Multiple Vehicle Detection and Classification

More complex approaches for vehicle classification have been applied to the problem of multi-target detection and classification [2, 57]. In most of these approaches, the problem was reduced into a single target classification as follows, a bearing estimation and tracking scheme was initialized followed by beamforming in the direction of each spatially isolated target. This does not mean that classification of multi-targets can be always reduced to the simple case of single target classification. Acoustic emission

Table 2.4: Popular Collaborative Signal Processing Schemes

Scheme[a]	Overhead Communication	Distribution (Span)	Fusion Level	Target Discrimination	Energy efficiency
SM-SMd-ME	None	Time	Decision	Single	Very High
SM-MMd-SE	None	Modality	Decision	Single	Moderate
MM-SMd-SE	High	Space	Data	Single & Multiple	Low
MM-SMd-SE	High	Space	Decision	Single & Multiple	Moderate
MM-MMd-SE	High	Space & Modality	Data	Single & Multiple	Very Low
MM-MMd-SE	High	Space & Modality	Decision	Single & Multiple	Moderate

[a]SM=Single module, MM=Multiple modules, SMd=Single modality, MMd=Multiple modalities, SE=Single event, ME=Multiple events

of wheeled and lightweight vehicles is affected by the loud noise coming from the heavy and tracked vehicles which increases the challenge in separating between such sources [39].

Military Vehicle Counting, Localization and Tracking

The dense deployment of sensor modules helps in performing additional functions, beside detection and classification, such as vehicle counting, localization and tracking. Performing target localization and tracking generally requires sensor array processing, in which an upper level module is responsible for the localization and tracking of targets. Signal based localization techniques using the direction of arrival or time delay estimation techniques require significant wireless resources [93]. In [94–96], acoustic energy was proposed as an appropriate choice for multiple source localization, exploiting the fact that acoustic signal intensity generally attenuates with distance from the source. The ultimate goal for localization algorithms is to reduce inter-sensor communication and complexity to extend the lifetime of the sensor module

by transmitting only useful information to the upper level modules. In [96], multi-modal energy detection is used prior to localization to reduce the false alarm rate and filter out any events that result from background noise. The localization algorithm starts with an estimate for candidate target locations when targets are detected in the covered region. Then an estimation algorithm, such as ML, is applied to determine the most probable target location from among all candidate locations. In [91], target tracking was proposed for a single target problem and can be extended to the case of multiple targets if there is sufficient separation in space and/or time. In [2], an algorithm has been developed for counting military vehicles using a group of sensors arranged in a predetermined Field of View (FOV).

Chapter 3

Time Domain Zero-Crossings: Bearing Fault Classification

3.1 Overview

ZC features have been successfully used in many signal processing and pattern recognition tasks such as speech recognition [97], vehicle classification [63], and biomedical applications [98]. The simplicity in extracting such features makes them very attractive when compared to spectral, wavelet, and statistical features.

Several speech recognition techniques have been built utilizing ZC features since the observations made by Licklider and Pollack [99] who showed that speech intelligibility is preserved in the speech signal's ZC. The time intervals between successive ZC are related in a complicated way to the frequencies present in the time domain

signal. The features extracted from the signal's ZC are time domain features but carry information that is apparent in the frequency domain. The information can be extracted from the conventional signal's ZC and/or the signal's shape between successive ZC. While the conventional ZC are actually the signal's real zeros, the shape/perturbations (e.g., number of minima or maxima between ZC) are proportional to a subset of the signal's complex zeros [100]. Unlike the duration between successive ZC, which can be measured directly from the binary signal, extracting the information from the shape of the signal requires preserving the signal's amplitude. TESPAR [101] proposed generating two descriptors per interval[1], where these descriptors are coded using a symbol alphabet. The generated code is used for the production of the characteristic features. The alphabet is very dependent on the application [41, 101, 102] and to reduce the computations needed to generate the characteristic features, our proposed approach will generate the features approximating only the original signal's real zeros. The proposed ZC features will be generated based on the duration between successive ZC as shown in Figure 3.1.

3.2 ZC characteristic features

Unlike spectral approaches, which require sampling of the vibration, acoustic, or current signal with a high resolution A/D converter to detect the presence of a specific characteristic frequency or modulated sidebands, the only information necessary for generating ZC features is extracted from the polarity of the time domain signal. The

[1]The interval represents the time period between two successive ZC.

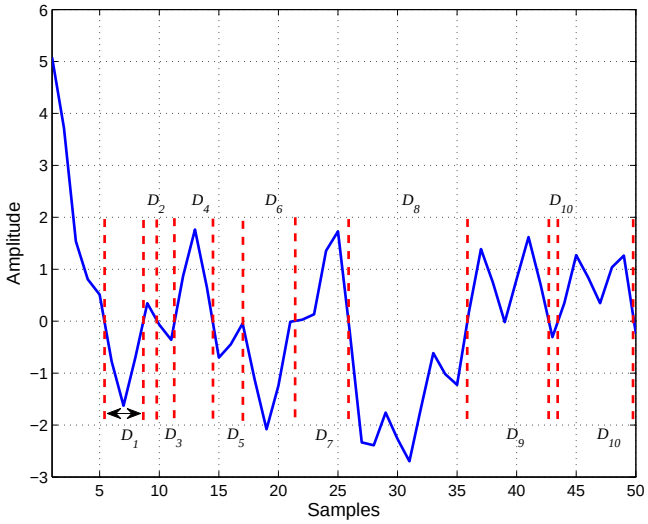

Figure 3.1: Duration between signal's ZC

ZC features are generated directly from either the count or the duration of the ZC intervals.

Two basic measurements are widely used for representing the information enclosed in the ZC signature [100]. The first is the density of the time intervals between successive ZC. The second is the excess threshold measurement, which is a function of the ZC intervals' durations. The second measurement can be directly obtained from the first measurement.

The time domain signal is processed in small observation windows, where the length of the window should be larger than the largest expected time duration between successive ZC. The generated feature vector is normalized by the duration of the observation window to be able to compare feature vectors generated using variable window lengths. Define T_Q as the longest expected time duration between successive ZC and T_L as the length of the observation window, where $T_L \geq T_Q$. Divide the

range $[0, T_Q]$ into Q intervals using $Q+1$ thresholds $(T_0, T_1, \ldots, T_Q)$, where $T_0 = 0$. The first measurement is simply generated by counting the number of ZC intervals in the current observation window whose duration falls between each pair of adjacent thresholds. Let c_j denote the count of ZC intervals whose duration falls between T_{j-1} and T_j, where $j = 1, 2, \ldots, Q$. The characteristic feature vector formed from the first measurement is

$$FV_{count} = \frac{1}{T_L}[c_1\ c_2\ \ldots\ c_Q], \tag{3.1}$$

The second measurement was developed by Thomas and Niederjohn [103] and it represents the fraction of the current time window for which ZC intervals exceed a certain threshold T_j, where $j = 0, 1, \ldots, Q-1$. The characteristic feature vector formed from the second measurement is

$$FV_{duration} = \frac{1}{T_L}[f(T_0)\ f(T_1)\ \ldots\ f(T_{Q-1})], \tag{3.2}$$

where

$$f(T_j) = \frac{1}{2}\sum_{k=j}^{Q-1} c_{k+1}(T_k + T_{k+1}). \tag{3.3}$$

where the product $\frac{1}{2}c_{k+1}(T_k + T_{k+1})$ estimates the sum of time intervals with duration falling in the range $T_k < T < T_{k+1}$, and c_{k+1} is the count of ZC intervals whose duration falls between T_k and T_{k+1}. The second measurement produces a descending feature vector with $f(T_0)$ equal to the duration of the observation window, T_L. This

reduces the effective feature vector length to $Q-1$ as compared to Q for the first measurement.

3.2.1 Computational Savings

A key benefit of the proposed approach lies in the computational power needed to generate the characteristic feature vectors as compared to other spectral or statistical features. The number of real multiplications required for generating the ZC features is $O(Q)$, where Q is the number of intervals. The number of complex multiplications required to obtain N FFT coefficients is $O(N \log_2 N)$ [104]. Each complex multiplication requires four real multiplications and two real additions. Spectral features will result in a larger feature vector and will either employ an optimization function to reduce the length of the generated feature vector or will require more computations by the classifier. When considering the memory requirements, ZC features are generated from the polarity of the acquired vibration signal (1 bit) which requires minimum data storage when compared to spectral and statistical approaches. ZC features are also characterized by generating a fixed length feature vector independent of the window length. Finally, careful implementation choices for the ZC features (e.g., multiplies by factors of two being replaced by shifts) will result in even greater simplifications.

3.2.2 Application to Bearing Fault Detection

ZC features are characterized by low computational and memory requirements. However, the question remains, to what extent are they useful in signal discrimination?

Independent of the selected classifier, simulation results in [63, 67] have proven that ZC features generated using only the duration between successive ZC are efficiently capable of discriminating between classes that have fewer variations among their signatures. Since ZC features select part of the signature for discrimination, the generated features should be easily separable. In this Chapter, we will assess the performance of the proposed ZC features in early identification of bearing faults using the vibration signals acquired by accelerometers mounted over the housing of induction motors.

3.3 Bearing vibration data model

Rolling bearings consist of two concentric rings, called the inner raceway and the outer raceway, with a set of rolling elements (balls) running in their tracks [85]. The balls in a bearing are guided in a cage that ensures uniform spacing and prevents mutual contact. The magnitude of the vibration data can be an indication of a potential defect and it can be utilized for differentiating between a motor with or without defects. Early detection and diagnosis requires additional processing of the vibration data since some defects will not produce a significant change in the vibration data until a complete failure is about to happen. A single point bearing defect will produce a harmonic series with a FF f_D equal to one of four characteristic frequencies depending on which bearing surface the defect occurs. Assuming that the inner race is rotating with the shaft while the outer race is fixed, these four characteristic frequencies are

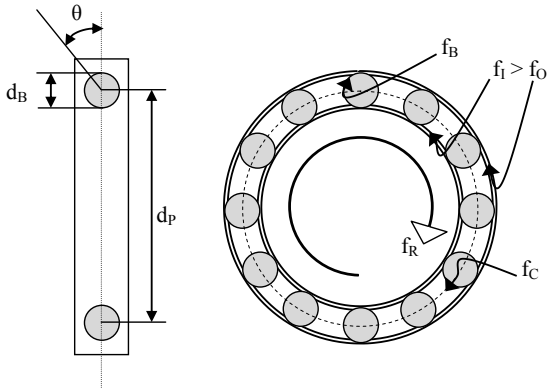

Figure 3.2: Bearing dimensions and characteristic defect frequencies

given as[2] [27, 85, 105]:

$$\begin{aligned}
\text{Ball defect:} \quad & f_B = \frac{d_p}{d_B} f_R (1 - \frac{d_B^2 cos^2(\theta)}{d_P^2}) \\
\text{Inner race defect:} \quad & f_I = \frac{N_B}{2} f_R (1 + \frac{d_B cos(\theta)}{d_P}) \\
\text{Outer race defect:} \quad & f_O = \frac{N_B}{2} f_R (1 - \frac{d_B cos(\theta)}{d_P}) \\
\text{Cage defect:} \quad & f_C = \frac{1}{2} f_R (1 - \frac{d_B cos(\theta)}{d_P})
\end{aligned} \tag{3.4}$$

where d_p is the pitch diameter, d_B is the ball diameter, f_R is the motor rotational frequency, θ is the ball contact angle (with the races), and N_B is the number of rolling elements. Figure 3.2 illustrates the bearings' dimensions and their relation to characteristic defect frequencies.

Since most bearing vibrations are periodic movements, it is more relevant to express the vibration signal in the form of harmonic signature. Motor vibrations can be

[2]The ball defect frequency is frequently represented as double the ball spin frequency, where the ball defect will hit both the inner and outer raceways each time the ball spins on its own axis [85,105].

modeled as the sum of three components. The first is a harmonic series approximating vibrations caused by mass imbalance or misalignment[3] having a FF f_R equivalent to the motor speed and M harmonics [105]. The second is a harmonic series approximating vibrations caused by mechanical shockwaves produced due to the occurrence of collisions when rolling elements pass over defect areas with a FF f_D[4] and L harmonics [27]. The third component represents the inter-modulation between the two harmonic series. A simple vibration model approximating motor vibrations can be expressed as

$$\begin{aligned} v(t) &= x_1(t) + x_2(t) + \alpha x_1(t) x_2(t), \\ \text{where } x_1(t) &= \sum_{k=1}^{M} A_{1k} cos(2\pi k f_R t + \varphi_{1k}), \\ \text{and } \quad x_2(t) &= \sum_{k=1}^{L} A_{2k} cos(2\pi k f_D t + \varphi_{2k}), \end{aligned} \tag{3.5}$$

where α is a scaling factor and $(f_R, A_{1k}, \varphi_{1k})$ and $(f_D, A_{2k}, \varphi_{2k})$ are the FF, and the amplitude and phase of the k^{th} harmonic for the first and second harmonic series, respectively. During normal operations and without any defects, only $x_1(t)$ can be detected. The second harmonic series appears with the start of a single point defect in the bearings and its harmonics' amplitudes increase giving rise to the sidebands of the third component with the increase in the severity of a defect [66]. The only two

[3]Mass imbalance is caused by the non-uniform distribution of mass about the center of rotation and misalignment occurs when the shafts of the driver and the driven machines do not align properly [105].

[4]f_D is equal to one of the characteristic defect frequencies.

components that can be estimated a priori in this harmonic model are f_R and f_D, while all other components are unknown to the analyst beside the fact that certain harmonics start to appear only at certain levels of defect severity. In order to illustrate the distribution of the vibration signal energy in the frequency domain with respect to bearing's condition, several spectra of the vibration data[5] acquired (in [105]) using an accelerometer are shown in Figure 3.3. It is clear from the spectra that during normal operation with no defects, the vibration energy is concentrated in the lower 2 kHz band with only three distinct harmonics. When a bearing defect starts to occur, the inter-modulation between the two harmonic series give rise to sidebands with most of the energy between $2 - 4$ kHz. In order to justify the occurrence of significant spectral peaks, the rotational speed should be precisely estimated and the bearings' dimensions and the number of balls must be known. This information may not be available in many situations [105].

The authors in [106] studied bearing fault induced failures and they observed that $> 80\%$ of the time, the variation in the vibration signature will follow the following four stages:

. 1^{st} stage, microscopic defects with almost no change in the spectrum since only shock pulses in the ultrasonic range (> 20 KHz) will appear.

. 2^{nd} stage, with the start of a single point defect, the impact between the rolling elements and the bearing races will excite the bearings' natural frequencies[6].

[5]The bearings have 9 balls, the pitch diameter to ball diameter $\frac{d_P}{d_B} = 4.9$, and the rotational speed $f_R = 29.13Hz$.

[6]The natural frequencies of most bearings lie between 850 and 1700 Hz [106].

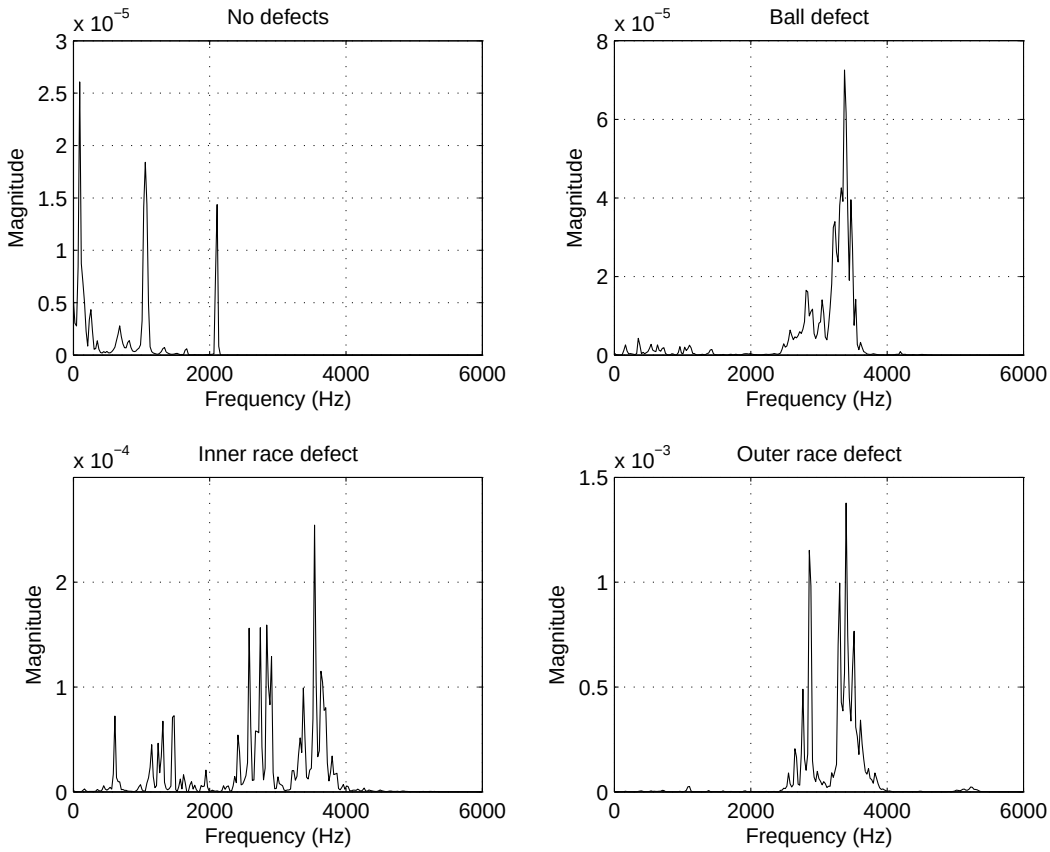

Figure 3.3: Vibration signal spectrum during normal operation and 3 single points defects

Later during the end of this stage, the natural frequencies become modulated with the motor's rotational frequency.

. 3^{rd} stage, defect characteristic frequencies begin to appear in the vibration spectrum and with the increase in the defect severity, those defect frequencies and their harmonics become modulated with the motor's rotational frequency.

. 4^{th} stage, the sidebands surrounding the defect frequency and its harmonics will grow higher in magnitude and additional sidebands (2x and 3x of the rotational frequency) will show up followed by random high frequency components just prior to complete failure.

In [106], it was shown through a series of experiments that starting from the 3^{rd} stage, other defects might show up in the vibration signature and the time it takes until complete failure is highly unpredictable. For more details regarding the variations in the vibration signature with increase in severity level, please refer to [106]. In figure 3.3, the rotational frequency was 29.53 Hz, and the characteristic frequencies for the ball defect, inner race defect, and outer race defect were 139.19 Hz, 159.91 Hz, and 105.86 Hz, respectively. It is clear from the spectra that during normal operation with no defects, the vibration energy is concentrated in the lower 2 KHz band with only three distinct harmonics. Although the rotating speed is at 29.53 Hz, there are two peaks at 1 KHz and 2.1 KHz, which can be related to the presence of a microscopic defect that caused the bearings to vibrate at the bearing system natural frequencies ($500 - 2000$ Hz for most bearing systems [106]). At the 3^{rd} stage, with the presence of a bearing defect, the defect frequency and its harmonics become

modulated with the rotational frequency, which give rise to sidebands with most of the energy between $2-4$ KHz. This is clearly seen in Figure 3.3. In order to justify the occurrence of significant spectral peaks and identify the defect type/source, the rotational speed should be estimated and the bearings' dimensions and the number of balls must be known [105].

In order to search among all candidate characteristic defect frequencies, the FFT can be computed with a spectral resolution around 1 Hz and the rotational speed of the motor should be precisely estimated. At a given rotational speed, bearings' dimensions, and number of balls, characteristic defect frequencies cannot be mistaken among each other but a subset of their higher harmonics might overlap [105]. Reliable fault identification using the vibration signal spectrum requires analyzing the whole spectrum which requires complex computations not suitable for low power microcontroller implementation on a sensor module. Even with precisely defining the bearings' dimensions, and fixing the load, the rotational speed will fluctuate because of external factors such as the performance of the controller, noise, and disturbance in the power system [85]. Figure 3.4 shows the spectrum of an inner race fault with an inner race defect frequency $f_I = 157.91$ Hz obtained with a spectral resolution of 0.75 Hz. It is apparent from the spectrum that distinct peaks represent only the inter-modulation sidebands. This relative level of signal harmonics has encouraged the development of amplitude demodulation (envelope) detectors for fault identification [27].

Figure 3.5 shows examples of the characteristic features extracted from accelerometer recordings for a rotating motor with no defect and three separately applied single point defects (Ball, Inner, and Outer). The observation window of length $T_L = 0.34$

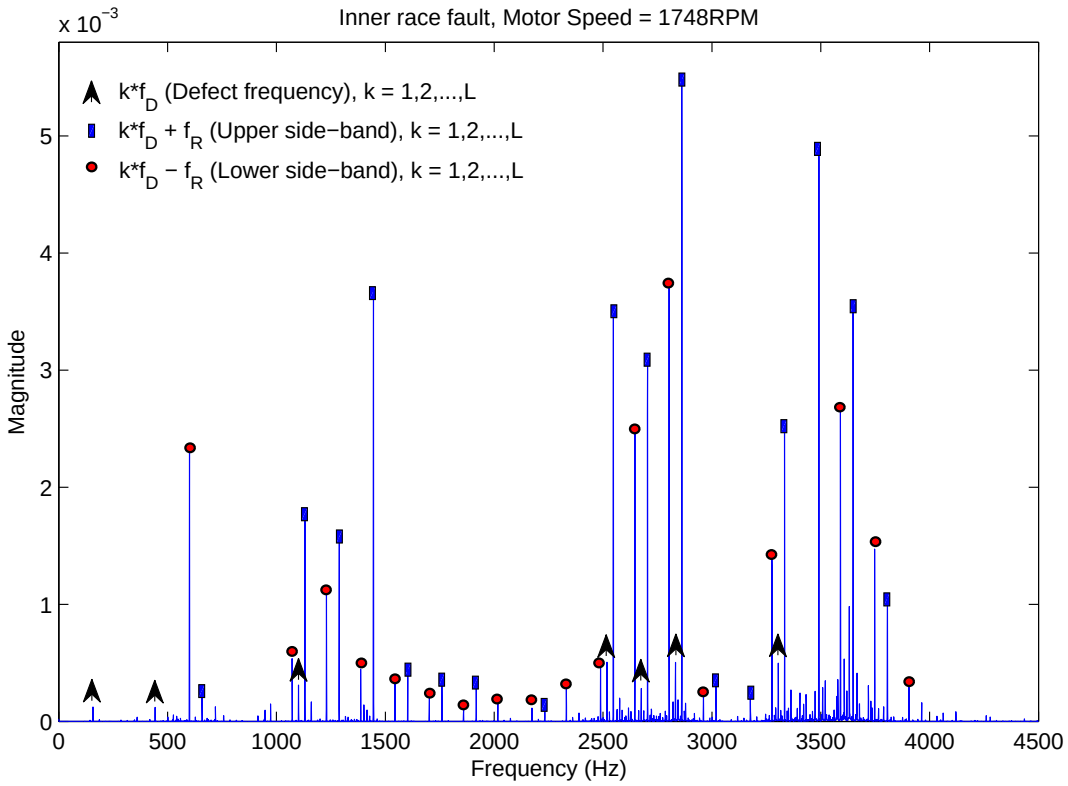

Figure 3.4: Spectrum for an inner race fault

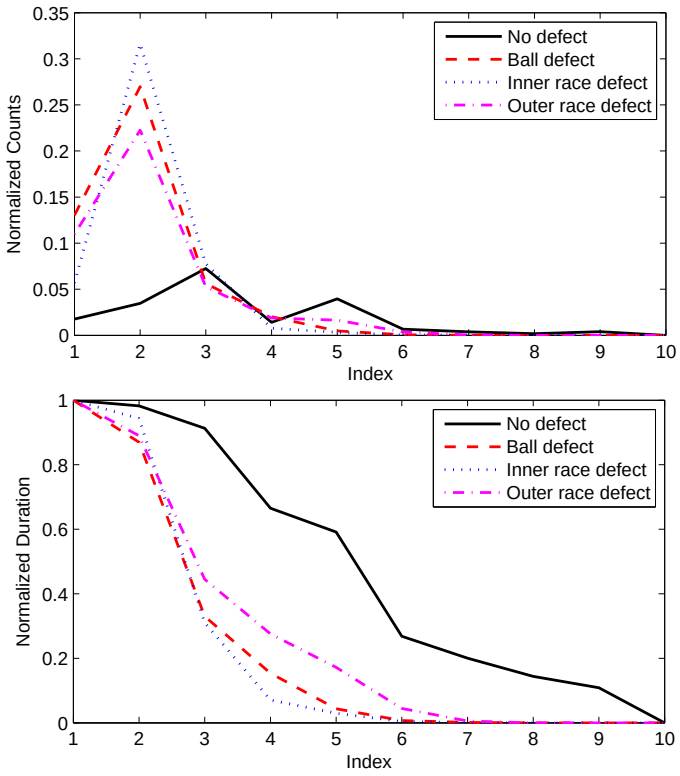

Figure 3.5: ZC characteristic features (FV_{count} - top and $FV_{duration}$ - bottom)

seconds is divided into $Q = 10$ intervals and the two previously described feature vectors are generated for each window. From the extracted features we can observe that for a motor with no defects either the first or the second type of feature vector is distinct compared to the features obtained from a motor with a single point bearing defect. The generated features for the ball defect and outer race defect are more similar to each other than to the inner race defect, which is consistent with the spectral representations shown in Figure 3.3.

3.4 Bearings data set

Vibration data was collected [28, 105] from the drive end bearing of an induction motor using accelerometers, which were attached to the housing with magnetic bases. Accelerometers were placed at the 12 O'clock position at the drive end of the motor housing. Vibration signals were collected using a 16 channel Digital Audio Tape (DAT) recorder. All data files are stored in Matlab (*.mat) format. Digital data was collected at 12000 samples per second. Speed and horsepower data were collected using the torque transducer/encoder and were recorded by hand.

Experiments were conducted using a 2 hp Reliance Electric motor (Figure 3.6). Data was collected for normal bearings and single-point defects. The bearings used are deep groove bearings manufactured by NTN (p/n 6205c3) with $f_B = 4.71$x, $f_I = 5.41$x, $f_O = 3.59$x, and $f_C = 0.40$x, where x is the rotational frequency. Motor bearings were seeded with faults using Electro-Discharge Machining (EDM). Faults ranging from 0.007 inches in diameter to 0.021 inches in diameter were introduced separately at the inner raceway, rolling element (i.e. ball) and outer raceway. Faulted bearings were reinstalled into the test motor and vibration data was recorded for motor loads of 0 to 3 horsepower (motor speeds of 1797 to 1720 RPM). The test stand consists of the motor, a torque transducer/encoder, a dynamometer, and control electronics. The test bearings support the motor shaft. SKF bearings were used for the 0.007, 0.014 and 0.021 inches in diameter faults. Drive end bearing specifications, including bearing geometry and defect frequencies are listed in Table 3.1.

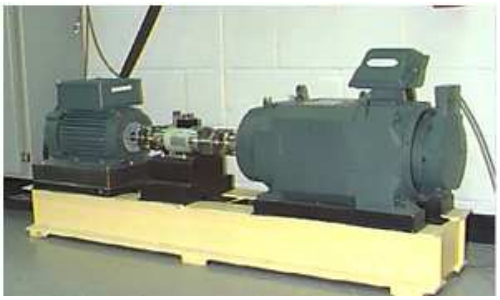

Figure 3.6: The test stand (the motor(left), the torque transducer/encoder (center), and the dynamometer (right))

3.5 Experimental results

A three-layer FNN with a sigmoid transfer function is utilized, where the number of neurons in the input layer is equivalent to the length of the characteristic feature vector and the number of neurons in the output layer is 4 corresponding to No-fault, ball fault, inner race fault, and outer race fault. The number of neurons in the hidden layer is usually selected to be sufficiently large so that the network is capable of adapting, as indicated by improved performance. Increasing the number of neurons requires storing more weights and biases and performing more calculations. So a trade-off exists between increased performance and reduced complexity. With these constraints, we varied the number of neurons in the hidden layer between 3 and 30 and we found that having fewer than 8 neurons in the hidden layer limited the ability of the ANN to adapt. In addition, there was no significant increase in discrimination capability with more than 12 neurons in the hidden layer. Therefore, the number of neurons chosen for the hidden layer in our simulations is 10. The ANN is trained for 200 epochs, where after each epoch, the weights and biases are updated once using a factor of 1.2 times the error derivative after calculating the ANN output using all

Table 3.1: The bearing specifications

Size	
Inside Diameter	0.9843 inches
Outside Diameter	2.0472 inches
Thickness	0.5906 inches
Ball Diameter	0.3126 inches
Pitch Diameter	1.537 inches
Number of balls	9
Defect frequencies (multiple of running speed in Hz)	
Inner Ring	5.4152
Outer Ring	3.5848
Cage Train	0.39828
Rolling Element	4.7135

available passes in the entire training data set. The performance function for the FNN is MSE. Figure 3.7 illustrates the decrease in the MSE with respect to training epochs. The FNN started to adapt to the variable data in the training set after just a few epochs and training for 100 epochs or more guarantees a good classification rate. Training up to 1000 epochs or more did not give any significant increase (or decrease) in the classification performance.

Classification results are given in the form of detection, false alarm and classification rates. The detection rate for each class is the ratio of the number of events correctly classified for the given class to the total number of events in the given class. The false alarm rate for each class is the ratio of the number events for all other classes classified as the given class to the total number of events in the given class.

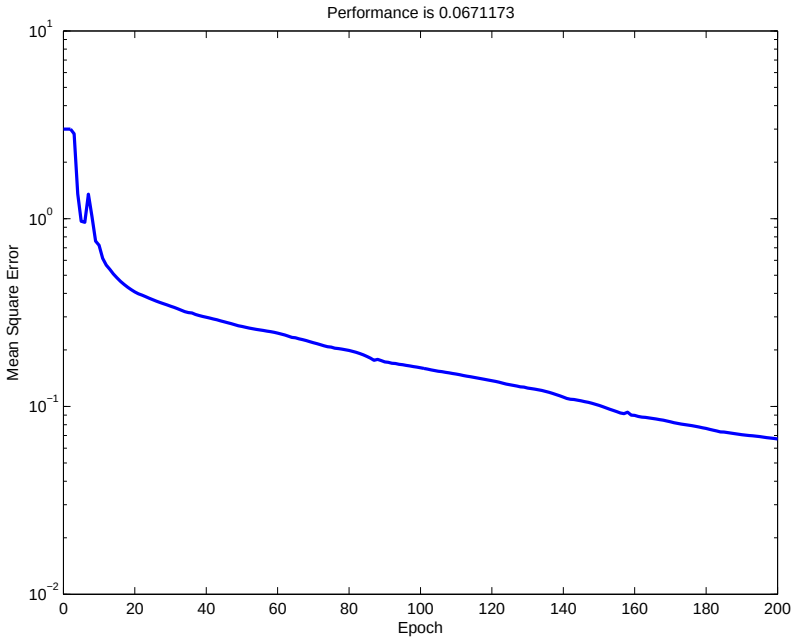

Figure 3.7: The FNN performance (MSE) vs. training epochs

The classification rate is the ratio of the number of events correctly classified for all classes to the total number of events.

The recorded vibration data was divided into observation windows, where each observation window indicates an event. Selecting a specific window length does not require a precise estimate of either the rotational frequency, or the bearings' dimensions. Since the length of the window should be larger than the largest expected time duration between successive ZC, the window length will be selected to acquire the lowest frequency that might show up in the vibration signal. From equation (3.4), and given the fact that $d_p > d_B$, and $N_B > 2$, the smallest frequency of interest will be the cage defect frequency. The cage defect frequency can be 0.25 or larger times the motor rotational frequency. Geometrical constraints limit the maximum number of balls that can fit in a bearing with a given pitch diameter to ball diameter ratio [105]. Thus, we can argue that for a given number of balls, the pitch

diameter to ball diameter ratio is at least a specific value which limits the expected cage defect frequency, which guarantees that the smallest frequency of interest cannot be lower than a specific value. This matches the argument in [106], that the cage defect frequency will always range from 0.33x rotational frequency and 0.48x rotational frequency and that the cage defect frequency will not normally appear as its fundamental defect frequency but will appear as a sideband of one of the other characteristic defect frequencies. Given that the rotational speed for the test motor is > 1700 RPM, the window length should be selected $> \frac{1}{2*0.25*f_r} = 0.07$ seconds. All of our window lengths exceeded 0.07 seconds.

Our first set of results are for Monte-Carlo simulations of 100 iterations. In each iteration $\frac{1}{3}$ of the total number of detected events are randomly selected for network training and the remaining $\frac{2}{3}$ are used for testing. For the given vibration recordings, the maximum time duration between successive ZC was 5.7 msec. The range $[0, 5.7]$ msec was divided into 10 and 20 intervals using the thresholds shown in Table 3.2. Different variations of threshold levels were tested and the best classification results were obtained for those with uneven ascending spacing as those shown in Table 3.2. This collection of thresholds outperformed both linear and logarithmic spacing.

Table 3.2: Thresholds used in defining ZC intervals

Intervals	Thresholds (msec)
10	0, 0.1, 0.2, 0.3, 0.4, 0.6, 0.8, 1.1, 1.4, 1.8, and 5.7
20	0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.1, 1.3, 1.5, 1.7, 1.9, 2.1, 2.3, 2.6, 2.9, 3.2, and 5.7

ZC features were evaluated for three different observation windows (0.085, 0.17, and 0.34 seconds)[7]. The two feature vectors described previously (FV_{count} and $FV_{duration}$) were generated for each event corresponding to the two measurements as illustrated in section 3.3. Tables 3.3 and 3.4 compare the results obtained for features generated by dividing the ZC intervals into 10 and 20 intervals, respectively.

Table 3.3: Bearing fault identification results for 10 ZC intervals - single event classification

Features		FV_{count}			$FV_{duration}$		
Observation window (seconds)		0.085	0.17	0.34	0.085	0.17	0.34
Detection rate%	No fault	99.97	99.71	99.88	99.97	99.82	99.29
	Ball fault	88.28	93.81	97.17	85.42	90.30	94.97
	Inner race fault	96.26	98.49	95.44	93.24	95.44	93.99
	Outer race fault	87.36	92.33	95.56	86.48	91.63	97.09
False alarm rate%	No fault	0.02	0.17	0.04	0.09	0.38	0.18
	Ball fault	4.00	2.32	1.45	4.81	3.02	1.81
	Inner race fault	0.51	0.46	0.78	1.38	1.52	1.00
	Outer race fault	4.33	2.02	1.46	4.77	2.28	1.64
Classification rate%		93.25	96.21	97.13	91.59	94.53	96.46

The classification results in Tables 3.3 and 3.4 demonstrate that the information contained within the ZC features was sufficient to specify whether the motor had a defect or not and, if so, to specify the defect type. The ANN was able to identify the kind of defect with high classification rates among all classes. Identifying a non defective motor was very high $> 99\%$ over all observation windows and using either

[7]The total number of events are 6000 for 0.085 second windows.

Table 3.4: Bearing fault identification results for 20 ZC intervals - single event classification

Features		FV_{count}			$FV_{duration}$		
Observation window (seconds)		0.085	0.17	0.34	0.085	0.17	0.34
Detection rate%	No fault	99.68	99.54	99.31	99.74	99.49	99.88
	Ball fault	87.73	91.75	95.63	83.56	88.94	92.10
	Inner race fault	93.88	95.35	93.46	92.80	93.02	93.25
	Outer race fault	87.11	92.48	95.46	86.60	90.75	94.48
False alarm rate%	No fault	0.28	0.35	0.04	0.13	0.24	0.04
	Ball fault	4.25	2.63	1.84	5.10	4.03	2.61
	Inner race fault	1.13	1.04	1.45	1.80	1.68	1.77
	Outer race fault	4.24	2.56	1.75	4.71	2.77	1.93
Classification rate%		92.45	94.99	96.11	91.06	93.34	95.15

of the generated ZC features. Increasing the duration of the observation window was more advantageous for identifying what kind of defect was present, where the classification rate for both ball and outer race defects increased by 4% or more each for doubling the window's width. The highest confusion is between ball and outer race defect, where 4% of ball defects were misclassified as outer race defects and vice versa for smaller observation windows. There was no gain for doubling the number of ZC intervals in the FV. Also note that having more intervals generates longer vectors and requires a larger neural network for classification. There is a slight improvement in the average classification rate using the feature vector generated using the first measurement (FV_{count}).

In order to evaluate the ability to identify defects outside the training set and

assess the ability of early detection of bearings' faults, the previous experiment was repeated using test features that were extracted from events with different defect levels than those used in training the neural network. Table 3.5 shows the classification results for a 0.34 second window, using FV_{count} obtained for the following:

1. Training using events with a smaller defect (0.007 inches) and testing using events with larger defects (0.014, and 0.021).

2. Training using events with a medium defect (0.014 inches) and testing using events with smaller and larger defects (0.007, and 0.021).

3. Training using events with a large defect (0.021 inches) and testing using events with smaller defects (0.007, and 0.014).

It is clear from Table 3.5 that, on average, training using events with a large defect is much better than training using events with a smaller defect when utilizing ZC features. Since the magnitude of the defect characteristic frequency and the intermodulation sidebands are proportional to the defect size, features extracted from events with larger defects are more easily separable from those extracted from events with no defects. The ANN adapts easily to separable features, which helps in identifying smaller, medium, and large defects without much degradation in performance. The ability to identify smaller defects is very important for early detection.

In order to evaluate the performance of the proposed ZC features in a more realistic scenario, where the classifier is not trained for all types of single point defects, the previous experiment was repeated using test features that were extracted from events

Table 3.5: Bearing fault identification results for 0.34 seconds observation windows and 10 ZC intervals - separate training and testing - single event classification

Features		FV_{count}		
Training defect size (inches)		0.007	0.014	0.021
Testing defect size (inches)		0.014 and 0.021	0.007 and 0.021	0.007 and 0.014
Detection rate%	No fault	99.89	99.96	99.79
	Ball fault	77.58	84.39	91.37
	Inner race fault	69.48	87.05	95.25
	Outer race fault	78.96	92.45	94.74
False alarm rate%	No fault	8.23	1.07	0.07
	Ball fault	5.90	1.95	1.65
	Inner race fault	5.13	3.81	1.16
	Outer race fault	4.19	4.40	2.92
Classification rate%		82.52	91.46	95.55

with different defect types than those used in training the neural network. The neural network is trained using features extracted from a no defect recordings and either a ball, inner race, or outer race defect and tested for features extracted from a no defect and the remaining two single point defects (using cross validation). In this case the neural network is trained to tell if the current event is for a normal bearing or a defected bearing. Table 3.6 shows the classification results for a 0.085 second window, using FV_{count}, and a 10 ZC intervals. It is clear from Table 3.6 that because the vibration signatures of defected bearings are distinct from normal bearings, ZC features are capable of preserving this distinction and the ANN was capable of identifying defected vibration signatures from normal non-defected ones with an average classification rate of 99.91%.

Table 3.6: Bearing fault identification results for 10 ZC intervals - single event classification - Unknown faults

Features		FV_{count}		
Training classes		No fault & Ball	No fault & Inner race	No fault & Outer race
Testing classes		No fault & Inner & Outer race	No fault & Ball & Outer race	No fault & Ball & Inner race
Detection rate %	No fault	100	100	100
	fault	100	100	98.91
False alarm rate %	No fault	0.00	0.00	0.91
	fault	0.00	0.00	0.00
Classification rate %		100	100	99.91

A simple way of increasing the detection rate of bearing faults is through decision fusion by combining individual decisions obtained from single events to produce a single decision. We selected a simple temporal decision fusion algorithm that combines consecutive individual decisions produced during a single recording that lasts for few seconds. The basic temporal decision fusion algorithm employs voting between individual decisions, which is a simple and efficient method with a minimum number of computations and minimal memory requirements. Table 3.7 shows the single event and decision fusion classification results obtained for training using events with

larger defects and testing using events with smaller defects. With decision fusion, the false alarm rate for the no fault class is zero, which increases the reliability in the identification system.

Table 3.7: Bearing fault identification results for 0.34 seconds observation windows and 10 ZC intervals - separate training and testing with fusion

Features		FV_{count}	
Training defect size (inches)		0.021	
Testing defect size (inches)		0.007 and 0.014	
Decision		Single	Fusion
Detection rate%	No fault	99.79	100.00
	Ball fault	91.37	97.72
	Inner race fault	95.25	98.86
	Outer race fault	94.74	100.00
False alarm rate%	No fault	0.07	0.00
	Ball fault	1.65	0.00
	Inner race fault	1.16	0.45
	Outer race fault	2.92	0.90
Classification rate %		95.55	99.02

Table 3.8 shows a comparison between the present work and three published works using different characteristic features generated from the vibration signal of rotating machinery and feedforward neural networks for classification. In [65], the authors used statistical parameters for classifying between normal and defective bearings. In [85] and [107], the authors used spectral features which requires fast processing, high precision, and a large number of complex multiplications compared to our approach using ZC based features. All four methods achieve a high classification rate,

but the actual benefit in the proposed approach lies in the simplicity in extracting the characteristic features directly from the polarity of the time domain sensor recordings without signal transformation or complicated analysis of the transformed coefficients. ZC characteristic features can be obtained for different window sizes as long as the window is long enough to acquire the largest expected time duration between successive ZC.

Table 3.8: Comparison between four different characteristic features used in identifying bearing faults

Reference	Features	Classified states	Classification rate%
[85]	FFT+Statistical	Normal, ball, inner, and outer race	86.8 - 96.2
[65]	Statistical	Normal and defective	98 - 100
[107]	Wavelet+PCA	Inner and outer	82.8 - 96.3
Present work Table 3.3	ZC (FV_{count})	Normal, ball, inner, and outer race	93.2 - 97.1
Present work Table 3.5 Table 3.7	ZC (FV_{count}) separate training & testing + fusion	Normal, ball, inner, and outer race	82.5 - 95.5 83.3 - 99.2
Present work Table 3.6	ZC FV_{count}	Normal and defective	99.1 - 100

Chapter 4

Time Domain Harmonics' Amplitudes: Military Vehicle Classification

4.1 Overview

There is a subset of harmonic signals that, although it includes multiple harmonic series, the energy for any given signal is highly concentrated in a specific series that can be used to characterize the source. With the focus of extracting features for classification, not for signal reconstruction, the proposed TDHA algorithm estimates the harmonics' amplitudes of a simple single harmonic structure. Since our motivation for developing the TDHA algorithm was primarily for the detection and classification of

military vehicles using their acoustic/seismic signatures for peacekeeping operations, it is appropriate to present the developed feature extraction algorithm in the context of the acoustic/seismic signature of military ground vehicles.

The chapter is organized as follows: Section 4.2 points out the need for an efficient monitoring system for peacekeeping operations, Section 4.3 describes the data sets used in the evaluation process, Section 4.4 presents the acoustic signature of ground vehicles, Section 4.5 illustrates the development of the TDHA algorithm and the computational savings compared to extracting features from the spectral domain, and Section 4.6 presents the experimental results.

4.2 Motivation: Bringing technology to peacekeeping operations

Peacekeeping is one among a range of activities undertaken by the United Nations (UN) and other international actors to maintain international peace and security throughout the world [108]. The UN has more experience monitoring cease-fire and peace agreements than any other organization. The UN peace operations have evolved considerably from early observer missions to armed complex multidimensional operations. The UN operations have relied almost exclusively on human observers, both military and civilian, for monitoring. It is difficult, if not impossible, to monitor large regions, and to maintain a permanent presence in distant locations relying only on human observers [30]. Furthermore, when violence rises, it becomes more dangerous

to maintain a human presence. Modern technology offers a number of alternatives to efficiently monitor vast territories when the presence of military observers is not recommended, especially during conflicts. Beside the extended range, unmanned monitoring offers the ability to record data and retrieve information which can be very useful for further analysis or as a reliable evidence for agreement verification. Modern monitoring technologies offer aerospace, ground and even underground continuous surveillance.

The UN has used some monitoring technologies in missions but mostly in an inconsistent manner. Simple digital and video cameras, often brought by the human observer, are now providing valuable photographic evidence, but they are not yet a regular part of the monitoring program [30]. The lack of a defined schema for an unmanned modern monitoring system makes it hard to rely on the collected data for monitoring agreements or even establishing the agreements [30].

Implementation of cease-fire or peace agreements is the principal motive for deployment of UN peacekeeping operations. Besides keeping peace, UN peacekeeping operations may also use force at the tactical level to defend themselves and their mandate, particularly in situations where the State is unable to provide security and maintain public order [108].

Since 1984 and there have been more than 60 UN peacekeeping operations around the world. As of July 2011, the number of peacekeeping operations is 15 [109], while the current operations directed and supported by the Department of Peacekeeping Operations (DPKO) is 16 (figure 4.1). The approved budgets for the period between July 1^{st}, 2010 and June 30^{th}, 2011 is about $7.83 billion [109]. More than 113,000

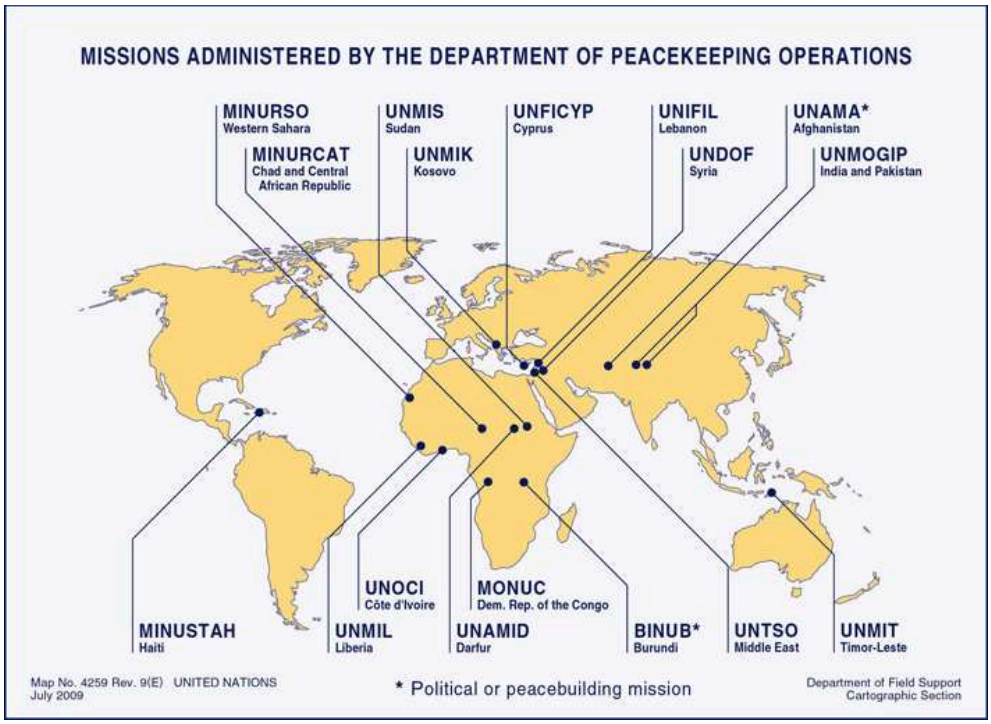

Figure 4.1: Current UN administered peacekeeping operations

persons are serving in these 15 operations. Troops and military observers represent more than 72% of the total numbers of persons involved in these operations [110]. Despite the huge increase in personnel and operations, peacekeeping remained an extremely cost-effective enterprise compared to its goals and achievements. The most recent successful peacekeeping operation was terminated on July 9^{th}, 2011 after the birth of the Republic of South Sudan as a result of a six-year peace process [109]. The annual budget for all UN peacekeeping is equal to approximately one half of one percent of global military spending [111].

The US Government Accountability Office (GAO) had reported to the Committee on Foreign Relations, U.S. Senate in [112] the challenges that face the UN peacekeeping unit when deploying any new operation. In their report, GAO pointed out that as

of September 2008, ongoing UN operations had about a 20% gap between troops and military observers authorized to carry out operations and actual deployments. The gap was about 34% for both police and civilians serving at these operations. These gaps reflect current UN difficulties in obtaining and distributing their resources to carry out operations. Lack of these resources and infrastructure would question the ability of the UN to provide peace and build confidence [112].

A commissioned study [30] for the UN-DPKO over the role of technology in improving monitoring and surveillance for peacekeeping operations proposed the rapid engagement of automated monitoring systems in all operations as a vital solution to fill the gap between mandates and capabilities. Ground sensing using a group of unmanned sensor nodes, for detection and identification of moving objects, represents a central element of every UN peacekeeping mandate; it is strange that monitoring technologies are missing from the organization agenda. This lack is emphasized by contrast to the UN-DPKO's use of state of the art communications systems, which are rapidly deployable anywhere on the globe, and capable of voice, video and data transmission at the operational level [30].

The UN does not have the state of the art technology or access to the high resolution satellite images and the information available to the US, UK, Russia, and China, hindering their efforts to maintain international peace and security. Many of the UN tools for monitoring territories are either too antiquated to perform efficiently or consume too much of their relatively scarce human resources. In summary, despite the rapid evolution of sensor technologies, the UN has been slow to apply sensors to the military or civilian domains of its peacekeeping operations. Several examples

from the UN-DPKO's reports show the importance of importing modern monitoring in current and future operations. For example in Rwanda 1994 [30], due to the lack of situational awareness, field personnel have found themselves in untenable situations. The force commander, General R. Dallaire, complained of being "deaf and blind" in the field.

Nefarious activities are much more likely to be carried out under the cover of darkness, rather than in the revealing light of day. So it is important for monitoring techniques to be able to operate equally well during day and night. Under darkness technical information complements human observation by creating a larger and more detailed picture of the area of operation. Unmanned ground sensors help peacekeepers perform better monitoring over larger areas, in rough territories, at night and in severe weather conditions. A dependable, reliable monitoring system will be very beneficial for all sides, the UN (or the monitoring organization) and the concerned parties. The UN will provide an efficient early warning system for the safety of its troops and observers and will efficiently deliver reports based on broad observations.

4.3 Data Sets

This section describes two different data sets that include recordings for different ground vehicles using a group of ground sensor stations. Our experiments use the raw time series data collected during both experiments. We will summarize the data collection procedures, available vehicles, and the deployed sensors for each data set.

4.3.1 The Bochum Verification Project (BVP) Data Set

The Bochum data set was collected in the context of the BVP [113], which investigates the potential of automatic sensor systems for verification of disarmament and peace keeping agreements. The project started in 1989 and several recordings and investigations have been performed in both Germany and the Netherlands. Their most recent data set of ground vehicles was collected during October 2000 on the site of the Bundeswehr Technical Center for Weapons and Ammunition (WTD 91) in Meppen, Germany [113]. The sensor deployment consisted of two sensor stations. Each station was equipped with a 24-bit A/D converter, a digital signal processor, 64 MB RAM and an EEPROM device emulating the hard drive in addition to a 10/100 Ethernet adapter for transferring the recorded data and results to a laptop PC. Each sensor station offered acoustic measurements sampled at 20 KHz and 3D seismic measurements sampled at 5 KHz using a microphone and three 3D geophone assembly (two horizontals and a vertical), respectively. The microphone nominal sensitivity was 50 mV/Pa and the geophone sensitivity was 26 V/(m/s).

Four parallel lanes were used at the site, one concrete lane and three sand lanes as depicted in Figure 4.2. The width of the concrete lane was 5 meters and the width of each sand lane was 6 meters. The length of the road was 400 meters, 200 m on each side of the main sensor line, with the two stations placed at the two different sides of the road half way between the start and the end of the road. Station 1 was placed 29.6 meters from the concrete road and station 2 was placed 71.8 meters from the concrete road on the other side (the separation between the two stations was 101.4 meters). Only one vehicle at a time was driven past the sensing stations while

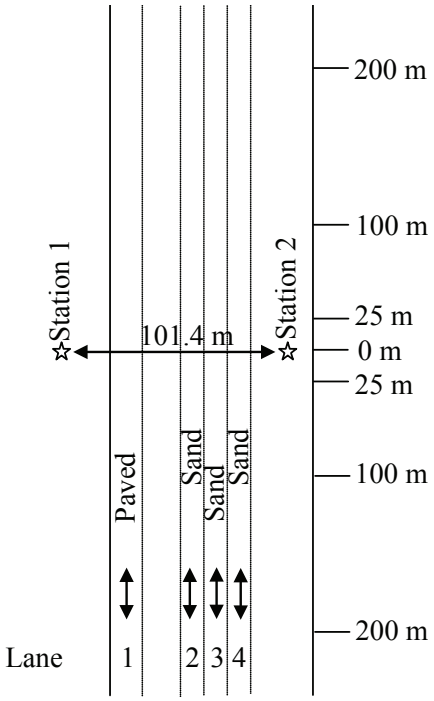

Figure 4.2: Layout of the sensor stations' positions

the data was collected. More than 365 runs ($>$ 19.5 Hours) were recorded at variable speeds with different directions and on different surfaces.

Five tracked vehicles (Leopard 1, Leopard 2, M48, Jaguar, and Wiesel) and five wheeled vehicles (Fuchs, Hermelin, MB 1017, Unimog, and a Volkswagen van) were used in generating the data set. Table 4.1 lists some of the important specifications of the vehicles. Each vehicle was allowed to pass several times in both directions with respect to the stations at the same speed. This process was repeated for seven different speeds. Although the system was constructed from only two stations, the large number of vehicles and several days of recordings allowed for the generation of a large data set with thousands of events. The vehicles can be classified into 4 different categories depending on type of track and the vehicle's weight; heavyweight

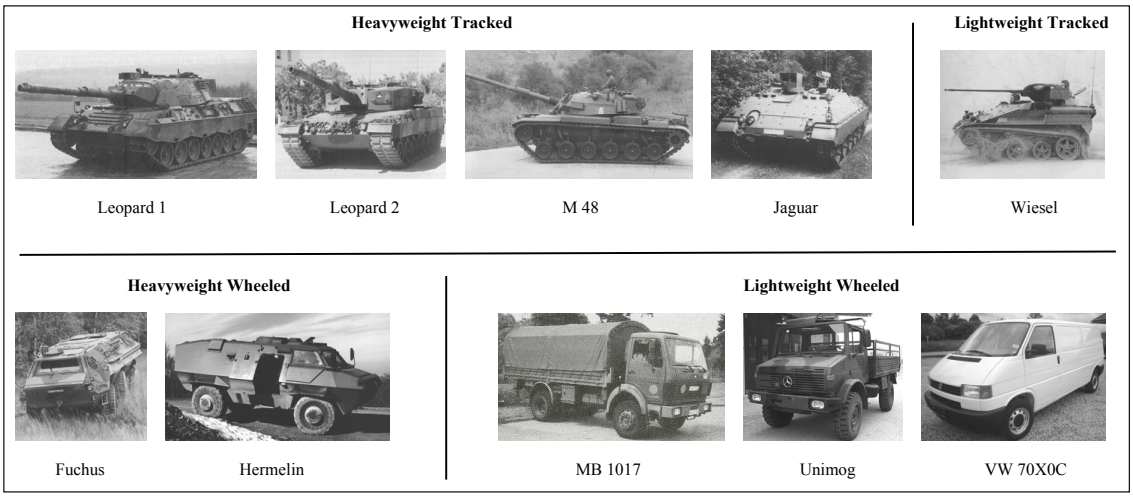

Figure 4.3: Four different vehicle classes

tracked, heavyweight wheeled, lightweight tracked, and lightweight wheeled. Figure 4.3 demonstrates the four categories with the types of vehicles considered in each category.

Table 4.1: Different vehicles' specifications

Vehicle	Type	Weight (Kg)	Number of Cylinders	Engine (cm^3)	Transmission
Heavyweight Tracked					
Leopard 1	Main Battle Tank	40200	10	37.4	auto
Leopard 2	Main Battle Tank	52000	12	47.6	auto
Jaguar	Tank Destroyer	23200	8	29.9	auto
M48	Main Battle Tank	45600	12	29.4	auto
Heavyweight Wheeled					
Fuchs	Armored General Purpose Carr.	14400	8	12.8	auto
Hermelin	Reconnaissance Vehicle	9745	6	5.7	manual
Lightweight Tracked					
Wiesel	Airborne Weapon carrier	1927	5	2.0	auto
Lightweight Wheeled					
MB 1017	Medium Truck	6800	6	5.7	manual
Unimog	Small Truck	5150	6	5.7	manual
VW 70X0C	Van	1713	5	2.4	manual

4.3.2 SensIT Project Data Set

The SensIT (SENSor Information Technology) data set was collected during the Wireless Distributed Sensor Network (WDSN) experiment, sponsored by DARPA, conducted at 29 Palms, CA, USA in November 2001 [40]. The purpose of the original experiment was to investigate the performance of a densely deployed sensor network in detection and classification of ground vehicles when a set of existing pattern classification methods are applied. The original experiment randomly deployed seventy five WINS (Wireless Integrated Network Sensors) NG 2.0 nodes [114] at the Marine Corps Air Ground Combat Center in 29 Palms, CA, USA. A map of the entire field with the distribution of the sensor nodes are shown in Figure 4.4. The field included three roads, the west to north road, the north to east road, and the east to west road. The sensor field was an area of approximately 900 by 300 meters and the nodes were deployed randomly along the sides of the roads with separation of adjacent nodes ranging from 20 to 40 meters.

The nodes were equipped with three different transducers, a microphone, a geophone and a polarized Infrared (IR) sensor, to enable nodes to record the acoustic, seismic, and infrared signals, respectively. Each node holds an A/D converter and an on-board programmable digital signal processor. We received the data collected at the third SensIT situational experiment (SITEX02), organized by DARPA/IXOs SensIT program. The acoustic, seismic, and infrared data were recorded with a rate of 4960 Hz. The recordings were for two heavyweight vehicles: a tracked vehicle, Amphibious Assault Vehicle (AAV), shown in Figure 4.5, and a wheeled vehicle, Dragon Wagon (DW), shown in Figure 4.6. Each recording (run) represents a single vehicle

Figure 4.4: SensIT field

moving along one of the three roads either with a constant or an accelerating speed. During SITEX00, research groups from BAE systems and Xerox PARC (Palo Alto Research Center) recorded low bandwidth seismic and broadband acoustic data for two types of tracked vehicles and five types of wheeled vehicles. The recordings at SITEX00 are not available for public use. The seismic data was sampled at 256 Hz. The BAE data was sampled at 10 KHz and the Xerox data was sampled at 20-40 KHz [73]. Although SITEX00 included seven vehicles, the classification schemes presented in [73] select between tracked and wheeled vehicles without determining the specific class of the vehicle.

The original recording was triggered manually which resulted in the duration of each run being several minutes. The set of detected events is expected to be relatively small, as these occur only when the vehicle is within a certain range with respect to the node. Usually an energy based detection algorithm is used for determining the

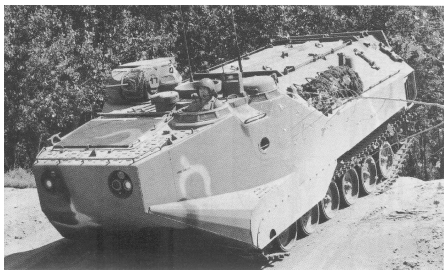

Figure 4.5: Assault Amphibian Vehicle

Figure 4.6: Dragon Wagon Vehicle

data that defines an event from the complete run recording. We have manually labeled individual events of duration 0.5 seconds to establish whether the acoustic signal reflects the presence of a vehicle or background noise. Manual labeling was performed by listening to the acoustic signal and labeling individual events. Listening was repeated several times before the final label was assigned. The purpose of labeling these short events was to provide a ground truth for events in the data set. In addition, this allows us to establish the efficiency of energy based detection schemes for different modalities. The labeling includes 16 randomly selected nodes over 6 complete runs, resulting in a total of 96 manually labeled runs for the two vehicles (over 3 hours of recordings).

The team that obtained the original recordings investigated the performance of three different classifiers (ML, SVM, and kNN) for the same set of characteristic features (FFT components) over two different modalities, acoustic and seismic [40]. They extracted features from the frequency band from DC to 1 KHz (acoustic) and from DC to 500 Hz (seismic) for events of duration 0.75 seconds. Their classifiers were trained to differentiate among three different classes (AAV, DW, and noise). The noise class was selected to identify events that show high acoustic or seismic activity but that were manually labeled as non-vehicle. Their results in [40] have

shown that acoustic features outperform seismic features for classification of these ground vehicles and that the ML classifier outperformed the other classifiers for the given set of vehicle classes. They proposed the use of multi-sensor data and decision fusion methods for enhancing the performance of the classifier [40].

4.3.3 Data set comparison

Both data sets were recorded separately at two different sites (USA and Europe) using different equipment, hardware settings and resolution. We find that the SensIT data set is characterized by the large number of distributed sensor nodes in the sensor field. Three different roads were included in the recordings and the efforts made in manually labeling individual events add reliability to the detector and allows for investigation of the performance of different detection parameters. The SensIT data set, which was recorded at the third site (SITEX02), was released for public use which makes the literature relatively rich with different approaches and results obtained based on these two specific vehicles (AAV and DW).

The Bochum project data set is characterized by the wide variety in vehicle classes and the availability of three different seismic transducers at the same node (station) with different orientations which helps in investigating modality usage for vehicle detection and classification. On the other hand, we found that the acoustic background noise was very low compared to the SensIT data set, so there is no need to include additional classes for background noise or perform manual labeling of individual events.

4.4 Vehicle's Acoustic Signature

The acoustic emissions of ground vehicles contain a wealth of information which can be used for vehicle classification [5]. Ground vehicles have two main sources of acoustic emissions: the engine and the propulsion mechanism [3, 48]. The acoustic signal can be modeled as a combination of two main components, a deterministic component and a non-deterministic component. The engine and drive train are responsible for the deterministic component and can be represented using a coupled harmonic signal model. Features extracted from this model can be very useful in defining the class or type of the vehicle. The non-deterministic component includes other acoustic noises generated by the engine that do not fit in the harmonic model and other acoustics produced due to the interactions of the tracks/tires with the road [29].

In the spectral domain, the acoustic emission is composed of a family of narrow-band harmonic lines with an FF directly related to the engine RPM [3,29,48] and other periodic components that arise from the tracks or the interaction between the tires and the road. For tracked vehicles, the track produces an additional series with another FF [42]. A popular coupled harmonic model has been proposed for modeling a large portion of the acoustic signature for military vehicles [3, 68]. Acoustic signals for a heavyweight tracked vehicle and a lightweight wheeled vehicle are shown in Figure 4.7. The time-frequency representations for both vehicles are shown in Figure 4.8.

The harmonic structure for both vehicles is clear from Figure 4.8. The vehicles' spectrographs include a set of harmonically related features that have been proven efficient in classification of military vehicles. The noise and a portion of the target's

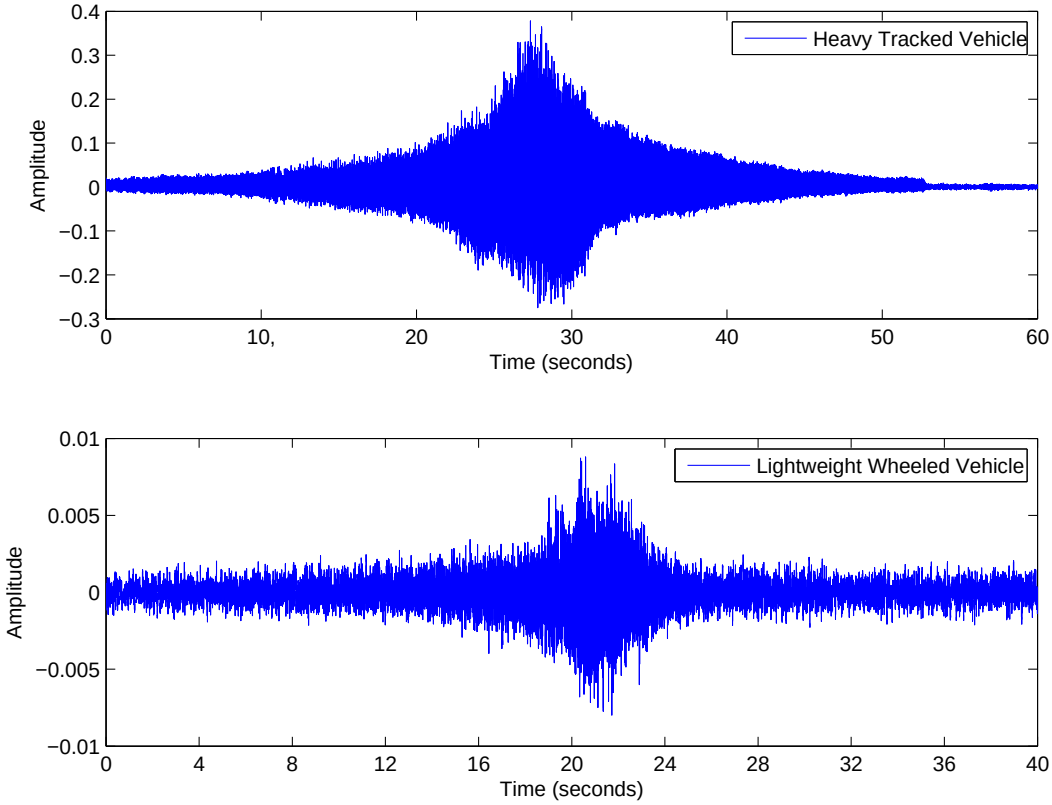

Figure 4.7: Time Domain Acoustic Signal (Fs $\sim$ 5 KHz)

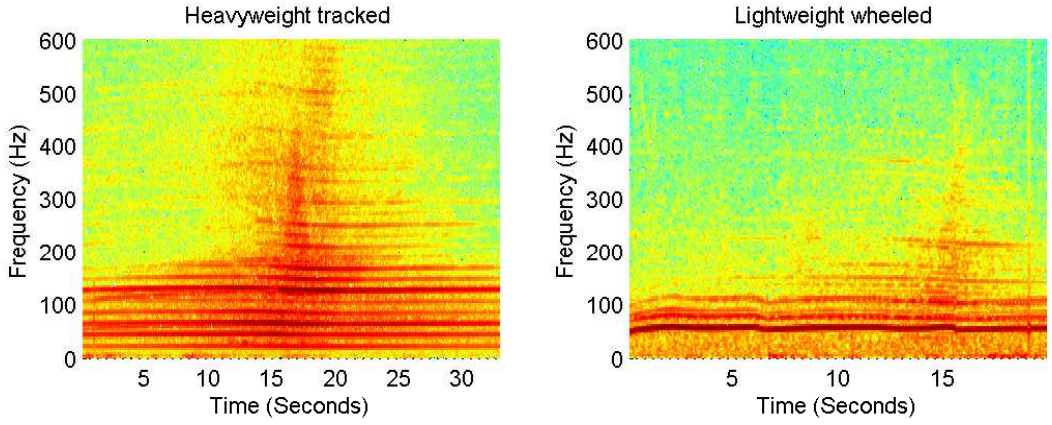

Figure 4.8: Spectral Representation (Fs $\sim$ 5kHz)

acoustic signature that do not fit into the coupled harmonic model are responsible for the remaining components in the signals spectra. The harmonic structure is more distinct in the heavyweight tracked vehicles than the lightweight wheeled ones. A sole dependence on a single set of harmonic features has proven to be a challenging problem due to the lack of uniformity for a specific vehicle with respect to the vehicle's speed, gear, direction of motion and the surrounding environment [76].

Vehicle detection and classification is an event driven application, where the process of feature extraction and classification is usually preceded with a detection phase. During target detection, sensor nodes decide whether the current event is equivalent to a target or just background noise. An event is the smallest interval through which sensor nodes are capable of producing feature vectors and, ultimately, decisions. Usually the average acoustic or seismic energy is computed and compared with an adaptive threshold that is updated according to the background noise level maintaining a Constant False Alarm Rate (CFAR) [40, 74]. Other approaches select a window of few seconds around the Closest Point of Approach (CPA) and treat it as a series of detected events [2, 80]. Although the detection phase appears to be a simple task, unreliable detection results in either missing a target or extracting features from a burst of background noise. Most sensor nodes stay in a low power mode (sleep) until the average energy of a certain number of successive events exceeds the background noise [42, 52].

4.5 TDHA characteristic features

The performance of every verification algorithm is largely determined by the feature extraction scheme and classification approach employed. Feature extraction represents a fundamental sensor module task that typically follows communication in consuming power [55]. The harmonic structure that appears in the vehicle's acoustic emission also appears in the generated seismic waves. The acoustic and seismic spectra of four different vehicle categories (heavyweight tracked, heavyweight wheeled, lightweight tracked and lightweight wheeled) are shown in Figures 4.9 and Figure 4.10, respectively. Averaging over nine different vehicles (Bochum data set), 96.64% of the energy of the acoustic emission was found to be in the range $0 - 250$ Hz and 98.27% of the energy of the seismic emission was found to be in the range $0 - 100$ Hz.

The concentration of the acoustic and seismic energy in the lower frequency bands motivates extracting features from this band. In the literature, several schemes [40, 73, 88] generate the characteristic features directly from the spectral domain. Other schemes search for the coupled harmonics from the peaks in the spectral representation of the signal and select a representative number of these harmonics to identify the detected vehicle [2–5,48,68,69]. The former schemes compute the spectral components and select a smaller number of components averaging both the deterministic and non deterministic components of the signature. The latter schemes require searching for the peaks that generate a coupled harmonic series and select their power in classification relying solely on the deterministic signature. However brute force searching for the harmonic series without knowing the FF is a computationally intensive operation for individual sensors to carry out at the sensor node. Although averaging spectral

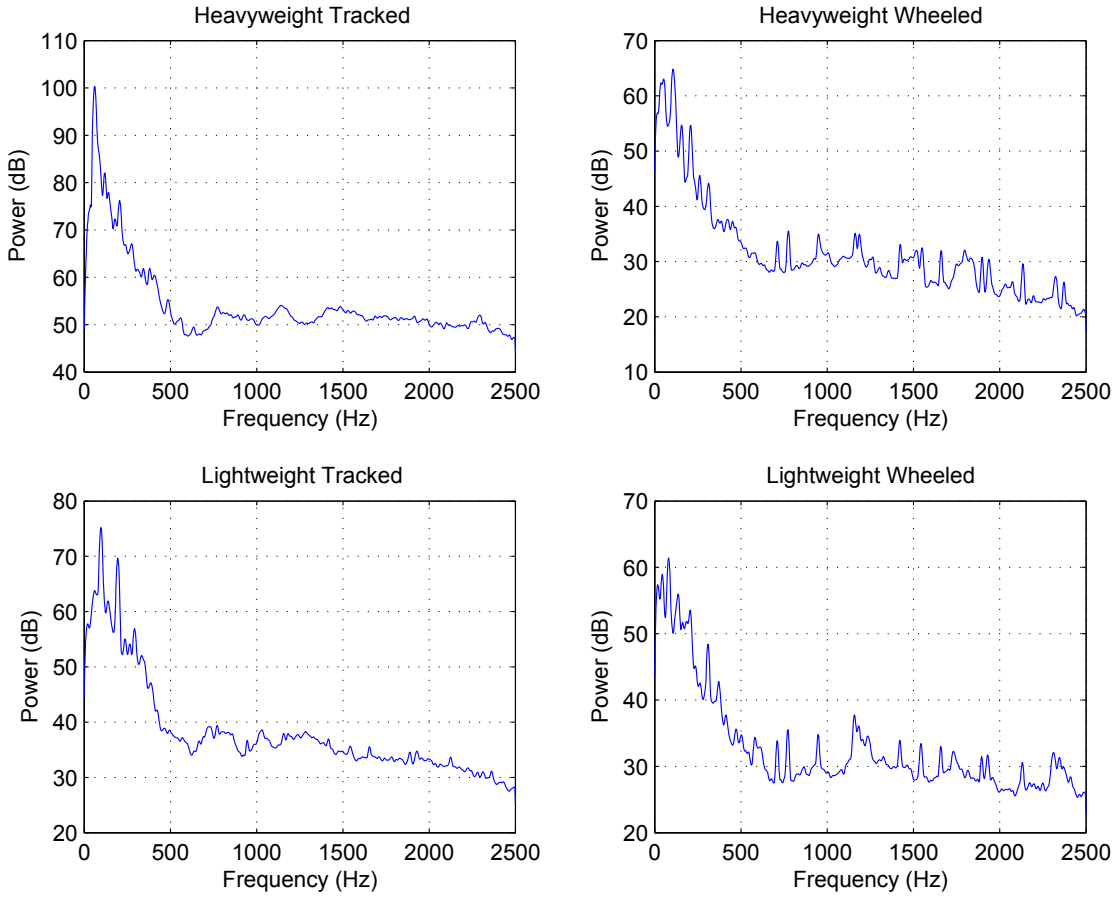

Figure 4.9: Acoustic spectrum of 4 different vehicle categories

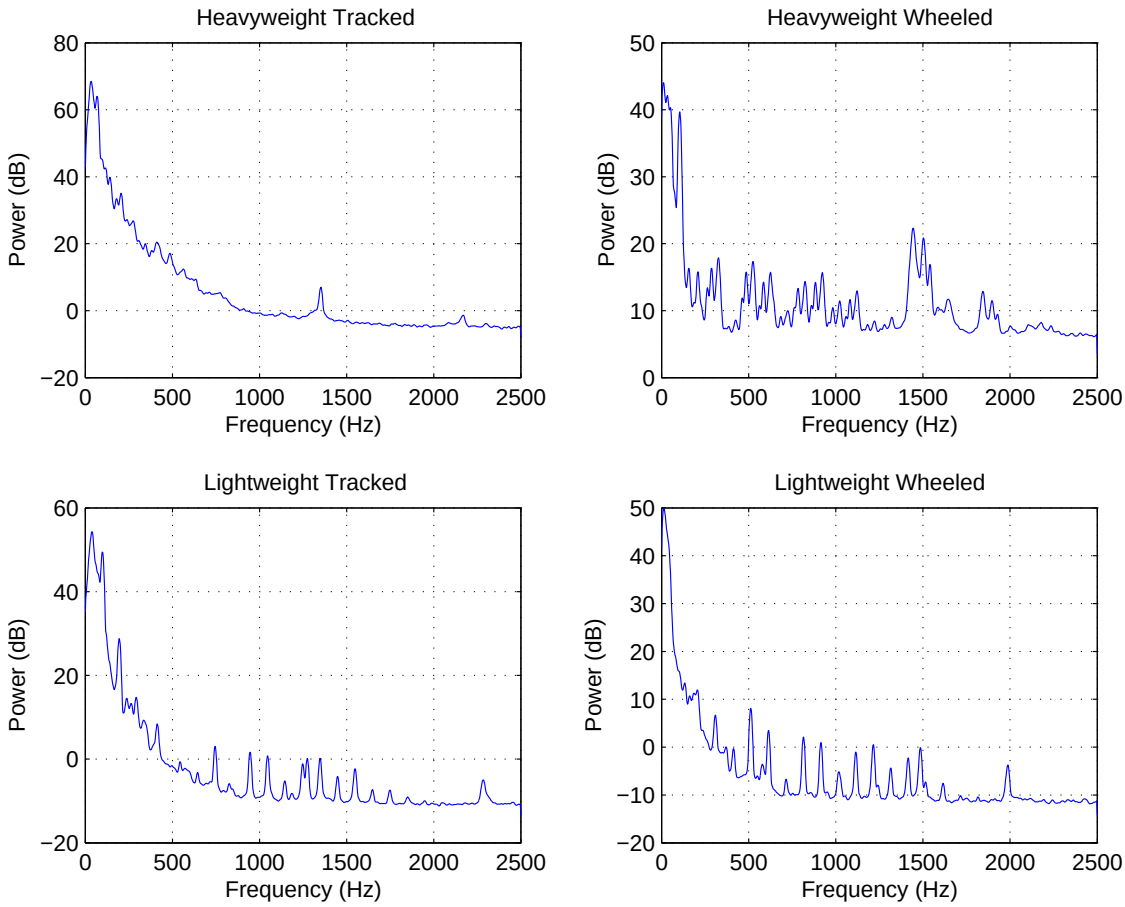

Figure 4.10: Seismic spectrum of 4 different vehicle classes

components consumes less power than searching for the harmonic series, selecting only the deterministic components has been proven to be more efficient for vehicle classification [2].

The acoustic and seismic spectrograms of four different vehicles (heavyweight tracked, heavyweight wheeled, lightweight tracked, and lightweight wheeled) are shown in Figure 4.11 and Figure 4.12, respectively. The harmonic structure is clear in the acoustic spectrogram of the four different vehicles and the structure becomes clearer when the vehicle is closer to the sensing node. The harmonic structure appears as well in the seismic emission of the vehicle, but not as distinctly as it appears in the acoustic emission. The following observations are found to be true regarding acoustic emissions of military ground vehicles:

1. The harmonic structure varies with time but it can be considered stationary during events of short durations ($\leq$ 1 seconds[1]) [52].

2. Not all of the harmonics are present during the whole run.

3. The strongest harmonic is not the fundamental component and it shows up in the band 50 - 150 Hz for most vehicles.

4. Harmonic lines (deterministic) are responsible for most of the energy when compared to the non-deterministic components.

Since the deterministic harmonic structure of the acoustic emission is more consistently present than it is in the seismic emission, we will focus more in our discussion

[1]In [52], a window size of $\leq$ 0.82 seconds is recommended such that the harmonic structure of ground vehicles will not change by more than 10% during the same event.

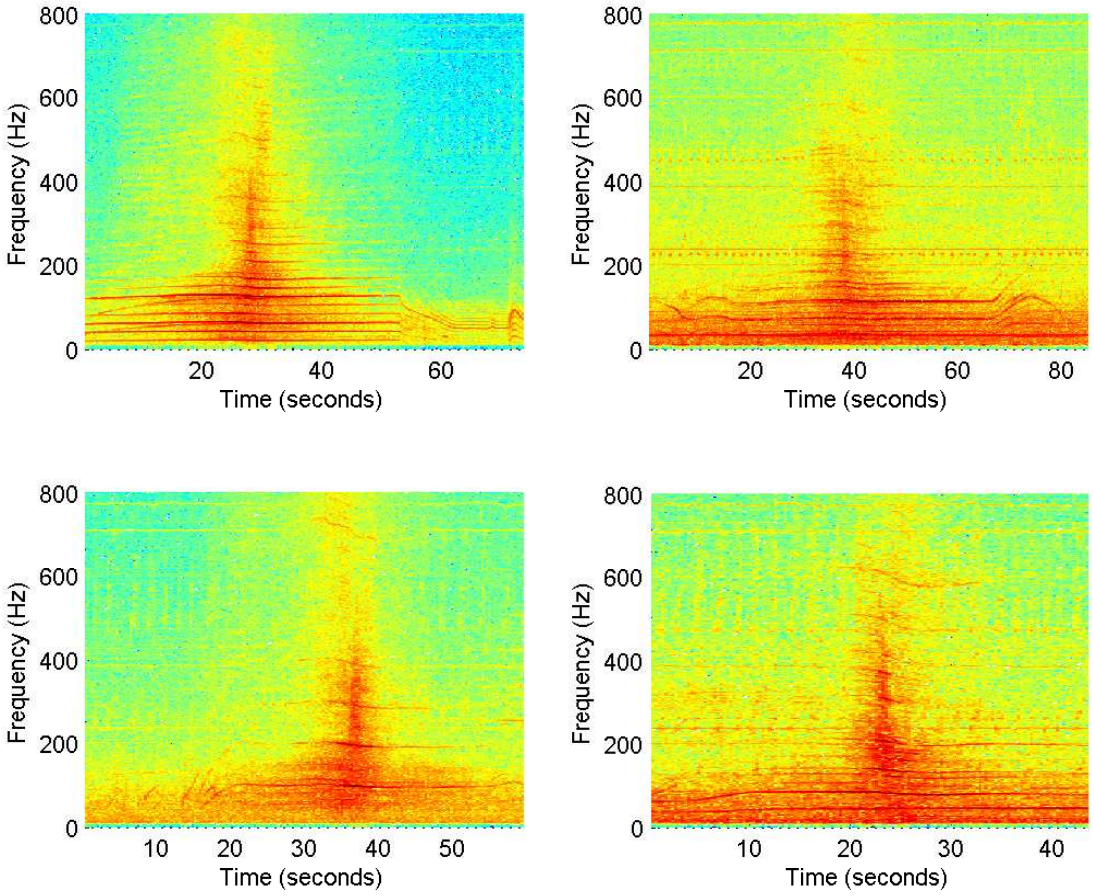

Figure 4.11: Spectrogram of the acoustic signal for four different vehicles (heavyweight tracked (top left), heavyweight wheeled (top right), lightweight tracked (bottom left), and lightweight wheeled (bottom right))

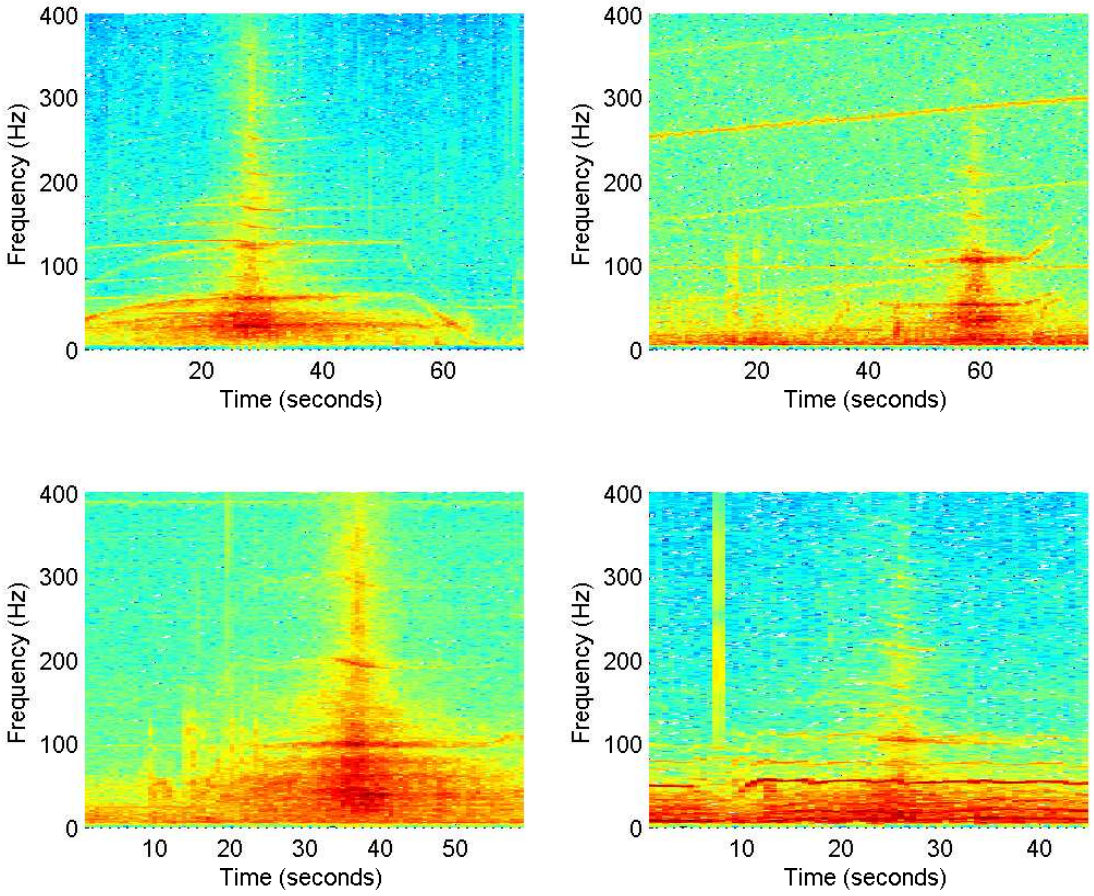

Figure 4.12: Spectrogram of the seismic signal for four different vehicles (heavyweight tracked (top left), heavyweight wheeled (top right), lightweight tracked (bottom left), and lightweight wheeled (bottom right))

in extracting the characteristic features from the acoustic emission. But experiments have been also performed on signatures from both modalities to evaluate the performance with respect to each modality. The seismic emission's utility for feature extraction has been demonstrated using multimodal detection and multimodal fusion (4.6.6).

Extracting characteristic features from the time domain signal is the core of the TDHA method, which eliminates the need to perform an FFT for the generation of characteristic FVs. Time and frequency domain features are evaluated according to the number of multiplications required for feature extraction, the dimensionality of the FV, and the amount of information carried by the characteristic feature that is crucial for discrimination.

The deterministic acoustic signature can be represented using a simple harmonic model. The coupled harmonic signal model can be described as [3]

$$x(t) = \sum_{k=1}^{M} \alpha_k \cos(2\pi f k t + \varphi_k), \tag{4.1}$$

where $\alpha_k \geq 0$ and φ_k are the amplitude and phase of the k^{th} harmonic respectively, f is the FF, and M is the total number of harmonics used in the model. It is clear that the number of parameters is $2M + 1$. Assuming that $\mathbf{y}$ is the vector of acquired acoustic sensor data such that

$$\mathbf{y} = \mathbf{x} + \boldsymbol{\epsilon}, \tag{4.2}$$

where $\boldsymbol{\epsilon}$ is assumed to be the noise vector with a multivariate Gaussian distribution such that $\boldsymbol{\epsilon} \sim \mathcal{N}(0, \Sigma)$. The noise vector includes both the non-deterministic acoustic signal and the background noise. The acquired signal may also be written in matrix form as:

$$\mathbf{y} = C\mathbf{u} + S\mathbf{v} + \boldsymbol{\epsilon}, \tag{4.3}$$

where the elements of the $\mathbf{u}$ and $\mathbf{v}$ vectors and the C and S matrices are given by

$$\begin{aligned} u_k &= [\alpha_k cos\varphi_k]^T \quad \text{and} \quad v_k = [-\alpha_k sin\varphi_k]^T \\ c_{nk} &= \cos(2\pi fkn) \quad \text{and} \quad s_{nk} = \sin(2\pi fkn), \end{aligned} \tag{4.4}$$

where $k = 1, ..., M$, $n = 0, ..., N-1$, and N is the number of samples. The set of parameters of this model will be denoted by $\theta = (A, \varphi, f)$. Among all harmonic parameters, the information in the harmonics' amplitudes is more capable in verifying the correct source among other candidates [2, 53, 62, 63]. The goal of the proposed feature extraction algorithm is to estimate the harmonics' amplitudes and use them in classification such that the features perform at least as well as the spectral features and can be extracted using less computational power.

4.5.1 Parameter estimation

At this point, we will assume that the FF and the number of harmonics are known. The MLE of the coupled harmonic parameters θ is found by minimizing the negative log-likelihood function given by

$$J(\theta) = \frac{N}{2}\ln(2\pi) + \frac{1}{2}\ln(|\Sigma|) + \frac{1}{2}(\epsilon(\theta))^T(\Sigma)^{-1}(\epsilon(\theta)), \tag{4.5}$$

The first term is a constant that does not affect the minimization procedures and can be omitted. Using the assumption of white noise, then we can substitute for Σ by $\sigma^2\mathcal{I}$, where $\mathcal{I}$ is the identity matrix. Under this assumption the parameter vector θ that minimizes $J(\theta)$ also minimizes $||\epsilon||^2$, which is the squared norm of the difference between the measurements and the signal model ($||y - x||^2$). The signal model in 4.2 can also be represented using a linear model as:

$$\mathbf{y} = \gamma\beta + \boldsymbol{\epsilon}, \tag{4.6}$$

where

$$\gamma = [C \;\; S] \text{ and } \beta = [\mathbf{u} \;\; \mathbf{v}]^T, \tag{4.7}$$

Under the white noise assumption, the ML estimation is given by the LS solution. In our case the LS solution of β is

$$\hat{\beta} = (\gamma^T\gamma)^{-1}\gamma^T y, \tag{4.8}$$

Even with the assumption that the FF and the number of harmonics are known, obtaining the MLE of the coupled harmonics' parameters is computationally expensive because of the matrix inversion. Precisely defining the phase in the middle of the observation window [3,68] simplifies the estimation process such that the elements of the C and S matrices become

$$\begin{aligned} c_{nk} &= \cos(2\pi fkn + 2\pi fk(N-1)/2) \quad \text{and,} \\ s_{nk} &= \sin(2\pi fkn + 2\pi fk(N-1)/2). \end{aligned} \tag{4.9}$$

In addition, defining the phase in the middle of the observation window guarantees that $C^T S = 0$ and $S^T C = 0$ (diagonalizing the information matrix) such that

$$\gamma^T \gamma = \begin{bmatrix} C^T C & S^T C \\ C^T S & S^T S \end{bmatrix} = \begin{bmatrix} C^T C & 0 \\ 0 & S^T S \end{bmatrix} \tag{4.10}$$

The amplitude vector estimate in equation 4.8 is given by

$$\hat{\boldsymbol{u}}_{\boldsymbol{LS}} = (\Lambda_C)^{-1} C^T \mathbf{y} \quad \text{and} \quad \hat{\boldsymbol{v}}_{\boldsymbol{LS}} = (\Lambda_S)^{-1} S^T \mathbf{y}, \tag{4.11}$$

where

$$\begin{aligned} \Lambda_C &= \text{diag}(C^T C) \quad \text{and,} \\ \Lambda_S &= \text{diag}(S^T S) \end{aligned} \tag{4.12}$$

Given a known FF and the appropriate number of harmonics, the harmonic structure guarantees that C^TC and S^TS are diagonal matrices which substitutes matrix inversion with just division reducing the overall computations needed to estimate the harmonics' amplitudes. Table 1 in the Appendix shows a comparison between the harmonics' amplitudes estimated using matrix inversion and those obtained using the reciprocal of the Λ_C and Λ_S diagonal matrices using the Bochum data set. Using the amplitude vector estimates, the coupled harmonic signal amplitude is

$$\hat{\alpha}_k = \sqrt{\hat{\boldsymbol{u}}^2 + \hat{\boldsymbol{v}}^2}, \tag{4.13}$$

and the coupled harmonic signal phase is

$$\hat{\varphi}_k = \arctan(\frac{\hat{\boldsymbol{v}}}{\hat{\boldsymbol{u}}}), \tag{4.14}$$

where $k = 1, ..., M$.

4.5.2 Strongest Harmonic Period

Without knowing the FF and the number of harmonics, the complexity in estimating the harmonics' amplitudes is large. Even when the complexity was discarded, several verification schemes in the literature [42,52] failed in defining the harmonic line series from the spectrum of the acoustic signal searching among all distinct peaks, while other schemes [3,68] found difficulty in precisely estimating the FF value, even when the number of harmonics was fixed.

Instead of searching for the actual FF, the TDHA approach estimates the strongest harmonic component. There are two reasons why we are interested in estimating the strongest harmonic component: first, it requires simple and few computations to be estimated; second, it will be combined with the required spectral resolution (ΔF) in determining the fundamental component of the approximated harmonic model for estimating the amplitude parameters. We will call this fundamental component the assumed FF ($\widetilde{FF}$). The time domain signal is processed in the form of short events (windows). Selecting events of short durations (0.5 seconds) guarantees that variation with time in the harmonic structure is minimum. The amplitude of the strongest harmonic component was previously used for normalization [57], in order to mitigate the variation in signature with range (i.e., the separation between sensor node and vehicle). In [42, 52], the strongest harmonic component was used in determining the vehicle's number of cylinders, where it was shown that for most events the strongest component appears at an integer multiple of the fundamental equivalent to the number of cylinders of the vehicle. Estimating the strongest harmonic period (τ) is accomplished through computing the correlation between the first half of the current event time signal and its shifted version. Computing the correlation of the acoustic signal represents a simple and efficient method for defining repeated patterns. The correlation maximum occurs at zero shift and the distinct correlation peaks are identified to occur at shifts equivalent to the strongest harmonic component period and its multiples[2]. If correlation peaks are located at displacements equal to d_1,d_2,d_3,...etc., the strongest harmonic period becomes

$$\tau = \frac{d_1}{F_s}, \tag{4.15}$$

[2]We have seen that this is consistently true for the military vehicle application

where F_s is the sampling frequency. The estimated period can be refined by considering all peaks or selecting those with amplitude above the noise level as follows,

$$\tau = \text{mean } (\frac{d_1}{F_s}, \frac{d_2}{2F_s}, \frac{d_3}{3F_s}, \ldots) \tag{4.16}$$

Examples of the correlation function are shown in Figure 4.13 for different signal to noise ratios. The original signal has a harmonic structure with the strongest harmonic period equal to 8.9 msec. The estimated harmonic period using either the first correlation peak (d_1) or the second correlation peak (d_2) is found to be 8.5 msec and 9 msec, respectively. The accuracy in estimating τ depends on two important parameters, the expected value of the strongest harmonic period and the sampling frequency. As the strongest harmonic period decreases or the sampling frequency decreases, the resolution in estimating the strongest harmonic period also decreases and vice versa. This limitation comes from the fact that we compute the correlation at positive integer shifts only[3].

Using the proposed simple estimation approach, the estimated strongest harmonic period might not be accurate for a number of events. As long as our approximated harmonic structure includes the band where the energy is highly concentrated, however, the exact value of the estimated period is not going to affect our generated features as will be shown in the results in sections 4.6, 5.4, and 5.5.

The simple harmonic model discussed in section 4.5 is defined by two components, the FF and the number of harmonics. The TDHA method uses the estimated

[3]Simulation results in [62] illustrated that refining the estimated strongest component period to non-integer values did not show much improvement in the classification rate.

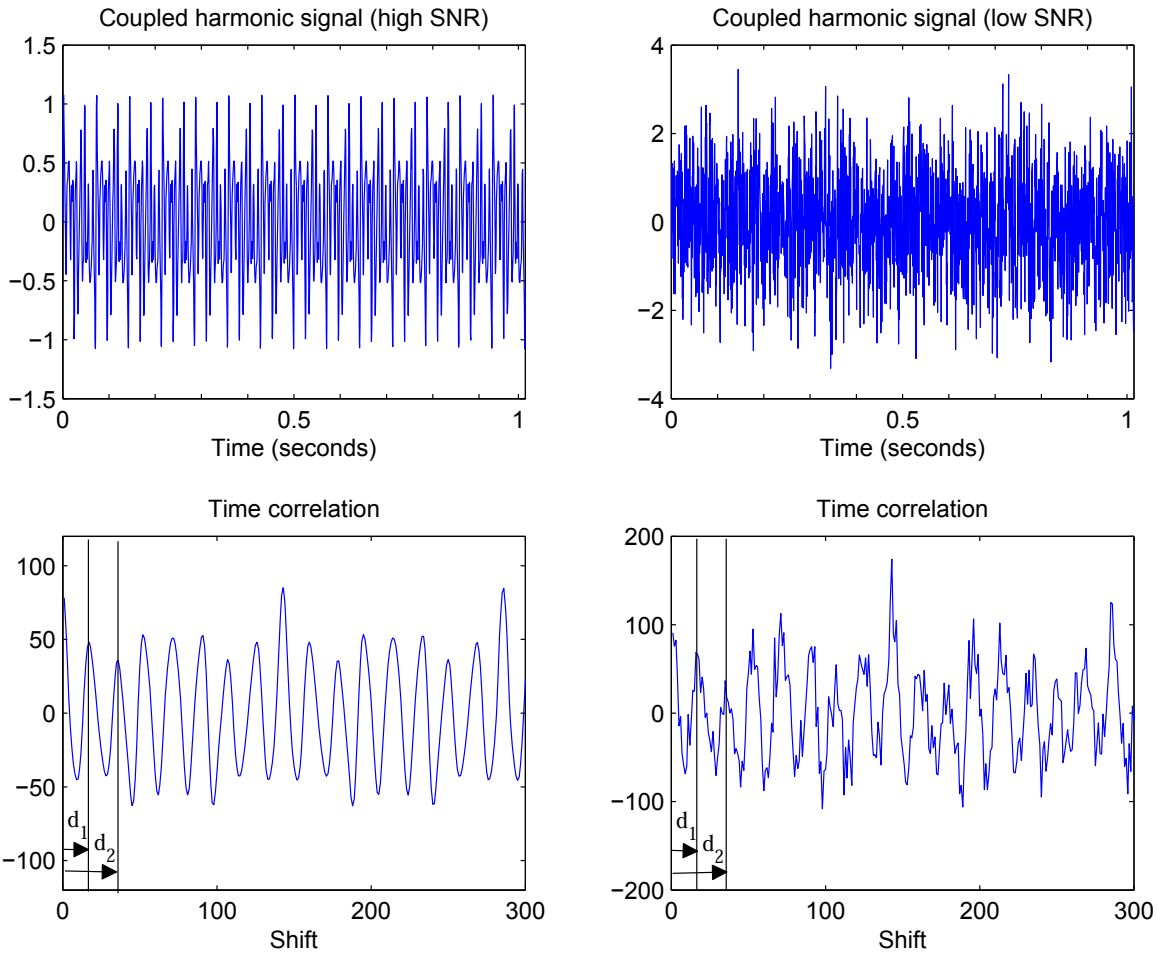

Figure 4.13: Sample coupled harmonic signals and their computed time correlation at high SNR (left) and low SNR (right)

strongest harmonic frequency ($\xi = \frac{1}{\tau}$) and ΔF (our desired frequency resolution) in computing $\widetilde{FF}$ such that

$$\widetilde{FF} = \frac{\xi}{\lceil \frac{\xi}{\Delta F} \rceil}, \tag{4.17}$$

where $\lceil \frac{\xi}{\Delta F} \rceil = \min\{D \in \mathbb{Z} \mid D \geq \frac{\xi}{\Delta F}\}$. The number of harmonics is fixed and is selected such that the generated coupled harmonic model approximates the acoustic signature up to the Bandwidth (BW) of interest (for vehicles, the BW is 250 Hz). The number of harmonics is equal to the length of the generated FV, and depending on the available resources and the classifier, the FV length is determined and combined with the BW of the highest energy concentration in specifying the spectral resolution such that

$$\Delta F = \frac{\text{BW}}{\text{FV length}} = \frac{\text{BW}}{\text{number of harmonics}}, \tag{4.18}$$

Note that averaging over nine different vehicles, 96.64% of the energy of the acoustic emission for more than 32000 detected events[4] (Bochum data set) was found to be in the range $0 - 250$ Hz. Using (4.17), the $\widetilde{FF}$ is always an integer divisor of the estimated strongest harmonic frequency with a defined spectral resolution, which guarantees that the generated harmonic model will cover this frequency in addition to others with higher energy concentration.

Figure 4.14 shows the time-frequency response for all detected events for a tank passing through a sensor node with both the actual and estimated strongest harmonic frequency. The acoustic signal was processed in 0.5 seconds windows, the total number of detected events for this specific run was 64 (32 seconds). It is clear from

[4]Each event represents 0.5 seconds

Figure 4.14 that another harmonic component dominates, most probably that coming from the tracks, when the vehicle is closer to the sensor node. Although the simple estimation of the strongest component slipped for few events, the use of (4.17) in estimating a fundamental helps in correctly estimating a quite good approximation of the actual signal spectrum. This is seen from the comparison in Figure 4.15, between the actual signal spectrum computed with 0.17 Hz resolution and the estimated harmonics' amplitudes using the simplified LS method at four different events (t_1, t_2, t_3, t_4). Table 4.2 shows the actual and estimated frequencies for the four events. τ was estimated using the first three peaks (d_1, d_2, d_3) at the correlation function and the harmonics' amplitudes were estimated for a harmonic model with 50 harmonics and an $\widetilde{FF}$ near 5 Hz.

Table 4.2: Actual and estimated frequencies for the 4 events indicated in Figure 4.14

Event	Actual strongest harmonic (Hz)	Estimated strongest harmonic (Hz)	Assumed fundamental (Hz)
t_1	126.49	126.41	4.86
t_2	128.47	128.42	4.93
t_3	129.08	129.69	4.98
t_4	64.23	96.15	4.80

The estimated harmonics' amplitudes are likely to have the same envelope as the original signal spectrum as shown in Figure 4.15. Although there was an error in estimating the strongest harmonic component at t_4, the estimated harmonics' amplitudes using the simplified LS method confirm that the strongest harmonic is at 62.95 Hz and not at 96.15 Hz. Although the change in the assumed fundamental is very small, this difference becomes more significant for higher harmonic components. It is also shown in Figure 4.15 that the estimated amplitudes of all harmonics that are not distinct are generally close to zero.

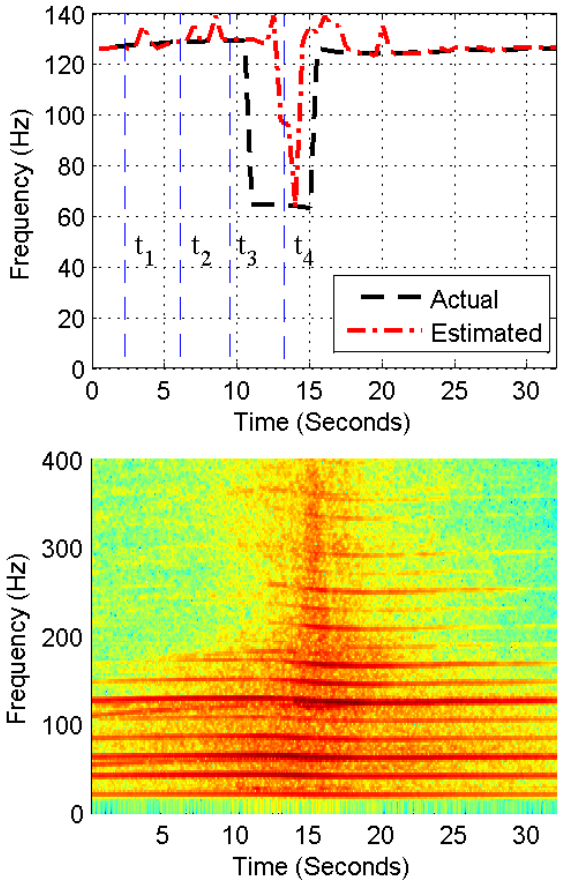

Figure 4.14: Time-frequency response and ξ estimation for the M48 tank

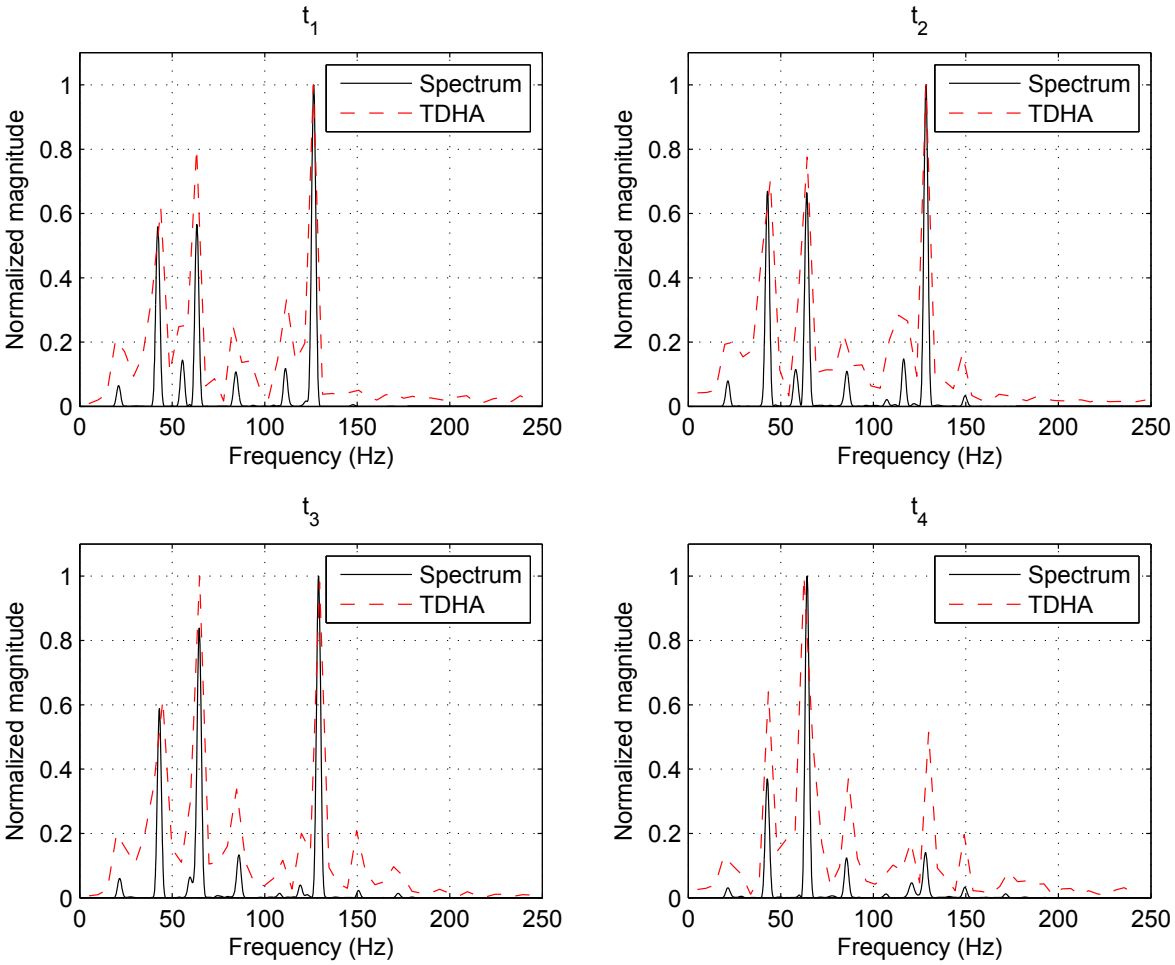

Figure 4.15: Estimated harmonics' amplitudes compared to the actual spectrum for the M48 tank at 4 different events

4.5.3 Computational savings

The $\widetilde{FF}$ represents the smallest harmonic component of the coupled harmonic model used for approximating the deterministic acoustic signature. A reduced number of samples equivalent to the period of the $\widetilde{FF}$ will be selected as a *template* of the original window. The selection is supported by the validity of the coupled harmonic model, and the fact that harmonic components change slowly [69]. By extracting the characteristic features from the template, we select the signature of this vehicle in the frequency range starting at $\widetilde{FF}$ to $F_s/2$. A simple illustration of the amount of computational savings due to data reduction using a shorter template is shown in Table 4.3 for different sampling frequencies and window lengths.

Table 4.3: Illustration of the amount of savings provided by using a *template* that corresponds to an $\widetilde{FF} \sim 5$ Hz

F_s (Hz)	Window length (seconds)	*Template* $= \frac{F_s}{\widetilde{FF}}$ (samples)	Percentage of the *template* size with respect to the original window size
1024	0.5	205	40%
	1.0		20%
2048	0.5	410	40%
	1.0		20%

The actual benefit lies in the computational power needed to extract the harmonics' amplitudes as compared to the FFT components. The number of real multiplications required for estimating the harmonics' amplitudes in the TDHA method is $O(M * NT)$, where M is the number of harmonics and NT is the template width. The number of complex multiplications required to obtain N FFT coefficients is $O(N \log_2 N)$ [104] . Each complex multiplication requires four real multiplications

and two real additions. Extracted features using a higher resolution leads to better classification over both methods, but the TDHA method requires a smaller number of multiplications as compared to the FFT method. The number of real multiplications required to estimate the harmonics' amplitudes compared to the number of real multiplications needed to obtain their equivalent FFT components for different spectral resolutions are shown in Figure 4.16. We implemented the proposed feature extraction algorithm in fixed point operation using 16 bit data values with 32 bit computational precision. A fixed point realization of the FFT algorithm, using twiddle factors from [104], for extraction of spectral features is implemented as well for comparison. Both algorithms use lookup tables: twiddle factor tables for the spectral features and sine and cosine tables for the time domain features. On average, the time consumed for extraction of the time domain features is less than 23% of the time required to obtain the spectral features for the same tested signal on the same platform.

4.5.4 Feature vector selection

The number of harmonics M is fixed and depends on the spectral resolution and the projected spectral band. The generated feature vector must include any useful information that preserves the vehicle's signature. After estimating the M harmonics' amplitudes using the simplified LS method as described in section 4.5, the characteristic FV that will be used in discrimination is constructed from the estimated strongest harmonic frequency, the event energy, and the M estimated harmonics' amplitudes $(\hat{\alpha}_1, \hat{\alpha}_2, \cdots, \hat{\alpha}_M)$. The length of the FV is $M + 2$. Although $\widetilde{FF}$ is proportional

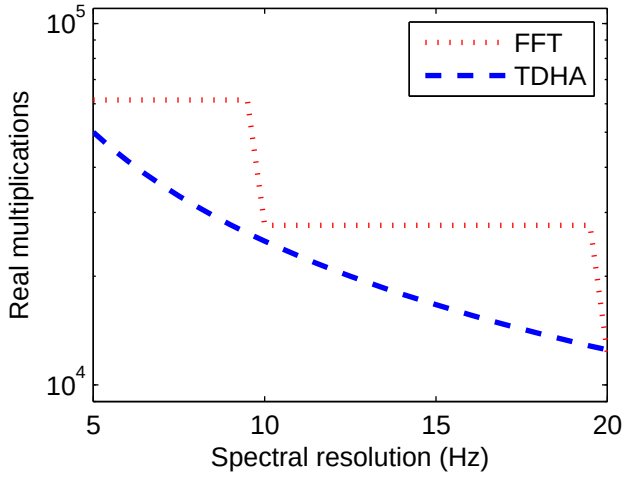

Figure 4.16: Real multiplications needed to estimate both the harmonics amplitudes and the FFT components at different ΔF

to the Estimated Strongest harmonic Component (ESHC), having the ESHC as an additional dimension to the FV is very useful since several ESHCs can result in the same $\widetilde{FF}$. The information in the signal's energy is very useful. Beside using the energy in detecting the vehicle, energy can be helpful in distinguishing tracked and wheeled vehicles since non-deterministic components in the acoustic/seismic emission of tracked vehicles are large when compared to wheeled vehicles and this information is stored in the signal's energy.

The generated FV at this level can be used with any classifier. An additional benefit for utilizing ANNs for classification is their ability to acquire FVs formed from different measurements that have different dynamic ranges/resolution without any adjustments. ANNs adjust the weights and biases such that certain weights and biases will have higher values than others depending on the impact of this specific entry and its relative value compared to the other entries. After training the neural network, the weights and biases are stored and the classification process becomes straight-forward

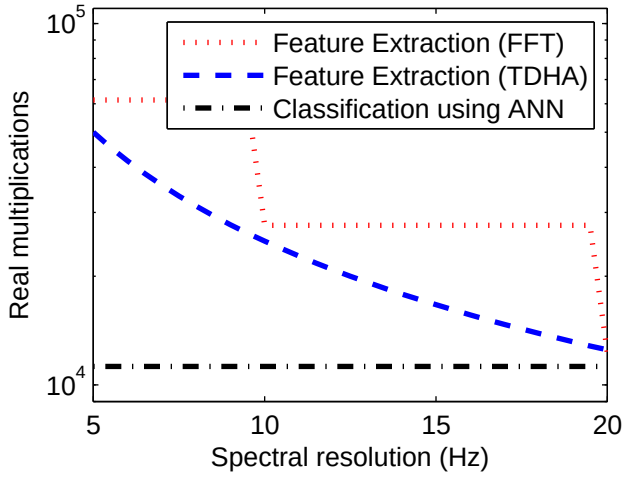

Figure 4.17: Real multiplications needed for classification versus feature extraction (TDHA and Spectral) at different ΔF

by propagating the input FV through the network weighted by the stored weights and biases. Figure 4.17 compares between the number of real multiplications required to estimate the characteristic FV and those needed for classification using ANN at different spectral resolutions. The number of real multiplications needed for classification depends on the FV length and the ANN connectivity. Simulations results in section 4.6 will prove that better classification is achieved at a smaller spectral resolution (5 Hz) than at a higher resolution (20 Hz) for the same FV length. The number of real multiplications in Figure 4.17 is for a full connected ANN[5]. Pruning FV entries (i.e., higher harmonics' amplitudes) or certain neurons at the hidden layer that have minimum contribution to the overall performance will result in having a significant reduction in the total number of real multiplications for classification and might have negligible effect on the overall system performance.

[5]All neurons at each layer are connected to the neurons in the preceding layer.

4.6 Experimental Data Processing Results

In this section, we will demonstrate the simulation results and compare them with published results in the literature for verification of ground military vehicles. All simulations were performed on the data sets previously described in section 4.3. Unless stated otherwise, all generated features were obtained for events of duration 0.5 seconds and all detected events were split randomly into two sets, one set (1/3 of the features) is for adapting/training of the classifier and the other set (remaining 2/3 of the features) is for testing the proposed scheme and the results are for Monte-Carlo simulations of 100 iterations. Simulations were performed using Matlab [115] on a PC (Intel dual core 2.1 GHz, 4 GB RAM). Detailed specifications of the CFAR detector and the ANN classifier can be found in the appendix in Table 2 and Table 3, respectively.

We attempted to represent all our simulation results in a logical and consistent order that will demonstrate our proposed approach and justify the selections we made for each of the detection algorithm, harmonic structure, modality, classifier, and fusion algorithms. Although the Bochum data set seems to be richer than the SensIT data set, some results will include both data sets to better emphasize the proposed approach. Section 4.6.1 illustrates the importance of a simple, yet reliable detection algorithm for event driven applications with limited resources. Section 4.6.2 compares the performance of features extracted from different harmonic approximations of the deterministic signature of military vehicles. Section 4.6.3 evaluates the performance of features extracted from different modalities. Section 4.6.4 compares between features extracted using the proposed TDHA algorithm with spectral features. Section 4.6.5

shows how ANNs outperform other classifiers for the same feature set. Section 4.6.6 points out to the benefits obtained for applying simple fusion algorithms on top of the proposed algorithm. Section 4.6.7 compares our simulation results with previously published results for verification of military vehicles using the same data sets.

4.6.1 Detection

Detection is the most frequent processing task performed by the sensor module, and although it is usually performed using a simple energy based scheme, the performance of the selected detection scheme has a great impact on the overall performance of the verification system. An event is the smallest interval through which sensor nodes are capable of producing decisions. During target detection, sensor nodes decide whether the current event is equivalent to an intruder or just background noise. The event length is usually pre-defined based on the proposed algorithm and the required classification rate. Energy detection uses minimal a priori information about the target, hence the average acoustic or seismic energy is computed and compared with an adaptive threshold that is updated according to the background noise level maintaining a CFAR [40]. Due to the inherent signal averaging, the noise component in the output of the detector may be modeled as a Gaussian random variable whose mean and variance can be determined from the statistics of the background noise. CFAR detection computes the energy of the current event and compares it with the adaptive threshold and announces the presence of the intruder if the average energy is higher than the threshold. The threshold is dynamically adjusted only when no

target is detected according to the noise variance so that the detector maintains a CFAR.

CFAR is very sensitive to the initial assignment of the adaptive threshold. If the initial threshold is higher than the background noise, then fewer targets will be detected, and the number of false alarms will be low. On the other hand, having a threshold lower than the background noise will result in detecting more targets, and raising the number of false alarms. Features extracted from non-target events will result in error in classification and will decrease the ability of discrimination if such events are involved in the classifier learning process. Since the features extracted from the acoustic or seismic signal entail the harmonic structure of the moving vehicle and this structure needed to be almost stationary during each event, the event length is usually selected to be ≤ 1 second for verification of military vehicles.

Vehicle verification for peacekeeping applications is an event driven application, where the rate of event detection is unpredictable. Although sensor nodes' duty cycle is very low, the effects of background noise on acoustic sensors and the consequences of any non-anticipated movement in the sensor field on seismic sensors make individual sensors test for targets more frequent. In the absence of strong wind, background noise, and other seismic activities, utilizing a running average of either the acoustic or seismic energy will yield to a similar set of detected events depending on the location of the sensor with respect to the road. Figure 4.18 and Figure 4.19 illustrate the CFAR detection output utilizing either acoustic or seismic energy for a heavyweight tracked vehicle and a heavyweight wheeled vehicle using SensIT data set. Detected events are those with energy higher than the adapted threshold, resulting in a number of events

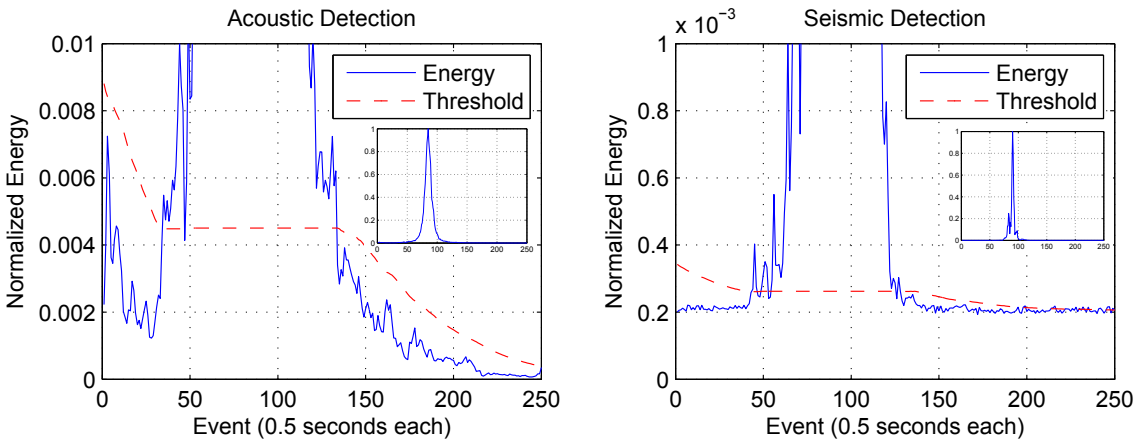

Figure 4.18: Acoustic and Seismic detection using single modality at low background noise (Heavyweight tracked - SensIT)

around the CPA. The x-axis is the event index with duration of 0.5 seconds per event where the duration of the whole run was just over 2 minutes. The y-axis represents the normalized energy. The full scale plot is minimized to illustrate the average energy detected at the sensor node and how it increases with the vehicle approaching the node, while the lowest energy band is magnified to reveal the adaptation of the threshold with the background noise. At low wind and background noise, either acoustic or seismic energy is sufficient to extract the required set of events for feature extraction and classification. For the heavyweight tracked vehicle, the acoustic and seismic energy produced by the tracks can be detected at a longer range with respect to the sensor node compared to the wheeled vehicle.

One of the first problems that faces the single modality detection scheme is the location of the vehicle with respect to the sensor node and the effect of gear shifting and the noises coming from the exhaust in the produced energy. In Figure 4.20, the

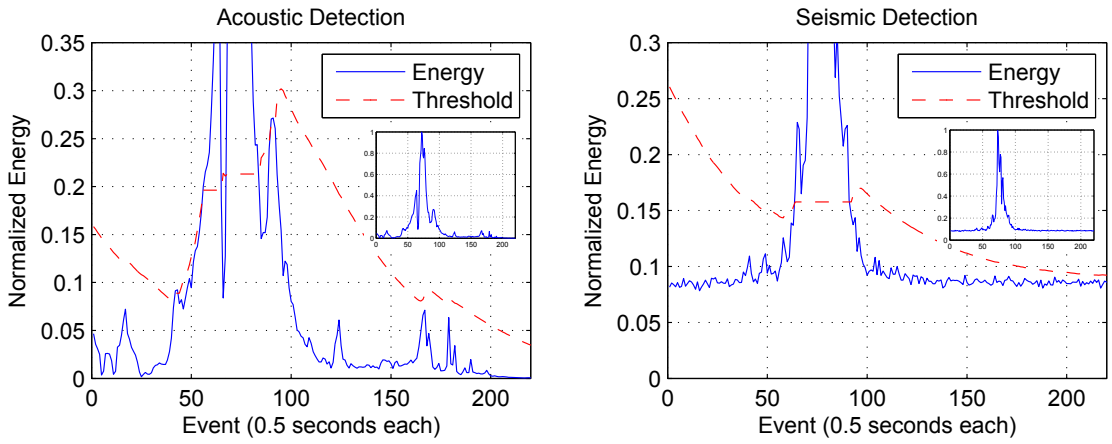

Figure 4.19: Acoustic and Seismic detection using single modality at low background noise (Heavyweight wheeled - SensIT)

heavyweight tracked vehicle is passing by the node and after the energy level had dropped for more than 15 seconds, the energy level shows some peaks although the vehicle is far from the node and there are no changes in the seismic energy at this time. The reason for this acoustic change is the acoustic energy coming from the exhaust, where the vehicle produces louder noises when the engine RPM is high. The features extracted from these events are very distracting to the classifier because the acoustic signature of the engine is not clear.

In the presence of high wind noise and when the initial threshold was obtained based on a lower background noise level, the acoustic detector will produce a large number of detected events compared to the actual number of events where the vehicle is responsible for the acoustic energy. Figures 4.20, 4.21, and 4.22 show three different situations where CFAR detection using acoustic energy fails to report the correct set of events compared to the seismic energy.

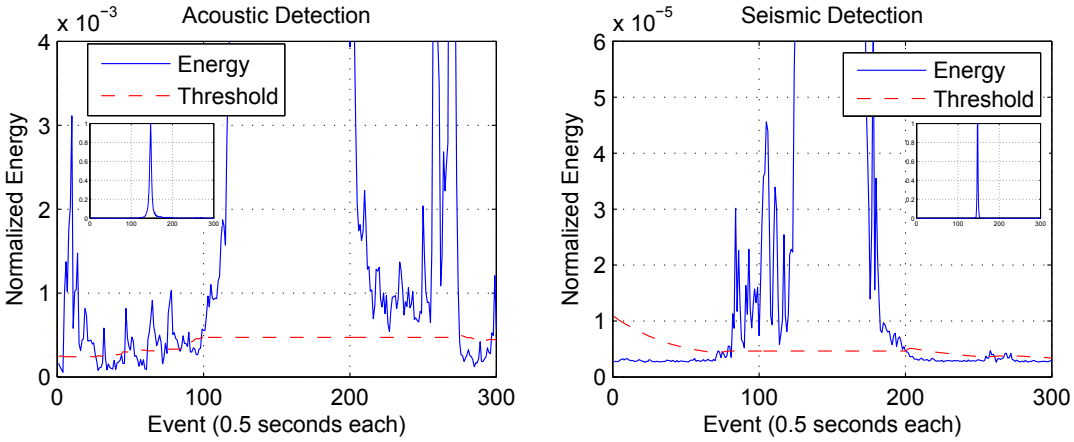

Figure 4.20: Acoustic and Seismic detection using single modality at high background/wind noise (Heavyweight tracked - SensIT)

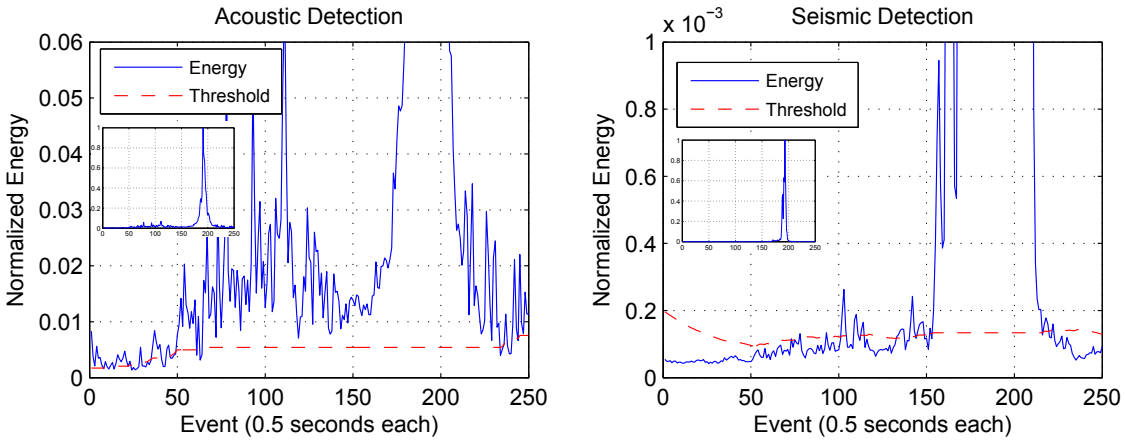

Figure 4.21: Acoustic and Seismic detection using single modality at high background/wind noise (Heavyweight tracked - SensIT)

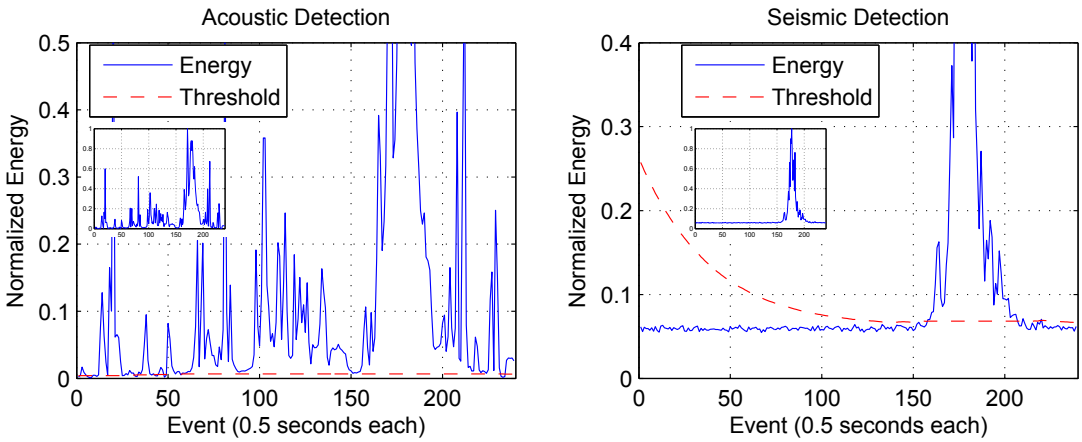

Figure 4.22: Acoustic and Seismic detection using single modality at high background/wind noise (Heavyweight wheeled - SensIT)

We selected SensIT data set to investigate the performance of the detector because we performed manual labeling of the acoustic signal for more than 3 hours of its complete runs. We listened to the run and labeled individual events using a noise canceling headphones. Through manual labeling, we differentiated between two sets of events, those that have distinct acoustic energy when the listener can tell there is a vehicle crossing and those that have just background noise or silence. Table 4.4 compares between the true number of detected events using acoustic energy and that using seismic energy compared to using manual detection. Although the total number of detected events using acoustic energy is close to the total number of manually labeled events, $\frac{1}{3}$ of these events represent background noise. The total number of detected events using seismic energy is almost half of those detected using acoustic energy because geophones can detect activities within a limited range compared to microphones and because of the high attenuation in the seismic signal with respect to distance and its dependence on the substrate. But less than 7% of these events

are not true events compared to over 32% using acoustic energy. Although seismic detection misses more than half of the events, geophones were capable of detecting at least a few valid events in every single run which implies that sensor nodes using seismic signals will not miss any intruding vehicle within their coverage range.

Table 4.4: Detected events using SensIT data set

Detection modality	Total number of detected events	Correctly detected events
Acoustic	19251	13239 (68.77%)
Seismic	8691	8123 (93.46%)
Manual	21290	21290 (100%)

When detection of military vehicles was investigated with respect to the second data set (Bochum), both seismic and acoustic sensors have shown similar performance with respect to the detected events. Vehicle's movements were along a straight path, where the site included 4 parallel roads that run through a forest which minimizes wind noise and makes acoustic detection sufficient for providing a reliable set of detected events for representing the moving vehicle. The total number of detected events using seismic energy is larger than the total number of detected events using acoustic energy which makes the selection of the best modality for detection very dependent on the site, selected hardware, and background noise.

Selecting other modalities for detection of military vehicles such as infrared sensors or camera sensors will not solve the problem since these sensors are highly affected by the weather and might be blocked by fog or heavy snow fall. Seismic sensors have a great advantage compared to the acoustic sensors, which is their vulnerability to spoofing [116]. Geophones are completely buried and less affected by weather conditions. While microphones could be spoofed easily by loudspeakers at close distances,

spoofing of seismic activities that covers large areas would be more difficult to achieve and would require much more energy

Our previous discussion with respect to single modality detection suggests that the need for an extended detection range requires deployment of acoustic sensors. Utilizing acoustic energy for detection implies that the system might suffer from a high false alarm rate depending on the background noise. On the other hand, solely relying on seismic sensors for detection will guarantee minimum detection error but will limit the range compared to acoustic detection. The problem of detection becomes more complicated with having lightweight military vehicles among different candidates because lightweight wheeled vehicles produce lower acoustic and seismic energy.

Peacekeeping operations are more interested in reliable reporting of trespassing military vehicles in disarmament territories. Seismic energy will provide sufficient information to the sensor node to detect approaching targets. But when the sensor network is deployed for providing an early warning system for peacekeeping camps and important assets, acoustic detection will be of great importance for distant detection. The network might benefit from the distributed processing of individual sensors and the availability of multiple modalities for vehicle detection but with limited resources and the possibility that what might affect one sensor most likely will affect the neighboring ones, multimodal detection using multiple transducers will yield the most reliable option. Multimodal detection requires having multiple transducers, measuring different modalities, delivering data to the same node. The network might rely on a single modality in most situations with the possibility of having both transducers active for detection when needed.

4.6.2 Harmonic Structure

The goal of our model is to assemble as much information regarding the incidence with minimum power consumption such that the extracted features are sufficient for discrimination. Since the extracted features are our primary concern and they represent the main contribution of our work, we will start with investigating the performance of the monitoring system with having different approximations for the harmonic structure of the acoustic signal. All models in this section apply the same detection and classification scheme, which are a CFAR acoustic detector and a feedforward multilayer perceptron neural network for classification, respectively. The only difference is in the set of extracted features.

Classification results are given in the form of *confusion matrices*, which classifies the events by their true classification, and the experimental classification results. The rows of the *confusion matrix* represent the true classification (manual labeling) and the columns represent the experimental classification. The detection, false alarm, and classification rate are as previously described in Chapter 3.

Three different simple harmonic approximations were investigated to assess the performance of the proposed TDHA algorithm. We will discuss the complexity and limitations of these harmonic approximations. The only information available to us is that the FF of military ground vehicles lies in the range $5 - 20$ Hz [3, 36], the FF is not the strongest harmonic component, and the energy of the acoustic signal is concentrated in the frequency band $5 - 250$ Hz. Given the resource limitations consistent with our attempt to develop a simple feature extraction algorithm, a harmonic

approximation that seeks to estimate the true FF and number of harmonics will not be considered in our simulations.

1. **Super and sub harmonics model**

   This model uses the ESHC divided by an integer and uses this sub-harmonic component as the base frequency for the harmonic model. Since we don't know the order of the ESHC with respect to the original FF, the selected harmonic structure might approximate the signature in a higher or lower band from our target 250 Hz band. The limitation of this model is in selecting a fixed integer independent on the ESHC and the fact that an error in estimating the ESHC will result in approximating a band different than the projected band.

2. **Proposed TDHA model**

   This model uses the ESHC divided by an integer that will force the approximation to be based upon the ESHC and spans our target band with a desired spectral resolution (ΔF) as previously discussed in section 4.5.

3. **Fixed harmonic model**

   This approximation uses a fixed harmonic structure for all signals. Estimating the harmonics' amplitudes using the simplified LS method for a harmonic model with a constant FF and number of harmonics. This model is the simplest and does not require the extra computations needed to estimate the strongest harmonic period and determine an assumed FF.

The characteristic features are the ESHC, the event energy and the estimated harmonics' amplitudes using the simplified LS method. The FV combines all three feature types in one vector. The number of harmonics is fixed for each model/approximation.

The number of harmonics is fixed at 50 except for the first approximation, we select only 20 harmonics, since we did not expect any harmonic components above 1500 Hz. Table 4.5 compares between the detection, false alarm and classification rate using the Bochum data set for the following approximations:

- Model 1: Base freq. $= \frac{ESHC}{1}$, 20 harmonics $\Rightarrow$ FV length = 22.

  Base freq. $= \frac{ESHC}{10}$, 50 harmonics $\Rightarrow$ FV length = 52.

  Base freq. $= \frac{ESHC}{20}$, 50 of harmonics $\Rightarrow$ FV length = 52.

- Model 2: Base freq. $= \widetilde{FF} \sim 5$ Hz, 50 harmonics $\Rightarrow$ FV length = 52.

- Model 3: Base freq. = 5 Hz, 50 harmonics $\Rightarrow$ FV length = 52.

The features extracted from all approximations have an average classification rate above 80% which is considered excellent for military vehicle classification [80] considering that we are differentiating among 9 different vehicles using time domain features extracted from the harmonic approximation of the deterministic signal. The proposed TDHA algorithm was able to produce better approximation, hence generating more representative features. The difference among the four approximations is distinct in all vehicles. The importance in the lower frequency band ($\leq$ ESHC) and the way of forcing the model to include our target band and being proportional to the strongest harmonic component (the only component we are quite definite that is present in the original signature) puts the proposed algorithm on top of our preferred approximations.

Table 4.6 shows a sample confusion matrix for the proposed TDHA algorithm with 50 harmonics approximating the harmonic acoustic signal in the band $5 - 250$ Hz.

Table 4.5: Different approximations for the harmonic signal - Bochum data set

Base Frequency →		$\frac{ESHC}{1}$	$\frac{ESHC}{10}$	$\frac{ESHC}{20}$	$\widetilde{FF} \sim$ **5Hz**	5Hz
Feature vector length→		22	52	52	**52**	52
Vehicle ↓						
Detection rate (%)	Leopard1	94.62	97.08	97.09	**97.49**	93.38
	Leopard2	85.18	88.82	93.18	**94.15**	89.52
	Jaguar	84.50	85.95	90.03	**92.26**	85.87
	M48	96.99	95.96	96.88	**97.94**	93.71
	Fuchs	73.32	74.43	77.57	**86.49**	83.58
	Hermelin	55.89	51.14	63.41	**74.87**	67.21
	Unimog	32.36	50.18	56.68	**66.54**	63.26
	MB1017	46.68	55.89	51.74	**70.23**	56.15
	Wiesel	72.36	72.41	78.66	**84.57**	80.66
False alarm rate (%)	Leopard1	0.99	0.36	0.24	**0.39**	1.13
	Leopard2	3.83	2.94	2.77	**0.39**	3.46
	Jaguar	3.34	3.37	1.67	**1.95**	2.92
	M48	0.53	0.66	0.15	**0.35**	0.92
	Fuchs	3.63	2.34	1.89	**1.26**	1.87
	Hermelin	1.98	1.83	2.01	**1.10**	1.50
	Unimog	0.82	1.34	1.37	**0.64**	0.71
	MB1017	2.39	2.91	1.93	**1.63**	1.78
	Wiesel	3.78	2.94	2.93	**1.61**	2.59
Classification rate (%)		81.14	83.34	85.58	**90.38**	85.20

Misclassification takes place among vehicles that are similar in their engine, weight or tracks. The only exception is the Wiesel, which is a lightweight tracked vehicle. In [42,52], the Wiesel has been excluded from their simulations because the harmonic structure was not distinct and the authors found difficulty in finding the harmonic series that represent the engine signature. From the confusion matrix in Table 4.6, we conclude the following:

1. The extracted features were capable of preserving the significance between ve-

hicles that are close to each other.

2. Misclassifying a heavyweight tracked vehicle as a lightweight wheeled vehicle is a very rare situation and vice versa.

Table 4.6: Confusion Matrix for the proposed TDHA algorithm

True → Experimental ↓	Leopard1	Leopard2	Jaguar	M48	Fuchs	Hermelin	Unimog	MB1017	Wiesel
Leopard1	**3776**	35	28	18	0	0	1	1	14
Leopard2	18	**4060**	159	35	1	0	2	0	37
Jaguar	18	176	**3458**	8	14	17	0	1	56
M48	9	38	9	**2723**	0	0	0	0	1
Fuchs	0	2	9	0	**1339**	34	9	106	49
Hermelin	0	7	37	1	39	**778**	33	75	69
Unimog	17	5	11	2	39	77	**539**	84	36
MB1017	0	1	4	0	114	70	51	**675**	46
Wiesel	6	52	80	0	38	22	33	61	**1601**

Although most of the acoustic energy is concentrated in the lower 250 Hz band, few higher harmonics show up in the signal spectrum. These higher harmonics are very distinct when the vehicle is at close distance to the sensor node. Since higher frequency components tend to attenuate faster with respect to separation, higher components show up only for few seconds around the CPA. Feature extraction consumes a significant amount of the network resources, therefore the approximated harmonic model must consider the fact that having a small frequency resolution and a longer FV might lead to better discrimination but it will consume more power for extraction and a complex classification scheme will be required to process the generated FV. Table 4.7 shows the results for varying the desired spectral resolution (ΔF) and the number of harmonics for the proposed TDHA algorithm. The generated FV combines the ESHC, the event energy and the estimated harmonics' amplitudes. The

results in Table 4.7 are for simulations using the Bochum data set for the following approximations:

- $\Delta F = 5$ Hz and 25 harmonics, approximating the freq. band $5 - 125$ Hz.
- $\Delta F = 10$ Hz and 25 harmonics, approximating the freq. band $10 - 250$ Hz.
- $\Delta F = 5$ Hz and 50 harmonics, approximating the freq. band $5 - 250$ Hz.
- $\Delta F = 10$ Hz and 50 harmonics, approximating the freq. band $10 - 500$ Hz.
- $\Delta F = 15$ Hz and 50 harmonics, approximating the freq. band $15 - 750$ Hz.

It is clear from table 4.7 that by approximating the same frequency band using more components will help in better discrimination. Between the second and third column, doubling the number of components for the 250 Hz band enhances the classification by 10% with the highest improvements in the wheeled and lightweight vehicles. This enhancement costs requires doubling the number of estimated harmonics' amplitudes and doubling the number of neural network input nodes and nodes at the hidden layer which means multiplying the number of weights and biases by 4 (for a fully connected network). Comparing between the last three cases, we find that having a large number of components approximating the target frequency band is much better than having a lower resolution and approximating a wider band. It has been proven in the literature [2] that a longer FV does not always mean better discrimination. On the same tone, the results in the first two columns in Table 4.7 show that although the energy content of the acoustic signal spans the whole 250 Hz,

Table 4.7: Different approximations for the harmonic signal - Bochum data set

Band(Hz)/Resolution(Hz) →		125/5	250/10	**250/5**	500/10	750/15
Feature vector length→		27	27	**52**	52	52
Vehicle ↓						
Detection rate (%)	Leopard 1	95.59	94.78	**97.49**	95.07	95.54
	Leopard 2	91.25	80.89	**94.15**	90.49	87.29
	Jaguar	86.69	79.98	**92.26**	88.52	84.60
	M48	97.20	92.22	**97.94**	96.80	94.99
	Fuchs	80.95	78.30	**86.49**	84.56	77.17
	Hermelin	73.15	61.77	**74.87**	73.69	71.20
	Unimog	61.09	52.03	**66.54**	62.39	58.67
	MB1017	61.00	57.06	**70.23**	62.26	57.32
	Wiesel	80.31	71.77	**84.57**	83.08	77.39
False alarm rate (%)	Leopard 1	0.53	1.44	**0.39**	0.88	0.83
	Leopard 2	2.51	4.09	**0.39**	2.50	3.22
	Jaguar	2.97	4.99	**1.95**	2.68	3.65
	M48	0.60	1.27	**0.35**	0.60	0.73
	Fuchs	1.75	2.40	**1.26**	1.95	2.02
	Hermelin	1.53	1.51	**1.10**	1.52	1.47
	Unimog	0.70	1.04	**0.64**	0.70	0.97
	MB1017	1.84	2.25	**1.63**	1.78	2.26
	Wiesel	2.51	3.18	**1.61**	1.92	2.39
Classification rate (%)		86.79	80.65	**90.38**	87.51	84.55

preserving a lower resolution (5 Hz) and approximating a lower band (125 Hz instead of 250 Hz) was proven to be more efficient over all vehicles when the number of harmonics was fixed to 25 harmonics.

The proposed TDHA algorithm combines the ESHC, the event energy, and the harmonics' amplitudes in a single FV for discrimination. In order to assess the selection of such components and the importance of having such combination for classification, we tested the capability of the ANN to discriminate between the nine vehicles

in the Bochum data set with different FVs. Table 4.8 shows the detection, the false alarm, and the classification rate of the proposed TDHA algorithm with $\widetilde{FF} \sim 5$ Hz and 50 harmonics for the following features:

1. The ESHC and the signal energy.
2. The harmonics' amplitudes.
3. The ESHC and the harmonics' amplitudes.
4. The ESHC, the signal energy, and the harmonics' amplitudes.

From Table 4.8, we can conclude that harmonics' amplitudes are the most effective set of features and the information they carry is valuable for discrimination. The energy and the ESHC are computed prior to estimating the harmonics' amplitudes during detection and feature extraction, respectively. The ESHC and the energy are very important as well but they are not recommended for use by themselves without the harmonics' amplitudes. The event energy and the ESHC can differentiate between different categories like wheeled versus tracked or heavyweight versus lightweight as we will indicate in section 4.6.7 but not to discriminate among two heavyweight tracked vehicles. The difference in performance between the last three columns is very small and the length of the generated FV and the total number of computations required are almost the same.

Table 4.8: Detection, False alarm and classification rate - Different feature vectors

Selected Feature → Vector		ESHC + Energy	H. Amp.	ESHC + H. Amp.	ESHC + Energy + H. Amp.
FV length		2	50	51	52
Vehicle ↓					
Detection rate (%)	Leopard 1	87.19	97.09	95.83	97.49
	Leopard 2	37.63	93.53	93.63	94.15
	Wiesel	59.40	85.85	86.97	84.57
	Jaguar	55.78	90.07	92.23	92.26
	M48	74.97	96.45	97.35	97.94
	Fuchs	64.85	84.84	88.65	86.49
	Hermelin	31.86	68.22	78.09	74.87
	Unimog	0.00	64.53	62.18	66.54
	MB1017	7.15	68.63	71.44	70.23
False alarm rate (%)	Leopard 1	6.63	0.61	0.36	0.39
	Leopard 2	8.33	2.17	1.78	0.39
	Wiesel	7.23	2.45	1.77	1.61
	Jaguar	11.23	1.82	1.99	1.95
	M48	8.26	0.58	0.48	0.35
	Fuchs	7.43	1.22	1.30	1.26
	Hermelin	1.84	0.93	1.21	1.10
	Unimog	0.00	0.71	0.50	0.64
	MB1017	0.62	1.79	1.49	1.63
Classification rate (%)		55.75	89.09	90.32	90.38

4.6.3 Modality

Since the harmonic structure is more distinct in the acoustic signature of ground vehicles than in their seismic signature, all of our previous simulations selected the acoustic modality for evaluating the selected set of characteristic features. In addition to the published results in the literature [2, 40], which showed that vehicle's acoustic signature is more efficient than its equivalent seismic signatures in discrimination among ground vehicles. Many sensor modules are equipped with multiple transducers to capture different modalities especially unattended ground sensors. In many situations the microphone might fail compared to the geophone since the later is buried beneath the ground and it is less affected by the weather and spoofing. The proposed TDHA algorithm can be used to extract the features from the seismic signature of ground vehicles. Tables 4.9 and 4.10 compare between acoustic and seismic modalities for detection and classification of vehicles using both data sets, Bochum and SensIT, respectively. Applying the same detection algorithm over both modalities where the energy of each event is compared to an adaptive threshold preserving a CFAR. The same classification technique is applied using the same ANN architecture, training scheme, and cost function. The features were extracted from the harmonic model approximating the frequency band $5 - 250$ Hz using 50 harmonics and $\Delta F = 5$ Hz. Since the harmonic structure is more distinct in the acoustic signature, the $\widetilde{FF}$ is estimated from the acoustic signal for both approximations (acoustic and seismic).

Although the total number of detected events using the seismic signal varies (higher for the Bochum data and lower for the SensIT data), the extracted features from the seismic signal did not show the same capability in discrimination as the

acoustic features in the two data sets. The variation in the total number of detected events is a confirmation for the high dependence of the seismic propagation on the substrate which has a large effect on the seismic wave propagation. We conclude that, acoustic features are more successful in vehicle verification than their equivalent seismic features. Although we again note the value of the seismic signal in detection.

Table 4.9: Acoustic versus seismic modality - Bochum data set

Modality →		Acoustic	Seismic
Total number of Detected events →		30000	36000
Vehicle ↓			
Detection rate (%)	Leopard 1	97.49	74.17
	Leopard 2	94.15	61.10
	Wiesel	84.57	69.12
	Jaguar	92.26	66.86
	M48	97.49	63.03
	Fuchs	86.49	51.24
	Hermelin	74.87	65.40
	Unimog	66.54	49.22
	MB1017	70.23	70.11
False alarm rate (%)	Leopard 1	0.39	3.80
	Leopard 2	1.89	3.63
	Wiesel	1.61	3.86
	Jaguar	1.95	5.00
	M48	0.35	4.51
	Fuchs	1.26	5.16
	Hermelin	1.10	5.87
	Unimog	0.64	4.55
	MB1017	1.63	4.74
Classification rate (%)		90.38	63.40

Table 4.10: Acoustic versus seismic modality - SensIT data set

Modality →		Acoustic	Seismic
Total number of Detected events →		36000	14000
Vehicle ↓			
Detection rate(%)	AAV	89.77	80.04
	DW	77.84	63.91
	Noise	84.48	36.77
False alarm rate(%)	AAV	8.91	32.26
	DW	5.97	19.82
	Noise	7.68	2.28
Classification rate(%)		85.24	71.60

4.6.4 Feature space

Most of the previous military vehicle verification schemes generate their characteristic features from the spectral domain (FFT components). We propose that time domain features generated using the proposed TDHA algorithm for approximating the deterministic signature of the acoustic signal are capable of performing as well as the spectral features with fewer computations.

In order to evaluate the discrimination ability of harmonics' amplitudes as compared to spectral features, FFT components were computed for three different spectral resolutions (5, 10, and 20 Hz for NFFT points = 1024, 512, and 256, respectively, at 5 KHz sampling frequency). NFFT spectral components were computed for each detected event with a duration of 0.5 seconds using a rectangular window. Three

different windows were tested (Hamming, Hanning, and rectangular) and the spectrum obtained using the rectangular window resulted in better classification than the other two windows. The FFT characteristic FV that will be used in discrimination is constructed from the event energy followed by the first 50 FFT coefficients approximating the spectral signature in the low pass band up to 250 Hz, 500 Hz, and 1000 Hz, respectively. This results in a FV of length 51. The TDHA method was applied for the same set of detected events for the same approximate spectral resolutions (ΔF values of 5, 10, and 20 Hz) and using 50 harmonics. Table 4.11 illustrates the average detection, false alarm and classification rate for the proposed TDHA algorithm versus the spectral features generated for the same set of spectral resolutions.

It is clear from table 4.11 that for a fixed FV length, the TDHA modeling of the deterministic signature of the acoustic emission of vehicles is at least as effective in vehicle classification as the equivalent spectral features from the FFT method. The classification rate of lightweight vehicles is less than the equivalent heavyweight vehicles as expected since the harmonic structure is less distinct. Although the variety of vehicles and unequal number of events per vehicle make the classification phase very challenging to the NN, the classification performance for the given FVs is quite good. For a fixed number of harmonics, constructing the harmonic model using a smaller base frequency and achieving more harmonic resolution (smaller frequency bands in spectral domain) allows for better approximation of the acoustic signal over both the time domain and spectral features (on condition that the total frequency band includes a significant amount of the acoustic energy, i.e. $0 - 250$ Hz).

A more realistic scenario is presented, where the ANN is trained using events that are detected using one of the two sensor stations (Bochum data set) and is

Table 4.11: Time domain features versus spectral features - Bochum data set

Features →		Harmonics' amplitudes			Spectral		
Model →		$\widetilde{FF\alpha}$ strongest component			FFT using a rectangular window		
Vehicle ↓ Resolution →		∼ 5Hz	∼ 10Hz	∼ 20Hz	5Hz	10Hz	20Hz
Detection rate (%)	Leopard 1	97.49	95.07	94.55	95.69	92.29	90.76
	Leopard 2	94.15	90.49	81.71	90.16	85.68	73.49
	M48	97.94	96.80	92.46	94.29	92.44	85.25
	Jaguar	92.26	88.52	83.89	89.79	86.40	74.56
	Fuchs	86.49	84.56	77.44	86.43	84.02	80.51
	Hermelin	74.87	73.69	63.77	72.86	72.94	64.67
	Unimog	66.54	62.39	53.14	66.33	61.20	51.70
	MB1017	70.23	62.26	53.62	69.49	60.85	53.27
	Wiesel	84.57	83.08	79.85	83.28	85.17	73.82
False alarm rate (%)	Leopard 1	0.39	0.88	1.13	0.83	1.47	2.86
	Leopard 2	1.89	2.50	3.68	2.67	3.22	5.33
	M48	0.35	0.60	1.29	0.87	1.25	2.22
	Jaguar	1.95	2.68	4.54	2.46	3.12	6.12
	Fuchs	1.26	1.95	2.49	1.20	1.66	2.45
	Hermelin	1.10	1.15	1.32	1.09	1.27	1.45
	Unimog	0.64	0.70	0.90	0.66	0.73	1.08
	MB1017	1.63	1.78	2.24	1.71	1.77	1.77
	Wiesel	1.61	1.92	2.83	2.12	2.46	3.58
Classification rate (%)		90.38	87.51	82.17	88.02	85.19	76.80

tested using events detected by the other station. Table 4.12 illustrates the average detection, false alarm and classification rate, where the training and testing of the ANN is performed using events that are detected by two separate stations that are 101 m apart and are on two different sides of the road with respect to the moving vehicles. Although the classification rate was reduced for both the proposed TDHA algorithm and the spectral features, the simpler TDHA approach again performs at least as well as the spectral features.

Table 4.12: Time domain features versus spectral features - Bochum data set - Different training and testing stations

Features →		Harmonics' amplitudes			Spectral		
Model →		$\widetilde{FF}$ α strongest component			FFT using a rectangular window		
Vehicle ↓ Resolution →		$\sim$ 5Hz	$\sim$ 10Hz	$\sim$ 20Hz	5Hz	10Hz	20Hz
Detection rate (%)	Leopard 1	89.82	91.94	91.00	87.98	85.20	86.90
	Leopard 2	90.48	92.05	89.34	88.51	92.46	86.18
	M48	78.96	77.46	68.39	78.86	76.63	63.54
	Jaguar	82.51	79.51	71.74	78.04	72.08	64.35
	Fuchs	80.37	77.91	69.05	75.80	71.51	61.87
	Hermelin	75.12	73.96	49.38	71.94	73.22	36.24
	Unimog	62.94	55.88	35.44	59.41	57.05	36.02
	MB1017	66.66	59.22	65.42	68.18	61.15	78.98
	Wiesel	80.10	79.52	77.44	75.14	82.25	79.23
False alarm rate (%)	Leopard 1	0.37	0.69	0.52	0.82	0.68	1.57
	Leopard 2	6.15	7.24	9.94	6.59	8.68	12.77
	M48	1.41	0.45	0.63	1.60	1.69	0.89
	Jaguar	1.91	2.02	3.46	1.99	1.65	4.11
	Fuchs	1.04	1.36	2.42	0.82	1.01	1.46
	Hermelin	3.11	3.05	1.68	3.55	3.19	0.76
	Unimog	1.13	1.52	1.67	1.30	1.12	0.68
	MB1017	3.07	3.36	5.28	3.95	3.62	6.11
	Wiesel	2.07	1.76	3.78	2.53	2.83	3.84
Classification rate (%)		82.10	81.10	78.04	79.50	78.50	72.16

In another realistic scenario, we tested the performance of the ANN for features extracted from events for certain runs that have been used in the training phase. In this setting, the ANN is trained using events detected by both stations from half of the recorded runs and is tested using events detected by the other half of the runs (and vice versa for cross-validation). For the TDHA approach with an $\widetilde{FF} \sim 5$ Hz and 50 harmonics, the average single event classification rate is 82.77%. These results are virtually identical to those shown in Table 4.12. This was true for other cases as

well.

4.6.5 Classifier selection

All our previous simulations use an ANN for classification. Although it has been shown in the literature [2, 4, 5, 62] that ANNs are the preferred classifier for features extracted from the vehicle's acoustic emissions, in this section we evaluate the performance of the TDHA algorithm combined with three different classifiers. Simulations results in Table 4.13 compares between different classification schemes for the same FV generated using the harmonics' amplitudes estimated from the approximation of the deterministic signal of the acoustic signature with an $\widetilde{FF} \sim 5$ Hz and using 50 harmonics. The characteristic FV is constructed from only the 50 estimated harmonics' amplitudes to avoid weighting individual FV dimensions according to the dynamic range of this feature for an LVQ classifier or a MAP classifier.

LVQ uses a splitting technique as previously discussed in chapter 2. Two different codebooks were generated for LVQ; the first assigns 64 code vectors per class, while the second assigns 128 vectors per class. Although increasing the size of the codebook improves the capability in vehicle discrimination, the computational complexity increases exponentially with the size of the codebook. The MAP classifier is memory and computational efficient for a small number of vehicles when compared to other classifiers, but its performance is limited when the number of candidate vehicles is large. From Table 4.13, it is clear that ANN classification rate is better than MAP and LVQ (64 vectors/class). Although the classification rate of the wheeled vehicles

using LVQ (128 vectors/class) is better than ANN, ANNs are characterized by better discrimination among heavyweight tracked vehicles and their ability to acquire characteristic features that have other features/measurements (such as the ESHC and the event energy) beside the harmonics' amplitude. The fact that the cost function for an LVQ or a MAP classifier is the Euclidean distance along all input vector dimensions, which requires finding a weighting scheme for input vectors that include diverse features, limits the capability of combining different measurements for generation of a single characteristic feature per event. In comparison, ANNs adjust their weights and biases such that the dynamic range of each dimension of the input vector is ineffective to the cost function (mean square error). Sensor nodes have limited storage capabilities and although ANNs require storage of the weights and biases, LVQ requires storing all of the code vectors. For a full connected ANN with a FV length of 50 and 30 neurons in the hidden layer and 9 neurons at the output layer, the total number of weights are 50*30 + 30*9, and the total number of biases are 30 + 9, which gives a total number of entries for the ANN of 1809. LVQ (128 vectors per class) requires storing all code vectors, making the total number of entries equal to 50*9*128 or 57600 (more than 30 times than that required by the ANN). There are many algorithms for storing the table in compressed format, but this will impact both the ANN and the LVQ classifiers' storage requirements.

4.6.6 Fusion

Multimodal decision fusion was previously proposed for reducing the noise effect by combining individual decisions regarding either temporal events on the same node or

Table 4.13: Different classifiers - Bochum data set

Feature vector →		Harmonics' amplitudes			
Classifier →		ANN	LVQ (64)	LVQ (128)	MAP
Vehicle ↓					
Detection rate (%)	Leopard 1	97.09	90.76	96.30	50.97
	Leopard 2	93.53	77.71	86.63	75.97
	Jaguar	90.07	86.85	90.72	60.36
	M48	96.45	87.61	94.41	54.81
	Fuchs	84.84	86.99	87.21	59.88
	Hermelin	68.28	79.98	78.62	63.02
	Unimog	64.53	69.06	72.65	65.99
	MB1017	68.63	76.46	75.46	73.98
	Wiesel	85.85	75.69	80.93	61.58
False alarm rate (%)	Leopard 1	0.61	1.31	0.75	0.07
	Leopard 2	2.17	2.35	1.23	13.03
	Jaguar	1.82	3.44	1.89	0.33
	M48	0.58	0.93	0.58	0.24
	Fuchs	1.22	1.96	1.69	0.72
	Hermelin	0.93	2.10	1.78	1.62
	Unimog	0.71	1.43	1.33	5.12
	MB1017	1.79	2.28	1.65	11.98
	Wiesel	2.45	2.97	2.18	9.21
Classification rate (%)		89.09	83.33	88.23	62.13

the same event on distributed nodes (spatial fusion). Although fusion might require extra processing, storage and/or communication, fusion can be achieved using simple and energy efficient schemes. In this section, we will investigate the benefits of applying multi-modal fusion at lower processing levels (relay fusion) and at higher processing levels (decision fusion).

We have shown that the extracted features from the acoustic signal are more efficient than their equivalent seismic signals in vehicle verification. Seismic features

can be helpful in two scenarios: very high wind and background noise or the damage of the acoustic transducers (microphones). Beside these two scenarios seismic energy has shown both a great potential in detection of events (not classification) when compared to the acoustic energy and a high resistivity to spoofing. Most UGS are equipped with multiple transducers. Data fusion requires defining a correlation or a coupling scheme between different modalities and consumes a considerable amount of the sensor power. Relay fusion as previously described in Chapter 2 is another fusion method that occurs at a low data level and does not require finding a correlation between different data signals. Relay fusion proposes utilizing the seismic data in the detection phase and the acoustic data in the feature extraction phase. Tables 4.14 and 4.15 compare between detection, false alarm, and classification rate for a single modality scenario compared to a relay fusion scenario using Bochum and SensIT data sets, respectively.

The results in Tables 4.14 and 4.15 prove that seismic energy can be very reliable in detecting moving vehicles independent on any background or wind noise. The only drawback in seismic detection is the limited detection range as compared to acoustic energy. The range for acoustic detection can exceed few hundred meters while that for seismic detection is limited to a maximum of 200 meters for heavyweight tracked vehicles. Seismic signals produced by either wheeled or tracked vehicles will be detected for at least few events during the run, which makes seismic detection suitable for implementing a reliable and early warning system for peacekeeping operations in specific territories that do not allow deployment of microphones. The microphones in SensIT data set captured a lot of wind and background noise which forced the acoustic

Table 4.14: Seismic-Acoustic relay fusion compared to a single modality scenario - Bochum data set

Detection Modality →		Acoustic	Seismic	Seismic
Classification Modality →		Acoustic	Seismic	Acoustic
Total number of Detected events →		30000	36000	36000
Vehicle ↓				
Detection rate (%)	Leopard 1	97.49	74.17	95.68
	Leopard 2	94.15	61.10	91.36
	Jaguar	92.26	66.86	91.83
	M48	97.49	63.03	96.60
	Fuchs	86.49	51.24	77.72
	Hermelin	74.87	65.40	52.25
	Unimog	66.54	49.22	70.88
	MB1017	70.23	70.11	71.46
	Wiesel	84.57	69.12	85.94
False alarm rate (%)	Leopard 1	0.39	3.80	0.29
	Leopard 2	1.89	3.63	1.86
	Jaguar	1.95	5.00	2.14
	M48	0.35	4.51	0.55
	Fuchs	1.26	5.16	1.58
	Hermelin	1.10	5.87	0.49
	Unimog	0.64	4.55	2.00
	MB1017	1.63	4.74	2.14
	Wiesel	1.61	3.86	3.19
Classification rate (%)		90.38	63.40	87.37

Table 4.15: Seismic-Acoustic relay fusion compared to a single modality scenario - SensIT data set

Detection Modality →		Acoustic	Seismic	Seismic
Classification Modality →		Acoustic	Seismic	Acoustic
Total number of Detected events →		36000	14000	14000
Vehicle ↓				
Detection rate (%)	AAV	89.77	80.04	94.22
	DW	77.84	63.91	81.43
	Noise	84.48	36.77	66.97
False alarm rate (%)	AAV	8.91	32.26	9.63
	DW	5.97	19.82	5.12
	Noise	7.68	2.28	4.44
Classification rate (%)		85.24	71.60	88.61

detector to declare more than double the events detected by the seismic detector as moving vehicles (Table 4.15), although both vehicles in this data set were heavyweight military vehicles. Most of the difference between the total number of events that has been detected using seismic energy and acoustic energy for SensIT data set is for the noise class. The wind and background noise in the Bochum data set is very low compared to the SensIT data set, which suggests that in this specific situation, detection and classification using only the acoustic emission of ground vehicles is preferable to relay fusion.

In many situations, high level fusion is a possible approach to increase the confidence in the monitoring system. In addition to increasing the confidence in the system, temporal decision fusion between individual events at the same sensor node reduces network traffic and produces single decisions per node over long periods instead of sending a decision every half second. Since decision fusion will be implemented over the sensor node, a simple and energy efficient scheme is required such as simple voting between individual decisions to decide upon a winning vehicle class (a winner takes all scheme). Table 4.16 shows the average classification rate for a single event decision and for employing decision fusion for the SensIT and the Bochum data sets.

Although the SensIT data set contains two heavyweight vehicles, but the average classification rate for individual decisions was less than that achieved for Bochum data set. The classification rate was improved by applying decision fusion using voting among all detected events at a sensor node regarding a single vehicle crossing the sensor field. It is clear from Table 4.16 that decision fusion was capable of enhancing

the classification rate which intern increases the confidence in the network decision. Decision fusion can be as well combined with other types of low level fusion such as data or relay fusion.

Table 4.16: Single decision versus decision fusion - SensIT data set(3 classes) and Bochum data set(9 classes)

Data set	SensIT	Bochum
Modality	Acoustic	
Algorithm	TDHA ($\widetilde{FF} \sim$ 5Hz and 50 harmonics)	
Classification rate - single event	85.24	90.38
Classification rate - decision fusion	93.72	92.05

We have considered processing events that were detected using either their seismic or acoustic energy. Peacekeeping operations and other monitoring applications might require more confidence before sending an alert message. One way of increasing the confidence in the detected events without missing these events and with minimum communication among sensor modules, is through multi-modal detection. This means that only events that are detected using both modalities (seismic and acoustic energy) will be considered for further processing. Table 4.17 shows the detection, false alarm, and classification results for features generated using the proposed TDHA algorithm for events that have acoustic and seismic energies above their adaptive threshold levels. The total number of events is much less than those detected by either acoustic or seismic modalities. Although we showed that the acoustic features are better than their equivalent seismic features in discrimination, Table 4.17 compares between acoustic and seismic features as well as a combined FV that concatenate features from the acoustic and seismic signature of the ground vehicle. Although combining

Table 4.17: Multi-modal detection and different characteristic features (Single event) - Bochum data set

Characteristic features →		Harmonics' Amplitudes		
Classification Modality →		Seismic	Acoustic	Seismic+Acoustic
FV length →		22	52	72
Approximated Band (Hz)		5-100	5-250	5-100 + 5-250
Vehicle ↓				
Detection rate (%)	Leopard 1	84.31	94.85	94.15
	Leopard 2	60.10	80.08	89.22
	Jaguar	72.43	8.50	87.56
	M48	68.78	95.37	95.38
	Fuchs	48.96	84.72	85.65
	Hermelin	67.50	76.53	82.97
	Unimog	54.05	71.15	74.67
	MB1017	70.38	70.84	80.27
	Wiesel	70.17	82.40	85.29
False alarm rate (%)	Leopard 1	2.67	0.48	0.54
	Leopard 2	3.78	1.42	1.25
	Jaguar	3.43	1.99	1.63
	M48	3.88	1.19	0.45
	Fuchs	5.30	2.49	1.94
	Hermelin	4.77	2.72	2.85
	Unimog	6.63	2.39	2.19
	MB1017	4.16	3.88	3.35
	Wiesel	3.23	3.10	1.56
Classification rate (%)		66.30	82.47	85.96

features increased the classification rate, but it requires a larger neural network and more computations which must be taken into consideration.

Table 4.18 shows the temporal decision fusion results for the single decisions obtained for the multi-modal detection scenario with enhancements over all classes independent on the characteristic features.

Table 4.18: Multi-modal detection and different characteristic features (Decision fusion) - Bochum data set

Characteristic features →		Harmonics' Amplitudes		
Classification Modality →		Seismic	Acoustic	Seismic+Acoustic
FV length →		22	52	72
Approximated Band (Hz)		5-100	5-250	5-100 + 5-250
Vehicle ↓				
Detection rate (%)	Leopard 1	95.50	98.52	100.00
	Leopard 2	70.00	95.18	98.78
	Jaguar	91.23	98.43	95.45
	M48	83.03	100.00	98.41
	Fuchs	66.66	95.00	100.00
	Hermelin	84.16	92.85	100.00
	Unimog	58.94	70.00	73.07
	MB1017	73.25	84.61	87.03
	Wiesel	90.99	92.64	96.92
False alarm rate (%)	Leopard 1	2.09	0.00	0.96
	Leopard 2	1.50	0.25	0.25
	Jaguar	1.62	0.47	0.00
	M48	4.33	0.95	0.00
	Fuchs	2.71	0.43	0.86
	Hermelin	3.40	1.92	2.57
	Unimog	3.89	1.62	0.23
	MB1017	1.41	1.39	1.17
	Wiesel	1.01	1.20	0.24
Classification rate (%)		80.32	92.53	94.16

4.6.7 Comparison with published results

The research group that conducted the original BVP experiment [42, 52] developed a processing tool for detection and classification of military vehicles using their acoustic and seismic signatures. They used the relative powers of the first 15 harmonically related components. They computed the FFT components and located the harmonic line series using the distinct peaks found in the spectrum followed by selection of the powers of the wining harmonic peaks among all detected peaks for generation of the feature vector. The wining peaks are the ones that form a harmonic line series with a fundamental between 8 and 20 Hz. They used learning vector quantization for classification of their weighted power feature vector. The weighting was proven to be helpful since the power of the first 15 harmonics can contain more than 99% of the total line series power [42]. After processing all recordings in 0.82 seconds intervals, a random selection of two out of each three feature vectors are used for learning, with the rest used for testing. In [42, 52] only 8 vehicles out of 10 were used for evaluation since the remaining two vehicles (Wiesel and VW van) did not show strong harmonic structure and the harmonic line series could not be extracted from the acoustic data. They also attempted performing multi-level classification by using the maximum amplitudes of both the acoustic and seismic signals in separating tracked vehicles from wheeled vehicles, then using the relative powers of the first 15 harmonic peaks in determining the detected vehicle.

We tested our proposed TDHA algorithm combined with a feedforward ANN for classification of military vehicles on the same 8 vehicles and by using the same ratio of trained/tested features, but when we attempted to perform multi-level classification,

we selected the ESHC, the acoustic energy, and the seismic energy in separating tracked vehicles from wheeled ones, followed by utilizing the harmonics amplitudes in determining the class of the vehicle. Table 4.19 compares between the results obtained by the group of researchers who collected the Bochum data set [42,52] with our results for the harmonics' amplitudes features approximating the deterministic signal in the band $[5 - 250]$ Hz. *Altmann* et. al. in $[42, 52]$ did not include the false alarm rate in their results, but when we compare the classification rate, we find that in both single level and multi-level classification, our algorithm is very competitive to their results and we achieved a higher average classification rate and a higher detection rate for the lightweight vehicles and most of the heavyweight tracked ones. In addition to the simplicity in extracting the characteristic features compared to the algorithm proposed in $[42, 52]$.

The researchers who collected SensIT data set proposed a distributed detection and classification system that is capable of producing single event decisions and multiple events' decisions utilizing decision fusion across distributed sensors. They have three classes (Assault Amphibious Vehicle, Dragon Wagon vehicle, and Noise) and use acoustic and seismic features for classification. Events of duration 0.75 seconds were used. For the acoustic modality, 100 FFT components with resolution of 9.6875 Hz were pair wise averaged into 50 components with resolution of 19.375 for generating the acoustic FV carrying the information for frequencies up to ~ 1 KHz. For the seismic modality, 50 FFT components with resolution of 9.6875 Hz were selected for generating the seismic FV carrying the information for frequencies up to ~ 0.5 KHz. They performed classification of both acoustic and seismic features using K-Nearest

Table 4.19: Comparison with the *BVP* published results

Features →		Harmonics' Amp.	Spectral [52]	Harmonics' Amp.	Spectral [52]
Classifier →		ANN	LVQ	ANN	LVQ
Classification		Single		Multi-level	
Vehicle ↓					
	Tracked	-	-	98.63	97.00
	Wheeled	-	-	96.65	89.56
Detection rate (%)	Leopard 1	96.18	99.00	97.83	98.00
	Leopard 2	94.65	88.00	95.79	96.00
	Jaguar	95.18	91.00	92.47	95.00
	M48	97.72	95.00	98.76	98.00
	Fuchs	88.86	86.00	89.56	95.00
	Hermelin	79.31	67.00	83.39	75.00
	Unimog	68.58	60.00	75.73	75.00
	MB1017	74.64	64.00	79.21	69.00
	Tracked	-	-	3.34	NA
	Wheeled	-	-	1.36	NA
False alarm rate (%)	Leopard 1	0.42	NA	0.28	NA
	Leopard 2	1.89	NA	1.07	NA
	Jaguar	1.96	NA	1.74	NA
	M48	0.28	NA	0.35	NA
	Fuchs	1.28	NA	1.01	NA
	Hermelin	1.14	NA	1.43	NA
	Unimog	0.68	NA	0.66	NA
	MB1017	1.44	NA	1.24	NA
Classification rate (%)		92.07	89.00	98.19	93.12

Neighbor, ML, and Support Vector machines. They used $\frac{1}{3}$ of the generated features for training and the remaining $\frac{2}{3}$ for testing.

We tested our proposed monitoring system for verification of military vehicles on the same classes and by using the same ratio of trained/tested features. Table 4.20 compares between the results obtained by the group of researchers who collected the SensIT data set [40] with our results for the harmonics' amplitudes features approximating the deterministic signal in the band 5 - 250 Hz. The only difference between the data used in SensIT published results and our simulations is the manual labeling of the noise class (detected vehicle that has the noise level higher than the vehicle sound). The original labeling was lost when the data was imported from one server to the other and we manually labeled the acoustic recordings. For our seismic simulations, we used the same labeling performed using the acoustic signal which results in a very small number of seismically detected events that are originally labeled as noise. It is clear from Table 4.20 that our approximated harmonic model and extracted features approximating the band 5 - 250 Hz have higher discrimination capability and very low false alarm rate compared to the original results obtained by the researchers who collected the original data. The only improvements in their results is in the noise class while using the seismic modality and the reason for that is the small number of detected noise events using seismic modality. ANNs are very sensitive to the training set size, as previously discussed in section 4.3. Finally, we must point out the remarkably low false alarm rate obtained using our proposed features compared to the spectral features used by the researchers who collected the SensIT data set.

Table 4.20: Comparison with *SensIT project* published results

Features →		Harmonics' Amp.		Spectral [40]					
Detection Modality →		Acoustic	Seismic	Acoustic			Seismic		
Classifier →		ANN	ANN	k-NN	ML	SVM	k-NN	ML	SVM
Vehicle ↓									
Detection rate (%)	AAV	89.77	80.04	67.48	73.63	66.54	55.32	67.04	58.01
	DW	77.84	63.91	61.18	65.39	60.09	49.95	40.05	56.76
	Noise	84.48	36.77	75.26	68.66	76.66	60.51	74.32	71.03
False alarm rate (%)	AAV	8.91	32.26	29.39	34.50	19.58	56.57	52.33	48.58
	DW	5.97	19.82	32.88	39.36	40.38	54.81	44.81	47.62
	Noise	7.68	2.28	30.07	22.72	29.35	23.81	22.08	19.56
Classification rate (%)		85.24	71.60	69.36	68.95	69.48	56.24	62.81	63.79

Chapter 5

Analog-to-Information: Classification of Harmonic Signals

5.1 Overview

In this chapter, an Analog-to-Information (ATI) approach is developed to extract representative features directly from acquired harmonic analog signals for signal classification. The reduction in the computational load is achieved by extracting crucial information that is relatively sparse in the frequency domain directly from the analog time domain harmonic signal. Sources that have rotating machinery emit acoustic or vibration signals that have a distinct harmonic structure with the majority of the signal energy lying in its harmonics. Harmonic signals can be defined from the amplitude of each harmonic component which can be obtained from the frequency domain. The ATI approach replaces a high resolution A/D converter followed by an FFT with a

multichannel analog projection and integration to extract the sparse spectral features for signal classification. The proposed approach does not require knowing the FF of the harmonic series or the number of harmonics or estimating the strongest harmonic component as was done in the TDHA method described in Chapter 4. Compared to commonly used feature extraction techniques (FFT and wavelets), the complexity in extracting the harmonic signal characteristics using the proposed ATI approach in a parallel processing approach is much smaller.

The proposed ATI approach requires a challenging analog electronic implementation. The complete analog design and error estimates has been performed by Daniel J. White, at the DSP Lab, University of Nebraska-Lincoln. In this Chapter we will describe the proposed approach and present the simulation results as a feasibility study that address the key design issues.

In Section 5.2, we will briefly review the motives and challenges for bringing the feature extraction phase to the analog front-end with emphasis on the power consumption and the overall system performance, Section 5.3 describes the basic ATI approach suitable for implementation on extremely low power sensor modules, Section 5.4 presents experimental results obtained for detection and classification of military vehicles using acoustic signals, and Section 5.5 presents experimental results obtained for identification of bearings faults in induction motors using the motor's vibration signal.

5.2 Background

One of the major components that consumes the node's power during signal processing is the A/D circuitry, especially at high speed. Another major component is the high speed digital signal processing circuitry. Compressed Sensing (CS) [117] was previously proposed as an efficient signal acquisition approach to extract sparse signal characteristics using random projection over orthogonal basis to avoid sampling at high rates. Signal characteristics obtained using CS are randomly distributed across projected data and cannot be used for discrimination and they are only useful for signal regeneration using a complex algorithm that is not suitable for low power sensor modules. The majority of signal detection and classification schemes transform the acquired signal into a favorable domain that can reveal significant characteristics in a relatively condensed feature vector. Spectral and wavelet features have been proven to be adequate for most signal detection and classification tasks such as speech recognition [97], vehicle detection [40, 52, 53, 62], fault detection [26–28, 66], etc. Spectral and wavelet coefficients are characterized by being able to represent signal characteristics using a compact number of components, but they require sampling the acquired signal at high rates prior to processing. Several low power monitoring schemes use statistical parameters such as signal mean, standard deviation, peak, etc., obtained from the time domain signal to generate signal characteristics for identification [118], but they are limited in the amount of information carried compared to spectral features.

Analog approaches are generally used whenever the size, weight and power consumption is of primary concern to the system designer [119]. The challenge is in

having an analog system that performs as well as its equivalent digital system irrespective of the common analog problems as mismatch, non-linearity, offset, and error propagation. While the art of designing low power analog computational devices for extracting the information of interest is usually not as automated as in digital designs, the resulting analog computational system can be far more power efficient than its equivalent digital realization [120]. Several approaches proposed increasing the analog processing share in a mixed signal approach to reduce the rate at which the A/D must operate [121].

Characteristic features can be evaluated based on two aspects; the first, is represented by the amount of resources (power and area) needed to extract those features. The second is the amount of information carried by those features that hold the signal's significant characteristics. A classic representation of the resource-precision relationship for both analog and digital systems for sub-threshold technology is shown in Figure 5.1 (as described in [103]) with similar curves obtained for the area as well. The precision is represented by the SNR with 10 bits equivalent to 60.2 dB. The exact location of the cross over point and the power values highly depend on the task, technology and the skills of the analog and digital designer. The noise gain, bandwidth, clock frequency, average switching capacitance, etc., chosen to produce a crossover that occurs at 55 dB or $<$ 10 bits. The comparison in [103] did not include the fact that the actual power consumption of the digital system is higher since an A/D will be needed for a digital system with an analog input, which is the case in most wireless sensor applications. The power and area costs to operate an A/D at high precision and speed are high.

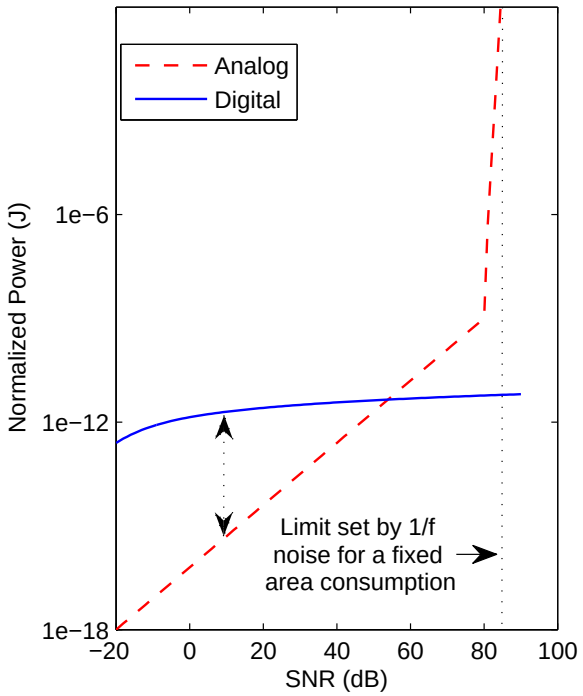

Figure 5.1: Resource-Precision for subthreshold technology

As Figure 5.1 shows, the low power requirements of analog systems at low SNR is attractive. However, this approach requires an ability to mitigate several non-ideal aspects of analog systems such as mismatch, offset, non-linearity, etc. In section 5.4, we will demonstrate ATI performance at SNRs of 10 and 20 dB, taking into consideration expected analog hardware errors obtained from the $0.13\mu m$ hardware models designed by Daniel White [122].

5.3 Analog-to-Information Front-End Approach

Acoustic and vibration signals emitted from rotating machinery acquired at several passive sensors can be modeled as a sum of a deterministic harmonic signal and a non deterministic component. It was shown in [52, 53, 62] that selective harmonic features extracted from the captured acoustic signal are sufficient for signal/source discrimination.

The ATI approach estimates the harmonics' amplitudes of the deterministic harmonic signal that extends over the band where the original signal energy is concentrated, without the need to estimate or track the FF or the number of harmonics. We are trying to approximate the complex harmonic structure to a single harmonic series and use its amplitudes in discrimination. The signal is processed in short windows to ensure a relatively constant harmonic signature. For each window, The ATI approach generates the set of harmonics' amplitudes that best approximate the deterministic signature within a defined band. These amplitudes will be the sole features

for identifying the underlying signal. The extracted information is sufficient for signal classification. It is not used for signal regeneration.

The FF and number of harmonics of the approximated harmonic signal are selected such that this signal has the highest spectral resolution given the frequency band of interest and the limited resources of the wireless sensor node. For example, if the acquired signal has most of its energy in the bandwidth $0 - 500$ Hz, and the system allows the use of a characteristic feature of length 100 per window, the ATI approach uses 5 Hz as the FF and 100 harmonics as the number of harmonics. In another situation, if the signal energy is mostly in the band of $500 - 1000$ Hz, and the length of the feature vector is 50, the ATI transformation uses 10 Hz as the FF and estimate the amplitude of the 51^{st} to 100^{th} harmonics for signal identification without the need to estimate the lower harmonics' amplitudes[1].

5.3.1 Analog Projection

A harmonic signal modeling the deterministic signature can be described as

$$x(t) = \sum_{k=1}^{M} \alpha_k \cos(2\pi k f t + \phi_k), \tag{5.1}$$

where $\alpha_k \geq 0$ and ϕ_k are the amplitude and phase of the k^{th} harmonic respectively, f is the FF, and M is the total number of harmonics. For estimation of each harmonic amplitude α_k, the input signal is projected onto a pair of basis functions with

[1]Unlike the FFT, which generates the spectral coefficients of the band $0 - F_s/2$, the ATI approach is designed to estimate the harmonics' amplitudes of selected higher bands without the need to estimate those of the lower band.

frequency kf and the mixed signal is integrated over $T = 1/f$ producing

$$y_{Il} = \int_0^T \sum_{k=1}^{M} \alpha_k \cos(2\pi k f t + \phi_k)\ \ \psi_{Il}(t)dt,$$
$$y_{Ql} = \int_0^T \sum_{k=1}^{M} \alpha_k \cos(2\pi k f t + \phi_k)\ \ \psi_{Ql}(t)dt, \tag{5.2}$$

where $l = 1, 2, \cdots, L$ and $L \geq M$. Each pair of basis functions represents even and odd square functions with respect to the midpoint of the time window. $\psi_{Il}(t)$ is the in-phase basis and $\psi_{Ql}(t)$ represents the quadrature basis given as

$$\psi_{I1}(t) = \begin{cases} 1 & 0 \leq t < \frac{T}{4} \\ -1 & \frac{T}{4} \leq t < \frac{3T}{4} \\ 1 & \frac{3T}{4} \leq t < T \end{cases} \quad \text{and } \psi_{Il}(t) = \psi_{I1}(lt - \lfloor \tfrac{lt}{T} \rfloor T), \quad l = 2, 3, \cdots, L$$
$$\psi_{Q1}(t) = \begin{cases} 1 & 0 \leq t < \frac{T}{2} \\ -1 & \frac{T}{2} \leq t < T \end{cases} \quad \text{and } \psi_{Ql}(t) = \psi_{Q1}(lt - \lfloor \tfrac{lt}{T} \rfloor T), \quad l = 2, 3, \cdots, L \tag{5.3}$$

where $\lfloor \frac{lt}{T} \rfloor = \max\{D \in \mathbb{Z} \mid D \leq \frac{lt}{T}\}$. Figure 5.2 shows the basis pairs for several harmonics up to $l = 10$ harmonics with $f = 1/T$.

5.3.2 Feature Extraction

The simple harmonic signal in (5.1) can be expressed using the in-phase and quadrature components as

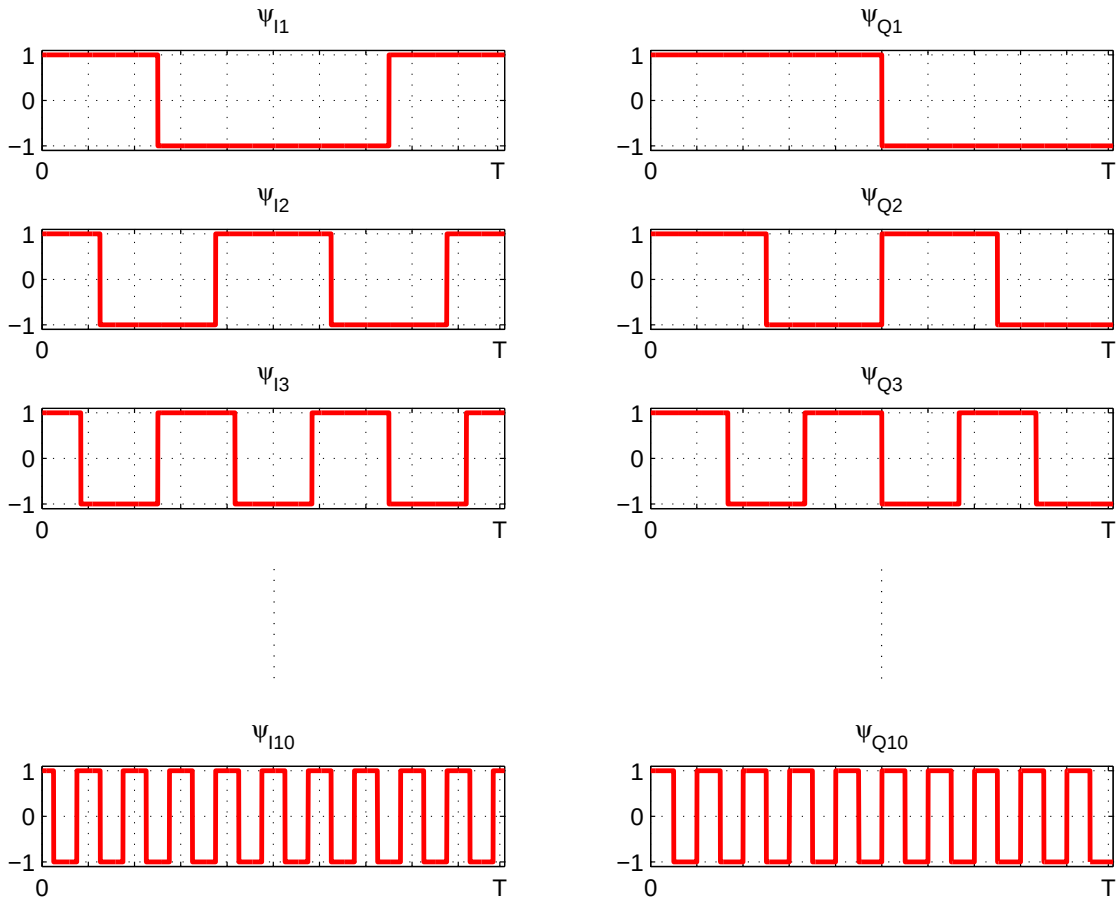

Figure 5.2: Basis Functions for $l = 1, 2, ..., 10$

$$x(t) = \sum_{k=1}^{M} \alpha_k \cos(\phi_k) \cos(2\pi k f t) - \sum_{k=1}^{M} \alpha_k \sin(\phi_k) \sin(2\pi k f t), \tag{5.4}$$

Substituting in (5.2) with the in-phase and quadrature representation and with the basis functions given in (5.3) and evaluating the integration assuming that $f = \frac{1}{T}$, we get

$$\begin{aligned} y_{Il} &= \frac{T}{2\pi} \sum_{k=1}^{M} \frac{\alpha_k \cos(\phi_k)}{k} \sum_{r=1}^{2l} (-1)^{r-1} \sin\left(\frac{(2r-1)\pi k}{2l}\right), \\ y_{Ql} &= \frac{-T}{2\pi} \sum_{k=1}^{M} \frac{\alpha_k \sin(\phi_k)}{k} \sum_{r=1}^{2l} (-1)^{r-1} \cos\left(\frac{2r\pi k}{2l}\right), \end{aligned} \tag{5.5}$$

where $l = 1, 2, \cdots, L$. The output of each in-phase and quadrature channel represents the sum of scaled in-phase and quadrature harmonics' amplitudes, respectively. In order to illustrate the simplicity in resolving back the harmonics' amplitudes, the $\boldsymbol{y}_I$ and $\boldsymbol{y}_Q$ can be represented in matrix format as

$$\underbrace{\begin{bmatrix} y_{I1} \\ y_{I2} \\ \vdots \\ y_{IL} \end{bmatrix}}_{\boldsymbol{y}_I} = \frac{T}{2\pi} \underbrace{\begin{bmatrix} 1 & 0 & -\frac{1}{3} & 0 & \frac{1}{5} & 0 & -\frac{1}{7} & 0 & \frac{1}{9} & 0 & \cdots \\ 0 & 1 & 0 & 0 & 0 & -\frac{1}{3} & 0 & 0 & 0 & \frac{1}{5} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & & \cdots & & 0 & 0 & 0 & 1 & \end{bmatrix}}_{U_I} \underbrace{\begin{bmatrix} \alpha_1 \cos(\phi_1) \\ \alpha_2 \cos(\phi_2) \\ \vdots \\ \alpha_M \cos(\phi_M) \end{bmatrix}}_{\boldsymbol{a}_I}, \tag{5.6}$$

and

$$\underbrace{\begin{bmatrix} y_{Q1} \\ y_{Q2} \\ \vdots \\ y_{QL} \end{bmatrix}}_{\boldsymbol{y}_Q} = \frac{-T}{2\pi} \underbrace{\begin{bmatrix} 1 & 0 & \frac{1}{3} & 0 & \frac{1}{5} & 0 & \frac{1}{7} & 0 & \frac{1}{9} & 0 & \cdots \\ 0 & 1 & 0 & 0 & 0 & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{5} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & & \cdots & & 0 & 0 & 0 & 1 & \end{bmatrix}}_{U_Q} \underbrace{\begin{bmatrix} \alpha_1 \sin(\phi_1) \\ \alpha_2 \sin(\phi_2) \\ \vdots \\ \alpha_M \sin(\phi_M) \end{bmatrix}}_{\boldsymbol{a}_Q}, \quad (5.7)$$

where U_I and U_Q are Unipotent matrices[2]. The U_I and U_Q are sparse matrices with most off-diagonal entries are zero. Given $\boldsymbol{y}_I$ and $\boldsymbol{y}_Q$, the in-phase and quadrature amplitude vectors $\boldsymbol{a}_I$ and $\boldsymbol{a}_Q$ can be obtained using

$$\boldsymbol{a}_I = U_I^{-1}\boldsymbol{y}_I, \qquad \boldsymbol{a}_Q = U_Q^{-1}\boldsymbol{y}_Q, \quad (5.8)$$

The sparsity of the Unipotent matrix reduces the computational power required for obtaining $\boldsymbol{a}_I$ and $\boldsymbol{a}_Q$[3], which is used to estimate the harmonics' amplitudes using

$$\alpha = \sqrt{a_I^2 + a_Q^2}. \quad (5.9)$$

where $l = 1, 2, \cdots, L$. Figure 5.3 illustrates the ATI approach for estimation of

[2]Unipotent matrices are those with all eignvalues = 1, and $(U - \mathcal{I})^n = 0$, where $\mathcal{I}$ is the identity matrix for $n > 1$

[3]The inverse of the Unipotent matrices is known and does not need to be calculated for each window.

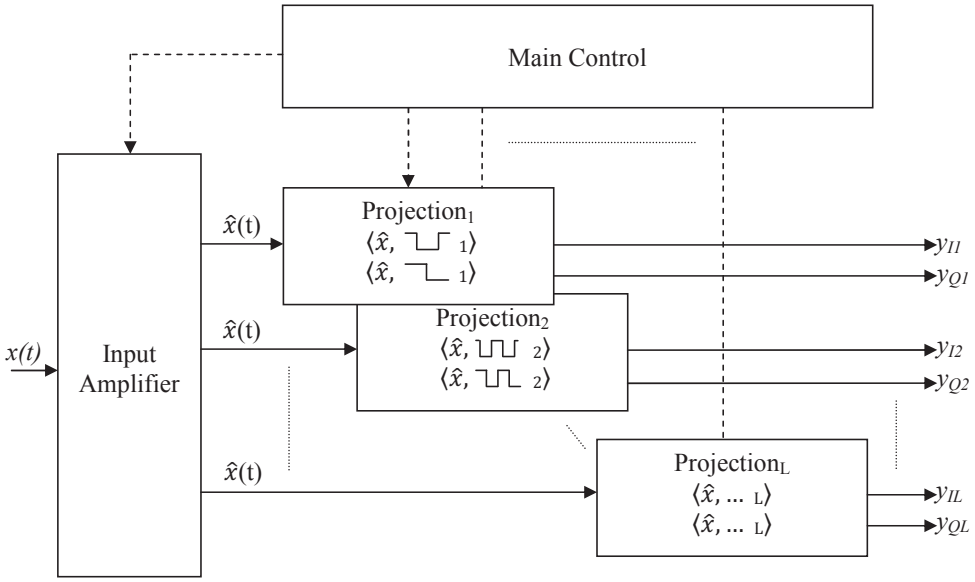

Figure 5.3: Analog-to-information front-end

the harmonics' amplitudes approximating the deterministic signature in the acquired analog signal.

5.3.3 Harmonic Parameters

The ATI approach assumes the acquired analog signal is band limited to $L \cdot f$, where $L \geq M$, M is the total number of harmonics in the original signal, and the fundamental of the input, f, equals to $\frac{1}{T}$, where T is the analysis window length. Projection over rectangular basis has one primary benefit: the fact that the analog hardware realization of these basis functions and the projection operation has the potential to be relatively simple and very power efficient.

A sample example is shown in Figure 5.4 for four harmonic signals with 10 harmonics and four different actual FFs (3.0 Hz, 2.9 Hz, 2.7 Hz and 2.4 Hz). The ATI

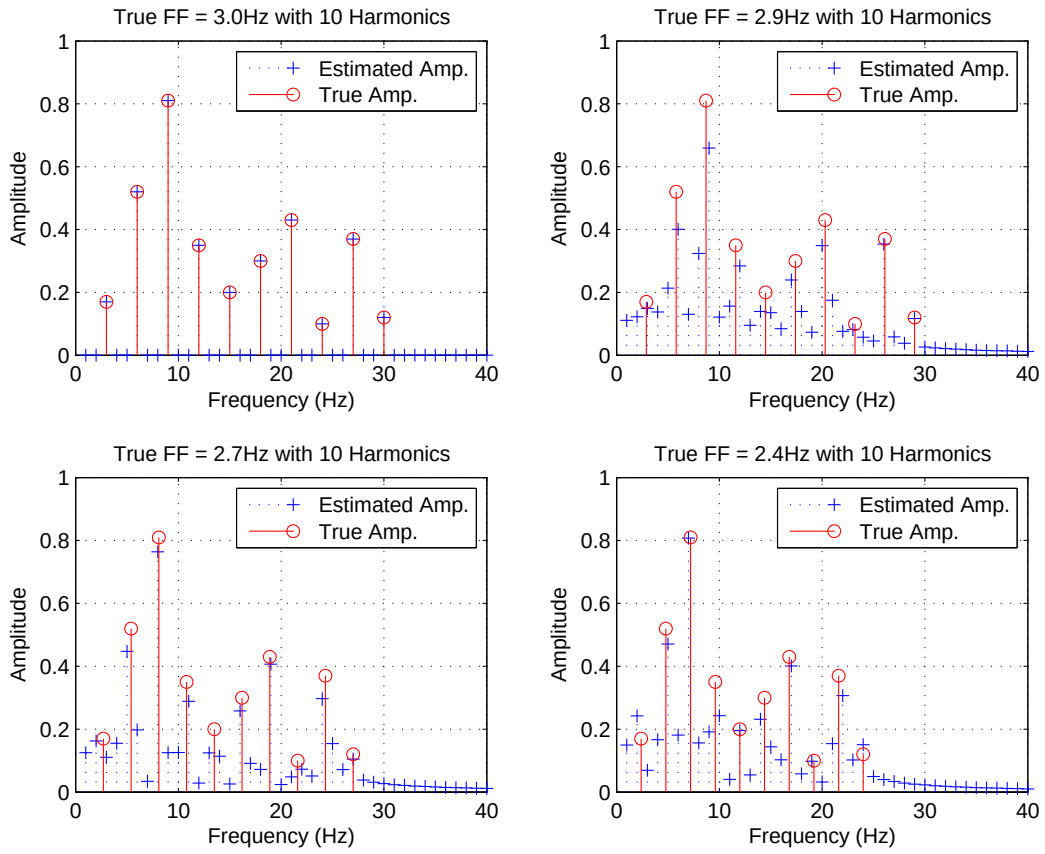

Figure 5.4: Analog-to-Information approach applied with different harmonic shifts.

approach uses an assumed FF of 1 Hz and 40 harmonics, which results in complete alignment between the actual harmonics and the proposed harmonic model in the first case and a harmonic shift that varies between $0 - 20\%$ in the other three cases depending on the true FF and the harmonic order. The Harmonics' estimates in Figure 5.4 shows that harmonics' amplitudes estimated using the ATI approach are very close to the original harmonics' amplitudes preserving the whole spectral envelope and the amplitudes of those harmonics that span the band with no energy are close to zero.

5.4 Case Study I: Classification of military vehicles

The acoustic data of 10 different vehicles were recorded by the BVP during verification experiments in 2000 [42,52], as previously discussed in Chapter 4. The ATI approach considers only the acoustic data since the harmonic structure is more apparent in the acoustic emission of vehicles. The original recording was triggered manually when the vehicle crosses the 200 m flag to acquire the whole run. In order to determine whether an event has occurred or not, the adaptive CFAR [53] outputs a decision every 0.5 second, based on the average energy level of the acoustic signal. The CFAR threshold is updated according to the noise variance of the detector output. Evaluating the performance of the CFAR detector is done by measuring any gaps that occurs in the detection process[4]. On average the implemented CFAR detector misses less than 7% of total number of detected events. The CFAR detector also detects events that do not have a clear signature, and those events and their presumably errant classifications are already included in the classification rates shown in the presented results. The total number of detected events per run depends on the vehicle itself and its speed. During normal speeds, medium and lightweight vehicles are detected within 50 m from the sensor station, while heavyweight tracked vehicles are detected beyond 100 m. Although the same number of runs were recorded for each vehicle, the variation in the detection range resulted in having the total number of detected events per

[4]A gap is defined as one or more events that are preceded and followed by at least 2 seconds or more of continuous detection.

vehicle range between $1200 - 5500$ events/vehicle. The total number of runs for the Volkswagen van were very small compared to the remaining military vehicles which resulted in having a very small number of generated FVs. An ANN was selected for classification. Since ANNs are very sensitive to the size of the training set and since we were trying to adapt a single network to differentiate among all vehicles, the Volkswagen van runs were dropped from our simulations. Averaging over the remaining nine vehicles, 96.64% of the energy of the acoustic emission for more than 32000 detected events[5] was found to be in the range $0 - 250$ Hz.

A three-layer FNN with sigmoid transfer function is utilized. The input layer represents the FV, the hidden layer consists of 40 hidden neurons, the output layer consists of 9 neurons (corresponding to 9 vehicles). The training (learning) of the network is according to the Resilient backpropagation algorithm (Rprop) developed from an analysis of the performance of the standard steepest descent algorithm [123]. The structure of the neural network is fixed for all evaluations. For all of our results, $\frac{1}{3}$ of the total number of detected events are randomly selected for network training and the remaining $\frac{2}{3}$ are used for testing.

The features extracted using the proposed ATI analog front-end approach utilizing three different harmonic models were evaluated for a FF $= 5$ Hz and L $= 25, 50$, and 100, approximating the deterministic signature of military vehicles in the band 125 Hz, 250 Hz, and 500 Hz, respectively. The detection, false alarm and classification rate results are shown in Table 5.1. Temporal decision fusion results are included as well in Table 5.1, where the decision is produced based on simple voting between individual decision on the same node for the same detected vehicle.

[5]Each event represents 0.5 seconds

Table 5.1: Detection, False alarm and classification rate - Military Vehicles

Number of harmonics		25	50	100
Vehicle ↓ FF →		5 Hz		
Detection rate (%)	Leopard 1	92.87	96.20	95.83
	Leopard 2	81.25	90.91	91.85
	Wiesel	77.42	82.95	86.62
	Jaguar	80.09	90.16	88.79
	M48	88.09	95.62	95.47
	Fuchs	81.89	88.79	87.66
	Hermelin	53.04	66.77	64.82
	Unimog	56.81	64.20	63.77
	MB1017	49.39	59.02	65.94
False alarm rate (%)	Leopard 1	1.33	0.89	1.09
	Leopard 2	5.15	2.48	2.69
	Wiesel	3.29	2.27	2.21
	Jaguar	3.87	2.30	2.14
	M48	2.08	0.63	0.55
	Fuchs	2.49	1.98	1.64
	Hermelin	1.58	1.26	0.89
	Unimog	1.04	0.56	0.59
	MB1017	2.10	1.54	1.70
Classification rate (%)		80.00	87.73	88.14
Decision Fusion (%)		87.60	89.38	91.53

From Table 5.1, we conclude that the ATI approach was capable of extracting distinctive features sufficient for vehicle discrimination without the need to estimate the engine FF or the exact number of harmonics. For the same spectral resolution (5 Hz), the capability to discriminate among military vehicles using harmonics' amplitudes extracted using the ATI approach will increase with more projections up to 250 Hz and the improvements from additional increase is not significant. We investigated the effect of changing the spectral resolution and found that having a smaller resolution (< 5 Hz) results in a slight improvement in discrimination but requires more projections to approximate the same band.

A quick comparison between the ATI approach (Table 5.1) and the TDHA algorithm (Table 4.5) discussed in Chapter 4 for the same spectral resolution and feature vector length shows a slight degradation in the average classification rate (from 90.38% for the TDHA to 87.73% for the ATI) with slightly more degradation in the lightweight vehicles. Although both algorithms employ the harmonics' amplitudes for classification, generating a harmonic structure based on the assumed FF ($\propto$ strongest harmonic component) enables the estimated amplitudes to be more effective than using a constant FF. However, the strength in the ATI approach lies mostly in the reduction in the power and area costs to implement the ATI approach compared to either the TDHA algorithm or FFT.

The sparsity in the Unipotent matrix and its inverse simplify the last step in resolving the harmonics' amplitudes from the projected vectors. However, the information needed for discrimination between vehicles is already present in the projected

vector (y_{Il}, and y_{Ql} where $l = 1, 2, ..., L$) and these vectors can be directly used as the characteristic feature vector for classification of military vehicles[6].

A hardware implementation of the proposed ATI approach introduces several sources of deviation from the formulation in (5.2). These may be combined into five classes of errors: timing, gain, offset, distortion, and random. For more details regarding the hardware realization and error modeling, please refer to [122]. These errors must be addressed in our simulations to evaluate the discrimination capability when the neural network is tested using features that are subjected to different errors. The average distortion effect can be estimated, and the generated features randomly selected for training of the ANN can be subjected to the estimated distortion. The three classes of errors that are crucial for the proposed analog approach are the gain, offset, and random noise. These errors will vary among the harmonic channels and can only be tuned within certain limits. Figure 5.5 addresses the effect of each of the gain, offset, and random noise in the classification process. Training of the ANN is performed using features subjected only to the estimated distortion function, while testing is performed using features subjected to the estimated distortion function, added noise, random gain, and random offset. The effect of the added offset in decreasing the average classification rate is more significant than the gain, and the decrease in the average classification rate for the same offset and gain is higher at lower SNR levels. The preliminary simulation results given analog hardware errors within certain levels of gain and offset at an SNR of 10 to 20 dB show we can achieve an average classification rate of 75% or higher. The average classification rate can always be increased by other means including temporal decision fusion.

[6]For a FF of 5 Hz, and 50 harmonics, the classification rate using the projected vector was 86.64% compared to 87.73% when resolving the harmonics' amplitudes (Table 5.1).

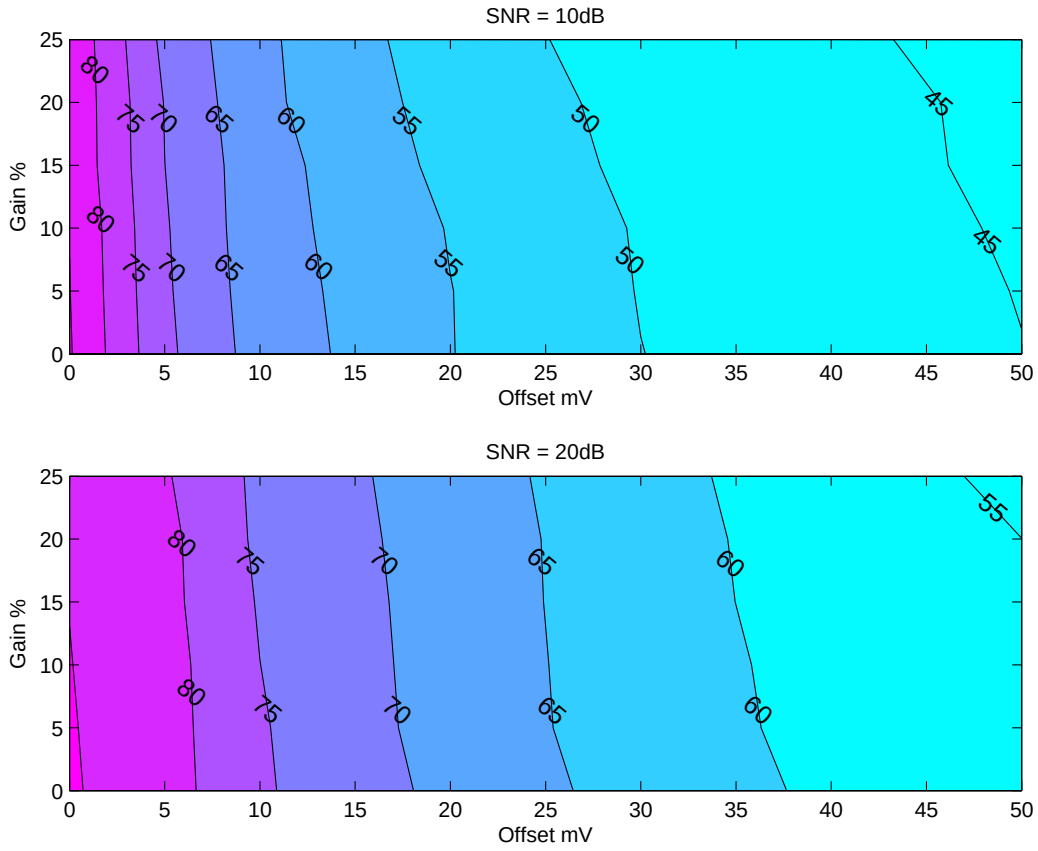

Figure 5.5: Average classification rate at Gain-Offset effect on at different SNR

5.5 Case Study II: Identification of bearing faults in rotating machinery

The major electric machines' faults are bearing, stator, broken rotor bar and end ring, and eccentricity-relates faults [26]. These faults may lead to increased vibration and noise levels. Fault detection can be realized by monitoring machine vibrations, acoustic emissions, or motor current. Machine vibrations are difficult to represent using a single harmonic model because the natural mechanical resonant frequencies of the machine are modulated by the defect frequency resulting in other spectral components that are not harmonics of either the defect frequency or the machine natural resonant frequencies [27]. It was shown in [66] that the amplitude of the defect frequency and its harmonics is an indication of the defect severity.

Instead of testing for each fault by itself, the ATI approach was employed to estimate the harmonics' amplitudes of the harmonic signal approximating the signature acquired from an accelerometer for an electric motor at different loads (0 to 3 horse power) and with different bearing faults. The accelerometer was mounted on the motor housing at the drive end of the motor. Single point faults were introduced to the test bearings with different fault diameters (0.007 inches, 0.014 inches, and 0.0021 inches) [28].

The vibration data was recorded at 12 KHz and since the information extracted should be sufficient for discriminating between normal operation, ball fault, inner race fault, and outer race fault, a single harmonic model with a FF = 100 Hz and 40 harmonics (approximating the deterministic signature in the $4K$ Hz band) is used in

the ATI approach to estimate the harmonics' amplitudes. A similar neural network to the one used in vehicle discrimination is used for identification of bearing faults. Table 5.2 shows the classification results for three cases, when the lower 10 or 20 amplitudes are used in discrimination as compared to the full 40 amplitudes. It is clear from Table 5.2 that the amplitudes of the lower 10 harmonics are sufficient for discriminating a normal bearing with no defects and a defective one, but it is not enough to identify the type of defect. If we need to identify what kind of defect is present, at least 20 harmonics are required to efficiently approximate motor vibrations. A harmonic model with more than 40 harmonics will not have any advantage since there is no vibration energy detected above 4 KHz.

Table 5.2: Detection, False alarm and classification rate - Bearing faults

Number of harmonics		10	20	40
Bearing Fault ↓ FF →		100 Hz		
Detection rate%	No fault	99.68	99.83	100.00
	Ball fault	83.80	94.34	98.79
	Inner race fault	86.70	96.04	98.93
	Outer race fault	76.73	94.90	98.49
False alarm rate%	No fault	0.10	0.03	0.00
	Ball fault	9.28	2.16	0.45
	Inner race fault	1.22	0.85	0.33
	Outer race fault	6.23	1.66	0.39
Classification rate (%)		87.20	96.42	99.09

When comparing the ATI approach results in Table 5.2 with those obtained using the ZC features in Table 3.3 we see that the ATI approach with 20 harmonics or more

is capable of achieving better classification results than utilizing ZC features, but it must be noted that the ATI approach generates a larger FV, which requires a larger neural network compared to the ZC algorithm previously discussed in Chapter 3.

Chapter 6

Dissertation Summary

6.1 Contributions

As mentioned in the introduction to this dissertation, the primary goals of the present work include developing simple and energy efficient feature extraction algorithms to enable low power sensor modules to reliably monitor harmonic sources. Specific contributions of this dissertation are listed as follows:

- The development of a simple, yet reliable, feature extraction algorithms that preserve the significant characteristics in the generated features for source verification.
- Evaluating the proposed approaches using different applications and actual recorded data sets using different sensor modules.

- Comparison of the proposed extracted features and spectral features for the same spectral resolution and applied on the same data set and using the same classifier.

- Applying a simple and robust detection algorithm for event driven applications.

- Applying several fusion algorithms that are very useful in reducing the effect of background noise and increasing the average detection and classification rate.

6.2 Potential Future Research Topics

The TDHA and ATI developed algorithms are very interesting in the sense that with minimum knowledge regarding the underlying incidents, we managed to extract a significant set of features. Extending the proposed algorithms to other applications brings many challenges to the development process, i.e.,

- Investigating the performance of the developed algorithms in more practical scenarios, for example a convoy of different vehicles. This might required developing the proposed algorithms to utilize the different transducers integrated on the same sensor module and those distributed among an array of sensors. The classifier might require some changes to accommodate the complexity of the classification algorithm and classification might be performed over several stages, for example, verifying the state of a convoy of vehicles followed by determining the type of the vehicles.

- Utilizing cueing theory to reduce the number of computations required to process data acquired by different transducers without losing the vital information carried in each modality.

- How can these features share in performing other tasks than classification, for example target counting, localization, and tracking, which will require developing an efficient array processing algorithm for low power sensor modules.

- Employing the extracted features for feedback control of the source may be very useful in remote applications, i.e., wind turbines, aircrafts, etc.

- Developing an automated monitoring system that utilizes time-folding techniques in increasing the recognition rate, i.e., building the confidence in the decision using the fact that the incident will last for a period of time.

Appendix

Table 1 shows a simple comparison between the estimated harmonics' amplitudes using the LS method when matrix inversion is replaced by division as a strong evidence that precisely defining the phase in the middle of the observation window as previously described in section 4.5 diagonalizes the information matrix.

Table 1: Sample harmonic model and its approximated structure

Original frequency components H(Hz)	Amplitudes	Harmonic structure for 50 components (strongest = 77.7659 Hz)	Estimated Harmonics' amplitudes using division	Estimated harmonics' amplitudes using matrix inversion
		4.8547	0.1040	0.1031
FF = 9.8	0.2	**9.7095**	0.3016	0.3005
		14.5642	0.1004	0.0995
2^{nd} = 19.6	0.02	**19.4190**	0.1254	0.1244
		24.2737	0.1112	0.1103
		29.1284	0.1219	0.1209
		33.9832	0.1443	0.1434
4^{th} = 39.2	0.5	**38.8379**	0.6134	0.6122
		43.6927	0.0883	0.0874
		48.5474	0.1246	0.1236
		53.4022	0.1582	0.1573

Original frequency components H(Hz)	Amplitudes	Harmonic structure for 50 components (strongest = 77.7659 Hz)	Estimated Harmonics' amplitudes using division	Estimated harmonics' amplitudes using matrix inversion
6^{th} = 58.8	0.22	**58.2569**	0.3811	0.3800
		63.1116	0.1867	0.1857
7^{th} = 68.6	0.22	**67.9664**	0.4577	0.4566
		72.8211	0.3946	0.3936
8^{th} **= 78.4**	3	**77.6759**	2.9356	2.9334
		82.5306	0.4524	0.4530
		87.3853	0.1682	0.1690
		92.2401	0.0754	0.0762
		97.0948	0.0225	0.0234
		101.9496	0.0239	0.0233
		106.8043	0.0751	0.0744
		111.6590	0.1753	0.1746
12^{th} = 117.6	1.1	**116.5138**	1.0189	1.0180
		121.3685	0.2651	0.2657
		126.2233	0.0723	0.0730
		131.0780	0.0388	0.0383
14^{th} = 137.2	0.7	**135.9328**	0.5603	0.5597
		140.7875	0.2657	0.2664
15^{th} = 147	0.001	**145.6422**	0.1235	0.1242
		150.4970	0.0766	0.0773
16^{th} = 156	0.01	**155.3517**	0.0402	0.0409
		160.2065	0.0233	0.0241
17^{th} = 166.6	0.13	**165.0612**	0.0825	0.0820
		169.9159	0.0618	0.0624
		174.7707	0.0143	0.0151
		179.6254	0.0429	0.0426
19^{th} = 186.2	0.34	**184.4802**	0.2593	0.2590
		189.3349	0.1479	0.1485

Original frequency components H(Hz)	Amplitudes	Harmonic structure for 50 components (strongest = 77.7659 Hz)	Estimated Harmonics' amplitudes using division	Estimated harmonics' amplitudes using matrix inversion
		194.1896	0.0324	0.0332
		199.0444	0.0500	0.0499
$21^{st} = 205.8$	0.56	**203.8991**	0.3718	0.3716
		208.7539	0.3225	0.3231
		213.6086	0.1413	0.1419
		218.4634	0.0909	0.0914
		223.3181	0.0588	0.0593
		228.1728	0.0226	0.0233
$24^{th} = 235.2$	0.3	**233.0276**	0.1383	0.1383
		237.8823	0.2367	0.2372
		242.7371	0.1220	0.1224

Table 2 shows the procedure executed for maintaining a CFAR during detection of an incident. The only assumption made about this procedure is that there should be no vehicles during the first window, where the initial threshold is determined based on the mean and standard deviation of background noise.

Table 2: CFAR detection

For the first window
    Compute the mean μ and the standard deviation σ.
    Initial $threshold = \mu + k \cdot \sigma$, where k is a constant.
For every other window
    Compute the normalized energy "κ.
    If $\kappa >$ $threshold$
        an incident is detected.
    else
        nothing is detected.
        update μ & σ & $threshold$

Table 3 lists the specific details of the constructed neural network used in all of our simulations for classification of characteristic feature vectors.

Table 3: ANN specification

Architecture	Multilayer perceptron - feedforward neural network
Number of hidden layers	1
# of neurons in the input layer	FV length
# of neurons in the hidden layer	FV length
# of neurons in the output layer	# of vehicles (9 Bochum & 3 SensIT)
Training function	Resilient Backpropagation
Transfer function	*tansigmoid (all neurons)*
Weight connect	Input-hidden & hidden-output
Bias connect	Hidden & output layer
Weight and bias initial values	Random initialization
Cost function	Mean Square Error
Training epochs	200

Bibliography

[1] A. Arora et. al., "A line in the sand: a wireless sensor network for target detection, classification, and tracking", *The International Journal of Computer and Telecommunications Networking*, Vol. 46(5), pp. 605-634, 2004.

[2] M.E. Hohil, J. Heberley, J. Chang and A. Rotolo, "Vehicle counting and classification algorithms for unattended ground sensors", *Proc. of SPIE in Unattended Ground Sensor Technologies and Applications V*, Vol. 5090, pp. 99-110, April 2003.

[3] D. Lake, "Efficient ML estimation for multiple and coupled harmonics", *Army Research Laboratory - Technical Report*, ARL-TR-2014, December 1999.

[4] A. Filipov, and N. Srour, "Distributed and disposable sensors at ARL and ASCTA", *IEEE Aerospace Conf.*, pp. 2065-71, 2004.

[5] N. Srour and J. Robertson, "Remote netted acoustic detection system: Final report", *Army Research Laboratory - Technical Report*, ARL-TR-706, May 1995.

[6] D. Puccinelli and M. Haenggi, "Wireless Sensor Networks: Applications and Challenges of Ubiquitous Sensing", *IEEE Circuits and Systems magazine*, Vol. 5(3), pp. 19-31, 2005.

[7] E.M. Yeatman, "Energy scavenging for wireless sensor nodes", *IEEE* 2^{nd} *International Workshop on Advances in Sensors and Interface (IWASI)*, pp. 1-4, june 2007.

[8] http://www.memsic.com (July 2011)

[9] J. H. Davies, "MSP430 Microcontroller Basics", Newnes - ELSIVIER, USA, September 2008.

[10] http://www.ti.com (July 2011)

[11] http://www.coolflux.com (July 2011)

[12] G.J. Pottie and W.J. Kaiser, "Wireless integrated network sensors", *Communications of the ACM*, Vol. 43(5), pp. 551-558, 2000.

[13] http://www.intel.com (July 2011)

[14] http://www.atmel.com (July 2011)

[15] http://www.btnode.ethz.ch (July 2011)

[16] http://www.eecs.berkeley.edu/ prabal/projects/epic/ (July 2011)

[17] A. Rowe, R. Mangharam, and R. Rajkumar, "Wireless Ad Hoc Networking: Personal-Area, Local-Area, and the Sensory-Area Networks", CRC Press, November 2006.

[18] http://www.xbow.jp/neokit-e.html (July 2011)

[19] S. T. S. Bukkapatnam, S. Mukkamala, J. Kunthong, V. Sarangan, and R. Komanduri, "Real-time monitoring of container stability loss using wireless vibration sensor tags", *IEEE Int. Conf. on Automation Science and Engineering - CASE*, pp. 221-226, August 2009.

[20] http://tinynode.com (July 2011)

[21] http://www.genetlab.com (July 2011)

[22] S. Corroy, J. Beiten, J. Ansari, H. Baldus, and P. Mhnen, "Energy efficient selection of computing elements in wireless sensor networks", *IEEE Proceedings of the* 2^{nd} *Int. Conf. on Sensor Technologies and Applications*, pp. 312-318, 2008.

[23] G.J. Pendock, L. Evans, and G. Coulson, "Wireless Sensor Module for Habitat Monitoring", *IEEE* 3^{rd} *Int. Conf. on Intelligent Sensors, Sensor Networks and Information (ISSNIP)*, pp. 699-702, December 2007.

[24] S. Kim, S. Pakzad, D. Culler, J. Demmel, G. Fenves, S. Glaser, and M. Turon, "Wireless sensor networks for structural health monitoring", *Proceedings of the* 4^{th} *Int. Conf. on Embedded networked sensor systems*, pp. 427-428, 2006.

[25] J. Pan and W. Tompkins, "A real time QRS detection algorithm", *IEEE Trans. On Biomedical Engineering*, Vol. 32 (3), pp. 230 - 236, March 1985.

[26] S. Nandi, H.A. Toliyat, and X. Li, "Condition monitoring and fault diagnosis of electrical motors - A review", *IEEE Trans. on Energy Conversion*, Vol. 20(4), pp. 719-729, December 2005.

[27] J.R. Stack, R.G. Harley, T.G. Habetler, "An amplitude modulation detector for fault diagnosis in rolling element bearings", *IEEE Trans. on Industrial Electronics*, Vol. 51(5), pp. 1097-1102, October 2004.

[28] X. Lou and K.A. Loparo, "Bearing fault diagnosis based on wavelet transform and fuzzy inference", *J. Mechanical Systems and Signal Processing*, Vol. 18(5) , pp. 1077-1095, September 2004.

[29] G. Prado, "Acoustic-seismic sensors: past experiences and future prospects", *SPIE Proc. in Unmanned/Unattended Sensors and Sensor Networks*, Vol. 5611, pp. 117-124, 2004.

[30] W. Dorn, "Tools of the Trade? Monitoring and Surveillance Technologies for UN Peacekeeping", *Peacekeeping Best Practices Unit, Department of Peacekeeping Operations (DPKO), United Nations*, New York, 2007. (Tabled with

and orally presented to the UN Special Committee on Peacekeeping Operations, 5 March 2007, as an independent commissioned study.)

[31] A.S. Palatnick and H.R. Inhelder, "Automatic vehicle identification systems-Methods of approach", *IEEE Trans. on Vehicular Technology*, Vol. 19(1), pp.128-136, February 1970.

[32] J.J. Reijmers, "On-Line Vehicle Classification", *IEEE Trans. on Vehicular Technology*, Vol. 29(2), pp.156-161, May 1980.

[33] N. Miyoshi, Y. Tachibana, T. Oda and Y. Imi, "Vehicle classification equipment for expressway toll collection station", *IEEE* 3^{rd} *Vehicular Technology Conf.*, Vol. 33, pp. 81-86, May 1983.

[34] J.L. Barker, "Radar, Acoustic, and Magnetic Vehicle Detectors", *IEEE Trans. on Vehicular Technology*, Vol. 19(1), pp.30-43, February 1970.

[35] N. Srour, "Unattended ground sensors a prospective for operational needs and requirements", *Army Research Laboratory - Technical Report Prepared for NATO*, October 1999.

[36] M.C. Wellman, N. Srour and D.B. Hillis, "Acoustic feature extraction for a neural network classifier", *Army Research Laboratory - Technical Report*, ARL-TR-1166, January 1997.

[37] R.B. Randall, "State of the art in monitoring rotating machinery - Part 1", *Sound and vibration magazine*, Vol. 38, pp. 10-16, March 2004.

[38] R.B. Randall, "State of the art in monitoring rotating machinery - Part 2", *Sound and vibration magazine*, Vol. 38, pp. 14-20, May 2004.

[39] T.R. Damarla and V. Mirelli, "Distributed Acoustic sensor data processing for target classification", *Proc. of SPIE in Unattended Ground, Sea and Air Sensor Technologies and Applications VIII*, Vol. 6231, 2006.

[40] M. Duarte and Yu-Hen Hu, "Vehicle classification in distributed sensor networks,", *Journal of Parallel and Distributed Computing*, Vol. 64(7), pp. 826-838, 2004.

[41] G.P. Mazarakis and J.N. Avaritsiotis, "Vehicle classification in Sensor Networks using time-domain signal processing and Neural Networks, *Journal of Microprocessors and Microsystems (ELSEVIER)*, Vol. 31(6), pp. 381-392, 2007.

[42] J. Altmann, "Acoustic and seismic signals of heavy military vehicles for cooperative verification", *Journal of Sound and Vibration (ELSEVIER)*, Vol. 273(4-5), pp. 713-740, June 2004.

[43] V. W. Porto, L. J. Fogel and D. B. Fogel, "Using evolutionary computation for seismic signal detection: A homeland security application", *IEEE Int. Conf. on Computational Intelligence for Homeland Security and Personal Safety*, pp.62-66, July 2004.

[44] L. Peck and J. Lacombe, "Seismic-Based Personnel Detection", *Annual IEEE 41st Int. Carnahan Conf. on Security Technology*, pp.169-175, October 2007.

[45] H.O. Park, A.A. Dibazar and T.W. Berger, "Protecting Military Perimeters from Approaching Human and Vehicle Using Biologically Realistic Dynamic Synapse Neural Network", *IEEE Conf. on Technologies for Homeland Security*, pp.73-78, May 2008.

[46] A.A. Dibazar, H.O. Park and T.W. Berger, "The Application of Dynamic Synapse Neural Networks on Footstep and Vehicle Recognition", *IEEE Int. Joint Conf. on Neural Networks (IJCNN)*, pp. 1842-1846, 2007.

[47] T.S. Anderson, M.L. Moran, S.M. Ketcham and J. Lacombe, "Technical validation of high-fidelity seismic signature simulations in support of FCS network ground sensors", *IEEE Proc. of User Group Conference*, pp. 62-67, 2003.

[48] D.A. Depireux, S. Varma, J. Baras, N. Srour and T. Pham, "Vehicle classification using acoustic data based on biology hearing model and multi-scale vector quantization", 4th *Annual Symposium, ARL Federal Laboratory*, pp.83-87, March 2000.

[49] A. Filipov and N. Srour, "Networked Micro-sensor research at ARL and the ASCTA ", *Sensors for Industry Conf. (Sicon)*, pp. 212-218, November 2002.

[50] W.M. Corcoll, N.C. Ellis and D.B. Simmons, "Evaluating three types of artificial neural networks for classifying vehicles with multisensory data", *Proc. of SPIE*

in Applications and Science of Artificial Neural Networks III, Vol. 3077, pp. 307-318, 1997.

[51] M. Brnnstrm, et. al., "Distributed data fusion in a ground sensor network", *Proc. of* 7^{th} *Int. Conf. on Information Fusion*, 2004.

[52] J. Altmann, S. Linev, A. Weisz, "Acoustic-seismic detection and classification of military vehicles-developing tools for disarmament and peace-keeping", *Journal of Applied Acoustics (ELSEVIER)*, Vol. 63(10), pp. 1087-1107, October 2002.

[53] P.E. William and M.W. Hoffman, "Classification of military ground vehicles using time domain harmonics' amplitudes", *IEEE Transaction on Instrumentation and Measurement*, in press (DOI:10.1109/TIM.2011.2135110).

[54] A. Giridhar and P.R. Kumar, "Toward a theory of In-Network computation in wireless sensor networks ", *IEEE Communications Magazine*, Vol. 4, pp. 98-107, April 2006.

[55] F. Pianegiani, M. Hu, A. Boni and D. Petri, "Energy-efficient signal classification in *Ad hoc* wireless sensor networks", *IEEE Trans. on Instrumentation and Measurements*, Vol. 57(1), pp. 190-196, January 2008.

[56] A.B. Doser, M.L. Yee, W.T. O'Rourke, M.E. Slinkard, D.C. Craft and H.D. Nguyen, "Distributed algorithms for small vehicle detection, classification, and velocity estimation using unattended ground sensors", *SPIE Proc. on Unat-*

tended Ground Sensor Technologies and Applications VII, Vol. 5796, pp. 374-385, February 2005.

[57] T.R. Damarla, T. Pham and D. Lake, "An algorithm for classifying multiple targets using acoustic signature", *Proc. of SPIE in Signal Processing, Sensor Fusion and Target Recognition XIII*, Vol. 5429, pp. 421-427, August 2004.

[58] N. Srour, "Back propagation of acoustic signature for robust target identification, *Proc. of SPIE in Enabling Technologies for Low Enforcement and Security*, Vol. 4232, pp. 399-408, 2001.

[59] J.C.R. Licklidder and I. Pollack, "Effects of Differentiation, Integration, and Infinite Peak Clipping upon the Intelligibility of Speech", *Journal of the Acoustical Society of America (JASA)*, Vol. 20(1), pp. 42-51, January 1948.

[60] E. Lupu, Z. Feher and P.G. Pop, "On the speaker verification using the TESPAR coding method Signals", *IEEE Int. Symposium on Circuits and Systems, ISCAS 2003*, Vol. 1, pp. 173-176, July 2003.

[61] G. P. Mazarakis and J.N. Avaritsiotis, "Lightweight Time Encoded Signal Processing for Vehicle Recognition in Sensor Networks", *Research in Microelectronics and Electronics 2006, Ph. D.*, pp. 497-500, June 2006.

[62] P.E. William, M.W. Hoffman, "Efficient sensor network vehicle classification using peak harmonics of acoustic emissions", *Proc. of SPIE in Unattended*

Ground, Sea, and Air Sensor Technologies and Applications X, Vol. 6963, March 2008.

[63] P.E. William, M.W. Hoffman, "Identification of battlefield vehicles using the zero crossings and the strongest harmonic component of the time domain acoustic signal", *Military Sensing Symposia (MSS), Battlespace Acoustic and Magnetic Sensors*, August 2008.

[64] F. Immovilli, M. Cocconcelli, A. Bellini, and R. Rubini, "Detection of generalized-roughness bearing fault by spectral kurtosis energy of vibration or current signals", *IEEE Trans. on Industrial Electronics*, Vol. 56(1), pp. 4710-4717, November 2009.

[65] B. Samanta and K.R. Al-Balushi, "Artificial neural network based fault diagnostics of rolling element bearings using time-domain features", *J. Mechanical Systems and Signal Processing*, Vol. 17(2) , pp. 317-328, March 2003.

[66] N. Gebraeel, M. Lawley, R. Liu and V. Parmeshwaran, "Residual life predictions from vibration based degradation signals: A neural network approach", *IEEE Trans. on Industrial Electronics*, Vol. 51(3), pp. 694-700, June 2004.

[67] Peter E. William and Michael W. Hoffman, "Identification of Bearing faults using time domain zero-crossings", *Journal of Mechanical Systems and Signal Processing*, Vol. 25(8), pp. 3078-3088, November 2011.

[68] D. Lake, "Harmonic phase coupling for battlefield acoustic target identification", *IEEE Intl. Conference in Acoustics, Speech, and Signal Processing, (ICASSP)*, Vol. 4, pp. 2049-52 May 1998.

[69] D. Lake, "Tracking fundamental frequency for synchronous mechanical diagnostic signal processing", *IEEE* 9^{th} *Signal Processing Workshop on Statistical Signal and Array Processing*, pp. 200-203, September 1998.

[70] X. Wang and H. Qi, "Acoustic target classification using distributed sensor arrays", *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, Vol. 4, 2002.

[71] S.S. Yang, Y.G. Kim and H. Choi, "Vehicle identification using wireless sensor networks", *IEEE Proc. of SoutheastCon.*, pp. 41-46, March 2007.

[72] D.A. Depireux, S. Varma, J. Baras, N. Srour and T. Pham, "Vehicle Classification Using Acoustic Data Based on Biology Hearing Models and Multi-scale Vector Quantization", *Technical Report - Federated Lab00, College Park, MD*, 2000.

[73] D. Li, K. Wang and Y.H. Hu, A.M. Sayeed, "Detection, classification, and tracking of targets", *IEEE Signal Processing Magazine*, Vol. 19(2), pp. 17-29, March 2002.

[74] R.H. Mgaya, S.Z. Sabatto, A. Shirkhodaie and C. Wei, "Vehicle identifications using acoustic sensing", *IEEE Proc. of SoutheastCon.*, pp. 555-560, March 2007.

[75] J.A. Robertson and B. Weber, "Artificial Neural Networks for Acoustic Target recognition", *Joint Report between ARL and ITT research*, 1993.

[76] B. Guo, M. S. Nixon and T. R. Damarla, "Acoustic Information Fusion for Ground Vehicle Classification", *Proceedings of the* 11^{th} *Int. Conf. on Information Fusion*, 2008.

[77] G. Succi and T.K. Pedersen, "Acoustic target tracking and target identification - recent results", *SPIE Proc. on Unattended Ground Sensor Technologies and Applications*, Vol. 3713, pp. 10-21, April 1999.

[78] P. Gaunard, C.G. Mubikangiey, C. Couvreur and V. Fontaine, "Automatic classification of environmental noise events by hidden Markov models", *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Vol. 6, pp. 3609-3612, May 1998.

[79] B. Malhotra, I. Nikolaidis and M.A. Nascimento, "Distributed and efficient classifiers for wireless audio-sensor networks", *Journal of Computer Networks (Elsevier)*, Vol. 52(13), pp. 2582-93, 2008.

[80] H. Wu and J.M. Mendel, "Classification of battlefield ground vehicles using acoustic features and fuzzy logic rule-based classifiers", *IEEE Trans. on Fuzzy Systems*, Vol. 15(1), pp. 56-72, February 2007.

[81] H. Wu and J.M. Mendel, "Multi-category classification of ground vehicles based on their acoustic emissions", *SPIE Proc. in Unattended/Unmanned Ground,*

Ocean, and Air Sensor Technologies and Applications VI, Vol. 5417, pp. 31-42, 2004.

[82] K. Sayood, "Introduction to data compression", second edition, Morgan Kaufmann: San Francisco, CA, USA, 2000.

[83] Y. Linde, A. Buzo, and R.M. Gray, "An algorithm for vector quantization design", *IEEE Transactions on Communications*, Vol. 28, pp. 84-94, January 1980.

[84] Z. Kamal, A. Gupta, A.A. Khokhar, and L. Lilien, "Power consumption analysis of maximum a posterior classification using LU decomposition and Jacobi iterations", *IEEE Trans. on Parallel and Distributed systems*, Vol. 19(11), pp. 1473-83 , November 2008.

[85] B. Li, Mo-Yuen Chow, Y. Tipsuwan and J.C. Hung, "Neural network based motor rolling bearing fault diagnosis", *IEEE Trans. on Industrial Electronics*, Vol. 47(5), pp. 1060-1069, October 2000.

[86] S. Szyszko and P.A. Payne, "Artificial neural networks for feature extraction from acoustic emission signals", *IEE Colloquium on Measurements, Modeling and Imaging for Non-Destructive Testing*, pp. 1-6, 27 March 1991.

[87] A.K. Jain, M. Jianchang, and K.M. Mohiuddin, "Artificial neural networks: a tutorial", *IEEE Computer Society*, Vol. 29(3), pp. 31-44, March 1996.

[88] P. Ramanathan, "Location-centric approach for collaborative target detection, classification, and tracking", *IEEE Proc. of CAS Workshop on Wireless Communication and Networking (Invited)*, September 2002.

[89] J. Llinas, E. Waltz, Multi-sensor Data Fusion. Boston, MA: Artech House, 1990.

[90] S.G. Iyengar, P.K. Varshney and T. Damarla, "On the detection of footsteps based on acoustic and seismic sensing", IEEE 41^{st} Asilomar Conf. on Signals, Systems and Computers (ACSSC), 2248-52 (2007).

[91] R.R. Brooks, P. Ramanathan and A.M. Sayeed "Distributed target classification and tracking in sensor networks", *Proceedings of the IEEE*, Vol. 91(8), pp. 1163-1171, 2003.

[92] P.E. William and M.W. Hoffman, "Acoustic classification of battlefield vehicles based on their seismic detection", 156^{TH} *Meeting of the Acoustical Society of America (ASA)*, Volume 124(4), pp. 2508-2508, November 2008.

[93] C. Meesookho, U. Mitra and S. Narayanan, "On Energy-Based Acoustic Source Localization for Sensor Networks", *IEEE Trans. on Signal Processing*, Vol. 56(1), pp. 365-377, January 2008.

[94] X. Sheng and Y.H. Hu, "Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks", *IEEE Tran. on Signal Processing*, Vol. 53(1), pp. 44-53, 2005.

[95] D. Li and Y.H. Hu, "Energy-based collaborative source localization using acoustic micro sensor array", *Journal of Applied Signal Processing (EURASIP)*, Vol. 4, pp. 321-327, 2003.

[96] X. Sheng and Y.H. Hu, "Sub-band energy based collaborative target localization in wireless sensor network system," *The First ACM Conference on Embedded Networked Sensor Systems*, November 2003.

[97] R.J. Niederjohn, "A Mathematical formulation and comparison of zero-crossing analysis techniques which have been applied to automatic speech recognition", *IEEE Trans. on Acoustics, Speech, and Signal Processing*, Vol. ASSP-23(4), pp. 373-380 August 1975.

[98] M. Nakagawa, H. Tsai, B. He, "3D-EEG rhythm mapping during flash stimulation using zero-crossings analysis", *Proc. of the* 19^{th} *Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society*, Vol. 4, pp. 1536-1539, October 1997.

[99] J.C.R. Licklider and I. Pollack, "Effects of differentiation, integration, and infinite peak clipping upon the intelligibility of speech", *J. Acoust. Soc. America*, Vol. 20, pp.42-51, 1948.

[100] M.R. Ito, "Relationship between zero-crossing measurements for speech analysis and recognition", *J. Acoust. Soc. Amer.*, Vol. 51, pp. 2061-2062, 1972.

[101] R.A. King and T.C. Phipps, "Shannon, TESPAR and approximation strategies", *Computers and Security*, Vol. 18(5), pp. 445-453, 1999.

[102] E. Lupu, Z. Feher, and P.G. Pop, "On the Speaker verification using the TESPAR coding method", *International Symposium on Signals, Circuits and Systems*, Vol. 1, pp. 173-176, 2003.

[103] I.B. Thomas and R.J. Niederjohn, "A preliminary analysis technique for speech classification", *Proc. Nat. Electron. Conf.*, Vol. 25, pp. 685-690, 1969.

[104] A.V. Oppenheim, R.W. Scafer, and J.R. Buck, "Discrete-Time Signal Processing", 2^{nd} edition, Prentice Hall, NY, January 1999.

[105] H. Ocak and K.A. Loparo, "Estimation of the running speed and bearing defect frequencies of an induction motor from vibration data", *J. Mechanical Systems and Signal Processing*, Vol. 18(3) , pp. 515-533, May 2004.

[106] J.E. Berry, "How to track rolling element bearing health with vibration signature analysis", *Sound and Vibration*, Vol. 25(11), pp. 24-35, November 1991.

[107] A. Malhi and R.X. Gao, "PCA-based feature selection scheme for machine defect classification", *IEEE Trans. on Instrumentation and Measurement*, Vol. 53(6), pp. 1517-1525, December 2004.

[108] UN department of peacekeeping operations, "United Nations Peacekeeping Operations: Principles and Guidelines", *Department of Peacekeeping Operations, United Nations Secretariat*, One UN Plaza, New York, NY 10017, January 2008.

[109] http://www.un.org/en/peacekeeping/ (July 2011)

[110] UN department of peacekeeping operations, "Background note on current peacekeeping operations", *The peace and security section of the UN department of public information, United Nations Secretariat*, One UN Plaza, New York, NY 10017, July 2009.

[111] UN department of peacekeeping operations, "United nations peace operations: Year in review 2008", *Peace and Security Section of the United Nations Department of Public Information*, February 2009.

[112] Government Accountability Organization (GAO), "United Nations Peacekeeping: Challenges obtaining needed resources could limit further large deployments and should be addressed in U.S. reports to Congress", *A report to the committee on foreign relations, U.S. Senate*, December 2008.

[113] S. Linev, J. Altmann, A. Wei, "Monitoring for Verification", *Acoustic-Seismic Military-Vehicle Measurements of 2000 at Meppen, Germany*, Research Reports, Vol. 11, Lenzen: Grnberg, 2001.

[114] W. Merrill, K. Sohrabi, L. Girod, J. Elson, F. Newberg, and W. Kaiser, "Open Standard Development Platforms for Distributed Sensor Networks", *Proceed-*

ings of SPIE - Unattended Ground Sensor Technologies and Applications IV, Vol. 4743, pp. 327-337, 2002.

[115] www.mathworks.com (July 2011)

[116] J. Altmann, "Acoustic-seismic detection of ballistic-missile launches for cooperative early warning of nuclear attack", *J. of Science and global security*, Vol. 13, pp.129-168, 2005.

[117] D. L. Donoho, "Compressed sensing", *IEEE Trans. on Information Theory*, Vol. 52(4), pp. 1289-1306, April 2006.

[118] F. Immovilli, M. Cocconcelli, A. Bellini and R. Rubini, "Detection of generalized roughness bearing fault by spectral kurtosis energy of vibration or current signals", *IEEE Trans. on Industrial Electronics*, Vol. 56(11), pp. 4710-4717, November 2009.

[119] E. Stern, "Comparison of New Analog Device Technologies for Signal Processing", *Ultrasonics Symposium*, pp. 129-136, 1983.

[120] R. Sarpeshkar, "Analog Versus Digital: Extrapolating from Electronics to Neurobiology", *Journal of Neural Computation*, Vol. 10(7), pp. 1601-1638, 1998.

[121] D. Healy and D.J. Brady, "Compression at the physical interface", *IEEE Signal Processing magazine*, Vol. 25(2), pp. 67-71, March 2008.

[122] D. White, P.E. William, M.W. Hoffman, and S. Balkır, "Analog sensing front end for harmonic signal classification applications", *IEEE Sensors Journal*, to be submitted.

[123] M. Riedmiller, H. Braun, "A direct adaptive method for faster backpropagation learning: the RPROP algorithm", *Proc. IEEE Int. Joint Conf. on Neural Networks, USA*, Vol. 1, pp. 586-591, 1993.