

UNIVERSITY OF CALIFORNIA

Los Angeles

**Signal Classification and Identification  
for  
Wireless Integrated Networked Sensors**

A dissertation submitted in partial satisfaction  
of the requirements for the degree  
Doctor of Philosophy in Electrical Engineering

by

Sivatharan Natkunanathan

2004

© Copyright by

Sivatharan Natkunanathan

2004

The dissertation of Sivatharan Natkunanathan is approved.

---

Mario Gerla

---

Bahram Jalali

---

William J. Kaiser

---

Kung Yao

---

Gregory J. Pottie, Committee Chair

University of California, Los Angeles

2004

*To my family.*

# TABLE OF CONTENTS

## CHAPTER 1

<b>INTRODUCTION</b> .....	<b>1</b>
1.1 WINS BACKGROUND .....	1
1.2 SIGNAL CLASSIFICATION AND IDENTIFICATION .....	5
1.2.1 <i>SSE Tree Building</i> .....	5
1.2.2 <i>SSE Training</i> .....	6
1.2.3 <i>Signal Detection and Confidence Measure Assignment</i> .....	7
1.2.4 <i>Signal State-Space Decomposition</i> .....	8
1.2.5 <i>Decision Fusion</i> .....	9

## CHAPTER 2

<b>SYSTEM LEVEL DESCRIPTION OF WINS</b> .....	<b>11</b>
2.1 APPLICATIONS OF WINS .....	11
2.2 PATH LOSS AND FADING ISSUES IN RF .....	13
2.2.1 <i>Effect of Communication vs. Battery Life</i> .....	15
2.2.2 <i>Maximum Bandwidth vs. Transmitted Power</i> .....	16
2.3 PROPAGATION MODELS - FADING EFFECTS ON RF SIGNALS FROM SENSOR NODES .....	28
2.3.1 <i>Communication Energy vs. Range</i> .....	31
2.4 DATA WAREHOUSING AND DECISION SUPPORT SYSTEMS .....	33
2.4.1 <i>Vehicle Monitoring Applications and Database Integration</i> .....	35
2.4.2 <i>Hierarchy Specifications found in WINS Data</i> .....	38
2.4.3 <i>Data Mining the WINS Data Warehouse to extract a Training Set</i> .....	43
2.5 DECISION SUPPORT SYSTEM OVERVIEW .....	47
2.6 DISTRIBUTED VS. CENTRALIZED SIGNAL PROCESSING / COLLABORATIVE DECISION MAKING .....	48
2.7 SIGNAL CLASSIFICATION / IDENTIFICATION ARCHITECTURES .....	48
2.8 SYSTEM LEVEL ARCHITECTURAL ALGORITHM DESCRIPTION .....	49
2.9 SIGNAL SEARCH ENGINE .....	51
2.10 DYNAMIC SYSTEM AND RECONFIGURABILITY .....	52
2.11 MULTI-SIGNAL SEARCH ENGINE (M-SSE) .....	53
2.12 MULTI-SIGNAL DETECTION AND SIGNAL EVOLUTION .....	54
2.13 SSE IMPLEMENTATION .....	55
2.14 SUMMARY .....	56

## CHAPTER 3

<b>DISTRIBUTED VS. CENTRALIZED SIGNAL PROCESSING AND DECISION MAKING</b> .....	<b>58</b>
3.1 SSE BACKGROUND .....	58
3.2 DISTRIBUTED (LOCALIZED) VS. CENTRALIZED SIGNAL PROCESSING FOR SSE .....	59
3.3 SSE APPLICATION ARCHITECTURES .....	67
3.4 COMMUNICATION VS. COMPUTATION COSTS .....	70
3.4.1 <i>Communication Cost</i> .....	70
3.4.2 <i>Comparison of SSE / M-SSE to Communication</i> .....	73
3.4.3 <i>Accuracy of Distributed vs. Centralized SSE &amp; M-SSE</i> .....	75

3.5	DECISION MAKING – (LOCALIZED VS. CENTRALIZED).....	79
3.5.1	<i>Decision Making (Localized)</i> .....	80
3.5.2	DECISION MAKING (DISTRIBUTED) .....	81
3.5.3	<i>Weighting Methods for Segmentation</i> .....	82
3.5.4	<i>Weighting Method for Different Classes of Signals</i> .....	86
3.5.5	<i>Lumped Weighting Method with Combined Class and Segmentation Based Decision</i> .....	93
3.6	<i>Sensor SCIA Control Criteria</i> .....	101
3.7	SUMMARY.....	102

## CHAPTER 4

<b>INVESTIGATION OF SSE TO SIGNAL VARIABLES .....</b>		<b>103</b>
4.1	TEST SIGNAL SET .....	104
4.2	CHANNEL AND ENVIRONMENTAL EFFECTS ON SIGNALS .....	113
4.2.1	<i>Acoustic Signals</i> .....	114
4.2.2	<i>Seismic Signals</i> .....	114
4.2.3	<i>Infrared Signals</i> .....	115
4.2.4	<i>All Signals</i> .....	116
4.3	CLASSIFICATION / IDENTIFICATION OF TEST SIGNAL SET .....	117
4.3.1	<i>Template Selection for Classification / Identification</i> .....	118
4.3.2	<i>Classification / Identification on SSE</i> .....	119
4.4	RESULTS ON TEST RUNS.....	120
4.4.1	<i>Identification</i> .....	120
4.4.2	<i>SSE Runs on Test Signals</i> .....	121
4.4.3	<i>SSE Run Results on Signal Sets</i> .....	133
4.5	<i>Comparison of ‘Music’ / ‘Pisarenko’ Method for Identification / Classification</i> .....	133
4.5.1	<i>Noise Variations</i> .....	135
4.5.2	<i>Amplitude Variation Effects</i> .....	136
4.5.3	PHASE VARIATION EFFECTS.....	138
4.5.4	<i>Multiple Source Identification</i> .....	140
4.5.5	<i>Summary of finding of ‘MUSIC’ and Pisarenko Identification</i> .....	140
4.6	SUMMARY.....	142

## CHAPTER 5

<b>CLASSIFICATION &amp; IDENTIFICATION ALGORITHM FOR SSE .....</b>		<b>144</b>
5.1	INTRODUCTION .....	144
5.2	MOVING SOURCE VARIABLES.....	148
5.2.1	<i>Doppler Shifts</i> .....	149
5.2.2	<i>ABRUPT FREQUENCY LOSS / GAIN</i> .....	153
5.2.3	<i>Directional Change in Path of Travel</i> .....	155
5.2.4	<i>Glitches and Saturation of Sensor Readings</i> .....	159
5.2.5	<i>Methods used to Mitigate Variables and Generalizations</i> .....	161
5.3	TIME-FREQUENCY OBSERVATIONS.....	161
5.4	TIME DOMAIN PROCESSING.....	163
5.4.1	<i>Signal Segmentation</i> .....	163
5.4.2	<i>Methodology for Signal Segmentation</i> .....	165
5.5	BUILDING THE TEMPLATE TREE .....	167
5.5.1	<i>Classification</i> .....	167
5.5.2	<i>Identification</i> .....	170

5.6	TRAINING THE SSE WITH TEMPLATE SELECTION .....	172
5.6.1	<i>Window Size Selection</i> .....	172
5.6.2	<i>Minimum Signal Length</i> .....	174
5.6.3	<i>Time Domain RMS Algorithm</i> .....	175
5.6.4	<i>Probability of Detection or False Alarm</i> .....	176
5.6.5	<i>State Space Approach</i> .....	180
5.6.6	<i>Detection in M-SSE</i> .....	182
5.7	SUMMARY .....	183

## CHAPTER 6

<b>SSE / M-SSE RESULTS ON REAL WORLD SIGNALS .....</b>		<b>184</b>
6.1	BACKGROUND .....	184
6.2	REAL WORLD SIGNAL DATABASE .....	185
6.3	SSE EVALUATION .....	188
6.4	SSE / M-SSE IMPLEMENTATION .....	194
6.5	WAVELET METHOD CLASSIFICATION .....	205
6.6	WAVELET DSP HARDWARE IMPLEMENTATION FEASIBILITY .....	206
6.7	SUMMARY .....	209

## CHAPTER 7

<b>CONCLUSION AND FUTURE WORK .....</b>		<b>210</b>
7.1	CONCLUSION .....	210
7.2	IMPLEMENTATION: HARDWARE SOFTWARE CO-DESIGN .....	213
7.3	FUTURE RESEARCH .....	214
<b>BIBLIOGRAPHY .....</b>		<b>216</b>

# LIST OF FIGURES

## CHAPTER 1

FIGURE 1. 1 WINS DEPLOYMENT FOR EVENT DETECTION.....	2
FIGURE 1. 2 WINS ARCHITECTURE WITH FUNCTIONAL MODULES. ....	4

## CHAPTER 2

FIGURE 2. 1 RECEIVED SIGNAL POWER VS. SEPARATION DISTANCE FOR SHORT/MEDIUM RANGE. ....	19
FIGURE 2. 2 RECEIVED SIGNAL POWER VS DISTANCE FOR LONG RANGE. ....	22
FIGURE 2. 3 MAXIMUM RECEIVER BANDWIDTH VS. TRANSMITTED POWER. ....	24
FIGURE 2. 4 MAXIMUM SEPARATION DISTANCE VS. TRANSMITTED POWER (FREE SPACE =10M). ....	25
FIGURE 2. 5 MAXIMUM RECEIVER BANDWIDTH VS TRANSMITTED POWER (FREE SPACE = 100M).....	26
FIGURE 2. 6 MAXIMUM SEPARATION DISTANCE VS. TRANSMITTED POWER (FREE SPACE = 100M).....	27
FIGURE 2. 7 ENERGY CONSUMPTION VS. TRANSMITTED DISTANCE (CHANNEL DEPENDENT $1/R^2$ ).....	32
FIGURE 2. 8 ENERGY CONSUMPTION VS. TRANSMITTED DISTANCE (CHANNEL DEPENDENT $1/R^4$ ).....	33
FIGURE 2. 9 SSE USAGE IN A DECISION SUPPORT AND DATA MINING WINS APPLICATION.....	34
FIGURE 2. 10 WINS DATA WAREHOUSING FOR SSE BASED DECISION SUPPORT SYSTEMS. ....	37
FIGURE 2. 11 AN EXAMPLE OF SET GROUPING OF SPEEDS WITH WINS DATA. ....	41
FIGURE 2. 12 M-SSE SYSTEM LEVEL BLOCKS AND MODULES. ....	50

## CHAPTER 3

FIGURE 3. 1 SYSTEM DIAGRAM OF SSE RELATIVE TO SENSOR NODE TRANSCIEVER. ....	60
FIGURE 3. 2. DISTRIBUTED LOCAL SIGNAL PROCESSING AND SCIA FOR WINS.....	61
FIGURE 3. 3. LOCALIZED SIGNAL PRE-PROCESSING AND LOCALIZED / CENTRALIZED SCIA FOR WINS.....	62
FIGURE 3. 4. SEGMENTATION BASED LOCALIZED SCIA ARCHITECTURE. ....	64
FIGURE 3. 5 DISTRIBUTED/LOCALIZED SCIA AND SENSOR TYPE BASED DECISION FUSION AT CLUSTER HEAD DATA. ....	65
FIGURE 3. 6. CONSUMER, AND INDUSTRIAL APPLICATIONS USING DISTRIBUTED / CENTRALIZED PROCESSING. ....	69
FIGURE 3. 7. CLASS BASED PROBABILITY ASSIGNMENT PROCEDURE. ....	78
FIGURE 3. 8. COLLABORATIVE 'DECISION MAKING' AT CLUSTER HEADS OR CENTRALIZED REMOTE PROCESSOR. ....	80

## CHAPTER 4

FIGURE 4. 1 A SAMPLE LOW FREQUENCY SIGNAL WINDOW, WITH AN SNR OF 0DB.....	106
FIGURE 4. 2 A HIGH FREQUENCY SIGNAL WITH A SNR OF 0 dB.....	109
FIGURE 4. 3 SIGNAL WITH CONTENTS FROM MULTIPLE ADJOINING FREQUENCY BANDS.....	110
FIGURE 4. 4 DISTINCT SIGNAL FEATURES SELECTED AS A CORRELATOR TEMPLATE.....	112
FIGURE 4. 5 HIGH FREQUENCY (180 Hz) PHASE OFFSET BEHAVIOR OF SCIA. ....	123
FIGURE 4. 6 LOW FREQUENCY (20 Hz) PHASE OFFSET BEHAVIOR OF SCIA. ....	124
FIGURE 4. 7 SSE IDENTIFICATION WITH A 64 SAMPLE LENGTH CORRELATOR. ....	125
FIGURE 4. 8 SSE IDENTIFICATION WITH A 32 SAMPLE LENGTH .....	126

FIGURE 4. 9 SSE PERFORMANCE WITH HIGH SNR ( 20dB ) SIGNALS. ....	128
FIGURE 4. 10 SSE PERFORMANCE WITH LOW SNR ( 0 dB ) SIGNALS. ....	129
FIGURE 4. 11 WINDOW STEPPING SAMPLE EFFECT ON RMS DURING SIGNAL CORRELATION. ....	131
FIGURE 4. 12 SSE PERFORMANCE WITH AN OPTIMIZED STEPPING FOR MINIMUM POWER COMPUTATION. ....	132
FIGURE 4. 13 RESULTS FROM RUNS WITH NOISE ON WELL-SEPARATED AND CLOSE FREQUENCY CONTENTS. ....	135
FIGURE 4. 14 ANALYSIS OF SIGNAL STRENGTH ON CLOSELY SPACED AND WELL SEPARATED SIGNAL SETS. .	138
FIGURE 4. 15 EFFECT OF PHASE DISTORTION ON SIGNAL SET. ....	139

## CHAPTER 5

FIGURE 5. 1. STATE-SPACE APPROACH TO SSE EVENT IDENTIFICATION / CLASSIFICATION. ....	146
FIGURE 5. 2. MOVING SOURCE WAVEFORM WITH STATE-SPACE DECOMPOSITION. ....	149
FIGURE 5. 3. TIME-FREQUENCY BEHAVIOR SHOWING DOPPLER EFFECTS ON HIGH-SPEED (30 KM/H) SOURCE RAW DATA. ....	151
FIGURE 5. 4. TIME-FREQUENCY BEHAVIOR SHOWING MINIMAL DOPPLER EFFECTS ON LOW-SPEED (15KM/H) SOURCE DATA. ....	152
FIGURE 5. 5. ABRUPT FREQUENCY LOSS / GAIN AS SEEN BY AN ACOUSTIC MICROPHONE IN A MOVING SOURCE. ....	154
FIGURE 5. 6. DOPPLER FROM SAME SOURCE MOVING IN DIFFERENT DIRRECTIONS. ....	156
FIGURE 5. 7. SHOWS AN ACOUSTIC MICROPHONE DETECTED TIME-FREQUENCY INFORMATION. ....	157
FIGURE 5. 8. SEISMIC GEOPHONE DETECTED TIME-FREQUENCY INFORMATION FOR THE SAME RUN AS ABOVE. ....	158
FIGURE 5. 9. SATURATED SIGNAL READINGS ARE AVOIDED DURING SIGNAL PRE-PROCESSING. ....	160
FIGURE 5. 10. SIGNAL PRE-PROCESSING ASSOCIATED WITH STATE SPACE DECOMPOSITION. ....	164
FIGURE 5. 11. CORRELATOR TEMPLATE SELECTION FOR TIER0, TIER1, TIER2. ....	168
FIGURE 5. 12. TIER STRUCTURE FOR CLASS BASED CLASSIFICATION AND IDENTIFICATION. ....	171
FIGURE 5. 13. 2D AND 3D DECISION BOUNDARIES FORMED DURING MULTI-DIMENSIONAL STATE-SPACE CLASSIFICATION AND IDENTIFICATION. ....	178
FIGURE 5. 14. STATE TYPE, SPACE, SUB-SPACE BASED DECOMPOSITION OF ACQUIRED SIGNALS. ....	181

## CHAPTER 6

FIGURE 6. 1. DATABASE CONTAINING REAL WORLD SENSOR READINGS PROVIDED BY ARL. ....	186
FIGURE 6. 2. SSE ACOUSTIC SIGNAL CLASSIFICATION RESULTS ARE SHOWN FOR APPROACH STATE IN DESERT TERRAIN. ....	189
FIGURE 6. 3. SSE ACOUSTIC SIGNAL CLASSIFICATION RESULTS ARE SHOWN FOR THE DEPARTURE STATE IN DESERT TERRAIN. ....	190
FIGURE 6. 4. ARCTIC TERRAIN SSE EVALUATION FOR AN ACOUSTIC DATA SET. ....	191
FIGURE 6. 5. NORMAL TERRAIN EVALUATION OF SSE FOR AN ACOUSTIC DATA SET. ....	192
FIGURE 6. 6. TIER STRUCTURE OF CLASS DIVISIONS AND BRANCHING BASED ON SIGNAL DATABASE. ....	195
FIGURE 6. 7. HIERARCHICAL SEARCH (TYPE ABSTRACTION HIERARCHY)-STEP 0. ....	196
FIGURE 6. 8. SSE RESULTS AFTER COLLABORATIVE DECISION FUSION. ....	197
FIGURE 6. 9. SSE RUNS TO CLASSIFY HEAVY VS. LIGHT VEHICLES. ....	199
FIGURE 6. 10. CLASSIFIED RESULTS AFTER COLLABORATIVE DECISION MAKING FROM MULTIPLE STATE SPACES. ....	201
FIGURE 6. 11. SSE IDENTIFICATION BY VEHICLE TYPES. ....	202
FIGURE 6.12. SSE IDENTIFICATION RESULTS AFTER COLLABORATIVE DECISION-MAKING. ....	204

# LIST OF TABLES

## CHAPTER 2

TABLE 2. 1	GENERALIZED WINS DATA SET.....	45
TABLE 2. 2	DERIVED RELATIONS WITH GENERALIZATION OF WINS DATA SET.....	46

## CHAPTER 3

TABLE 3. 1.	CLASS/TYPE BASED MAXIMUM POLLING.....	83
TABLE 3. 2.	CLASS/TYPE BASED WEIGHTED AVERAGING.....	85
TABLE 3. 3.	SEGMENTATION BASED MAXIMUM POLLING.....	87
TABLE 3. 4.	SEGMENTATION BASED MAXIMUM POLLING WITH A TIE.....	88
TABLE 3. 5.	SEGMENTATION BASED WEIGHTED AVERAGING.....	91
TABLE 3. 6.	LUMPED MAXIMUM POLLING.....	95
TABLE 3. 7.	LUMPED WEIGHTED AVERAGING.....	98
TABLE 3. 8.	SUB-SYSTEM LEVEL DECISION MAKING.....	100

## ACKNOWLEDGMENTS

A most sincere thank you goes to my advisor Dr. Greg Pottie. Professor Pottie inspired in streamlining my research while providing valuable feedback on my algorithms. He is always there to talk on academic and all matters. He is a great advisor to be with and provided valuable feedback that covered new avenues of research and other breakthrough areas of development.

I would also like to express my gratitude to Dr. W. J. Kaiser who was instrumental during the initial phases of my research work. Dr. Kaiser's initial motivation and enthusiasm was instrumental in stimulating my interest with wireless sensor networks. Overall, he is a great advisor on academic and research matters. I am grateful to have known him early during my undergraduate years to develop a strong academic relationship.

I would like to thank Dr. Yao for all the expert advice given in the signal processing field. He is always available for all advice. Thanks go to Dr. Jalali with whom I worked to obtain a great start in my career. He is always concerned about issues faced by his students and researchers giving good advice always. Finally yet importantly thanks to Dr. Gerla for being in my committee and making me write the dissertation faster.

Finally, this work would not have been possible without the love and support of all family members. In particular, I would like to single out my parents as being the inspiration behind any good that I might have done.

A sincere thank you and indebtedness to all.

## VITA

- 1969 Born, Manipay, Northern Province, Sri Lanka
- 1998 B.S., Electrical Engineering  
University of California, Los Angeles
- 1999 M.S., Electrical Engineering  
University of California, Los Angeles
- 1998-2002 Graduate Student Researcher  
Electrical Engineering Department  
University of California, Los Angeles
- 2000-2001 Sr. Member of Technical Staff / Design Engineer  
Cognet Microsystems  
West Los Angeles, California
- 2001-2003 Sr. CAD Engineer  
Intel Corporation  
Calabasas, California & Chandler, Arizona

## PUBLICATIONS

1. S. Natkunanathan, J. Pham, W. Kaiser, G. Pottie, "Embedded Networked Sensors: Signal Search Engine for Signal Classifications" SECON conference October, 2004.

2. S. Natkunanathan, A. Krishnaswami, S. Seetharaman, Y. Tsai, "Optimization and Back End verification of on-chip Inductors" Intel DTTC Conference July, 2003.
3. S. Natkunanathan, G.J. Pottie, W. Kaiser, "Implementation and optimization of Signal Search Engine for target recognition in Wireless Integrated Networked Sensors," ARL Annual Symposium (Demo) March, 2001.
4. S. Natkunanathan, A. Sipos, S. Avery, G.J. Pottie, W. Kaiser, "A Signal Search Engine for Wireless Integrated Networked Sensors," ARL Annual Symposium, March, 2000.
5. G. Asada, I. Bhatti, T.H. Lin, S. Natkunanathan, F. Newberg, R. Rofougaran, A. Sipos, S. Valoff, G. J. Pottie, W. J. Kaiser, "Wireless Integrated Network Sensors (WINS)," Proceedings of the SPIE – The International Society for Optical Engineering, vol.3673, SPIE - Int. Soc. Opt. Eng, p.11-18, 1999.
6. A. Sipos, S. Natkunanathan, W. Kaiser, "A Time Domain Signal Classification and Identification Architecture for Wireless Integrated Networked Sensors," ARL Annual Symposium, March, 1999.

**ABSTRACT OF THE DISSERTATION**

**Signal Classification and Identification  
For  
Wireless Integrated Networked Sensors**

*by*

**Sivatharan Natkunanathan**

Doctor of Philosophy in Electrical Engineering

University of California, Los Angeles, 2004

Professor Gregory J. Pottie, Chair

Wireless Integrated Network Sensors (WINS) incorporate the latest advances in wireless technology, integrated circuits, networking protocols and computing to provide compact intelligent sensor nodes for many distributed monitoring applications. [5, 13, 14, 36, 85] From battlefield intelligence to personnel monitoring, WINS nodes will play an integral part in national security, the manufacturing industry, transportation and health care. These distributed wireless

sensors require signal-processing technology to enable local detection, classification, and identification of events. In this dissertation, a time domain Signal Search Engine (SSE) has been developed for event processing, allowing events to be classified, identified, and communicated with minimum data payload requirements. This SSE was shown to be able to resolve signal time-evolution behavior. This dissertation presents time-domain signal classification and identification algorithms, and the fusion of classified or identified sensor information for distributed and centralized decision-making. One of the highlights of this dissertation includes a modular design using the technique of signal data decomposition into state spaces. The SSE implementation has shown excellent results for acoustic and seismic vehicle signal classification and identification. Furthermore, excellent classification and identification performances are demonstrated for a variety of vehicle signal sources and environments.

The SSE algorithms were benchmarked with two parametric methods, MUSIC and the Pisarenko algorithm, for created test signal models. Further benchmarking with real world signals was done with a wavelet method. It was found that system level processing and functions contained in the signal pre-processing modules is the reason for the high accuracy rates.

The performance of the SSE depends on the assigned template tree and its span. The more complete the span of the tree to sensed signals, the better the accuracy rates and associated confidences measures. The type abstraction hierarchy (TAH) concept is used in building the classification / identification tree. Type abstraction hierarchy is a methodology where similar signals are grouped together within abstract types for selecting a common template. TAH brings a methodical hierarchical classification and identification structure to the SSE. The obtained results are fused for collaborative decision-making with distributed and centralized decision-making architectures. The fusion of these different types (i.e. acoustic, seismic) and state-space sensor results not only enhances the performance and throughput of the SSE or Multi-Signal Search Engines (M-SSE) but also makes it more robust in noisy environments.

# CHAPTER 1

## Introduction

### 1.1 WINS Background

Wireless Integrated Networked Sensors (WINS) form a new distributed information technology based on compact, low power nodes that include sensors, actuators, computation and communication systems. [5,13,14,36,85] WINS distributed nodes form autonomous, self-organized, wireless sensing and control networks (see Figure 1.1). WINS nodes include microsensors, signal processing, computation and low-power wireless networking.

WINS have been developed over the last few years for basic scientific inquiry, defense applications in battlefield security, condition-based maintenance, and applications in industrial automation and healthcare. WINS technology includes MEMS micropower sensor and interface systems, micropower signal processing, and micropower RF

communication circuits. [15,31,58,69] WINS also includes novel signal processing algorithms for event recognition, network communication protocols for low power operation, and methods for bridging low power WINS networks with conventional long-range tactical networks. Most of the signal processing and event detection of the complicated signals takes place in a remote node that has minimal or no power constraints. The purpose of this SSE is to research the possibility of on-node implementation or use in off-node data-mining and decision-support systems.

By supporting a dense distribution of low cost measurement nodes, the distributed sensor network increases the effectiveness of surveillance by increasing the density of sampling points. In addition, current applications exploit the dense sensor distribution to enable multi-hop networking and reduced path loss resulting in dramatic communication energy reduction. Finally, WINS wireless networks also take advantage of distributed sensors to introduce fault tolerance into sensing and communication.

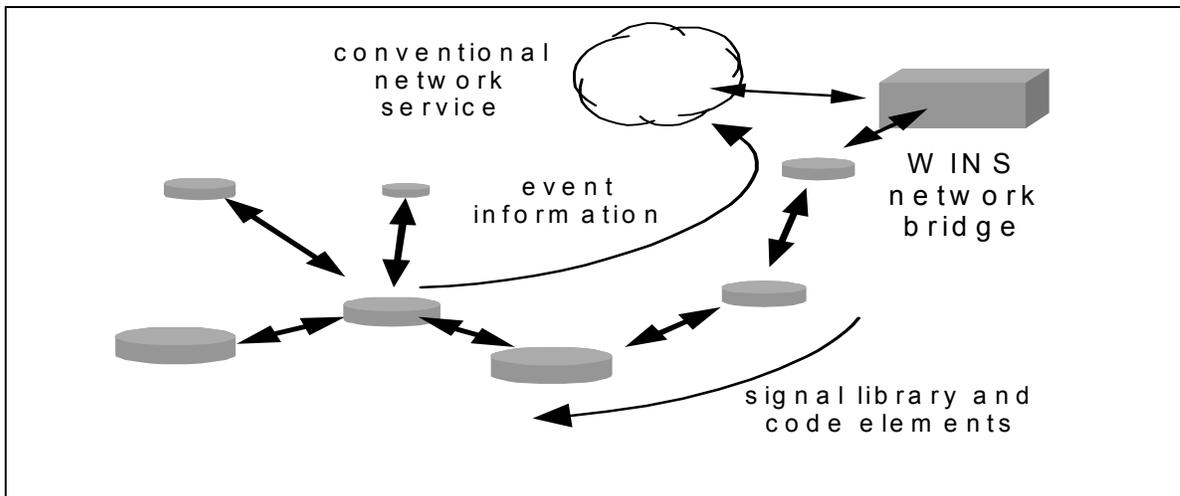


Figure 1. 1. WINS deployment for event detection.

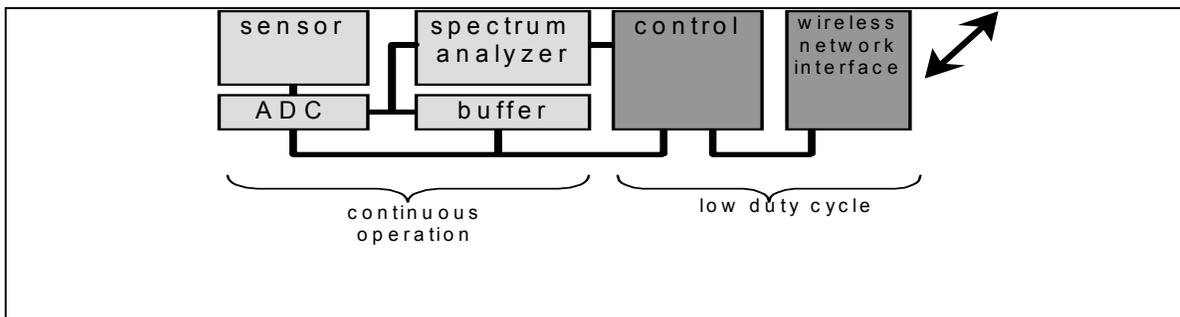
WINS have been demonstrated in distributed surveillance applications in many exercises with the Department of Defense. In addition, WINS vibration monitoring sensors have been deployed for condition-based measurement applications on a Navy ship. WINS are now being developed for a wide variety of applications.

Figure 1.1 shows WINS nodes (shown as disks) that are distributed at high density in an environment to be monitored. Multi-hop communication permits low power operation of dense WINS sensor networks. WINS node data is transferred over an asymmetric wireless link to an end user or to a long-range network service through a WINS network gateway. The WINS gateway manages the network and provides a network protocol translation. SSE library and code information flows to the nodes, while derived events and data flow to the remote user.

Recent advances in integrated circuit technology have enabled the construction of far more capable yet inexpensive sensors, radios, and processors, which has led to the mass production of sophisticated systems that link the physical world to digital data networks. Compact geometry and low cost allow WINS to be embedded and distributed at a fraction of the cost of conventional wireline sensor and actuator systems.

WINS opportunities depend on the development of a scalable, low-cost, sensor-networked architecture. Such applications require delivery of sensor information to the

user at a low bit rate through low-power transceivers. Continuous sensor signal processing enables the constant monitoring of events in an environment in which short message packets will suffice. Future applications of distributed embedded processors and sensors will require vast numbers of devices. Conventional methods of sensor networking represent an impractical demand on cable installation and network bandwidth. Processing at the source will drastically reduce the financial, computational, and management burden on communication system components, networks, and human resources.



**Figure 1. 2. WINS architecture with functional modules.**

In this dissertation, we limit ourselves to a security application designed to detect and identify threats within some geographic region and to report the decisions concerning the presence and nature of such threats to a remote observer via threat codes indexed in the signal search engine. The SSE is designed to perform event recognition and signal processing at the node itself to minimize the communication and power constraints inherent in wireless sensor nodes.

Figure 1.2 shows a wireless integrated network sensor (WINS) node architecture that includes sensor, data converter, signal processing, and control functions. Micropower RF communication provides bi-directional network access for low bit-rate, short-range communication. The micropower components operate continuously for event recognition, while the network interface operates at a low duty cycle. The prototype SSE is resident in the node spectrum analyzer. Future implementations of the SSE may operate as an independent coprocessor with signal pre-processing, signal identification, and classification modules.

## **1.2 *Signal Classification and Identification***

The dissertation includes the following contributions. The results are benchmarked with created test signals along with real world signals recorded during field trials. The ‘ACIDS’ database provided by the Army Research Laboratory (ARL) is used for benchmarking the SSE with real world signals. [92]

### **1.2.1 SSE Tree Building**

The SSE contains a correlator template tree that is vital for classification and identification. Template trees are updated with the introduction of new signals or when erroneous decisions or low confidence decisions are present. Battery power limitations require the building of template trees off-line to increase battery life of individual nodes. Research done for this dissertation concentrates on the building of a generalized tree

structure (generic template tree for all sensors in a cluster) and is performed using the ‘ACIDS’ database. [92] This database was obtained (created) in an extensive real world data gathering exercise, with real world ‘tracked’ and ‘wheeled’ vehicles. Correlator templates are selected after grouping training signals according to terrain, vehicle type etc and then made ready for template selection. Efficient, generalized correlator template tree assignments are presented in this dissertation.

### **1.2.2 SSE Training**

The SSE is trained using an existing database depending on applications, needs, and signals present in the database. A well-structured database will enhance the building of a proper template tree.

a. Pre-processing and identification of variables: Training is initially done for investigating signal variables and parameters. These variables and parameters depend on the sensed signals and the class / type of sensors used in sensing. Identified signal variables and parameters are mitigated in signal pre-processing modules. The investigation of signal variables and parameters are signal dependent. Investigating signal variables and parameters are presented in this dissertation for moving vehicle waveforms, which tend to have many complex variables in comparison to other signals. (i.e.: Bio-medical signals, stationary signals). Signal pre-processing modules are assigned and customized depending on the variables present. A state-space decomposition is used during pre-processing with individual signal processing operations

for each signal that is then fed into the signal classification and identification algorithm (SCIA).

b. Feature Extraction: It is during the training phase that distinct signal features for classification and identification are obtained. Individual signal features are extracted and used in building a template tree. These distinct features are then used in building tree structures and form the core of the template tree. Distinct features are selected depending on the chosen signals and on the class / type of the considered signal with its associated data-gathering phase.

c. Probability assignment for correlator templates: Training is used not only for selecting correlator templates to be used in the template tree, but also to assign probabilities to each selected correlator in the template tree. Probabilities are assigned depending on post-selection of the template tree, with extensive use of datasets gathered in real world scenarios. The assignment of probabilities is more stable and reliable when the training database has a large signal set with most variables present as experienced in real world deployment.

### **1.2.3 Signal Detection and Confidence Measure Assignment**

The detection of signals is separated into classification and identification. The Signal Classification and Identification Algorithm (SCIA) performs time-domain signal classification and identification. This dissertation concentrates on the proper classification and identification of signals using time domain SCIA algorithms. Proper classification and identification depend on the 'SCIA' modules that have application

specific modules that are based on the gathered signal set. The time domain SSE (inclusive of signal pre-processing modules, and 'SCIA') was benchmarked with test signals and real world signals. Based on the successful classification and identification results obtained by the use of the SCIA, it is suggested that attempts should be made at implementing this algorithm using hardware-software co-design or programmable hardware.

#### **1.2.4 Signal State-Space Decomposition**

Signal pre-processing includes signal decomposition into state-spaces (which is exploited for proper classification and identification). Signal variables are reduced or eliminated during state-space decomposition and forms a novel and critical methodology for signal classification and identification. The time domain SSE algorithm has to include signal state-space decomposition without which signal classification or identification is not possible. State space assignments not only give more accuracy in classified or identified results, but also give more control features and present a modular approach to classification and identification. Various environmental effects on signals are mitigated by the use of state-space decomposition that is used as switching modules utilized for different sensor types or signal segments.

#### **1.2.5 Decision Fusion**

Decision fusion is an important methodology for collaborative decision-making. Distributed and centralized decision-making are possible depending on the WINS design

architecture. The implemented architecture is based on the application and is presented in detail in this dissertation. Decision-making and confidence assignments are performed depending on the decision-making criteria. Decision-making is based on state-space and state-class / type. Modular control of sensed signals is obtained following the decision-making criteria that are presented extensively for different design architectures. The following methods are derived for WINS decision-making architectures.

**Maximum Polling:** A process where polling is done on individual nodes and the highest number of polled decisions is taken as the result. Associated confidence is calculated for the result on decisions used to obtain the final decision and their interim individual confidences.

**Weighted Averaging:** Weighted averaging makes decisions based on a list of decisions obtained from each module / decision node. The list of decisions is used depending on the class / type, or state-space. Decision-making is separated from weighted averaging depending on the design architecture. Associated confidence measures are obtained using the confidence calculation criteria detailed in this dissertation.

**Lumped Decision:** Lumped decision-making is used to make collaborative decisions, but at the cost of modular control. Lumped decisions are used when the decision-making process needs to be simple. However, losing modularity in lumped decision-making makes it difficult to track errors when error tracking is required.

System Level Decisions: When decisions arrive from sensor sub-systems, each containing multiple sensors, the decision making criteria is weighted with the number of sensors from each cluster / sub-system. An introduction to system level decision making is provided with examples, and should be tested in real world situations.

This dissertation concludes with a comparison and benchmarking of the algorithm with created test signals and real world signals giving excellent results for the time-domain signal classification / identification algorithm.

## **CHAPTER 2**

### **System Level Description of WINS**

#### **2.1 Applications of WINS**

Wireless Integrated Networked Sensors (WINS) will enable new applications for event detection, condition-based maintenance, manufacturing automation, and various environmental and bio-medical applications. Further, applications of intelligent sensing with additions of the latest data-warehousing and data-mining techniques aid the evolution of integrated sensor based decision support systems. These applications have a fundamental requirement of the need to properly classify, identify, catalog, and use data warehousing of sensed signals and decisions. One means to obtain the classification and identification of signals is the time domain Signal Search Engine (SSE), which is the core of this dissertation.

Wireless sensor nodes and systems are energy constrained due to battery power limitations. [76] These limitations restrict processing of SSE modules. Wireline use of SSE enables computationally intensive signal processing modules (extended modules attached to the time domain SSE) that may otherwise be excluded due to the computational power limitations in a wireless sensor system.

In this chapter, an exploration of wireless communication cost to that of transmitting data and decisions for short, medium, and long range is investigated. A new application of decision support systems based on the SSE is presented for wireline applications. An introduction to the time domain SSE / M-SSE is presented in this chapter with a system level description of the signal pre-processing, SCIA, and decision making modules of the SSE. SSE requirements of scalability, reconfigurability, and deployability are taken up on the chapters that form the basis of the dissertation.

A successful outcome of wireless node sensing is to properly classify and identify sensed signals with the given constraints. Detected signals are then used by the end user as information, to support a decision, or to take appropriate counter measures or for accumulation of intelligence. Accurate classification and identification is the goal of SSE modules. Once accurately classified / identified the sensed signals need to be sent to the user with low power wireless transmission. It is therefore necessary to look at the detection problem with

wireless data communication, since it is found that wireless communication dominates the energy budget. The detection system design must thus cope with limits in transmitting data for data fusion, and decision fusion. Here we explore the wireless communication cost of data and battery life to get a quantitative analysis of wireless sensor network lifetime and wireless communication issues.

## **2.2 Path Loss and Fading Issues in RF**

Signal transmission between nodes experience many forms of environmental interferences similar to what source signals experience during propagation through a transmission medium. Here we look at the case when sensor nodes are transmitting (with RF antennas) between each other (for cases of short range ( $\leq 30\text{m}$ ), medium long range ( $30 < D \leq 100\text{m}$ ), and long range ( $100 < D \leq 500\text{m}$ )). These scenarios occur when nodes in a cluster transmit among each other (short range), or transmit to a centralized node (medium long range), and when a cluster head transmits to a base station or remote node (long range) that is located much farther away compared to nodes in a local cluster.

Path loss occurs during data communication between a transmitter and receiver, where electromagnetic waves go through spreading, as well as energy loss due to interaction of electromagnetic waves. [83] A generalized formula for path loss is

given below, with a mean path loss  $\overline{PL}(d)$ , measured in dB, at the transmitter-receiver separation distance  $d$ . The mean path loss is conveniently computed with respect to a reference distance  $d_0$  (which is the free space distance). For sensor node clusters we choose  $d_0$  to be 10m, while for transmission to a cluster head  $d_0$  is set to 100m. The path loss exponent  $n$  can be in the range  $2 \leq n \leq 4$ , with  $n=2$  for free space, and  $n=4$  for near ground propagation. Then

$$\overline{PL}(d) = \overline{PL}(d_0) + 10 * n * \log_{10}\left(\frac{d}{d_0}\right) \text{ (dB)} \quad (2.1)$$

where  $\overline{PL}(d_0)$  is given by the following formula

$$\overline{PL}(d_0) = \frac{(4\pi)^2 d_0^2}{\lambda^2} \quad (2.2)$$

or in log form expressed as

$$\overline{PL}(d_0) = 20 \log_{10}\left(\frac{4\pi d_0}{\lambda}\right) \quad (2.3)$$

where  $\lambda = 3 * 10^8 \text{ m/s} / f_{\text{in Hz}}$ .

The above formulas present cases of wireless node communication for sensor networks. Having formulated the path loss macro model, let us calculate the effects of wireless transmission on communication parameters. These parameters have direct effects on battery life that are critical for battery power constrained sensor nodes.

### 2.2.1 Effect of Communication vs. Battery Life

Sensor nodes wirelessly communicate decisions based on detected events. These events may occur randomly or routinely depending on sensor network topology and implementation locality. [3,54,81] Threshold detectors are used to determine whether an event occurred to trigger the SSE / M-SSE. When an event is classified or detected, communications are required to a cluster-head or a centralized node through wireless RF transceivers to alert the user / client. A study of transmission time to battery life would give an idea of how much energy is lost with time. [48,82]

Assume the sensor node has the following battery power = 100mA-hour;  
Conventional transceiver with an operating frequency: 902 – 928 MHz;  
Current drain: 3V, 200mA (Transmission/Receive), 35mA on Receive;  
Available systems: 3V, 1mA peak current; [66]

Calculating the battery life for the above system:

Average wireless communication rate: 10 Events / Day;  
Event length: 32 bits/Event at the rate of 1 Kbits/s;  
Total transmission time: 0.8 ms/hour;

Battery lifetime T to transmit 10 events every Day:

$$T = \frac{100 * 60 * 60 * 60(mA - ms)}{(0.8 * 200)(mA - ms) / hour} = 135,000hours = 15.41years \quad (2.4)$$

Assuming that the battery usage for wireless communication of decisions consumes 33% of battery power, it is observed that the battery lifetime would be of the range of 4-6 years depending on the deployed scenario. Above calculations, are quite stringent and would in many scenarios average less than 10 events a day depending on sensor applications and deployment. Here calculations assume sleep mode when transmission does not occur. The above calculation takes in decision transmission only and does not assume data transmission for collaborative data fusion. Therefore, it could be seen that communicating decisions only would result in long battery life, compared to transmitting data for collaborative data fusion.

### **2.2.2 Maximum Bandwidth vs. Transmitted Power**

In situations when signal data needs to be transmitted to a cluster-head / centralized node for data fusion or beamforming purposes, transmission between sensor nodes could be optimized for bandwidth with power considerations. [6,62,66] Consider scenarios where distances between the centralized and cluster

head nodes are set to be a constant. These calculations are in an environment with the following sensor node specifications:

Assuming maximum transmitter receiver separations of:

30m (node to node) - short range; single cell;

100m (node-to-node) medium long range; single cell;

500m (cluster head to base station) – long range; dual cell;

Transmitter power consumption: 0.06 W @ 915MHz;

Transmitter power consumption: 22 mW to 80 mW @ 915MHz; [66]

Assume a 10dB noise figure and a required SNR of 25 dB.

Let us assume unity gain antennas at both ends.

Assume free space or near-ground propagation.

Calculating the Path Loss from equation 2.2:

$$\lambda = \frac{c}{f_c} = \frac{3 * 10^8}{9.15 * 10^8} = 0.3278m \quad (2.5)$$

Therefore mean path loss at  $d_0 = 10m$  and  $100m$  respectively are,

$$\overline{PL}(d_0 = 10m) = \frac{(4\pi)^2 d_0^2}{\lambda^2} = \frac{(4\pi)^2 * (10)^2}{(0.3278)^2} = 1.469 * 10^5 = 51.67dB \quad (2.6)$$

$$\overline{PL}(d_0 = 100m) = \frac{(4\pi)^2 d_0^2}{\lambda^2} = \frac{(4\pi)^2 * (100)^2}{(0.3278)^2} = 1.469 * 10^7 = 71.67dB \quad (2.7)$$

From Equation 2.1 above, for free space  $n=2$ , at  $d= 30m, 100m,$  and  $500m$  we obtain;

$$\begin{aligned}\overline{PL}(d = 30m) &= \overline{PL}(d_0 = 10m) + 10 * n * \log_{10}\left(\frac{d}{d_0}\right) (dB) \\ &= 51.67 + 10 * 2 * \log_{10}\left(\frac{30}{10}\right) = 61.21dB\end{aligned}\quad (2.8)$$

$$\begin{aligned}\overline{PL}(d = 100m) &= \overline{PL}(d_0 = 10m) + 10 * n * \log_{10}\left(\frac{d}{d_0}\right) (dB) \\ &= 51.67 + 10 * 2 * \log_{10}\left(\frac{100}{10}\right) = 71.67dB\end{aligned}\quad (2.9)$$

$$\begin{aligned}\overline{PL}(d = 500m) &= \overline{PL}(d_0 = 100m) + 10 * n * \log_{10}\left(\frac{d}{d_0}\right) (dB) \\ &= 71.67 + 10 * 2 * \log_{10}\left(\frac{500}{100}\right) = 85.649dB\end{aligned}\quad (2.10)$$

From the above calculations, it is obtained that path loss for 30m, 100m, and 500m are

$$\overline{PL}_{(30m)} = 61.21dB; \quad \overline{PL}_{(100m)} = 71.67dB; \quad \overline{PL}_{(500m)} = 85.649dB; \quad (2.11)$$

Having obtained the path loss at a distances d=30m, 100m, and 500m we obtain the received power during transmission with the following formulae

$$P_r(d)_{(dBm)} = (P_t)_{dBm} + (G_t)_{dB} + (G_r)_{dB} - (\overline{PL}(d))_{dB} \quad (2.12)$$

Substituting,  $\overline{PL}_{(30m)}=61.21dB$ ,  $\overline{PL}_{(100m)}= 71.67 dB$ ,  $\overline{PL}_{(500m)}= 85.639 dB$ , and antenna gains of 0dB in equation 2.12

$$P_r(30m)_{(dBm)} = 17.78 + 0.0 + 0.0 - 61.21 = -43.43dBm \quad (2.13)$$

$$P_r(100m)_{(dBm)} = 17.78 + 0.0 + 0.0 - 71.67 = -53.89dBm \quad (2.14)$$

$$P_r(500m)_{(dBm)} = 17.78 + 0.0 + 0.0 - 85.649 = -67.869dBm \quad (2.15)$$

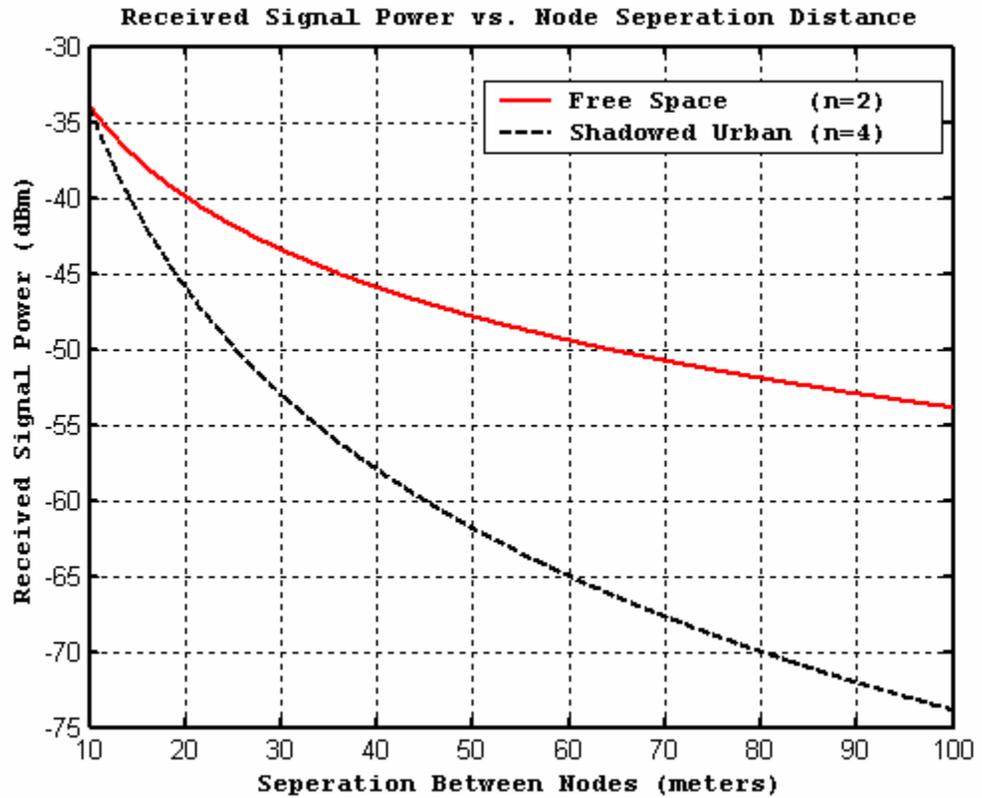


Figure 2.1 Received Signal Power vs. Separation Distance for short/medium range.

<sup>1</sup>Figure 2.1 shows received signal power for separation distances up to 100m. It is assumed that the wireless RF transceiver transmits at a power level of 60mW at a

---

<sup>1</sup> Near Ground Propagation is defined as Shadowed Urban in Plots.

carrier frequency of 915MHz. Both complete free space, and near-ground propagation environments (with free space reference distances  $d_0=10m$ ) are plotted. Calculation assumes unity gain antennas at the transmitter and receiver.

Assuming that SNR should be 25dB for reliable transmission [66], then the maximum noise power is required to be

$$N = P_r(d)_{(dBm)} - SNR_{dB} \quad (2.16)$$

$$N_{30m} = P_r(30m)_{(dBm)} - SNR_{dB} = -43.43 - 25 = -68.43dB \quad (2.17)$$

$$N_{100m} = P_r(100m)_{(dBm)} - SNR_{dB} = -53.89 - 25 = -78.89dB \quad (2.18)$$

$$N_{500m} = P_r(500m)_{(dBm)} - SNR_{dB} = -67.869 - 25 = -92.869dB \quad (2.19)$$

If only thermal noise is considered, noise power N is given by [83]

$$N_{(dBm)} = KT_0BF \quad (2.20)$$

or

$$N(dBm) = -174(dBm) + 10 \log_{10} B + F_{(dB)} \quad (2.21)$$

where  $K = 1.38 \cdot 10^{-23}$  J/K is Boltzmann's constant, and  $T_0 = 290$  K the standard temperature. B is receiver bandwidth in Hz, while the noise figure of the receiver in dB is given by F.

Using equation 2.21, we obtain the following

$$N_{30m} = -174\text{dBm} + 10\log_{10}B + 10 = -68.43 \text{ dBm} \quad (2.22)$$

$$N_{100m} = -174\text{dBm} + 10\log_{10}B + 10 = -78.89 \text{ dBm} \quad (2.23)$$

$$N_{500m} = -174\text{dBm} + 10\log_{10}B + 10 = -92.869 \text{ dBm} \quad (2.24)$$

Solving for B we have maximum bandwidth,

$$B_{30m} = 3.605 \text{ GHz}; \quad (2.25)$$

$$B_{100m} = 324.34\text{MHz}; \quad (2.26)$$

$$B_{500m} = 12.975\text{MHz}; \quad (2.27)$$

Repeating the above calculations for the case below, we find the following:

From Equation 2.1 above, for free space  $n=4$ , at  $d= 30m$ ,  $100m$ , and  $500m$  respectively;

$$\overline{PL}_{(30m)} = \overline{PL}_{(d_0 = 10m)} + 10 * n * \log_{10}\left(\frac{d}{d_0}\right) (dB) = 51.67 + 10 * 4 * \log_{10}\left(\frac{30}{10}\right) = 70.754dB \quad (2.28)$$

$$\overline{PL}_{(100m)} = \overline{PL}_{(d_0 = 10m)} + 10 * n * \log_{10}\left(\frac{d}{d_0}\right) (dB) = 51.67 + 10 * 4 * \log_{10}\left(\frac{100}{10}\right) = 91.67dB \quad (2.29)$$

$$\overline{PL}_{(500m)} = \overline{PL}_{(d_0 = 100m)} + 10 * n * \log_{10}\left(\frac{d}{d_0}\right) (dB) = 71.67 + 10 * 4 * \log_{10}\left(\frac{500}{100}\right) = 99.628dB \quad (2.30)$$

Substituting,  $\overline{PL}(d = 30m) = 70.754\text{dB}$ ,  $\overline{PL}(d = 100m) = 91.67\text{dB}$ ,  $\overline{PL}(d = 500m) = 99.628\text{dB}$ , with antenna gains of 0dB in equation 2.12,

$$P_r(d = 30m)_{(dBm)} = 17.78 + 0.0 + 0.0 - 70.754 = -52.974\text{dBm} \quad (2.31)$$

$$P_r(d = 100m)_{(dBm)} = 17.78 + 0.0 + 0.0 - 91.67 = -73.89\text{dBm} \quad (2.32)$$

$$P_r(d = 500m)_{(dBm)} = 17.78 + 0.0 + 0.0 - 99.628 = -81.848\text{dBm} \quad (2.33)$$

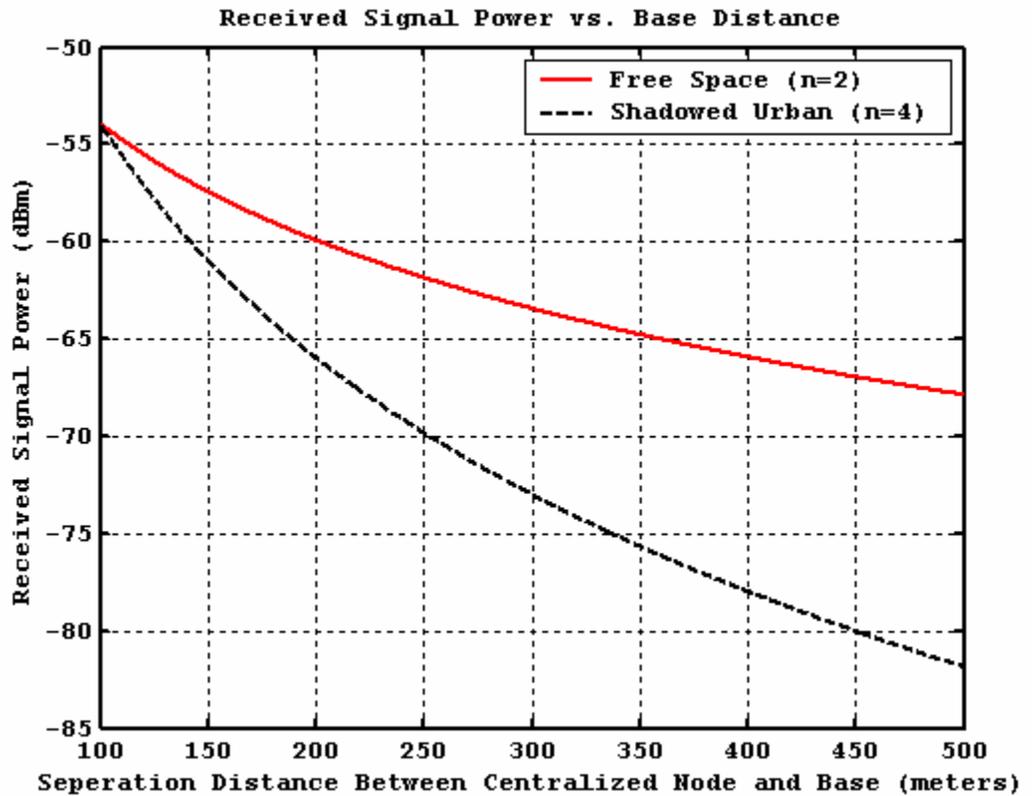


Figure 2. 2 Received Signal Power vs Distance for long range.

Figure 2.2 shows received signal power for separation distances up to 500m used for transmission from cluster heads to a remote centralized node or client. It is assumed that the wireless RF transceiver transmits at a power level of 60mW but could be adjusted to a higher power level to increase received noise power. The carrier frequency of 915MHz is used for complete free space, and near ground propagation environments (with free space reference distances  $d_0=100m$ ). The calculation assumes unity gain antennas at the transmitter and receiver.

Assuming that SNR should be 25dB, then the maximum noise power is required to be

$$N_{30m} = P_r(d = 30m)_{(dBm)} - SNR_{dB} = -52.974 - 25 = -77.974dB \quad (2.34)$$

$$N_{100m} = P_r(d = 100m)_{(dBm)} - SNR_{dB} = -73.89 - 25 = -98.89dB \quad (2.35)$$

$$N_{500m} = P_r(d = 500m)_{(dBm)} - SNR_{dB} = -81.848 - 25 = -106.848dB \quad (2.36)$$

From equation 2.21, we obtain the following

$$N_{30m} = -174dBm + 10\log 10B + 10 = -77.974 \text{ dBm} \quad (2.37)$$

$$N_{100m} = -174dBm + 10\log 10B + 10 = -98.89 \text{ dBm} \quad (2.38)$$

$$N_{500m} = -174dBm + 10\log 10B + 10 = -106.848 \text{ dBm} \quad (2.39)$$

Solving for B we have maximum bandwidth,

$$B_{10m} = 400.497\text{MHz}; \quad (2.40)$$

$$B_{100m} = 3.243\text{MHz}; \quad (2.41)$$

$$B_{500m} = 519.039 \text{ KHz}; \quad (2.42)$$

Figures 2.3 – 2.6 display bandwidth / transmitted power relationships for different values of  $d_0$ . The figures indicated that for most sensor network applications of interest, transmission rate would not be the limiting factor; rather it will be the

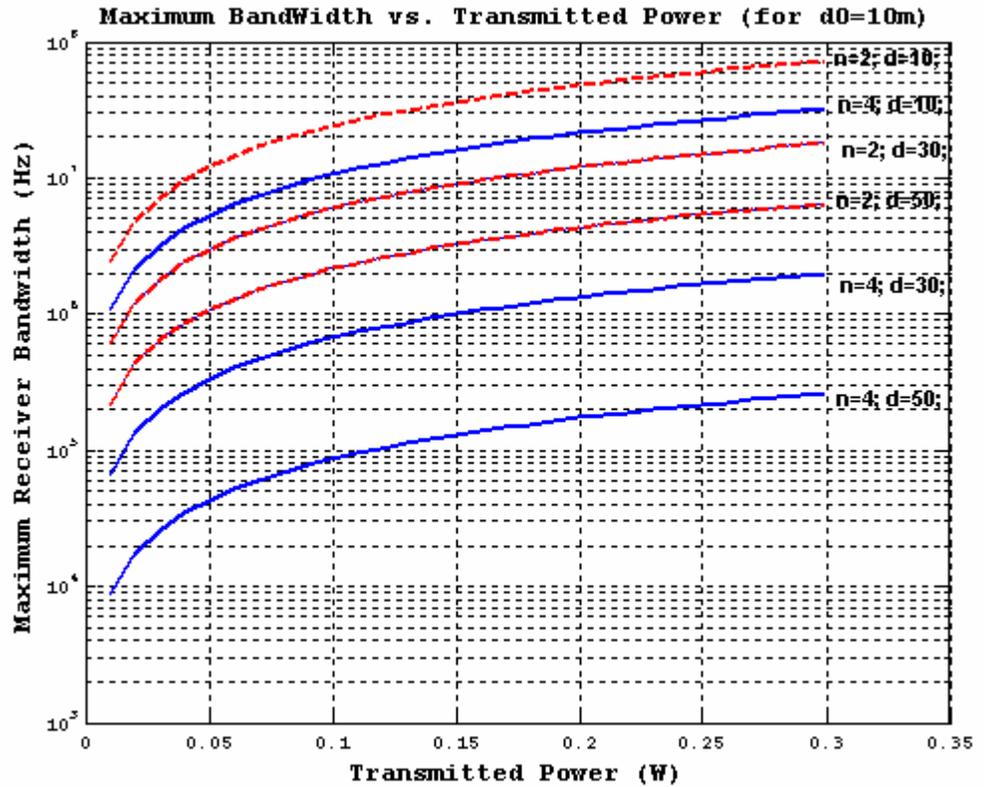


Figure 2.3 Maximum Receiver Bandwidth vs Transmitted Power.

energy to sustain transmissions. However, as will be seen in the next section, a more sophisticated propagation model indicates substantially greater difficulties for wireless communication.

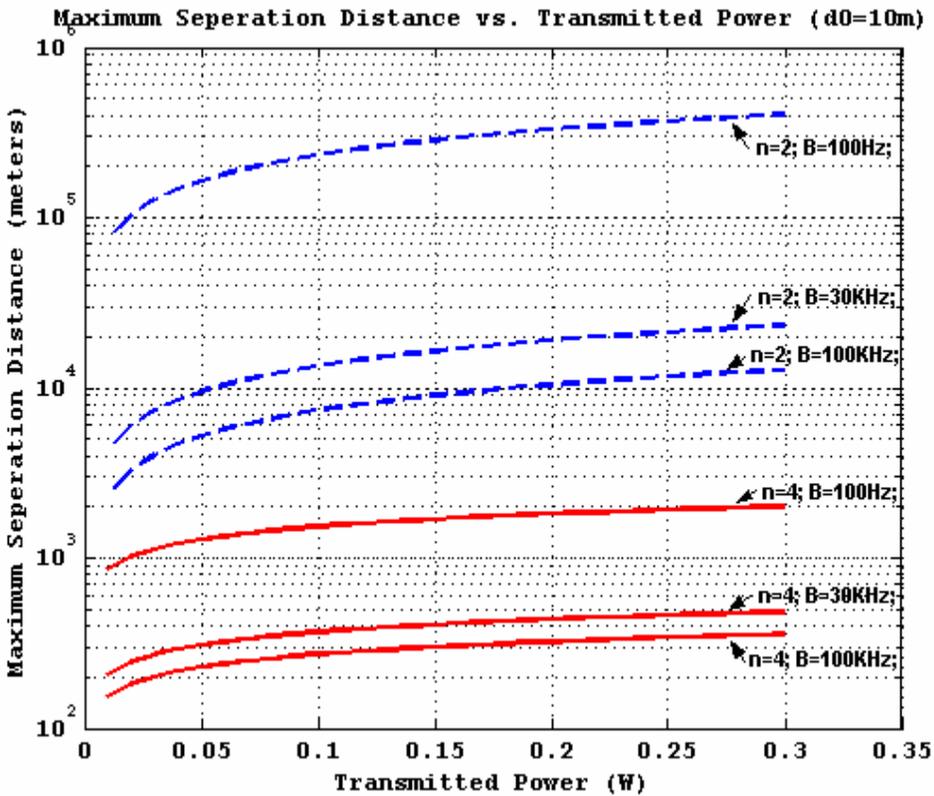


Figure 2. 4 Maximum Separation Distance vs. Transmitted Power (Free Space 10m).

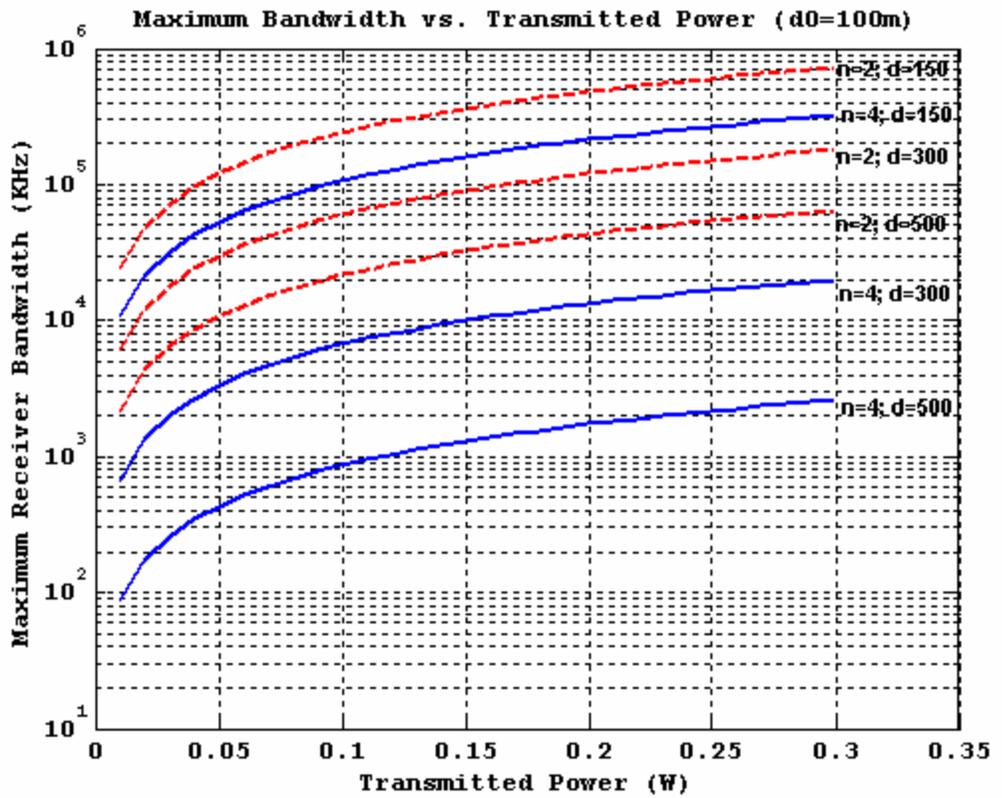


Figure 2.5 Maximum Receiver Bandwidth vs Transmitted Power (Free Space = 100m)

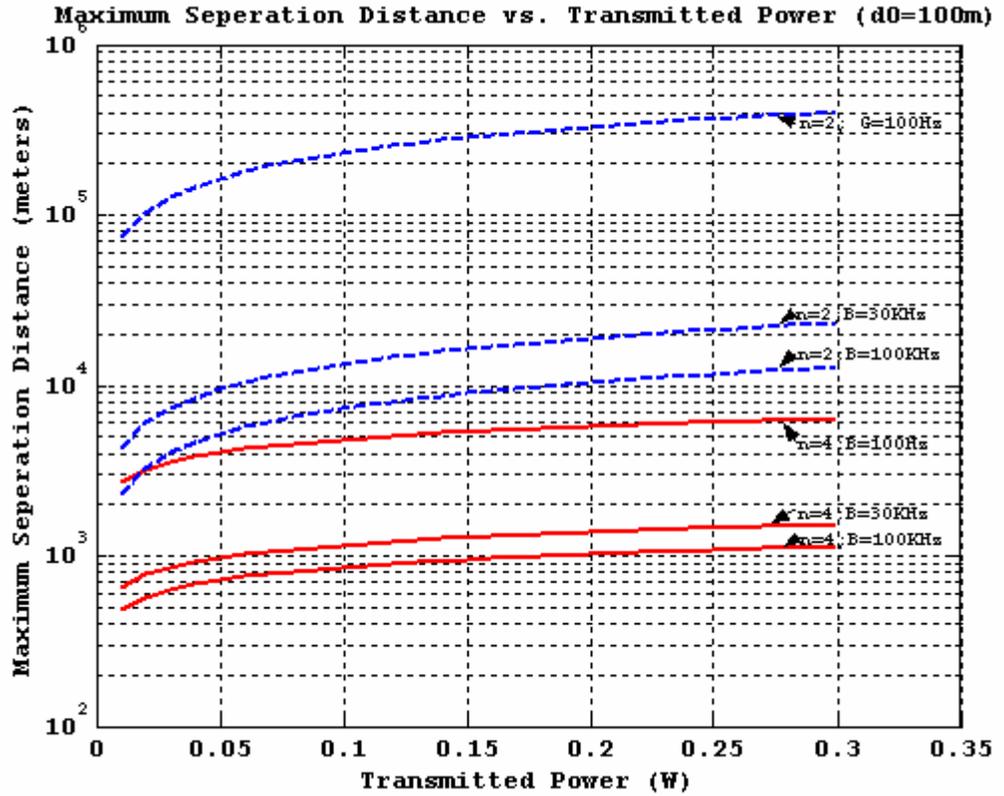


Figure 2. 6 Maximum Separation Distance vs. Transmitted Power (Free Space of 100m).

The calculations above yield a good approximation of wireless bandwidth in free space ( $n=2$ ), and for near-ground propagations ( $n=4$ ), absent large obstructions. Thus, when wireless communication cost intensive data fusion methods are used the above criteria determine communication bandwidth limitations. Sensor networks implemented for ground to ground links will for example obtain  $1/25^{\text{th}}$  the data rate compared to transmission during elevated links at 500m range to obtain same error rates.

## 2.3 Propagation Models - Fading Effects on RF Signals from Sensor Nodes

Two types of fading effects are considered for analysis in this section. The two methods are common in outdoor wireless sensor environments. Indoor wireless environments do go through similar environmental effects and have different propagation constants associated with the propagation formula.

Environmental effects such as multi-path fading, shadowing, and fast fading affect sensed wave (i.e. acoustic signal) propagation, in addition to transmitted signals from RF transceivers on nodes. Additionally other environmental effects could be added and incorporated into fading channel models to attain accurate theoretical calculations. The sections below give a detailed outlook of fading channel models used in sensor networks. [26]

### - *Shadowing (Slow Fading)*

Path loss models incorporate the following random variable as the shadowing component. The shadowing component is a zero-mean Gaussian random variable with a standard deviation of  $\sigma_L$  – representing the location variability.  $\sigma_L$  varies with environmental effects, along with antenna heights, and frequency of transmission. The probability density function (PDF) of the shadowing component  $L_S$  is often modeled by the following:

$$p(L_s) = \frac{1}{\sigma_L \sqrt{2\pi}} \exp\left[-\frac{L_s^2}{2\sigma_L^2}\right] \quad (2.43)$$

The above variable is used to obtain sensor propagation calculations and graphs while comparing relative costs associated with transmission in slow fading environments.

- *Rayleigh Fading (Fast Fading)*: Rayleigh fading is the common fading component for narrow band RF signals used to model fast signal variation due to scattering, diffraction, and refraction. Here the signal varies rapidly compared to path-loss, and shadowing with respect to time. Fast fading is often modeled in sensor networks with a Rayleigh distribution with initial conditions measured from node tests and real world measurements. The probability distribution function (PDF) of the signal amplitude is

$$p(r) = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}} \quad (2.44)$$

where  $\sigma$  is the standard deviation of the narrow band signal. The above model is used in obtaining energy data for transmission in a NLOS (non-line of sight) sensor network environment.

- *Rician Fading(Fast Fading)*

In LOS (line of sight) situations, where sensor nodes are in direct line of sight, the amplitude from the LOS dominates the power and is modeled by the Rician distribution given below:

$$p(r) = \frac{r}{\sigma^2} e^{-\left(\frac{r^2+s^2}{2\sigma^2}\right)} I_0\left(\frac{rs}{\sigma^2}\right) \quad (2.45)$$

where  $s$  is the direct path gain and  $I_0$  is the modified Bessel function of the first kind. The above model is used for proper estimation and calculation of the data payloads in sensor networks and is shown below.

*With Different Modulation Schemes:*

Sensor data is transmitted between sensor nodes using simple modulation schemes in sensor networks. [80] The following BPSK scheme is utilized for determining energy calculations for wireless transmission. A general BPSK modulation scheme is represented by

$$S_m(t) = S_{m1}f_1(t) + S_{m2}f_2(t) \quad (2.46)$$

where  $S_{m1}, S_{m2}$  are finite energy waveforms, with  $f_1(t), f_2(t)$  given as unit energy signal waveforms represented by

$$\begin{aligned}
f_1(t) &= \sqrt{2/\epsilon_g} g(t) \cos(2\pi f_c t) \\
f_2(t) &= -\sqrt{2/\epsilon_g} g(t) \sin(2\pi f_c t)
\end{aligned}
\tag{2.47}$$

where  $f_c$  is the carrier frequency of the signal. Using the above BPSK signal representation we obtain the following graphs shown in Figure 2.7 and 2.8 for propagation energy measurements.

### 2.3.1 Communication Energy vs. Range:

Sensor node wireless transmission costs were calculated for various sensor nodes placements along with obtaining energy requirements for communication. Here a communication energy vs. range calculation to transmit wirelessly a 1 Kb data payload in conventional free-space systems with a  $1/R^2$  propagation constant is presented. We present this case for the above-mentioned BPSK transmission with a required bit error rate (BER) of  $10^{-6}$ .

The same calculations were done for sensor networks for surface-to-surface transmission in both Gaussian, and Rayleigh channels with  $1/R^4$  propagation. Energy required for 1Kb data payload with a BPSK transmission and a bit error rate of  $10^{-6}$  is shown in Figures 2.7 and 2.8. The corresponding energy to transmit a distance of 100m in Rayleigh, and Gaussian channels are 3 Joules and 0.5mJ respectively. It is observed that the energy required to transmit increases steeply

with distance and is observed to be exceeding the low power limitations of sensor nodes after a 100 m range. [79]

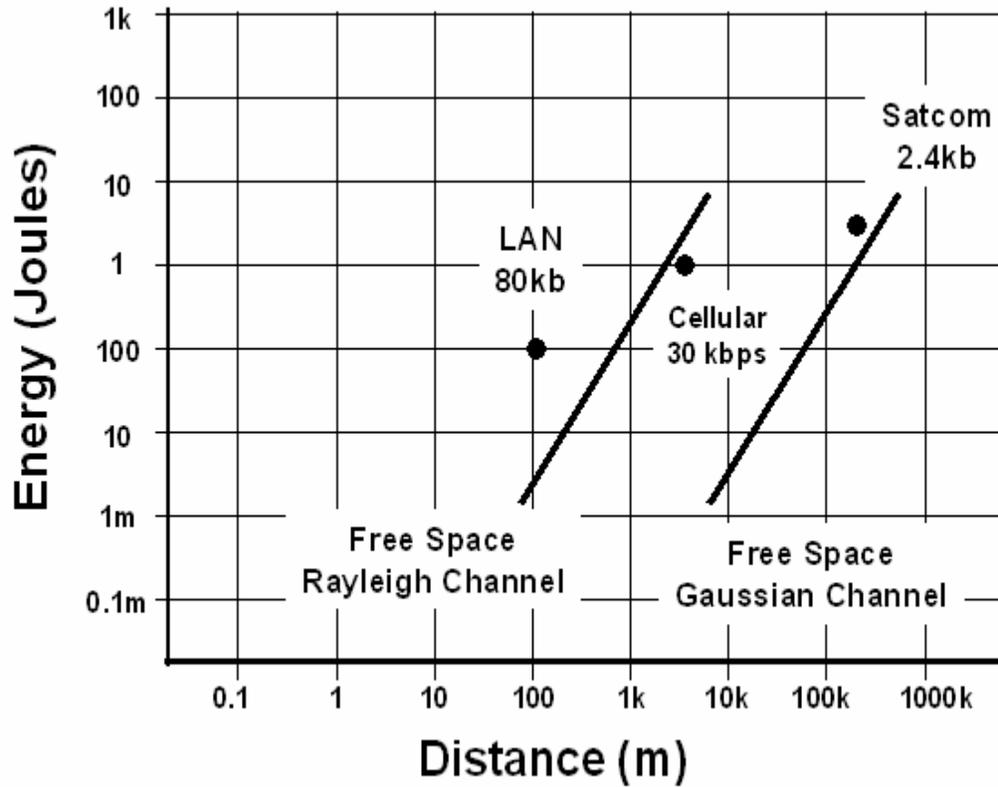


Figure 2. 7 Energy Consumption vs. Transmitted Distance (Channel Dependent  $1/r^2$ ).

Thus the prospects of both fading and ground propagation losses limit the range and bandwidth available in low power communications. This reinforces the need for in-network processing to reduce the number of instances of long-range communications.

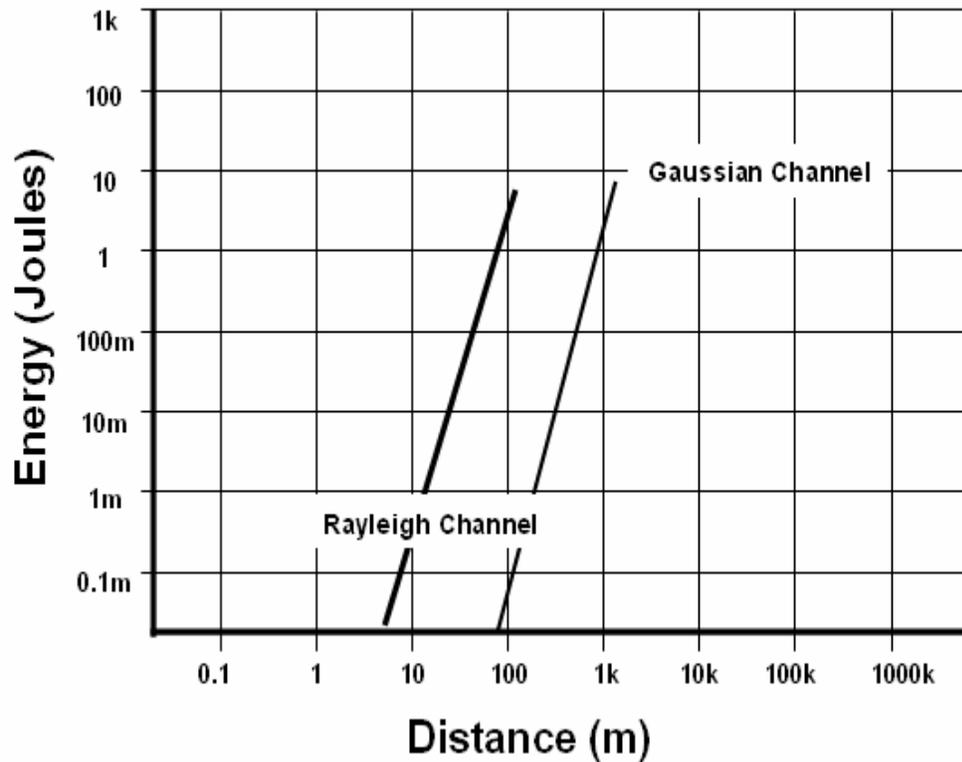


Figure 2. 8 Energy Consumption vs. Transmitted Distance (Channel Dependent  $1/r^4$ ).

## 2.4 Data Warehousing and Decision Support Systems

SSE applications include decision support systems and intelligent tracking with pervasive computing. [89,94,95] An initial well structured database would enable in decision support systems where any new signal that arrives at the sensor node would be identified and classified and use a look-up table or a codebook built

from past signals enabling decision support. The method of data warehousing and decision support systems is explained below with a macro level diagram used with intelligent sensor systems.

WINS defense applications include battlefield situational awareness for security and tactical advantage and many other intelligent information applications. [8] These features in WINS require large amounts of information (data) to be acquired, indexed, stored, retrieved, and used to classify potential targets. A methodical and structured way of storing the acquired data (data warehousing) and an intelligent method of analyzing patterns (data mining) are looked at from a

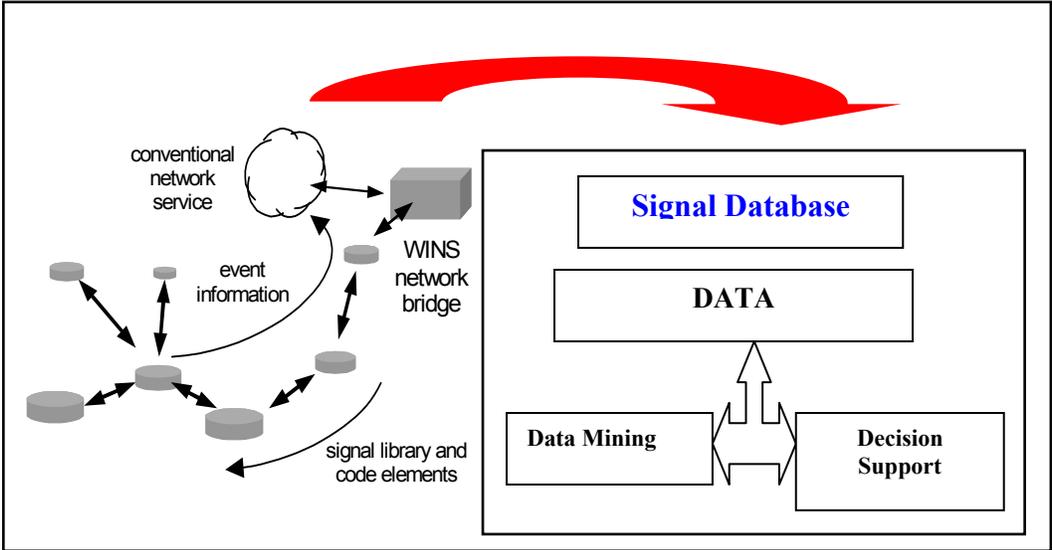


Figure 2.9 SSE usage in a Decision Support and Data Mining WINS application.

system level perspective [53] for the application of vehicle event detection and transportation monitoring. Figure 2.9 shows an application of SSE where WINS nodes (shown as disks) are distributed at high density in an environment to be monitored. WINS node data is transferred over the asymmetric wireless link to an end user or to a conventional wire-line or wireless network service through a WINS network Gateway. The WINS signal search engine (SSE) library and code information flows to the nodes, while derived events and data flow to the remote user. The signal database is at the remote user end, which stores, retrieves and gives intelligence information to signals and decision support systems. Various application specific data warehousing and data mining systems [43,96] help in providing information to the WINS Gateway while interacting with each other.

#### **2.4.1 Vehicle Monitoring Applications and Database Integration**

Vehicle monitoring has been studied in detail from traffic monitoring on highways to detecting threat vehicles around sensitive locations and institutions. Whatever the application of vehicle monitoring is, the information obtained by the sensors (WINS) look similar in the attributes they inherit. [27] Sensor nodes monitoring an institution may be randomly placed in the locality of the building. The location of the sensor becomes important, especially for locating the data source and further for signal processing of obtained data. Date and time of sensor installments would be critical information when it comes to finding the lifetime of

the battery. Therefore a sensor deployment information source (S\_DIS) becomes necessary.

Sensors themselves have useful information without any vehicles present. Some sensor-inherent specification parameters would help a great deal in giving meaning to the acquired data. Different types of sensors, distinguishable by product identification (PID which has SN1, SN2, SN3, SN4, SN5,...), have different values for parameters such as gain, battery life and detection range to name a few. These parameters would be encapsulated in the sensor product identification information source (S\_PID).

The data warehouse should also contain information on events that have been recorded earlier so that a user query of past history or a decision support query could be explored and validated. Therefore an event history information source (PEH) will contain an event (i.e.: bicycle rider, animal in surroundings, helicopter present ...) and additional information as time, date, speed and, location or sensor ID.

Further, an event classifier information source (EC) is needed for real-time autonomous event/threat detection or classification. Here, selected classifier tree-structures (i.e. 'templates' in template matching classification) could be included with associated confidence or other classifier method parameters. Therefore,

WINS needs an easy and efficient access to integrated information from multiple, heterogeneous, semi-autonomous, distributed information sources. This integrated information can be extracted from independent information sources and integrated to form a data warehouse. Significant use of WINS nodes in defense and consumer applications has made it a necessity to form a constructive and meaningful warehouse DBMS, primarily for decision support systems (DSS) and data mining for discovering patterns of behavior.

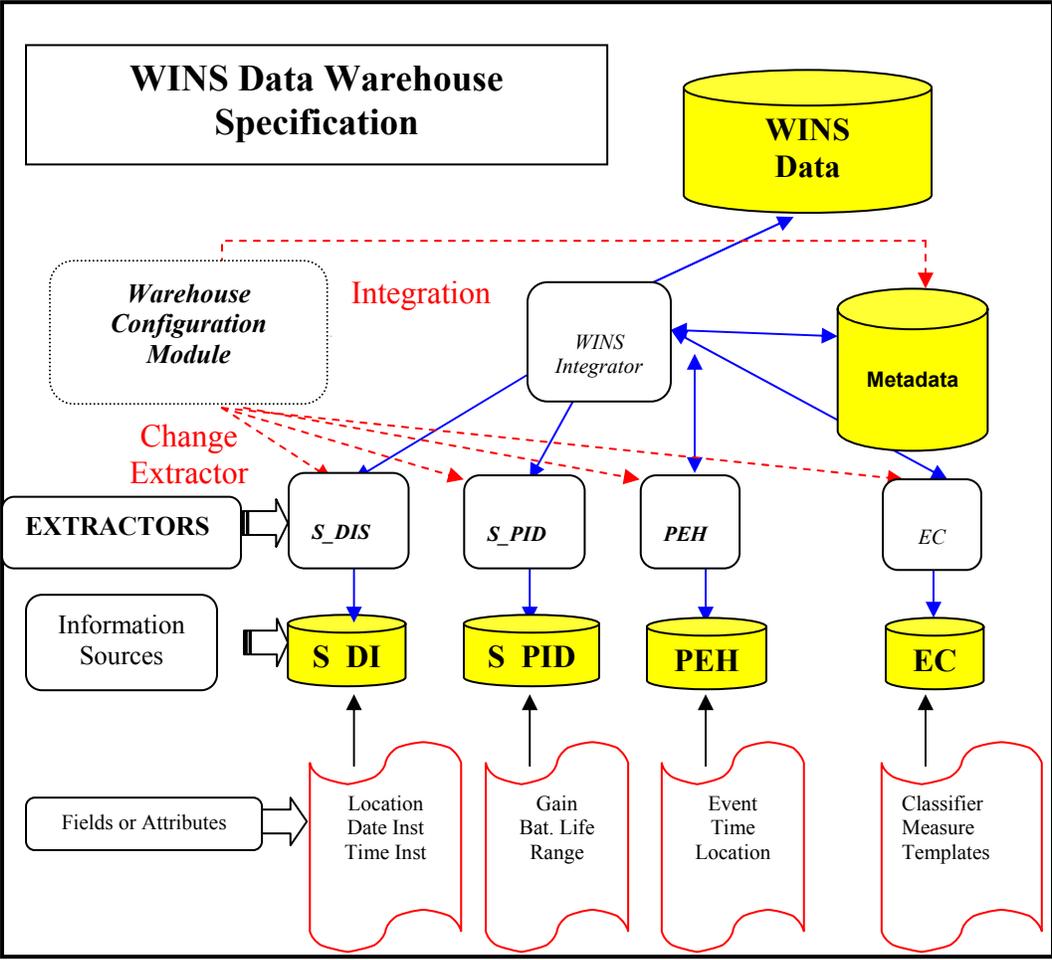


Figure 2. 10 WINS Data warehousing for SSE based Decision Support Systems.

A system level specification diagram on Figure 2.10 shows the WINS data warehouse model. The extractor and integrator play a major role in that all data in the data warehouse is first extracted with detection requirements or constraints, and then formatted to fit the warehouse model and attributes. Unwanted or useless data is deleted during the extraction process and thus does not reside in the warehouse. Extraction of data from information sources can be further constrained by application specific extraction and application specific warehousing. One such application specific function is mining a ‘training’ data set of signal waveforms and associated data for ‘learning’ and classification purposes. In that case ‘learning’ specific extraction constraints are used whereby the S\_PID information source could be completely omitted during the extraction process.

#### **2.4.2 Hierarchy Specifications Found in WINS Data**

If the vehicle monitoring case of the WINS application is taken into consideration, hierarchy specification becomes essential considering the variety and amount of information contained in the data warehouse. [27] Hierarchy specification and groupings further enhance data mining tasks by providing a structured knowledge base for the query at hand. It also helps in cutting down run-time, and unwanted tree - parsing when tree structures are present. Further,

these specifications help in rule-based querying and finding optimality for data mining tasks.

**Schema Hierarchy :** S\_DIS installment date and time would come under schema hierarchy either individually or as a combination for an application specific task. Use of individual or combined date and time measures depend on the mining task and relaxation constraints. A mining task which asks for an event data history would require both date and time in individual forms. However, a relaxed mining task, which requires hour and date of installation, to study installation performance of a worker or installing vendor, would like to have combined data (some parameters in each date and time variables). Thus, the schema hierarchy would look like the following macro design:

Define hierarchy **date\_hierarchy** on **date** as

**[date, month, quarter, year] (individual)**

Define hierarchy **time\_hierarchy** on **time** as

**[sec, minute, hour] (individual)**

Define hierarchy **date\_time\_hierarchy** on **date\_time** as

**[hour, date, month, year] (combined)**

**Set-grouping Hierarchy :** A data mining task for a traffic monitoring system using WINS would require speed clustering to derive knowledge out of raw data. The knowledge derived from this data could be used to study and monitor high-speed areas and detect traffic congestions and jams and even report in real time over the internet to possible travelers about traffic conditions. The PEH information source would then be built with a set grouping of speeds of vehicles. A level hierarchy of speed groupings is given below.

Define hierarchy **speed\_hierarchy** for **speed** on vehicle as:

Level 1: {very slow, slow, average, fast, over-limit} < level 0: **all**

Level 2: {0,...,15} < level 1: **very slow**

Level 2: {16,...,30} < level 1: **slow**

Level 2: {31,...,45} < level 1: **average**

Level 2: {46,...,70} < level 1: **fast**

Level 2: {>70} < level 1: **over-limit**

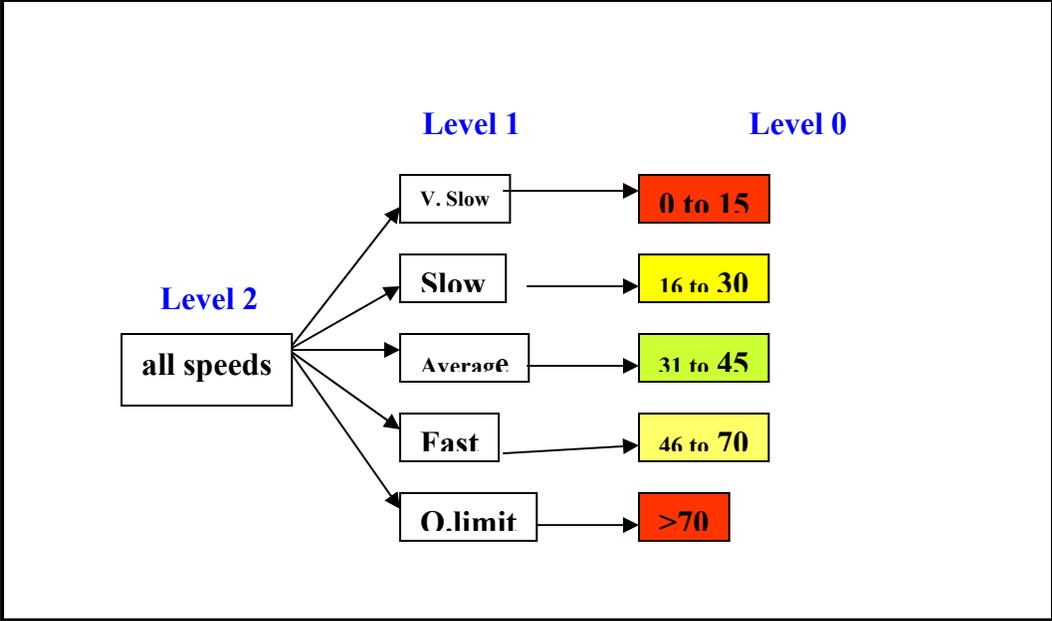


Figure 2. 11 An example of set grouping of speeds with WINS data.

The groupings and eventual tree structure would not only enhance a specific data mining task but also would help to form rules for rule induction purposes. Further statistical measures could be attached to the groupings and would give a better structure in acquiring knowledge through a direct user query. The knowledge bases would enhance the system intelligence and could be used for real-time monitoring of traffic and other data.

**Operation Derived Hierarchy:** Data mining also benefits from operation derived hierarchies namely categorizing sensor gains into a specified number of clusters.

Define hierarchy **gain\_hierarchy** for **gain** on sensor as

$$\{\text{gain\_category}(1), \dots, \text{gain\_category}(10)\}$$
$$:= \text{cluster}(\text{default}, \text{gain}, 10) < \text{all}(\text{gain})$$

This definition will categorize the sensor gains into ten different clusters whereby a user query would be optimized when a specific gain lookup is on for sensors. This would reduce run-time significantly since product\_IDs would be given for a specific gain category with a short run-time in this instantaneous query and solution process.

**Rule based Hierarchy:** Mining tasks can also benefit from rule based hierarchies. A person trying to find sensor installation priorities for different locations from a preliminary data source would want to prioritize the procedure. This would mean a location with high occurrence of over speeding would fall into the highest priority cluster. A rule hierarchy, which would enhance this procedure, is given below.

Define hierarchy **priority\_hierarchy** on **location** as

$$\text{Level}_1: \text{high\_priority} < \text{level}_0: \text{all}$$
$$\text{If} ( \text{speed} > 70 \ \&\& \ \#\text{occurrences} > 20 )$$
$$\text{Level}_1: \text{medium\_priority} < \text{level}_0: \text{all}$$
$$\text{If} ( \text{speed} > 70 \ \&\& \ ( 8 \leq \#\text{occurrences} \leq 20 )$$
$$\text{Level}_1: \text{low\_priority} < \text{level}_0: \text{all}$$
$$\text{If} ( \text{speed} > 70 \ \&\& \ ( \#\text{occurrences} < 8 )$$

This priority hierarchy rule would cluster the locations into high, medium and low priorities from which the user can determine which location has sensor installation priority.

In addition to having hierarchies, they could have confidence measures, and relaxation constraints attached to them. A user thereby could either tighten or loosen the relaxation constraints depending on the data mining task. In addition, confidence measures could give more knowledge as to how good or vague the results are for that particular mining task. Data mining could also be more user driven by interactive multilevel mining whereby in each level the user could use any of the hierarchical types mentioned above which would enhance performance of the mining task at each level.

### **2.4.3 Data Mining the WINS Data Warehouse to Extract a Training Set**

One of the main functions of the WINS system is to identify and classify events autonomously. This would require a training data set for learning purposes from the existing data warehouse that would be extracted by data mining. The training set would not only have attributes and information existent in a typical

information gathering node, but would also contain previously known data or information that would not be present on a real-time implemented WINS system.

A look at the vehicle monitoring WINS system would reveal particular information present on a training set and not in a typical system. The defense community would like to track vehicular movements to characterize them (e.g. as threat or not). Weight, wheel type, and speeds help to characterize the event. Therefore, data generalization and attribute removal from an existing database helps to build a training set. The information of a vehicle weight could be obtained from parsing a training set information source, by trying to define the vehicle weight category as Heavy or Light from the vehicle\_ID attribute. A threshold of  $tw$  (weight threshold) could be given a value for a particular application and could be characterized as

Heavy Vehicle := When Vehicle\_ID  $\rightarrow$  weight  $>$   $Wt$

Light Vehicle := When Vehicle\_ID  $\rightarrow$  weight  $<$   $Wt$

$Wt = 5$ tons for the example given below on Tables 2.1 & 2.2
---

Additionally, a vehicle base structure would be important to have since it is known that vehicles with wheels give a different waveform than those with chain

tracks in the nature of the seismic and acoustic waves produced by moving sources. This signal processing knowledge would be used to classify whether a vehicle is wheeled or tracked. Further, signal waveforms are terrain dependent. Therefore, terrain knowledge is important since the detection process is environment dependent. Terrain knowledge could be mined by mapping the sensor location to the terrain. (i.e: desert, arctic, normal, hilly etc...)

A blind learning algorithm would not need the exact vehicle but would have a vehicle Id which could be named as

M-55	:	Heavy	:	Tracked	→	HT1
M-56	:	Heavy	:	Tracked	→	HT2
T-9	:	Light	:	Wheeled	→	HW1
Tata-12	:	Light	:	Wheeled	→	LW1
Isuzu-6	:	Light	:	Wheeled	→	LW2

**Table 2.1 Generalized WINS Data Set**

Here the generalization has been done having prior knowledge of the behavior of the waveforms from signal processing. Thus, these generalizations and attribute

Vehicle Type	Speed	Location	Data File Name
HT1	Slow	Desert	df2033.dat
HT2	Slow	Desert	df2045.dat
HT1	Slow	Desert	df2022.dat
HW1	Slow	Normal	df3003.dat
HW2	Slow	Normal	df2526.dat
LW1	Fast	Normal	df0034.dat
LW2	Fast	Normal	df0035.dat
LW3	Fast	Normal	df0036.dat
:	:	:	:

**Table 2.2 Derived relations with generalization of WINS data set.**

induction helped in clustering data into classes that would now be used as a training set for autonomous identification and classification purposes. Tables 2.1 and 2.2 shows the derived relations with generalization and attribute removal for a WINS vehicle monitoring ‘training’ data set. Note that this table has only 4 fields compared to more fields as observed in the original data set.

The previous table gives an example of vehicle characterization for a training dataset obtained by mining the WINS data warehouse. The use of expert knowledge of signal processing and military intelligence was needed in arriving at this generalized, attribute removed data set. Similar procedures could be used to derive application specific data sets for learning task specific data sets.

## **2.5 Decision Support System Overview**

Necessary precautions, constraints and planning needs to be specified before building a data warehouse for application specific systems. [38] Data mining of the built data warehouse would require expert knowledge and user inputs in extracting the data needed to apply in specific applications for learning and classification purposes. Here the WINS data warehouse and WINS data mining techniques were discussed at arriving at a training data set for classification and identification purposes. This derived data set is presently used in many academic institutions and defense research labs to study vehicle monitoring and intelligent information gathering. Though the WINS technology is not fully developed yet, the mentioned system level specifications would build the core of the WINS data warehouse and decision support systems.

## **2.6 Distributed vs. Centralized Signal Processing / Collaborative Decision Making**

Sensor systems that collaborate with data and decisions give superior results in noisy conditions and abnormal situations. A distributed decision making architecture is hence explored for collaborative decision making on different architectures related to sensed signal processed decisions. [8,12] Benchmarking of the accuracy levels are done with that of the MUSIC and Pisarenko parametric methods and wavelet method. Results indicate that collaborative decision-making enhances decision accuracy especially during abnormal environmental and noisy conditions.

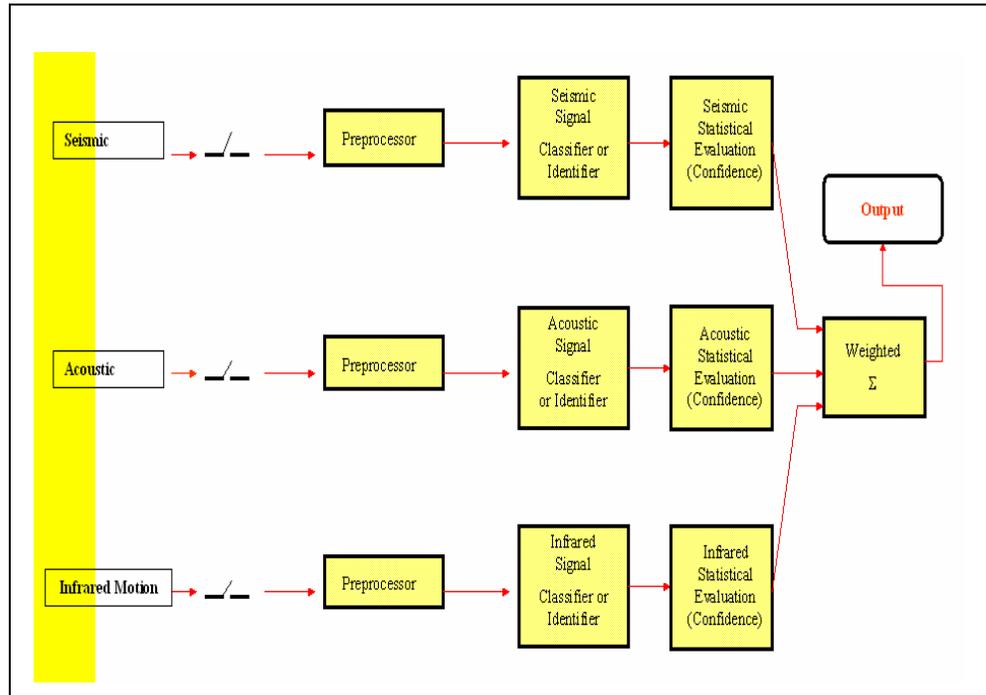
## **2.7 Signal Classification / Identification Architectures**

Many detection schemes have been researched in the area of classifying and identifying moving vehicle signatures. Neural networks, deconvolution and source separation, and wavelet methods are a few of the promising methods. [63,70,100] The time domain signal search engine approach has been to classify and identify moving sources with the end goal being the same as those of the above mentioned methods concentrating more on low power signal processing.

Moving source waveforms were studied with wavelet methods and is used to benchmark the time domain SCIA with real world signals using the ‘ACIDS’ database. [92]

## **2.8 System Level Architectural Algorithm Description**

We now give an overview of the classification / identification approach to be explained in detail in subsequent chapters. Wireless sensor networks require a robust SSE that can accommodate numerous consumer, and client/user applications. Classification / identification of moving sources, traffic monitoring, intruder detection along perimeters of buildings and other similar consumer applications with WINS require robust signal pre-processing or conditioning based on application needs. Signal pre-processing or conditioning is included in the SSE architecture, and is a vital module that is application specific. The following diagram gives a system level architecture of the WINS SSE. The system is divided into a multitude of system and sub-system blocks or modules, each with a specific function and algorithm as explained in the SSE system architecture block level description below. The division of sub-systems reduces the number of signal variables while achieving low power signal processing. This approach also enhances accuracy levels while providing modularity for decision making methods.



**Figure 2. 12 M-SSE system level blocks and modules.**

A system level diagram of an M-SSE for multi-sensing nodes is given in Figure 2.12. Multi-sensing nodes provide acoustic, seismic, and infrared signal waveforms. A micropower event detector awakens the preprocessors, which segments and feeds signals containing unique features while discarding low SNR and featureless waveforms. The preprocessing block is critical in that it minimizes the computational power of the classifier / identifier block. Preprocessed signals enter each of its classifier / identifier block for classification (generalization) and identification (specialization). The independent classifier trees are built with the help of training data sets in TAH and super-TAH structure.

Separately classified events are then combined (fused) with their relative confidence. Combining classification results are done by previously chosen decision making methods. The methods effectively incorporate individual type based or segmentation based result fusion. Output is then sent to the user with a fused confidence measure.

## **2.9 Signal Search Engine**

The signal search engine performs intelligent processing of sensed signal waveforms to detect a threat or an event. [65, 85] An event is then classified and / or identified using the time domain algorithm embedded on WINS nodes. This mode of local processing at the node drastically reduces the power required for wireless transmission of complete time series data sets. However, this architecture places demands on the capability of local, low power signal processing and event detection and identification. Independent processing relieves networking complexities and enhances scalability by limiting communication requirements.

A system design of the Multi-signal search engine (M-SSE) is given in Figure 2.12. It incorporates multiple signal types in the form of acoustic, seismic, and infrared waveforms obtained from multi-sensing nodes. The M-SSE is more robust compared to the single signal search engine (S-SSE) in that a choice could

be made to include or exclude any signal type depending on signal to noise ratio (SNR) and environmental conditions. An initial implementation of the M-SSE uses independent classifier tree structures for different types of signal waveforms. Therefore, the tree structure built for classification and identification is independent in that each signal type traverses its unique tree structure. Once the results are obtained, a confidence measure is given for each type of signal and combined by weighted averaging to output the event.

## **2.10 Dynamic System and Reconfigurability**

The SSE architecture employs a signal correlation engine that operates on incoming, unidentified input signals with a library of stored signal waveform templates (referred to as "correlators"). [85] Template selection is completed by extracting distinct features from previously acquired data. Once the template correlator library is formed it is arranged in blocks or trees to group similar event classes together. A hierarchical type abstraction hierarchy (TAH) is considered in building this tree structure. This preparation may be completed prior to deployment of the wireless sensor node or remotely uploaded after deployment. With this done, the SSE and the sensor node are ready for real-time classification and identification.

For any distributed signal processing system, it is required that the system be modified whenever new events, previously unknown, occur. For the SSE, this modification is conveniently accomplished through template additions. Apart from new events occurring, users of the SSE may want to classify previously indexed events with new classification algorithms. Further, new indexing schemes may be needed depending on environmental changes in the localities of WINS nodes. When new events occur or templates are added, the tree structure is rearranged in the critical node or rebuilt altogether on off-node computing platforms. These examples introduce the importance of the dynamic system specification needed for the SSE.

The SSE classification algorithm therefore could be implemented on application specific processors or on field programmable gate array (FPGA) modules. [36] Dynamic programmability features existent in these modules makes them the best choice for sensor nodes.

## **2.11 Multi-Signal Search Engine (M-SSE)**

Multi-signal search engine concept was taken since there is a need to make decisions based on the various types of sensing present in integrated sensors. Collaborative decisions obtained from multi-type sensors enhance accuracy levels of the sensed signals while giving control of the obtained decisions. Modular control is obtained during decision fusion within a sensor or at the cluster-head.

Various control features inherent in decision fusion architectures are shown in Chapter 3 of this dissertation.

An implementation of M-SSE with acoustic and seismic signals gave better accuracy levels compared to the S-SSE (Single-Signal Search Engine). Signals that had saturated, or unidentifiable / unclassifiable signal segments (state-spaces) attained further robustness. This was achieved by replacing unworthy state-spaces with that of the different type state-space. Environment and circuit issues were overcome with the M-SSE that is robust to variations of signal parameters, while giving results that are more accurate.

## **2.12 Multi-Signal Detection and Signal Evolution**

The application of the SSE to multiple sensor data streams or to the time evolution of sensor signals offers the possibility of enhanced identification accuracy. With continued success for SSE identification with signal classification and identification the SSE has expanded into multisensor signal analysis.

Evaluation of SSE operation has also demonstrated the ability for the SSE to resolve signal evolution. Signal evolution behavior has been examined experimentally for seismic and acoustic vehicle data sets. Examination of the time series waveforms for vehicle motion reveals that these waveforms evolve

during the periods when a moving vehicle is approaching the sensor system, during the period of closest approach (CPA), and during the departure phase.

The SSE system was applied to the identification of vehicle seismic and acoustic signal waveforms for signals obtained during these periods. It is a goal that these data sets would be combined to provide enhanced measurement and signal identification opportunities for vehicle signals. As a first step, segments of the approach, CPA, and departure phase of the vehicle were considered to form a super TAH similar to the longitude and latitude combination used on COBASE implementation. [27] A combination of these segments not only gave enhanced performance but also excluded the constraint that all three segments need to be extracted for classification. An example scenario would be when a vehicle stops before passing the sensor whereby not giving the departure segment or a vehicle begins moving after a stop near to the sensor whereby not giving the approach segment. The use of super TAHs effectively reduced constraints to having at least one segment rather than the need to have all signal segments for classification / identification with the SSE.

## **2.13 SSE Implementation**

The SSE correlator template operates as a matched FIR digital filter. The complete correlator library forms a complete filter bank. [67, 68] Correlators are

short data time series segments extracted directly from field data. The SSE operates on unknown signals by forming an inner product between each data point of a correlator and the data set (or properly sub-sampled versions) for each correlator in the correlator library. By translating the correlator along the unknown signal time series, a new waveform is generated. The RMS value of this waveform, computed over a window, can be treated as a scoring value. The application of a family of correlators to the unknown data set provides a scoring spectrum and permits classification or identification. Due to the limitations or lack of models for the generation and propagation of seismic and acoustic signal sources, the SSE is developed to rely on actual data acquired in the field for its identification codebook. The SSE development is directed to exploiting the variability of terrain and condition to identify a signal source as well as its operating location. With information on operating location present, effective beamforming techniques could be applied for fusion of data between closely located sensor arrays therefore enhancing the SNR of the unknown signal waveforms.[22]

## **2.14 Summary**

This chapter analyzed data and decision communication costs to that of battery life and other communications criteria as bandwidth, and distance for low power

operation and longevity of the wireless sensor system. The tradeoff of communication vs. computation is essential for battery power constrained wireless sensor nodes. Battery power determines sensor lifetime and is crucial for microsensors that are remotely deployed, due to re-deployment problems. The basis for the time domain SSE was set with an analysis of competing classification and detection schemes. An application of WINS classification and identification was explained for decision support systems with the use of data warehousing and traffic monitoring.

The SSE concept was formulated with the system diagram for triggering, classification and identification, and decision support for WINS systems. The need for M-SSE was explained in detail with an emphasis on multi-sensing sensors. SSE implementation was discussed with a system level description of the blocks present for identification and classification. These explanations form the basis for this dissertation in the form of the time domain signal search engine ( SSE ).

## CHAPTER 3

# Distributed vs. Centralized Signal Processing and Decision Making

### 3.1 SSE Background

The Signal Search Engine (SSE) is implemented to accommodate either distributed or centralized signal processing and decision-making depending on application needs [2,33,34,59,64,101,102]. A thorough system architecture study and description enable optimal implementation of SSE (Signal Search Engine) or M-SSE (Multi-Signal Search Engine) for distributed and centralized signal processing and decision-making. In this chapter, different system architectures for signal processing, and decision-making are considered. Signal processing is divided into two modules:

1. Signal Pre-Processing: Signal conditioning as required for SCIA.
2. Signal Classification / Identification Algorithm (SCIA): Time domain algorithm for Classification / Identification.

Each module has its own variant depending on the selected SSE / M-SSE architecture.

The decision-making architecture follows the SCIA architecture [47], also separated into two modules:

1. Decision Methods: Maximum Polling, Weighted Averaging.
2. Statistical Confidence Measure: Assignment is dependent on the weighting scheme and decision making architecture.

This chapter begins with a description of different application specific robust implementation requirements, and compares the computational power for distributed vs. centralized signal processing for SSE / M-SSE.

### **3.2 Distributed (Localized) vs. Centralized Signal Processing for SSE**

Signal processing for SSE and M-SSE is separated into two modules. Initially acquired signals go through a signal pre-processing step, where various signal-conditioning operations are performed in order for the raw signal to be in an acceptable form for the SSE / M-SSE's SCIA module. Signal pre-processing is an extremely critical functional module of SSE that strongly influences accuracy and optimizes the SCIA algorithm, while reducing signal parameter variability in the acquired signals. Hence, the signal pre-processing module is vital, enabling the SSE algorithm to be applicable to wideband signals. Figure 3.1 below shows a system level view of SSE relative to the RF transceiver of the sensing node. Received signals go through analog and digital

processing after which the SSE follows. The SSE module is separated into signal pre-processing / conditioning module and SCIA and decision making modules. Once signal pre-processing is complete, segmented signals are input to the SSE's SCIA module.

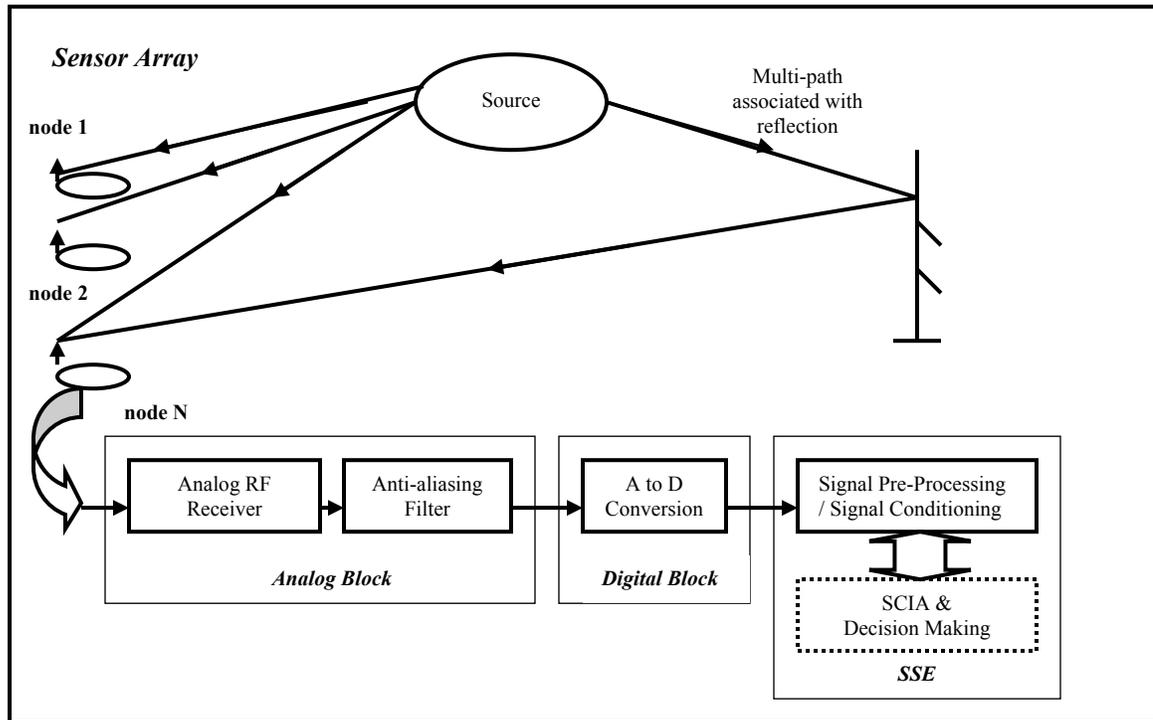


Figure 3. 1. System diagram of SSE relative to sensor node transceiver.

The SCIA module contains a pre-assigned template tree with associated confidence on each leaf that is obtained by the process of template tree building using a training set. The SCIA module is independent of the pre-processing module and therefore could be architecturally separate.

In Figure 3.2 the signal processing architecture shows both the signal pre-processing and SCIA resident in local nodes within each cluster. Signals sensed at individual nodes go

through signal pre-processing and signal classification / identification modules using a localized signal processing architecture. Once classified or identified, decisions are transmitted wirelessly to a local cluster head that does collaborative decision making based on information received from local nodes. The above architecture is preferred due to its low communication cost, scalability, and ease of reconfigurability.

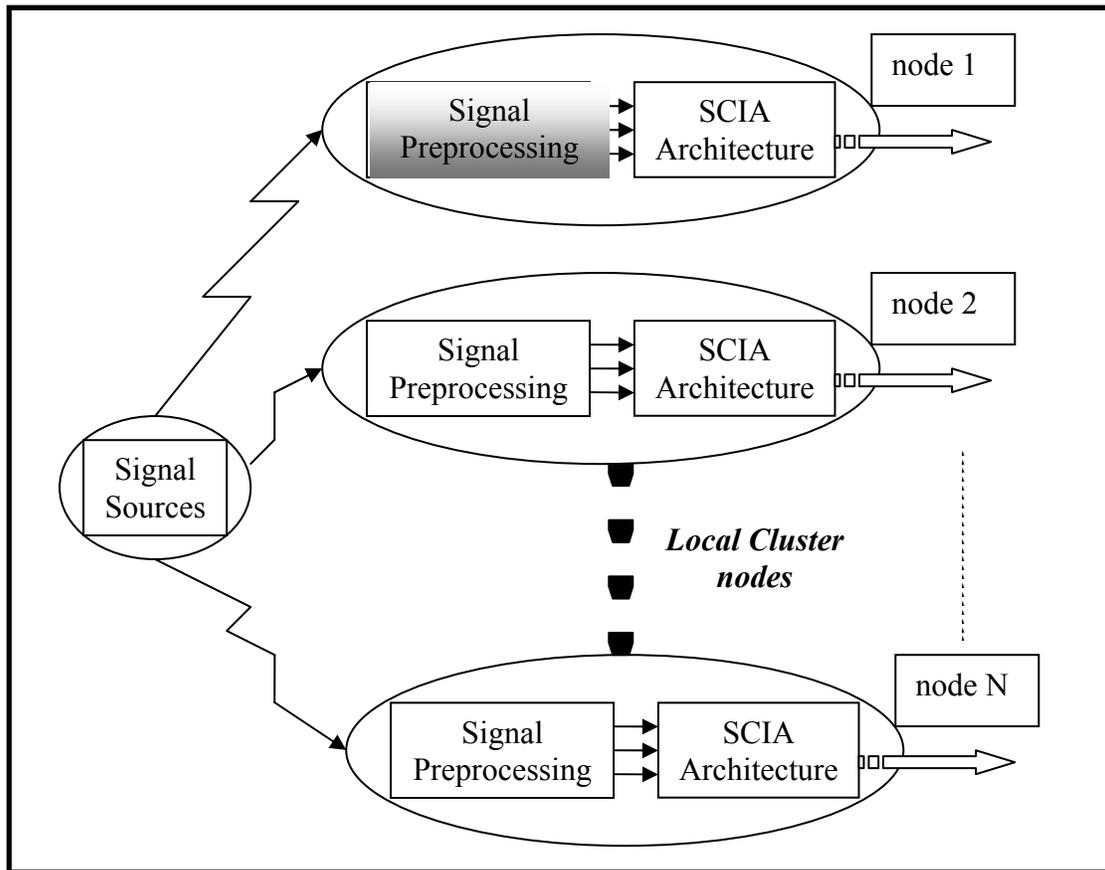


Figure 3. 2. Distributed local signal processing and SCIA for WINS.

Figure 3.3 below shows the signal processing architecture with localized signal pre-processing and local cluster head based SCIA. Distributed local nodes within each cluster sense and perform signal conditioning / pre-processing operations. Signal pre-processing is application specific and performs functions of sample down conversions, filtering, signal variable reductions, and state space decomposition. Selected signal segments are then wirelessly transmitted to a local cluster head for signal classification and identification. Cluster head based SCIA is used when data fusion

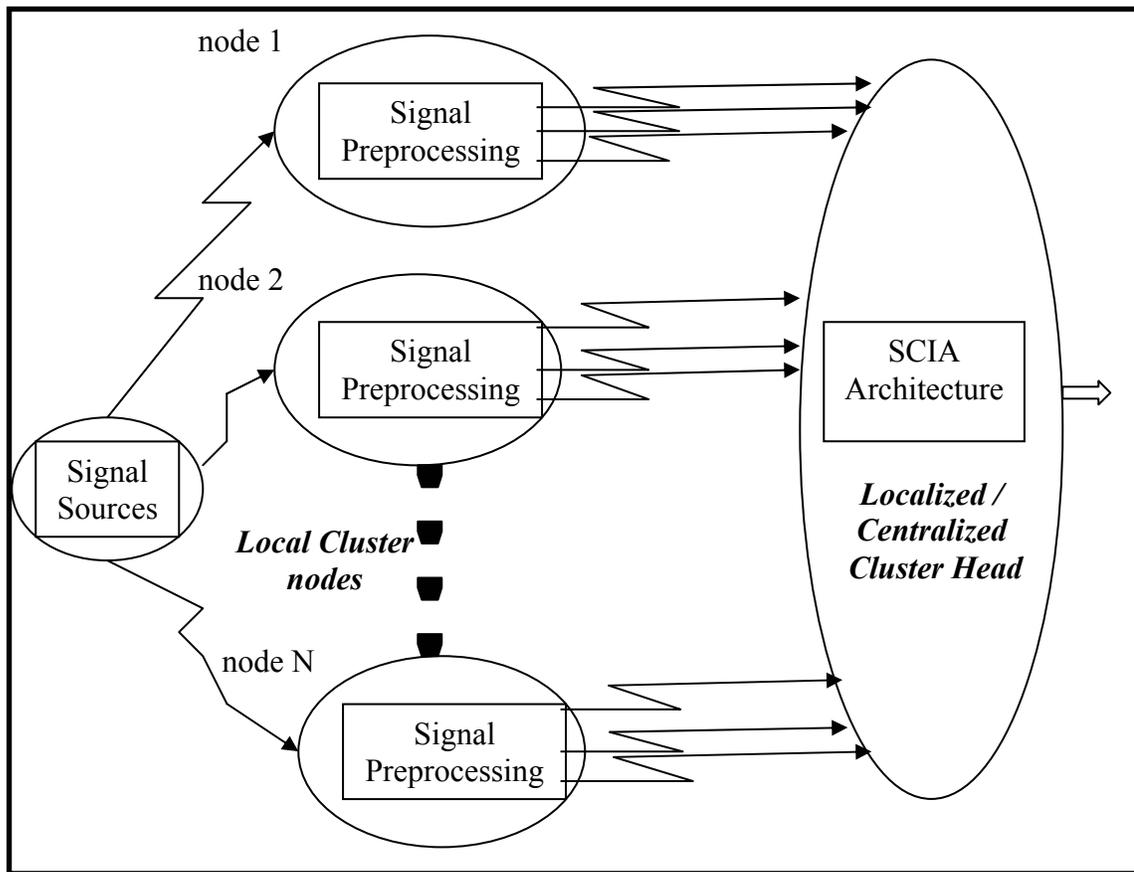


Figure 3. 3. Localized signal pre-processing and localized / centralized SCIA for WINS.

algorithms (i.e. beamforming or array processing) are incorporated for location finding or source tracking. [22] Added information, obtained during beamforming increases accuracy of the SCIA without adding wireless communication overhead. The above architecture is preferred for mobile ‘source’ tracking, identification, and classification. This architecture is concurrently used with the architecture of Figure 3.2 and is activated in critical circumstances, either to validate localized identification / classification or employed when localized identification / classification falls in the grey / overlapping area of decision boundaries. This architecture is preferred due to its ability for verification, and source tracking. However, wireless transmission costs, and high calculation costs for data fusion makes this architecture possible for selective WINS applications. [39]

Results obtained from SCIA are input into the ‘decision making’ module that consists of two parts, a statistical confidence measure assignment module, and a weighting module, each of which could be independent depending on the WINS application. Confidence measures are assigned to each classification and/or identification result, from a-priori probability knowledge obtained during training possibly from a large database. These confidence measures are updated frequently as and when more signals are acquired and re-used for training. A final module consisting of weighting schemes is incorporated to make a final weighted decision coming from individual sensor nodes or sensor node clusters.

Figure 3.4 shows the architecture for signal segmentation based localized SCIA. Decision making and confidence calculations are performed in two steps where

segmentation based decisions and confidence are fused initially in a local cluster head. This information is then either processed for collaborative decision making on the same local cluster head, or transmitted to a centralized (remote) node. Here, signals are partitioned into approach, arrival, and departure segments. This method of segment-based decision-making gives more control over segment states and eliminates low confidence decisions.

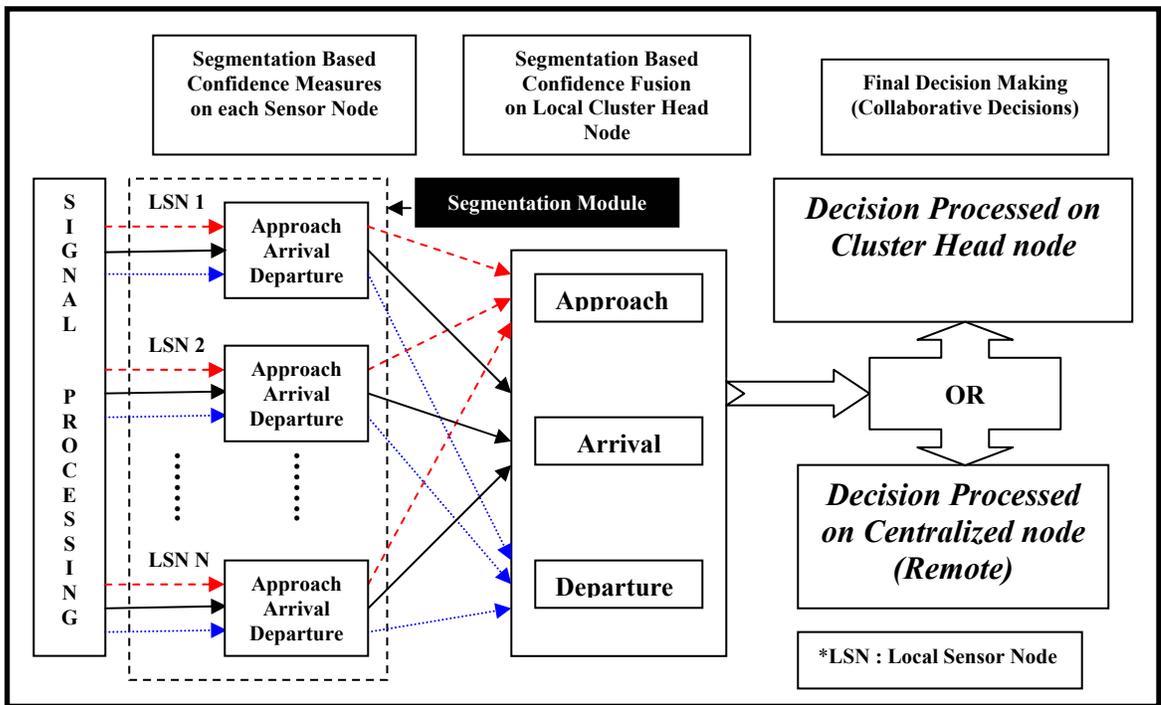


Figure 3. 4. Segmentation based localized SCIA Architecture.

Interim decisions and associated confidences are obtained on each sensor node (segmentation module). The local cluster head obtains information from the segmentation module and fuses decisions and confidence measures according to

segmental states. These decisions are then fused to derive a final decision on the cluster head or transmitted to a remote centralized node for final decision-making. Confidence measures are calculated following decision-making criteria as explained in Figure 3.7. The above architecture is robust in that segmental state control for decision-making is attained. Segments that have low confidence measures or segmental state changes that introduce new variables into signals, could be excluded in decision making (i.e. exclude departure segments if road conditions change in the departure locality (state) of the sensed signals / or if arrival segments are saturated due to circuits not being tuned properly.)

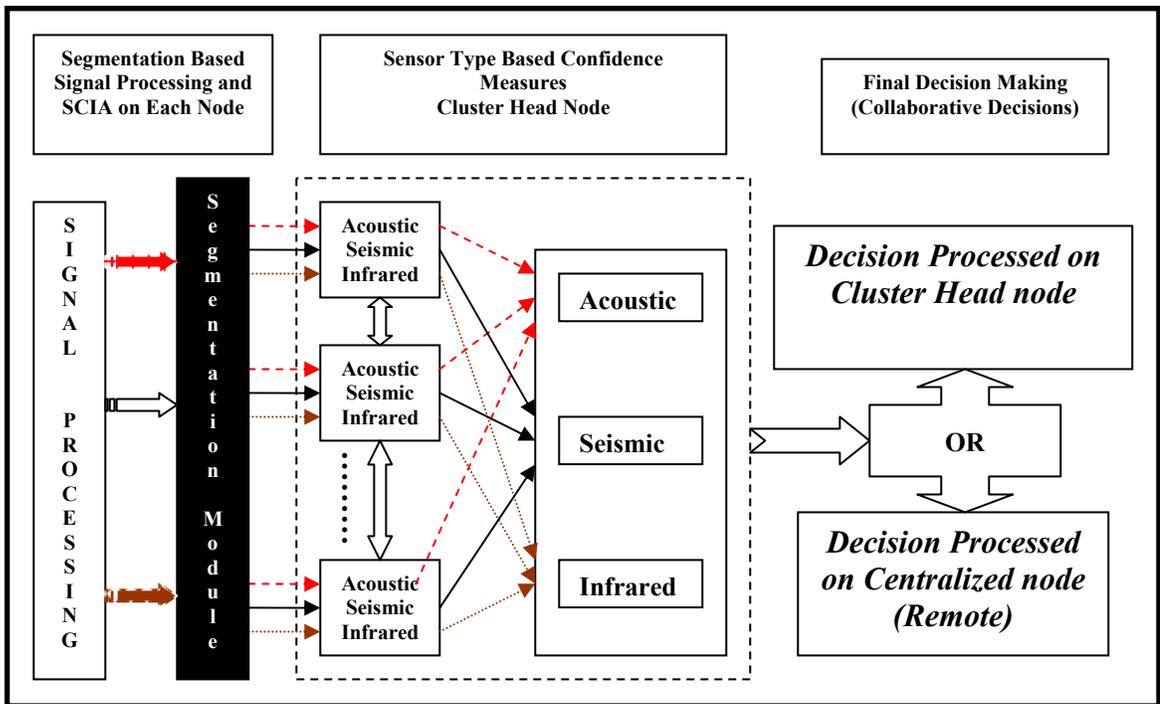


Figure 3. 5. Distributed / Localized SCIA and Sensor Type based Decision Fusion at cluster head.

Figure 3.5 shows the signal segmentation based localized SCIA architecture within the segmentation module and a sensor-type-based decision fusion within a cluster head for M-SSE. Decision making and confidence calculations are performed in three steps. Segmentation based decisions and confidence are fused initially within each node for each sensor type. Fused segmentation based signals are then transmitted to a cluster head for sensor type based decision fusion. The above diagram shows sensor type based decision fusion in the form of acoustic, seismic, and infrared signal types. This module gives control over sensor types that may have low confidence due to deteriorating environmental or geographical conditions. The sensor type based decisions and confidence information is then either transmitted for collaborative decision making to a cluster head or to a remote centralized node.

The above architecture is robust in that sensor-type based control for decision-making is achieved. Segments as well as sensor types that have low confidence measures, or segmental state changes that introduce new variables into signals and sensor types that have low confidence could be excluded from decision making (i.e. exclude acoustic sensors in high wind environments, while excluding seismic sensors during ground condition changes due to rain, melting snow etc.).

SSE / M-SSE can be scaled to distributed localized (signal processing and decisions on each sensor node), or local head node (local signal processing with decisions processed on local neighboring head-node within a geographical cluster), or centralized (local signal

processing with decisions processed on a centralized node) architectures for a combination of signal processing and decision making depending on applications. Application needs determine these architectural designs. This requires consideration of wireless communication and computing costs in addition to scalability, deployment issues, and throughput of the SSE / M-SSE system.

At this stage of processing, any special environmental conditions that may impact different types of signals (i.e. acoustic, seismic, IR) or classification / identification accuracies may be taken into account by turning particular sensor readings on / off if they have not already been acted upon much earlier during signal acquisition or pre-processing. Again, statistical confidence measures and weighting could be in the same SSE architectural block or in different blocks in the case of M-SSE. A careful analysis of application specific signal parameters, wireless communication and computational costs, accuracy, throughput, and feasibility of SSE / M-SSE will yield a suitable architecture as shown in the study given below.

### **3.3 SSE Application Architectures**

Wireless sensor network applications determine SSE / M-SSE architectures for signal processing and decision-making. Many implementations are feasible while making the SSE and M-SSE more robust, less complex, and suitable for a particular application. [47]

Sensor networks that deal with high data densities require localized (on node) signal processing, or localized signal pre-processing and local cluster head node based signal processing (due to signal fusion needs such as beam forming) while decision-making is done at a local or remote head node. Interim decisions are made at localized nodes, while final decisions are made at a local cluster head node.

Applications that require minimal data transfer and communication overhead can use centralized decision-making, as scalability and communication resource costs become less important issues. Temperature monitoring nodes embedded on walls of buildings are a classic example. Other applications that can follow this architecture include intelligent sensor nodes embedded in bridges, and biomedical applications in medical and research facilities.

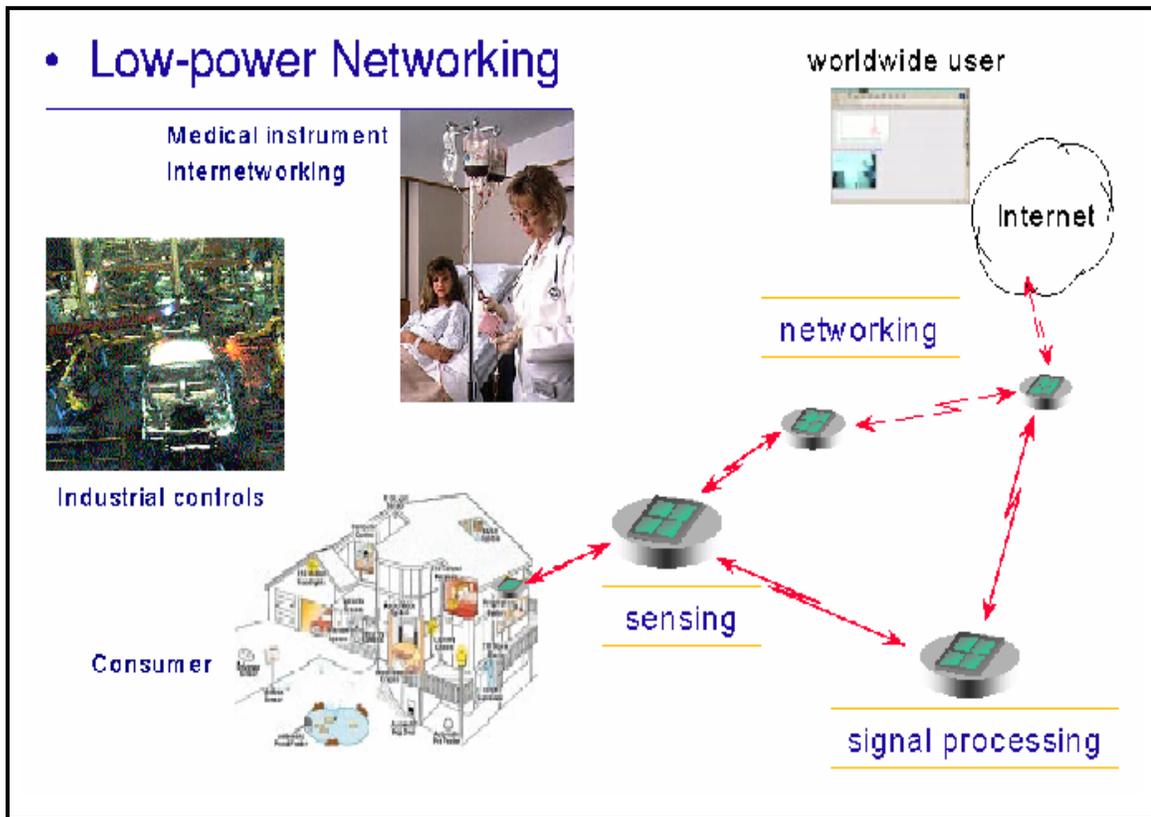


Figure 3. 6. Consumer, and industrial applications using distributed / centralized processing.

SSE system architectures are determined initially with application specific needs that make use of criteria given in section 3.4 of this chapter. Analysis done in this section helps to determine whether to implement distributed / localized or cluster head /centralized sub-systems within the SSE / M-SSE systems architecture. Figure 3.6 shows consumer and industrial applications that use distributed, and centralized processing. Specific application implementations strive to determine a simple architecture that achieves necessary performance while consuming the least power for each wireless node. Low complexity and low power operation are formulated by analysis given in section 3.4 that is done during architectural design phase.

### **3.4 Communication vs. Computation Costs**

Power consumed by wireless nodes for wireless data communication is typically much higher than the computation power requirements for identification / classification. [93] Much of the signal pre-processing and often the bulk of the signal processing is done on-board each node after data acquisition, in order to minimize wireless data communication costs. However, advanced signal data fusion algorithms require signal data to be transmitted to a cluster head node for further processing, in which case wireless communication costs become important. The following section analyzes wireless communication vs. computation costs as experienced in real world sensor network scenarios.

#### **3.4.1 Communication Cost:**

A number of methods are now detailed for reducing communication costs. These methods concentrate on optimization methods for low power, and performance.

##### ***a. Local Communication of Decisions to a Cluster Head / Centralized node:***

In this architecture, SSE / M-SSE have signal pre-processing and SCIA on-board each node. Results from each node are transmitted to a cluster head or centralized node for collaborative decision-making. A study of communication cost for transmitting interim-decisions and associated confidence is formulated and calculated for a WINS network.

We assume that sensor nodes are used to transmit decisions to a cluster head or centralized node at a distance of 10m, with the following transmission environment criteria [83]:

- Surface node to Surface node (flat terrain) transmission in Rayleigh Channel @  $1/R^4$  propagation. Calculations also assume BPSK transmission at a target bit error rate of  $10^{-6}$ .

Assume transmitted decision has 32-bit word length;

Assume Symbol Rate of 10 bits / Sample;

For 10m node separations,

Propagation Limit @ 10m: 1000 bits / 0.3 mJ;

Energy required to transmit 32 bit detection result = 9.6 nJ;

Similarly for 100m node separations,

Propagation Limit @ 100m: 1000 bits / 3 J;

Energy required to transmit 32 bit detection result = 96 mJ;

The above calculations show the energy required to transmit data to a local cluster head or centralized node, on a per-detected result basis. Different Classification / Identification schemes in SSE / M-SSE architectures would require variations of these

detected results (32-bit word lengths) to be transmitted with multiplicative factors of the calculated energy needs.

**b. Local Communication of Data to Local Cluster Head after Pre-Processing:**

Assuming that the sensor node is used to transmit signal segments to a cluster head or local centralized node 10m in separation with the following wireless communication environment criteria:

- Surface-node to Surface-node (flat terrain) transmission in Rayleigh Channel @  $1/R^4$  propagation. Calculations also assume BPSK transmission with a required error rate of  $10^{-6}$ .

Assume Sample Signal Length = X Samples \* 10 bits/Sample = 10X bits;

Propagation Limit @ 10m = 1000 bits / .3 mJ;

Energy required to transmit 1000 samples of segmented signal = 3.0 mJ;

Summarizing, evaluations of decision communication vs. data communication costs, it is observed that to transmit pre-processed data to a local cluster head based SSE / M-SSE architecture consumes high energy and is used only in critical and selective circumstances. [76] This is the case when computationally intensive data fusion algorithms are implemented and used before classification / identification by SCIA. Here a calculation for a quantity of 1000 samples is given to quantify the energy requirements as needed for wireless transmission with this design of SSE / M-SSE architecture. It is

therefore evident that communicating decisions only would prolong battery life and hence node life. Example 3.1. shows the needs of battery energy for transmitting decisions and data. [87]

**Example 3. 1. *Direct Communication of Data / Event:***

The following calculations are performed for a channel with the following propagation and transmission criteria:

- Surface Node to Surface Node, Rayleigh Channel with  $1/R^4$  propagation and BPSK transmission at a required error rate of  $10^{-6}$  bits / sample.

Sample Word Length = 32 bits; (Assumption)

Propagation Limit @ 100m: 1000 bits / 3 Joules;

Propagation Limit @ 100m: 100 samples / 3 Joules;

Number of samples transmitted per unit energy = 10.4 Samples / Joule;

### **3.4.2 Comparison of SSE / M-SSE to Communication**

It is crucial to make a comparison of communication cost to SSE / M-SSE calculation costs to decide upon the appropriate strategy. Example 3.2 looks at a sample case and compare the associated power consumption for a SSE/M-SSE implemented algorithm.

### **Example 3. 2. SSE / M-SSE Event Processing Power Cost Calculations:**

Let us look at the following SSE/M-SSE event processing calculations:

Number of Samples used in the Signal Segment : 1000 Samples

Number of Correlators parsed during a search : 100 Correlators

Number of points per correlator : 100 Points per Correlator

Number of operations per sample : 10 ops / sample

Assuming a worst case general purpose processor

Processing Rate of Processor – 100 MIPS/W and 100 MIPS Processing Rate;

Cost associated with detection / identification of this system: 1 Joule;

Now, let us calculate cost associated with direct communication of event [79,83]:

Number of Samples used for Transmission : 1000 Samples@10 bits/Sample;

Wireless Transmission Distance (wireless link): 100 m;

Cost associated for the above transmission : 30 Joules;

Example 3.2 above shows a simple calculation in energy requirement for event detection and wireless transmission. It is observed that energy required to transmit samples wirelessly is 30 times more compared to detecting a signal with the same number of samples. This simple calculation shows that data communication needs to be done only for obtaining the highest resolution, and the architecture should more usually be limited to communicating decisions only, for low power operation.

Summarizing the results of example 3.2, we derive the following for a comparison of computation to communication energy:

$\sim 10^5$  Samples Computed : 1 Sample Transmitted @ 100 m

The above case provides a notion of how transmission cost dominates overwhelmingly in comparison to computation cost. The following were assumed for qualification of the above calculations:

- Propagation calculations assume 100% efficiency for conversion from base band to radiating energy.
- Propagation calculations assume no shadowing.
- Realistic communication system analysis will further degrade communication link and further favor computation.

### **3.4.3 Accuracy of Distributed vs. Centralized SSE & M-SSE**

Accuracy of results obtained during classification and / or identification in distributed and centralized SSE, M-SSE are very similar. However, there is a slight decrease in confidence measures associated with centralized classification or identification (without data fusion), due to the phase offset carried by signal sets located at different distances from the source, and time delayed versions of source signals arriving at nodes due to multi-path. Further errors are incurred due to generalized correlator templates forming the tree structure for centralized processing. It is found that the associated probabilities

attached to each leaf correlator template contributes in forming of lower confidence measures when class based decisions are made. Other reasons for the increase in error rates are due to limited signal data in order to reduce wireless communication costs whereby only limited signal segments are obtained and transferred for centralized node processing.

The above findings were obtained using sensor data from different neighboring sensors that was excluded from the template choosing training set. Signal segments obtained from different neighboring sensors were used for segmental identification. Given the above findings, it is concluded that without any additional data fusion algorithms it is better to use distributed processing rather than centralized processing for accuracy, communication costs, and low power operation. [78]

Figure 3.7 shows a class based confidence assignment procedure for classification decisions. [7,42] The confidence assignment procedure for identification is handled similarly. An initial training data set is used for confidence measure assignment on each leaf of the tree that is remotely updated intermittently. Correlator templates are initially used for assigning confidence using the existing training data set. Once each confidence measure is obtained, it is tagged to each correlator. An example of assignment of a final confidence measure for a classification result of class  $\{B_{12}\}$  would be  $P_{b_1} * P_{b_{11}} * P_{b_{112}}$ . Similarly for identification decisions, confidence is tagged with 'source ID', with final confidence calculated as shown in the Figure 3.7. Centralized processing with data

fusion is therefore called for during some critical decision-making instances, and may be used to validate or reinforce decisions when it has a low confidence measure associated with classification / identification results.

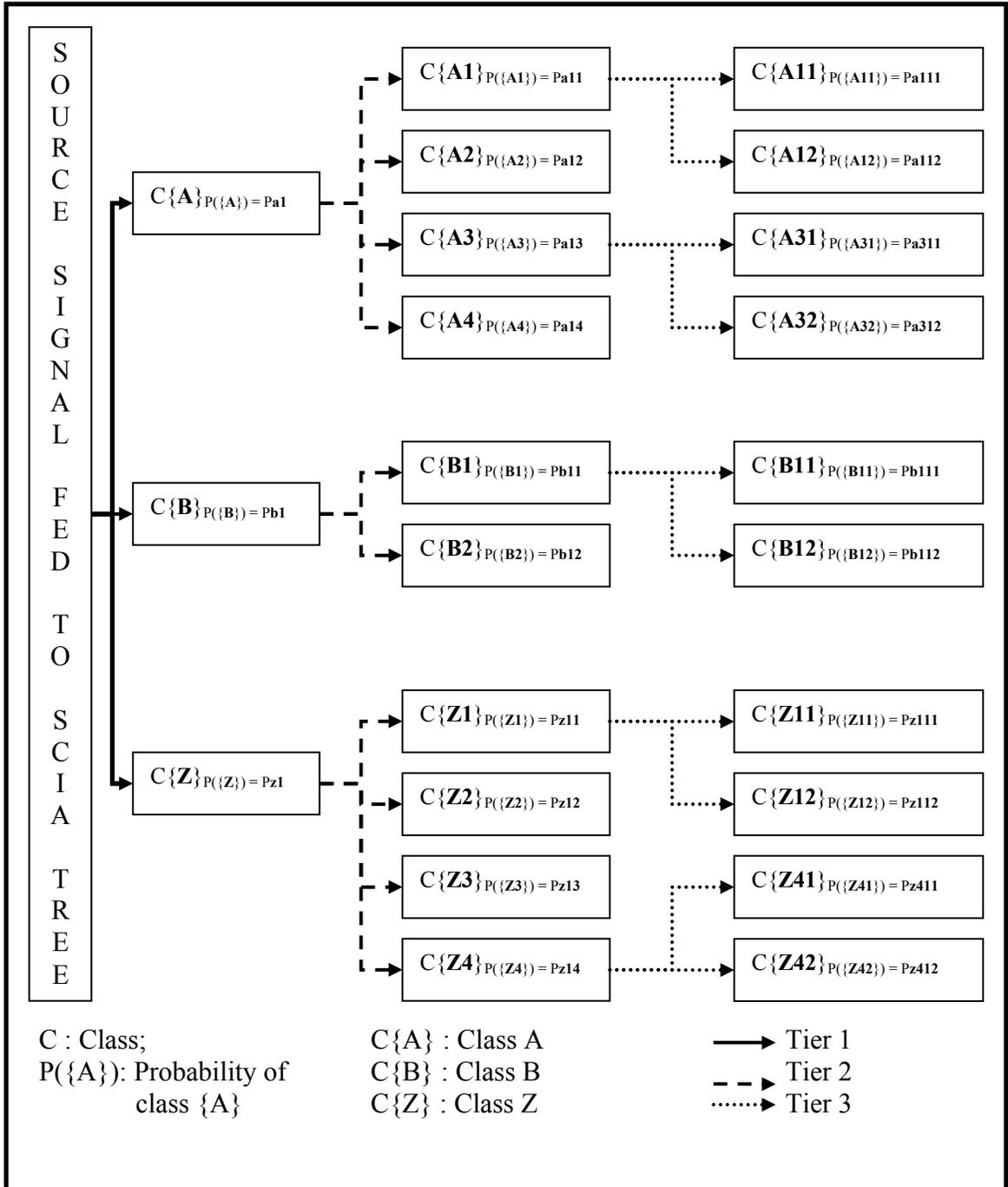


Figure 3. 7. Class based probability assignment procedure.

### **3.5 Decision Making – (Localized vs. Centralized)**

Localized decision-making is performed when the SCIA is local, with the results transmitted to a cluster-head or centralized node for final decision making with results obtained from other neighboring sensor nodes. [4, 9] Results output from the SCIA would consist of the classified / identified source or source listing, along with its confidence measures. Cluster head based or Centralized SCIA on the other hand has decisions made in their respective domain where the SCIA is resident. In centralized decision-making, a centralized node would obtain results from all individual nodes, or cluster heads of sub-systems to consider weighting of decisions.

Figure 3.8 is a system level diagram for ‘decision making’ with individual WINS nodes, cluster heads, and a remote centralized node. Decision-making may be local cluster head based with individual WINS nodes transmitting decisions and confidence to cluster heads that perform collaborative decision making to achieve final identification / classification. Alternatively, a centralized decision-making process receives sensor cluster head processed interim decisions and makes decisions and confidence measures based on received cluster head information. This module based approach helps in avoiding low confidence decisions by giving more control to avoid low confidence decision origins.

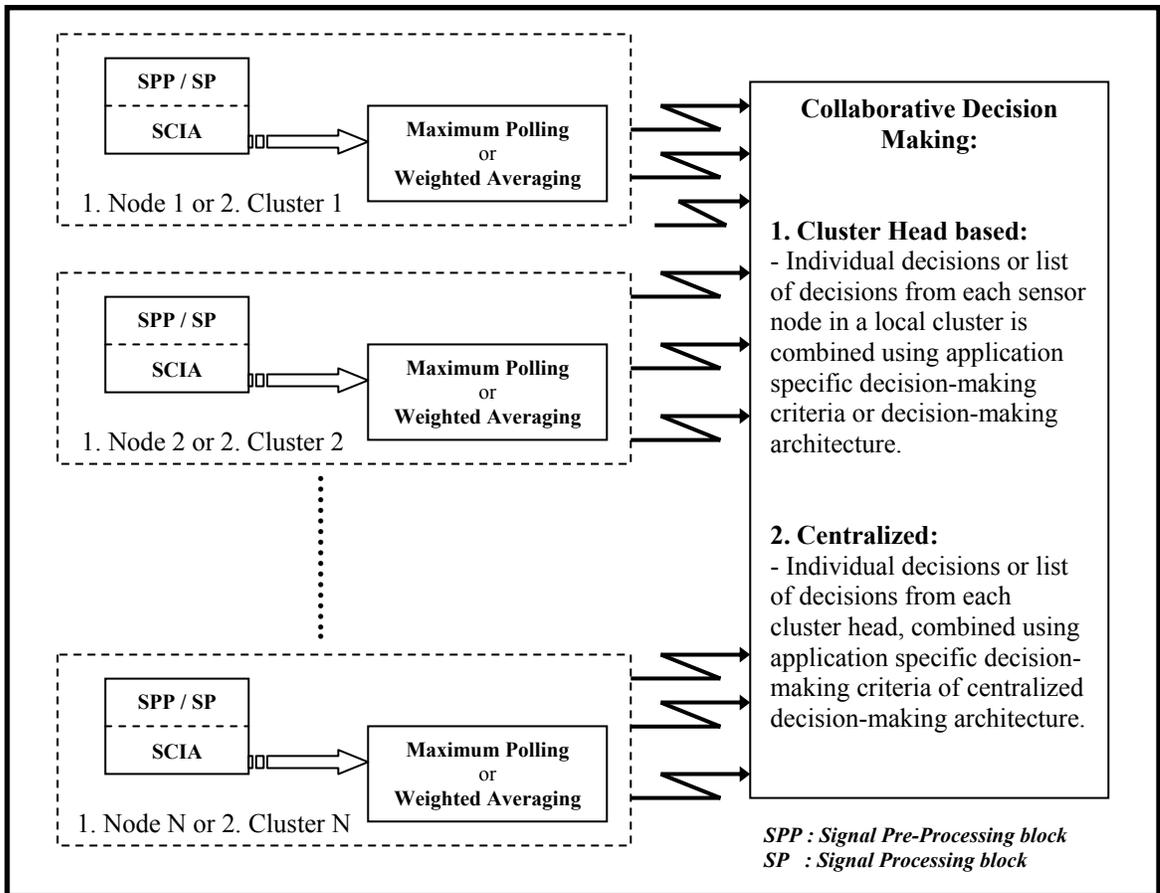


Figure 3. 8. Collaborative ‘Decision Making’ at Cluster Heads or Centralized Remote Processor.

### 3.5.1 Decision Making (Localized)

SSE / M-SSE architectures containing local, or local interim decision making have a statistical confidence assignment and weighting at each sensor node. The decision making process is similar in both cases with information passed to the localized or centralized node for final decision-making, depending on the architecture and would consist of:

- identified or classified source with its confidence (i.e. Source A – 95% confidence)

- identified or classified source order with its confidence (i.e. Source B – 35% confidence, Source C – 31% confidence, Source D – 32% confidence etc.)

A collaborative decision-making is involved at the sensor node cluster head to achieve a final result. This result is assigned weights according to one of several methods that may include for example lower weighting due to environmental effects. These scenarios happen when ground conditions are soggy on a rainy day, altering the channel of the source signal drastically for seismic signatures. In these situations, seismic identification / classification is excluded from decision-making and assigned a lower weight than in regular conditions.

### **3.5.2 Decision Making (Distributed)**

Distributed decision-making is collaborative decision making, where individual or cluster head node identification / classification results and associated confidence measure is transmitted to a cluster head, or centralized node. These results are weighted using different methods mentioned below to attain a final decision. Weighting could be assigned depending on environmental conditions as in local decision-making. The final decision is then transmitted to the user / client.

### 3.5.3 Weighting Methods for Segmentation

Various weighting methods are considered for decision making with signal segmentation. Signal segmentation is used to reduce variables of signals, especially waveforms of moving objects. The following methods are used for decision making along with obtaining a final confidence measure on one type of signal set (i.e. acoustic, seismic, IR etc).

- **Class based decision-making:** Methods where decisions are made with the same class / type (i.e. Seismic, Acoustic, and IR) of signals with different segmental states (i.e. approach, arrival, departure.)
- **Maximum Polling:** Signals of the same type are classified / identified in parallel with SCIA and would consist of single source result along with its confidence measure for each identification / classification. The method of maximum polling considers only the selected source from each segment and selects interim final, or final decision without consideration of confidence. However, confidence for the final decision is calculated independently after obtaining the maximum polling result.

Maximum Polling Source Identification / Classification

segment 1 – S(1) - Selected Source;	Confidence – $P_1(S_{(1)})$ ;
segment 2 – S(2) - Selected Source;	Confidence – $P_2(S_{(2)})$ ;
segment N – S(N) - Selected Source;	Confidence – $P_N(S_{(N)})$ ;

Decision:

$$\text{Maximum } \left\{ \sum_{i=1}^N \delta ( S(i) = Xk ) \right\}; \quad \text{for } k = 1,2,\dots,N; \quad \forall Xk \in \mathfrak{R};$$

where  $\mathfrak{R} : \{ \text{all possible events / sources for the WINS application} \};$

Confidence:

$$P( S = Xk ) = \left\{ \left\{ \sum_{j=1}^N P_j( S(j) = Xk ) \right\} / n \right\};$$

where,

$$n = \sum_{j=1}^N \delta ( S(j) = Xk ); \quad \text{for } k = 1,2,\dots,N \text{ where } Xk \in \mathfrak{R};$$

Maximum polling is implemented for decision making with signal segments acquired with the same sensor type. These segments will each have independent confidence measures that are used for calculating the final confidence of the decision. The following table shows decisions and confidence derived by using the above formulas.

	<b>S<sub>(i)</sub></b>	<b>P(S<sub>(i)</sub>)</b>
approach	tv1	0.75
arrival	wv2	0.39
departure	tv1	0.69

Decision:	tv1 (2/3)
Confidence:	0.72

**Table 3. 1. Class/Type based Maximum Polling.**

Maximum polling on a single class / type is used where 2 out of 3 segmented decisions on a signal results in an identification of ‘tv1’. Confidence is calculated as mentioned in the above with a measure of 0.72.

- **Weighted Averaging:** Weighted averaging is considered here for segmented signals acquired with the same sensor type. However, this method selects a list of sources as results for identification / classification as opposed to single source selections with maximum polling. However, confidence for the final decision is calculated with dependence on a selected list of results as output from the SCIA.

#### Weighted Polling Source Identification / Classification

segment 1 – S<sub>1</sub>(1) - Selected Source 1; Confidence – P<sub>11</sub>(S<sub>(1)</sub>);  
segment 1 - S<sub>2</sub>(1) - Selected Source 2; Confidence – P<sub>12</sub>(S<sub>(1)</sub>);  
segment 1 – S<sub>n</sub>(1) - Selected Source n; Confidence – P<sub>1n</sub>(S<sub>(1)</sub>);  
segment 2 – S<sub>1</sub>(2) - Selected Source 1; Confidence – P<sub>21</sub>(S<sub>(2)</sub>);  
segment 2 – S<sub>2</sub>(2) - Selected Source 2; Confidence – P<sub>22</sub>(S<sub>(2)</sub>);  
segment 2 – S<sub>n</sub>(2) - Selected Source n; Confidence – P<sub>2n</sub>(S<sub>(2)</sub>);  
segment N – S<sub>1</sub>(N) - Selected Source 1; Confidence – P<sub>N1</sub>(S<sub>(N)</sub>);  
segment N – S<sub>2</sub>(N) - Selected Source 2; Confidence – P<sub>N2</sub>(S<sub>(N)</sub>);  
segment N – S<sub>n</sub>(N) - Selected Source n; Confidence – P<sub>Nn</sub>(S<sub>(N)</sub>);

Decision:

$$Maximum \left\{ \sum_{i=1}^n \sum_{j=1}^N \{ \delta (S_i(j) = Xk) * P_{ji}(S_{(j)}) \} \right\};$$

*for k = 1,2,...,N;  $\forall Xk \in \mathfrak{R}$ ;*

where  $\mathfrak{R}$  : { all possible events / sources for the WINS application };

Confidence:

$$P(S = Xk) = \left\{ \left\{ \sum_{j=1}^N \sum_{l=1}^n P_{jl}(S_{\theta}) * \delta(S_l(j) = Xk) \right\} \right\} / n;$$

where,

$$n = \sum_{j=1}^N \sum_{l=1}^n \delta(S_l(j) = Xk); \text{ for } k = 1, 2, \dots, N \text{ where } Xk \in \mathfrak{R};$$

Table 3.2 shows an example of segmentation based weighted averaging. Here the top three results are selected for averaged decision-making. Decisions achieved from the above equation results in an identification of ‘tv1’ with a confidence measure of 0.60.

	S <sub>(i)</sub>	P(S <sub>(i)</sub> )
approach	tv1	0.75
	tv2	0.41
	tv3	0.21
arrival	wv2	0.39
	tv1	0.35
	tv3	0.11
departure	tv1	0.69
	wv2	0.31
	tv2	0.22

Decision:	tv1
Confidence:	0.60

**Table 3. 2. Class/Type based Weighted Averaging.**

### 3.5.4 Weighting Method for Different Classes of Signals

Similar to weighting methods used for decision making with the same type of signals, acquired signals coming from different types / classes of sensors are combined for decision making with methods shown below along with calculation of associated confidence measures.

- **Segment based decision making:** Principle where decisions are made with different classes of signal on an identical segment basis.

The above principles are derived and shown for maximum polling and weighted averaging for node-based decision-making.

- **Maximum Polling:** Signals of different types are classified / identified in parallel with SCIA. It results in a single source along with its confidence measure for identification / classification for each segment and class. The method of maximum polling considers only the selected source for each class segment and selects interim final, or final decision with confidence calculations to follow soon. However, confidence for the final decision is calculated independently for maximum polling.

Maximum Polling Source Identification / Classification

Type 1, Segment 1 –  $S(11)$  - Selected Source; Confidence –  $P_{11}(S_{(1)})$ ;

Type 2, Segment 1 –  $S(21)$  - Selected Source; Confidence –  $P_{21}(S_{(2)})$ ;

Type M, Segment 1 –  $S(M1)$  - Selected Source; Confidence –  $P_{M1}(S_{(M)})$ ;

Decision for Segment 1:

$$\text{Maximum} \left\{ \sum_{i=1}^M \delta ( S(i1) = Xk ) \right\}; \text{ for } k = 1,2,\dots,N; \quad \forall Xk \in \mathfrak{R};$$

where  $\mathfrak{R} : \{ \text{all possible events / sources for the WINS application} \};$

Confidence:

$$P(S = Xk) = \left\{ \frac{\left\{ \sum_{j=1}^M P_{j1}(S_{\theta}) * \delta (S(j1) = Xk) \right\}}{n} \right\};$$

where,

$$n = \sum_{j=1}^M \delta ( S(j1) = Xk ); \text{ for } k = 1,2,\dots,N \text{ where } Xk \in \mathfrak{R};$$

	<b>S<sub>(i)</sub></b>	<b>P(S<sub>(i)</sub>)</b>
acoustic	tv8	0.77
seismic	tv8	0.63

Decision:	tv8 (2/2)
Confidence:	0.70

**Table 3.3. Segmentation based Maximum Polling.**

Table 3.3 shows maximum polling for different classes / types of signals on a single segment. All classes / types result in identifying ‘tv8’ with a confidence of 0.70 calculated using the above method.

Decision for Segment 2:

$$\text{Maximum } \left\{ \sum_{i=1}^M \delta ( S(i2) = Xk ) \right\}; \quad \text{for } k = 1,2,\dots,N; \quad \forall Xk \in \mathfrak{R};$$

where  $\mathfrak{R} : \{ \text{all possible events / sources of the WINS application} \};$

Confidence:

$$P( S = Xk ) = \left\{ \left\{ \sum_{j=1}^M P_{j2}(S_{(j)}) * \delta ( S(j2) = Xk ) \right\} \right\} / n;$$

where,

$$n = \sum_{j=1}^M \delta ( S(j2) = Xk ); \quad \text{for } k = 1,2,\dots,N \text{ where } Xk \in \mathfrak{R};$$

Table 3.4. shows maximum polling on a different segment. Here results are tied with an acoustic identification resulting in ‘tv8’ and the seismic identification resulting in ‘wv6’. In cases where maximum polling results in a tie, confidence measures are inferred for decision-making. Acoustic identification of ‘tv8’ resulting in a higher confidence (confidence 0.79) is chosen as the decision over ‘wv6’ (confidence of 0.6). If ties are frequent occurrences during training, weighted averaging is utilized for the particular decision making sub-system.

	<b>S<sub>(i)</sub></b>	<b>P(S<sub>(i)</sub>)</b>
acoustic	tv8	0.79
seismic	wv6	0.66

Decision:	tv8 (1/2)
Confidence:	0.79

**Table 3. 4. Segmentation based Maximum Polling with a tie.**

Decision for Segment N:

$$\text{Maximum } \left\{ \sum_{i=1}^M \delta ( S(iN) = Xk ) \right\}; \quad \text{for } k = 1,2,\dots,N; \quad \forall Xk \in \mathfrak{R};$$

where  $\mathfrak{R} : \{ \text{all possible events / sources for the WINS application} \};$

Confidence:

$$P(S = Xk) = \left\{ \left\{ \sum_{j=1}^M P_{jN}(S_{ij}) * \delta ( S(jN) = Xk ) \right\} \right\} / n;$$

where,

$$n = \sum_{j=1}^M \delta ( S(jN) = Xk ); \quad \text{for } k = 1,2,\dots,N \quad \text{where } Xk \in \mathfrak{R};$$

Maximum polling in this case is considered in signals where the signal types are different but segments are identical (i.e. approach, arrival, departure) with confidence measure calculations following soon.

- **Weighted Averaging:** Weighted averaging is considered here for segmented signals of different types. This method selects from a list of selected sources (results of SCIA) for identification / classification decision making relative to individual selections with maximum polling. However, confidence for the final decision is calculated with dependence to selected list of results from SCIA.

## Weighted Polling Source Identification / Classification

Type 1; Segment 1 –  $S_{11}(1)$  - Selected Source 1; Confidence –  $P_{11}(S_{(1)})$ ;

Type 1; Segment 1 -  $S_{11}(2)$  - Selected Source 2; Confidence –  $P_{11}(S_{(2)})$ ;

Type 1; Segment 1 –  $S_{11}(n)$  - Selected Source n; Confidence –  $P_{11}(S_{(n)})$ ;

Type 2; Segment 1 –  $S_{21}(1)$  - Selected Source 1; Confidence –  $P_{21}(S_{(1)})$ ;

Type 2; Segment 1 -  $S_{21}(2)$  - Selected Source 2; Confidence –  $P_{21}(S_{(2)})$ ;

Type 2; Segment 1 –  $S_{21}(n)$  - Selected Source n; Confidence –  $P_{21}(S_{(n)})$ ;

Type M; Segment 1 –  $S_{M1}(1)$  - Selected Source 1; Confidence –  $P_{M1}(S_{(1)})$ ;

Type M; Segment 1 –  $S_{M1}(2)$  - Selected Source 2; Confidence –  $P_{M1}(S_{(2)})$ ;

Type M; Segment 1 –  $S_{M1}(n)$  - Selected Source n; Confidence –  $P_{M1}(S_{(n)})$ ;

Decision:

$$\text{Maximum} \left\{ \sum_{i=1}^M \sum_{j=1}^n \{ \delta ( S_{i1}(j) = Xk ) * P_{i1}(S_{(j)}) \} \right\};$$

*for*  $k = 1, 2, \dots, N; \quad \forall Xk \in \mathfrak{R};$

where  $\mathfrak{R} : \{ \text{all possible events / sources for the WINS application} \};$

Confidence:

$$P(S = Xk) = \left[ \left\{ \sum_{j=1}^M \sum_{l=1}^n P_{j1}(S_{(l)}) * \delta (S_{j1}(l) = Xk) \right\} \right] / n;$$

where,

$$n = \sum_{j=1}^M \sum_{l=1}^n \delta ( S_{j1}(l) = Xk ); \quad \text{for } k = 1, 2, \dots, N \text{ where } Xk \in \mathfrak{R};$$

	$S_{(i)}$	$P(S_{(i)})$
acoustic	tv8	0.77
	wv2	0.66
	tv3	0.63
seismic	tv8	0.63
	tv3	0.44
	wv4	0.32

Decision:	tv8
Confidence:	0.70

**Table 3. 5. Segmentation based weighted averaging.**

Weighted averaging is utilized on acoustic and seismic signals with a listing of the best three results. Using decision-making method shown above a decision of ‘tv8’ is obtained with a confidence of 0.70.

Similarly, for the case of Segment 2;

Decision:

$$Maximum \left\{ \sum_{i=1}^M \sum_{j=1}^n \{ \delta ( S_{i2}(j) = Xk ) * P_{i2}(S_{(j)}) \} \right\} ;$$

*for*  $k = 1,2,\dots,N; \quad \forall Xk \in \mathfrak{R}$

where  $\mathfrak{R} : \{ \text{all possible events / sources for the WINS application} \};$

Confidence:

$$P(S = Xk) = \left\{ \left\{ \sum_{j=1}^M \sum_{l=1}^n P_{j2}(S_{(l)}) * \delta(S_{j2}(l) = Xk) \right\} / n \right\};$$

where,

$$n = \sum_{j=1}^M \sum_{l=1}^n \delta(S_{j2}(l) = Xk); \text{ for } k = 1, 2, \dots, N \text{ where } Xk \in \mathfrak{R}$$

A general case for Segment N is given by;

Decision:

$$\text{Maximum } \left\{ \sum_{i=1}^M \sum_{j=1}^n \{ \delta(S_{iN}(j) = Xk) * P_{iN}(S_{(j)}) \} \right\};$$

for  $k = 1, 2, \dots, N; \quad \forall Xk \in \mathfrak{R};$

where  $\mathfrak{R} : \{ \text{all possible events / sources for the WINS application} \};$

Confidence:

$$P(S = Xk) = \left\{ \left\{ \sum_{j=1}^M \sum_{l=1}^n P_{jN}(S_{(l)}) * \delta(S_{jN}(l) = Xk) \right\} / n \right\};$$

where,

$$n = \sum_{j=1}^M \sum_{l=1}^n \delta(S_{jN}(l) = Xk); \text{ for } k = 1, 2, \dots, N \text{ where } Xk \in \mathfrak{R};$$

### **3.5.5 Lumped Weighting Method with Combined Class and Segmentation based Decision Making**

Similar to weighting methods used for decision making with the same type or same segmentation of signals, weighting methods are combined to have both types and segments lumped together for generalized weighted decision-making and confidence building. This is the super set of independent segmentation based, and class based weighting. However, this lumped generalized method loses control features present in independent class based and segmentation based weighting. Thus, it cannot discard all departure segments if situational / environmental changes cause signal variations. It may also not be able to discard seismic signatures due to environmental soil condition changes. With less control features, this set of segmented signals acquired by different types / classes of sensors are combined for decision making as well as associating confidence measures.

- **Lumped Segment and Class based decision making:** Principle where decisions are made with different classes and segments of signals combined together with confidence assigned with lumped weighting;

Both criteria for Maximum Polling and Weighted Averaging for lumped segment and class-based decision-making are shown in the following derivations.

- **Maximum Polling:** Signals of different types and different segments are classified / identified in parallel with the SCIA. Results obtained from the SCIA would constitute a single source along with its confidence measure for each segment and class. The method of maximum polling considers only the selected source for each class segment and selects interim final, or final decision with confidence measure calculations following soon. However, confidence for the final decision is calculated independently for maximum polling.

Maximum Polling Source Identification / Classification

- Type 1, Segment 1 –  $S_{11}(11)$  - Selected Source; Confidence –  $P_{11}(S_{(1)})$ ;
- Type 1, Segment 2 –  $S_{12}(12)$  - Selected Source; Confidence –  $P_{12}(S_{(1)})$ ;
- Type 1, Segment N –  $S_{1N}(1N)$  - Selected Source; Confidence –  $P_{1N}(S_{(1)})$ ;
- Type 2, Segment 1 –  $S_{21}(21)$  - Selected Source; Confidence –  $P_{21}(S_{(2)})$ ;
- Type 2, Segment 2 –  $S_{22}(22)$  - Selected Source; Confidence –  $P_{22}(S_{(2)})$ ;
- Type 2, Segment N –  $S_{2N}(2N)$  - Selected Source; Confidence –  $P_{2N}(S_{(2)})$ ;
- Type M, Segment 1 –  $S_{M1}(M1)$  - Selected Source; Confidence –  $P_{M1}(S_{(M)})$ ;
- Type M, Segment 2 –  $S_{M2}(M2)$  - Selected Source; Confidence –  $P_{M2}(S_{(M)})$ ;
- Type M, Segment N –  $S_{MN}(MN)$ -Selected Source; Confidence –  $P_{MN}(S_{(M)})$ ;

Lumped Decision with all Segment and types included:

Decision:

$$\text{Maximum} \left\{ \sum_{i=1}^M \sum_{j=1}^N \delta ( S_{ij}(ij) = Xk ) \right\};$$

*for*  $k = 1, 2, \dots, N$  sources;  $\forall Xk \in \mathfrak{R}$ ;

where  $\mathfrak{R} : \{ \text{all possible events / sources for the WINS application} \}$ ;

Confidence:

$$P(S = Xk) = \left\{ \left\{ \sum_{i=1}^M \sum_{j=1}^N P_{ij}(S_{ij}) * \delta ( S_{ij}(ij) = Xk ) \right\} / n \right\};$$

where,

$$n = \sum_{i=1}^M \sum_{j=1}^N \delta ( S_{ij}(ij) = Xk ); \text{ for } k = 1, 2, \dots, N \text{ sources; where } Xk \in \mathfrak{R};$$

		$S_{(i)}$	$P(S_{(i)})$
acoustic	approach	tv8	0.88
	arrival	tv8	0.55
	departure	tv5	0.81
seismic	approach	tv8	0.71
	arrival	wv8	0.44
	departure	wv1	0.51

Decision: tv8 (3/6)

Confidence: 0.71

**Table 3. 6. Lumped maximum polling.**

Table 3.6 shows an instance when multi-sensing is utilized with results from multiple state spaces. Here maximum polling yields a decision of ‘tv8’ with a confidence of 0.71 obtained using the above methods.

Maximum polling in this case is considered for signals where the sensor types are different along with their segments (i.e. approach, arrival, departure). The confidence measure is now calculated with all segments with all types of signals that were used for decision-making.

- **Weighted Averaging:** Weighted averaging is considered here with all segments and sensor types simultaneously. This method selects from a list of selected sources (results of SCIA) with identification / classification results. Lumped decision-making is obtained with individual results for maximum polling. However, confidence for the final decision is calculated following depending to selected list of results from the SCIA, with association to corresponding results.

#### Weighted Polling Source Identification / Classification

Type 1; Segment 1 –  $S_{11}(1)$  - Selected Source 1; Confidence –  $P_{11}(S_{(1)})$ ;

Type 1; Segment 1 -  $S_{11}(2)$  - Selected Source 2; Confidence –  $P_{11}(S_{(2)})$ ;

Type 1; Segment 1 –  $S_{11}(r)$  - Selected Source n; Confidence –  $P_{11}(S_{(n)})$ ;

Type 1; Segment 2 –  $S_{12}(1)$  - Selected Source 1; Confidence –  $P_{12}(S_{(1)})$ ;

Type 1; Segment 2 -  $S_{12}(2)$  - Selected Source 2; Confidence –  $P_{12}(S_{(2)})$ ;

Type 1; Segment 2 –  $S_{12}(r)$  - Selected Source n; Confidence –  $P_{12}(S_{(n)})$ ;

Type 1; Segment N –  $S_{1N}(1)$  - Selected Source 1; Confidence –  $P_{1N}(S_{(1)})$ ;

Type 1; Segment N -  $S_{1N}(2)$  - Selected Source 2; Confidence –  $P_{1N}(S_{(2)})$ ;

Type 1; Segment N –  $S_{1N}(r)$  - Selected Source n; Confidence –  $P_{1N}(S_{(n)})$ ;

Type 2; Segment 1 –  $S_{21}(1)$  - Selected Source 1; Confidence –  $P_{21}(S_{(1)})$ ;

Type 2; Segment 1 -  $S_{21}(2)$  - Selected Source 2; Confidence –  $P_{21}(S_{(2)})$ ;

Type 2; Segment 1 –  $S_{21}(r)$  - Selected Source n; Confidence –  $P_{21}(S_{(n)})$ ;

Type 2; Segment 2 –  $S_{22}(1)$  - Selected Source 1; Confidence –  $P_{22}(S_{(1)})$ ;

Type 2; Segment 2 -  $S_{22}(2)$  - Selected Source 2; Confidence –  $P_{22}(S_{(2)})$ ;

Type 2; Segment 2 –  $S_{22}(r)$  - Selected Source n; Confidence –  $P_{22}(S_{(n)})$ ;

Type 2; Segment N –  $S_{2N}(1)$  - Selected Source 1; Confidence –  $P_{2N}(S_{(1)})$ ;

Type 2; Segment N -  $S_{2N}(2)$  - Selected Source 2; Confidence –  $P_{2N}(S_{(2)})$ ;

Type 2; Segment N –  $S_{2N}(r)$  - Selected Source n; Confidence –  $P_{2N}(S_{(n)})$ ;

Type M; Segment 1 –  $S_{M1}(1)$  - Selected Source 1; Confidence –  $P_{M1}(S_{(1)})$ ;

Type M; Segment 1 –  $S_{M1}(2)$  - Selected Source 2; Confidence –  $P_{M1}(S_{(2)})$ ;

Type M; Segment 1 –  $S_{M1}(r)$  - Selected Source n; Confidence –  $P_{M1}(S_{(n)})$ ;

Type M; Segment 2 –  $S_{M2}(1)$  - Selected Source 1; Confidence –  $P_{M2}(S_{(1)})$ ;

Type M; Segment 2 –  $S_{M2}(2)$  - Selected Source 2; Confidence –  $P_{M2}(S_{(2)})$ ;

Type M; Segment 2 –  $S_{M2}(r)$  - Selected Source n; Confidence –  $P_{M2}(S_{(n)})$ ;

Type M; Segment N –  $S_{MN}(1)$  - Selected Source 1; Confidence –  $P_{MN}(S_{(1)})$ ;

Type M; Segment N –  $S_{MN}(2)$  - Selected Source 2; Confidence –  $P_{MN}(S_{(2)})$ ;

Type M; Segment N –  $S_{MN}(r)$  - Selected Source n; Confidence –  $P_{MN}(S_{(n)})$ ;

Decision:

$$\text{Maximum} \left\{ \sum_{i=1}^M \sum_{j=1}^N \sum_{p=1}^r \delta ( S_{ij}(p) = Xk ) * P_{ij}(S_{(i)}) \right\};$$

*for k = 1,2,...,N sources;  $\forall Xk \in \mathfrak{R}$ ;*

Confidence:

$$P(S = Xk) = \left\{ \left\{ \sum_{i=1}^M \sum_{j=1}^N \sum_{p=1}^r P_{ij}(S_{(p)}) * \delta ( S_{ij}(p) = Xk ) \right\} \right\} / n$$

where,

$$n = \sum_{i=1}^M \sum_{j=1}^N \sum_{p=1}^r \delta ( S_{ij}(p) = Xk );$$

*for k = 1,2,...,N sources; where  $Xk \in \mathfrak{R}$ ;*

		$S_{(i)}$	$P(S_{(i)})$			$S_{(i)}$	$P(S_{(i)})$
<b>acoustic</b>	approach	tv8	0.88	<b>seismic</b>	approach	tv8	0.71
		wv4	0.91			tv5	0.55
		tv3	0.11			wv3	0.22
	arrival	tv6	0.55		arrival	wv3	0.44
		tv8	0.53			tv8	0.41
		wv4	0.21			tv3	0.51
	departure	tv5	0.51		departure	wv1	0.31
		tv8	0.49			wv3	0.30
		wv1	0.11			tv8	0.29

Decision: tv8

Confidence: 0.55

**Table 3. 7. Lumped weighted averaging.**

Lumped weighted averaging is taken up above with a list of results with associated confidence for the case of multi-sensing sensors with multiple state spaces. From the above method, a decision of ‘tv8’ is obtained with a confidence of 0.55. With this form of lumped decision-making modularity is lost, but it has a simple hardware implementation of multipliers and adders.

These weighting methods are used with SSE / M-SSE following the SCIA to obtain statistical analysis and calculate confidence intervals. These weighting schemes are studied for localized, cluster head based, or centralized weighting schemes depending on SSE / M-SSE architecture. Decisions derived from sub-systems or cluster heads are assigned one of the above weighting schemes on a system-to-system, or sub-system to sub-system basis. The number of nodes taking part in the decision making process needs to be included with this method of decision-making.

System 1 => S(1)=Xi    Confidence => P(S(1)=Xi)    Number of nodes => n1  
 System 2 => S(2)=Xj    Confidence => P(S(2)=Xj)    Number of nodes => n2  
 System N =>S(N)=Xk    Confidence => P(S(N)=Xk)    Number of nodes => nN

Decision:

$$\text{Maximum} \left\{ \sum_{t=1}^N \sum_{u=i,j,k} \{ \delta(S(t)=X(u)) * P(S(t)=X(u)) * nt \} \right\};$$

*for u = i,j,k.... sources;  $\forall Xu \in \mathfrak{R}$ ;*

Confidence:

$$P(S = Xk) = \left\{ \left\{ \sum_{i=1}^M \sum_{j=1}^N P(S(t) = Xk) * nt * \delta (S(t) = Xk) \right\} / z \right\};$$

where,

$$z = \sum_{t=1}^N nt * \delta (S(t) = Xk);$$

where  $n$  = number of occurrences of  $S = Xk$ ;

for  $k = 1, 2, \dots, N$  systems ; where  $Xk \in \mathfrak{R}$ ;

An example of decision making with results from sensor sub-systems follows:

	System Name	$S_{(i)}$	$n_i$	$P(S_{(i)})$
Sensor sub-systems	system 1	tv8	3	0.66
	system 2	wv5	2	0.55
	system 3	tv8	3	0.78

Decision:	tv8
Confidence:	0.72

**Table 3. 8. Sub-system level decision making.**

The above method is utilized at the system and sub-system levels depending on WINS applications. These methods offer a novel way of combining statistical weighting schemes towards the final decision making with the SSE / M-SSE architecture. Table 3.8 shows results from three sensor sub-systems, and is shown here as an example when cluster head relay decisions to a central processing node.

### **3.6 Sensor SCIA Control Criteria**

Robustness of SCIA is achieved by having different modules activated and de-activated depending on the nature and quality of sensed signals. Signals are often corrupted, or distorted due to environmental as well as circuit effects (as in the case of random glitches and saturation at CPA). These discrepancies cause low probability of detection, while resulting in low confidence on classification / identification results. Therefore, a pre-emptive low power optimization is used to discard these undesirable signal segments. This is achieved by detecting saturation points and glitches with the help of a threshold. Deactivating seismic sensing during rainy days, or to deactivate acoustic sensors during windy conditions or in mountainous terrain could mitigate these discrepancies. Decision-making process, and confidence calculations follow the previously mentioned methods. SCIA modules are utilized efficiently by deactivating modules assigned to identify and detect signal types or segments, or a combination of both. The module based concept of SSE / M-SSE architecture facilitates the process whereby energy constrained nodes will prolong life with selective SCIA processing.

### **3.7 Summary**

This chapter investigated architectures for distributed vs. centralized signal processing. Signal pre-processing, and signal processing (SCIA algorithm), along with decision-making methods were described. SSE / M-SSE architectures were described along with requirements and schemes for decision-making. Various collaborative decision making schemes were discussed along with obtaining associated confidence measures with each decision. Decision-making architectures and methodology are determined during training of application specific signals. It is decided during training as to which decision fusion method to implement for a particular application. Additional new signal variables (either introduced after deployment, or not present in signals in the training database) will introduce low confidence measures and errors that call for updates on the template tree along with associated confidence measures. In this case re-training is enforced either off-line with the new training database or process on-board.

Energy constraints were calculated and compared with communication and computation costs to analyze wireless data communication cost. The need to perform on-board training and decision-making is emphasized by the large energy requirements in transferring data between nodes.

## **CHAPTER 4**

### **Investigation of SSE to Signal Variables**

The Signal Search Engine (SSE) algorithm was investigated extensively for many theoretical and real-life signal sets. A test signal set was supplemented with additional wideband signals added from real life situations to obtain full coverage. This chapter looks at SSE performance based on this full coverage signal set.

These test signals were self-contained in that they were initially created with varying amplitude, frequency, phase, and noise content. The signals were divided into the following classes: Single narrow-band (discrete frequency), multiple narrow-band (multiple discrete frequencies), and wideband signals for extensive testing on the SSE. These signals formed the fundamentals for SSE event classification and identification.

Created signals were input to the SSE algorithm that was coded with MATLAB modules, for template selection, and VBasic, and Visual C++ modules for identification and classification. An interface with MS 'Excel' application software was used to obtain

plots of AARMS / MARMS scores for a graphical representation. Frequency features from created test signals were used as the classification / identification criteria that closely associated with that of real world signal detection.

## **4.1 Test Signal Set**

Narrow band signals correspond to mutually exclusive events, which are discrete in frequency and often periodic. Wideband frequencies on the other hand are discrete in time and continuous or semi-continuous (wideband with discontinuities) in frequency. The following are examples of the cases of wideband signals: clock-ticks with sensors in the vicinity of the source; non-periodic events such as gun shots from repeater rifles (signals modified to obtain non-periodic or random occurrence by introduction of dirac delta / sinc function in time); semi-periodic or random events such as sounds from robots (used in mine clearing operations), and sudden explosions, and automatic fire. Such signals are created to investigate the wideband / discontinuous wideband characteristics of the SSE. Semi-periodic or random events were created by either manually adjusting the periodic signal without loss of phase or by modifying the amplitude content.

Additionally, narrow band signals were created and tested for SSE behavior. Wideband frequencies or discontinuous wideband frequencies can be modeled with continuous narrowband frequencies in the discrete frequency domain. Therefore, narrowband behavior of the SSE is critical for accurate detection of wideband signals / discontinuous

wideband signals in the discrete frequency domain. The following gives information on the created test signal set.

**4.1.1. Single Narrow Band:** Narrow band sinusoidal signals with single discrete frequencies corresponding to low frequency acoustic and seismic signals were created. High frequency acoustic, and seismic signals were created for a sanity check since the SSE is robust to narrow band signals in the acoustic and seismic frequency spectrum. Zero mean additive white Gaussian noise (AWGN) with variance  $\sigma^2$  was added to have signal SNRs of 0 dB to 20 dB. These signals were used to investigate and test the robustness of the SSE and its response to noise. Amplitudes were varied discretely for different signals throughout the signal set to test amplitude robustness of the algorithm.

The following represents a single frequency narrow band signal

$$S_s(n) = A_1 * e^{j(2\pi n \frac{f_l}{f_s})} + \psi_s \quad (4.1)$$

where  $A_1$  is the signal amplitude and  $f_l$ , the narrow band frequency. Here we sample the signals at  $f_s=1$  KHz for both acoustic and seismic sensor nodes. Frequencies are selected within close proximity with separation increasing many fold to see the robustness of the SSE in classifying / identifying close frequency narrowband signals to those with frequencies separated wide apart. [10,29]  $\psi_s$  is zero mean AWGN with a variance of  $\sigma^2$  introducing signal SNRs of 0 to 20dB. The effect of Doppler is modeled with the following equation and is used in all preceding equation with step functions triggering to model closest point of approach.

$$f_{doppler} = \alpha^2 f_l \quad (4.2)$$

A phase shift component is added to the above narrow band signals to investigate the tolerance to phase offset and performance of SCIA with a generic template tree. With the added phase offset, the signal is represented by,

$$S_s(n) = A_i * e^{j(2\pi n \frac{f_1}{f_s} + \phi_s)} + \psi_s \quad (4.3)$$

where  $\phi_s$  is the phase component associated with signal  $S_s$ . Phase is varied from  $0 - \pi$  radians, covering source presence from aligned far-field to closest point of approach (CPA) and encompassing non-aligned far-field ( $\pi$  radians off sensor beam direction).

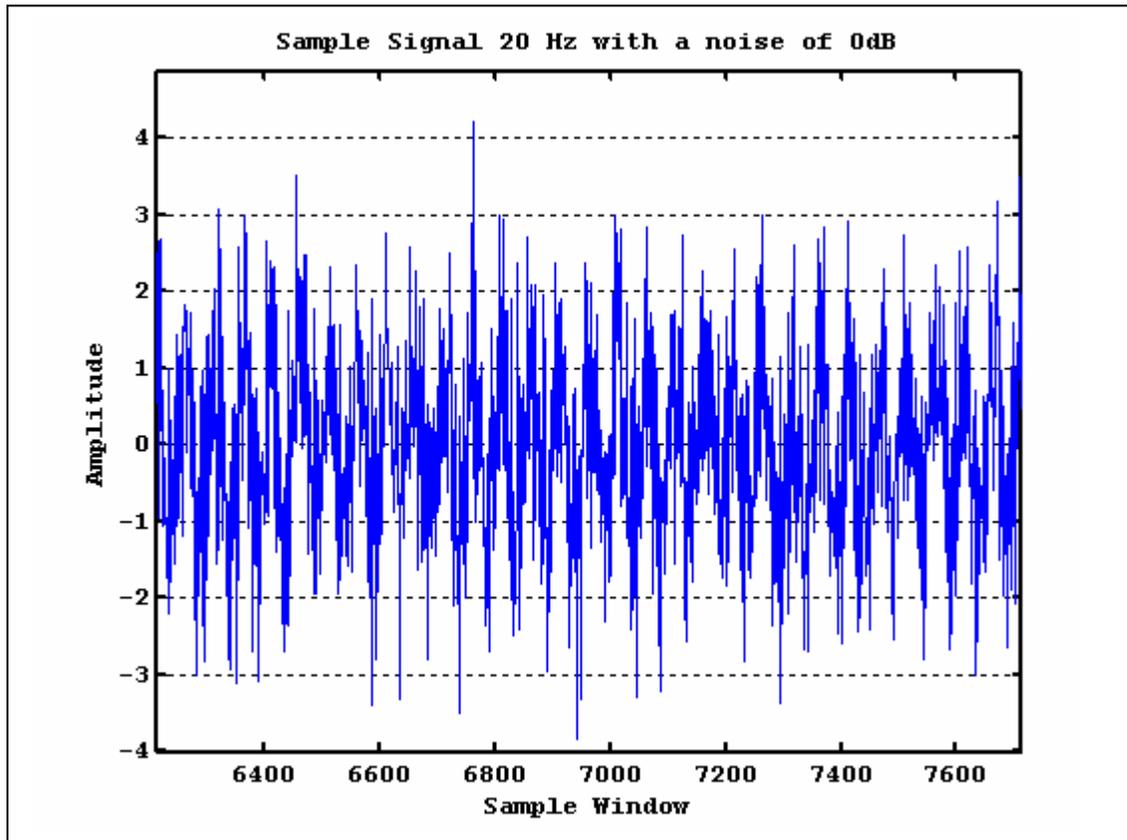


Figure 4. 1 A sample low frequency signal window, with an SNR of 0dB.

This is the case when stringent power constraints make computation intensive beamforming algorithms unusable, whereby each sensor (with a generic template tree embedded on the SCIA) in the sensor array would receive the same copy of the signal with phase-shifts. This investigation gives insight, as to how an array of sensors would receive the same signal with different phase offsets. Successful results would yield validity of a generic SSE template tree for distributed SCIA reducing the need in customized template tree assignments for each sensor node in an array. Its robustness to different phase and accuracy levels would validate generic SSE for distributed signal processing. This will provide validity to a generic template tree for individual clusters, and will prove independence of sensor location for the SSE's SCIA template tree.

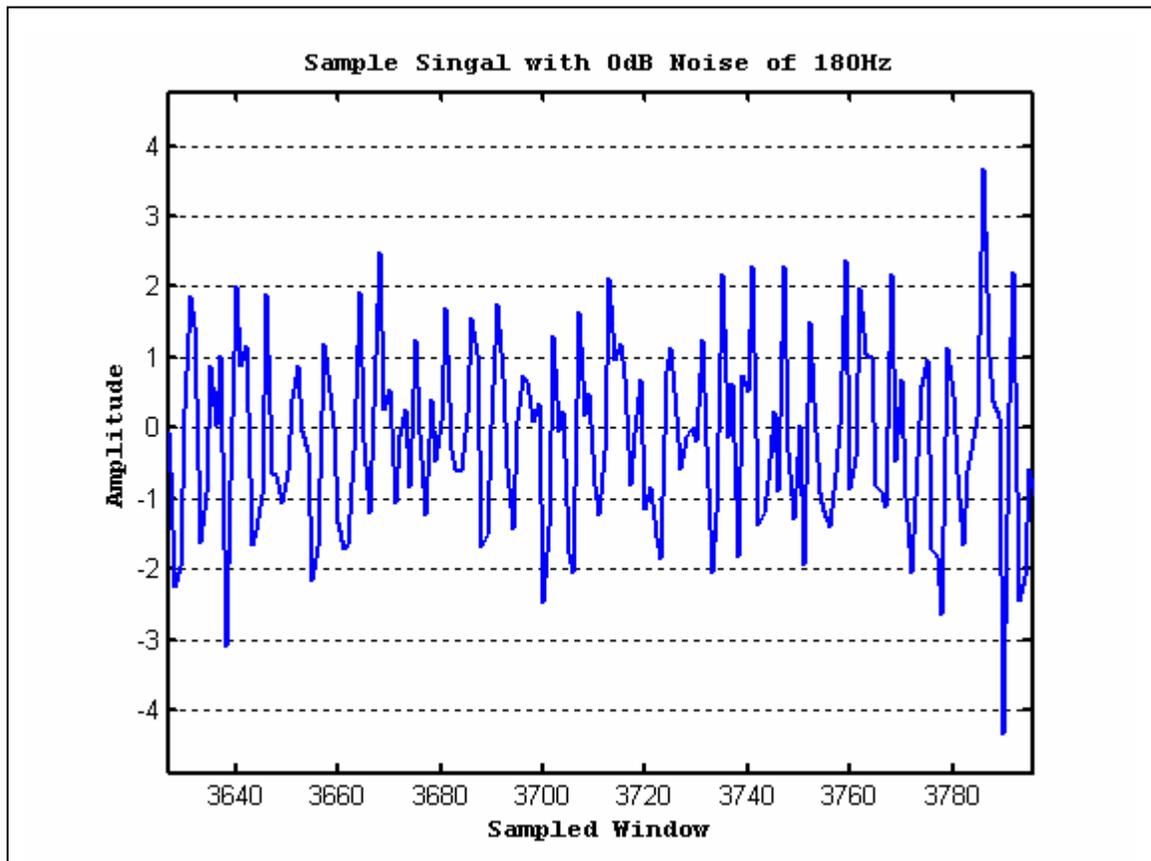
**4.1.2. Multiple Discrete Frequency Narrowband:** Multiple narrow band signals were created using a similar procedure for multiple discrete frequencies. The multiple discrete frequencies were selected with random spacing and had at least two discrete frequencies per signal. A set of discrete frequencies in a continuous block was created to see the effect of a comb train (when wideband signals are digitized into segments of continuous discrete signals in the time domain) and its response to SSE classification / identification. Varying amplitudes and phases were introduced in each of these multiple-discrete frequency signals corresponding to simultaneous periodic and semi-periodic, mutually exclusive events. Amplitudes were varied for each segment, along with their frequencies for a low frequency band of signals. Higher frequency signals were created on high frequency acoustic / seismic bands for a sanity check of

their effect on the SSE. Investigation of low and high frequency signals were checked similarly with input to the SCIA modules. Phase offsets on these events were introduced to include readings from different sensors in a sensor array, or to include multi-path components (delayed versions of the same signal arriving at individual sensors due to reflection). Relative phase offsets ranged between  $\phi_1 = \pi/180$  and  $\phi_2 = \pi$  radians to cover directional antennas with a coverage of a  $180^\circ$  plane. Zero mean AWGN with variance  $\sigma^2$  was introduced to the signals at a level ranging from 0 dB to 20 dB. The generic formula below shows a multi frequency narrow band signal:

$$S_m(n) = A_1 * e^{j(2\pi n \frac{f_1}{f_s})} + A_2 * e^{j(2\pi n \frac{f_2}{f_s})} + \dots + A_m * e^{j(2\pi n \frac{f_m}{f_s})} + \psi_s \quad (4.4)$$

where  $A_1, A_2 \dots A_m$  are the respective amplitudes, and  $f_1, f_2 \dots f_m$  the irrespective narrow band frequencies.  $\psi_s$  is the zero mean AWGN noise component with variance  $\sigma^2$  resulting in signal SNRs of 0dB to 20dB noise levels.

The above signals present situations where multiple mutually exclusive events occur simultaneously. Real life situations consist of periodic, semi-periodic, random, and / or chaotic events occurring during a given time interval. These events would happen at the same time at different locations. An acoustic sensor in a closed environment may receive periodic clock ticks from multiple clocks located at different distances from the sensor, each with different signal strengths and phase offset from each other. A battleground



**Figure 4.2 A high frequency signal with a SNR of 0 dB.**

may consist of multiple automatic / semi-automatic gunfire events often coming from different locations. Gunfire is bursty in nature, and may co-exist with artillery, or intermittent rocket fire corresponding to different signal strengths modeled by Dirac delta functions. Such cases are modeled by multiple discrete narrowband frequencies, with different amplitudes, and phases. Frequencies for the various cases are modeled by periodic, semi-periodic, or random frequencies along with their phase offsets. The following formula shows a multi-narrow band signal with phase offsets related to each other.

$$S_m(n) = A_1 * e^{j(2\pi n \frac{f_1}{f_s} + \phi_1)} + A_2 * e^{j(2\pi n \frac{f_2}{f_s} + \phi_2)} + \dots + A_m * e^{j(2\pi n \frac{f_m}{f_s} + \phi_m)} + \Psi_s \quad (4.5)$$

where  $A_1, A_2, \dots, A_m$  are the respective amplitudes, and  $f_1, f_2, \dots, f_m$  the respective narrow band frequencies. The phase offset on each frequency is given by  $\phi_1, \phi_2, \dots, \phi_m$  respectively.

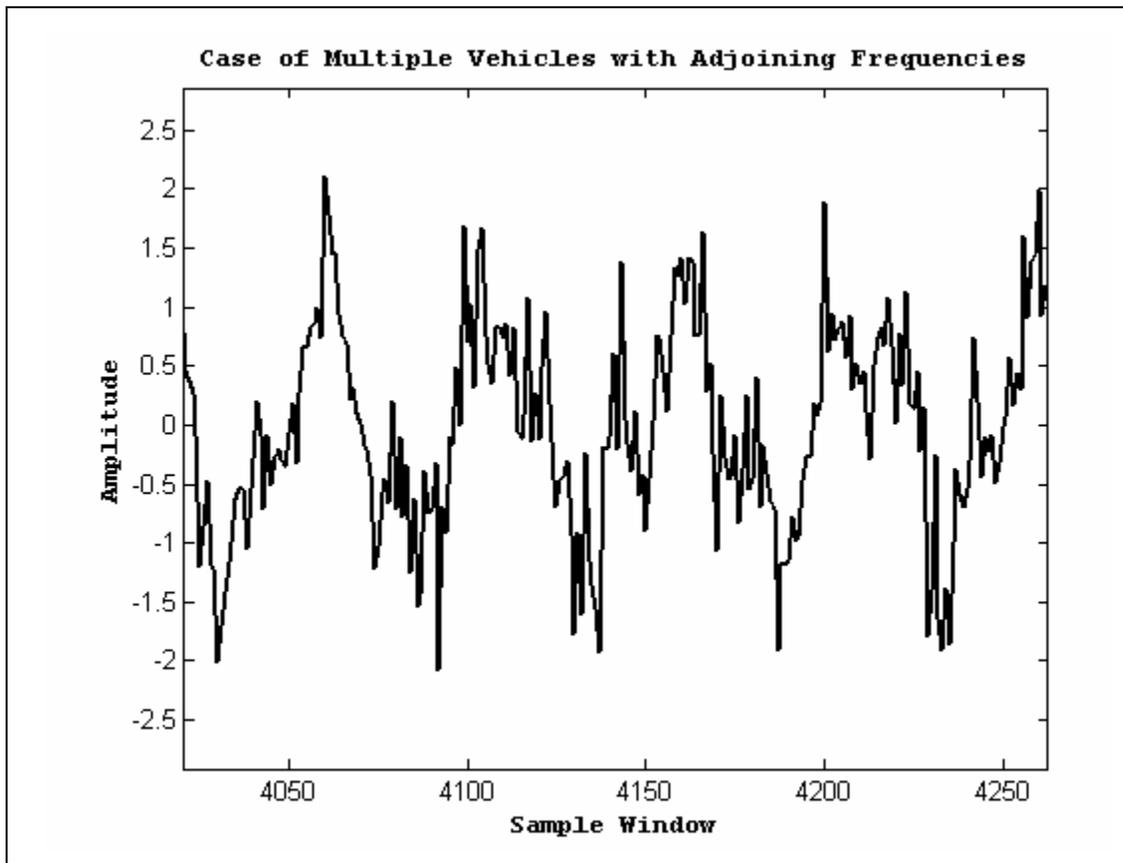


Figure 4.3 Signal with contents from multiple adjoining frequency bands.

These signal features from Figure 4.3 are exploited to obtain templates during feature extraction (signal pre-processing) to build a classification / identification tree structure. Distinct features are embedded on the template tree structure along with associated confidence measures obtained during training.

### 4.1.3. Wideband Signal Set:

Application specific SSEs should cater to wideband signals, since many present-day signals detected by acoustic, and seismic sensors are wideband in nature. It should be noted that unlike the radar communication problem where the signal may be a random process with known statistics or drawn from a known finite alphabet, the acoustic/seismic wideband source signals are most likely deterministic but unknown (e.g. waveforms generated by a passing vehicle). [6] Therefore, real world signals are used to test wideband source signals on the SSE and results presented in Chapter 6. These wideband signals are given by the following formula, [23] and take into account multiple source presence:

$$S_w(n) = \sum_{m=1}^M A_w^{(m)} S_w^{(m)}(n - t_w^{(m)}) + \psi_w(n) \quad (4.6)$$

where  $n = 0 \dots L-1$  is the number of received samples,  $A_w^{(m)}$  is the amplitude of the signal of the  $m^{\text{th}}$  source, and  $m = 1, \dots, M$  the number of sources.  $S_w^{(m)}$  is the wideband source signal from the  $M^{\text{th}}$  source, with  $t_w^{(m)}$  which is the fractional time delay in samples, along

with  $\psi_w$  which is the zero mean AWGN with variance  $\sigma^2$ . These sensor readings could further be modeled to include multiple sensor readings with associated spacing between them, and to characterize sensor readings to model signals for sensor arrays. [91] However, this phenomenon is not looked at in the SSE testing.

Test signals were introduced with zero mean AWGN with variance  $\sigma^2$  to include environmental and circuit noise as sensed by the receiver to achieve various levels of

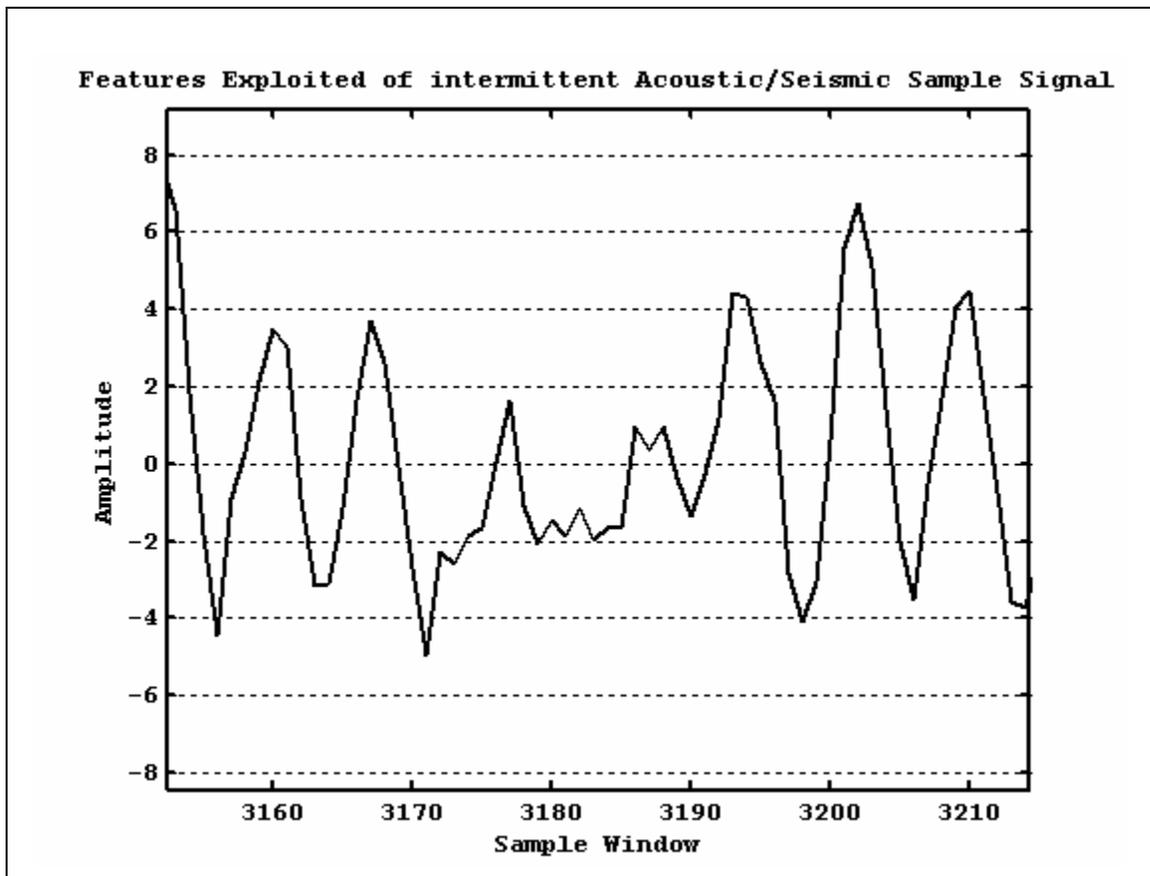


Figure 4. 4 Distinct signal features selected as a Correlator template.

SNRs. It was observed that the SSE was robust to noise effects up to 0 dB SNR, with AARMS / MARMS variations of up to 65% at 0 dB, falling within the boundaries of correct identification / classification. Having obtained excellent identification / classification results, a look at noise and distortion in acoustic, seismic, and infrared signals is given below.

## **4.2 Channel and Environmental Effects on Signals**

Seismic, acoustic, and infrared signals traveling through a channel suffer distortion and channel specific degradation. [74] The channel is modeled by convolving the channel transfer function with that of the wideband signal. However, the SSE algorithm was robust to minimum channel nulling effects of the time varying acoustic and seismic channels. This is achieved by statistical averaging that is used over the segments under consideration, in addition to averaging RMS values obtained by window stepping. A more accurate method would be to identify the channel and deconvolve the channel to obtain an approximation of the source signal without channel-induced distortion. [51,60,61,97] However channel estimation and deconvolution of the signal is computationally intensive and therefore avoided for purposes of the SSE.

### 4.2.1 Acoustic Signals

#### **Noise due to Environmental Conditions:**

**Wind (Flat Open Terrain):** Noise due to wind or turbulence distorts acoustic signals. Examples of such distortion include dilation effects. Dilation is a form of noise that introduces frequency, and amplitude offsets that may give erroneous classification / identification. It is therefore best to turn acoustic sensing off in windy situations since confidence measures would be low.

**Echo (Mountainous Terrain):** Acoustic signals sensed in mountainous terrain have replicas of reflected signals that arrive at sensor nodes with a time delay. These multi-path components are aligned and combined by array-processing or beamforming algorithms. However, when these computationally intensive methods are not present, a combination of time-frequency information along with the time domain SSE could be utilized to increase confidence in results.

### 4.2.2 Seismic Signals

**Rain / Snow (Atmospheric Effects):** Rain causes ground conditions to change drastically especially on seismic sensors located in clay ground, or situated in dense foliage that would keep water resident much longer than desert sand. Channel characteristics also change in snowy conditions that would distort and have nulling

effects on sensors buried in the ground. [50] In these situations, sensors should be looked at closely with time-frequency effects along with accuracy readings obtained during classification / identification. Therefore to have a generic tree structure, it is essential to train the algorithm, and create the template tree with signals obtained from different environmental conditions as performed in Chapter 6 of this dissertation.

### **4.2.3 Infrared Signals**

**Thermal gradients (Atmospheric Effects):** Thermal gradients existent in the channel medium introduce noise in low frequency infrared signals causing feature set loss. When contrasting thermal gradients are present in LOS (line of sight) signals, the obtained IR signals are fuzzy or distorted. Presently infrared signals are used as a triggering mechanism to awake the SSE from sleep mode. When thermal gradients cause LOS IR signals to be distorted, other modalities such as seismic signals should be used to awake the SSE by a threshold trigger mechanism.

**Foliage ( Line Of Sight ) :** Low frequency infrared signals form the core for triggering the LOS (line of sight) signals and helps in the detection of thermal gradients originating from the source. However, in dense foliage, signals are distorted considerably due to environment sensitive vegetation conditions that make the signals hard to detect or to be used as a trigger mechanism. The above characteristics would make IR sensors hard to

use and therefore other means such as seismic signals should be used to trigger sensing from sleep state.

#### 4.2.4 All Signals

**Lightning:** Lightning and other circuit effects introduce glitches on the signals that need to be eliminated by the SSE algorithm. Glitches are either detected ahead of time and eliminated during calculations of average RMS or excluded during pre-processing of the signal. This could be obtained by introducing a threshold discard during average AARMS / MARMS calculation.

**Glitches:** Many sensor readings exhibit random glitches that need to be acted upon during signal pre-processing. These random occurrences caused RMS variations, and special handling of the glitches were derived and implemented in the SSE algorithm. Circuit effects, or EMI and environmental effects such as lightning cause glitch. Test signals with intentional glitches were used to understand SSE's SCIA algorithm's behavior. However, classification / identification results remained accurate even in the presence of glitches. This is due to the following feature in the SSE:

- threshold elimination of windowed signal segments during RMS calculation.
- Maximum Absolute RMS (MARMS) algorithm instead of AARMS. (definitions of MARMS and AARMS are given in Chapter 5, Section 5.6.3)
- threshold elimination of low RMS values during MARMS.

These fixes were performed in the signal-processing module. However, threshold elimination of adjacent low amplitude continuous samples gave even more robustness and fewer computations in the signal-processing module.

#### **Effect of Circuit Noise:**

Low frequency signals are dominated by flicker noise that is inversely proportional to the frequency of the sensed signal. At the higher frequencies, thermal noise begins to limit performance, since it dominates over flicker noise and introduces a relatively flat and linear frequency noise effect. Circuit system analysis already accounts for noise effects in designs; therefore, we use a macro system model and include these noise effects with a zero mean AWGN with variance  $\sigma^2$ . Noise is introduced depending on the frequency of the signal.

### **4.3 Classification / Identification of Test Signal Set**

The test signal set that we created was used to test for classification / identification using SSE's SCIA algorithm. Occasionally, we used acquired real world signals to obtain full-coverage of the test signal set. The test signal set was initially clustered into individual sets as described in 4.1.1 – 4.1.3, with results obtained by initially picking correlator templates. Once correlator templates were selected, testing was done by separating the

signals into different classes and thereby obtaining classification and identification on the signal sets.

#### **4.3.1 Template Selection for Classification / Identification**

Signals used for selection of correlator templates were from a training set that mutually excluded test signals. Correlator segments were initially chosen for classification, and later for identification. Template selection algorithms, described in Chapter 5, were used in selecting correlator templates.

Template selection for classification of the test signal set was done separately for single frequency, multiple frequency, and wideband signals. Following the selection of templates, test signals were performed on the selected correlator templates with mutually exclusive signal sets from a combination of single, multiple, and wideband frequency signals. Correlator templates were picked from a subset of created signals comprising a training set. The signal used to pick the template was excluded from the test signal set. Investigated signal sets comprised of phase and SNR variations of

$$0 \leq \Delta\phi \leq \pi \quad (\text{phase});$$

$$0 \text{ dB} \leq \Delta\psi \leq 20 \text{ dB} \quad (\text{zero mean AWGN noise with variance } \sigma^2);$$

which were used to measure the robustness of SSE to frequency, phase, and noise variations.

Confidence on the classification is derived from previous runs on initial training sets, and is updated based on performance of known test signal sets, following repeated runs using those signals. The selection of correlator templates, along with identification / classification boundaries are obtained with exhaustive tests on created signals (now included in the training set).

Selected correlator templates were investigated with histogram plots output from the SSE in order to determine the optimal choice of correlator window for different signal parameters and signal variables. In the initial implementation, a graphical view of the effects of the algorithm is provided, without any signal pre/post processing. This is essential to minimize signal parameter dependence on correlator templates. A complete set of MATLAB scripts were developed for comprehensive full coverage during exhaustive correlator picks and SSE test runs.

#### **4.3.2 Classification / Identification on SSE**

A multi-tier correlator pick algorithm is followed using a multi-step (module) procedure, where the first tier is for classification and the leaf templates for class subset based classification or identification purposes. Template picks followed algorithms mentioned in Chapter 5. Identification could further be extended with a broader signal category such as real-life signals as demonstrated in Chapter 5.

## **4.4 Results on Test Runs**

Created signals were used to test the SSE extensively. Different cases were tested methodically, while SSE modules were optimized to improve performance. A summary of the key results of the SSE is shown below along with the methodology.

### **4.4.1 Identification**

Identification of signals was performed with one, and two hierarchies, and could be extended to multiple hierarchies in real word identification SSEs. Initial one tier correlator templates were created by clustering similar signals together (i.e. single frequency, multiple frequency, wideband) taking into consideration similarities in frequency range / separation, harmonics, bandwidth, and event repetition.

First Tier selection: First tier selection of correlator templates is for generalized classification (i.e., single frequency narrow band, multiple frequency narrow band, wideband in the particular test signal set.) Correlator template picks were made for each training signal with auto or self correlation over the signal. The top three-correlator templates were chosen and used for classification algorithm runs on known (created) test signals. Each test signal from the training set had its top three template picks, for the particular class of signals.

Second Tier selection: Second tier correlator template picks are for the first step in identification. First tier templates could be contained as a subset of the second tier

correlator template superset. A tree can be built from the 1<sup>st</sup> tier, and 2<sup>nd</sup> tier correlator signal sets. In real world SSE implementation, this is the first step in building a tree as discussed in Chapter 6.

#### **4.4.2 SSE Runs on Test Signals**

Test signal runs were performed on signal sets systematically, to investigate behavior of the following on the classification / identification results:

- i. Amplitude variation of signals.
- ii. Phase offset / noise effects on signals.
- iii. Template window size variation.
- iv. Channel medium noise effects on signals.
- v. Stepping size of template window.

A summary of results of the investigation is given below. Customized algorithm modules for classification / identification of moving objects is explained in Chapter 5, while run results are investigated in Chapter 6 on real world moving object waveforms from acquired acoustic and seismic signals.

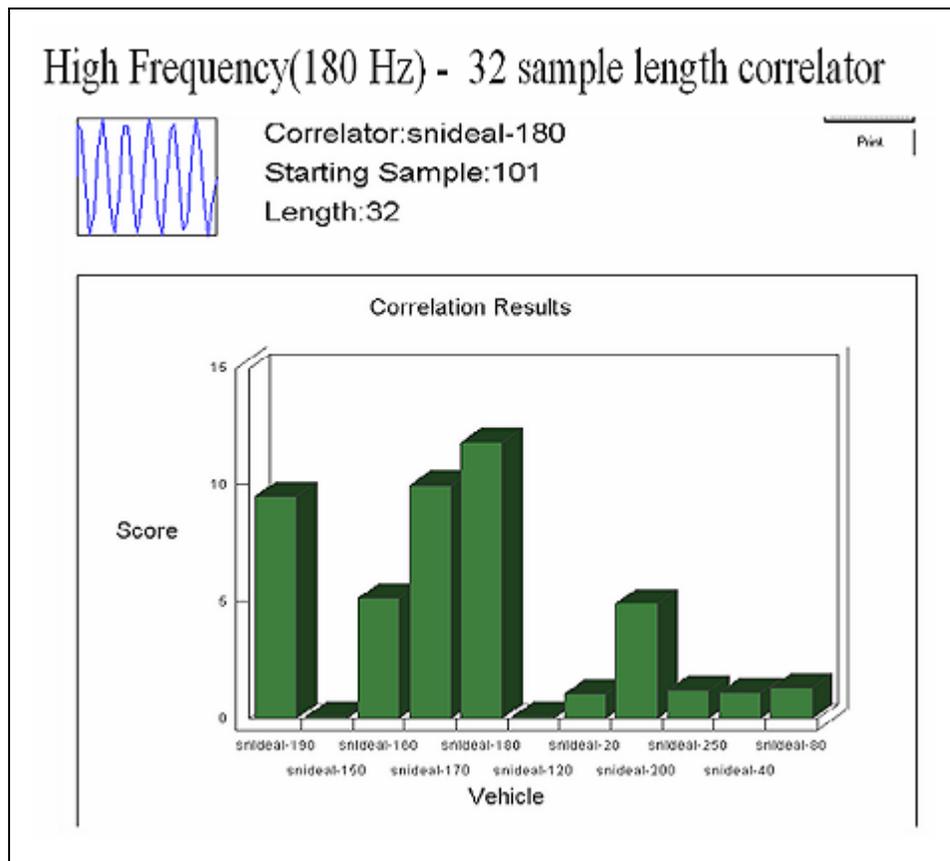
- i. Amplitude Variation.

Signals were modeled to include amplitude variations as observed by the sensors with amplitude dependence corresponding to signal strength. This is the case when sources are present at different distances from sensor nodes. These variations occur during signal

propagation through the channel, or due to multiple time delayed replicas arriving at the sensor node due to multi-path where time delayed version of the signals contribute to amplitude. This effect also occurs when signal sources are at different distances from sensor nodes. SSE robustness was investigated to see whether the signal amplitude variation had any effects on the classification or identification results (CIR). It was observed that the windowed signal normalization module in the SSE mitigated the amplitude variation of the signal. This normalization step provided the same values for averaged RMS values that resulted within the CIR boundaries. The SSE results were consequently robust to amplitude variations.

ii. Phase offset / Noise Effects on Signals.

Many classification / identification methods are sensitive to phase offset / noise effects on signals. This is an important phenomenon and therefore needs to be investigated thoroughly. Test signals with phase offset included in the test signal set gave no difference in the RMS values of the SSE. This is due to the overlap of window stepping during RMS calculations and eventual averaging during AARMS / MARMS calculations. It was concluded that averaging and stepping the signal mitigates dependence in phase. This result indicates that the SSE is robust to phase offset, and therefore any phase noise occurring in sensor arrays has no effect regarding phase in a generic template. Thus the SCIA utilized in distributed classification / identification could have the same templates in adjoining nodes.



**Figure 4.5 High frequency (180 Hz) phase offset behavior of SCIA.**

Results shown on the histogram plot of Figure 4.5 have a 32-sample length correlator template. Phase offset is introduced on signals to observe phase mismatch behavior of the SCIA. Results show correct identification of signals with frequencies within a 10 Hz range to that of the correlator template, depicted by AARMS scores that are close to the correct frequency content of the signal.

The histogram of Figure 4.6 shows results of SSE identification for a low frequency signal. It is inferred here that correlator templates could have a window of less than one wavelength. This is a crucial finding since it eliminates long correlator templates for low frequency signals minimizing computational costs. Phase offset is investigated here with a 32-sample template that identifies 20Hz signals correctly.

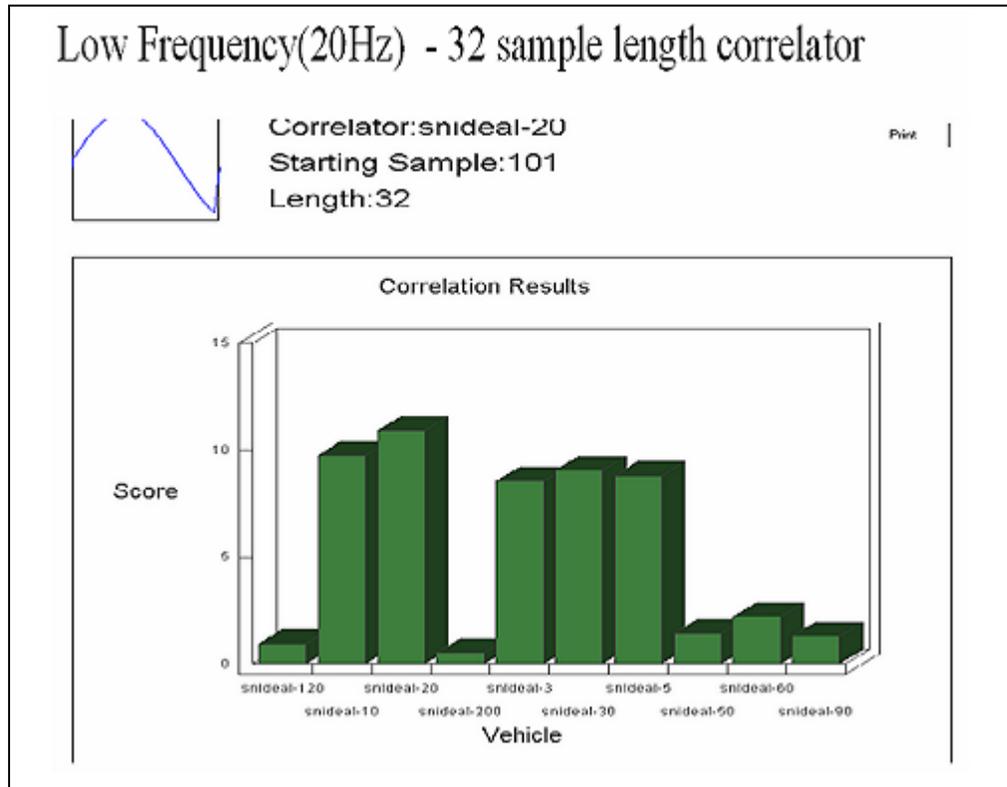


Figure 4. 6. Low frequency (20 Hz) phase offset behavior of SCIA.

iii. Template Window Size Variation

SSE window sizes need to contain as few samples as possible to be low power. The SSE algorithm was investigated to find the minimum number of samples on templates with the mentioned sample parameter variables. Tests were carried out methodically to see what

the effects of reducing the window length would be to increasing noise on signals, along with phase offset and wavelength. The following observations were found:

- RMS scores are scaled according to the size of templates.
- Number of template samples required for correct identification / classification increased with increasing noise.
- It was found that a template could have less than one wavelength; however, results were quite distinguishable as more wavelengths were added to the template, and fell well within the range of decision boundaries.

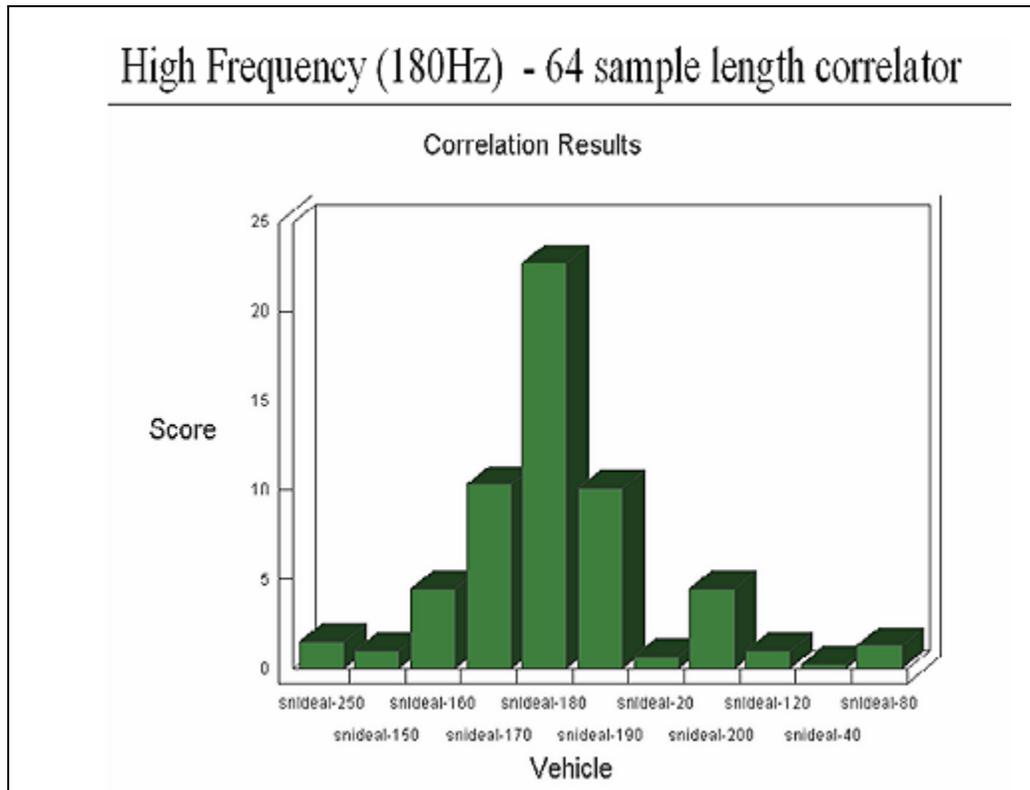
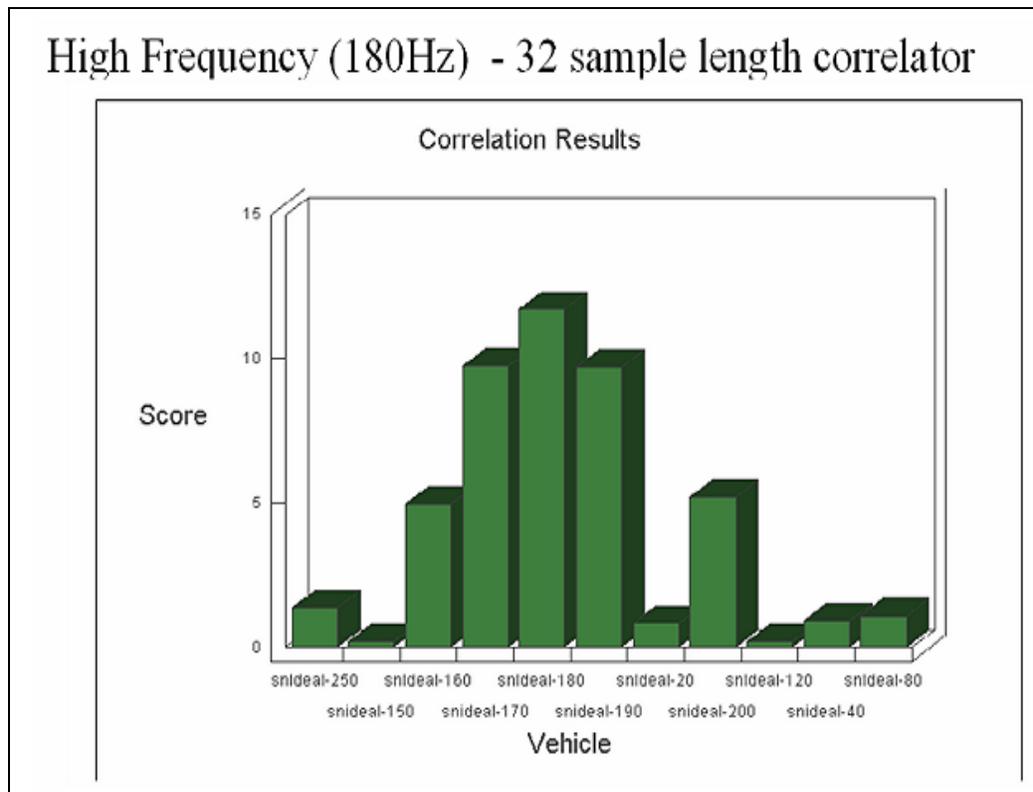


Figure 4. 7 SSE identification with a 64 sample length correlator.

Window size variation effects were investigated on test signals to study the identification performance while minimizing computational costs of the SCIA. Figure 4.7 results show SSE identification with a 64-sample correlator template. RMS scores clearly distinguish this signal from training signals in the same frequency neighborhood. This particular correlator template is good for identification rather than classification since it clearly distinguishes the correct signal from that of adjoining frequency signals, clearly identifying the target. An RMS score of 23 is obtained for this identification.



**Figure 4. 8 SSE identification with a 32 sample length correlator.**

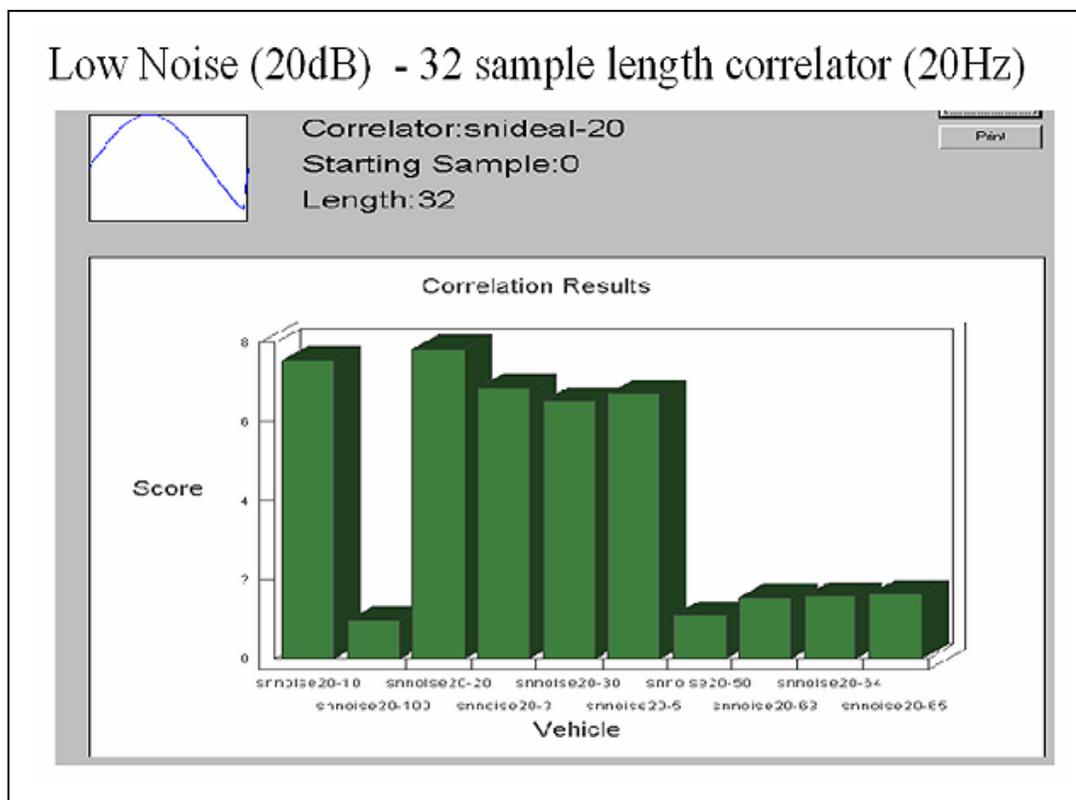
Figure 4.8 shows a 32-sample correlator template with a RMS score of 12.5. It could be seen that when the window size is made smaller, the RMS score is scaled accordingly. Decision boundaries are scaled similarly for different correlator

template sizes. It is inferred from the above study that signal window length is signal type specific and therefore should be optimized for each signal type (i.e. acoustic, seismic, infrared).

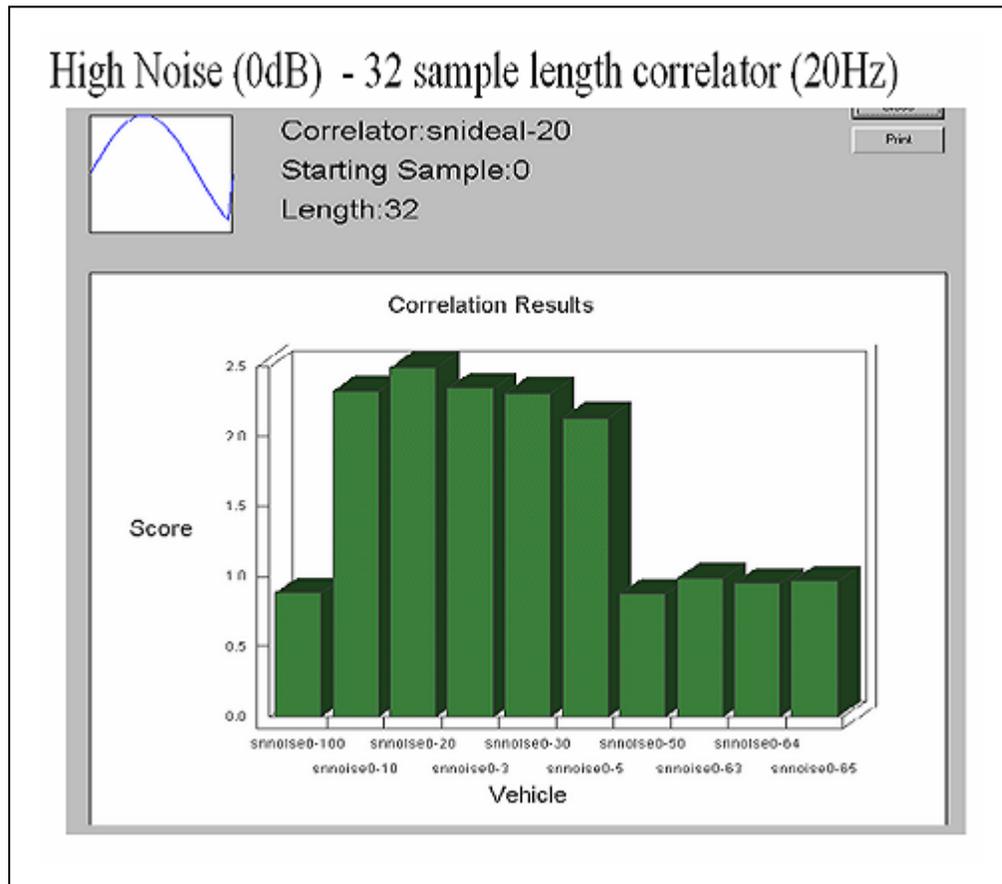
#### iv. Channel Medium Noise Effects on Signals.

Signal channel and medium noise effects were investigated with window sizes, and stepping. The investigated signals showed good results with noise up to 3 dB. However, with signals spaced widely in frequency, SNRs of up to 0 dB could be identifiable / classifiable. Required correlator template sample sizes were twice that of what was required for 4 dB SNR. It was also found that SSE was more robust to noise compared to many other parametric identification schemes, and was formidable in properly identifying signals which had wideband characteristics. Noise performance on the SSE is superior due to segmentation that eliminates low SNR components of the signal while reducing variable signal parameters. AARMS / MARMS of the windowed signal RMS values further makes the SSE robust to noise. Results in Figure 4.9 show SSE performance to signal noise effects. Histogram plots show low noise signal identification using the SSE. It is seen that for a high SNR signal, RMS scores are much higher than that of its low SNR counterpart. The signal conditioning / pre-processing module effectively handles this variable by selecting high SNR signal samples, while avoiding low SNR samples. This is performed in many ways by the signal pre-

processing modules. A combination of threshold range selection and signal windowing eliminates low SNR signals.



**Figure 4. 9SSE performance with high SNR ( 20dB ) signals.**



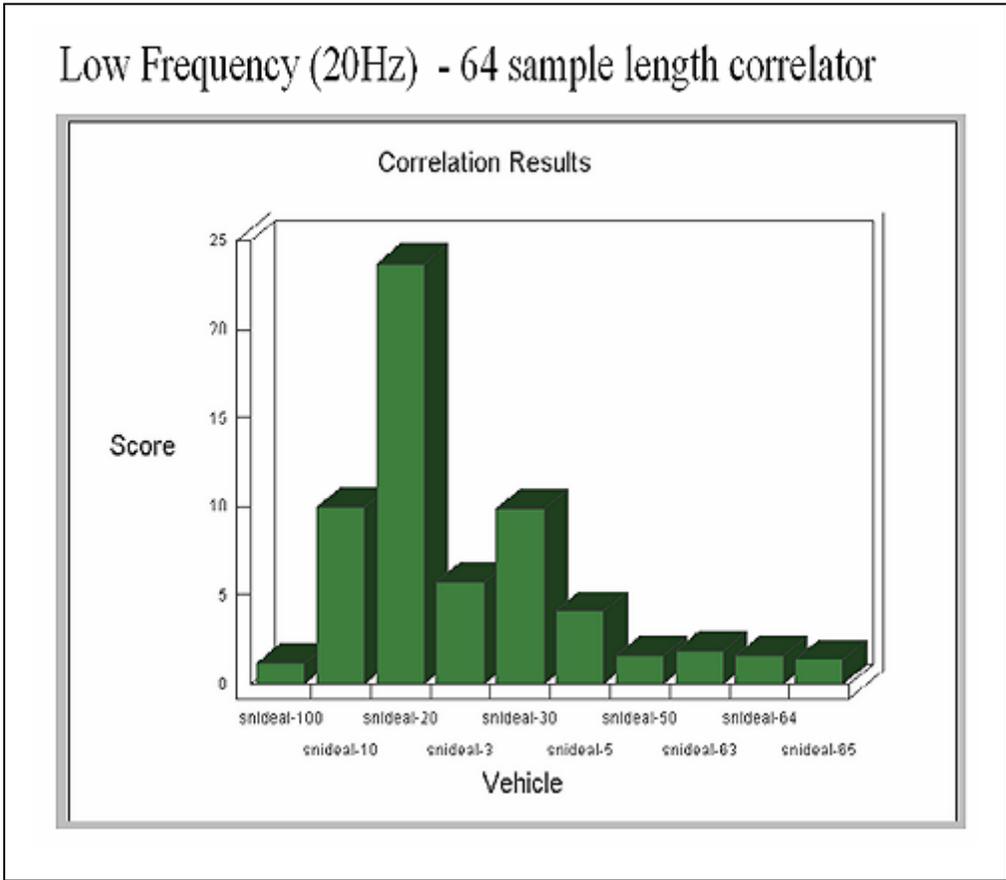
**Figure 4. 10. SSE performance with low SNR ( 0 dB ) signals.**

The histogram plot of Figure 4.10 gives results of signals with SNR of 0 dB. It could be seen that during identification of lower SNR signals, RMS scores are lower compared to higher SNR signals. It is also evident that the results are close and falls in the boundary regions when low SNR signals are present.

v. Stepping Size of Template Windows.

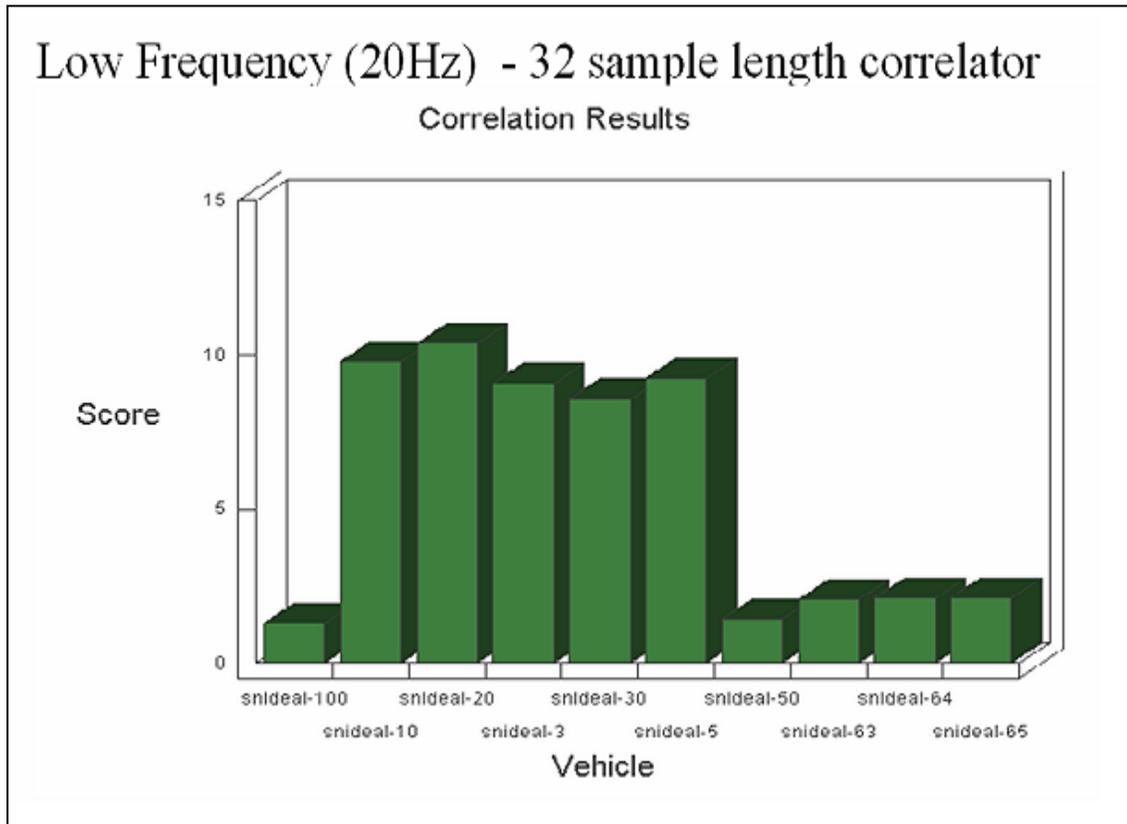
Window stepping distances are varied and their results observed. Window movements with 1 window stepping were observed to give a complete set of identification / classification at the cost of high computational power. Variable step size results were then compared to that from the 1-step RMS score as the reference. Arrays of steppings are observed and results compared. Steppings were set at a rate of  $\{1 \leq \text{step size} \leq (\text{Window Length of Template} - 1)\}$  samples. It was found that the smaller the spacing the better it was in avoiding phase offsets, as well as in detecting special features present in samples.

The following shows SSE classification runs with different stepping distances. The test results indicate that stepping distances are more application specific. This can also be observed in real world signal runs in Chapter 6. The following graphs show results on test signals.



**Figure 4. 11 Window stepping sample effect on RMS during signal correlation.**

Figure 4.11 shows identification with different window sample steppings. To obtain minimum computation cost application specific maximum stepping sample sizes are obtained. Stepping of correlator template window along signal samples depends on signals types under consideration. It is a necessity that window stepping be customized for low power depending on sensed signal characteristics for specific applications.



**Figure 4.12 SSE performance with an optimized stepping for minimum power computation.**

Figure 4.12 shows SSE identification with a smaller correlator template and different stepping compared to Figure 4.11. It was observed that the bigger the template sample size and the smaller the stepping size, the better the results at the cost of computational power. For low power operation, it is optimal to have a smaller template size compared to that of a bigger template window sample size with optimized maximum stepping.

### **4.4.3 SSE Runs on Real Signal Sets**

Results from test runs showed a very superior performance in detecting signals. SSE was found to be robust to many signal parameters namely signal strength, phase variation, and channel noise exhibiting errors of 2-5% during low SNR signal identification / classification but otherwise showing almost 100% accuracy. A combination of signal segmentation into state spaces, in addition to SCIA modules that perform AARMS / MARMS contributed to these excellent results.

Having obtained excellent results from SSE runs, the MUSIC and Pisarenko parametric methods were tested with a similar signal set to compare with SSE identification.

## **4.5 Comparison of ‘Music’ / ‘Pisarenko’ Method for Identification / Classification**

SSE performance with a test signal set was compared and benchmarked with the parametric methods namely “Pisarenko” and “MUSIC”. The Pisarenko and MUSIC methods are two of the earliest forms of parametric spectral analysis for complex signals. [57] They are used for time-frequency, spatial-transform domain spectral analysis. One

of the features of identifying and classifying data is a study of the signal spectral contents and its match it with a known spectral feature existing in a codebook. Since the signals received at a sensor node or monitoring antenna have many variables, the robustness of the identifying or classifying method needs to be studied with the above parametric methods similar to that of the SSE signal tests.

The studies of both methods were conducted with signals having various SNRs, amplitude variations, phase shifts, and different number of sources present as parameters. Each parameter is considered separately at particular noise levels similar to the SSE tests, and their limits were obtained.

It has also been found that these parametric approaches have often outperformed the interpolated FFT algorithms for short data records. [30] The two methods considered were taken for a particular sample length (25 samples) with 10 time averages with much lower computation power requirements compared to the SSE signal set. A smaller sample set than that used with the SSE was used to test the parametric methods. A sampling frequency of 1 KHz was used to create signal data with distinct frequencies in the range of 300-400 Hz. The limiting parameters were obtained by running tests on the algorithms many times. A sample set similar to the SSE signal set is used for parametric identification analysis of the 'MUSIC' and Pisarenko methods.

### 4.5.1 Noise Variations

Signals with two frequency components are considered to evaluate robustness in distinguishing spectral contents from noise variation. Two different cases were considered. In one case, the signals had frequency components that were far apart; in the other case, the signal had closely spaced frequency components. The signal-to-noise ratio (SNR) ranged from 20dB to -3dB.

When the frequency components were well separated ( $f_1=200\text{Hz}$ ,  $f_2=400\text{Hz}$ ) the noise variation did not affect the detection all the way up to an SNR of 0dB. Fig.4.13 shows

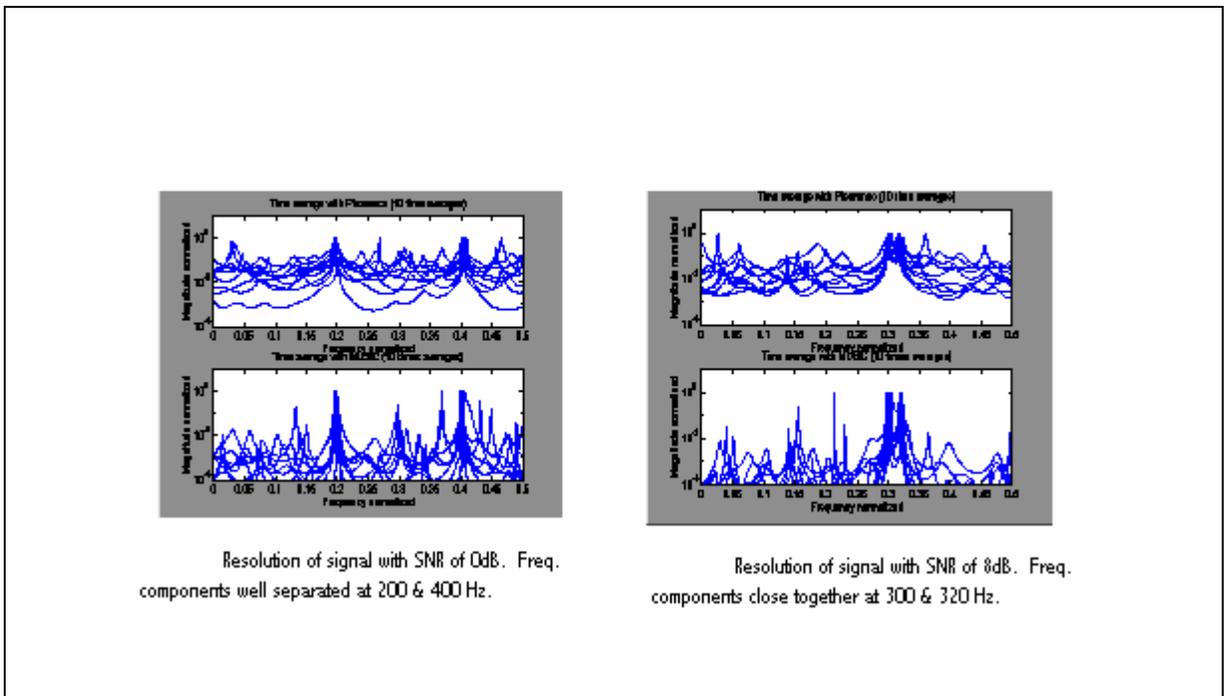


Figure 4. 13 Results from runs with noise on well-separated and close frequency contents.

the distinguishable spectral contents that can be clearly detected with high precision. The noise tolerance of the MUSIC method for well-separated frequencies even at very low SNR is evident, allowing for proper detection even in noisy environments.

Having obtained the results for the case with well separated frequency components, the second case investigated was that of a signal that had two different frequency components close together ( $f_1=300\text{Hz}$ ,  $f_2=320\text{Hz}$ ). The detection process was effective up to an SNR of 8 dB as illustrated by Figure 4.14 on the comparison of the 'Pisarenko' and MUSIC methods. With further degradation of SNR, the MUSIC method gave a slightly better outcome than the Pisarenko method, though the spectral contents were not entirely prominent from inspection of test results.

From the above results, it is concluded that this approach is a good technique to detect narrow-band signals but would not be useful for wideband signals, with closely separated frequencies.

#### **4.5.2 Amplitude Variation Effects**

The second interesting case that was studied is that of two signals, each with independent frequency components, but having different amplitudes. The detection of multiple sources with different SNR was tested using the MUSIC and Pisarenko methods. The test cases were built with amplitude ratios of 20:1, a dominant signal having a 20-fold

strength as compared to that of the less dominant signal. This is the case when one of the sources is close to the sensor while the other is farther away. When the source frequencies were far apart with a SNR of 8dB, the two distinct frequencies were clearly detectable. However when the dominant amplitude is five times higher than the non-dominant one, both frequencies could be detected with the Pisarenko method, whereas the MUSIC method did not allow for proper detection. When the amplitude ratio was further increased, only one frequency component was detectable, while the other was lost.

The detection process with the amplitude ratio as a variable was considered for two close frequency components as well. For the case of 0 dB, it was possible to detect the two distinct frequencies when the amplitude ratio was 2:1. The Pisarenko method gave a better resolution though the MUSIC method gave a peak range between the two distinct frequencies as evident in Figure 4.15. When the amplitude ratio was 5:1, both methods are limited by their resolution, allowing detection of only the dominant frequency as seen in Figure 4.15 with Pisarenko and MUSIC methods. Thus, it can be concluded that detecting sources with differing strengths under low-SNR conditions will cause problems for the weak signal. For this reason, detecting multiple sources would not be trivial if the signal strengths vary.

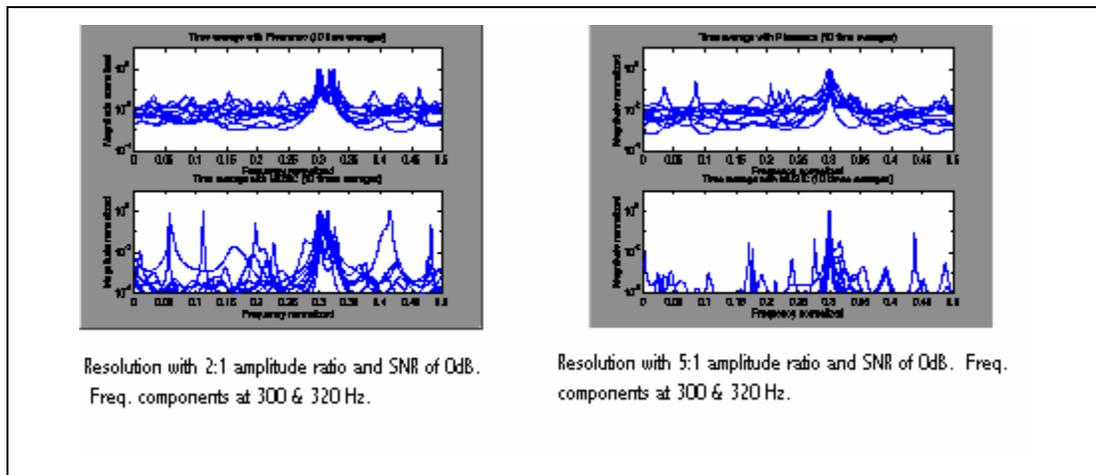
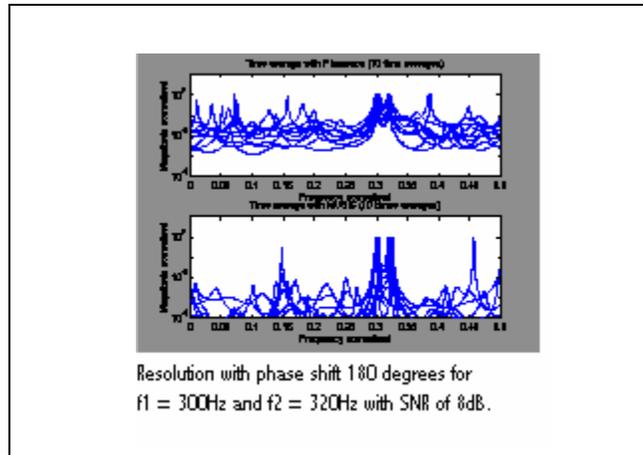


Figure 4. 14 Analysis of signal strength on closely spaced and well separated signal sets.

### 4.5.3 Phase Variation Effects

Phase variation in the signal is an important parameter, which would significantly influence the proper identification of the source. Most spectral detection methods are very sensitive to phase changes. Hence, robustness of the Pisarenko and MUSIC methods to phase variations was evaluated. A phase difference ranging from  $\pi/4$  to  $\pi$  was used for the phase offset. [30]

When the signals were far apart, it was evident that the two distinct frequencies could be detected without any frequency resolution being lost, even with a low 0 dB SNR.



**Figure 4.15** Effect of phase distortion on signal set.

The performance of the signal with close frequencies was more dependent on noise than the phase variation. When the signals were shifted by  $\pi$ , the ones with 8 dB SNR gave the correct distinct frequency, whereas for 0 dB SNR, both methods gave a frequency band between the two distinct frequencies of the source. Thus, it is inferred that though the MUSIC and the Pisarenko methods are very robust to phase shifts, its performance is based on the SNR. In comparison, SSE's close frequency signals were detected even with a low SNR of 4 dB. Thus, the SSE is more robust to phase shifts, and noise content compared to that of parametric signal identification, using MUSIC and Pisarenko methods.

#### **4.5.4 Multiple Source Identification**

Some of today's common applications include the detection of the presence of multiple sources. Since the number of sources present is usually unknown, the need to find methods to identify the frequency contents of all the sources is high. It resulted in an erroneous detection, for cases with both close and widely separated frequencies. When the frequencies were far apart, it was still not possible to see any distinct frequency. This incorrect prediction is as expected, since the algorithm used was designed to detect only when two or fewer sources are present.

When the two frequencies were close together, it gave an incorrect detection of a single frequency that was not part of the three frequencies present in the input signal. These erroneous results indicate that the Pisarenko and MUSIC methods can only be used when the number of input sources is known. This design can be implemented using more matrix properties of Pisarenko and MUSIC methods. If the numbers of sources are unknown, the design should be implemented with a larger number of sources in mind. The results validate the fact that the parametric approach does not work when the model order is incorrect, at which time adaptive ordering algorithms should be considered.

#### **4.5.5 Summary of finding of 'MUSIC' and Pisarenko Identification**

The Pisarenko and the MUSIC methods were shown to be efficient tools for the detection of discrete frequencies in low-noise situations. Thus the behavior needs to be tested under

such realistic conditions, to check their feasibility specifically for acoustic, seismic, infrared, and EKG applications. One method is to decompose these frequency band signals into discrete frequency components and use the Pisarenko and MUSIC methods to find the threshold of frequencies present, thereby giving a good classification of the incoming data. Though this concept may be conceptually correct, its feasibility is yet to be seen.

This detection method is mainly for stationary signals and it may be extended to non-stationary signals with the help of snapshots, at time instances for it to achieve stationary behavior. Further, this method should be extended to be robust relative to the number of sources present by following adaptive ordering techniques especially to be tried for wideband signals where the discrete frequency spectrum could be modeled with a comb train with variable amplitudes. One form of source finding would be to use beamforming techniques and then separate the signals and focus on an individual source rather than changing the algorithm altogether for multi-source identification.

The above parametric methods are weaker compared to the SSE in identification of closely spaced signals. They further have erroneous results for low SNR signals. It was also observed that the dominant signal could be easily detected whereas a signal with less strength could be lost as a noise component. SSE performance was more robust to low SNR signals, and had higher accuracy levels in comparison to parametric identification. It should also be pointed out that the parametric methods could not detect multiple

sources when the number of sources was more than the algorithm initial design assumed. However, the SSE algorithm performed well in detecting multiple source model signals, while being more robust to noise.

## **4.6 Summary**

Test signals were created with various parameters and tested extensively on the SSE. Performance of the SSE to amplitude, noise, phase offset, template window size variations, and window stepping distance was investigated. It was concluded that the SSE accuracy was very good for the test signal models, with a 2-5% error on low SNR (0dB) signals and correct identification of signals with higher SNR. SSE performed similarly for phase distortion. Signal strengths were irrelevant to the SSE due to it having a signal window normalization module. Template window size did give a scaling on the RMS values. However it fell within the range of classification / identification boundaries. Investigations also revealed that the templates consisting of less than one wavelength (with distinct features) were successful in identifying or classifying the signals successfully.

A comparison of signals with parametric approaches revealed that the SSE performed much better in noise, with high identification accuracies. Benchmarking of created test signals was done with the Pisarenko and MUSIC methods. SSE's performance was a lot more accurate in detecting multiple source presence. It was also concluded that

'MUSIC', and Pisarenko methods were not able to detect wideband, or closely separated narrowband signals (low SNR), whereas the SSE was more suited for this case. Benchmarking of SSE with that of the competing signal identification methods was also done with real world signals and is presented in Chapter 6.

## **CHAPTER 5**

### **Classification & Identification Algorithm for SSE**

#### **5.1 Introduction**

The Signal Search Engine (SSE) for Wireless Integrated Networked Sensors (WINS) receives signals with many environmental, situational, and conditional signal effects. Therefore, it is necessary to pre-process signals before the Signal Classification / Identification Algorithm (SCIA) becomes active. Many of today's wireless sensor networks require low power, high capacity, high-throughput, robust algorithms to incorporate a multitude of signals for classification and / or identification. This chapter includes a description of how pre-processing of signals is undertaken before activating the SCIA along with the variables that are acted upon to reduce signal variables. The SCIA algorithm performed well in identifying and classifying moving sources, as well as a multitude of narrowband sources with the addition of the pre-processing module.

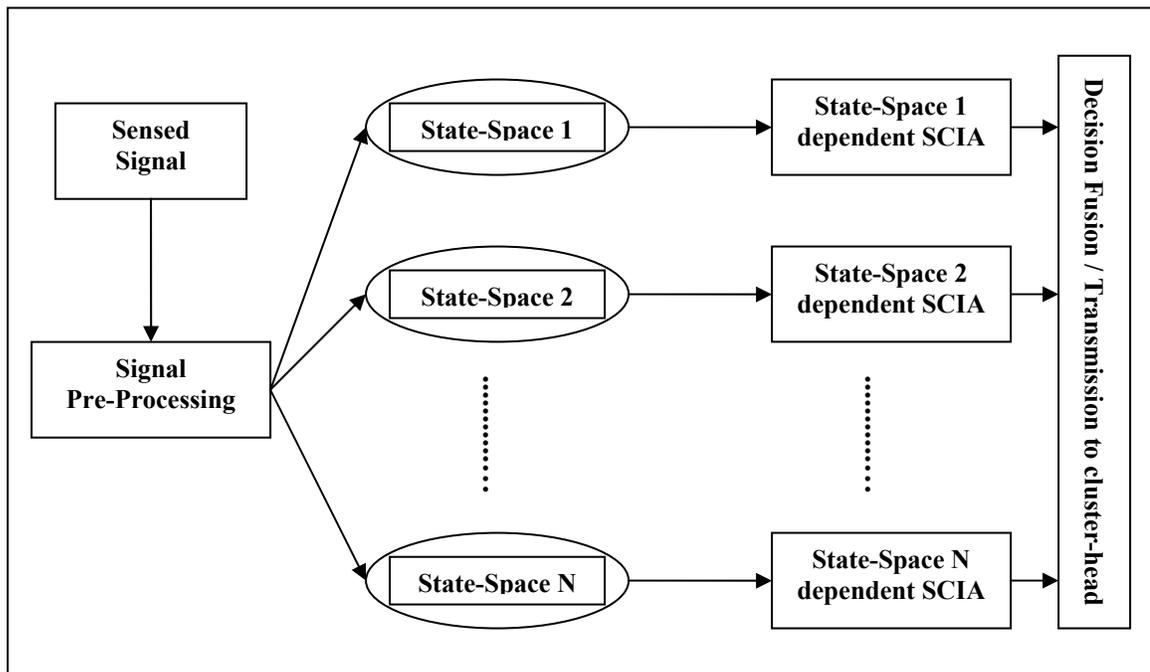
Results obtained from real world signals are presented in Chapter 6, along with correlator template selection, classification / identification, and decision-making.

The SSE algorithm deals with a broad range of signals that include customized application specific modules. Here we concentrate on classification and identification of moving sources, that has the most complex waveforms with a multitude of signal variables. Other less complex signal waveforms can be classified and identified with this SSE algorithm with the addition or elimination of application specific modules. This has been proven with created signal models as seen in the previous chapter.

A two-step approach is taken in explaining the signal classification / identification scheme. The initial step is signal pre-processing. In this module signals go through noise reduction, segmentation and are assigned to state spaces. Signal segmentation is a crucial module of the SSE, and is designed using a state-space approach. The SSE algorithm exploits state-space information for proper classification / identification that was unrealistic in the time domain before the addition of this method. Once signals are decomposed into different state-spaces, they are fed into the SSE and are parallel piped into state-dependent processing modules. This parallel processing activity is more intense when signal state types are present along with state spaces as in cases of M-SSE. Multi-sensing sensors are assigned state spaces dependent on state types for acquired signals. State space assignments were a crucial breakthrough for proper classification / identification with enhanced accuracy levels as seen from results presented below.

Once these signals are submitted for parallel processing, they undergo signal classification or identification according to the customized SSE / M-SSE modules. The SCIA algorithm properly classifies and identifies signals according to ‘Type Abstraction Hierarchy’ (TAH), or ‘Type Identification Hierarchy’ (TIH) depending on client / user requirements. It properly traverses the TAH or TIH tree designed with a training set. A method of choosing the correlator templates and building the tree is shown in Section 5.5 of this chapter.

State space assignments are accomplished in many different forms. Figure 5.1 shows an example of state-space dependent signal pre-processing along with parallel SCIA for



**Figure 5. 1. State-Space approach to SSE event Identification / Classification.**

source identification and classification. Signal pre-processing assigns state spaces to signals depending on application needs. A moving vehicle waveform will have state-spaces such as approach, arrival, and departure assigned to it. Furthermore, condition-based maintenance sensor readings associate lifetime with the observed phenomena (e.g. brake change / maintenance levels assigned state1: up to 10 K miles, state 2: 10-20K miles ...etc.). The state-space based approach to SSE gives more modular control, and methodically divides and conquers signal redundancies.

Once classified, the results undergo confidence measure assignment as shown in Section 3.4.3 of Chapter 3. The decision making process is explained in detail for both ‘distributed’ and ‘centralized’ architectures in section 3.5 of Chapter 3.

The implemented time-domain algorithm was obtained and modified while experimenting with complex signals that had a multitude of signal variables. The complexity of signals becomes more pronounced especially with moving sources. This is the case when sensor readings do not have information such as location, or other variables such as speed. When such information is not available, these complex signals cannot have high precision-specification requirements such as the identification of vehicular speeds, or angle of approach. Contrary to the above, if there is a requirement for determining vehicular speeds or angle of arrival, time-frequency information can be exploited to obtain the requested information. The SSE algorithm by itself gave excellent results from SSE runs without time frequency information. Present day tracking

algorithms obtain speed information using source localization and identification. An explanation of why time / frequency information needs to be exploited for high precision identification / classification is shown with time frequency spectrogram results obtained from real-world sensor readings. These waveforms are studied to investigate sensed waveform behavior in sections 5.2.1 – 5.2.5, and were initially used to find moving object waveform characteristics.

## **5.2 Moving Source Variables**

Special concentration is directed at identifying and classifying moving object waveforms because numerous variables are present in the acquired signals. [46, 56] These waveforms are methodically segmented and decomposed into state spaces to diminish signal variability while optimizing for low power. Signal variables that are dominant in moving source waveforms are described below with an introduction to the state space concept for SSE. Spectrogram plots are used to study signal characteristics and model them mathematically, while observing time –frequency behavior. However, spectrogram plots are not included in the SSE, where time –frequency behavior is calculated in the time-domain.

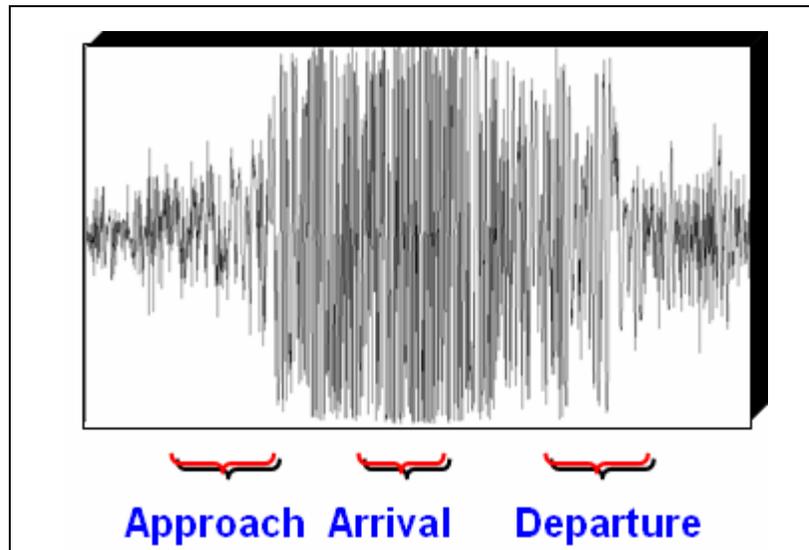


Figure 5. 2. Moving source waveform with state-space decomposition.

### 5.2.1 Doppler Shifts

Moving objects exhibit frequency dependence relative to time. Frequency variation relative to time is dependent on the velocity of the moving object (direction relative to the sensor gives scaling in speed of the vehicle) and causes Doppler effects. Doppler frequency shifts need to be detected initially and mitigated by pre-processing to diminish the effect of rapidly varying frequencies over short periods of time. One of the main functions of the signal pre-processing algorithm is to reduce these redundancies. Fast frequency variations cause widespread and non-conclusive variation in algorithm results, and thus need to be acted upon before the signal reaches the SCIA.

Moving objects present additional signal variables that are mitigated by the state space approach shown in Figure 5.1 and by the use of signal segmentation and state space assignment shown in Figure 5.2. Such pre-processing and signal segmentation not only mitigates the Doppler effect to a certain extent, but also eliminates other dominant variables that are discussed in section 5.2.2 – 5.2.5. When signal segmentation is not present, frequency variation within the same signal is dominant and therefore produces significant effects that make classification and identification unreliable. The following methods are considered to reduce Doppler shifts due to vehicular movements:

- Signal segmentation into approach, arrival, and departure state spaces for moving sources.
- Choosing ‘correlator’ templates in the highest SNR region within average frequency limits, between the highest and lowest frequency band for each individual state-space. (i.e. using time–frequency information for the arrival phase (CPA: Closest Point of Approach) to find frequency regions to choose the correlator template where Doppler shift is minimal.)
- Including phase variation mitigation efforts in the SCIA.
- Including time-frequency spectrogram information during pre-processing (before signal segmentation) where any abrupt frequency changes are identified and segmented into sub-spaces. However, use of time frequency information is feasible only with sensors with a reliable energy supply or

during off-line training. Low energy wireless sensor nodes to the contrary use time domain equivalents corresponding to time-frequency information. Time – frequency information is further decomposed into sub-spaces where source gearshift information could be included in the correlator template tree for more precise identification.

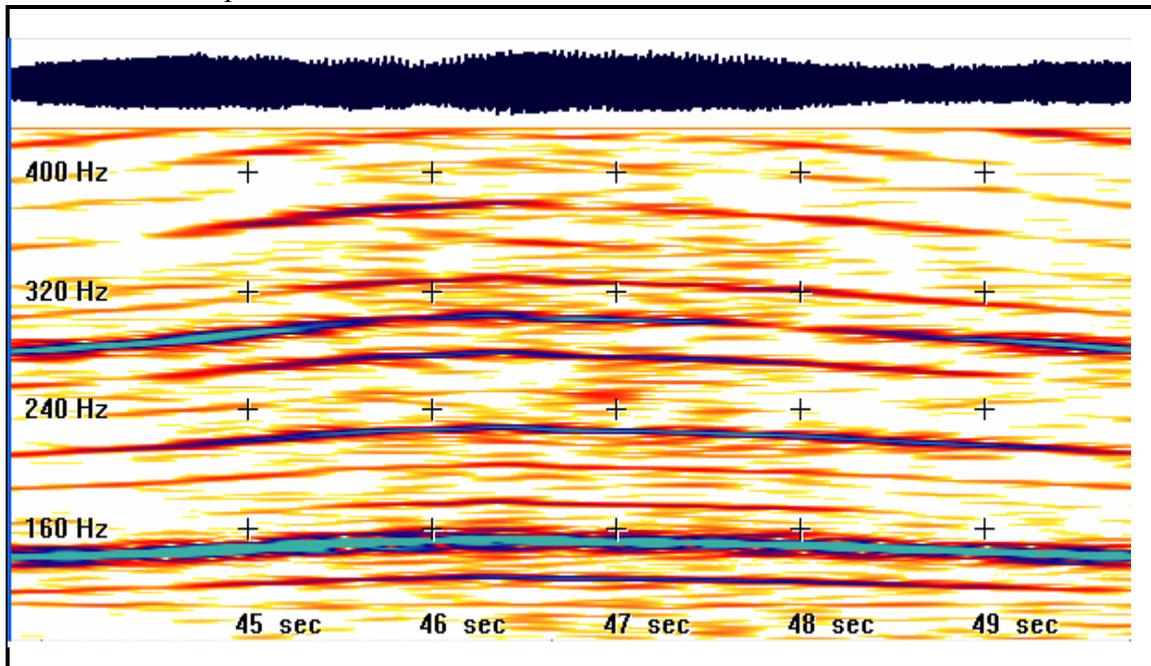


Figure 5. 3. Time-Frequency behavior showing Doppler effects on high-speed (30 km/h) source raw data.

The above methods were used to pre-process and post-process signals obtained during real world signal acquisition. A training set was utilized in an extended study for features present in moving sources that helped in refining the SCIA.

Figure 5.3 shows the presence of Doppler on a moving heavy wheeled vehicle at 30 km/h. Acoustic microphones sensing these signals were located at a distance of 25-100m from the moving source. The presence of rapid frequency variation is reduced by the

state space approach implemented in the pre-processing module. The same vehicle traveling at a lower speed of 15 km / h shows minimal Doppler as depicted in figure 5.4. However, even the presence of minimal Doppler is averted during state-space segmentation, and is one of the fundamental building blocks for accurate classification and identification in the time domain.

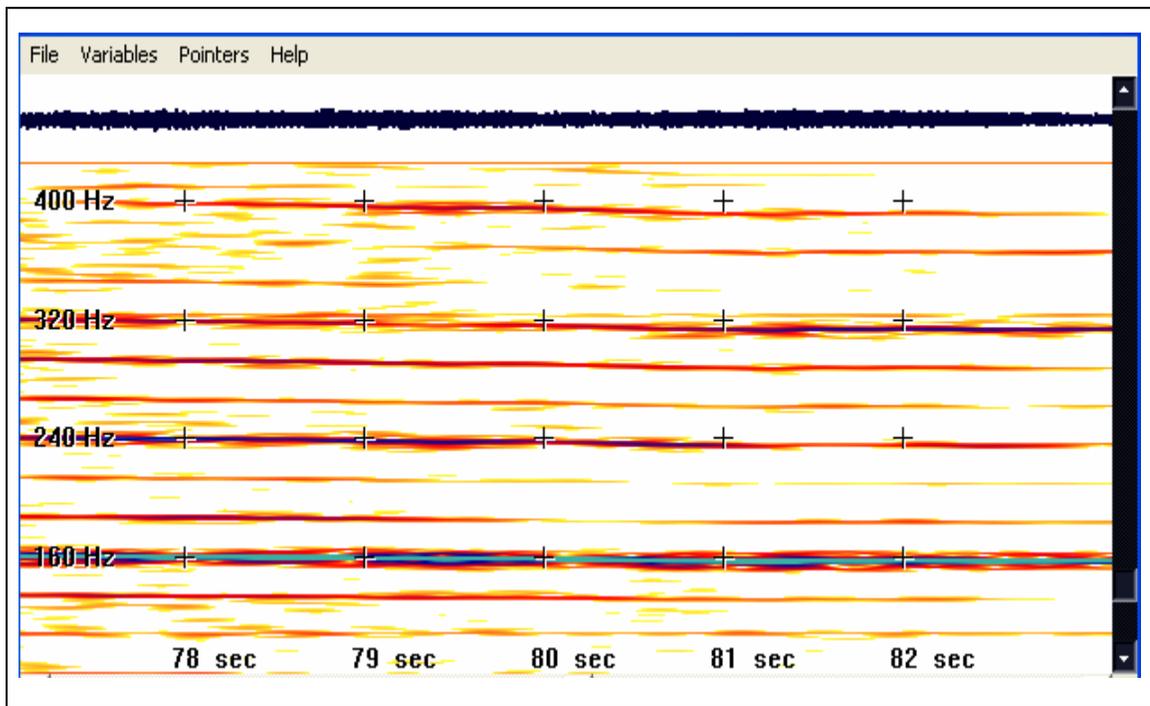


Figure 5. 4. Time-Frequency behavior showing minimal Doppler effects on low-speed (15km/h) source data.

## 5.2.2 Abrupt Frequency Loss / Gain

Abrupt frequency changes in source waveforms create perturbation or loss of continuity in the signal waveform that causes inconclusive or erroneous results to show up during classification / identification with SSE / M-SSE. Abrupt frequency changes are common during gearshifts in moving sources due to changes in road conditions or obstacles in the path of travel. When a moving source encounters a random obstacle in the form of a ditch or bump, this obstacle causes either gearshifts or slowing of the source that affects signal parameters. Some of these situational circumstances prompt the moving source to make adjustments in motion in the form of gearshifts (sometimes introducing some multiple quantity gear down/up shifts) that introduces multiple Hz of frequency loss or gain.

A study of time frequency information points to a 3 – 7 Hz frequency drop during one gearshift (up or down) with linear increases for multiple shifts. These abrupt changes call for further state-space decomposition of the signals into sub-spaces. Segmentation of signals into sub-spaces is a necessity for precision classification / identification (i.e. classification / identification with vehicular speeds). Figure 5.5 shows abrupt losses and gains in frequency during the 99-100 sec. interval and again during the 101.5 – 102.5 sec. time interval.

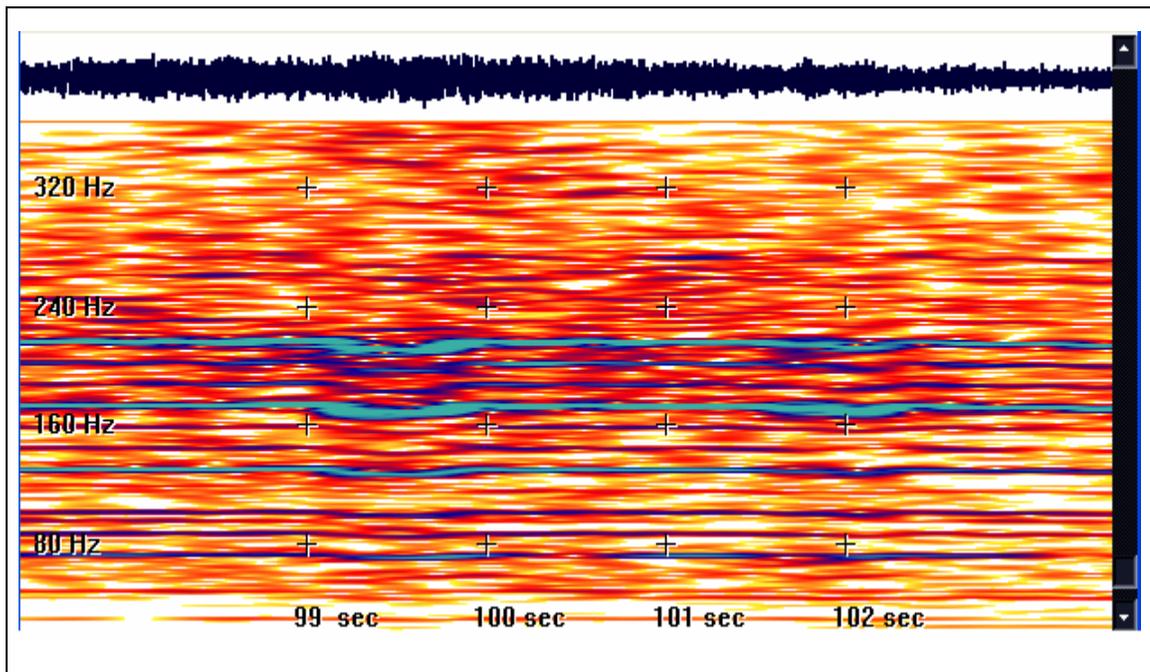


Figure 5. 5. Abrupt frequency loss / gain as seen by an acoustic microphone in a moving source.

Time / Frequency spectrogram information (on raw data) is exploited to find abrupt frequency changes due to the above-described scenarios including Doppler shift. These situations show a 3 - 7 Hz or more drop in frequency within a short time period ( $\sim \Delta(t) \rightarrow 0$ ). Frequency drops in raw signal data are segmented into state-space components in the classification tree. Subdivisions of these signals are separated into state sub-spaces for identification. Though minor in affecting the outcome, segmenting signals when abrupt frequency changes occur enhances the confidence measure considerably and minimizes errors in classification and identification results.

### 5.2.3 Directional Change in Path of Travel

The speed of the vehicle does relate to frequency variation through Doppler shifts. However, the effects of directional change of motion in path of travel further affect parameters such as angle of signal arrival at each sensor and scaling of source frequency. Vehicular speeds are closely coordinated with vehicle gear segmentations, and the axis of source travel. When sensors are stationary, (unlike mobile sensors) moving sources having the same speed but taking different paths exhibit path dependent Doppler frequency shifts. [26] Figure 5.6 shows the same vehicle taking different paths. The angle of arrival of a wideband source waveform is shown for cases where the distance of travel is the same ( $d_1 + d_2$ ). Doppler for the above case is derived as shown below:

**Case 1:** When the source is approaching the sensor with a speed of  $v$  km/h, the wideband Doppler shift  $\Gamma_d$  is:

$$\Gamma_{d1} = \frac{v}{\lambda_w} \sin\theta_1 \quad (5.1)$$

where  $\lambda_w$  is the wavelength of a dominant narrowband frequency and  $\theta_1$  is the angle of wideband signal waveform arrival at the sensor node.

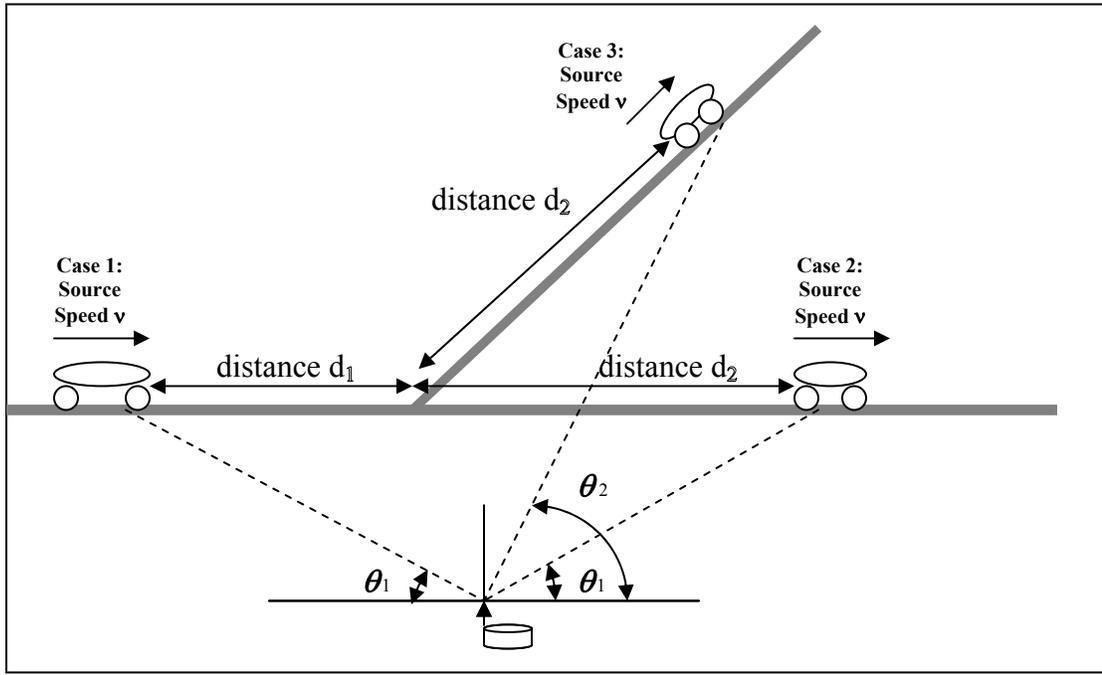


Figure 5. 6. Doppler from same source moving in different directions.

**Case 2:** When the source is away from the sensor at a distance  $(d_1 + d_2)$  traveling at  $v$  km/h:

$$\Gamma_{(d_1 + d_2)} = - \frac{v}{\lambda_w} \sin \theta_1 \quad (5.2)$$

The above equation is symmetric with equation (5.1) above, but shows a decay when the source travels away from the sensor.

**Case 3:** This case pertains to a source that travels in a different path, but has moved the same distance as that of case 2. Here, the reference angle contributes to a slower frequency decay than that of case 2 and is given by the following equation

$$\Gamma_{(d1 + d2)} = - \frac{v}{\lambda_w} \sin \theta_2 \quad (5.3)$$

Equations (5.1) – (5.3) show path dependent Doppler shifts that were observed in analyzed signals. State space segmentation was used effectively to eliminate these redundancies and was effective as seen in results to be presented in Chapter 6.

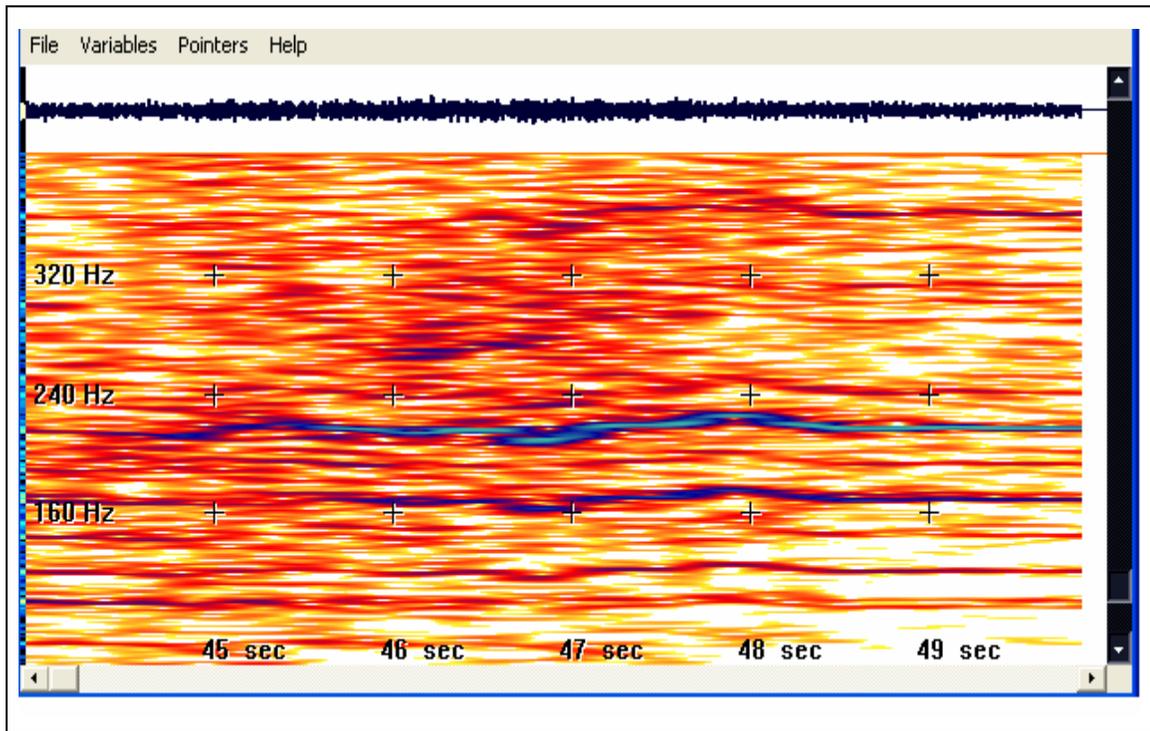
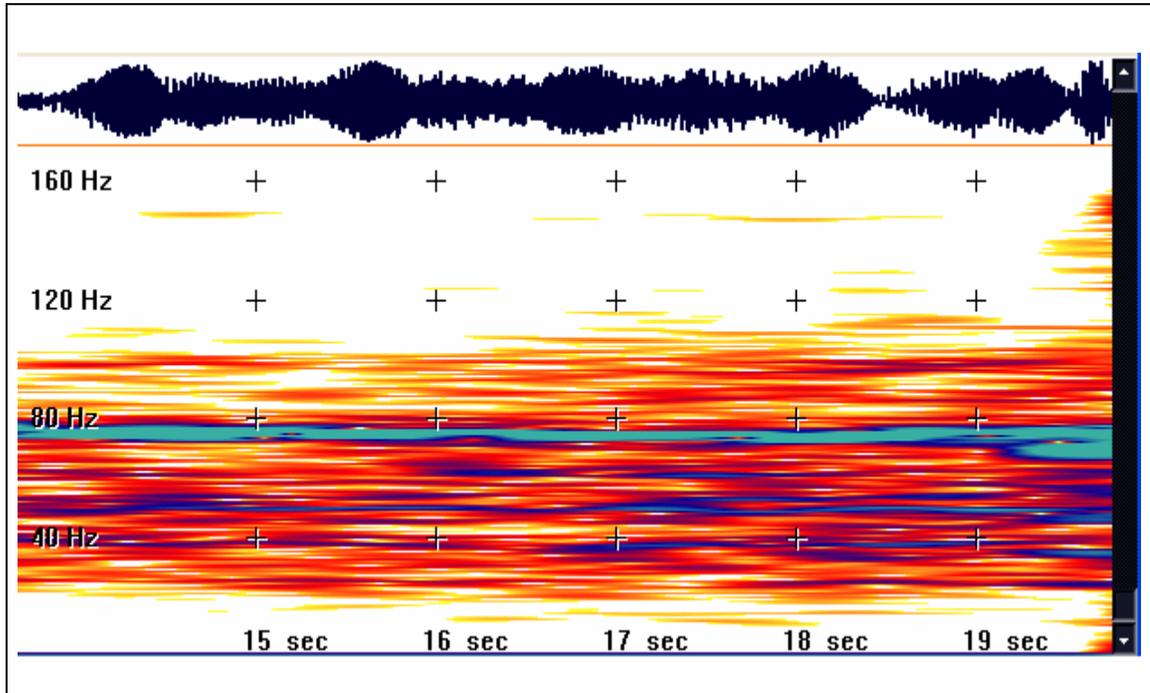


Figure 5.7. Shows an acoustic microphone detected time-frequency information.



**Figure 5. 8. Seismic geophone detected time-frequency information for the same run as above.**

Figure 5.7 and 5.8 show readings for an acoustic and seismic sensor for the same run. A heavy tracked vehicle with a constant velocity of 15 km / h traveling in normal terrain is illustrated. Acoustic signals were more sensitive to Doppler shifts and abrupt frequency loss / gain compared to their seismic counterparts. The above readings are for the vehicle at the same distance away from both acoustic and seismic sensors with seismic readings showing a delay due to the threshold for sensing low pass seismic signals.

## 5.2.4 Glitches and Saturation of Sensor Readings

Other occurrences of abnormal signal behavior are due to circuit effects on sensor readings. Glitches, saturation, or other circuit disruptions may occur during sensor time-outs or if sensing circuits are not tuned accordingly. These glitches or discontinuities cause the algorithm to generate low threshold RMS values during classification / identification that fall in the gray area of decision boundaries. This effect is corrected by the following methods that are used during calculation of the final threshold value.

- Avoid data obtained during the presence of glitches and exclude them in correlator window calculation of average data. Eliminating glitches is done with a thresholding operation observing a sudden loss of continuous signal amplitude gain followed by a sudden increase in continuous amplitude gain.
- Have a signal conditioning system module to minimize the effects of glitch and discontinuities, without introducing phase offset, abrupt frequency gain / loss, or amplitude effects.
- Detect uneven spikes or data corruption during the pre-processing stage by having a minimum and maximum threshold value for windowed segments to be process worthy. This will help in determining signals that are above the

average threshold and in detecting uneven spikes (above a maximum threshold) that may exist in collected data sets.

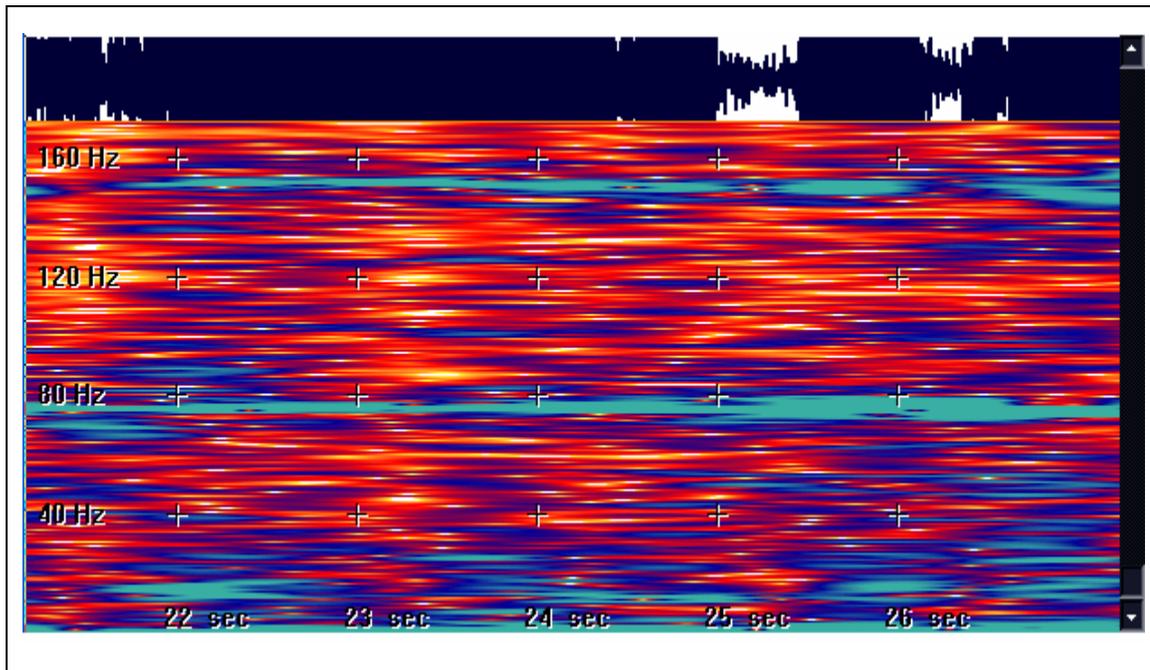


Figure 5. 9. Saturated signal readings are avoided during signal pre-processing.

- Avoid initial noise and initial signal accumulation segments where the SNR is low. Have a threshold of 0 to 4 dB minimum SNR where signals are considered process worthy.

Having discussed methods that mitigate the effects of glitches, we now present steps to avoid saturated signals at the CPA as exhibited in real world sensor readings. Saturation occurs when the vehicle is nearest to the sensor and signal strengths are at peak level. Fine-tuning sensors is needed when this effect is observed. However, saturated signal segments are excluded during signal pre-processing due to fine frequency features of signals being lost. A threshold value of signal strengths for a maximum value is set and

any segment that shows signal amplitudes of greater than this maximum threshold is discarded during signal pre-processing. Figure 5.9 shows a saturated seismic signal. It is evident that the signal cannot be used to obtain features for identification / classification.

### **5.2.5 Methods used to Mitigate Variables and Generalizations**

The above-mentioned variables are mitigated or handled by various pre-processing steps. Effective initial signal conditioning, along with various signal segmentation methods help to create a robust SSE. Both frequency information in the time-domain, and time-domain signal data are exploited to achieve signal uniformity. An explanation of block diagrams and their functionality give a description of what steps are taken during signal pre-processing.

## **5.3 Time-Frequency Observations**

Spectrogram plots were initially created to investigate and determine time frequency characteristics of the sensed signals. However they are not used in the SSE. These spectrogram plots were used to investigate frequency behavior relative to time. This information is more useful in real-world signals where signal frequencies tend to vary depending on source mechanics and dynamics. These observations were transformed into mathematical formulas to model real world signal behavior. [2, 20, 22, 23] The following equations model abrupt frequency gain / loss as experienced during situational

conditions in the path of travel. A wideband source signal  $S_w(n)$  sensed at the sensor is given by:

$$S_w(n) = \sum_{m=1}^M A_w^{(m)} S_w^{(m)}(n - t_w^{(m)}) + \mathbf{A}_{gu}^{(m)} * \Phi_{gu}^{(m)} * U(n - \tau_{gu}^{(m)}) - \mathbf{A}_{gd}^{(m)} * \Phi_{gd}^{(m)} * U(n - \tau_{gd}^{(m)}) + \psi_w(n) \quad (5.4)$$

where  $\mathbf{A}_{gu}^{(m)}$ ,  $\mathbf{A}_{gd}^{(m)}$  is the number of gear up-shift and downshift component matrices of  $m$  sources.  $\mathbf{A}_{gu}^{(m)}$ ,  $\mathbf{A}_{gd}^{(m)}$  are represented in equations 5.5 and 5.6 where superscripts (1,1), (1,2),...,(1,r) depict the 1<sup>st</sup>, 2<sup>nd</sup>, and r<sup>th</sup> instances respectively for source 1, and (m,1), (m,2),...,(m,t) depicting the 1<sup>st</sup>, 2<sup>nd</sup>, and t<sup>th</sup> instances respectively for source  $m$ . It is also evident that the matrix has varying row sizes with  $r, s, \dots, t$  instances for different sources.

$$\mathbf{A}_{gu}^{(m)} = \begin{Bmatrix} a_{gu}^{1,1} & a_{gu}^{1,2} & \dots & a_{gu}^{1,r} \\ a_{gu}^{2,1} & a_{gu}^{2,2} & \dots & a_{gu}^{2,s} \\ \vdots & \vdots & \vdots & \vdots \\ a_{gu}^{m,1} & a_{gu}^{m,2} & \dots & a_{gu}^{m,t} \end{Bmatrix}; \quad (5.5) \quad \mathbf{A}_{gd}^{(m)} = \begin{Bmatrix} a_{gd}^{1,1} & a_{gd}^{1,2} & \dots & a_{gd}^{1,u} \\ a_{gd}^{2,1} & a_{gd}^{2,2} & \dots & a_{gd}^{2,v} \\ \vdots & \vdots & \vdots & \vdots \\ a_{gd}^{m,1} & a_{gd}^{m,2} & \dots & a_{gd}^{m,w} \end{Bmatrix}; \quad (5.6)$$

Individual frequency gain and loss for  $m$  sources are given by matrices  $\Phi_{gu}^{(m)}$  and  $\Phi_{gd}^{(m)}$  respectively. Again, the frequency gain and loss are variables that are dependent on the distance and speed of the source relative to the sensor.

$$\Phi_{gu}^{(m)} = \begin{Bmatrix} f_{gu}^{1,1} & f_{gu}^{1,2} & \dots & f_{gu}^{1,r} \\ f_{gu}^{2,1} & f_{gu}^{2,2} & \dots & f_{gu}^{2,s} \\ \vdots & \vdots & \vdots & \vdots \\ f_{gu}^{m,1} & f_{gu}^{m,2} & \dots & f_{gu}^{m,t} \end{Bmatrix}; \quad (5.7)$$

$$\Phi_{gd}^{(m)} = \begin{Bmatrix} f_{gd}^{1,1} & f_{gd}^{1,2} & \dots & f_{gd}^{1,u} \\ f_{gd}^{2,1} & f_{gd}^{2,2} & \dots & f_{gd}^{2,v} \\ \vdots & \vdots & \vdots & \vdots \\ f_{gd}^{m,1} & f_{gd}^{m,2} & \dots & f_{gd}^{m,w} \end{Bmatrix}; \quad (5.8)$$

The time when these gear-up and down shifts occur is given by the time matrix  $\tau_{gu}^{(m)}$  and  $\tau_{gd}^{(m)}$  for m sources. Gearshifts are random depending on situational conditions and therefore are given as a variable with step functions that result in time-dependent matrix values.

$$\tau_{gu}^{(m)} = \begin{Bmatrix} t_{gu}^{1,1} & t_{gu}^{1,2} & \dots & t_{gu}^{1,r} \\ t_{gu}^{2,1} & t_{gu}^{2,2} & \dots & t_{gu}^{2,s} \\ \vdots & \vdots & \vdots & \vdots \\ t_{gu}^{m,1} & t_{gu}^{m,2} & \dots & t_{gu}^{m,t} \end{Bmatrix}; \quad (5.9)$$

$$\tau_{gd}^{(m)} = \begin{Bmatrix} t_{gd}^{1,1} & t_{gd}^{1,2} & \dots & t_{gd}^{1,u} \\ t_{gd}^{2,1} & t_{gd}^{2,2} & \dots & t_{gd}^{2,v} \\ \vdots & \vdots & \vdots & \vdots \\ t_{gd}^{m,1} & t_{gd}^{m,2} & \dots & t_{gd}^{m,w} \end{Bmatrix}; \quad (5.10)$$

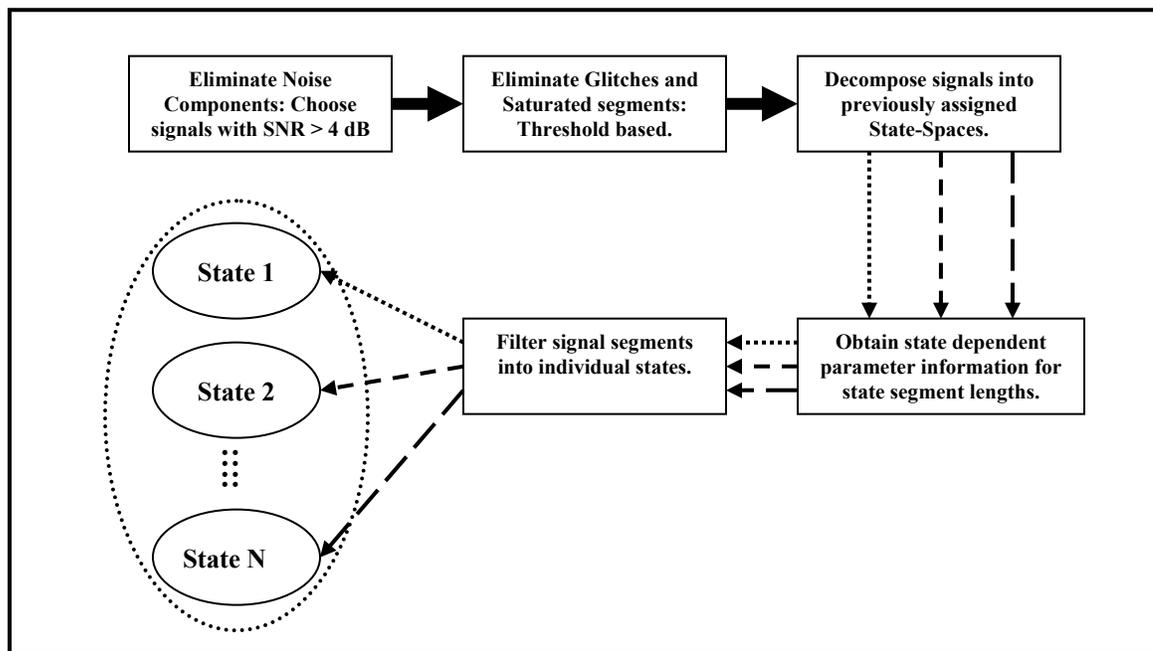
Wideband signals consist of dominant narrowband frequencies as observed in time-frequency information with the above information taken into consideration during signal decomposition into state spaces.

## 5.4 Time Domain Processing

### 5.4.1 Signal Segmentation

An important feature of the SSE is signal segmentation. Signal segmentation is required to reduce many signal variables and to create much needed uniformity and alignment

between ‘correlator’ templates, and corresponding signals to be detected. Many of the signal variables are mentioned in sections 5.2.1 – 5.2.5 above. Initial efforts to determine proper classification / identification were not successful in the time-domain SSE without signal segmentation since redundant parameters existed in the acquired signals that caused non-convergence. [40, 41]



**Figure 5. 10. Signal Pre-Processing associated with State Space decomposition.**

Signal segmentation is handled methodically using results acquired from both time-domain time-frequency information and time-series data during offline processing of time series raw data from each wireless node. Figure 5.1 presents block diagrams of the signal segmentation process. Figure 5.10 shows decomposition of signals into state spaces by methodical signal elimination and segmentation. State-dependent filter banks are associated with the windowing operation as a final step. More precise classification and

identification could consist of sub-state decomposition that is beyond the scope of this dissertation.

Once initial data processing is complete, each acquired signal undergoes data cleanup of each signal segment. Acquired signals undergo elimination of noise components and the retention of signals with a higher SNR. SNR requirements depend the sensor types and the associated source distance from sensors. The signal envelope is detected initially for all acquired data samples. The envelope is then used to clean data that have low SNR's. A full set of signal samples is then processed for time-frequency information in the time domain. This set also undergoes further time domain pre-processing before the SCIA algorithm is exercised.

## **5.4.2 Methodology for Signal Segmentation**

### **Time-Frequency Information**

Pre-processed data undergoes processing of time-frequency information in the time domain using time evolving snapshots. Analysis of time-frequency information was done with plots as seen in Figures 5.3-5.5, 5.7-5.9. Moving source gear shifts corresponding to

$$\Delta f \rightarrow \geq 3-7 \text{ Hz}; \quad \delta(t) \rightarrow 0;$$

were taken into consideration and avoided during state-space decomposition when gearshift associated state sub-spaces were not present in the classification / identification tree. In addition, any abrupt frequency losses were avoided during decomposition into signal states.

**Envelope Detection:** Detection of arrival, approach, and departure could also be clearly characterized or formulated by information present during the processing of time-frequency information in the time domain. Frequency shifts relative to time are minimal during the approach and departure state of sources, which is dependent on the distance of sources from the sensors:

$$\Delta f \leq 5 \text{ Hz; ( for } 0 \leq L \leq \text{MASSL) \quad Approach State: } \delta(t) \rightarrow 0;$$

$$\Delta f \geq -5 \text{ Hz; ( for } 0 \leq L \leq \text{MDSSL) \quad Departure State: } \delta(t) \rightarrow 0;$$

***MASSL** : Maximum approach state signal length.*

***MDSSL** : Maximum departure state signal length.*

An additional state defined (other than approach, and departure) is the arrival state (when sources arrive at the “closest point of approach” - CPA) of the sensor nodes. These states are classified separately since signal variables present in this state are sensitive to many variables. The arrival state consists of sensitive signal transformations to slight source movements along with the nature of fluctuating data with high SNR. The arrival state can be further narrowed depending on the application, and sensitivity of the source data.

## 5.5 Building the Template Tree

The SSE / M-SSE contain a template tree that is used for classification / identification. State spaces and state types each contain separate template trees and classify / identify signals coming from each state accordingly. Building a template tree and updating it as the signal database increases in size enhances accuracy levels of the SSE considerably. Template trees are differentiated into increasing ‘tiers’ which contain more precise identification with increasing hierarchy. The following section shows how classification and identification trees are built.

### 5.5.1 Classification

Classification of signals is the superset of identification [98]. Classification trees are built according to information obtained from a training set. Individual correlator templates are picked according to the method shown below. These templates are then chosen to represent the correlator tree structure. The classification tree is built according to the method shown below.

```
{  
    Choose a correlator template from a training set.  
    Choose signals from the same state space/type with different but known  
    classes.  
    Input signals into SCIA and obtain AARMS / MARMS.  
    From the results, obtain binary decision matrix as shown in figure 5.11.  
}  
    Perform above operation with other correlator templates from same class.
```

Correlator templates are selected for classification and identification using the following matrix table. Selected correlators are trained with a previously known signal set using the SCIA algorithm. Obtained RMS values are checked for p-n coverage (positive classifications are assigned “**p**” and erroneous classifications are assigned “**n**”). Binary assignments are subsequently used to pick tier 0, tier 1, and tier 2 correlator templates respectively. The correlator selection algorithm was used with a real world training set and is elaborated on Section 6.3 of Chapter 6.

	ST <sub>1</sub>	ST <sub>2</sub>	SW <sub>3</sub>	...	SW <sub>M</sub>
CT <sub>1</sub>	p	p	n	..	<b>n</b>
CT <sub>2</sub>	p	p	n	..	<b>n</b>
CW <sub>1</sub>	n	n	p	..	<b>n</b>
:	:	:	:	..	:
CW <sub>N</sub>	<b>n</b>	<b>n</b>	<b>p</b>	..	<b>p</b>

	ST <sub>1</sub>	ST <sub>2</sub>	SW <sub>3</sub>	...	SW <sub>M</sub>
CT <sub>1</sub>	1	1	1	..	<b>1</b>
CT <sub>2</sub>	1	1	1	..	<b>1</b>
CW <sub>1</sub>	1	1	1	..	<b>0</b>
:	:	:	:	..	:
CW <sub>N</sub>	<b>1</b>	<b>1</b>	<b>1</b>	..	<b>1</b>

Figure 5.11. Correlator template selection for tier 0, tier 1, and tier 2.

Figure 5.11 shows correlator template selection with ‘p’, ‘n’ assignments and eventual binary assignments for the selection. It is observed that correlator template ‘CW1’ gives an erroneous classification with the signal S<sub>M</sub> that is of the same class. These situations

are corrected by the selection of a correlator template that gives more classification hits than the template in question. When the training databases are large, probability assignments for each correlator template are taken with a combination of correct classifications along with the scaling of a reference signal.

Classes for tier-0 are stated below with major classes listed as A, B, and Z:

Class **A** = {Class a-1, Class a-2,..., Class a-20}

Class **B** = {Class b-1, Class b-2,..., Class b-25}

Class **Z** = {Class z-1, Class z-2,..., Class z-5}

The initial class tree contains tier-0 as the initial template in the tree structure. Tier-0 templates are a super set of tier-1, tier-2 ... classes. Therefore, when selecting tier-0 templates it is necessary to obtain all necessary information of sub-classes from the known training set.

Tier-1 classes stated below follow the tier-0 class assignments shown above:

Class **a-1** = {Class a-1\_1, Class a-1\_2,..., Class a-1\_20}

Class **a-2** = {Class a-2\_1, Class a-2\_2,..., Class a-2\_25}

Class **a-20** = {Class a-20\_1, Class a-20\_2,..., Class a-20\_5}

The above tier-1 class assignment contains the superset of identification formed in tier-2. Correlator templates form the superset to classify signals according to sub-classes, and

form the superset that includes all signal identification to be performed in tier-2. Therefore it is necessary to include all signals encompassed in each sub-class while choosing correlator templates from a previously known database.

### 5.5.2 Identification

Identification is performed in tier-2 of the classification / identification tree. Here templates are picked precisely for the identification of sources. Identification is done on source basis without consideration for situational variables (i.e. speed of moving vehicles / direction of movements etc.). SSE identification consists of the following as performed on each state space and type separately. The following definitions show identification based on vehicle types.

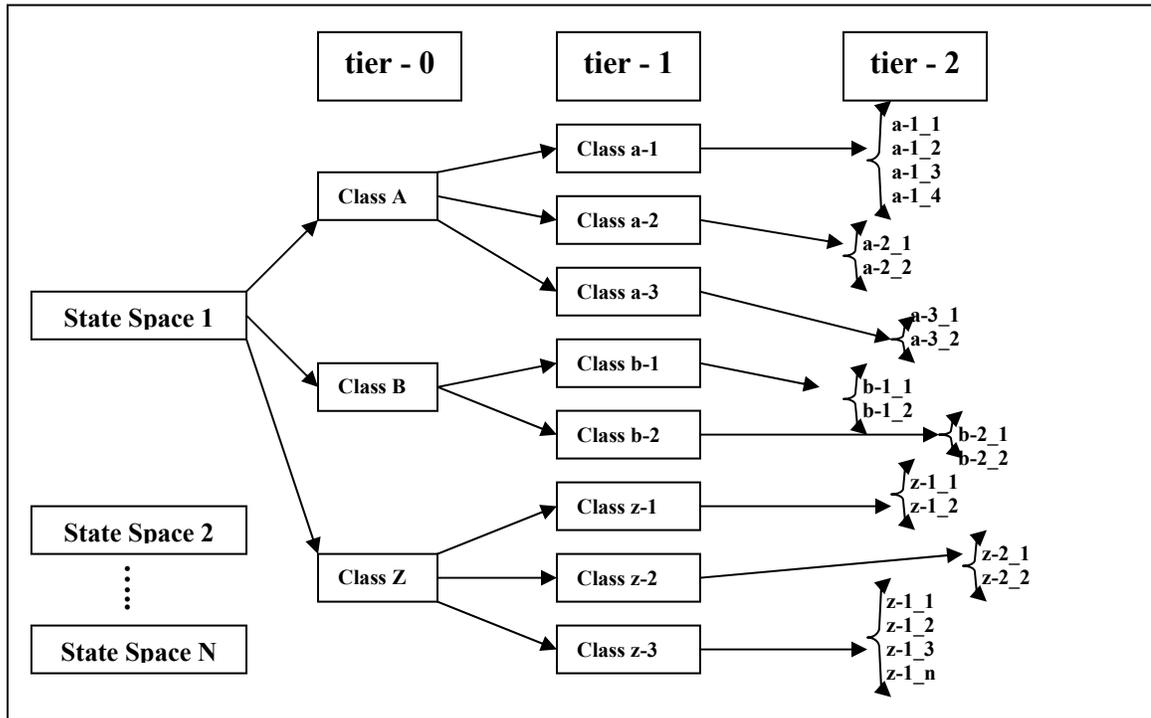
$$ID\ tv1 = \{ tv1-10; tv1-15; tv1-40, \dots \}$$

$$ID\ tv5 = \{ tv5-15; tv5-18; tv5-20, \dots \}$$

$$ID\ wv1 = \{ wv7-16; wv7-18; wv7-40, \dots \}$$

The identification set contains the letters “t”, “w” which represent tracked and wheeled vehicles, respectively. Different identifications for tracked vehicles 1, 5, and wheeled vehicle 1 are shown above. Each vehicle is shown with its associated speed in the example above. Associating precise identification with speeds was not performed with the SSE. Further decomposing the tree according to gearshift information provided by

the time-frequency information enables more precise identification with vehicle types along with speed groupings. (i.e. speed class 1 (1-10 MPH), speed class 2 (11-20 MPH) etc...). However, precision identification is not studied in this dissertation and is left as an issue that requires future investigation.



**Figure 5. 12. Tier structure for class based classification and identification.**

Figure 5.12 shows a tier structure with three hierarchies in depth. Each defined state-space has a parallel structure independent of the others. Tier-0 shows class divisions and tier-1 shows class sub-divisions. Identification is performed in tier-2, with each element representing an identification of sources. Classification and identification run results on real world signals gave excellent results in correctly classifying and identifying signals

from separate state spaces and sensor types. The methods mentioned above were implemented, and results obtained from real world signal runs are shown in Chapter 6 of this dissertation.

## **5.6 Training the SSE with Template Selection**

### **5.6.1 Window Size Selection**

Correlator templates were picked methodically with varying window sizes. An automated correlator selection algorithm was used to obtain this variable. Signals from the training set were selected and auto-correlated with window sizes varying from  $2^5$  samples to  $2^{10}$  samples for a sampling rate of 1 KHz and 512 Hz. Selected template windows are auto-correlated with the same signal segment to find the optimal correlator template. A stepping of one sample is used for correlator template selection to have a full set without losing feature sets. Each selected correlator template is parsed with the same training signal that the correlator template was picked. Maximum, and average RMS values of the parsing are stored to pick the best three templates for that particular template. This procedure is performed for each chosen window length.

```

Choose signal to pick correlator template from a training set
Choose a window length {is looped for different window lengths}
{
for{ 1 < stepping size < (training signal length - (window length)) }
{
    perform correlation on the signal and include averaged /
    maximum absolute RMS (AARMS / MARMS) for each stepping
    window in a matrix(similar to SCIA);
}
selected window length for template = highest average {AARMS /
MARMS} for each window length
}

```

Results obtained as AARMS / MARMS values from the SCIA algorithm are then grouped together to find a suitable window length for the particular state-space. These window lengths are grouped together to derive a common window length, with each state-space having its unique window length. Variable window lengths within a state-space are possible under critical circumstances with scaling used during confidence measure calculations. However, for simplicity without the loss in generality, template window lengths are kept constant throughout different state-spaces.

## 5.6.2 Minimum Signal Length

Signals to be classified / identified by the SSE are required to have minimum length for low power operation. Therefore, there is a need to have minimum number of samples in a signal for optimal performance of the SSE. This is performed by taking a reference AARMS / MARMS value for the chosen template – signal combination during correlator template picks. This reference value is used to compare scaled AARMS / MARMS values obtained from different signal lengths used with the picked correlator template. This process is followed for multiple signals from the training set individually. Once a common signal sample size is chosen, this chosen value becomes the signal length input to the SSE for template tree building as well as for reference AARMS / MARMS value selections.

**Choice of Correlator Templates:** The proper choice of correlator templates is essential for optimal detection and for the building of an efficient tree structure called “Type Abstraction Hierarchy” (TAH). Such a selection results in efficient use and parsing of the tree with minimal hierarchical levels along with the highest confidence in classified / identified results. Proper selection of the ‘correlator’ template was investigated with a GUI and then used in an automated correlator selection implementation.

### 5.6.3 Time Domain RMS Algorithm

The time domain RMS algorithm is separated into two different modules and implemented depending on applications. Choices of these modules depend on signal features as well as algorithm requirements. The selected methods are:

- **AARMS** (Average Absolute Root Mean Squared): This correlation, takes the root mean squared of the absolute correlation values. This method takes into account all absolute RMS values without discriminating between high or low RMS values for each stepping.
- **MARMS** (Maximum Absolute Root Mean Squared): This correlation takes the maximum, or averages the maximum three values of the correlation result on the signal to be classified / identified.

Both the above methods gave excellent results and did not differ in result. However, their confidences were different given the decision-making probabilities associated with each leaf. Both these methods are similar in their signal conditioning / pre-processing. Once these signals are input into the SCIA, the algorithm goes through steps of signal mean centering, normalization, absolute root mean cross correlation, window stepping, confidence scaling, and parsing the tree structure associated with classification / identification. These steps are performed in modules that have more control on the SSE algorithm.

**Signal Normalization:** Each windowed signal is initially centered (after obtaining its mean), and then normalized relative to its maximum value. The centering operation, and normalization of each value is needed for having a uniform operation and to obtain a normalized RMS value for scaling and association of probabilities. Without this internal operation on each windowed signal, results obtained during SCIA classification / identification will have non-coherent AARMS / MARMS values. This in turn will have probabilities that are skewed with a loss in uniformity.

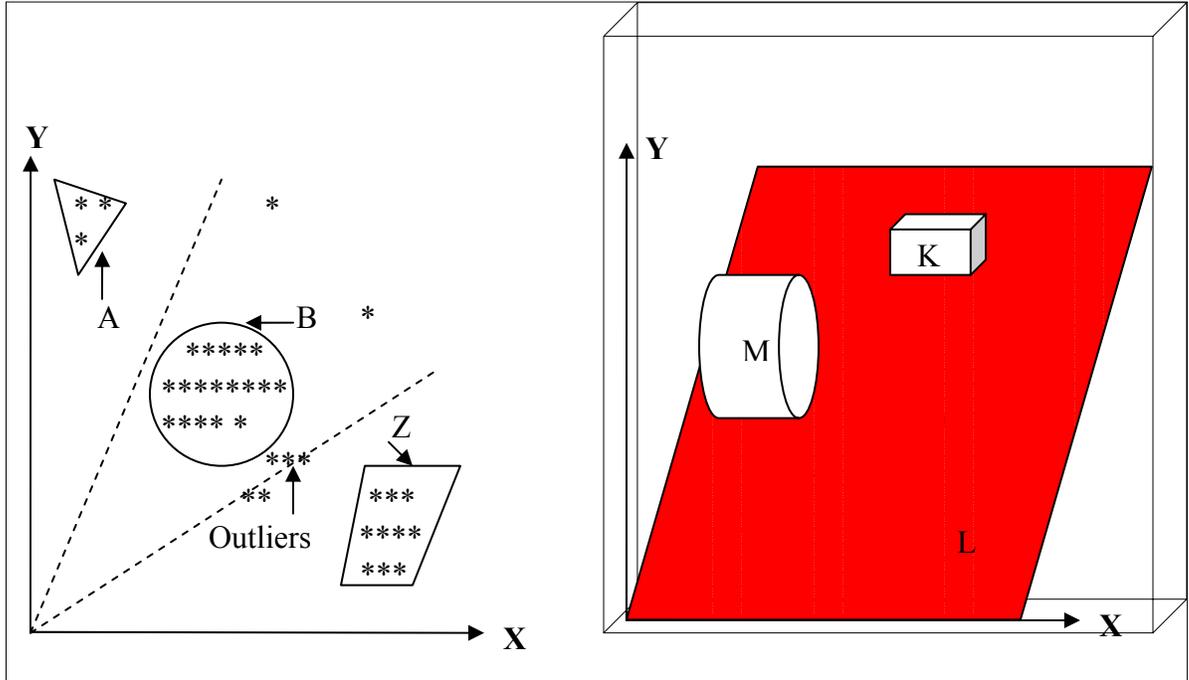
**Length of Signal:** As described in section 5.6.2 above, signals input to the SCIA follow minimum length requirements. However, when signals exceed the minimum requirements, the classification / identification results were similar to those with minimum signal length for each state-space with the loss of more power due to more processing required by the signals. When signals are divided into state spaces, and state types as in the case of classifying and identifying moving sources, each signal coming from each of these different state spaces and state types will follow this requirement.

#### **5.6.4 Probability of Detection or False Alarm**

Statistical weighting was used extensively for decision-making based on the decision boundaries. [32] Probabilities were assigned according to clusters. Training set signals with known parameters were used for obtaining identification / decision boundaries based on results from the SCIA. Decision clustering was multi-dimensional based on the

number of state-spaces and state-types. [88] Figure 5.13 presents decision clustering as shown for 2D and 3D graphs below cluster shapes.

Multi-dimensional state spaces result in multi-dimensional decision boundaries. It is evident from Figure 5.13 that when there is a 2D state space clustering of decision boundaries are 2 dimensional. This results in classification decision boundaries to be in class sub-spaces that have two dimensional probability statistics. Here, class A, B, and Z are shown with cluster boundaries with outliers present close to decision boundaries. The 3D diagram shows a three dimensional state space / sub-space as is evident in local clusters being formed for classes M, K, and L. The cluster boundaries are planes or three-dimensional spaces. For multi-dimensional state spaces ( $> 3$ ), the same decision boundaries are formed mathematically with the distance from the boundaries showing classification confidence scaling. [18, 35, 49, 84, 99]



**Figure 5. 13. 2D and 3D decision boundaries formed during multi-dimensional state-space classification and identification.**

Decisions are made according to the following criteria. For the case of 2 dimensional decisions, a signal  $S_{ij}$  is a member of class  $k_1$  if,

$$S_{ij} \in C_{k_1} : \frac{1}{\sqrt{2\pi}\sigma_{k_1}} \exp \left[ -\frac{1}{2} \left( \frac{X_{ij} - \mu_{k_1}}{\sigma_{k_1}} \right)^2 \right] > \frac{1}{\sqrt{2\pi}\sigma_{k_2}} \exp \left[ -\frac{1}{2} \left( \frac{X_{ij} - \mu_{k_2}}{\sigma_{k_2}} \right)^2 \right] \quad (5.11)$$

where,  $\{\mu_{k_1}, \sigma_{k_1}\}$  and  $\{\mu_{k_2}, \sigma_{k_2}\}$  are the mean and variance of predefined classes  $k_1$  and  $k_2$  respectively.

For multi-dimensional (N dimensional) decision boundaries the following criteria is used in associating a signal to a class. A signal  $S_{ij}$  is a member of class  $k_1$  iff,

$$S_{ij} \in C_{k_1}: \frac{1}{\sqrt{2\pi}\sigma_{k_1}} \exp\left[-\frac{1}{2}\left(\frac{x_{ij} - \mu_{k_1}}{\sigma_{k_1}}\right)^2\right] > \left\{ \frac{1}{\sqrt{2\pi}\sigma_{k_m}} \exp\left[-\frac{1}{2}\left(\frac{x_{ij} - \mu_{k_m}}{\sigma_{k_m}}\right)^2\right] \right\}$$

for all  $m = 2, 3, 4, \dots, N$  &  $m \neq 1$

(5. 12)

where,  $\{\mu_{k_1}, \sigma_{k_1}\}$  and  $\{\mu_{k_m}, \sigma_{k_m}\}$  are the mean and variance of predefined classes  $k_1$  and  $k_m$  respectively, where  $m = 2, 3, 4, \dots, N$  different classes. Confidence measure of each class is calculated relative to the variance of the signal to that particular class.

### Re-Configuration of Template Tree

Initial implementation of the template tree is from an available known signal set contained in a database. When more signals are acquired and stored in the signal database, these signals will constitute a broader class of signals for correlator template selection and tree assignment. Probabilities obtained during training from a larger database gave more credibility to confidence because a larger number of signals are used for probability assignment. The obtained probabilities are assigned to each leaf in a tree. Once this new tree and probabilities are obtained using off-node processing, this tree structure along with its probabilities is updated on the SSE architecture, either remotely or manually (if accessible). In many instances it is easier to deploy a new node with the

latest correlator template tree (along with its probabilities) than to reconfigure already deployed sensor nodes. This decision is based on the sensor network application and deployed topology given its cost and the number of sensor nodes in deployment.

### **5.6.5 State Space Approach**

One of the fundamental concepts of the SSE is the state-space approach to fragment signals into different state spaces. Separating signals into state-spaces reduces the number of variables as mentioned above and gives a methodical structure to classifying or identifying signals in the time domain. The SSE did not perform as expected without the state-space approach. However with the implementation of the state-space approach, result accuracy was enhanced tremendously. Moving vehicle signals were initially studied with time-frequency contents to investigate signal characteristics of wideband signals as shown in section 5.2.1-5.2.5 of this chapter. It was observed that moving source signals did have time dependent frequency characteristics and designing with a state-space approach mitigated or diminished these effects while enhancing the performance of the time-domain SSE.

The approach is to separate signals into state spaces, and sub-spaces. Subspaces will experience further separation of signals based on variables or parameters as observed in sensed signals, and therefore will consist of further division as observed in these sub-spaces. Figure 5.14 shows a diagram for acquired signal decomposition. Here, signals

are separated into state types, each of which has its own state-space and sub-space. State types depend on sensor types and are assigned labels of “acoustic”, “seismic”, and “infrared” signals based on the sensed signal types. These signal types are then decomposed into state spaces dependent on the distance and locality of moving sources relative to the sensors. [24] Each of these state spaces are further decomposed into state sub-spaces during feature extraction. Figure 5.14 shows approach, arrival, and departure as state spaces, and has feature 1, and feature 2 as state sub-spaces.

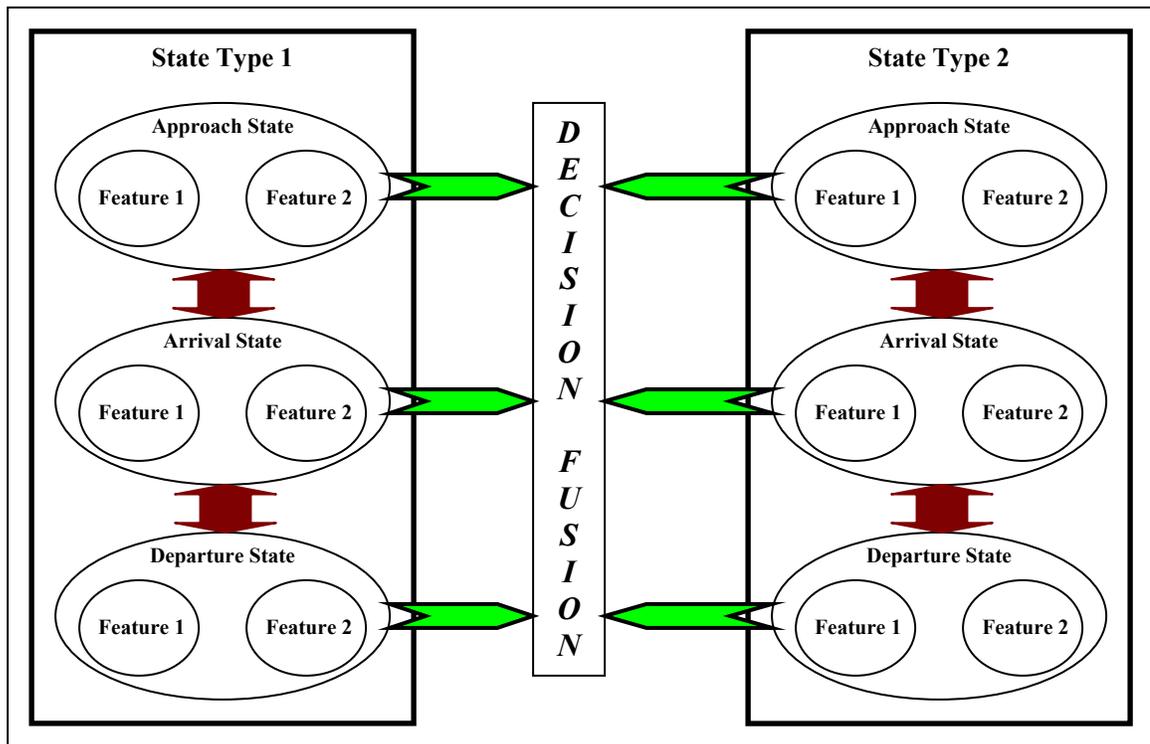


Figure 5. 14. State Type, Space, Sub-Space based decomposition of acquired signals.

### 5.6.6 Detection in M-SSE

The multi-signal search engine (M-SSE) concept originated when multiple types of signals were gathered with multiple sensor types for signal source classification / identification. These signals were processed in parallel and were combined after statistical evaluation (associated probabilities were obtained from a training set). Weighting and confidence measures were assigned independently, relative to each type of signal and were used to classify and identify based on classification / identification systems specified in section 3.3. The combining of signal results was done according to methods mentioned in section 3.5. A Multi - Type Abstraction Hierarchy (M-TAH) was used to obtain classification / identification when multiple signals were present. M-TAHs were used with two different methods: once during classification / identification, and the other during confidence measure calculations or weighting.

**Low Power:** A time domain signal classification / identification algorithm takes into effect a novel methodology that is lower in power in comparison to present day algorithms used for the classification / identification of signals. It is more efficient in computational power while it attains higher percentage of accuracy.

**Accuracy:** The accuracy of results from the time domain SSE has been quite outstanding as seen in training signals and test signal sets. Furthermore, classification / identification with real world signals have shown error rates of  $< 10\%$  for the SSE. Data of

comprehensive analysis and results yield even better error rates (as shown in the data run results in the following chapter).

## **5.7 Summary**

The state-space approach to signal classification / identification in the time domain SSE was presented. The importance and need for this approach is explained, and the signal parameters that were acted upon for structural decomposition explained. One of the dominant moving source parameters, the ‘Doppler’, is detailed with graphs from real world signal sets. Spectrograms were used during training to determine and investigate real signal characteristics. However, SSE modules do not contain spectrogram calculations / plots, and instead use time domain frequency analysis.

The time domain SCIA algorithm is presented along with methods for selecting correlator templates and building a tree. Choosing a proper correlator tree template is vital for event identification and classification. This was detailed with an explanation of optimization of window size for building a correlator template ‘tier’ structure.

## CHAPTER 6

### SSE / M-SSE Results on Real World Signals

#### 6.1 Background

All SSE / M-SSE modules were used to obtain behavior of the implementation to real world signals gathered in different environmental and situational conditions. Signal database included raw sensor readings from an array of acoustic microphones and seismic geophones. Correlator templates were selected for classification and identification separately. [55] Signals were decomposed into state-types and state-spaces during signal pre-processing / conditioning operation. These state spaces were input into the M-SSE for signal classification / identification. Obtained results were then fused for decision-making. The obtained decisions gave excellent results to moving source real- world sensor readings. It is observed that relative to other forms of classification M-SSE is more robust to classification / identification with an ‘open’ signal set, with comparable accuracy rates, and additional features. Benchmarking SSE results to that obtained

during wavelet method points to comparable accuracy, with the SSE having an extra feature in confidence measure calculation that detects a new signal previously unknown and not present in the training database. A comparison of computational cost for the wavelet method and the SSE is given later in this chapter.

## 6.2 Real World Signal Database

One of the signal databases used for real world signal testing was provided by the Army Research Laboratories (ARL) and consisted of a systematically acquired data set. [92] The data designated as the Acoustic-seismic Classification Identification Data Set (ACIDS) was used as real world signals. The data was collected by two co-located acoustic/seismic sensor systems designated as Sensor System 1 and Sensor System 2. There were over 270 data runs (single target only) from nine different types of ground vehicles in three different environmental conditions. The ground vehicles were traveling at constant speeds from one direction toward the sensor systems passing the closest point of approach (CPA) and then away from the sensor systems.

The ground vehicles repeated the same test runs from the opposite direction. The CPA to the sensor systems varied from 25 m to 100 m. The speed varied from 5 km/h to 40 km/hr depending upon the particular run, the vehicle and the environmental conditions. The acoustic data is from a 3-element equilateral triangular microphone array with an

equilateral length of 15 inches. In the x-y plane with the positive x-axis as the reference and the center of the array as the origin, Mic 1 is at 210 degrees, Mic 2 is at 330 degrees and Mic 3 is at 90 degrees. For all of the test runs, microphone 1 was positioned perpendicular to the path of travel of the vehicle as shown on Figure 6.1. There is a seismic sensor located at the center of the array. The data included raw digitized acoustic signatures from three piezo-ceramic microphones. Seismic data

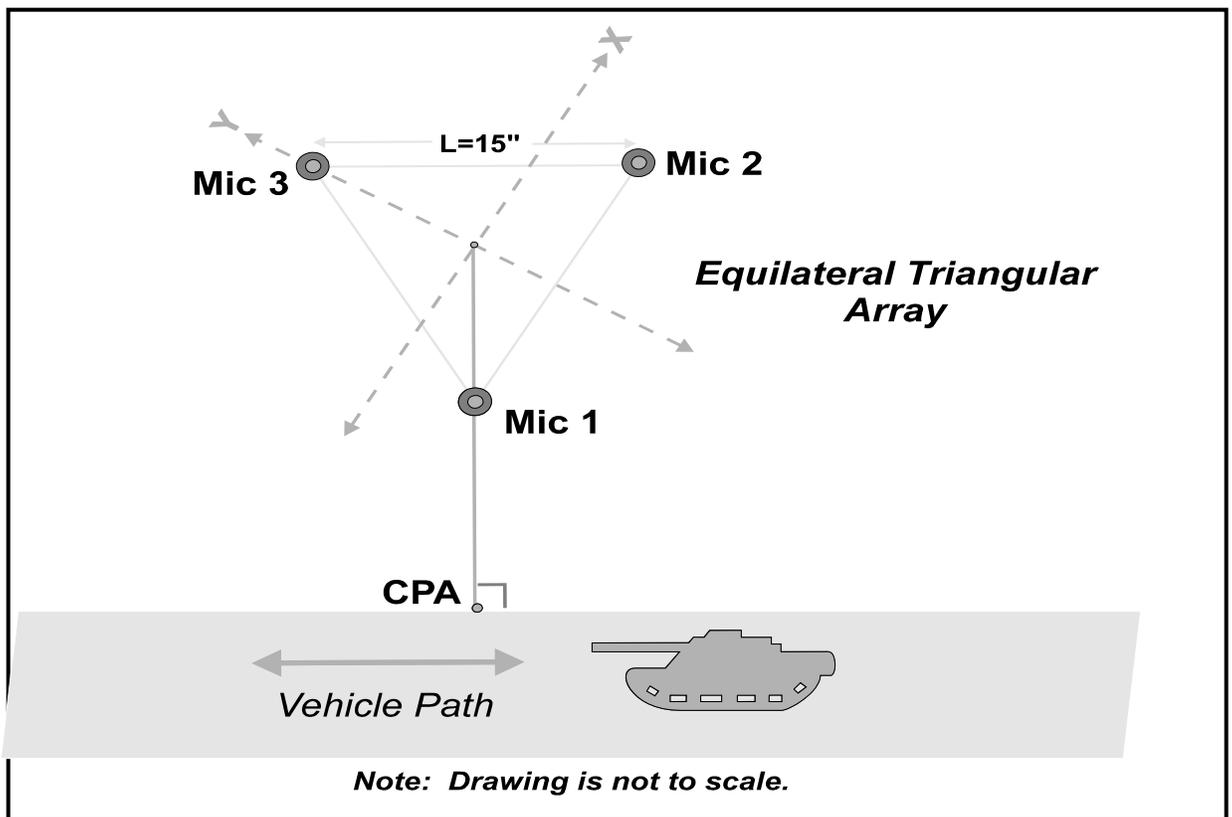


Figure 6. 1. Database containing real world sensor readings provided by ARL.

from the same runs were obtained from a co-located geophone sensor. The microphone acoustic data is low-pass filtered at 400 Hz via a 6<sup>th</sup> order filter to prevent spectral

aliasing and high-pass filtered at 25 Hz via a 1<sup>st</sup> order filter to reduce wind noise. The data is digitized by a 16-bit A/D at the rate of 1025.641 Hz. [92]

Seismic data is from a single vertical-axis geophone that is mounted on the bottom of the sensor system with the weight of the sensor system serving to provide coupling to the ground. The geophone is located at the center of the microphone array as shown on Figure 6.1. The data contained seismic signatures from the low gain channel (36 dB) and high-gain channel (62 dB) of the single axis geophone. The two channels accommodate the large dynamic range of the seismic signals without the need for an automatic gain control circuitry. The geophone data is low-pass filtered at 200 Hz via a 4<sup>th</sup> order Chebyshev filter to prevent spectral aliasing. The seismic data is digitized at 512 Hz. [92]

The dataset contained different classes of vehicles namely ‘heavy track vehicle’, ‘heavy wheel vehicle’, ‘light track vehicle’, ‘light wheel vehicle’. Data gathered came from the following environmental conditions:

- Site Desert: Desert like similar to conditions in the desert southwest in the summer (typically dry and very hot).
- Site Arctic I: Arctic like similar to conditions in Alaska in the winter (typically cold, windy, with and without snow cover).

- Site Normal: Normal, similar to conditions in Mid-Atlantic states in the spring and fall (typically mild, with and without humidity).
- Site Arctic II: Arctic like similar to conditions in Alaska in the winter but different from Arctic I. (typically cold, windy, with and without snow cover).

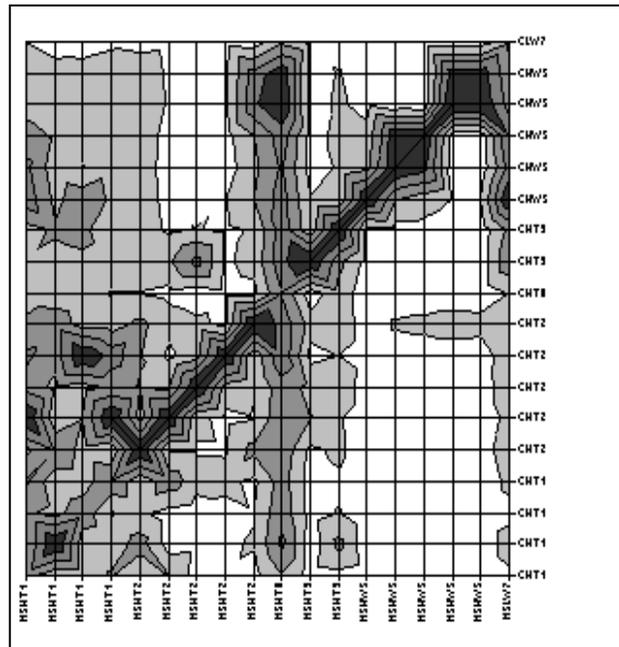
SSE evaluation was performed initially by correlator template selection and eventual classification / identification of signals.

### **6.3 SSE Evaluation**

The SSE is evaluated with the above database for a selection of tracked and wheeled vehicles with the selected signals used to "train" and observe SSE performance. [11] Signal content used for training the SSE was removed from the signal sets used for evaluation. Then, a matrix as shown in section 5.5.1 of chapter 5, is formed with the matrix element values corresponding to the AARMS / MARMS values of the template - signal inner products. The SSE matrix may now be examined to determine and predict SSE performance.

SSE operation on signals during the vehicle approach state-space is depicted here with signals and templates in the horizontal and vertical axis respectively. Windowed

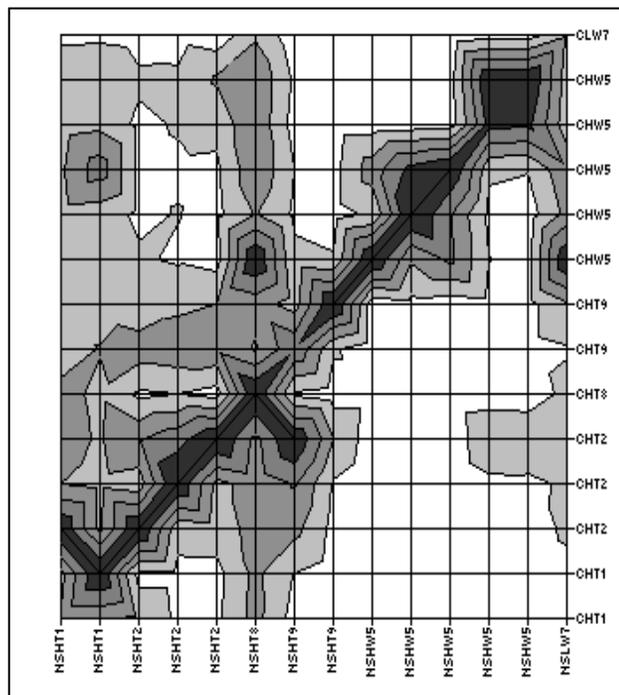
AARMS correlation values are plotted in a surface with darker grey scales corresponding to higher values. [65] This case shows near ideal performance with a SSE AARMS template-signal matrix that is nearly diagonal. Only one vehicle signal shows an error. The error rate obtained is a low 5.6%.



**Figure 6. 2. SSE acoustic signal classification results are shown for approach state in desert terrain.**

Figures 6.2 – 6.5 show examples of these results. In these Figures, the AARMS / MARMS amplitude of the template-signal inner products are shown for inner products between all signals and correlators. The set of these inner product RMS amplitudes forms a matrix with the matrix element scalar value being the AARMS / MARMS signal amplitudes.

Ideal SSE operation occurs where the RMS value of the inner product between a signal and a template are maximal for the case where the correlator correctly identifies the signal. [17, 75] Thus, ideal operation yields a matrix where the diagonal elements are greater than any off-diagonal element in the same row or column as any diagonal element. This would correspond to identification template selection, as opposed to classification template selection where off-diagonal elements within the same class will give high AARMS/MARMS values.



**Figure 6. 3. SSE acoustic signal classification results are shown for the departure state in desert terrain.**

Figure 6.3 shows a template-signal matrix with an acoustic signal for the departure state in the desert terrain. The SSE operation for signals obtained during the departure state is





Figure 6.5 shows SSE evaluation for acoustic signal identification for tracked and wheeled vehicles. This data corresponds to normal terrain and the vehicle approach state. The SSE template-signal inner products form a matrix. Here, the matrix element values, the MARMS value of the inner product, are plotted for each template-signal pair. The template and signals for each vehicle record are ordered along the matrix rows and columns. Since one off-diagonal matrix element is greater than a diagonal matrix element in its row, this result displays one error for a 5.6 percent error rate.

Ideal SSE identification occurs where the AARMS / MARMS value of the inner product between a signal and template are maximal for the case where the template correctly identifies the signal. Thus, ideal operation yields a matrix where the diagonal elements are greater than any off-diagonal element in the same row or column as a diagonal element.

The SSE has demonstrated the ability to identify the differences between signals generated by a vehicle for the states when a vehicle passing the sensor is:

- 1) approaching the sensor
- 2) at the point of closest approach (CPA)
- 3) departing from the sensor

In the event that proper templates are available in the SSE library, this sensitivity to the evolution of a signal is exploited further for enhancing identification accuracy.

The results of Figures 6.2 – 6.5 are shown as an initial step to prove excellent identification capability of the SSE algorithm in two different terrain conditions with the approach and departure states of vehicle travel.

## **6.4 SSE / M-SSE Implementation**

Having obtained extremely good results for the time-domain SCIA algorithm we proceeded to test the ability of the algorithm when implemented as a search engine. The SSE designed with a three-tier structure [44] whereby the signals are classified in the first two tiers and are identified in the third tier. Figure 6.6 shows system design diagram for the SSE hierarchical tree structure with branches converging for identification. System design diagram for SSE using ‘hierarchical search’, clusters signals into the following tier structure.

Step 0 : Corresponds to classification as tracked vs. wheeled.

Step 1 : Corresponds to classification as heavy vs. light.

Step 2 : Corresponds to identification of vehicles by distinct vehicle types.

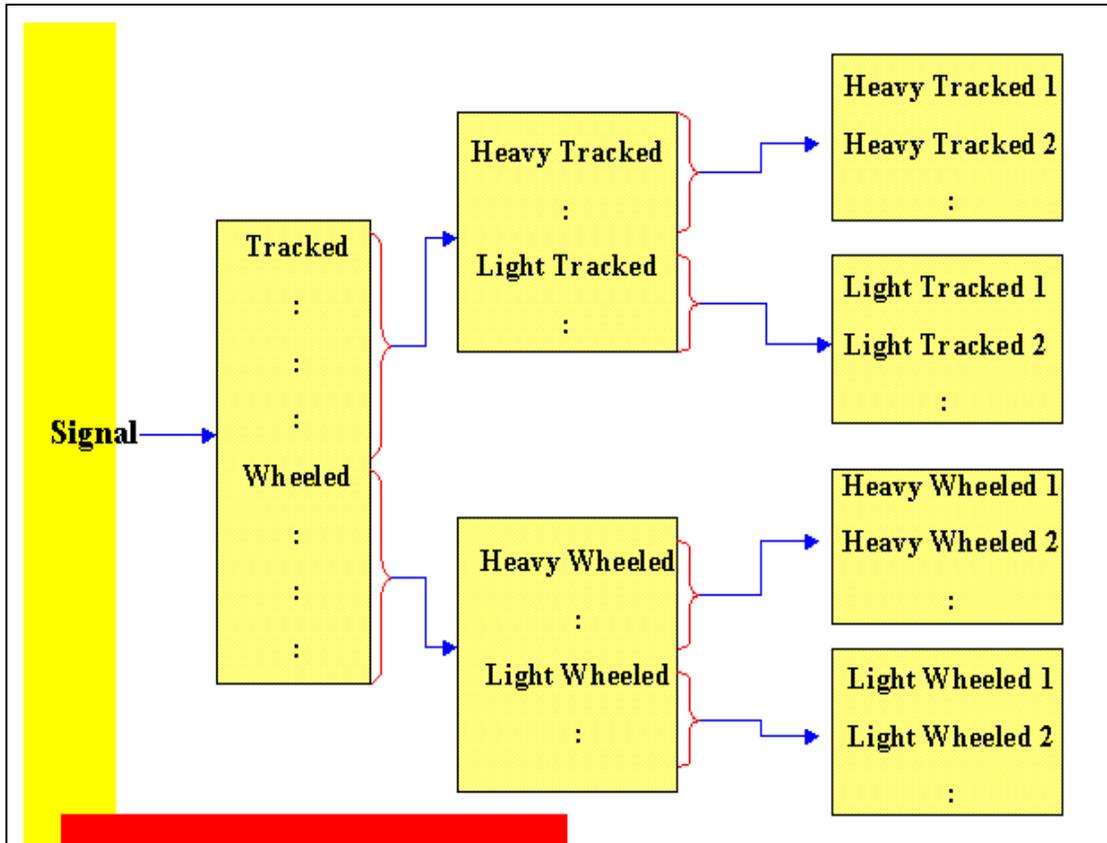
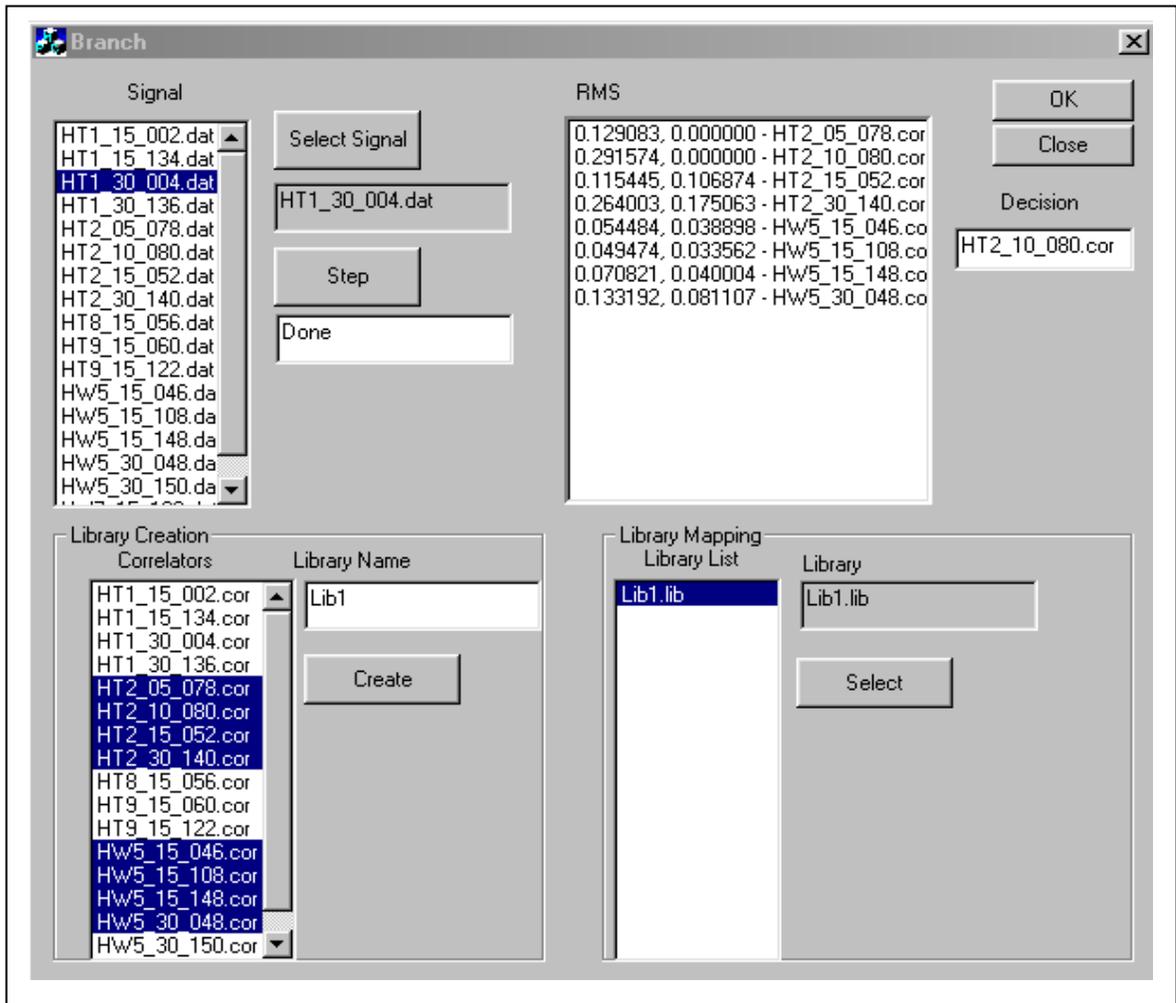


Figure 6. 6. Tier structure of class divisions and branching based on signal database.

The above structure for first tier classification is to group the vehicles into a meta-class as Tracked vs. Wheeled. Second tier classification is to group them into a sub-class as Heavy vs. Light. Results obtained from the groupings are then input into the third tier to identify vehicles with distinct vehicle types. Each of these blocks contains templates in a library, which give the best set for classification / identification for each tier. The following figures give test results and an analysis of the SSE / M-SSE for acoustic and seismic signal pairs.



**Figure 6. 7. Hierarchical Search (Type Abstraction Hierarchy) – Step 0.**

Figure 6.7 above shows a tier 1 classification results. The database of a subset of all signals present in the ‘ACIDS’ database is shown in the ‘Signal’ menu. One signal of this displayed subset can be chosen from the ‘Signal’ menu to represent a signal to be classified. Here we have chosen a subset of acoustic signals. Signal templates or ‘correlators’ are grouped into libraries for each meta-class. These previously grouped

templates perform the time-domain SCIA algorithm calculation and display the values for each segment in the RMS display menu. The menu lists calculated values for each

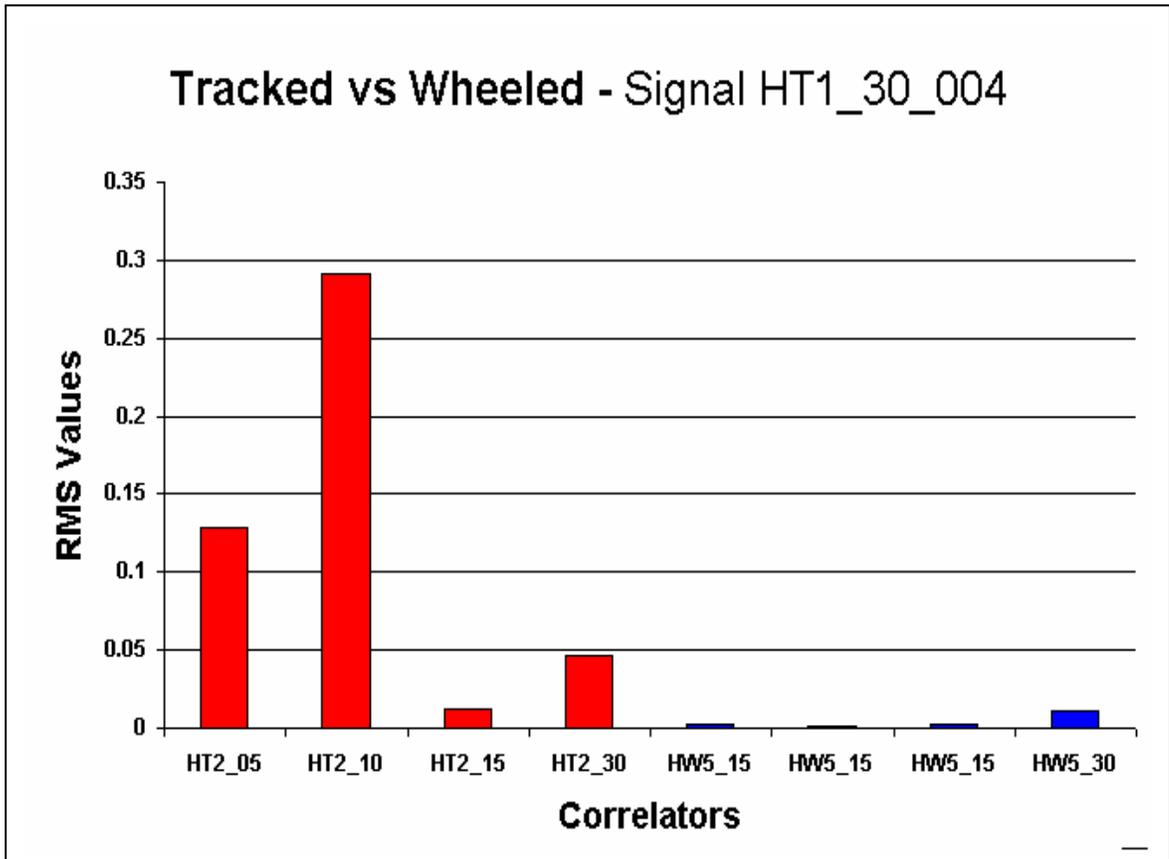


Figure 6. 8. SSE results after collaborative decision fusion.

segment (approach, arrival, and departure) and its corresponding template name. Statistical fusion of the results yields a combined weighted value of the initial results for each template as discussed in section 3.5 of Chapter 3. The highest possible value obtained from the decision-making fusion formulae gives the chosen class for the first tier.

Above classification chooses eight library templates, four of which corresponds to tracked vehicles and the rest to wheeled vehicles. A signal 'HT1\_30\_004' is chosen as the signal to be classified. (This signal corresponds to a tracked vehicle as seen by the second letter of the name 'T', rather than 'W', which corresponds to wheeled vehicles.) Note that the signal to be classified does not have any templates picked from that particular vehicle run, or for that matter from that vehicle make/model. Fusion of initial AARMS yields a correct classification of 'Tracked' as given in the 'Decision' displaying HT2\_10\_080 (Tracked – 'T' is the second letter in the decision name). Here we have shown a successful implementation of a tracked vs. wheeled TAH (Type Abstraction Hierarchy) classification

The decision rule after weighted fusion calculation performs a maximum-polling decision-making selection. Therefore, the decisions would have more confidence when the values of the decision class have a bigger margin than the incorrect decision class that is directly proportional to associated confidence measure. Here we see that the correct decision class templates that are 'tracked' (left; columns 1-4) give much larger magnitudes than those of the incorrect decision class of 'wheeled' (right; columns 5-8) templates. It is shown on Figure 6.8 with great certainty and confidence that the classified decision is a 'tracked' vehicle in that all tracked vehicle templates give larger decision values than each wheeled vehicle template.

With correct classification shown for tier-1, a sub-class based decision for the tier-2 structure is performed to classify sub-classes in hierarchical tree structure. Figure 6.9 shows classification

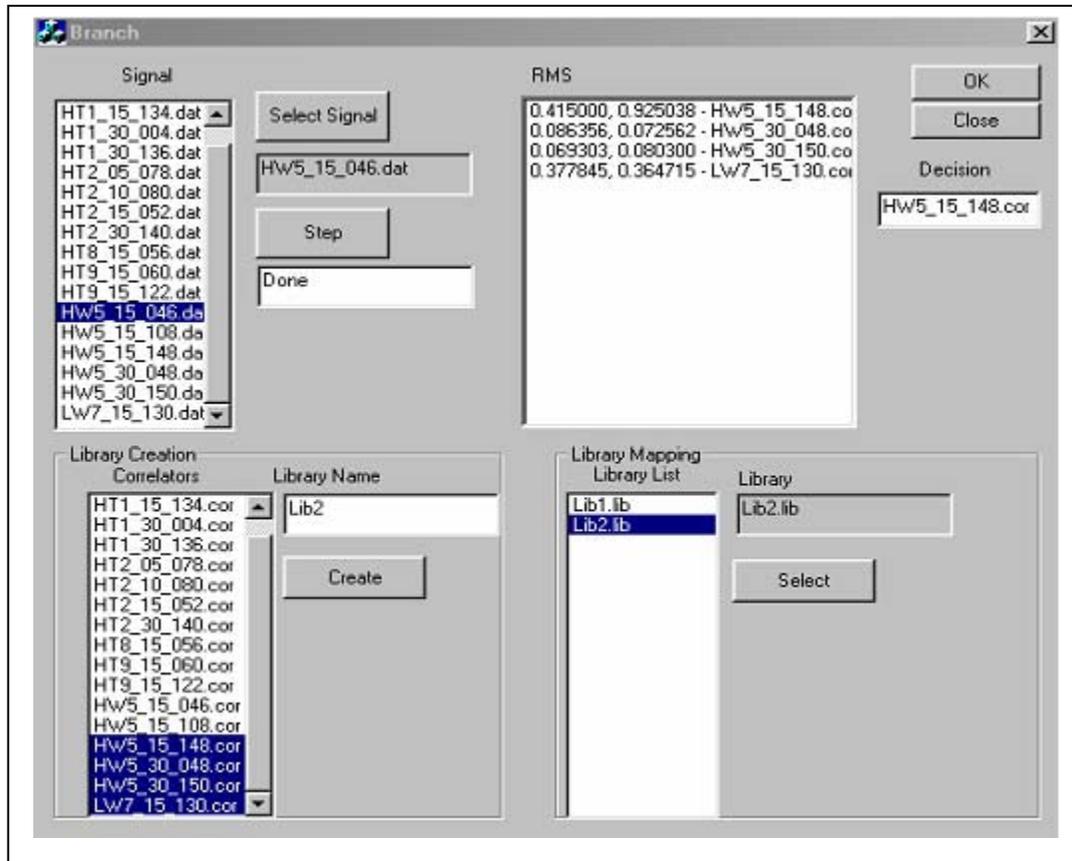


Figure 6.9. SSE runs to classify Heavy vs. Light vehicles.

step 1 of the hierarchical search (TAH). Tier 2 classification results are shown in the above diagram. The database of a subset of all signals present in the 'ACIDS' database is shown in the 'Signal' menu. One signal of this displayed subset is chosen from the 'Signal' menu. Here we have chosen a subset of acoustic signals. Signal templates or 'correlators' are grouped into libraries for each sub-class. These previously grouped templates perform the time-domain SCIA calculation for AARMS and display the values

for each segment in the RMS display menu. The menu lists the calculated values for each segment (approach, arrival, and departure) and its corresponding template name. Statistical decision-making fusion of the result yields a combined weighted value of the initial result for each template. The highest possible AARMS value obtained from the decision - fusion formulae gives the chosen sub-class for the second tier.

Here we choose four library templates one of which corresponds to a light vehicle and the rest to heavy vehicles. A signal 'HW5\_15\_046' is chosen as the signal to be classified into a sub-class. (This signal corresponds to a heavy vehicle as seen by the first letter of the name 'H', rather than 'L', which corresponds to light vehicles.) Note that the signal chosen to classify does not have any templates picked from that particular vehicle run.

Fusion of initial calculation yields a correct classification of 'Heavy' as given in the 'Decision' displaying HW5\_15\_148 (Heavy - 'H' is the first letter in the decision name). Here we have shown a successful implementation of a Heavy vs. Light TAH sub-class classification.

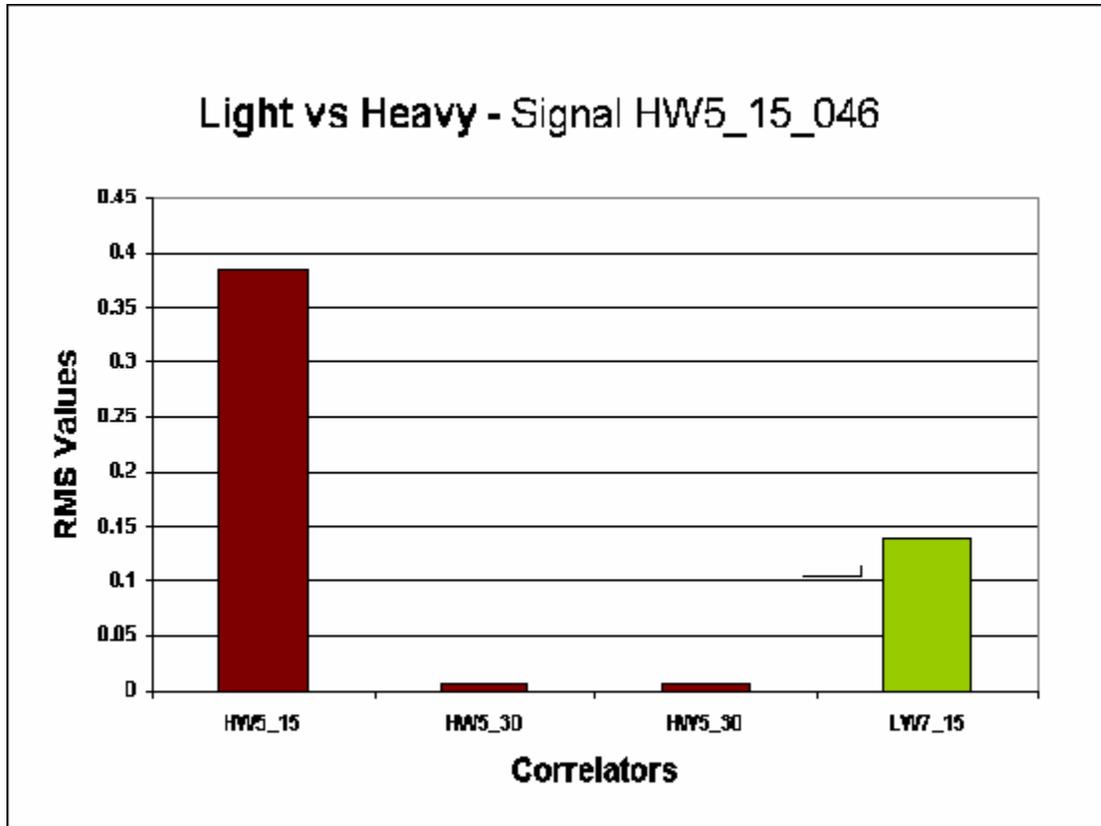


Figure 6.10. Classified results after collaborative decision making from multiple state spaces.

Figure 6.10 shows the decision values for hierarchical search step – 1 of TAH. The decision rule after weighted fusion calculation performs a maximum AARMS value selection. Therefore, the decisions would have more confidence when the values of the decision sub-class have a bigger margin than the incorrect decision class. Here we see that the correct decision class template, which is ‘Heavy’ (left; columns 1-3), gives a bigger magnitude than that of the incorrect decision class of ‘Light’ (right; column 4) class template. However, we see that only one heavy vehicle template outcores the light vehicle template while the others are low in amplitude. Therefore, the confidence of the result is less compared to the result obtained in Figure 6.8.

Tier 3 (TAH) identification is now taken up in step-2. Here identification is performed for correct identification of the vehicle type. The identified vehicle obtains MARMS values during SCIA processing. The MARMS value is then fused for decision-making and scaling for confidence measure calculations. Figure 6.10 shows SSE's hierarchical search step-2 of the Type Identification Hierarchy (TAH). The database of a subset of all signals present in the 'ACIDS' database is shown in the 'Signal' menu. One signal of this displayed subset is chosen from the 'Signal' menu to be identified by SSE. Here we

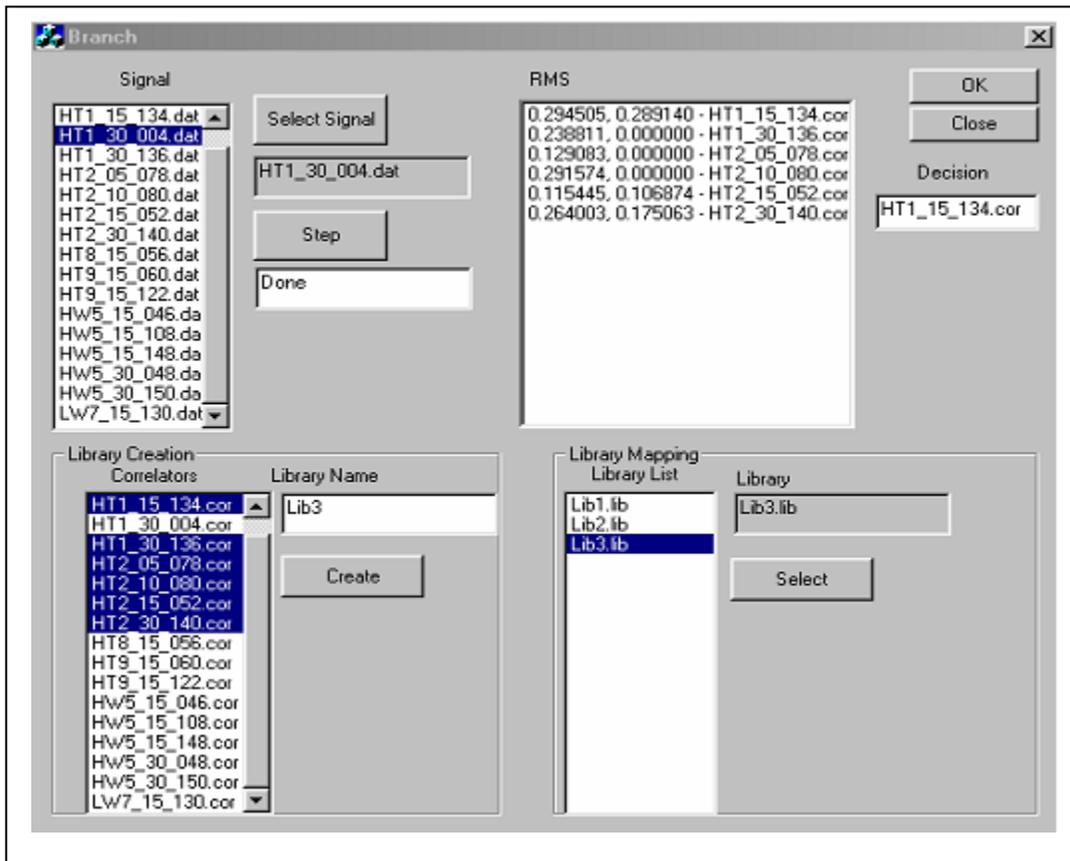


Figure 6. 11. SSE identification by vehicle types.

have chosen a subset of acoustic signals. Signal templates or ‘correlators’ are grouped into libraries for each identification-class. These previously selected template libraries perform the time-domain MARMS correlation calculation and display the values for each segment in the RMS display menu. This menu lists the calculated values for each state (approach, arrival, and departure) and its corresponding template name. Statistical fusion of the result yields a combined weighted value of the initial results for each template. The highest possible value obtained from the decision-making fusion formulae gives the identification for the third tier.

For identification we choose six library templates two of which corresponds to Heavy Tracked Vehicle#1 (HT1) and the rest to Heavy Tracked Vehicle #2 (HT2). A signal ‘HT1\_30\_004’ is chosen as the signal to be identified. (This signal corresponds to heavy tracked vehicle #1 as seen by the first three letters of the name ‘HT1’, rather than ‘HT2’ which corresponds to heavy tracked vehicle #2.) Note that the signal selected to be classified does not have any templates picked from that particular vehicle run forming a mutually exclusive signal from the template library. SCIA MARMS results decision-fusion of initial calculations yield a correct identification of ‘Heavy Tracked Vehicle#1’ as given in the ‘Decision’ displaying HT1\_15\_134 (Heavy Tracked Vehicle #1– ‘HT1’ are the first three letters in the decision name). Here we have shown a successful implementation of identification.

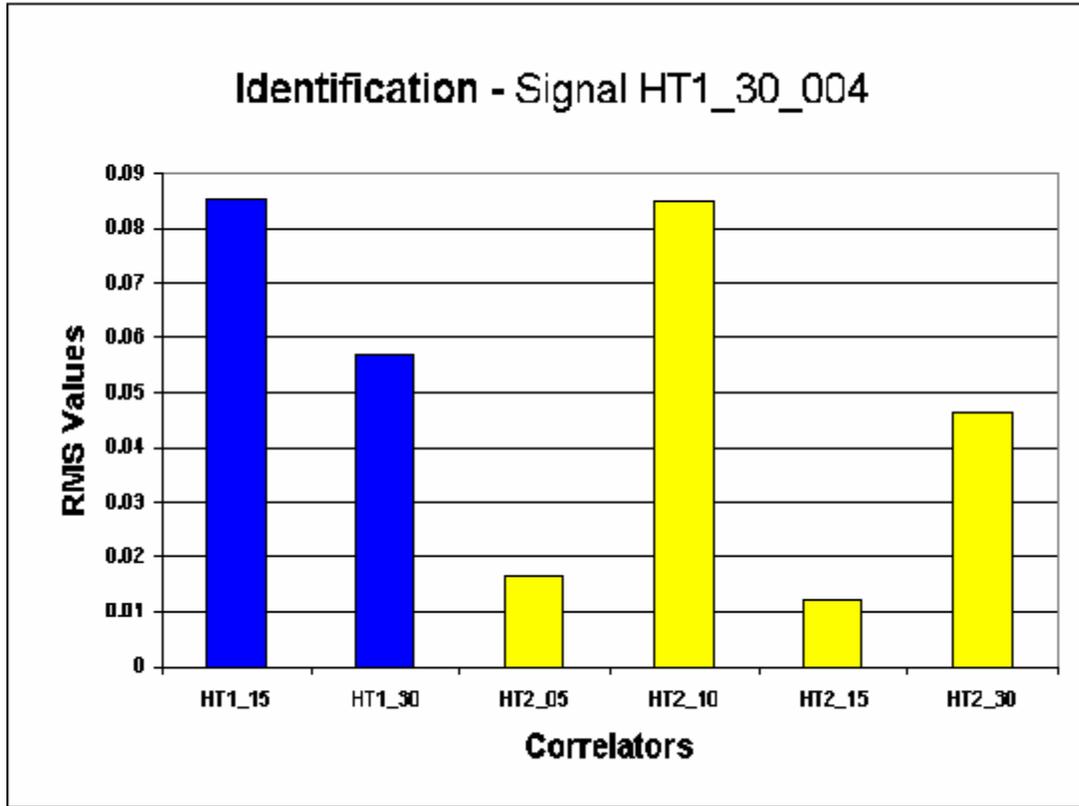


Figure 6.12. SSE identification results after collaborative decision-making.

Figure 6.12 shows decisions for hierarchical search step 2 in the type abstraction hierarchy. The decision-making procedure after weighted fusion calculation selects the maximum value. Therefore, the decisions would have more confidence when the values of the decision identifiers have a bigger margin than the incorrect decision identifiers, with normalized weighting and scaling according to MARMS values. Here we see that the correct decision identifier templates (left; columns 1-2) which is ‘Heavy Tracked Vehicle #1’ gives a very small margin than that of incorrect decision identifier ‘Heavy Tracked Vehicle #2’ (right; columns 3-6) identifier templates. SSE’s SCIA gave excellent identification decisions for identification in the tier 3 (TAH) with results fused

with decision-making criteria as shown in Section 3.5 of Chapter 3. Results gave more than 90% correction identifications in all test cases. Having obtained great results a comparison of computational calculation requirements to a wavelet method is done with the same signal set.

## **6.5 Wavelet Method Classification**

Wavelet methods were used for signal classification with the same data set. Accuracy levels were similar to the SSE with correct classification in the 90% range. It was found that the CSM algorithm used with the above signals gave the above accuracy levels.

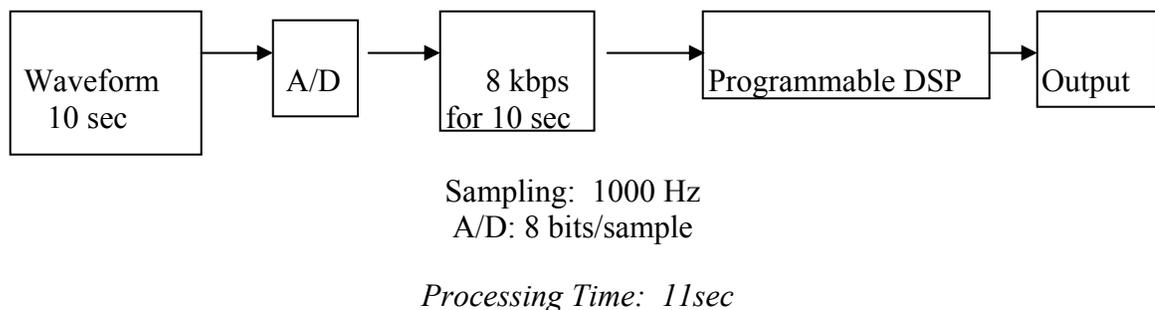
In comparison to SSE, the wavelet method decomposes signals into orthogonal sub-spaces obtained during training. Selected wavelets then perform signal energy decomposition into these orthogonal sub-spaces for signals to be classified. In comparison to the SSE that decomposes signals, the wavelet approach takes the whole signal without segmenting into vehicle location relative to the sensor for signal energy decomposition. [16, 52, 71, 72, 73]

A fundamental difference is that the wavelet signal subspace assignments are based on a closed form, where only signals used during training could be classified correctly. Any new signals being sensed by the sensors would be classified into an existing class of

signal thereby giving an erroneous decision. Further, wavelet methods need to be trained repeatedly when new signals are acquired further increasing the number of orthogonal subspaces. Thus, with increasing classes of signals there is a linear dependence in the number of orthogonal wavelet sub-spaces causing more calculations to be done for each classification. However, accuracy levels without confidence measures may facilitate an implementation with the wavelet method in parallel to the SSE.

## 6.6 Wavelet DSP Hardware Implementation Feasibility

SSE is more suitable for implementation with programmable DSPs. The need for sensor classification / identification to be reconfigurable makes it more suitable for this application to be implemented with DSP hardware. SSE hardware implementation requirements were shown in section 3.4 of Chapter 3. Here a specification and requirements are given at the macro-level for an understanding of implementation requirements for the wavelet method. It is assumed that implementation to be parallel to that of the time domain SSE.



For proper wavelet implementation in parallel with the SSE, the programmable DSP module is to contain two blocks DSP1 (pre-processing / signal conditioning module), and DSP II (wavelet algorithm). The following calculations show sample calculations to be performed on these two blocks and an analysis of the wavelet hardware implementation.

### **DSP I - Pre-processing**

Input : 8Kbps bit stream  
Time : 10 seconds  
Function : Would find Closest Point of Approach (CPA) and extract 1024 sample data.  
Algorithm : Find max(data samples) and extract 512 samples to the left and right of it.

Problems: I. The max(data samples) may not be the CPA. It may happen that the value may be due to :

- a. bumps on the path of travel (seismic)
- b. glitches (acoustic, seismic)

II. 10sec may not be enough to gather CPA data, especially if the vehicle is moving slowly.

Computation : Computationally least intensive.  
Cost : The drawback of having DSPI is that it may have an added cost of a microcontroller.

### **DSP II – Algorithm**

Input : Preprocessed Data (1024samples/node)  
Time : 1seconds  
Function : Would run the Algorithm with the received preprocessed data  
Algorithm : Do Wavelet Subspace decomposition on the preprocessed data of 1024 samples.  
Find the minimum distance (L2 norm) from the decomposed signal to the existing wavelet subspace existent in the database.

Problems : I. Should be remotely programmable/configurable (To update database when new signals arrive).  
II. Needs a cost function or off chip processing to detect new signals.

Computation : data = 1000  
decomposition =  $2 * N$  classes \* 1000  
minimum distance =  $k$  dimensional vector space (subtractions)  
+  $k$  squares + 1 sqrt +  $N$  comparisons  
approximately : 300,000 instructions for  $N = 100$ ;

**Go For >2MIPS microcontroller.**

### **Wavelet DSP applied to Node Networks**

- I. Centralized DSP to Process ALL nodes (All nodes DSPI and Base Station has Microprocessor)
- II. Each node has own DSP but is integrated (All nodes DSPI and 1/10 nodes DSPII)
- III. Each node has independent DSP (All nodes DSP II)

The above study looked at aspects of feasibility of wavelet method in DSP hardware for sensor networks. It is a macro-estimate to compare with that of the SSE power requirements and its concurrent implementation with time domain SSE. The above study was conducted based on results obtained from the wavelet method classification and identification. [70, 90]

## **6.7 Summary**

The SSE was implemented and evaluated using a database of tactical seismic and acoustic signals. Evaluation methods have been developed to characterize the SSE and optimize the selection of its matched filter correlators. Excellent signal identification performance has been obtained for acoustic signal and seismic signals with accuracy levels of greater than 90% were achieved from the ACIDS database.

The SSE has been used to resolve time evolution in acoustic and seismic signals. Multi-sensor classification and identification was performed with M-SSE with decision fusion between sensor types. Finally, the SSE architecture is well adapted to implementation in low power systems. In particular, future development effort may enable the SSE to be implemented in efficient, reconfigurable logic to permit high speed, and micro-power operation.

# CHAPTER 7

## Conclusion and Future Work

### 7.1 Conclusion

In this dissertation, a signal search engine (SSE) for signal classification and identification was presented. This method of time domain signal classification and identification has many potential applications in defense and commercial fields. Though many of the present day WINS applications are still in their infancy, the SSE can serve as a critical block for many wireless sensor applications.

Broad ranges of wireless applications were presented in the Introduction of Chapter 1. Many of these applications are already deployed with expanding commercial applications. Research is being carried out to optimize WINS sensor system lifetime with concentration on low-power, robust, and reconfigurable hardware – software co-design implementations. Chapter 2 shows battery power tradeoffs in transmitting signals and decisions for distributed and centralized signal processing and decision- making. An

analysis of battery power to lifetime was performed for variable power transceivers. Additional attention was given to power tradeoffs for transmission of decisions and data between sensor nodes for designs of cluster head based signal processing and transmission of decision to a remote centralized micro-sensor. An introduction to SSE / M-SSE was presented with applications in the military and commercial fields.

Distributed and centralized signal processing, and decision-making were presented in Chapter 3 with the introduction to possible system level architectures for WINS SSE applications. Decision-making was presented with maximum polling and weighted averaging for various cases used in WINS classification and identification. Decision – making based on sensor class / type, or state-spaces were discussed with the derivation of decision-making criteria, and confidence measure calculations for each case. It was concluded that the decision making criteria to implement on sensor nodes depends on the specific application, and should be decided during training with an application specific signal data-base. Signal pre-processing and ‘SCIA’ modules were presented from an architectural design perspective. These should be designed depending on application specific power, throughput, scalability, and complexity constraints.

Investigation of the SSE response to test signals was presented in Chapter 4 along with the creation and testing of possible real-world signal models. Test signals were created with single and multiple narrowband frequency signals, including real world wideband frequency signals with variable parameters as observed by sensors. The created signals

were used to investigate SSE performance to input SNR (variable SNRs introduced on signals with addition of variable levels of additive white gaussian noise (AWGN)), different levels of phase noise, presence of multiple-narrow band sources, effect of Doppler on moving vehicles, and environmental and circuit noise effects (such as glitches). Benchmarking of the time domain SSE to test signals was done with the MUSIC and Pisarenko parametric methods. It was concluded that the time domain SSE performed much better than ‘MUSIC’, and ‘Pisarenko’ methods. These two were noticed to have issues with closely spaced frequencies and low SNR signals. The behavior of the SSE to test signals was overwhelmingly positive, with classification and identification error percentages of 2-5%.

The time domain SSE was presented in Chapter 5, with a description and application of signal detection and classification to moving sources that have the most complex form of signals sensed. These signals contain a multitude of variables. These variables were presented with time-frequency spectrogram plots that were studied initially to obtain moving signal waveform characteristics from the training database. Spectrogram information was used extensively to study the signal behavior, but the software implementation has only time-domain analysis of time-frequency behavior. A method of reducing moving source variables was shown with segmentation of signals into state spaces using the segmentation module. This method of state-space decomposition of signals into state-spaces, and a state-space based SSE were explained with a description of template tree building, along with the SCIA algorithm.

Performance of the time domain SSE was presented with real world signals to validate time domain SSE / M-SSE for moving source classification and identification in Chapter 6. Time domain SSE modules were evaluated for selecting of correlator templates and confidence assignments for classification and identification. Tier based classification and identification was presented for real-world signals that give overwhelmingly accurate results, which are comparable with other classification and identification methods. The time domain SSE was benchmarked with the wavelet method classification and identification with a comparison of pros and cons along with possible co-implementation and techniques to enhance the accuracy of each scheme.

Based on the excellent accuracy levels obtained with both test signals and complex real-world wideband moving source signals during the validation and benchmarking of the SSE, we conclude that the SSE and its extension, the M-SSE, could be used with multi-sensing sensors.

## **7.2 Implementation: Hardware Software Co-Design**

With applications of WINS expanding in the consumer domain, SSE usage on WINS and other consumer applications has many possibilities. SSE and M-SSE can be used with the customization of application specific modules. Therefore, it is required to implement the SSE with hardware-software co-design or programmable FPGA hardware modules.

The dissertation research concentrated on a software-only implementation of the SSE / M-SSE. A software-only implementation is possible for applications with high computational processor powered systems. [19, 45] However, implementation of hardware-software modules and application specific IC's are necessary for low-power, integrated operation especially for WINS nodes. These applications will concentrate on reconfigurable, low-power hardware implementation that has to be initially tested and prototyped with an FPGA.

Building better classification / identification tree structures enhances SSE / M-SSE accuracy levels and throughput. Therefore, efforts need to concentrate on building re-programmable tree-structures in hardware, and application-specific IC's. It is important to have correlator templates built in a generic form for all nodes within a cluster as emphasized in this dissertation.

### **7.3 Future Research**

As a pre-cursor to future research, hardware-software co-design and low power hardware only implementation should be emphasized. [1, 25, 37, 77] It is inevitable that the integration of the SSE / M-SSE based on the requirements and specifications of the applications needs to be implemented.

Decision support systems will require SSE / M-SSE usage that need to be implemented with links to a well-established datawarehouse. Concentration needs to be focused on building a structured data warehouse, integrating the SSE / M-SSE, and building a decision support system that will cater to particular surveillance operations [86], condition based maintenance, and bio-medical monitoring systems. It is clear that with expanding applications in the consumer domain, a prototype automated correlator-template tree building algorithm (along with optimizing window sizing, and stepping) needs to be investigated.

Additional consideration should be given to investigating SSE / M-SSE accuracy levels to multi-source wideband source presence. SSE decision-making architecture should be exploited to curb low SNR signal-states with multi-vehicle presence. The presence of multiple sources are modeled and tested with a created test signal set that consists of up to three narrowband sources. However, a mixture of narrowband and wideband signals should be created and tested for SSE generalization.

Automation of threshold for triggering the re-building of the correlator template tree should be investigated further along with automated tree building for specific applications (along with dividing it into state spaces.) We conclude finally, that the time domain SCIA would be used extensively with various signal processing applications, both military and commercial.

## Bibliography

- [1] A. A. Abidi, G. J. Pottie, W. J. Kaiser, "Power-Conscious Design of Wireless Circuits and Systems", Proceeding of the IEEE, Vol. 88, No. 10, pp. 1528-1545, October 2000.
  
- [2] M. Ahmed, G. J. Pottie, "Information Theory of Wireless Sensor Networks: The n-helper Gaussian Case", IEEE International Symposium on Information Theory, pp. 436, June 2000.
  
- [3] J. Agre, L. Clare, "An Integrated Architecture for Cooperative Sensing Networks", Computer, Vol. 33, No. 5, pp. 106-108, May 2000.
  
- [4] S. A. Aldosari, J. M. F. Moura, "Fusion in Sensor Networks with Communication Constraints", IEEE Information Processing in Sensor Networks (IPSN 04), pp.108-115, April 2004.
  
- [5] G. Asada, M. Dong, T. S. Lin, F. Newberg, G. Pottie, H.O. Marcy, and W. J. Kaiser, "Wireless Integrated Network Sensors: Low Power Systems on a Chip", Proceedings of the 1998 European Solid State Circuits Conference, pp 15-20, October, 1998.
  
- [6] P. Bauer, M. Sichitiu, R. Istepanian, K. Premaratne, "The Mobile Patient: Wireless Distributed Sensor Networks for Patient Monitoring and Care", Proceedings of IEEE EMBS International Conference, pp.17-21, November 2000.

- [7] P. Bergamo, S. Asgari, H. Wang, D. Maniezzo, L. Yip, R. E. Hudson, K. Yao, D. Estrin, “Collaborative Sensor Networking Towards Real-Time Acoustical Beamforming in Free-Space and Limited Reverberance” IEEE Transactions on Mobile Computing, Vol. 3, No.3, pp.211-224, July – September 2004.
- [8] S. Bhattacharya, T. Abdelzaher, “Data Placement for Energy Conservation in Wireless Sensor Networks”, Master’s Thesis Presentation, University of Virginia, May 2002.
- [9] C. Boller, W. Staszewski, K. Worden, G. Manson, G. Tomlinson, “Structure Integrated Sensing and Related Signal Processing for Condition-Based Maintenance”, Tutorial, <http://www.arofe.army.mil/Conferences>.
- [10] A. L. Brandao, D. C. McLernon, L. B. Lopes, “Adaptive Blind Diversity Combining for Antenna Arrays”, IEEE Colloquium on Smart Antennas, Digest No. 1994/182, pp. 2/1 – 2/6, December 1994.
- [11] R. R. Brooks, P. Ramanathan, A. M. Sayeed, “Distributed Target Classification and Tracking in Sensor Networks”, IEEE Proceedings, Vol. 91, No. 8, pp. 1163-1171, August 2003.
- [12] W. D. Bryan, H. G. Nguyen, D. W. Gage, “Man-Portable Networked Sensor System”, <http://www.spawar.navy.mil/robots/pubs/spie3394.pdf>.
- [13] K. Bult, A. Burstein, D. Chang, M. Dong, M. Fielding, E. Kruglick, J. Ho, F. Lin, T. H. Lin, W. J. Kaiser, H. Marcy, R. Mukai, P. Nelson, F. Newberg, K. S. J. Pister, G. Pottie, H. Sanchez, O. M Stafsudd, K. B.

- Tan, C. M. Ward, S. Xue, J. Yao, "Wireless Integrated Microsensors", Proceedings of the 1996 IEEE Solid-State Sensor and Actuator Workshop.
- [14] K. Bult, A. Burstein, D. Chang, M. Dong, M. Fielding, E. Kruglick, J. Ho, F. Lin, T.-H. Lin, W. J. Kaiser, H. Marcy, R. Mukai, P. Nelson, F. Newberg, K. S. J. Pister, G. Pottie, H. Sanchez, O. M. Stafsudd, K. B. Tan, C. M. Ward, S. Xue, J. Yao, "Low Power Systems for Wireless Microsensors", International Symposium on Low Power Electronics and Design, pp. 17-21, August 1996.
- [15] A. Chandrakasan, R. Amirtharajah, S. H. Cho, J. Goodman, G. Konduri, J. Kulik, W. Rabiner, A. Wang, "Design Considerations for Distributed Microsensor Systems", Proceedings of the IEEE Custom Integrated Circuits, pp. 279-286, May 1999.
- [16] J. O. Chapa, M. R. Raghuveer, "Matched Wavelets – Their Construction, and Application to Object Detection", Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol, 5, pp. 2987-2989, May 1996.
- [17] C. Chatterjee, V.P. Roychowdhury, "An Adaptive Stochastic Approximation Algorithm for Simultaneous Diagonalization of Matrix Sequences With Applications", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 3, March 1997.
- [18] P. Cheeseman, J. Kelly, M. Self, J. Stutz, W. Taylor, D. Freeman, "AutoClass: A Bayesian Classification System", Proceedings of the Fifth International Workshop on Machine Learning, 1988.

- [19] A. Chen, R. R. Muntz, S. Yuen, I. Locher, S. I. Park, M. B. Srivastava, "A Support Infrastructure for the Smart Kindergarten", *Pervasive Computing*, pp. 49-57, April- June 2002.
- [20] J. C. Chen, K. Yao, "Wideband Signal Detection and Maximum-Likelihood Source Localization" *Asilomar Conference on Signals, Systems and Computers*, Vol. 2, pp. 931-935, November 2001.
- [21] J. C. Chen, R. E. Hudson, K. Yao, "Fast Frequency-Domain Acoustic Channel Estimation with Interference Cancellation", *IEEE International Conference on Acoustic, Speech, and Signal Processing*, Vol. 2, pp. 1709-1712, 2002.
- [22] J.C. Chen, K. Yao, R.E. Hudson, "Source Localization and Beamforming" *IEEE Signal Processing Magazine*, Vol. 19, No. 2, pp.30-39, March 2002.
- [23] J.C. Chen, "Wideband Source Localization Using Passive Sensor Arrays". Ph.D. Dissertation, University of California, Los Angeles, CA, 2002.
- [24] J. C. Chen, L. Yip, J. Elson, H. Wang, D. Maniezzo, R. E. Hudson, K. Yao, D. Estrin, "Coherent Acoustic Array Processing and Localization on Wireless Sensor Networks", *Proceedings of the IEEE*, Vol. 91, No. 8, pp.1154-1162, August 2003.
- [25] J. C. Chen, L. Yip, H. Wang, D. Maniezzo, R. E. Hudson, J. Elson, K. Yao, E. Estrin, "DSP Implementation of a Distributed Acoustical Beamformer on a Wireless Sensor Platform", *Proceedings of the IEEE International Conference on Acoustics, speech, and Signal Processing*, Vol. 2, pp. 597-600, April 2003.

- [26] Da-Ching Chen, “Blind Space-Time Processing for Wireless Communication and Sensor Systems”, Ph.D. Dissertation, University of California, Los Angeles, CA, 2002.
  
- [27] W. W. Chu, “Intelligent Information Systems”, Course Reader Material, Winter 2001.
  
- [28] W.T. Chu, J. Laufer, K. Kao, “Noise Source Distribution in Subsonic Jets”, International Conference on Noise Control, Washington, D.C., pp. 472-476, 1972.
  
- [29] J. S. Colburn, Y. Rahmat-Samii, M. A. Jensen, G. J. Pottie, “Diversity Performance of Dual Antenna Personal Communication Handsets”, International Symposium of Antennas and Propagation Society, Vol. 1, pp. 730-733, July 1996.
  
- [30] M. Coulon, J.Y. Tournet, M. Ghogho, “Detection and Classification of Spectrally Equivalent Processes: A Parametric Approach”, Proceedings of the IEEE Signal Processing Workshop on Higher-Order Statistics, pp.410-414, July 1997.
  
- [31] D. Culler, J. Hill, P. Buonadonna, R. Szewczyk, A. Woo, “A Network-Centric Approach to Embedded Software for Tiny Devices”, White Paper, Intel Research Berkeley, IRB-TR-01-001, Jan. 2001.
  
- [32] R. Davis and A. Prieditis, “Designing Optimal Sequential Experiments for a Bayesian Classifier”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 21, No. 3, pp. 193-201, March 1999.

- [33] A. D'Costa, A. M. Sayeed, "Data Versus Decision Fusion in Wireless Sensor Networks", Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 4, pp.832-835, April 2003.
  
- [34] A. D'Costa, V. Ramachandran, A. M. Sayeed, "Distributed Classification of Gaussian Space-Time Sources in Wireless Sensor Networks", IEEE Journal on Selected Areas in Communications, Vol. 22, No. 6, pp. 1026 – 1036, August 2004.
  
- [35] A. Djouadi and E. Bouktache, "A Fast Algorithm for the Nearest-Neighbor Classifier", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No. 3, pp. 277-282, March 1997.
  
- [36] M. J. Dong, G. Yung, and W. J. Kaiser, "Low Power Signal Processing Architectures for Network Microsensors", International Symposium on Low Power Electronics and Design, Digest of Technical Papers (1997), pp. 173-177, 1997.
  
- [37] R. T. Edwards, G. Cauwenberghs, "Mixed-Mode Correlator for Micropower Acoustic Transient Classification", IEEE Journal on Solid State Circuits, Vol 34, No. 10, pp. 1367-1372, October 1999.
  
- [38] J. C. Eidson, S. P. Woods, "A Research Prototype of a Networked Smart Sensor System" White Paper, Hewlett-Packard Company HPL-95-91, August 1995.
  
- [39] D. Estrin, L. Girod, G. Pottie, M. Srivastava, "Instrumenting the World with Wireless Sensor Networks," Proceedings of the IEEE International

Conference on Acoustic, Speech, and Signal Processing, vol.4, pp. 2033-2036, May. 2001.

- [40] S. Femmam, N.K. M'sirdi, "Application of Change Detection Technique to Ultrasound Signal Processing: Use of Spectral Parametric Estimation, Time-Frequency and Wavelet", Proceedings of the Southeastern Symposium on System Theory, pp.203-207, March 1998.
- [41] M.J. Fisher, K.R. Holland, "Measuring the Relative Strengths of a Set of Partially Coherent Acoustic Sources", Journal of Sound and Vibration, p.p. 103-125, 1997.
- [42] V. Gonzalez, E. Sanchis, G. Torralba, G. Torralba, and J. Martos, "Comparison of Parallel Versus Hierarchical Systems for Data Processing in Distributed Sensor Networks", IEEE Transactions on Nuclear Science, Vol. 49, No. 2, pp. 394-400, April 2002.
- [43] J. Han, M. Kamber, "Data Mining: Concepts and Techniques", Morgan Kaufmann, August 2000.
- [44] K. Harmanci, J. L. Krolik, "Robust Beamformer Weight Design for Broadband Matched-Field Processing", IEEE International Conference on Acoustic, Speech, and Signal Processing, Vol. 1, pp.471-474, April 1997.
- [45] J. L. Hill, D. E. Culler, "MICA: A Wireless Platform for Deeply Embedded Networks", IEEE Micro, Vol. 22, No. 6, pp.12-24, November-December 2002.

- [46] K.D. Hurst, T.G. Habetler, "A Comparison of Spectrum Estimation Techniques for Sensorless Speed Detection in Induction Machines", IEEE Transactions on Industry Applications, Vol.33, No. 4, pp. 898-905, July-August 1997.
- [47] S. S. Iyengar, Q. Wu, "Computation Aspects of Distributed Sensor Networks", Proceedings of the International Symposium on Parallel Architectures, Algorithms and Networks, pp. 19-26, May 2002.
- [48] A. Kansal, M. B. Srivastava, "An Environmental Energy Harvesting Framework for Sensor Networks", International Symposium on Low Power Electronics and Design, pp. 481-486, August 2003.
- [49] J. Kittler, M. Hatef, R.P.W. Duin, J. Mathas, "On Combining Classifiers", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20, No. 3, pp. 226-239, March 1998.
- [50] F. Kobayashi, Y. Tanabe, T. Fukuda, F. Kojima, "Acquisition of Sensor Fusion Rule Based on Environmental Condition in Sensor Fusion System", Joint IFSA World Congress and NAFIPS International Conference, Vol. 4, pp. 2096 – 2101, 2001.
- [51] R. Kohno, "Spatial and Temporal Communication Theory Using Adaptive Antenna Array", IEEE Personal Communications, Vol. 5, No. 1, pp.28-35, February 1998.
- [52] R. E. Learned, W. C. Karl, A. S. Willsky, "Wavelet Packet Based Transient Signal Classification", Proceedings of the IEEE-SP International

Symposium on Time-Frequency and Time-Scale Analysis, pp.109-112, October 1992.

- [53] W. Lee, S. J. Stolfo, K. W. Mok, “A Data Mining Framework for Building Intrusion Detection Models”, IEEE Symposium on Security and Privacy, pp. 120-132, May 1999.
- [54] M. Leopold, M. B. Dydensborg, P. Bonnet, “Bluetooth and Sensor Networks: A Reality Check”, First International Conference on Embedded Networked Sensor Systems, pp. 103-113, November 2003.
- [55] B. Li, R. Chellappa, “Simultaneous Tracking and Verification via Sequential Posterior Estimation”, IEEE Conference on Computer Vision and Pattern Recognition, Vol. 2, pp. 110-117, June 2000.
- [56] D. Li, K. D. Wong, Y. H. Hu, A. M. Sayeed, “Detection, Classification, and Tracking of Targets”, IEEE Signal Processing Magazine, pp. 17-29, March 2002.
- [57] B. Ma, “Parametric and Non Parametric Approaches for Multisensor Data Fusion”, Ph.D. dissertation, University of Michigan, 2001.
- [58] A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler, “Wireless Sensor Networks for Habitat Monitoring “, White Paper, Intel Research Berkeley, IRB-TR-02-006, June 2002.
- [59] C. Meesookho, S. Narayanan, C. S. Raghavendra, “Collaborative Classification Applications in Sensor Networks”, Proceedings of Sensor

Array and Multichannel Signal Processing Workshop, pp.370-374, August 2002.

- [60] I. Maravic, M. Vetterli, K. Ramachandran, "High Resolution Acquisition Methods for Wideband Communication Systems", IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 4, pp. 133-136, April 2003.
- [61] I. Maravic, M. Vetterli, K. Ramachandran, "Channel Estimation and Synchronization with Sub-Nyquist Sampling and Application to Ultra-Wideband Systems", Proceedings of the 2004 International Symposium on Circuits and Systems, Vol. 5, pp. 381-384, May 2004.
- [62] W. M. Merrill, F. Newberg, K. Sohrabi, W. Kaiser, G. Pottie, "Collaborative Networking Requirements for Unattended Ground Sensor Systems", Proceedings of the IEEE Aerospace Conference, vol. 5, pp. 2153-2165, March 2003.
- [63] R. Mukai, "Neural Networks for Correction of Pointing and Focal Errors on Large Deep Space Network Antenna at Ka-Band", Ph.D Dissertation, University of California, Los Angeles 2003.
- [64] D. Nadig, S. S. Iyengar, D. N. Jayasimha, "A New Architecture for Distributed Sensor Integration", IEEE Proceedings Southeastcon, pages: 8, April 1993.
- [65] S. Natkunanathan, A. Sipos, S. Avery, G. Pottie, W. Kaiser, "A Signal Search Engine for Wireless Integrated Networked Sensors", ASFL Annual Symposium March 2000.

- [66] F. N. Newberg, "Wireless Sensor Networks", Ph. D Dissertation, University of California, Los Angeles, 2002.
- [67] W. Ng, J. P. Reilly, T. Kirubarajan, "A Bayesian Approach to Tracking Wideband Targets Using Sensor Arrays and Particle Filters", IEEE Workshop on Statistical Signal Processing, pp. 510-513, September-October 2003.
- [68] W. Ng, J. P. Reilly, T. Kirubarajan, "Application of Particle Filters for Tracking moving Receivers in Wireless Communication Systems", 4<sup>th</sup> IEEE workshop on Signal Processing Advances in Wireless Communications, pp. 575-579, June 2003.
- [69] A. Perring, R. Szewczyk, V. Wen, D. Culler, J. D. Tygar, "SPINS: Security Protocols for Sensor Networks", White Paper, Intel Research Berkeley, IRB-TR-01-004, April 2001.
- [70] J. N. Q. Pham, "An Operator Theoretic Approach to Wavelet Theory and Applications", Ph. D Dissertation, University of California, Los Angeles, 1999.
- [71] S. Pittner, S.V. Kamarthi, "Introduction of Multiple Signal Processing Engines: Feature Extraction from Wavelet Coefficients for Pattern Recognition Tasks", IEEE Transactions on Pattern Analysis and Machine Intelligence, Jan 1999.

- [72] S. Pittner, S. V. Kamarthi, "Feature Extraction from Wavelet Coefficients for Pattern Recognition Tasks", IEEE transactions on Pattern Analysis and Machine Intelligence, Vol. 21, No. 1, pp.83-88, January 1999.
- [73] S. Pittner, S. V. Kamarthi, "Feature Extraction from Wavelet Coefficients for Pattern Recognition Tasks", International Conference on Neural Networks, Vol. 3, pp. 1484-1489, June 1997.
- [74] R. Pon, M. Batalin, M. Rahimi, Y. Yu, D. Estrin, G. J. Pottie, M. Srivastava, G. Sukhatme, W. J. Kaiser, "Self-Aware Distributed Embedded Systmes", IEEE International Workshop on Future Trends of Distributed Computing Systems, pp. 102-107, May 2004.
- [75] V. Popovici, J. P. Thiran, "Pattern Recognition using Higher-Order Local Autocorrelation Coefficients", Proceedings of the IEEE Workshop on Neural Networks for Signal Processing, pp. 229-238, September 2002.
- [76] G. J. Pottie, W. J .Kaiser, "Wireless Integrated Network Sensors", Communication of the ACM, vol.43, no.5, pp.51-58, May 2000.
- [77] G. J. Pottie, "Wireless Sensor Networks", IEEE Information Theory Workshop, pp. 139-140, June 1998.
- [78] G. J. Pottie, "Hierarchical Information Processing in Distributed Sensor Networks", Proceeding of the IEEE International Symposium on Information Theory, pp. 163, August 1998.
- [79] G. J. Pottie, Private F2F Communications.

- [80] J. G. Proakis, "Digital Communications", New York: McGraw-Hill 1995.
- [81] V. Raghunathan, C. Schurgers, S. Park, M. B. Srivastava, "Energy-Aware Wireless Microsensor Networks", IEEE Signal Processing Magazine, pp. 40-50, March 2002.
- [82] R. Rao, S. Vrudhula, D. N. Rakhmatov, "Battery Modeling for Energy-Aware System Design" IEEE Computer Society, pp. 77-87, December 2003.
- [83] T. S. Rappaport, K. Blankenship, H. Xu, "Propagation and Radio System Design Issues in Mobile Radio Systems for the GloMo Project" Tutorial Released January 31, 1997, Mobile and Portable Radio Research Group, Bradley Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University.
- [84] J. W. Shavlik, T. G. Dietterich, "Readings in Machine Learning", Morgan Kaufmann, Palo Alto, California, 1990.
- [85] A. Sipos, S. Natkunanathan, W. Kaiser, "A Time Domain Signal Classification and Identification Architecture for Wireless Integrated Network Sensors", ASFL Annual Symposium March 1999.
- [86] S. Slijepcevic, M. Potkonjak, V. Tsiatsis, S. Zimbeck, M. B. Srivastava, "On Communication Security in Wireless Ad-Hoc Sensor Networks", Proceedings of the 11<sup>th</sup> IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE '02).

- [87] K. Sohrabi, J. Gao, V. Ailawadhi, G. J. Pottie, "Protocols for Self-Organization of a Wireless Sensor Network", IEEE Personal Communications, pp. 16-27, October 2000.
- [88] K. Sohrabi, W. Merrill, J. Elson, L. Girod, F. Newberg, W. Kaiser, "Methods for Scalable Self-Assembly of Ad hoc Wireless Sensor Networks", IEEE Transactions on Mobile Computing, Vol. 3, No. 4, pp. 317-331, October-December 2004.
- [89] N. M. Su, H. Park, E. Bostrom, J. Burke, M. B. Srivastava, D. Estrin, "Augmenting Film and Video Footage with Sensor Data", IEEE Annual Conference on Pervasive Computing and Communications (PERCOM '04), pp. 3-12, March 2004.
- [90] Q. M. Tieng, W. W. Boles, "Object Identification Using the Dyadic Wavelet transform and Indexing Techniques", International Conference on Acoustics, Speech, and Signal Processing, Vol. 4, pp.2475-2478, May 1995.
- [91] M. N. Toksoz, D. H. Johnston, A. Timur, "Attenuation of Seismic Waves In Dry Saturated Rocks", *Geoph.* Vol44, p 681-690, April 1979.
- [92] U. S. Army Research Labs, "Acoustic and Seismic Classification and Identification Data Set", Adelphia, Maryland.
- [93] A. Wang, A. Chandrakasan, "Energy-Efficient DSPs for Wireless Sensor Networks", IEEE Signal Processing Magazine, Vol. 19, No. 4, pp. 68 – 78, July 2002.

- [94] M. Weiser, "The Computer for the 21<sup>st</sup> Century", IEEE Pervasive Computing, Vol.1, No.1, pp. 19-25, December 1995.
- [95] M. Weiser, "Some Computer Science Issues in Ubiquitous Computing", CACM, July 1993.
- [96] J. Widom, "Research Problems in Data Warehousing", Proceedings of the International Conference on Information and Knowledge Management, pp. 25-30, November 1995.
- [97] J. H. Winters, "Smart Antennas for Wireless Systems" IEEE Personal Communications, Vol.5, No.1, pp. 23-27, February 1998.
- [98] S. Yang, "Nonlinear Signal Classification Using Geometric Statistical Features in State Space", Electronic Letters, Vol. 40, No. 12, pp.780-781, June 2004.
- [99] K. Yao, J. C. Chen, R. E. Hudson, "Maximum-Likelihood Acoustic Source Localization: Experimental Results", Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 3, pp. 2949-2952, May 2002.
- [100] T. Yu, Target Identification Processor for Wireless Sensor Network, Ph. D. Dissertation, University of California, Los Angeles, 1999.
- [101] Z. Zhang, C. Zhang, "Result Fusion in Multi-Agent Systems Based on OWA Operator", 23<sup>rd</sup> Australasian Computer Science Conference, pp.234-240, Feb 2000.

- [102] F. Zhao, J. Liu, J. Liu, L. Guibas, J. Reich, “Collaborative Signal and Information Processing: An Information – Directed Approach”, Proceedings of IEEE, vol. 91, no.8, pp.1199-1209, August 2003.