A Data-Driven Framework for Ultrasonic Structural Health Monitoring of Pipes

Submitted in partial fulfillment of the requirements for

the degree of

Doctor of Philosophy

in

Civil and Environmental Engineering

Yujie Ying

B.S., Civil Engineering, Zhejiang University of Technology
M.S., Civil and Environmental Engineering, Carnegie Mellon University

Carnegie Mellon University
Pittsburgh, PA

August, 2012

ACKNOWLEDGEMENTS

This research would not have been possible without the support of many people. I would first like to express my deepest gratitude to my advisors, Professor James H. Garrett, Jr., Professor Irving J. Oppenheim, and Professor Lucio Soibelman. Thank you for your invaluable advice, support, and encouragement. The past four years' research experience under your guidance was a tremendous asset to my career and life. I am very fortunate and grateful to have you as my advisors, colleagues, and lifelong mentors.

My sincerest gratitude is also due to my dissertation committee, Professor Bhiksha Raj and Dr. Warren R. Junker, whose insightful advice and comments served as valuable inputs to the scientific significance and practical engineering values of this research.

I have been very fortunate to work in an interdisciplinary and collaborative research team with Mr. Joel B. Harley, Mr. Chang Liu, Mr. Abhinav Agrawal, Dr. Nicholas O'Donoughue, Mr. Jun Shi, Mr. Yanqing Bao, Professor José M. F. Moura, Professor Mario Bergés, Professor David W. Greve, and Professor Yuanwei Jin. I would not have completed this dissertation without the abundant help and resources from my team members.

My keen appreciation goes to Dr. Debaditya Dutta, Mr. Fernando A. Cerda Carrizo, Mr. Xuan (Peter) Zhu, Professor Lawrence Cartwright, Professor Jennifer E. Michaels, Professor Piervincenzo Rizzo, and Mr. Michael T. Stift, for their intellectual discussions and suggestions on my research. My sincerest gratitude to Professor Shiming Wang, Professor Yang Yang, and Professor Jianjun Zheng in my college, without whose inspiration and encouragement I would not have chosen to pursue a doctoral degree.

I would also like to acknowledge the financial support of Westinghouse Electric Company, National Energy Technology Laboratory (NETL) through Concurrent Technologies Corporation (CTC), Pennsylvania Infrastructure Technology Alliance (PITA), and Carnegie Institute of Technology Dean's Fellowship. Special thanks extend to IBM Corporation for the support of the IBM Smarter Infrastructure Lab at Carnegie Mellon University (CMU); this work used the IBM computing and data storage servers for conducting the analyses.

I am heartily thankful to my friends, especially Mr. Marc Bilder, Ms. Likun Tan, Mr. Chunyu Bai, Mr. Suvrajit Maji, Dr. Sourav Chatterjee, Mr. Xuesong (Pine) Liu, Dr. Sharad V. Oberoi, Dr. Daniel Oliveira, Dr. Ya-Mei Yang, Mrs. Amy King, and Ms. Amy Yu for their help and support in my research and life.

I also wish to express my cordial appreciation to all the faculty and staff in the Department of Civil and Environmental Engineering (CEE) at CMU for creating a warm, family-like climate and culture.

Finally, I am forever indebted to my parents, Zhenming Ying and Huiying Sun, for their endless love, understanding, patience, support and encouragement.

.

ABSTRACT

Cylindrical shells serve important roles in broad engineering applications, such as oil and natural gas pipelines, and pressurized industrial piping systems. To ensure the safety of pipe structures, various inspection equipment and platforms have been developed based on non-destructive testing (NDT) technologies. However, most existing approaches are time and labor intensive, and are only conducted intermittently. Drawbacks of current NDT methods suggest a proactive, automated and long-term monitoring system. Structural health monitoring (SHM) techniques continuously assess structural integrity through permanently installed transducers, allowing condition-based maintenance to replace the current practice of economically inefficient schedule-based maintenance.

Ultrasonics is an appealing SHM technology in which guided waves interrogate long stretches of a pipe with high sensitivity to damage, and can be generated by a surface-mounted, small-size piezoelectric wafer transducer (PZT). The challenges of implementing ultrasonic SHM with PZTs as active sensing devices lie in: (1) the wave pattern is complex and difficult to interpret; (2) it is even more difficult to differentiate changes produced by damage from changes produced by benign environmental and operational variability.

The ultimate goal of this research is to develop an ultrasonic sensing and data analysis system for continuous and reliable monitoring of pipe structures. The objective of this dissertation is to devise a data-driven framework for effective and robust analysis of guided wave signals to detect and localize damage in steel pipes under environmental and operational variations. The framework is composed of a three-stage SHM scheme: damage detection,

damage localization and damage characterization, supported by a multilayer data processing architecture incorporating statistical analysis, signal processing, and machine leaning techniques.

The data-driven methodology was first investigated through laboratory experiments conducted on a pipe specimen with varying internal air pressure. The sensed ultrasonic data were characterized and mapped onto a high dimensional feature space using various statistical and signal processing techniques. Machine learning algorithms were applied to automatically identify effective features, and to detect and localize a weak scatterer on the pipe. The reliability and generality of the data-driven framework was further validated through field tests performed on an in-service hot-water pipe under large, complex and uncontrollable operating conditions.

This data-driven SHM methodology involves an integrated process of sensing, data acquisition, statistical analysis, signal processing, and pattern recognition, for continuous tracking of the structural functionality in an adaptive and cost-effective manner. The techniques developed in this dissertation are expected to have broader applications related to the regular inspection, maintenance, and management of critical infrastructures not just limited to pipes.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ii

ABSTRACT iv

TABLE OF CONTENTS vi

LIST OF FIGURES x

LIST OF TABLES xvii

1 INTRODUCTION 1

1.1 Background and Motivation 1

1.2 Objectives and Research Questions 6

1.3 Thesis Organization 11

2 ULTRASONIC GUIDED WAVES FOR PIPE MONITORING 12

2.1 Guided Wave Propagation in a Cylindrical Shell 12

2.2 Evaluation of Ultrasonics for Pipe Monitoring 15

3 A DATA-DRIVEN PIPE MONITORING FRAMEWORK 21

3.1 Structural Health Monitoring Paradigm 21

3.2 Multilayer Data Processing Architecture 22

4 STATISTICAL ANALYSIS AND SIGNAL PROCESSING FOR CHANGE DETECTION 27

4.1 Fundamental Statistical Features 27

4.2 Time Reversal Signal Processing Features 30
4.2.1 Time Reversal Focusing 31
4.2.2 Time Reversal Change Focusing 32
4.2.3 Experimental Results of Time Reversal for Change Detection 35
4.3 Evaluation of Change Detectors 39
4.4 Alternative Change Detectors based on Learning Process 43
4.5 Summary 43
5 MACHINE LEARNING FOR SCATTERER DETECTION AND LOCALIZATION UNDER ENVIRONMENTAL AND OPERATIONAL VARIATIONS 45
5.1 Experimental Methods for Data Acquisition 45
5.2 Experimental Observations 47
5.3 Machine Learning Test Formulation 51
5.4 Feature Extraction using Combined Signal Processing and Machine Learning 54
5.4.1 Feature Domain Construction 55
5.4.2 Extracted Features 67
5.4.2.1 Baseline-Free Features 67
5.4.2.2 Baseline-Dependent Features 68
5.5 Integrated Feature Selection and Evaluation 70
5.5.1 Integrated Filter-Wrapper Feature Selection Method 70
5.5.2 Results of Feature Selection 74

5.6 Pattern Recognition for Scatterer Detection and Localization 83
5.6.1 Adaptive Boosting Algorithms 83
5.6.2 Support Vector Machines 85
5.6.3 Metrics for Machine Learning Classifiers 89
5.6.4 Results of Scatterer Detection 91
5.6.5 Results of Scatterer Localization 98
5.7 Summary 102
6 FIELD VALIDATIONS 103
6.1 Field Environment and Test Setup 103
6.2 Field Test Observations 106
6.3 Scatterer Detection Using Data-Driven SHM Framework 109
6.3.1 Field Test 1: Detecting a Scatterer at Different Locations 110
6.3.1.1 Data Collection and Machine Learning Test Formulation 110
6.3.1.2 Random Testing 113
6.3.1.3 Systematic Testing 121
6.3.2 Field Test 2: Detecting a Scatterer at One Location 126
6.3.2.1 Data Collection and Machine Learning Test Formulation 126
6.3.2.2 Random Testing 128
6.3.2.3 Systematic Testing 135

6.3.3 Field Test 3: Localizing a Scatterer at Two Different Locations 140
6.3.3.1 Data Collection and Machine Learning Test Formulation 140
6.3.3.2 Testing Results 142
6.3.4 Field Test 4: Detecting Different Scatterers 145
6.3.4.1 Data Collection and Machine Learning Test Formulation 145
6.3.4.2 Testing Results 147
6.4 Summary 149
7 CONCLUSIONS 150
7.1 Summary and Remarks 150
7.2 Contributions and Broader Impacts 155
7.3 Future Research 159
REFERENCES 162
APPENDICES 170
Appendix A: Comparison of Non-Destructive Evaluation Technologies 170
Appendix B: List of Relevant Publications 173

LIST OF FIGURES

Figure 1.1 Natural gas transmission pipelines in the United States (EIA 2009) 1

Figure 1.2 Pipeline hazard profile summary (GBRA 2005). 2

Figure 2.1 (a) Longitudinal modes; (b) torsional modes (c) flexural modes (Agrawal et al. 2009). 13

Figure 2.2 Dispersion curves of low-order wave modes in a steel pipe with 60.3 mm outside diameter and 3.6 mm wall thickness, generated by PCdisp. 14

Figure 2.3 Snapshot of wave propagation in a pipe, generated by PZFlex finite element simulation. 15

Figure 2.4 (a) A surface-mounted PZT wafer (photo courtesy of Joel B. Harley), and (b) an National Instruments PXI system. 17

Figure 2.5 (a) 200 kHz sinc windowed sinusoidal excitation, and (b) received signal, a 10 ms time history. 17

Figure 2.6 (a): Received signal before a saw-cut is induced, (b): received signal after a saw-cut is induced, (c): difference of signal (a) from baseline subtraction, and (d): difference of signal (b) from baseline subtraction (Ying et al. 2010). 19

Figure 2.7 Two signal responses from an undamaged pipe on two different days. 19

Figure 2.8 Three challenging issues in ultrasonic SHM implementation for pipes. 20

Figure 3.1 A data-driven structural health monitoring framework. 24

Figure 4.1 Scheme of the time reversal experiments. 36

Figure 4.2 (a) Received signal, and (b) signal obtained after time reversal focusing 36

Figure 4.3 (a): Received signal before a saw-cut is induced, (b): received signal after a saw-cut is induced, (c): difference of signal (a) from baseline subtraction, (d): difference of signal (b) from baseline subtraction, (e): signal obtained after applying time reversal to (c), and (f): signal obtained after applying time reversal to (d) (Ying et al. 2010). 37

Figure 4.4 Comparison of features extracted from the time domain and the TRCF-domain for damage detection. 38

Figure 4.5 (a): Received signal under 0 PSI, (b): received signal under 83 PSI, (c): difference of signal (a) from baseline subtraction, (d): difference of signal (b) from baseline subtraction, (e): signal obtained after applying time reversal to (c), and (f): signal obtained after applying time reversal to (d). 40

Figure 4.6 Comparison of features extracted from the time domain and the TRCF-domain under varying internal air pressure..... 42

Figure 5.1 (a) Schematic of the pressurized steel pipe, and (b) photo of the mass scatterer on the pipe..... 47

Figure 5.2 (a): Sinc pulse excitation, energy-normalized Collection I data: (b): a baseline signal recorded under 0 PSI, (c): received signal under 110 PSI, with the absence of the mass scatterer, (d): received signal under 0 PSI, with the presence of the scatterer, (e): differential signal of (c) and (b), and (f): differential signal of (d) and (b). 49

Figure 5.3 L2-norms of differential signals: (a) Collection I, (b) Collection II, and (c) Collection III. (Dots: undamaged instances; crosses: damaged instances.) 50

Figure 5.4 Received signals with the absence of the mass scatterer, (a) a Collection II record, and (b) a Collection III record. 50

Figure 5.5 Absolute value correlation matrix of (a) the measurements and (b) the differential measurements of the three collections. 51
Figure 5.6 Procedure of machine learning process. 52
Figure 5.7 Flowchart of feature domain construction. 56
Figure 5.8 Feature domains of a scatterer record from Collection I: (a) pre-processed time domain signal, (b) detailed wavelet coefficients, (c) de-noised signal after wavelet processing, (d) signal envelope, (e) frequency domain, (f) time-reversal domain, (g) scale domain, (h) PCA reconstruction error in a subspace constructed using undamaged records from a basis set, (i) PCA reconstruction error in a subspace constructed using scatterer records from the basis set, and (j) FLD reconstructed signal using a single FLD basis computed from the basis set. 64
Figure 5.9 Procedure of the integrated filter-wrapper feature selection. 72
Figure 5.10 Results of the Welch's *t*-statistics filter feature selection in a multi-day systematic test. (a) *p*-values of the features in the original order, (b) histogram of the *p*-values of the 1500 features. 75
Figure 5.11 Absolute value feature correlation matrices of the features selected by different filters (plots in the left column), and the correlation coefficients of these selected features and the labels of the training data (plots in the right column), in a multi-day systematic test: (a)(b) results after the *t*-test, (c)(d) results after the R-filter, and (e)(f) results after the S-filter. (Features are sorted in an ascent order of *p*-value.) 76
Figure 5.12 Examples of a normalized feature space defined by the first two features selected in the systematic testing on: (a) Collection I, (b) Collection II, (c) Collection III, and (d) combination of the three collections. 79

Figure 5.13 Projections of (a) the wavelet coefficients and (b) the time domain signals in a PCA subspace (Collection I data). 80
Figure 5.14 Sample variance of peak values greater than 20% of the maximum amplitude in the Mellin domain of wavelet coefficients: (a) Collection I, (b) Collection II, and (c) Collection III. 82
Figure 5.15 Signal variation due to (a) pressure change, and (b) the mass scatterer. 83
Figure 5.16 Formulation of the maximum margin optimization (Burges 1998). 86
Figure 5.17 Classification becomes easier by proper feature mapping. 87
Figure 5.18 Comparison of the average performance of the classifiers for scatterer detection in the random testing (left column) and the systematic testing (right column): (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM. 96
Figure 5.19 Comparison of the average performance of the feature selectors in random testing (left column) and systematic testing (right column): (a) R-filter, (b) S-filter, and (c) wrapper. .. 97
Figure 5.20 Scatterer localization results of the random testing on Collection IV, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM. 100
Figure 5.21 Scatterer localization results of the systematic testing on Collection IV, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM. 101
Figure 6.1 (a) The pipe indicated with a dashed line is the hot water return pipe which was instrumented in our experiments, (b) PZT transducer mounted on the pipe wall, and (c) a cylinder mass with grease (or honey) attached to the pipe used to simulate damage (C. Liu et al. 2012). 104
Figure 6.2 Received signal before and after applying a 125 kHz high pass filter (C. Liu et al. 2012). 107

Figure 6.3 Correlation coefficients of 20-hr measurements with respect to the first measurement (C. Liu et al. 2012). 108
Figure 6.4 Correlation coefficients variation aligned with temperature and flow rate fluctuation (C. Liu et al. 2012). 109
Figure 6.5 Energy variation over time of (a) Collection F-I, and (b) Collection F-II. 111
Figure 6.6 Performance of the first selected feature by a wrapper method (associated with linear SVM) in (a) a random test on Collection F-I, (b) a random test on Collection F-II, and (c) a random test on the combined Collections F-I and F-II. 114
Figure 6.7 Scatterer detection results of the random testing on Collection F-I, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM. 116
Figure 6.8 Scatterer detection results of the random testing on Collection F-II, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d. RBF-SVM. 118
Figure 6.9 Scatterer detection results of the random testing on combined Collections F-I and F-II, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM. 120
Figure 6.10 Performance of the first selected feature by wrapper method (associated with linear-SVM) in (a) the first systematic test on Collection F-I, (b) the second systematic test on Collection F-I, (c) the first systematic test on Collection F-II, and (d) the second systematic test on Collection F-II. 122
Figure 6.11 Scatterer detection results of the systematic testing on Collection F-I, with Collection F-II as a basis set, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM. 124

Figure 6.12 Scatterer detection results of the systematic testing on Collection F-II, with Collection F-I as a basis set, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM. 125
Figure 6.13 Energy variation over time of (a) Collection F-III, and (b) Collection F-IV. 127
Figure 6.14 Performance of the first selected feature by a wrapper method (associated with linear-SVM) in (a) a random test on Collection F-III, (b) a random test on Collection F-IV, and (c) a random test on the combined Collections F-III and F-IV..... 129
Figure 6.15 Scatterer detection results of the random testing on Collection F-III, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM. 131
Figure 6.16 Scatterer detection results of the random testing on Collection F-IV, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM. 132
Figure 6.17 Scatterer detection results of the random testing on combined Collections F-III and F-IV, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM..... 134
Figure 6.18 Performance of the first selected feature by a wrapper method (associated with linear-SVM) in (a) the first systematic test on Collection F-III, (b) the second systematic test on Collection F-III, (c) the first systematic test on Collection F-IV, and (d) the second systematic test on Collection F-IV..... 136
Figure 6.19 Scatterer detection results of the systematic testing on Collection III, with Collection IV as a basis set, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM. 137
Figure 6.20 Scatterer detection results of systematic testing on Collection F-IV, with Collection F-III as a basis set, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM. 139

Figure 6.21 Energy variation over time of Collection F-V. 141
Figure 6.22 Scatterer localization results of the random testing on Collection F-V, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM. 143
Figure 6.23 Scatterer localization results of the systematic testing on Collection F-V, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM. 144
Figure 6.24 An aluminum bar (with the cylinder mass on the top) grease-coupled to the pipe to simulate larger damage (C. Liu et al. 2012)..... 146
Figure 6.25 Energy variation over time of Collection F-VI. 146
Figure 6.26 Performance of the first selected feature by a wrapper method (associated with linear-SVM) in the systematic testing on Collection F-VI: (a) measurements of the first toggling cycle (Scatterer 1 data) as the basis set, (b) measurements of the second toggling cycle (partial of the Scatterer 2 data) as the basis set. 148
Figure 7.1 Threefold contributions of this research. 157

LIST OF TABLES

Table 1-1 Severe pipeline accidents (Parfomak 2011). 3

Table 2-1 Advantages and disadvantages of guided waves for pipe monitoring. 20

Table 4-1 Comparison of time-domain features between two signals before and after a saw-cut is created. 29

Table 4-2 Comparison of TRCF-domain features between two signals before and after a saw-cut is created. 38

Table 4-3 Comparison of time-domain features between two signals before and after the pipe is pressurized. 41

Table 4-4 Comparison of TRCF-domain features between two signals before and after the pipe is pressurized. 42

Table 5-1 A breakdown of the datasets in the three collections. 47

Table 5-2 Machine learning test formulation for the laboratory studies. 54

Table 5-3 Procedure of PCA subspace generation, and construction of PCA-based feature domains and features..... 65

Table 5-4 Procedure of FLD subspace generation, and construction of FLD-based feature domains and features..... 66

Table 5-5 List of features extracted (1500 in total). 69

Table 5-6 Range of the number of features selected in the random testing..... 77

Table 5-7 Range of the number of features selected in the systematic testing 77

Table 5-8 Common types of kernel functions (Fan et al. 2005). 88

Table 5-9 Confusion matrix. 89

Table 5-10 Summary of the applied algorithms. 92
Table 5-11 The best algorithm pairs of a feature selector and a classifier in the random testing for each collection and the mixed set, and the resulting average AUC and ACC. 93
Table 5-12 The best algorithm pairs of a feature selector and a classifier in the systematic testing for each collection and the mixed set, and the resulting average AUC and ACC. 93
Table 6-1 Comparison of the laboratory experiments and the field tests. 106
Table 6-2 Field data collection details (Field Test 1). 111
Table 6-3 Machine learning testing formulation for scatterer detection (Field Test 1). 113
Table 6-4 Comparison of the feature selectors and the classifiers in the random testing on Collection F-I (bold font indicates the best performance). 117
Table 6-5 Comparison of the feature selectors and the classifiers in the random testing on Collection F-II (bold font indicates the best performance). 118
Table 6-6 Comparison of the feature selectors and the classifiers in the random testing on the combined Collections F-I and F-II (bold font indicates the best performance). 120
Table 6-7 Comparison of the feature selectors and the classifiers in the systematic testing on Collection F-I using Collection F-II as the basis set (bold font indicates the best performance). 124
Table 6-8 Comparison of the feature selectors and the classifiers in the systematic testing on Collection F-II using Collection F-I as the basis set (bold font indicates the best performance). 126
Table 6-9 Field data collection details (Field Test 2). 127
Table 6-10 Machine learning testing formulation for scatterer detection (Field Test 2). 128

Table 6-11 Comparison of the feature selectors and the classifiers in the random testing on Collection F-III (bold font indicates the best performance). 131
Table 6-12 Comparison of the feature selectors and the classifiers in the random testing on Collection F-IV (bold font indicates the best performance). 133
Table 6-13 Comparison of the feature selectors and the classifiers in the random testing on the combined Collections F-III and F-IV (bold font indicates the best performance). 134
Table 6-14 Comparison of the feature selectors and the classifiers in the systematic testing on Collection F-III using Collection F-IV as the basis set (bold font indicates the best performance). 138
Table 6-15 Comparison of the feature selectors and the classifiers in the systematic testing on Collection F-IV using Collection F-III as the basis set (bold font indicates the best performance). 139
Table 6-16 Field data collection details (Field Test 3). 141
Table 6-17 Machine learning testing formulation for scatterer localization (Field Test 3). 141
Table 6-18 Field data collection details (Field Test 4). 146
Table 6-19 Machine learning test formulation for detection of different scatterers (Field Test 4). 147
Table 7-1 Generality of the data-driven SHM framework. 154
Table 7-2 Remarks on data-driven methods for pipe monitoring. 154
Table 7-3 Summary of the detailed impacts of this research. 158
Table A-1 List of common non-destructive evaluations techniques (Shull 2002). 170

1 INTRODUCTION

1.1 Background and Motivation

Cylindrical shells or pipe-like structures serve important roles in broad engineering applications including tubing, piping, and pressure vessels for transportation and storage of gases, liquids or solids, closely related to energy supply and critical infrastructures like nuclear power plants, refineries and petrochemical facilities, coal mines, airports, seaports, and military bases (Parfomak 2011).

Considering natural gas pipes as an example, over 320 thousand miles of transmission and gathering pipelines transport large volumes of natural gas from production wells to major processing or storage facilities across the United States (Figure 1.1) and more than 2 million miles of distribution pipelines distribute gas to residential and commercial buildings, according

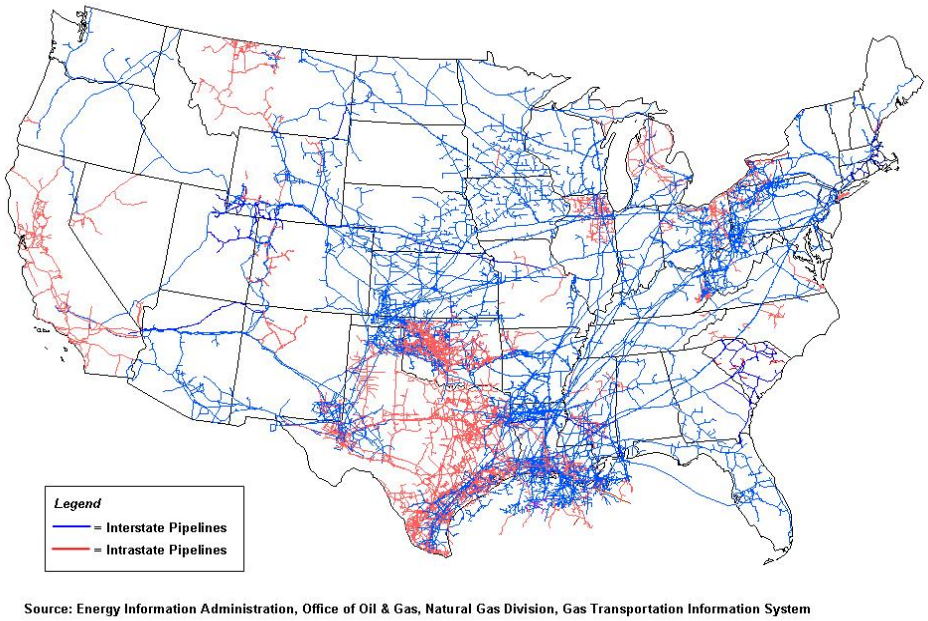


Figure 1.1 Natural gas transmission pipelines in the United States (EIA 2009)

to the pipeline mileage statistics provided by Department of Transportation (DOT) Pipeline and Hazardous Materials Safety Administration (PHMSA 2011c).

While supporting almost every aspect in industry, business and domestic life, the pipe network is also vulnerable to a variety of risks from third party excavations, eroding surrounding environments, and natural forces (Parfomak 2011). Moreover, the majority of the pipes are aging, because more than 60% of them were installed 40 years ago and are more likely subject to flaws and anomalies, such as corrosion, cracks, dents and gouges (e-CFR 2011; Nestleroth and Bubenik 1999; PHMSA 2011a). PHMSA reports that in 2010, there were 79 natural gas transmission and gathering pipeline accidents, 54 natural gas distribution accidents, and 111 hazardous liquid pipeline accidents. The major causes of significant pipeline failures include damage from digging, corrosion, overload, and material fatigue (PHMSA 2011b). Although pipeline accidents occur occasionally, each failure has only 0 to 3 hours average warning time and tends to cause substantial disruption in civic facilities, loss of property and even deaths, as illustrated in Figure 1.2 (GBRA 2005). Parformak 2011 lists a few recent severe pipeline accidents in a Congressional Research Service report (Table 1-1). The catastrophic incidents have highlighted the necessity of early detection and early warning of defects and imperfections in pipe infrastructures. As a result, the U.S. DOT has recently announced a new Pipeline Safety Action Plan to emphasize and urge regular inspection and maintenance, as well as replacement

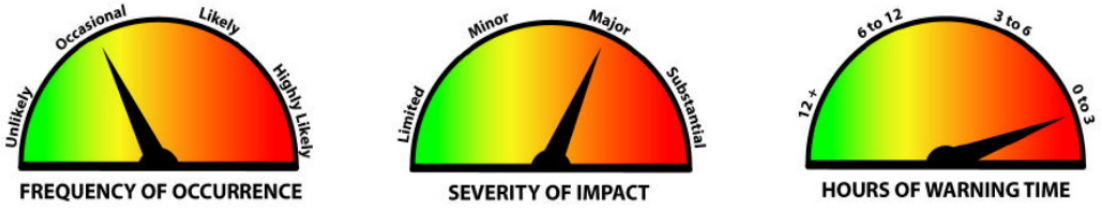


Figure 1.2 Pipeline hazard profile summary (GBRA 2005).

of aging pipes (DOT 2011).

In order to ensure the safety of pipe systems, many inspection tools and platforms have been developed based on a variety of non-destructive testing (NDT) technologies, such as acoustic emission, eddy current testing, magnetic flux leakage, vibration method, ultrasonic testing. The advantages and limitations of common NDT techniques are summarized in Appendix A.

Table 1-1 Severe pipeline accidents (Parfomak 2011).

Year	Accidents
1999	*A gasoline pipeline explosion in Bellingham, WA, killed three people and caused $45 million in damage to a city water plant and other property.*
2000	*A natural gas pipeline explosion near Carlsbad, NM, killed 12 campers.*
2006	*Corroded pipelines on the North Slope of Alaska leaked over 200,000 gallons of crude oil in an environmentally sensitive area and temporarily shut down Prudhoe Bay oil production.*
2007	*An accidental release from a propane pipeline and subsequent fire near Carmichael, MS, killed two people, injured several others, destroyed four homes, and burned over 70 acres of grassland and woodland.*
2010	*A pipeline spill in Marshall, MI, released 819,000 gallons of crude oil into a tributary of the Kalamazoo River.*
2010	*A natural gas pipeline explosion in San Bruno, CA, killed 8 people, injured 60 others, and destroyed 37 homes.*
2011	*A natural gas pipeline explosion in Allentown, PA, killed 5 people, damaged 50 buildings, and caused 500 people to be evacuated.*

Two common types of equipment for pipe inspection are intelligent mobile sensing devices and standalone inspection instruments (Piro 2010). Intelligent mobile sensing devices, often referred as pipeline inspection gauges (pigs), are intelligent moving robots equipped with sensors, Explorer (NETL 2007) as an example. Pigs can be propelled down pipes to measure temperature, strain, and/or wall-thickness, and then the recording data is analyzed to reflect the pipe conditions. However, pigs are limited by pipe geometries due to over- or under-sized valves, radius or miter bends of pipes, etc. (Piro 2010). Standalone inspection systems are traditionally probe-type instruments with a pulse-echo configuration that scan over the external surface of pipes checking for defects, but such equipment is typically only suitable for spot checking rather than long-range inspection. Lately, phase-array-type standalone instruments have made use of ultrasonic waves that can travel long distances in pipes. A wave beam can be formed to travel along any direction in the pipe by tuning a collar or a comb configuration of multi-element transducers with phase delays, and echoes from damage are recorded as indications of the existence and location of the damage (Cawley 2003; Li and J. L. Rose 2001b; Lowe et al. 1998; J. L. Rose et al. 1998). This type of system requires elaborate setups for mobility and mode selection, and has to be operated under high power (Piro 2010).

To summarize, although NDT technology has been widely deployed to inspect pipe systems, most existing approaches are time and labor intensive and are only conducted intermittently. The constraints of pipe inspection make clear the need for a proactive, automatic, and continuous monitoring system to perform structural health monitoring (SHM).

SHM is an "*integration of sensing and possibly also actuation devices to allow the loading and damaging conditions of a structure to be recorded, analyzed, localized, and*

predicted in a way that nondestructive testing becomes an integral part of the structure and material", as defined in the book of Encyclopedia of Structural Health Monitoring (Boller et al. 2009). In SHM, permanently installed transducers continuously assess the structural integrity, allowing condition-based maintenance to replace the current practice of economically inefficient schedule-based maintenance.

A number of NDT techniques have been converted into SHM applications, among which guided wave ultrasonics is an appealing technology. Ultrasonic elastic waves propagating into a medium guided by its geometric boundary are known as guided waves, whose favorable characteristics are long coverage and high sensitivity to damage (Rose 1999; Shull 2002). In addition, it has been demonstrated that guided waves can be easily generated in plate-like structures by using a small-size and low-power piezoelectric wafer made of Lead Zirconate Titanate (PZT) (Crawley and De Luis 1987; Giurgiutiu and Zagrai 2000; Nienwenhuis et al. 2005). Such guided waves, known as Lamb waves (Lamb 1917), have been extensively studied for damage detection and characterization in plates (Alleyne and Cawley 1992a; Ihn and Chang 2004; Pierce et al. 2000). For a pipe, the cylindrical shell serves as an excellent waveguide for ultrasonic waves propagating over a long range with full illumination through the pipe wall thickness (Demma et al. 2003; Lowe et al. 1998; Rose 1999). Our prior research has shown the feasibility of using PZT wafers attached to a pipe surface to produce guided waves in cylindrical shells (Harley et al. 2009; Ying et al. 2010). Consequently, ultrasonic technology using PZT wafers as active sensing devices is a promising technique for continuous pipe monitoring.

However, guided waves in a cylindrical shell consist of an infinite number of modes, most of which are highly dispersive and travel along multiple paths (Harley et al. 2009; Lowe et

al. 1998; Ying et al. 2010). As a result, changes produced by small damage are usually difficult to recognize in the complicated sensed waveform (Harley et al. 2009; Ying et al. 2010). In addition, a universal challenge for SHM lies in distinguishing the changes induced by damage from the changes produced by benign environmental and operational conditions, such as temperature, surface moisture (Y. Lu and J. E. Michaels 2009), and fluid flow or air pressure in pipes (Ying et al. 2011, 2012). An increasing number of researchers have considered this to be the main challenge that must be overcome before a SHM system can be successfully deployed in the real world (Sohn 2007). Only limited work has been directed at pipes, where the waveforms are more complex than Lamb waves in plate-like structures.

1.2 Objectives and Research Questions

The ultimate goal of this research is to develop an ultrasonic sensing and data analysis system for continuous and reliable monitoring of metallic pipe structures with permanently installed PZT wafer transducers. The challenges of implementing SHM of pipes with PZTs as active sensing devices are summarized as follows:

- Effective signal analysis is required to compensate for the dispersive, multi-modal, and multi-path effects of guided waves in a pipe.
- The damage-induced change in the sensed signal needs to be characterized to expose useful information associated with damage existence and location.
- A robust or adaptive methodology must be designed against environmental and operational disturbances in order to deliver real-world applications.

Data interpretation methods in SHM research and development include model-driven approaches and data-driven approaches. The model-driven approaches compare the responses of

a structure with those of a predicted model (such as an analytical model or a finite element model) that is supposed to reflect the actual structural behaviors. The predicted model is based on physical and mechanical characteristics of a structure, and therefore, in theory, can directly identify the cause of a change and quantify the influence of that change on the structural performance. However, reliably simulating the performance of structures is substantially difficult and expensive. Due to the uniqueness and uncertainties associated with structures or segments of a structure, many attributes cannot be modeled (such as structural imperfection) or are inaccurately modeled (such as environmental effects). Given many different sizes of pipes, different types of damage, and immeasurable or unknown environmental factors, model-based methods are prone to degrading adaptability and predictability for pipe monitoring (Catbas et al. 2011; Posenato et al. 2010).

The second type of data interpretation approach is data-driven and performed without forming a physical model of the structure, rather utilizing computer algorithms to automatically recognize damage-oriented patterns from measurement datasets in an information mining fashion. The data-driven approaches are easier to implement, generally less expensive, thus appealing for continuous monitoring (Catbas et al. 2011; Posenato et al. 2010). The difficulty of data-driven methods lies in finding the physical meanings behind the data-driven outcomes.

The objective of this research is to devise a data-driven framework for effective and robust analysis of guided wave signals to detect and localize damage in steel pipes under environmental and operational variations. In view of the aforementioned challenges and for practical SHM needs, three research questions are defined in this work, as described below along with relevant literature review.

- **Research Question 1**: What approaches in the statistical and signal processing domains can be used to quantify a change in the ultrasonic response obtained from a pipe under two different conditions?
- **Research Question 2**: How does the user determine whether the change in the ultrasonic response is produced by structural damage or produced by benign environmental and operational variations?
- **Research Question 3**: Once the damage is detected, what techniques can be used to localize the damage in the pipe (again, under varying environmental and operational conditions)?

A significant amount of research has been conducted to automate damage detection techniques in the context of SHM. One key requirement for such automation is that the changes in the signal caused by damage should be quantified. Techniques range from evaluating the norm of direct subtraction of time domain signals as a straightforward means to computing the change in magnitude of nonlinear harmonics (Buck et al. 1978; Hillis et al. 2006; Rivola 2000), to other measures that include changes computed from the real part of electromechanical impedance (Bhalla and Kiong Soh 2003; Giurgiutiu et al. 2002; G. Park et al. 2003), the local temporal coherence (J. E. Michaels and T. E. Michaels 2005), the coefficients of wavelet transforms of measured signals (Rizzo et al. 2005, 2009; Taha et al. 2006), the matching pursuit wavelet decomposition (Y. Lu and J. E. Michaels 2008), the amplitude of mode converted signal components (Kim and Sohn 2008; Sohn et al. 2011), and so forth.

Recently, several researchers have attempted to develop a quantitative index from the change in time reversal ultrasonic signal (Harley et al. 2009; O'Donoughue et al. 2009; H. W.

Park et al. 2007; Sohn et al. 2007a; Ying et al. 2010). Time reversal acoustics is a signal processing technique that can provide spatial and temporal wave focusing (Fink 1992). Previous research has shown that time reversal focusing is beneficial in highly scattering environments (Jin and Moura 2009; Moura and Jin 2007; Nunez and Negreira 2005; Prada and Fink 1998). Time reversal techniques have been exploited for Lamb waves in thin plates, leading to an enhanced signal level and compensation for multiple modes and dispersion (Ing and Fink 1998; Park et al. 2007; Prada and Fink 1998; Sohn et al. 2007b; Wang et al. 2004). In this thesis, time reversal was applied to neutralize the multi-modal, multi-path and dispersive effects in a pipe medium. Time-reversal-based change detectors were developed in this study (as presented in Section 4.2); such change detectors were found to be highly sensitive to damage while also sensitive to benign variations (causing false alarms).

Quantifying the change in ultrasonic signals is hardly sufficient for SHM. In practice, significant changes in ultrasonic signals are produced by benign operational and environmental effects. Distinguishing damage-related changes from changes due to harmless factors, such as temperature, pressure and fluid flow, is a major challenge for SHM. Among various benign ambient variations, temperature has been given the most attention. Model-based approaches have been explored to compensate for the temperature effect by taking into consideration the effect of temperature on material properties (Croxford et al. 2007; Goulet et al. 2010; Konstantinidis et al. 2006; Sepehry et al. 2011). However, such techniques are rather sensitive to the structural geometry and composition as well as uncertainties in the environments, and thereby may not provide correct behavior models in field applications. An alternative manner to handle temperature related issues is to apply data-driven methods. Michaels et al. (2004) have applied feature-based classifiers to differentiate structural damage from environmental disturbances in an

aluminum plate. Pattern recognition and machine learning techniques have also been tried with some success for certain structures (Oh and Sohn 2009; Posenato et al. 2008, 2010; Sohn et al. 2002; Worden and Manson 2007) , but few researchers have reported their studies on pipe structures under environmental and operational variations. Furthermore, limited research has been conducted to suppress the effects of more than one benign variable. Although temperature plays a large role in distorting ultrasonic signals, there are many other, possibly unidentifiable, effects that must be taken into account, for instance, internal air pressure or fluid flow conditions.

In this research, machine learning algorithms are investigated and combined with statistical and signal processing method to create robustness to environmental and operational variations. The techniques should allow the system to learn to discriminate between damage and benign variations. Questions that should be addressed are the selection of appropriate features that will ensure maximal robustness, the development of algorithms that will differentiate changes from different sources, and the design and fulfillment of the experiments/field tests to acquire training data as well as to validate the approaches.

Once damage is detected with confidence, it is desirable to locate the damage with reasonable resolution to enable fast response. Existing techniques include phased array beam-forming (Giurgiutiu and Bao 2004; Luo and Rose 2007; Yan et al. 2010) which relies on precise knowledge of structural geometry, material properties and wave velocities. Another technique to locate damage on plates requires multiple sensor pairs to interrogate a given point on the structure (Sohn et al. 2007b). However, this technique requires a large number of sensors for a rather small area on the structure and is especially difficult to implement in cylindrical structures because of the numerous guided wave paths. Several researchers have applied Artificial Neural

Networks and Genetic Algorithm to localize damage based on vibration methods (Hernández-Gómez et al. 2005; Worden and Staszewski 2000), but again, most approaches have merely been implemented for simple structures, and no environmental and operational conditions are taken into account in the course of damage localization.

This research explores a machine learning based damage localization method, in which zones of a pipe are labeled beforehand and the damage localization task is formulated as a multi-class classification problem, each class representing damage in a certain zone. The major challenges lie in extracting and selecting effective features as signatures for identifying the zone of interest within which the damage is located, while suppressing the effects of varying environmental and operational conditions, and in designing suitable experiments/field tests to acquire ample data as well as to validate the approaches.

1.3 Thesis Organization

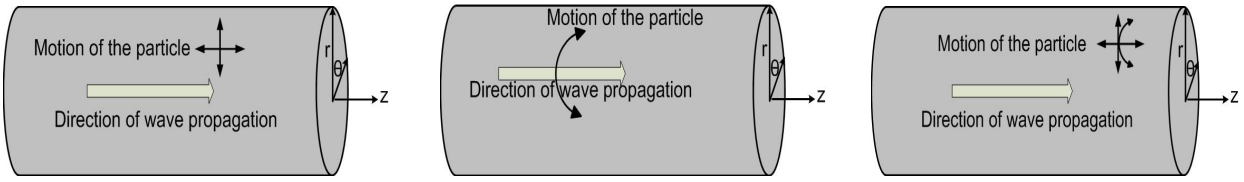
The remainder of the thesis is organized as follows. Chapter 2 describes the characteristics of guided wave propagation in a pipe, serving as a basis for this research. The ultrasonic pipe monitoring procedure is formulated in Chapter 3 as a multi-layer data-driven framework integrating statistics, signal processing and machine learning techniques. This framework is demonstrated through laboratory studies for quantifying a change in ultrasonic signals (discussed in Chapter 4) and identifying the cause of the change (Chapter 5). The data-driven methodology is further validated through field tests presented in Chapter 6. This thesis is concluded in Chapter 7 with remarks, contributions and future research.

2 ULTRASONIC GUIDED WAVES FOR PIPE MONITORING

2.1 Guided Wave Propagation in a Cylindrical Shell

Guided waves are ultrasonic waves (with frequency above 20 kHz) propagating through a medium with geometric boundaries. The boundaries form a waveguide, such as a plate, rod, or pipe. Guided waves are multi-modal and dispersive in nature: multiple wave modes can be generated at different frequencies; dispersion is a characteristic that the velocity of different modes depends on the frequency (Auld 1990; Cawley 2003; Graff 1975; Li and J. L. Rose 2001a).

Guided waves in a pipe or a cylindrical shell contain three fundamental classes of wave modes: an infinite number of longitudinal modes, an infinite number of torsional modes, and a doubly infinite number of flexural modes (see Figure 2.1). Longitudinal modes and torsional modes are axisymmetric, while flexural modes are non-axisymmetric (Li and J. L. Rose 2001a). In addition to these three classes of modes, there also exists a class of circumferential modes in which the waves propagate along the circumference of the cylinder. Two indices M and N are commonly used to denote the fields of the guided wave modes. The index N, named circumferential order, gives the order of symmetry around the axes of the cylinder, so that for axisymmetric modes, N equals zero; the index M, named family order, sorts the modes for a given family. Therefore, longitudinal modes, torsional modes, and flexural modes are often written as L (0, M), T (0, M) and F (N, M) (Seco et al. 2002).



(a) (b) (c)

Figure 2.1 (a) Longitudinal modes; (b) torsional modes (c) flexural modes (Agrawal et al. 2009).

In a cylindrical coordinate system of a pipe (r; θ; z), the response at a given position **x** at a given time t, can be defined as the superposition of all the wave modes excited (O'Donoughue et al. 2009):

$$u(\mathbf{x},t) = \sum_{M=0}^{\infty} a_M(r,\omega) e^{j(\omega t - \beta_M(\omega) z)} + \sum_{N=0}^{\infty}\sum_{M=0}^{\infty} a_M^N(\mathbf{x},\omega) e^{j(\omega t - \beta_M^N(\omega) z)} \tag{2-1}$$

where ω is the angular frequency, a_M^N and β_M^N are the amplitude and wave number due to the Mth mode of the Nth family, respectively, and denoted as a_M and β_M when N = 0, i.e. for axisymmetric modes.

Multiple guided wave modes can be excited at each frequency; meanwhile, the velocity of each mode varies with frequency. The relationship between the wave speed and the frequency is usually represented in the form of dispersion curves, exemplified in Figure 2.2 for a steel pipe with 60.3 mm outside diameter and 3.6 mm wall-thickness, generated by PCdisp (Pochhammer-Chree dispersion), an online package for simulating wave propagation in cylindrical waveguides (Seco et al. 2002). The theoretical dispersion curves provide a clear indication of the multi-modal and dispersive effects of guided waves in a pipe.

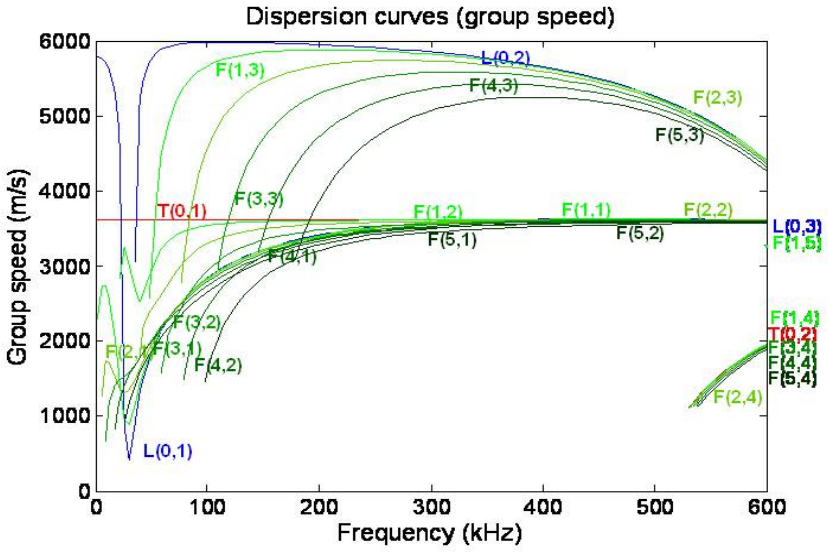


Figure 2.2 Dispersion curves of low-order wave modes in a steel pipe with 60.3 mm outside diameter and 3.6 mm wall thickness, generated by PCdisp.

To visualize the wave propagation in a pipe, numerical simulations were carried out using PZFlex (Weidlinger Associates Inc. Los Altos, CA), a finite element software package designed to analyze the response of continuous media subjected to piezoelectric or mechanical loads. Figure 2.3 shows a steel pipe model, 700 mm long, 120 mm in outside diameter and 6 mm in wall-thickness, with a 12 mm by 11.5 mm piezoelectric patch bonded on the outer surface close to one end of the pipe. The actuator was driven by a single 5-cycle Gaussian windowed sinusoid pulse with 35 μs duration. Figure 2.3 takes a snapshot of the simulated waves when the wave front of the first mode just reaches the other end of the pipe. The guided waves fully illuminate the pipe (both outer and inner surfaces), as a result of the multiple wave modes traveling along multiple longitudinal, circumferential, and helical paths.

The wave propagation theory and numerical simulation show that the guided waves in a pipe are multi-modal, highly-dispersive and multi-path. These effects distort the ultrasonic

signals, leading to difficulty in data interpretation. Nonetheless, these effects are also beneficial in the sense that a small-size PZT wafer is sufficient to excite waves illuminating the entire pipe; the multi-modal complex waveform that is received contains rich information about how different modes are affected by a defect (if present) that might allow one to characterize the defect with great precision. In the following section, physical experiments are described to evaluate the capability of ultrasonic waves for damage detection in pipes.

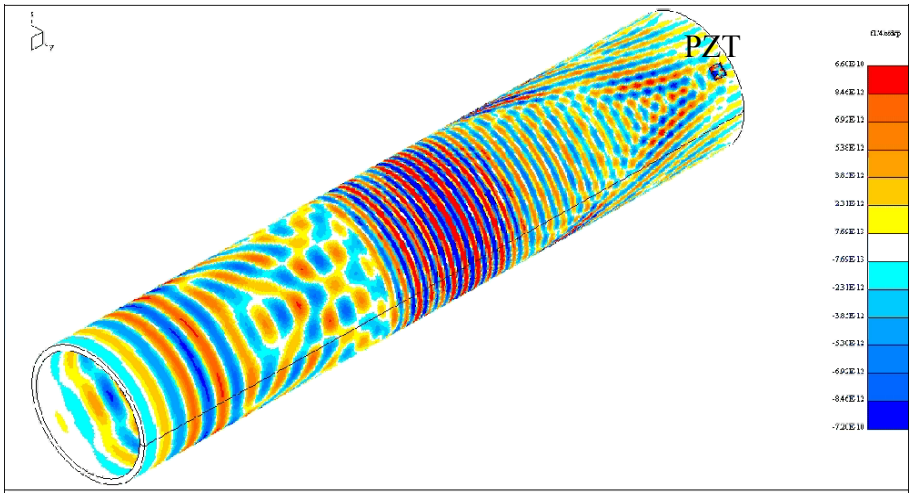


Figure 2.3 Snapshot of wave propagation in a pipe, generated by PZFlex finite element simulation.

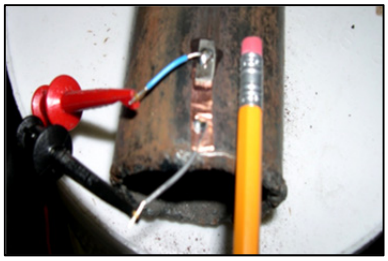
2.2 Evaluation of Ultrasonics for Pipe Monitoring

A common approach to implementing SHM is to collect signals from the current state of a structure and compare them with those previously taken of the pristine condition (baselines). Changes between the current data and the baselines are attributed to damage in the structure (Dutta 2010). In ultrasonic SHM, defects in a structure are essentially changes in the waveguide that produce changes in the wave modes. In principle, defects can be detected and assessed by analysis of wave mode characteristics.

Many researchers have investigated the characteristics of wave modes interacting with damage and have developed various techniques for damage detection (Alleyne and Cawley 1992b; Cegla et al. 2008; Cho 2000; Giurgiutiu 2005; Lee and Staszewski 2003; Su et al. 2006), however most of these techniques have been restricted to thin plate-like structures where a small number of Lamb wave modes can be selected and controlled at relatively low values of the product of frequency and thickness. Subsequently, each mode can be clearly identified according to dispersion curves. However, these mode-analysis-based methods are difficult to apply to pipes, because the guided waves in a hollow cylinder are more complicated than Lamb waves in a thin plate. Consequently, the signal produced by a short excitation pulse at an actuator becomes a long and complex sequence of arrivals at any downstream sensor (Harley et al. 2009; Ying et al. 2010).

As an illustration, a laboratory test was performed on a pipe specimen with length 3050 mm, outside diameter 60.3 mm and wall thickness 3.6 mm. Two PZT wafers (PSI-5A4E, Piezo Systems, Inc., Cambridge, MA) were mounted on the surface, at the two ends of the pipe using cyanoacrylate adhesive. Each wafer was 10 mm long and 5 mm wide (Figure 2.4a). One PZT transducer functioned as a transmitter and the other as a receiver in a pitch-catch configuration. The signal transmission and data acquisition unit was a National Instruments PXI system with an arbitrary waveform generator and a digitizer (Figure 2.4b). The waveform generator drove the PZT transmitter by a 200 kHz sinc windowed sinusoidal signal of 0.1 ms duration (Figure 2.5a). The digitizer was connected to the PZT receiver to record ultrasonic responses, sampled at 1 MHz. Figure 2.5b shows 10,000 voltage records at the receiver, corresponding to a 10 ms time history. The complex pattern renders the wave modes barely distinguishable. On the other hand, the experimental result also verifies that the guided waves can travel for a long range in the pipe;

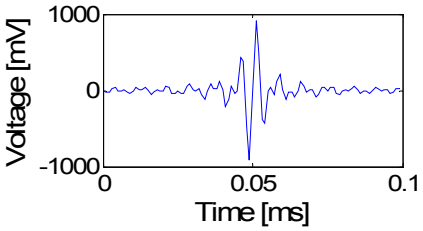
Figure 2.5b shows that the measured voltage remains above the noise level at 10 ms, equivalent to 30-55 m propagating distances for wave modes with velocities ranging from 3000 m/s to 5500 m/s.



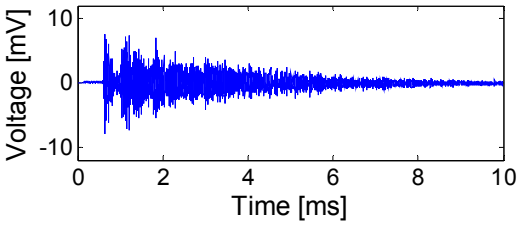
(a)

(b)

Figure 2.4 (a) A surface-mounted PZT wafer (photo courtesy of Joel B. Harley), and (b) an National Instruments PXI system.



(a)



(b)

Figure 2.5 (a) 200 kHz sinc windowed sinusoidal excitation, and (b) received signal, a 10 ms time history.

Conventional NDT approaches for ultrasonic pipe inspection aim to simplify the waveforms by tuning a collar or a comb configuration of multi-element transducers with phase delays, to selectively excite a single wave mode while suppressing unwanted modes (Cawley 2003; Li and J. L. Rose 2001b; Lowe et al. 1998; J. L. Rose et al. 1998). However, this type of

instrument is not feasible for long-term monitoring purposes, because of the requirements for sensor density and operating power, and the cost of modeling, sensing device and its installation. In addition, limiting the number of wave modes and the frequency band also limits the information received, and thus weakens the detectability of damage to certain extent.

The second experiment was conducted on the same pipe specimen, after a partial-thickness circumferential saw-cut was created, 1 mm wide, 1 mm deep and 25 mm in arc dimension, 590 mm from one end of the pipe. Figure 2.6 compares the ultrasonic responses of a pipe before and after the saw-cut was induced. The two received signals appear to be nearly identical. After baseline subtraction, the change produced by the saw-cut can be observed, indicating that the guided waves are sensitive to small damage. However, the residual signal is weak and may easily be lost in the noise, if the noise level is high in the pipe system. It should be noted that the many guided wave modes in the cylinder illuminate all downstream locations over a wide range of helical incidence angles, and therefore flaws of any orientation will scatter ultrasonic energy. Ying et al. (2010) have shown that small cuts oriented longitudinally are easily detected, as are small areas of diminished (corrosion-like) thickness.

A third experiment evaluates the behavior of guided waves under different benign environmental variations. Figure 2.7 shows the significant difference in two ultrasonic signals collected from the same undamaged pipe specimen but on two different days when the temperature, sensor bonding conditions, and unknown effects might varied. Environmental and operational variability tends to distort signals and masquerades as damage, preventing many SHM techniques moving from laboratory to field applications.

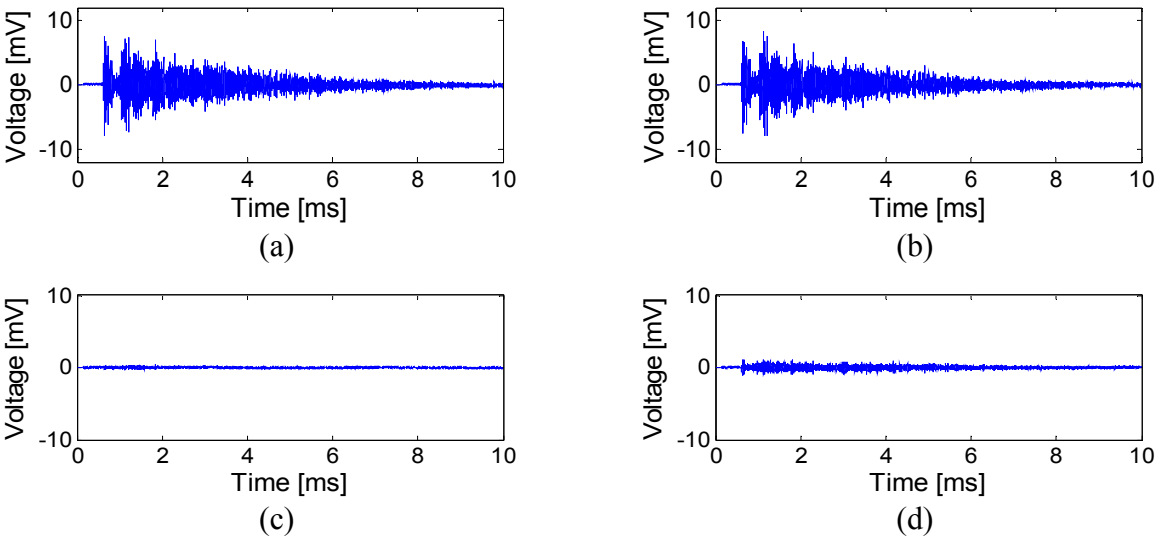


Figure 2.6 (a): Received signal before a saw-cut is induced, (b): received signal after a saw-cut is induced, (c): difference of signal (a) from baseline subtraction, and (d): difference of signal (b) from baseline subtraction (Ying et al. 2010).

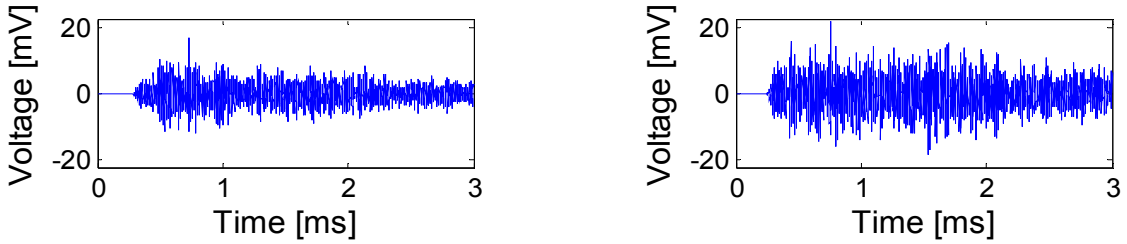


Figure 2.7 Two signal responses from an undamaged pipe on two different days.

The advantages and disadvantages of using ultrasonics for pipe monitoring are summarized in Table 2-1. The three difficulties correspond to the three challenges identified in Section 1.2. It should also be noted that these three problems are associated with each other in a pipe monitoring implementation (Figure 2.8). In addition to the subtle changes produced by damage, the ambient conditions create significant changes that obscure real structural damage or be mistaken as damage. Hence, it is inevitable to carefully analyze the complicated waveforms

resulting from the cylindrical geometry of a pipe. Conversely, the complexity in the signals renders it even more difficult to distinguish the causes of the changes. That is said, an effective and robust methodology must be designed to extract useful information of damage (the presence of damage) from an insignificant change embedded in a complicated wave pattern with considerable interference from benign variations.

Table 2-1 Advantages and disadvantages of guided waves for pipe monitoring.

Pros	Cons
• Long-distance propagation • Full illumination • Sensitivity to small damage • Easy generation using a PZT wafer • Complex wave pattern (information aspect)	• Complex wave pattern (interpretation aspect) • Subtle change produced by damage • Significant change produced by benign environmental and operational variations

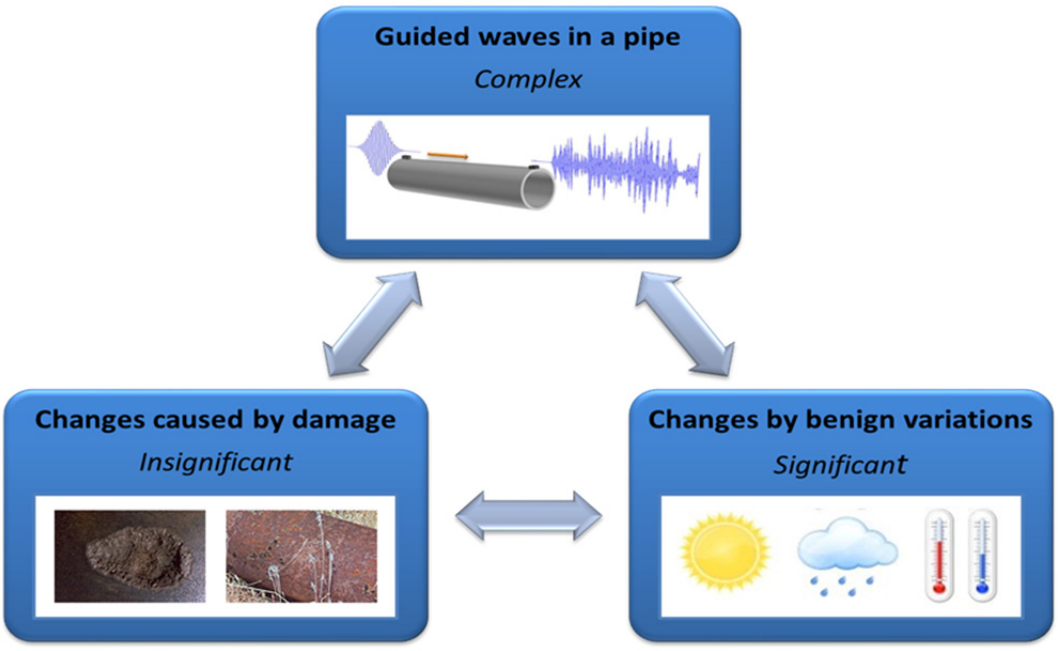


Figure 2.8 Three challenging issues in ultrasonic SHM implementation for pipes.

3 A DATA-DRIVEN PIPE MONITORING FRAMEWORK

To establish a reliable and practical ultrasonic system for continuous pipe monitoring with mounted PZT wafers, this research focuses on developing effective data interpretation approaches, and formulates a computer-assisted SHM methodology based on data-driven analysis of sensed ultrasonic signals. This chapter presents the data-driven SHM framework, bridging between SHM practices and integrated statistical, signal processing and machine learning algorithms.

3.1 Structural Health Monitoring Paradigm

The data-driven SHM framework is composed of a three-stage SHM schema, supported by a multi-layer data processing architecture, as depicted in Figure 3.1. Similar to the SHM statistical pattern recognition paradigm summarized by Farrar et al. (1999), the pipe monitoring process is organized into the following three stages.

- **Damage detection**: to recognize the presence of damage in a pipe. This stage is executed in two steps. The first step determines whether there is a change in the signal other than random electrical noise (Ying et al. 2010). Second, the source of the change, if one exists, is distinguished as structural damage or benign variation (a false alarm).
- **Damage localization**: to identify the location of the damage in the pipe. In this work, determining the presence of a defect in a particular pipe segment is of interest. The resolution of localization can be adjusted by dividing (virtually) the pipe into a different number of zones and assign different labels to each zone.

- **Damage characterization**: to classify the type and the severity of the damage in the pipe. Common types of defects and anomalies in pipes include areas of corrosion, cracks, dents and gouges (e-CFR 2011; Nestleroth and Bubenik 1999; PHMSA 2011c). The severity is evaluated by, for instance, the size of a corrosion area, the depth of a crack, etc.

The first two stages of the SHM schema correspond to the research questions addressed in Section 1.2, i.e., change quantification and detection, change identification (damage detection), and damage localization. Damage characterization, prognosis and prevention are not addressed in the research described in this dissertation.

3.2 Multilayer Data Processing Architecture

The multilayer data processing architecture is devised to uncover damage-related information and to support decision making in the course of pipe monitoring. The data processing schema consists of six modules:

- **Data acquisition**: to acquire ample signals from sensors installed on the pipe under a wide range of environmental and operational conditions. The sensors include PZT transducers and any other available sensing devices (e.g., thermocouples, flow meters, and pressure gauges).
- **Preprocessing**: to improve the signal quality through fundamental signal processing, such as averaging (to enhance signal-to-noise ratio), filtering (to eliminate outbound noise), and energy-normalizing (to reduce signal output variability).
- **Feature extraction**: to extract potentially good features from the ultrasonic signals. A feature is a single-valued quantitative characterization of a signal measurement (a

time-history record in this study, e.g., the 10 ms signal shown in Figure 2.5b), and can be constructed using various statistical, signal processing, and machine learning techniques. A change in the ultrasonic response obtained from a pipe under two different conditions can be represented in terms of the discrepancy of the feature computed under the two different conditions. Therefore, a feature is a change detector. A good feature is sensitive to damage, while unaffected by benign variations.

This feature extraction layer can be implemented in two steps. In the first step, a few features that do not require great computational effort are computed in real-time, and are regarded as metrics of change detection to monitor the pipe integrity continuously. In the second step, once any change is detected, a large number of additional features are extracted from the received signals for thorough examination in the higher-level modules described below. A considerable library of features (described in Section 5.4) has been created for use in this framework in order to be adapted to different scenarios, since it is extremely hard to develop a "super feature" that is sensitive to all types of damage, unaffected by all environmental effects, and effective in all of the three SHM stages.

- **Feature selection**: to rank and select effective features from all the feature candidates generated in the preceding feature extraction module. Various statistical and machine learning methods are used to conduct the selection process automatically. The output is a multi-feature vector with more discriminatory and less redundant information. The feature selection module can also feed information back to the feature extraction layer, by recommending good features as change detectors and by removing apparently inefficient features from the feature library. Feature selection is essential

to improve the computational efficiency and the performance of learning algorithms. In addition, the features marked as effective/ineffective by the machine learning algorithms may provide physical insights into the sensor-structure system.

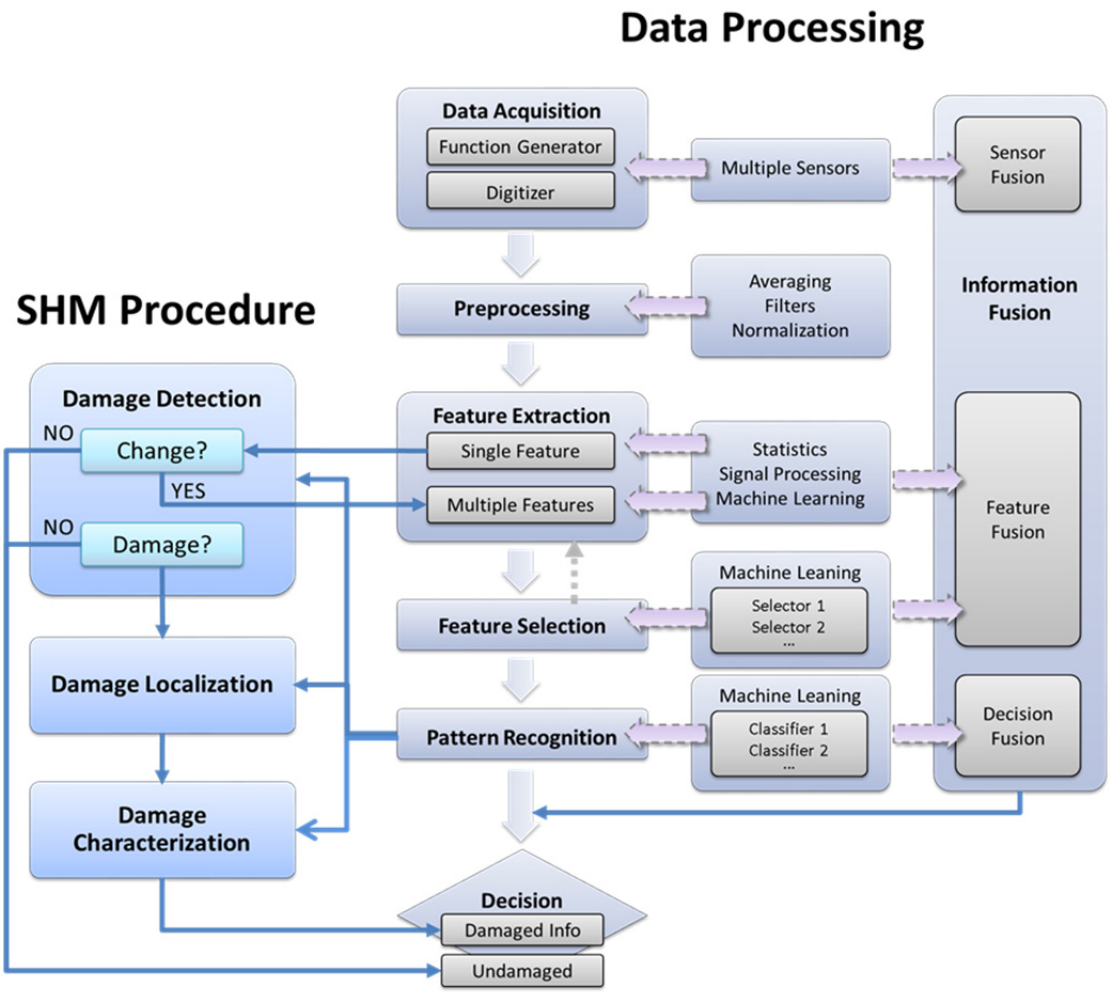


Figure 3.1 A data-driven structural health monitoring framework.

- **Pattern recognition**: to classify datasets into different categories based on the objective of a particular SHM stage: determining whether the pipe is damaged or

undamaged for damage detection; determining a specific zone of interest for damage localization; and determining a type and a certain degree of severity for damage characterization.

- **Information fusion**: to fuse information learned from diverse sources so as to achieve more comprehensive, unified or even new knowledge and to improve decision making. Information fusion techniques are operated at three levels: the sensor level, the feature level, and the decision level (Faundez-Zanuy 2005; Guo et al. 2008; Ross and Jain 2003).
  - **Sensor fusion** relies on deployment of multiple PZTs on the pipe. Distributed sensors can capture different characteristics of guided waves propagating along multiple channels of the pipe and interacting with damage at different angles. Multiple sensors can also allow for adaptive probing and learning about the pipe medium to constantly optimize the understanding of the damage response without necessarily making hard decisions after each transmission (Jin et al. 2010). In addition, wrong decisions made due to a malfunctioning sensor can be largely eliminated. Moreover, data recorded from other sensors, such as thermocouples, pressure gauges and flow meters, can also be utilized along with the ultrasonic signals collected by the PZTs.
  - **Feature fusion** refers to integrating different feature sets. For instance, one feature set is extracted from time domain signals, and another is extracted from frequency domain signals. Instead of drawing individual decisions using each set of features, multiple sets of features are integrated to produce a fused-feature vector. Two fusion methods can be considered. One is based on feature selection

by first grouping multiple feature sets into one feature pool, and then applying appropriate selection algorithms to form a feature vector of selected good features. The second method is based on feature combination by first combining multiple feature sets into one union vector (serial feature fusion) or one complex vector (parallel feature fusion) and then feeding that vector into a feature extractor that is usually a projection method, such as principal component analysis transforming the feature vector into a new vector of uncorrelated variables with reduced dimensionality (Sun et al. 2005; Yang et al. 2003).

- **Decision fusion** is to reconcile and merge redundant or complementary results from multiple detectors and classifiers into a final global decision (Dasarathy 1994; Jeon and Landgrebe 1999).

The work described in this thesis contributes to most of the data-driven modules, except the sensor fusion and the decision fusion modules, which are not elaborated in this document. The engineering design, deployment, and maintenance of sensing systems were not addressed by the research described in this dissertation.

4 STATISTICAL ANALYSIS AND SIGNAL PROCESSING FOR CHANGE DETECTION

The studies presented in this chapter attempted to generate change detectors for pipe monitoring using easily-implemented statistical analysis and signal processing methods. Section 4.1 provides several fundamental statistical measures. Section 4.2 describes time reversal signal processing technique that can focus wave energy and potentially improve the performance of change detection. Section 4.3 presents an evaluation of the developed change detectors in terms of damage sensitivity and environmental robustness. Alterative change detectors are then discussed in Section 4.4.

4.1 Fundamental Statistical Features

This section presents some basic and straightforward statistical features that can be used to depict certain characteristics of an ultrasonic signal collected from a pipe. A change in the feature value corresponds to a change in the ultrasonic response.

Mean and variance. Mean and variance are the most fundamental, but important, statistical measures of the center and the spread of the data, respectively. A discrete signal $x[i]$ is an *n*-sample array, whose sample mean and variance are given by,

$$\mu = \frac{1}{n}\sum_{i=1}^{n}\left(x[i]\right) \tag{4-1}$$

$$Var = \frac{1}{n-1}\sum_{i=1}^{n}\left(x[i]-\mu\right)^{2} \tag{4-2}$$

Kurtosis. Kurtosis is generally used as a descriptor of the shape or peakedness of a distribution. Any mode conversion may change the shape of the peaks. Kurtosis of a discrete signal domain $x[i]$ is defined as the standardized fourth central moment (Balanda and MacGillivray 1988):

$$K = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(x[i]-\mu\right)^4}{\left\{\frac{1}{n}\sum_{i=1}^{n}\left(x[i]-\mu\right)^2\right\}^2} \tag{4-3}$$

Correlation coefficient. The Pearson product-moment correlation coefficient (PCC) is a common metric for similarity between two signals. The PCC of a discrete signal domain $x[i]$ and a baseline domain $x_b[i]$ (a baseline is a reference measurement of the undamaged or pristine state of a structure; a baseline domain is a signal domain, such as a time domain or a frequency domain, of the baseline measurement) is obtained by,

$$PCC = \frac{\sum_{i=1}^{n}\left(x[i]-\mu\right)\left(x_b[i]-\mu_b\right)}{\sqrt{\sum_{i=1}^{n}\left(x[i]-\mu\right)^2}\sqrt{\sum_{i=1}^{n}\left(x_b[i]-\mu_b\right)^2}} \tag{4-4}$$

Peak level. Peak level is the maximum amplitude of the waveform, particularly, a wave mode (or a superposition of several modes) with the highest strength. The change in the peak infers the arrival, reflection, or conversion of wave modes.

$$peak = \max\left(\left|x[i]\right|\right) \tag{4-5}$$

As a demonstration, the above five features were computed for the two signals shown in Figure 2.6a and b, obtained in a 3050 mm long pipe before and after a 25 mm long partial-thickness saw-cut was induced (see computed feature values in Table 4-1). To quantify how much one feature can distinguish the damaged signal from the undamaged one, Equation (4-6) defines a change index due to damage as the difference of the feature between the two signals (measuring the change produced by damage), normalized by the difference between the undamaged signal and a baseline signal (measuring the change produced by random electrical noise during data collection):

$$CI_d = \left| \frac{F_d - F_u}{F_u - F_b} \right| \tag{4-6}$$

The change index has to be greater than one to indicate the discrepancy between a physical change and random noise. The larger the index, the larger the change caused by damage.

Table 4-1 Comparison of time-domain features between two signals before and after a saw-cut is created.

Feature	Baseline	Undamaged	Damaged	CI_d
Mean	1.0435×10^{-7} $[V]$	9.9082×10^{-8} $[V]$	9.5916×10^{-8} $[V]$	0.6
Variance	1.1262×10^{-6} $[V^2]$	1.1277×10^{-6} $[V^2]$	1.1337×10^{-6} $[V^2]$	4.09
Kurtosis	9.53	9.5391	9.8786	37.85
Corr. coef.	1	0.9986	0.9791	13.38
Peak level	0.0075 $[V]$	0.0076 $[V]$	0.0083 $[V]$	5.2

This simple example shows the basic idea of how to use a feature to quantify a change in the time domain signals, and how to compare the performance of those different features. However, as discussed in Chapter 2, the time domain signals are complex due to multi-modal and dispersive effects. Therefore, the same set of features extracted from other signal domains may better characterize the damage. That is to say, combining statistical methods with certain signal processing techniques may lead to improved features. The following section introduces the application of time reversal signal processing to generate features with increased discriminant power.

4.2 Time Reversal Signal Processing Features

Time reversal is a signal processing technique used to provide spatial and temporal wave focusing (Fink 1992). Time reversal methods have been exploited for Lamb waves in thin plates, leading to an enhanced signal level and compensation for multiple modes and dispersion (Ing and Fink 1998; Park et al. 2007; Prada and Fink 1998; Sohn et al. 2007b). Previous research has also shown that time reversal is applicable and beneficial in highly scattering environments (Jin and Moura 2009; Moura and Jin 2007; Nunez and Negreira 2005; Prada and Fink 1998). Therefore, by taking advantage of multiple modes and dispersion of guided waves in pipes, the time reversal technique was applied for ultrasonic wave focusing and resulted in the Time Reversal Change Focusing technique for damage detection in pipes (Harley et al. 2009; O'Donoughue et al. 2009; Ying et al. 2010).

In this section, a mathematical description of the time reversal process is first provided, followed by a description of laboratory tests that were conducted to evaluate the capabilities of time reversal to compensate for wave dispersions and to detect changes in pipes.

4.2.1 Time Reversal Focusing

The time reversal processing assumes that the waveguide is reciprocal and stationary (Fink 1992). Reciprocity ensures the same signals are received when the waves propagate backward; stationary means the signals do not vary over time. Consider a pitch-catch measurement, using two in-line transducers attached on the surface of a pipe, one as a transmitter and the other as a receiver. A received response $R(\omega)$ at the receiver due to an input excitation $S(\omega)$ at the transmitter can be expressed in the frequency domain through a pipe transfer function $H(\omega)$ (Prada et al. 1996):

$$R(\omega) = S(\omega) H(\omega) \tag{4-7}$$

The received signal is time-reversed and energy-normalized to form a time-reversal probing signal. Time reversal in the frequency domain is equivalent to the negation of angular frequency ω, thus the probing signal becomes:

$$S^{TR}(\omega) = kR(-\omega) = kS(-\omega) H(-\omega) \tag{4-8}$$

where k is the energy normalization term, ensuring the time-reversed signal retains the same amount of energy as the original input excitation, expressed as:

$$k = \sqrt{\frac{\int_{-\infty}^{\infty} |S(\omega)|^2 \, d\omega}{\int_{-\infty}^{\infty} |R(\omega)|^2 \, d\omega}}. \tag{4-9}$$

As the new probing signal is transmitted back through the pipe, the received signal becomes:

$$\begin{aligned} Y(\omega) &= S^{TR}(\omega) H(\omega) \\ &= kS(-\omega) H(-\omega) H(\omega) \end{aligned} \quad (4\text{-}10)$$

We note that the signal S(-ω) has a real-valued Fourier Transform, thus S(-ω) = S*(ω), and assume that phase velocity is even symmetric with respect to ω. Therefore, negating the angular frequency in H(ω)is equivalent to taking its complex conjugate (Harley et al. 2009; O'Donoughue et al. 2009):

$$H(-\omega) = H^{*}(\omega) \quad (4\text{-}11)$$

Therefore, Equation (4-10) can be rewritten as:

$$\begin{aligned} Y(\omega) &= kS^{*}(\omega) H^{*}(\omega) H(\omega) \\ &= kS^{*}(\omega) \left| H(\omega) \right|^{2} \end{aligned} \quad (4\text{-}12)$$

Equation (4-12) shows that the received signal has the same phase-profile as the conjugate or time-reversed excitation signal. This is an indication of a large reduction in dispersion and multi-modal effects after time reversal is applied. The guided wave modes focus at the original source at a single point in time, leading to an enhanced signal level and compensation for multiple modes and dispersion in pipe environment. As a result, the time domain signal is expected to be symmetric with a large peak at the center (Harley et al. 2009; Ying et al. 2010).

4.2.2 Time Reversal Change Focusing

A change in the ultrasonic response after baseline subtraction is often subtle and exhibits unapparent patterns due to the dispersive pipe environment (for example, as shown in Figure

2.6c and d). Time Reversal Change Focusing (TRCF) is developed by applying time reversal to the change, so that the change becomes focused, presenting as a significant peak in the time reversal focused signal (Harley et al. 2009; Ying et al. 2010).

Assume that the response of the pipe before any damage occurs is measurable and stationary over time, thus the baseline response is given by Equation (4-13) (Harley et al. 2009; Ying et al. 2010):

$$R_b(\omega) = S(\omega) H_b(\omega). \tag{4-13}$$

where $S(\omega)$ is the known excitation, and $R_b(\omega)$ can be measured at the PZT receiver. Therefore, the baseline transfer function $H_b(\omega)$ can be obtained from Equation (4-13).

When damage is induced, the received signal becomes:

$$R(\omega) = S(\omega) H(\omega) = S(\omega)\left(H_b(\omega) + H_d(\omega)\right). \tag{4-14}$$

where $H_d(\omega)$ is the transfer function of the change due to damage in the pipe.

The change is the difference between the received signal and the baseline response, therefore:

$$\begin{aligned} R_d(\omega) &= R(\omega) - R_b(\omega) \\ &= S(\omega) H_d(\omega) \end{aligned}. \tag{4-15}$$

The change is then time-reversed and energy-normalized to form new excitation for sending backwards, expressed as:

$$\begin{aligned} S_d^{TR}(\omega) &= kR_d(-\omega) \\ &= kS(-\omega)H_d(-\omega) \\ &= kS^*(\omega)H_d^*(\omega) \end{aligned} \tag{4-16}$$

where k is the energy normalization term given by Equation (4-9).

The new input signal is transmitted back in the pipe, and the received response is obtained:

$$\begin{aligned} Y(\omega) &= S_d^{TR}(\omega)H(\omega) \\ &= kS^*(\omega)H_d^*(\omega)\left(H_b(\omega)+H_d(\omega)\right). \\ &= kR_d^*(\omega)H_b(\omega)+kS^*(\omega)\left|H_d(\omega)\right|^2 \end{aligned} \tag{4-17}$$

After the time reversal process, this received signal consists of two components: the response to the baseline, and the response to the change. The baseline response can be determined, since $H_b(\omega)$ is known from Equation (4-13) and $R_d(\omega)$ is obtained from Equation (4-15). Therefore, we subtract the baseline component from Equation (4-17) and obtain the equation of the change in the pipe as:

$$\begin{aligned} Y_d(\omega) &= Y(\omega)-kR_d^*(\omega)H_b(\omega) \\ &= kS^*(\omega)\left|H_d(\omega)\right|^2 \end{aligned} \tag{4-18}$$

Equation (4-18) is similar to Equation (4-12), but the focused signal is now caused by the change in the pipe. A large peak is expected at the center of the time domain signal as an indication of the existence of damage in the pipe. Therefore, the TRCF-domain is likely to produce better features for change detection in a pipe (Ying et al. 2010).

4.2.3 Experimental Results of Time Reversal for Change Detection

To demonstrate the time reversal methods experimentally, pitch-catch experiments were conducted on steel pipe specimens considering three scenarios (a baseline pipe, a pipe with a welded butt joint, and a pipe with internal air pressure), with the aim of detecting three types of defects (a lateral defect, a longitudinal defect and a corrosion-like defect). A comprehensive description of the experimental setup and results can be found in Ying et al. (2010). This section discusses typical experimental results to demonstrate that time reversal can effectively compensate for multiple modes and dispersion in a pipe, resulting in focused wave energy and high damage detectability. The scheme of the experiments is shown in Figure 4.1.

First, time reversal focusing was applied to a received signal under the same experimental setup described in Section 2.2. The received response appears to contain a large number of guided wave modes that are very complex and difficult to interpret (Figure 4.2a). However, after performing the time reversal focusing process, different wave modes are compressed, presenting as a large peak at the center of the plot shown in Figure 4.2b. This peak is formed as a consequence of focusing wave modes as explained by Equation (4-12). The peak level of the received signal is increased from 7.5 mV to 55.4 mv, a difference of 7.4 times or 17.4 dB.

Second, the TRCF technique was applied to a change produced by the shallow partial-thickness circumferential saw-cut (width 1 mm, depth 1 mm, and arc length 25 mm). The difference signal caused by the physical saw-cut scattering reverses to produce a significant peak easily recognized in Figure 4.3f, while the random noise produces almost no visible peak (Figure 4.3e). As a quantitative evaluation, the peak levels of the TRCF-domains of the undamaged and damaged instances are 1.1 mV and 10.4 mV, respectively, a difference of 9.1 times or 19.2 dB.

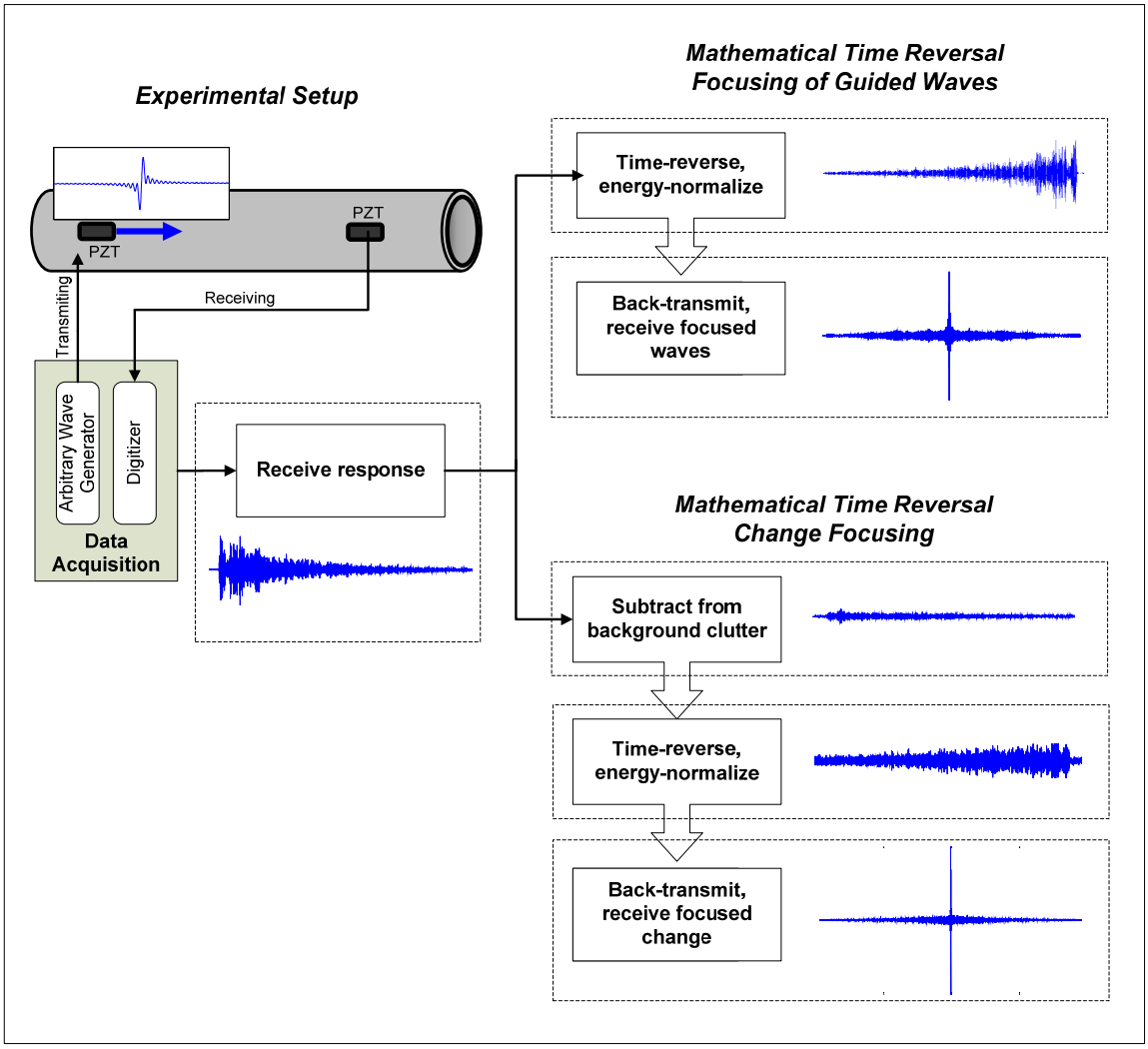
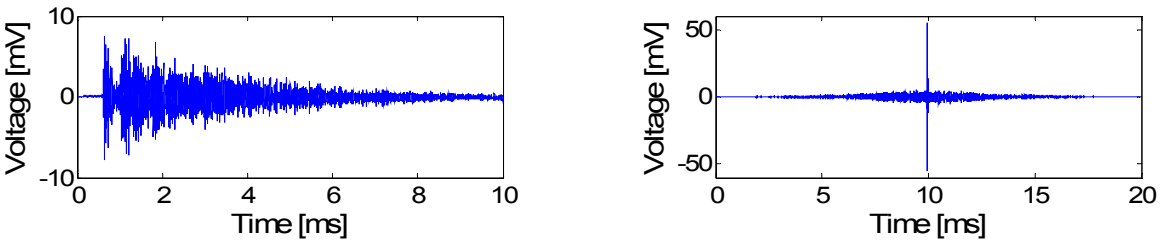


Figure 4.1 Scheme of the time reversal experiments.



(a) (b)

Figure 4.2 (a) Received signal, and (b) signal obtained after time reversal focusing

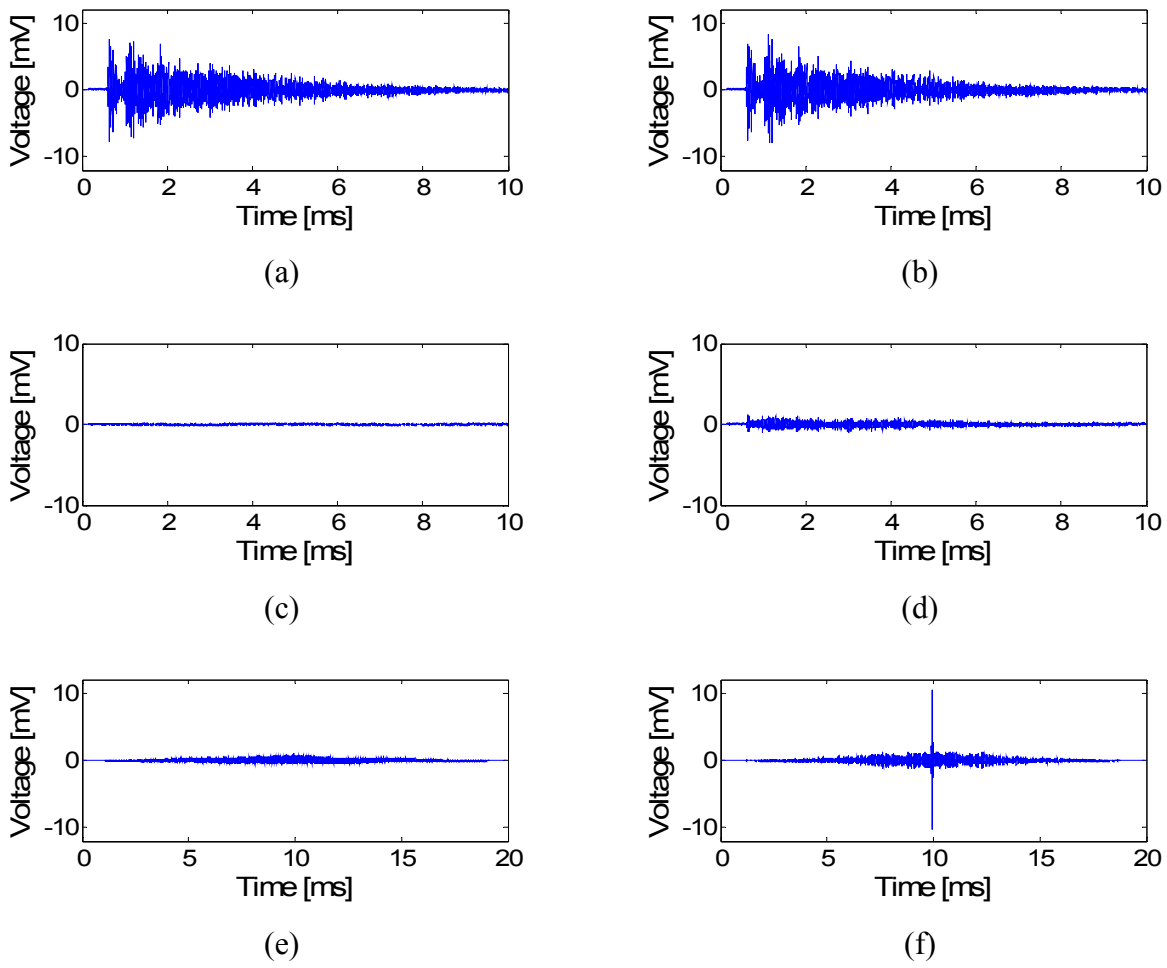


Figure 4.3 (a): Received signal before a saw-cut is induced, (b): received signal after a saw-cut is induced, (c): difference of signal (a) from baseline subtraction, (d): difference of signal (b) from baseline subtraction, (e): signal obtained after applying time reversal to (c), and (f): signal obtained after applying time reversal to (d) (Ying et al. 2010).

Furthermore, the peak level of the TRCF-domain has been shown to increase in direct proportion with the severity of the physical damage, which was reported in Harley et al. (2010). Hence, the peak level feature extracted from the TRCF-domain exceeds the performance of that from the time history. In addition to the peak level, other statistical features described in Section

4.1 are computed in the TRCF, and the change indices of those TRCF features are compared with those of the time-domain features (Figure 4.4). The results show that the variance, kurtosis and peak level features from the TRCF-domain greatly outperform those from the time domain. In other words, applying TRCF can generally enhance the feature discriminant power in distinguishing a change produced by a defect.

Table 4-2 Comparison of TRCF-domain features between two signals before and after a saw-cut is created.

Feature	Baseline	Undamaged	Damaged	CI_d
Mean	8.7861×10^{-10} $[V]$	1.5104×10^{-9} $[V]$	5.3378×10^{-10} $[V]$	0.41
Variance	4.8161×10^{-8} $[V^2]$	4.7167×10^{-8} $[V^2]$	1.2143×10^{-7} $[V^2]$	74.73
Kurtosis	4.2719	3.3764	136.1638	148.27
Corr. coef.	1	0.8877	0.1512	6.56
Peak level	0.0014 $[V]$	0.0011 $[V]$	0.0104 $[V]$	34.78

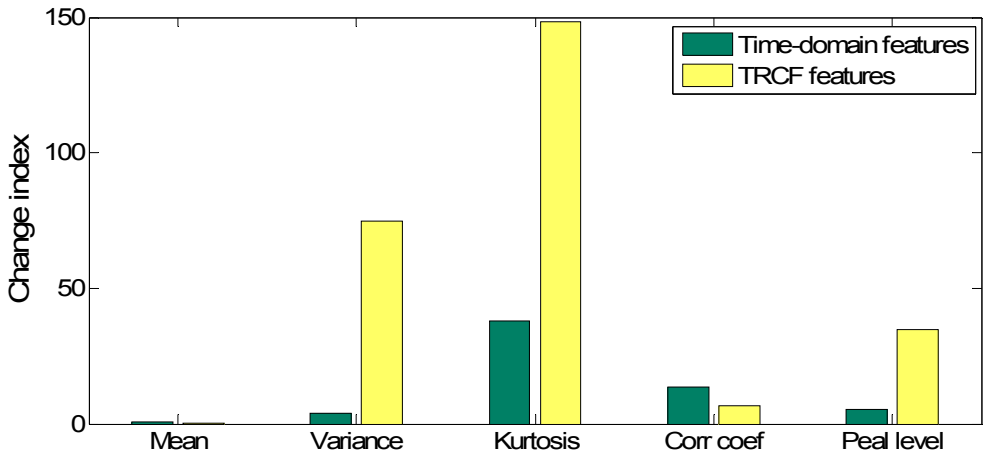


Figure 4.4 Comparison of features extracted from the time domain and the TRCF-domain for damage detection.

4.3 Evaluation of Change Detectors

To evaluate a change detector, two criteria must be taken into account: sensitivity to damage and robustness to environmental and operational variations. First, a change detector must reflect any changes caused by even a small defect in a pipe. Second, the change detector should produce zero or few false alarms caused by ambient environmental interferences; although it is impractical to find one "super" feature insensitive to all varieties of benign variations, one change detector less unaffected by such variations is favorable, and the cause of the change can be further identified using machine learning. The preceding sections focus on the first criterion when examining the change detectors. This section evaluates the robustness of the time-domain and TRCF features against benign variations through an experiment performed on a pipe with varying internal air pressure.

The pipe specimen, with length 1833 mm, outer diameter 70 mm and wall thickness 4 mm, had a welded cap and a pressure gauge at each end to control the internal pressure. Two PZT wafers, each 20 mm long and 8 mm wide, were mounted 1200 mm apart on the pipe. A sinc pulse centered at 200 kHz was chosen as the input signal. Similar to the results of detecting a saw-cut (Figure 4.3), Figure 4.5 shows that TRCF also produces a distinguishable peak when applied to the response difference caused by pressurizing the pipe by 83 pounds per square inch (PSI). The time reversal peak value also measures benign variations other than damage, increasing from 2 mV to 64.8 mV, a 58.8 dB difference. Admittedly, this characteristic can be useful in certain engineering control circumstances where stable internal pressure is required to be monitored. However, for defect detection, it leads to false alarms.

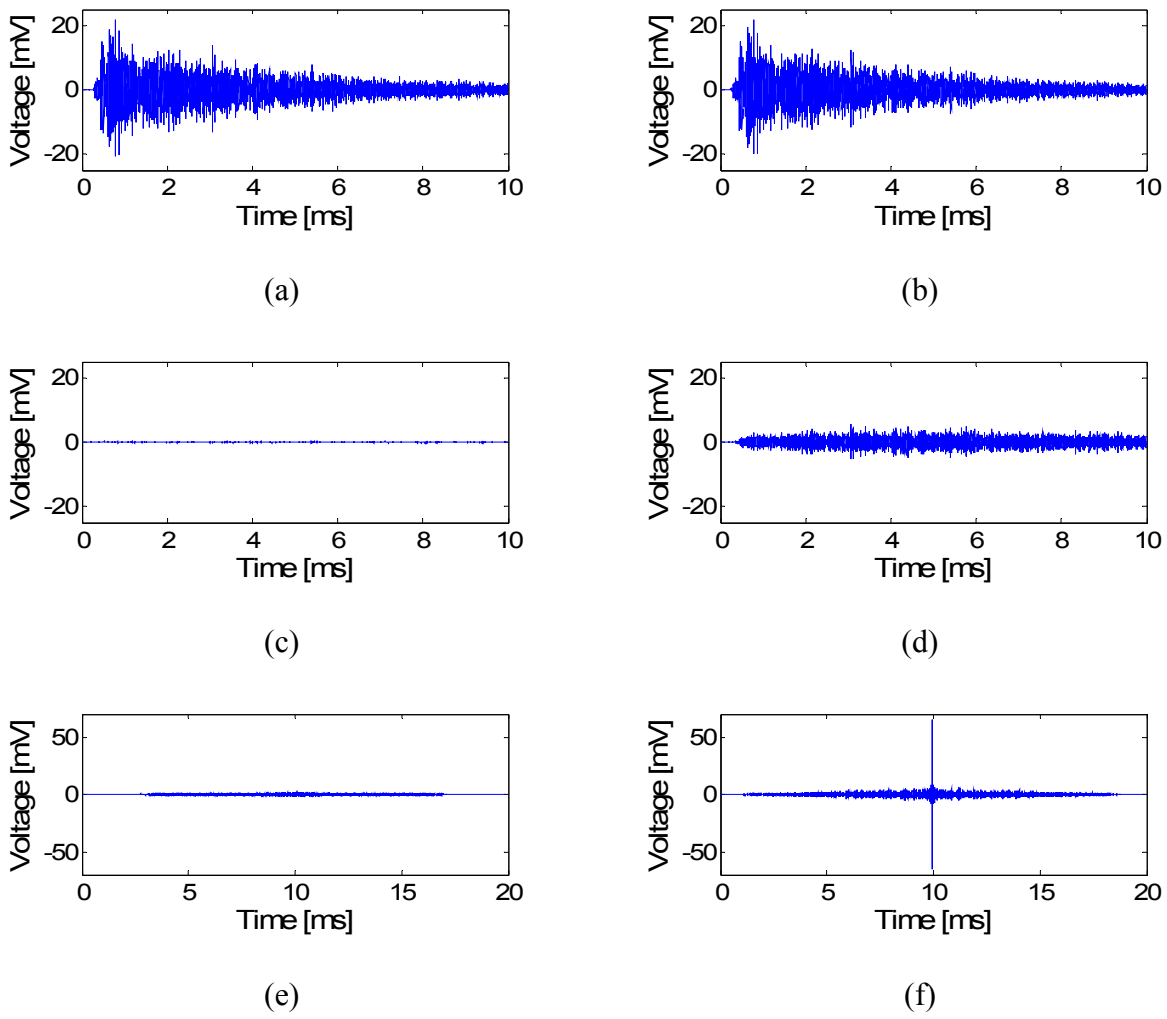


Figure 4.5 (a): Received signal under 0 PSI, (b): received signal under 83 PSI, (c): difference of signal (a) from baseline subtraction, (d): difference of signal (b) from baseline subtraction, (e): signal obtained after applying time reversal to (c), and (f): signal obtained after applying time reversal to (d).

A change index due to pressure variations is defined in the same form as the change index due to damage given by Equation (4-6), to quantify how much one feature can digress in the presence of a benign variation:

$$CI_p = \left| \frac{F_p - F_u}{F_u - F_b} \right| \tag{4-19}$$

The smaller the index is, the more robust a feature is.

The change indices are computed for both the time-domain and TRCF-domain features (as shown in Table 4-3, Table 4-4, and Figure 4.6). The results show that most of the features are sensitive to benign variations. These features, although effective to detect and quantify a change in the ultrasonic response, are limited in their applicability for robust pipe monitoring, and therefore machine learning approaches were investigated in this research to further distinguish an actual defect from a false alarm.

Table 4-3 Comparison of time-domain features between two signals before and after the pipe is pressurized.

Feature	Baseline	Undamaged	Pressurized	CI_p
Mean	1.9437×10^{-9} $[V]$	2.2803×10^{-9} $[V]$	6.7917×10^{-8} $[V]$	16.62
Variance	9.0216×10^{-6} $[V^2]$	9.0047×10^{-6} $[V^2]$	8.5481×10^{-6} $[V^2]$	2
Kurtosis	7.374	7.372	7.6484	27.13
Corr. coef.	1	0.9997	0.9146	249.49
Peak level	0.02179$[V]$	0.02175 $[V]$	0.02184 $[V]$	2.35

Table 4-4 Comparison of TRCF-domain features between two signals before and after the pipe is pressurized.

Feature	Baseline	Undamaged	Pressurized	CI_p
Mean	1.1198×10^{-10} $[V]$	4.8275×10^{-11} $[V]$	4.6377×10^{-11} $[V]$	1.49
Variance	1.0982×10^{-7} $[V^2]$	2.7798×10^{-7} $[V^2]$	2.231×10^{-6} $[V^2]$	11.6
Kurtosis	3.6099	3.6931	495.52	5908.6
Corr. coef.	1	-0.2050	0.012	0.27
Peak level	$0.0012[V]$	0.002 $[V]$	0.0648 $[V]$	83.04

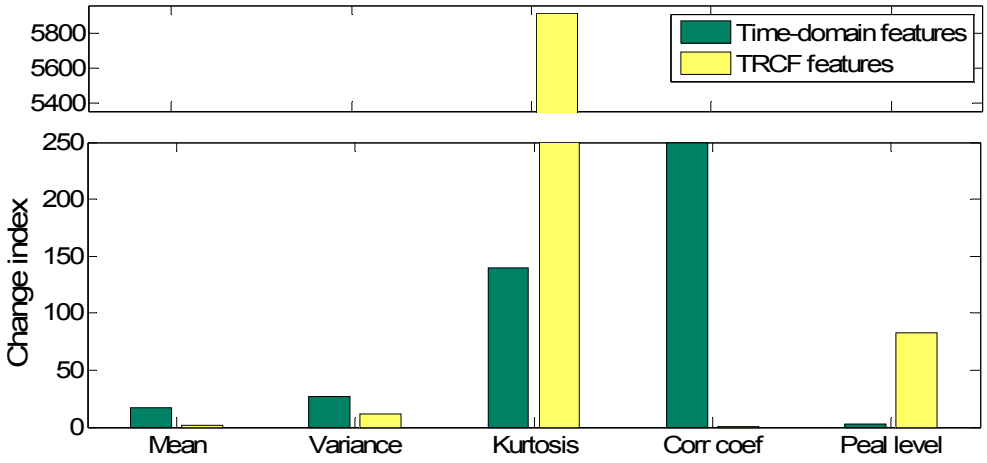


Figure 4.6 Comparison of features extracted from the time domain and the TRCF-domain under varying internal air pressure.

4.4 Alternative Change Detectors based on Learning Process

The preceding studies have shown that combining statistical analysis with appropriate signal processing can produce features with enhanced discriminant power for change detection, while an individual feature is usually inadequate to achieve robustness against varying environmental effects. As a result, this research endeavors to incorporate various algorithms to generate a sufficient number of features, with the aim of achieving a generalizable methodology for robust pipe monitoring.

In addition to time reversal techniques, many other signal processing methods, statistical approaches and machine learning algorithms can be integrated to construct thousands of features. The challenge is to choose one or a few of the most suitable features as the change detectors for online pipe monitoring. The feature selection module of the data-driven monitoring framework described in Chapter 3 outputs a multi-feature vector able to distinctively separate different classes, undamaged and damaged, or a particular damage zone, in spite of the existence of environmental and operational variation. Hence, these features or a combination of the features have higher probability to provide effective and reliable change detection. Such change detectors learned through the machine learning process are explored in laboratory studies described in Chapter 5 and field tests in reported in Chapter 6.

4.5 Summary

Statistical measures are employed to quantify changes in the ultrasonic signals collected in pipes. The change detectability can be improved by incorporating signal processing techniques. In particular, time reversal methods are discussed in this chapter. Time reversal acoustics has been shown to be a useful tool for revealing damage information from complex wave patterns,

due to its compensation for multiple modes and dispersion, and enhancement of signal-to-noise ratio. Time Reversal Change Focusing has experimentally demonstrated its effectiveness for change detection, but the applicability for monitoring is limited because it is prone to false positives caused by changing environmental and operational conditions. As a result, machine learning algorithms are expected to provide robust monitoring by differentiating damage from changes caused by benign variations.

5 MACHINE LEARNING FOR SCATTERER DETECTION AND LOCALIZATION UNDER ENVIRONMENTAL AND OPERATIONAL VARIATIONS

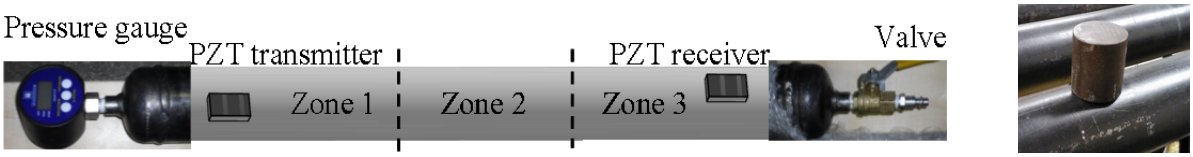
This chapter presents the findings from a comprehensive application of machine learning, signal processing, and statistical methods to detect and localize damage in a pipe while distinguishing changes produced by ambient environmental and operational variations. Damage was simulated physically by a mass scatterer grease-coupled to the surface of the pipe. Benign variations were simulated by varying the internal pressure of the pipe and the ambient temperature over time (Sections 5.1 and 5.2). To characterize the ultrasonic data, a library of 1500 features was extracted using various signal processing techniques (Section 5.4). An integrated filter-wrapper feature selection method was developed to automatically refine a small set of the most informative features (Section 5.5). With the selected features, adaptive boosting algorithms and support vector machines were applied to recognize the presence and the location of the scatterer. The classifiers were evaluated using a confusion matrix and a receiver operating characteristic curve (Section 5.6).

5.1 Experimental Methods for Data Acquisition

The data-driven SHM framework descried in Chapter 3 was demonstrated experimentally, using piezoelectric transducers affixed to a steel pipe specimen, 1833 mm in length, 70 mm in outside diameter and 4 mm in wall thickness. The pipe had a welded-on cap connected to an air compressor, and a pressure gauge at one end to control the internal atmospheric pressure from 0 pounds per square inch (PSI) up to 110 PSI. Two PZT wafers (PSI-5A4E, Piezo Systems, Inc.,

Cambridge, MA) were mounted on the surface of the pipe using cyanoacrylate adhesive, located 1.5 m apart. Each wafer was 12 mm long and 6 mm wide (Figure 5.1a). One PZT transducer functioned as a transmitter and the other as a receiver in a pitch-catch configuration. The signal transmission and data acquisition unit was a National Instruments PXI system with an arbitrary waveform generator and a synchronized digitizer. The waveform generator drove the PZT transmitter by a ±10 V sinc pulse, with 300 kHz center frequency, 200 kHz bandwidth, and 1 MHz sampling rate; the digitizer was wired to the PZT receiver and recorded responses, averaging 200 times, at the same sampling frequency. Each record of data consisted of 10,000 samples, corresponding to a 10 ms time history.

Three collections of data were recorded on three different days, with a one-month interval between Collections I and II, and four months between Collections II and III. This was done to induce uncertain variations in ambient environmental effects and in sensor bonding conditions, in addition to the pressure variations inside the pipe. During each collection of the experiments, a metal mass scatterer (38 mm in diameter, as shown in Figure 5.1b) was grease-coupled at different locations on the pipe surface (within a 120 degree offset with respect to the transmitter) to simulate a defect at those locations. Six locations were chosen in Collection I: three near the transmitter (Zone 1), and three close to the receiver (Zone 3); nine different positions were chosen in Collection II and Collection III, respectively: three locations at each of the three zones shown in Figure 5.1a. For every mass position, 20 scatterer records were measured as the pipe was randomly pressurized or discharged between 0 and 110 PSI; 20 undamaged records were collected, also at varying pressure levels, each time after removing the mass from the pipe and before placing it to the next position. Table 5-1 lists a breakdown of a total of 1000 records collected on the three days.



(a) (b)

Figure 5.1 (a) Schematic of the pressurized steel pipe, and (b) photo of the mass scatterer on the pipe.

Table 5-1 A breakdown of the datasets in the three collections.

	UNDAMAGED	SCATTERER		
		Zone 1	Zone 2	Zone 3
Collection I (10/15/2010)	140	60	0	60
Collection II (11/19/2010)	180	60	60	60
Collection III (3/24/2011)	200	60	60	60

5.2 Experimental Observations

Figure 5.2 shows a sinc transmission waveform, three energy-normalized Collection I measurements taken under different experimental setups, and the corresponding changes created from baseline subtraction. It can be seen that the short excitation pulse at the actuator produces a long sequence of complex arrivals at the downstream sensor, as a result of the presence of multiple dispersive wave modes in the pipe (Ying et al. 2010). In addition, the change in the ultrasonic signals caused by the mass scatterer is rather subtle and difficult to interpret visually (Figure 5.2f), whereas a 110 PSI internal pressure increment generates a more significant change

in the signal (Figure 5.2e) and makes it even more difficult to extract visually the pattern of the scatterer. Moreover, the L2-norm (root-of-sum-of-squares) of the difference between each measurement and the mean of the first 20 undamaged records in each collection is calculated, as shown in Figure 5.3 (dots represent the measurements with the absence of the scatterer; crosses represent the measurements with the presence of the scatterer). No apparent distinction between the undamaged instances and the damaged ones could be recognized in any of the three collections. Rather, the norms of the differential signals nearly trace the dynamic pressure variations over time.

In addition, the measurements show distinct variation from day to day. A Collection II record and a Collection III record shown in Figure 5.4 are seemingly different waveforms compared to the Collection I record (Figure 5.2), indicating that the changes caused by ambient conditions on different days, and/or the sensor bonding weakening over a long period of time, greatly exceed the changes produced by the scatterer. Likewise, two absolute value correlation matrices shown in Figure 5.5 illustrate the similarity between each pair of signals and of differential signals, respectively, with the Pearson product-moment correlation coefficient of one indicating two identical signals and zero indicating no similarity. Three bright squares along the diagonal suggest strong similarities among the signals collected on the same day, while the dark off-diagonal regions indicate significant discrepancy between three collections. The goal was then to detect the subtle changes produced by the mass scatterer in spite of the large changes produced by environmental and operational conditions. Adopting the data processing procedure of the data-driven framework, an integrated signal processing and machine learning approach was investigated and is described in the following sections.

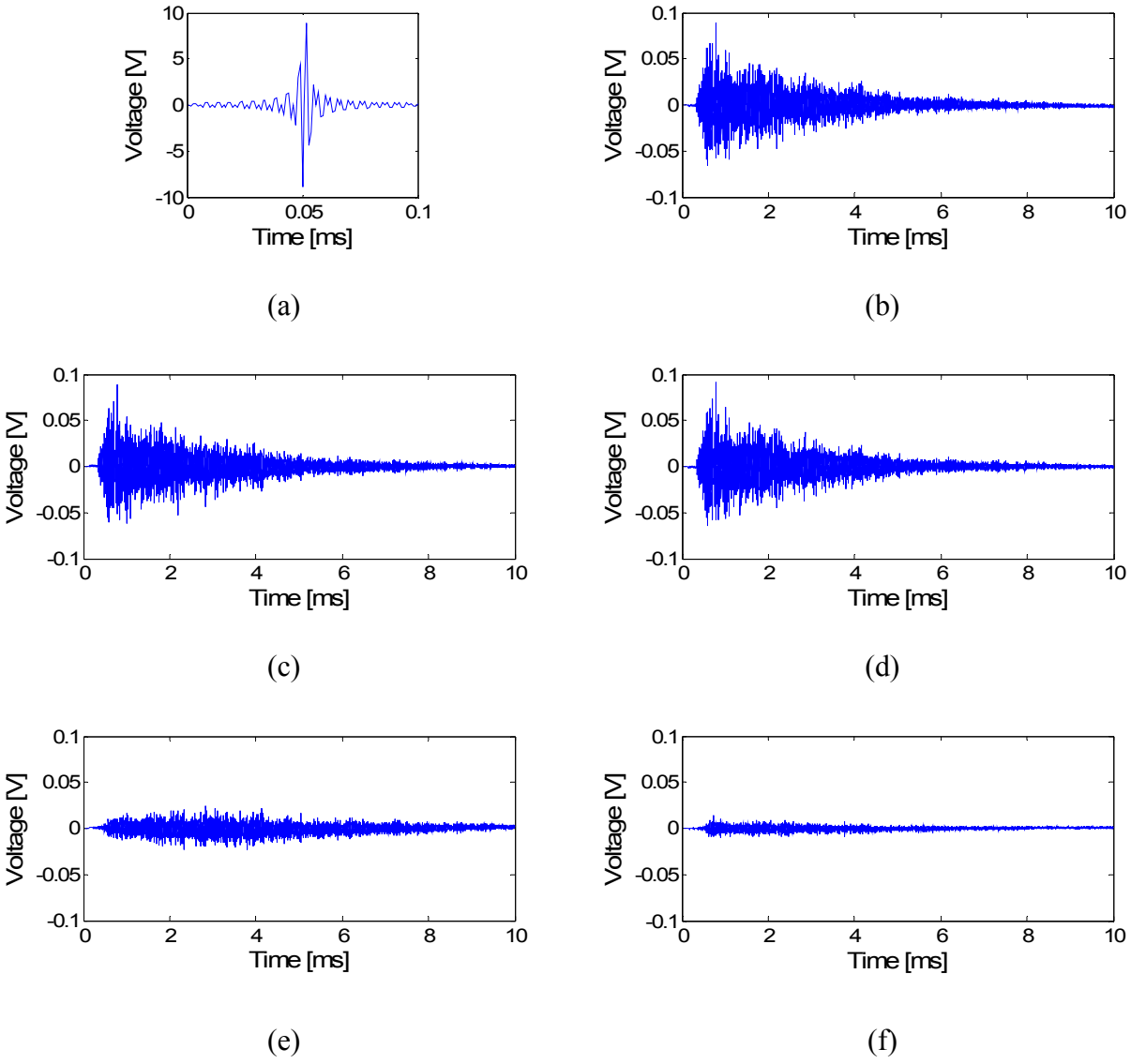
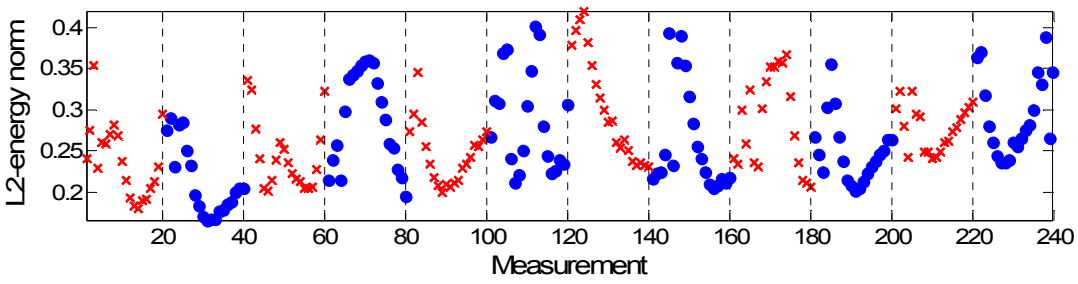
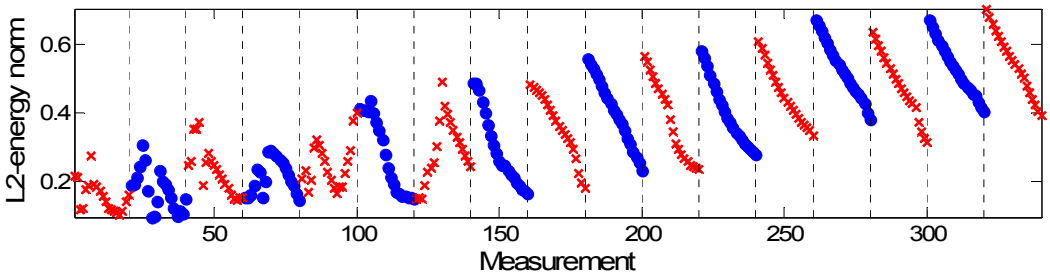


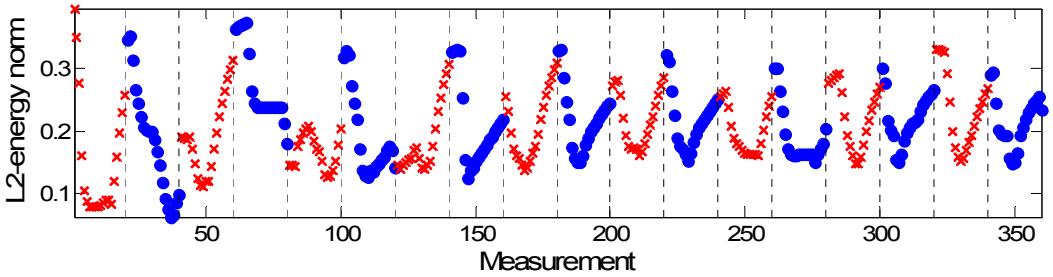
Figure 5.2 (a): Sinc pulse excitation, energy-normalized Collection I data: (b): a baseline signal recorded under 0 PSI, (c): received signal under 110 PSI, with the absence of the mass scatterer, (d): received signal under 0 PSI, with the presence of the scatterer, (e): differential signal of (c) and (b), and (f): differential signal of (d) and (b).



(a)

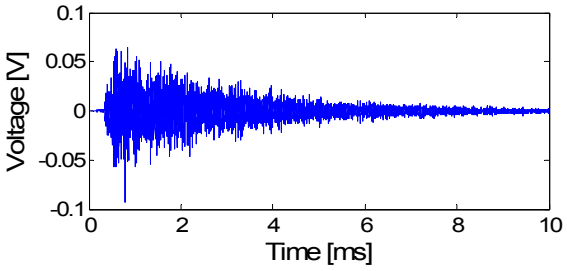


(b)

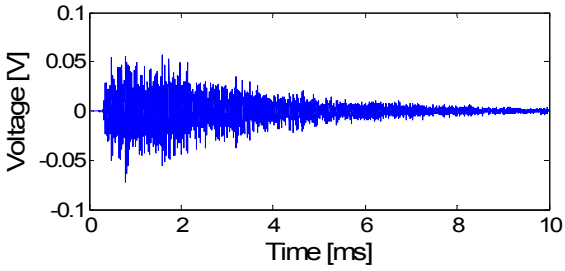


(c)

Figure 5.3 L2-norms of differential signals: (a) Collection I, (b) Collection II, and (c) Collection III. (Dots: undamaged instances; crosses: damaged instances.)

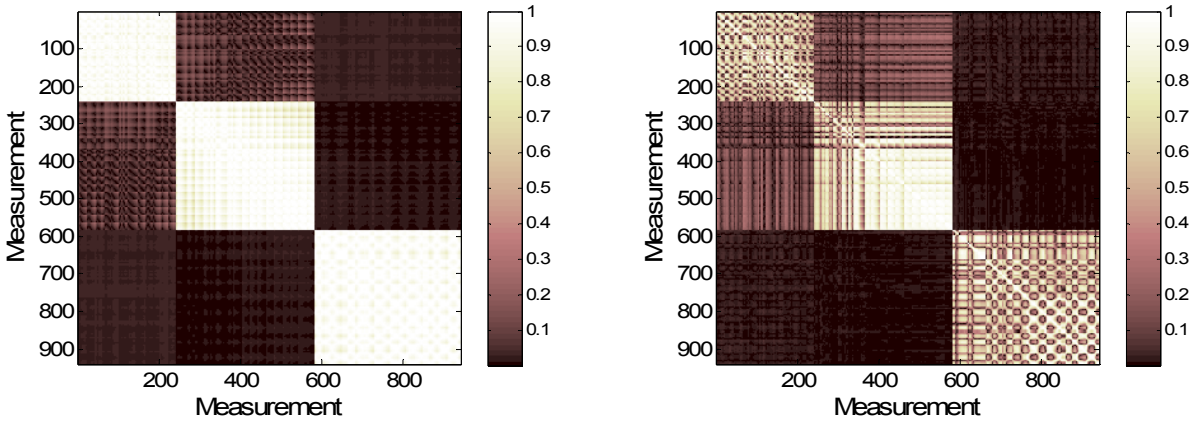


(a)



(b)

Figure 5.4 Received signals with the absence of the mass scatterer, (a) a Collection II record, and (b) a Collection III record.



(a) (b)

Figure 5.5 Absolute value correlation matrix of (a) the measurements and (b) the differential measurements of the three collections.

5.3 Machine Learning Test Formulation

Supervised learning (inferring a predictive model from labeled data samples) approaches were investigated in the research presented in this thesis. In the execution of machine learning for scatterer detection and localization, three sets of data were defined: a basis set, a training set, and a testing set. The basis set in this study was used to construct linear projection subspaces for generating certain features (presented in Section 5.4); the training data set was used to train a model representing the relationship between labels and input features; the predictability of the learned classifier was then assessed against the testing data set. The term "basis set" is not to be confused with "baseline". A baseline is an undamaged measurement taken under the pristine condition of a structure, while a basis set consists of both undamaged and scatterer measurements, which were used to generate features for the training set and the testing set.

The procedure of the machine learning process executed in this research is shown in Figure 5.6. First, different types of features are extracted from the training set. Second, a set of effective features are identified by applying feature selection algorithms to the features of the

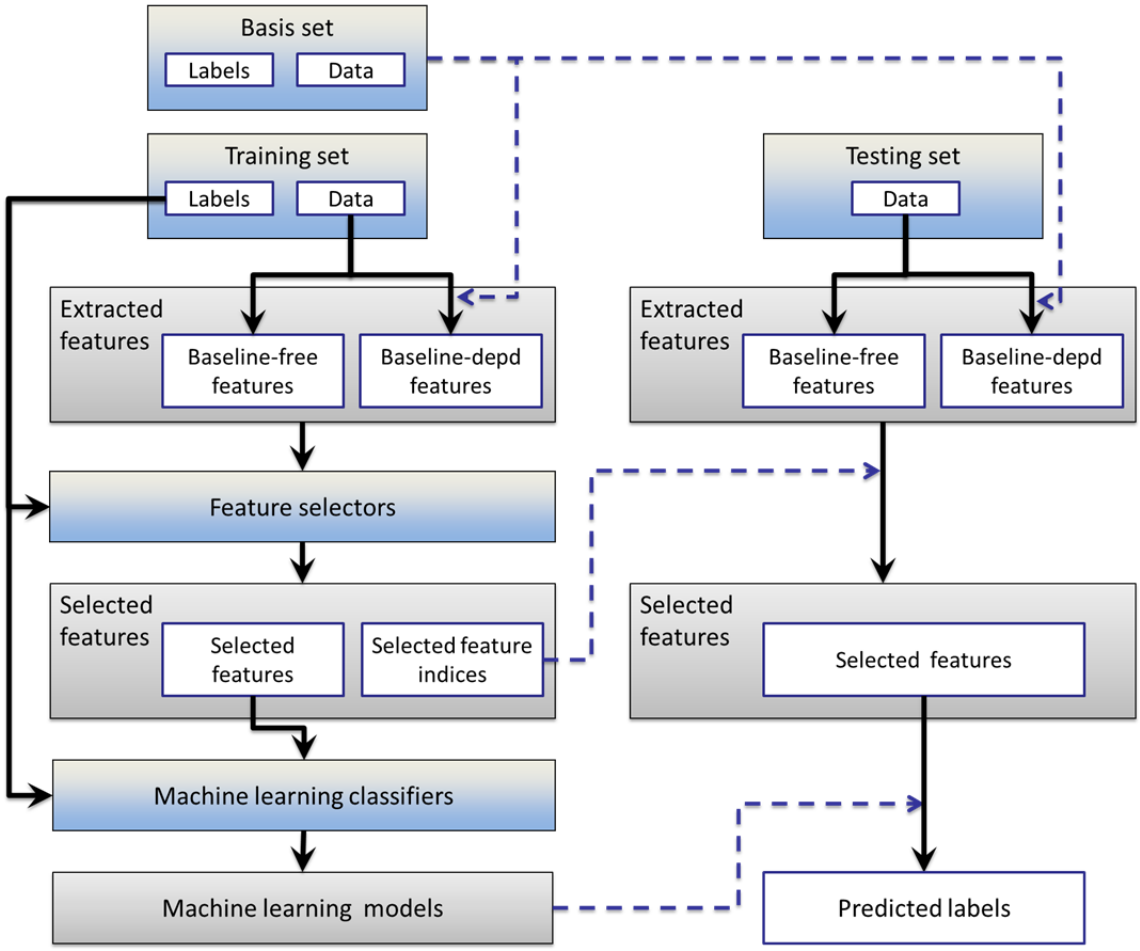


Figure 5.6 Procedure of machine learning process.

(Light blue modules are user-defined: including the input data and the chosen machine learning algorithm(s); gray modules are automated computational procedures; white blocks are input or output data; solid arrows indicate the executive flow; and dashed arrows indicate data or model feeding).

training set along with the known training labels. Third, the selected features and the labels of the training set are used to learn suitable machine learning classifiers. After these three steps, machine-learning models are prepared to predict the labels of any new measurements from the pipe. In the testing process, the same types of features are extracted from a different set of data. The selected features are fed into the trained classifiers so that the labels of the testing set can be determined. The accuracy and the reliability of the machine learning models can be evaluated by comparing the predicted labels with the ground truth.

In this research, random testing and systematic testing were designed to assess the machine learning methods for scatterer detection and localization (as shown in Table 5-2). In the random testing, a collection of data was randomly divided into three equal (or approximately equal) sized subsets, to function as a basis set, a training set and a testing set. By permuting the roles of the three subsets, six random tests were conducted for each given collection of data. The random testing was performed on each single-day data collection, as well as a combined set consisting of the three collections. In the single-day systematic testing, a collection of data was divided into three subsets based on the different scatterer locations; the multi-day systematic testing separated the data sets based on the three experimental dates. By permuting the roles of the three subsets, six systematic tests were conducted on each collection, as well as on the combined three-day data set.

It should be noted that there was no overlap between the training data and the testing data used for any of the testing. The training and testing sets in a random testing might contain instances corresponding to a same mass position, but the measurements were taken under

different environmental conditions. A systematic test is a stricter evaluation of the predictability of a classifier; it answers two practical questions:

- Is a machine learning model, built upon the instances of the scatterer located at one zone of the pipe, capable of recognizing the presence of the scatterer at different locations/zones?
- Is a classifier built upon the datasets collected months ago still effective to detect the scatterer at the current time?

Table 5-2 Machine learning test formulation for the laboratory studies.

	Single-day data	**Multi-day data**
Random testing	3 randomly divided, equal (or approximately equal) sized sets Total: 18 tests	3 randomly divided, equal (or approximately equal) sized sets Total: 6 tests
Systematic testing	3 measurement sets differentiated according to the zone/location where the mass scatterer was located Total: 18 tests	Collections I, II, and III (differentiated according to the experimental date) Total: 6 tests

5.4 Feature Extraction using Combined Signal Processing and Machine Learning

The performance of machine leaning algorithms relies on the effectiveness of features. In practice, it is difficult to find one universally good feature that is sensitive to all forms of damage, while robust in the presence of all types of ambient variations, for all types of structures.

Therefore, a large number of potentially good features were extracted in this work to create a library, and then feature selection methods were utilized to automatically search over the feature candidate pool for the most effective features for a specific task.

5.4.1 Feature Domain Construction

Various signal processing and machine learning methods were applied and combined to generate different signal domains (such as the time domain and the frequency domain) serving as bases from which to extract features. Such signal domains are referred as feature domains in this thesis. The construction of the feature domains follows three steps: wavelet processing, additional signal processing, and linear projection analysis, as illustrated in Figure 5.7 and detailed below.

Before the wavelet transform is implemented, a received ultrasonic signal is first passed through a 200-400 kHz band-pass filter to eliminate low-frequency mechanical vibrations and high-frequency electrical noise. The filtered signal is then energy-normalized to reduce ambient effects to certain degree.

Wavelet Processing. Wavelet techniques have been demonstrated to be effective for damage detection and classification in the areas of NDE and SHM. The most prominent advantages of wavelet processing are de-noising and detecting local discontinuities of a signal (Rizzo et al. 2005). Discrete wavelet decomposition was applied to the energy-normalized signal in this research, due to its computational efficiency.

Analogous to the discrete Fourier transform, where a signal is decomposed into orthogonal complex exponentials, the wavelet transform decomposes a signal into orthogonal wavelets that are scaled and translated from a mother wavelet (a fast-decaying oscillated

analyzing function). The difference between the two transforms lies in the basis function: the sinusoidal bases in the Fourier transform are functions of frequency, while the wavelet bases are functions of scale (corresponding to frequency) and dilation (corresponding to time).

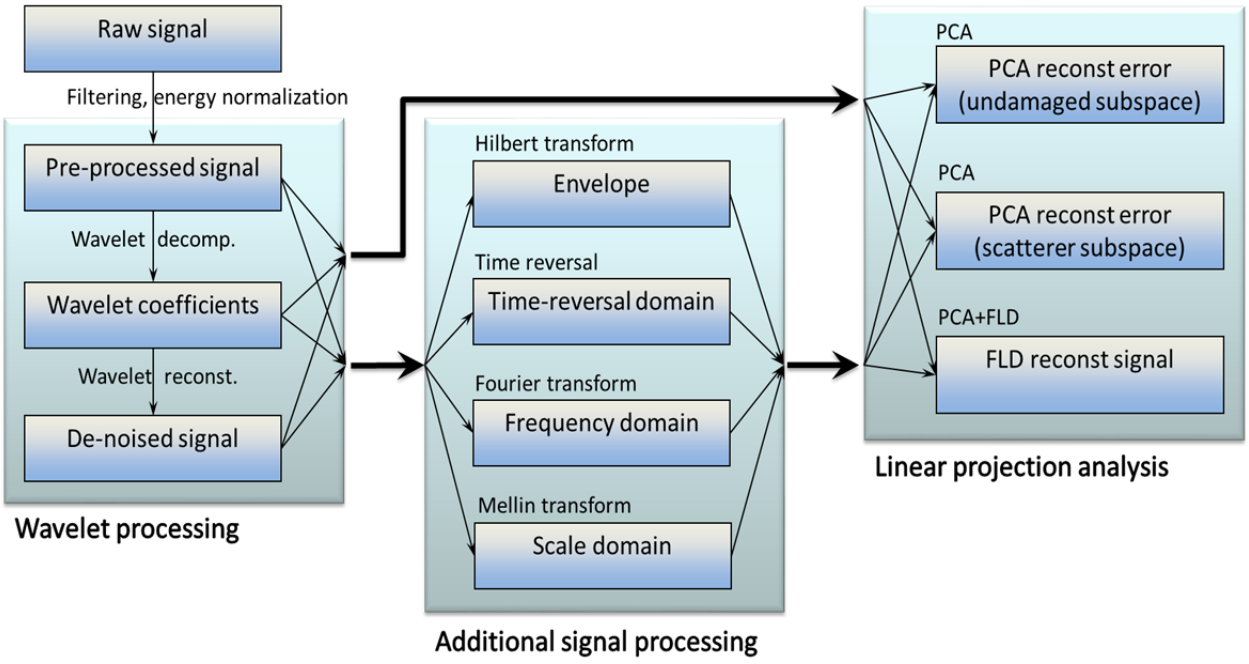


Figure 5.7 Flowchart of feature domain construction.

Equation (5-1) gives the wavelet functions, where *t* refers to time, and *s* and *l* are integers that scale and translate the mother wavelet to create a series of wavelet basis functions (Mallat 1999).

$$\psi_{s,l}(t) = \frac{1}{\sqrt{2^s}} \psi\left(\frac{t}{2^s} - l\right) \tag{5-1}$$

A time domain signal *x(t)* can then be represented as a weighted sum of those wavelets. The weights $W_{s,l}$ are the wavelet coefficients emphasizing the different contributions of certain frequency and time component of the signal:

$$x(t) = \sum_{s,l=-\infty}^{+\infty} W_{s,l}\psi_{s,l}(t) \tag{5-2}$$

The wavelet coefficients can be computed using an integral transform of the inner product of *x(t)* and the complex conjugate of the analyzing wavelet at each scale and dilation (Rizzo et al. 2005) :

$$W_{s,l} = \int_{-\infty}^{+\infty} x(t) \cdot \psi_{s,l}^{*}(t)\, dt \tag{5-3}$$

In a fast discrete wavelet decomposition, the wavelet coefficients can be obtained by applying two-channel conjugate quadrature mirror filters (Mallat 1989). The outputs of high-pass filters and low-pass filters are wavelet detail coefficients characterizing the high-frequency components of the signal, and approximation coefficients characterizing the low-frequency components of the signal, respectively.

The wavelet decomposition can be implemented hierarchically, where each level, *s*, corresponding to a scale 2^s with a resolution of 2^{-s}, is a process of high-pass and low-pass filtering and subsampling, leading to half the number of samples and half the frequency band spanned. For analysis purposes, the frequency of interest should lie on the frequency band at the selected level. For the signals collected in this study, with 300 kHz center frequency and 200 kHz bandwidth, sampled at 1 MHz, only the detail coefficients at Level 1 were required. The approximation coefficients at Level 1 corresponded to mostly noise and therefore were eliminated (Ying et al. 2012).

Wavelet reconstruction is the reversed process of decomposition, involving upsampling and filtering. Selecting and thresholding the wavelet coefficients before the wavelet

reconstruction results in a de-noised signal. For a discrete time domain signal *x[i]*, the reconstructed signal is obtained by

$$\tilde{x}[i] = \sum_{s} \sum_{l} W_{s,l} \psi_{s,l} \tag{5-4}$$

In this study, the detail coefficients at the first level after soft-thresholding (Donoho 1995) were used. This de-noised signal, the wavelet coefficients and the pre-processed time domain signal were all treated as feature domains.

It is noted that selection of the mother wavelet can affect the performance of wavelet processing. The mother wavelet function was chosen based on two criteria: first, the wavelet function should have a shape that is similar to that of the analyzed signal; second, the center frequency F of the mother wavelet should make the reconstructed frequency f_s at the decomposition level s as close as possible to the center frequency of the analyzed signal. The relation between F, f_s and s, is given by (Rizzo et al. 2005):

$$f_s = \frac{\Delta \times F}{2^s} \tag{5-5}$$

where f_s is the reconstructed frequency at Level s, Δ is the sampling frequency, and F is the center frequency of the mother wavelet. Symlets of order 5 (Mallat 1999) with the center frequency of 333 kHz at Level 1 was chosen as the mother wavelet, as it was close to the 300 kHz center frequency used for the laboratory measurements (Ying et al. 2012).

Additional signal processing. The Hilbert transform, the time reversal focusing method, the Fourier transform, and the Mellin transform are applied to the pre-processed time domain

signal and the wavelet feature domains, to construct an envelope, a time-reversal domain, a frequency domain, and a scale domain, respectively, as additional feature domains.

The frequency domain and the signal envelope are commonly used in signal analysis. The envelope is obtained by taking the magnitude of an analytical signal computed by the Hilbert transform (A. V. Oppenheim and Schafer 1989). The time-reversal domain has been demonstrated to compensate for wave dispersion and to increase the signal-to-noise ratio (as was discussed in Chapter 4). The Mellin transform is scale-invariant, and therefore can generate more robust features against environmental factors that change the ultrasonic wave velocity (Harley et al. 2011). The Fast Mellin Transform algorithm was implemented and used in this work; the detailed procedure has been presented by De Sena and Rocchesso (2007).

All the signal domains are multiplied by a rectangular window before feature extraction in order to preserve the most significant portions while reducing the noise contributions as well as the dimensionality of each domain. Figure 5.8a-g exemplify the various windowed feature domains extended from a time domain signal after wavelet processing and other additional signal processing. Each such discrete signal domain can be treated as a vector of samples from which certain statistical features can be extracted. It should be noted that the feature domains in the "additional signal processing" module are also generated for the wavelet coefficients and the de-noised signal. Therefore, integrated signal domains are also created to take advantage of different signal processing techniques, such as a frequency domain of the wavelet coefficients, and a time-reversal domain of a de-noised signal.

Linear projection analysis. Linear projection methods are commonly used in machine learning for feature dimensionality reduction, by projecting a data point with M features into a

lower dimensional space through a linear transform matrix. The vector bases in the new space are functions over all the *M* features (Rao 1964). However, applying the linear projection methods to all the features extracted in the data-driven framework would not provide useful physical meaning for pipe monitoring. Instead, the linear projections are applied to the various feature domains generated in the preceding signal processing modules, in order to de-couple intrinsic structures of the changes in the ultrasonic response produced by a scatterer source and by non-scatterer environmental factors.

Two linear projection methods were applied in this thesis, principal component analysis (PCA) and Fisher's linear discriminant (FLD), which are analogous to the Eigenface and Fisherface approaches that are well-established for face recognition (Belhumeur et al. 1997). PCA searches for a direction that shows the largest variance in the entire data set; and the subsequent directions have the highest variance among all the possible directions orthogonal to the preceding identified directions. These orthogonal directions are known are principal components (PCs). In pipe monitoring, it is likely that the first PC or the first few PCs would identify the large environmental variations, and some succeeding PC(s) might correspond to the scatterer changes. Unlike PCA, FLD is a supervised linear transform that searches for a set of orthogonal directions that maximize the discrepancy between different classes.

Consider the basis set consisting of *N* n-sample data points, $\{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N\}$, which can be categorized into *C* classes $\{X_1, X_2, \cdots, X_C\}$ with the number of data points $N_1, N_2, \cdots, N_C$, respectively. $\boldsymbol{\mu}$ is an n-by-1 mean vector of the N data points, and $\boldsymbol{\mu}_i$ are mean vectors of the C classes, where $i = 1, 2, \cdots, C$. The total variance matrix is given by

$$V_T = \sum_{k=1}^{N} (\mathbf{x}_k - \boldsymbol{\mu})(\mathbf{x}_k - \boldsymbol{\mu})^T . \tag{5-6}$$

The between-class variance matrix and the within-class variance matrix are defined in Equations (5-7) and (5-8), respectively (Belhumeur et al. 1997).

$$V_B = \sum_{i=1}^{C} N_i (\boldsymbol{\mu}_i - \boldsymbol{\mu})(\boldsymbol{\mu}_i - \boldsymbol{\mu})^T . \tag{5-7}$$

$$V_W = \sum_{i=1}^{C} \sum_{\mathbf{x}_k \in X_i} (\mathbf{x}_k - \boldsymbol{\mu}_i)(\mathbf{x}_k - \boldsymbol{\mu}_i)^T . \tag{5-8}$$

Under the above notations, PCA can be mathematically expressed as an optimization problem: to maximize the determinant of the total variance matrix of the projected data points (Belhumeur et al. 1997):

$$\begin{aligned} E_{pca} &= \arg\max_E \left| E^T V_T E \right| \\ &= \left[\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_K \right] \end{aligned} \tag{5-9}$$

where $\left[\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_K \right]$ is the set of *n*-dimensional eigenvectors of V_T corresponding to the *k* largest eigenvalues.

FLD maximizes the ratio of the determinant of the between-class variance matrix to the determinant of the within-class scatter matrix (Belhumeur et al. 1997):

$$E_{fld} = \arg\max_E \frac{\left| E^T V_B E \right|}{\left| E^T V_W E \right|} . \tag{5-10}$$

To avoid the singularity of the within-class variance matrix caused by the larger number of data dimensions compared to the number of data points, Belhumeur et al. (1997) developed a Fisherface approach by combining PCA and FLD. More specifically, PCA is first applied to the data for dimensionality reduction, and then FLD is implemented in the PCA subspace. Therefore, Equation (5-10) can be rewritten as

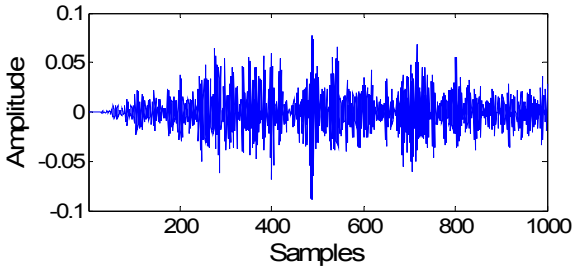
$$E_{fld} = \arg\max_{E} \frac{\left| E^T E_{pca}^T V_B E_{pca} E \right|}{\left| E^T E_{pca}^T V_W E_{pca} E \right|}. \tag{5-11}$$

In the following text, the combined PCA and FLD is referred as FLD for simplicity.

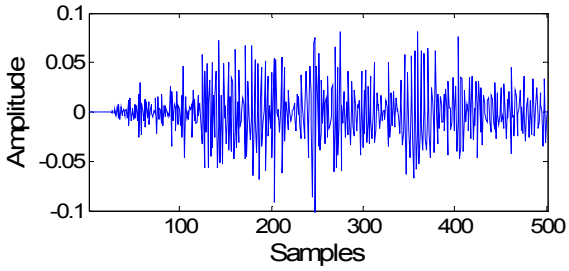
In this thesis, PCA and FLD were used to construct feature domains described in this section, and were also used to generate features presented in the succeeding section. The feature domains include a PCA reconstruction error in an undamaged subspace, a PCA reconstruction error in a scatterer subspace, and a FLD reconstructed signal domain. The procedures of generating these domains are descried in Table 5-3 and Table 5-4. It should be noted that Table 5-3 provides the general implementation of the PCA algorithm. To generate an undamaged PCA subspace, only undamaged data should be used. Similarly, only scatterer data instances were used to generate the reconstruction error domain in the scatterer subspace.

Figure 5.8h-j shows examples of the three linear projection domains constructed corresponding to a scatterer record from Collection I. It can be seen that the residual error in the scatterer subspace has an amplitude approximately 1/3 of that in the undamaged subspace. The undamaged subspace does not contain any information regarding the scatterer characteristics, and therefore the reconstruction of the scatterer record results in larger error. This observation

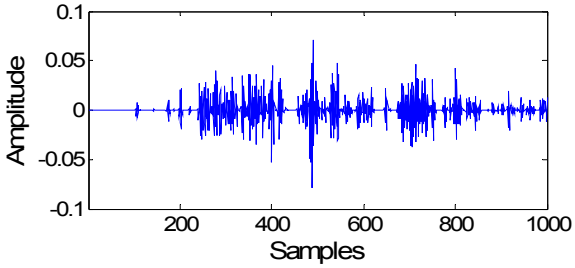
implies that features extracted from such domains can indicate the presence of the scatterer. The linear projection methods were also combined with different signal processing techniques, by applying PCA or FLD to various signal domains, such as the wavelet coefficients and the Mellin domain. The combined analysis may lead to features that are more effective.



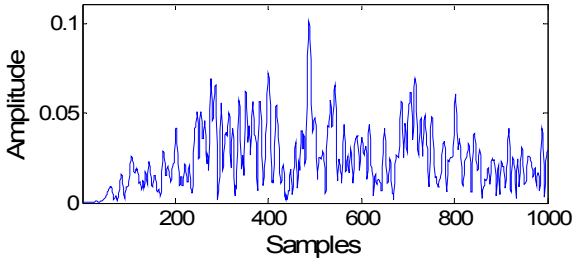
(a)



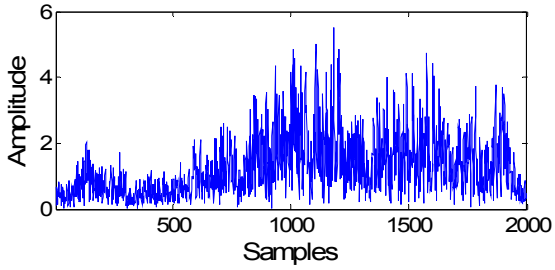
(b)



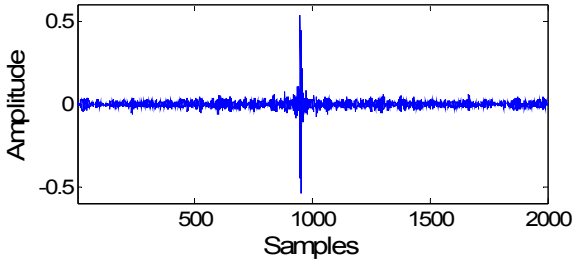
(c)



(d)



(e)



(f)

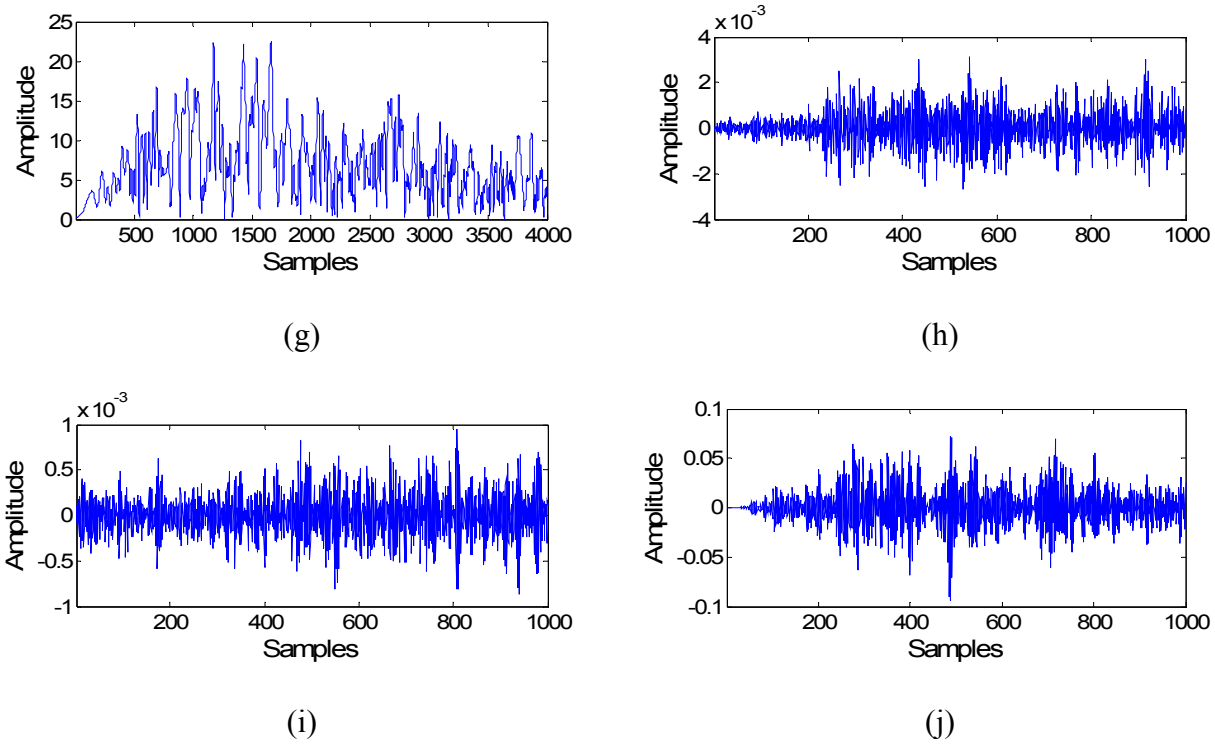


Figure 5.8 Feature domains of a scatterer record from Collection I: (a) pre-processed time domain signal, (b) detailed wavelet coefficients, (c) de-noised signal after wavelet processing, (d) signal envelope, (e) frequency domain, (f) time-reversal domain, (g) scale domain, (h) PCA reconstruction error in a subspace constructed using undamaged records from a basis set, (i) PCA reconstruction error in a subspace constructed using scatterer records from the basis set, and (j) FLD reconstructed signal using a single FLD basis computed from the basis set.

Table 5-3 Procedure of PCA subspace generation, and construction of PCA-based feature domains and features.

Generation of a PCA subspace:

Step 1. Compute a mean vector $\boldsymbol{\mu}$ of N data points from the basis set

$$\{\boldsymbol{\mu}\}_{n\times 1} = \frac{1}{N}\left(\mathbf{x}_1 + \mathbf{x}_2 +, \cdots, +\mathbf{x}_N\right)$$

Step 2. Subtract each of the N signals by $\boldsymbol{\mu}$, and align the centered signals into a n-by-N matrix

$$\{M\}_{n\times N} = \left[\mathbf{x}_1 - \boldsymbol{\mu},\ \mathbf{x}_2 - \boldsymbol{\mu},\ \cdots,\ \mathbf{x}_N - \boldsymbol{\mu}\right]$$

Step 3. Apply singular value decomposition to M

$$\{M\}_{n\times N} = USV^*$$

Where U is n-by-n unitary matrix, known as left singular vectors; S is a n-by-N rectangular diagonal matrix, the diagonal entries of which are singular values; and V^*, the conjugate transpose of V, is a N-by-N unitary matrix, known as right singular vectors.

Step 4. Select K ($K \le N$) column vectors in U with the first K largest singular values. Each column is a PCA basis vector

$$\left\{E_{pca}\right\}_{n\times K} = \left[\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_K\right]$$

Construction of PCA-based feature domains and features:

Step 5. For any new data point $\mathbf{x}_{new}$ (not from the basis set), project centralized $\mathbf{x}_{new}$ onto the K–dimensional PCA subspace, obtaining a K-by-1 projection

$$\{\mathbf{w}\}_{K\times 1} = E_{pca}^T\left(\mathbf{x}_{new} - \boldsymbol{\mu}\right)$$

Step 6. Reconstruct $\mathbf{x}_{new}$ using the K PCA bases

$$\left\{\widetilde{\mathbf{x}_{new}}\right\}_{n\times 1} = E_{pca}\mathbf{w} = E_{pca}E_{pca}^T\left(\mathbf{x}_{new} - \boldsymbol{\mu}\right) + \boldsymbol{\mu}$$

Step 7. Compute the residual error after reconstruction

$$\{\boldsymbol{\varepsilon}\}_{n\times 1} = \mathbf{x}_{new} - E_{pca}E_{pca}^T\left(\mathbf{x}_{new} - \boldsymbol{\mu}\right) - \boldsymbol{\mu}$$

Table 5-4 Procedure of FLD subspace generation, and construction of FLD-based feature domains and features.

Generation of a FLD subspace (Belhumeur et al. 1997):

Step 1. Generate a *K*-dimensional PCA subspace using all the data points in the basis set, following *Step 1-4* described in Table 5-3

Step 2. Compute a mean vector $\boldsymbol{\mu}_i$ of N_i data points for each of the C classes $\{X_1, X_2, \cdots, X_C\}$ within the basis set, where $i = 1, 2, \cdots, C$, and *C* is the total number of classes. For scatterer detection, *C* equals to 2

Step 3. Calculate the between-class variance V_B using Equation (5-7), and within-class variance V_W using Equation (5-8)

Step 4. Compute the generalized Eigen decomposition

$$\left(E_{pca}^T V_B E_{pca}\right) U = \left(E_{pca}^T V_W E_{pca}\right) US$$

Where the columns of *U* are generalized eigenvectors, and the diagonal entries of *S* generalized eigenvalues.

Step 5. Select *C*-1 column vectors in *U* with the first *C*-1 largest eigenvalues. The FLD basis vectors are obtained by

$$\left\{E_{fld}\right\}_{n\times(C-1)} = E_{pca}\left\{U\right\}_{K\times(C-1)}$$

Step 6. Normalize the FLD basis vectors to unit vectors. The normalized FLD bases is denoted as $\overline{E_{fld}}$

Construction of FLD-based feature domains and features:

Step 7. For any new data point $\mathbf{x}_{new}$ (not from the basis set), project centralized $\mathbf{x}_{new}$ onto the *(C-1)*–dimensional FLD subspace, obtaining a (C-1)-by-1 projection

$$\left\{\mathbf{w}_f\right\}_{(C-1)\times 1} = \overline{E_{fld}}^T \left(\mathbf{x}_{new} - \boldsymbol{\mu}\right)$$

Step 8. Reconstruct $\mathbf{x}_{new}$ using the *(C-1)* FLD bases

$$\left\{\widetilde{\mathbf{x}_{new}}\right\}_{n\times 1} = E_{fld}\mathbf{w}_f = \overline{E_{fld}}\,\overline{E_{fld}}^T \left(\mathbf{x}_{new} - \boldsymbol{\mu}\right) + \boldsymbol{\mu}$$

5.4.2 Extracted Features

Two types of features were considered in this research: one type that requires a baseline or the basis set, and the other type that is independent of the baselines. The features are extracted from the different feature domains by analysis of peaks (local maxima), statistical methods and correlation, resulting in 1500 different features in total (listed in Table 5-5).

5.4.2.1 Baseline-Free Features

In addition to the mean, variance and kurtosis statistical features, and the peak level feature, described in Section 4.1, three other baseline-free features were also considered in this study:

Energy. For an *n*-sample discrete signal $x[i]$, the energy is defined as the sum of the squared magnitude of the samples:

$$E = \sum_{i=1}^{n} x[i]^2 \tag{5-12}$$

Local maxima features. The peaks of a complex signal indicate the arrival, reflection, or conversion of wave modes. In addition to the greatest peak, certain peaks can be affected differently from others when damage is introduced. Local maxima are searched for over different feature domains. Features are constructed from the first three greatest peaks, and the peaks with the amplitude larger than 20% and 60% of the greatest peak.

Curve length. The curve length of a signal is useful for describing the signal complexity (Lu and Michaels 2009). A variation in curve length may be caused by changes in the modal

amplitudes or locations of waves. The curve length is also robust to time-scale changes since the signal's shape remains the same. The curve length of a discrete signal domain $x[i]$ is defined by

$$L = \sum_{i=2}^{n} \left| x[i] - x[i-1] \right| \tag{5-13}$$

5.4.2.2 Baseline-Dependent Features

Correlation coefficients. The Pearson product-moment correlation coefficient described in Section 4.1 is computed with respect to 20 basis vectors, including a FLD basis and 19 PCA basis vectors with the 19 largest eigenvalues computed from the entire basis set.

Linear projections. The projections onto the aforementioned 20 basis vectors are used as features. The computation of the PCA projections is given by *Step 5* in Table 5-3, and the FLD projection is computed by *Step 7* in Table 5-4.

Baseline-free type features extracted from linear transformation domains. The baseline-free type features (described above) are extracted from the baseline-dependent PCA and FLD domains that are created through the Linear Projection Analysis module presented in Section 5.4.1.

Table 5-5 List of features extracted (1500 in total).

	Types of Features	Number
Baseline-free	○ Energy	15
	○ Curve length	15
	○ Kurtosis	15
	○ Amplitude of the 1st greatest peak	15
	○ Amplitude difference between the 1st and the 2nd greatest peaks	15
	○ Amplitude difference between the 1st and the 3rd greatest peaks	15
	○ Location of the 1st greatest peak	15
	○ Location distance between the 1st and the 2nd greatest peaks	15
	○ Location distance between the 1st and the 3rd greatest peaks	15
	○ Energy of peaks greater than 60% of the maximum amplitude	15
	○ Energy of peaks greater than 20% of the maximum amplitude	15
	○ Mean of peak values greater than 60% of the maximum amplitude	15
	○ Mean of peak values greater than 20% of the maximum amplitude	15
	○ Sample variance of peak values greater than 60% of the maximum amplitude	15
	○ Sample variance of peak values greater than 20% of the maximum amplitude	15
Baseline-dependent	○ Correlation coefficients with respect to 20 basis vectors (19 eigenvectors with the largest eigenvalues, and 1 FLD basis computed from the basis set)	300
	○ Projections onto the 20 basis vectors	300
	○ Baseline-free type features extracted from a PCA reconstruction error in a scatterer subspace	225
	○ Baseline-free type features extracted from a PCA reconstruction error in a undamaged subspace	225
	○ Baseline-free type features extracted from a FLD reconstructed signal	225

5.5 Integrated Feature Selection and Evaluation

Creating a library of features is necessary for generalization purposes. However, applying machine learning classifiers to all the 1500 features is not computationally efficient, and some features may adversely affect the classifier performance. Feature selection can also suggest good change detectors and may provide insights into the physical characteristics of the sensor-structure system.

In this section, an integrated filter-wrapper feature selection method developed as part of this research is described. First, statistical approaches were utilized to remove irrelevant and redundant features. Next, a wrapper method was applied to refine a small set of features associated with classification learning algorithms. Finally, the performance of the selected features was evaluated.

5.5.1 Integrated Filter-Wrapper Feature Selection Method

Given a substantial number of features generated using the different techniques, feature dimension reduction is desirable especially for field applications. Good features should be highly correlated with labels of different classes, while uncorrelated with other features (Hall 1999). Thus, some simple and fast-implemented filter selection methods were considered to score the features based on certain criteria, such as class separability and correlation measures. By choosing a threshold for the scores, some apparently irrelevant and redundant features could be screened out before passing them over to machine learning classifiers.

The filter selection methods usually rank the features, but do not determine an optimal number of features that should be selected. In addition, the number of optimal features is likely

to vary for different classifiers as well as under different test conditions. Kohavi and John (1997) have reported those limitations of such filter selection methods based on criteria unassociated with the learning algorithms. By contrast, Kohavi and John (1997) have developed wrapper methods that adaptively identify a subset of salient features by assessing them through the specific learning algorithms. However, the wrapper selection method is much less computationally efficient compared to the filter approaches, because the learning algorithm is heavily involved in the feature searching process that examines the performance of all the possible combinations of feature sets. In order to allow for an efficient implementation, different search strategies have been devised, including forward search and backward search, to evaluate each feature subset by sequentially adding or removing features (Guyon and Elisseeff 2003; Kohavi and John 1997). The evaluation of feature subsets is often done by cross-validation to search for a local minimum of the misclassification error. Given that wrapper methods are not cost-effective for a large number of features, it is advisable to use wrapper methods on pre-processed features with reasonably good ordering.

Based on the preceding discussion, the filter methods and the wrapper selection method are incorporated to utilize the advantages of both while compensating for their limitations. As shown in Figure 5.9, the integrated feature selection method consists of three steps: Welch's *t*-test, correlation-based filters, and wrapper backward sequential selection.

First, Welch's *t*-test is applied to each feature to test the null hypothesis that the feature means of the data points in the two classes are equal. Welch's *t*-test is a generalized Student's *t*-test applicable to two samples with unequal variances. The statistic *t* is defined by (Welch 1947):

$$t = \frac{\mu_1 - \mu_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}} \tag{5-14}$$

where μ_i, s_i^2 and N_i are sample mean, variance and size, respectively. The variance is estimated using the Welch-Satterthwaite equation (Welch 1947).

The t-test yields a *p*-value indicating the confidence of the null hypothesis. For a discriminative feature, a lower *p*-value is desirable corresponding to the rejection of the null hypnosis. Therefore, the *p*-values can be used to rank the features based on the criterion of class separability, and the features producing *p*-values above a significance level (e.g. 0.05) are irrelevant features that should be removed from the feature pool.

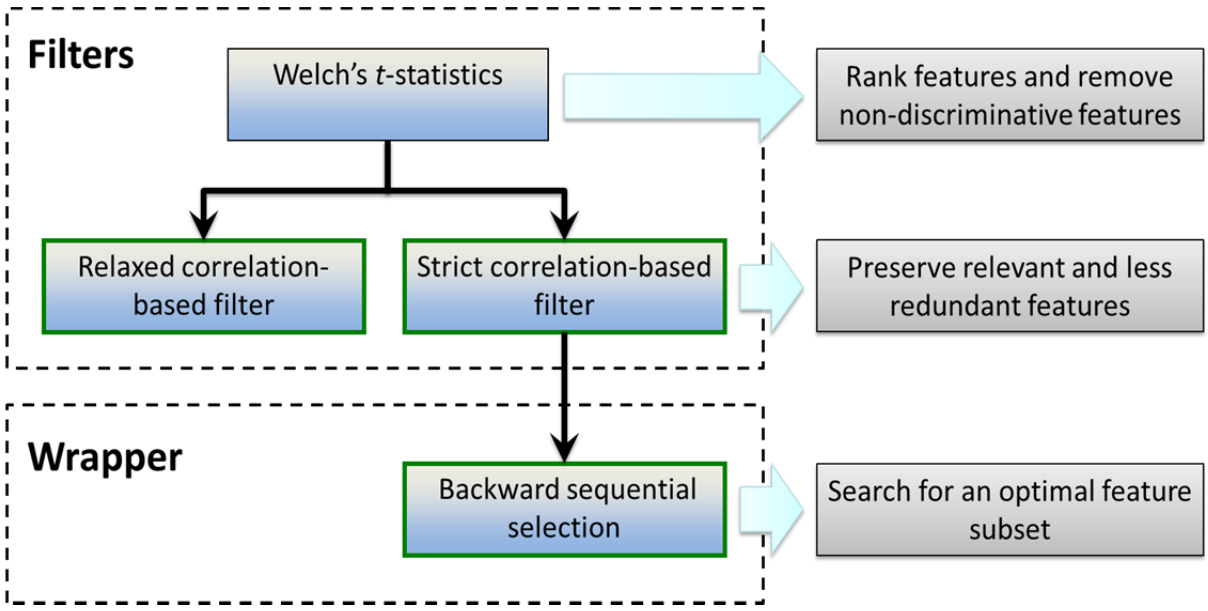


Figure 5.9 Procedure of the integrated filter-wrapper feature selection.

Second, a correlation filter is applied to the remaining features after the *t*-test filtering, based on pairwise correlation coefficients among those features and the correlation coefficient between each feature and the labels. The absolute value correlation coefficient thresholds are composed of a lower bound that is the lowest allowed correlation coefficient with respect to the labels, and an upper bound that is the maximum allowed correlation coefficients among all the retained features. To satisfy the upper threshold, only the feature with the lowest *p*-value is preserved among a group of highly correlated features. As a result, the output feature set has reduced dimensionality, less redundancy, and relatively high correlation with the labels of different classes.

Adjusting the correlation thresholds yields a relaxed filter (R-filter) and a strict filter (S-filter). In this study, a 0.9 upper bound was considered for the R-filter, provided that features with high correlation may serve as complements in classification tasks (Guyon and Elisseeff 2003), and a lower bound of 0.1 was used to avoid over fitting, because ineffective features for one data set may be good features for another. The S-filter is a stricter feature evaluation, so thresholds that were more restrictive were used: (0.6, 0.85). The S-filter also serves as a pre-processing step of the wrapper method that is computationally costly if applied to a relatively large number of features.

For the third step, a backward sequential search is applied to the features output from the S-filter. The features, ordered in terms of the *p*-values, are evaluated sequentially starting from the end with the largest *p*-value to determine whether to remove or keep the feature. The evaluation metric used was an average misclassification error of a stratified 5-fold cross validation test on the training set by a learning algorithm (the same algorithm was also applied to

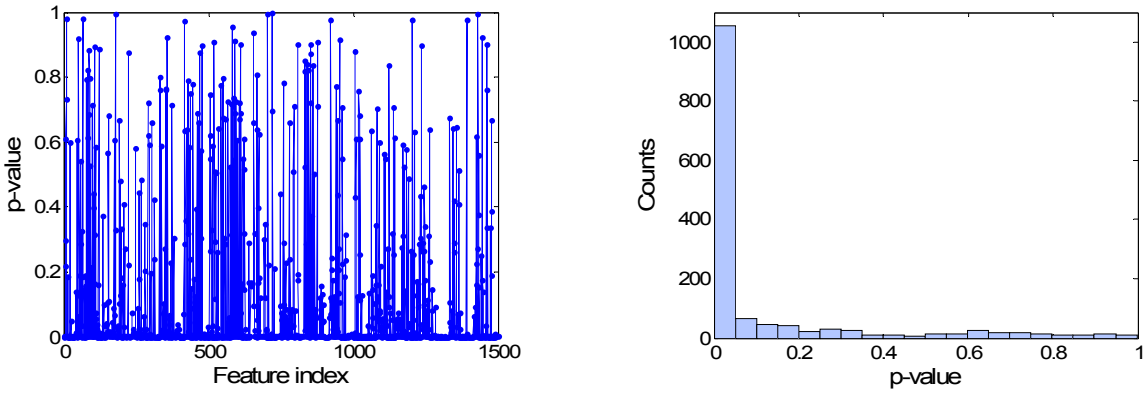
predict the labels of the test set). Four classification methods were used in this thesis (detailed in Section 5.6): adaptive boosting (AdaBoost), modified AdaBoost (M-AdaBoost), support vector machines with a linear kernel function (linear-SVM), and support vector machines with a radial basis function as kernel (RBF-SVM).

5.5.2 Results of Feature Selection

Following the procedure of the integrated feature selection, the *t*-test was performed on each feature, producing a *p*-value as a measure of the effectiveness of the feature. An example result of this process using a multi-day systematic test is illustrated in Figure 5.10a, and the distribution of the *p*-values of the 1500 features is plotted in a histogram (Figure 5.10b). Considering a significance level of 0.05, the first bin of the histogram corresponds to the features selected by the *t*-test, a total of 1053 features. The features are ranked in ascending order of their corresponding *p*-values.

Before passing the 1053 features to the correlation-based filters, the correlation among the feature pairs, and the correlation between each feature and the labels of the training data were evaluated, the results of which are shown in Figure 5.11a and Figure 5.11b, respectively. In general, the 1053 features do not appear to show much redundancy, whereas there is room for feature dimensionality reduction without degrading the performance of the feature set. For instance, a large portion of the first 400 features (corresponding to the upper left corner of the feature correlation matrix with high intensity) could be removed. By applying the R-filter, another 541 features could be eliminated from the feature candidates. The resulting feature correlation matrix shows much lower correlation among the features (Figure 5.11c). The feature dimensionality can be decreased to 18 by using the S-filter as illustrated in Figure 5.11e. The

correlation coefficients between each pair of the 18 features are below 0.85 and the correlation coefficients with respect to the training labels are above 0.6 (Figure 5.11e and Figure 5.11f). The wrapper selection was then applied to further refine a smaller set of features supposedly more suitable for certain learning algorithms. The numbers of the selected features in multiple random tests using different algorithms are summarized in Table 5-6; Table 5-7 gives the numbers of the features selected in multiple systematic tests.



(a) (b)

Figure 5.10 Results of the Welch's *t*-statistics filter feature selection in a multi-day systematic test. (a) *p*-values of the features in the original order, (b) histogram of the *p*-values of the 1500 features.

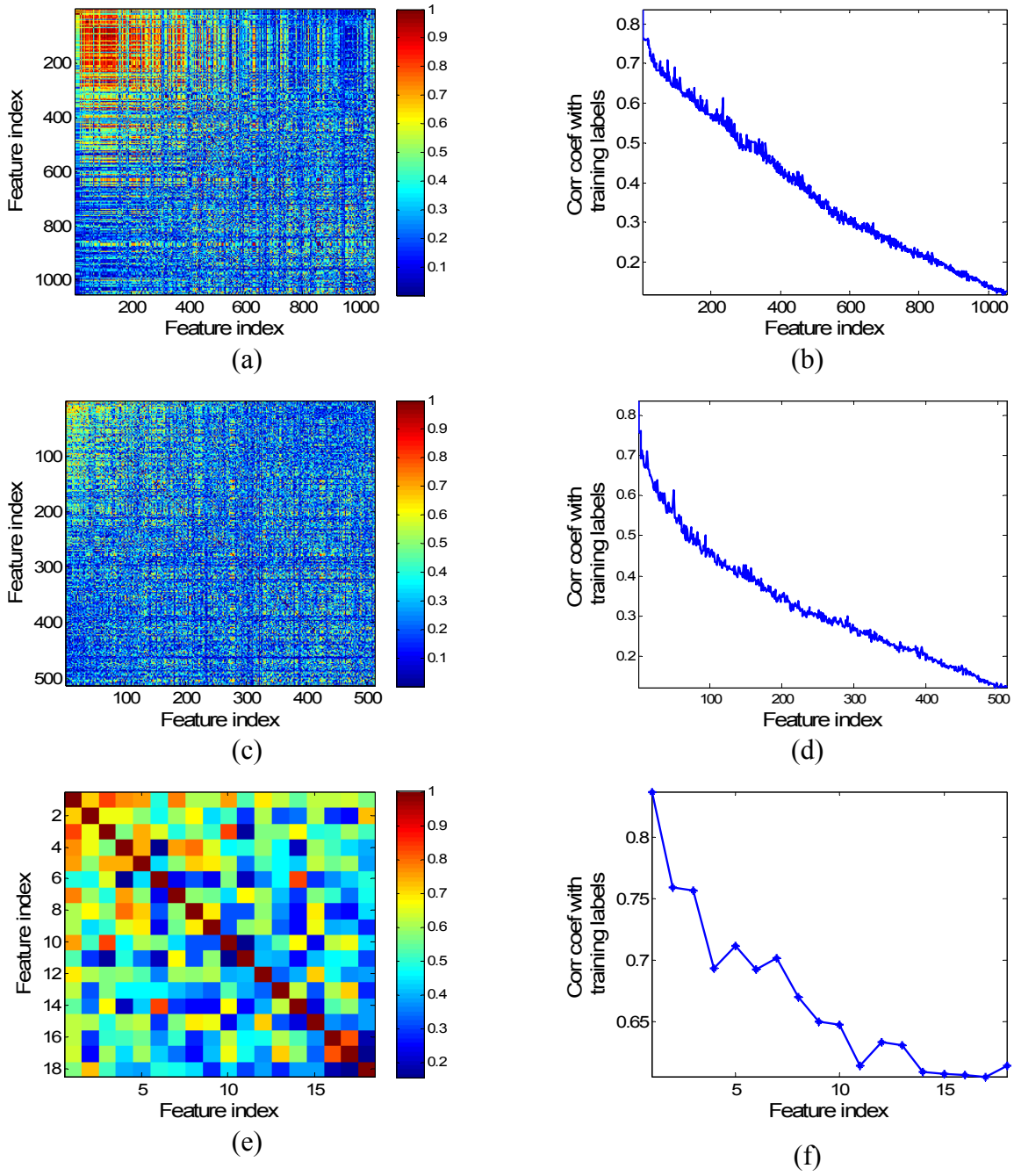


Figure 5.11 Absolute value feature correlation matrices of the features selected by different filters (plots in the left column), and the correlation coefficients of these selected features and the labels of the training data (plots in the right column), in a multi-day systematic test: (a)(b) results after the *t*-test, (c)(d) results after the R-filter, and (e)(f) results after the S-filter. (Features are sorted in an ascent order of *p*-value.)

Table 5-6 Range of the number of features selected in the random testing

		Collection I	Collection II	Collection III	Combined set
***t*-test**		900~975	906~1000	904~955	816~909
R-Filter		236~249	264~290	271~304	277~307
S-Filter		21~33	15~34	22~28	6~14
Wrapper-	**AdaBoost**	1	1	1	1
	M-AdaBoost	1	1	1	1
	Linear-SVM	2~11	1	1	1~5
	RBF-SVM	2~5	1	1	1~8

Table 5-7 Range of the number of features selected in the systematic testing

		Collection I	Collection II	Collection III	Combined set
t-test		1035~1163	1082~1124	1058~1104	945~1062
R-Filter		264~329	302~339	322~366	400~533
S-Filter		30~43	18~33	23~48	8~27
Wrapper-	**AdaBoost**	1	1	1	7~14
	M-AdaBoost	1	1	1	5~12
	Linear-SVM	1~6	1~2	1	3~12
	RBF-SVM	1~12	1~2	1	4~11

To visualize the performance of the selected features, Figure 5.12 provides examples of normalized 2-dimensional feature spaces defined by the selected features in the systematic testing on each of the data collections as well as the combined set of the three collections. In

these examples, the axes are defined by the first two feature selected by the wrapper method trained on linear-SVM; if a single feature is identified by the wrapper method, the y-axis is then defined by the first feature selected by the S-filter. As shown in Figure 5.12, the selected features significantly differentiate the scatterer instances (crosses) from the undamaged instances (dots). The multi-day systematic testing is the most difficult task, due to the large variation between different days. As a result, an effective feature for one collection might not provide similarly good class-separation for another collection. The two classes as shown in Figure 5.12d are still distinguishable, but with some overlapping occurring, indicating the necessity of using multiple features and feature selectors in the data-driven approach.

In addition, hybrid features extracted by combining different signal processing techniques and machine learning methods generally provide improved performance in distinguishing different classes, such as the features created by PCA, wavelet analysis, and the Mellin transform. To interpret the effectiveness of the PCA-wavelet features, as an example, we apply PCA to the wavelet coefficients of the Collection I data. The projections of the wavelet coefficients onto the first 10 PCs (10 orthogonal directions showing the largest data variations) are plotted in Figure 5.13a. Each curve represents the 260 measurements projected onto one PC. We observe that the projections onto the first PC shows a cyclic pattern, similar to Figure 5.3a. As discussed previously, such cyclic pattern reflects the pressure variations in the pipe. By contrast, we see "toggling steps" in the projections onto the seventh PC, which are consistent with the experimental procedure of applying and removing of the mass scatterer. Similar results are observed by applying PCA directly to the time domain signals (as shown in Figure 5.3). Therefore, we show that PCA can reveal the intrinsic dimensions of the data sets. Moreover, the PCA projections of the wavelet coefficients show clearer patterns with less noise.

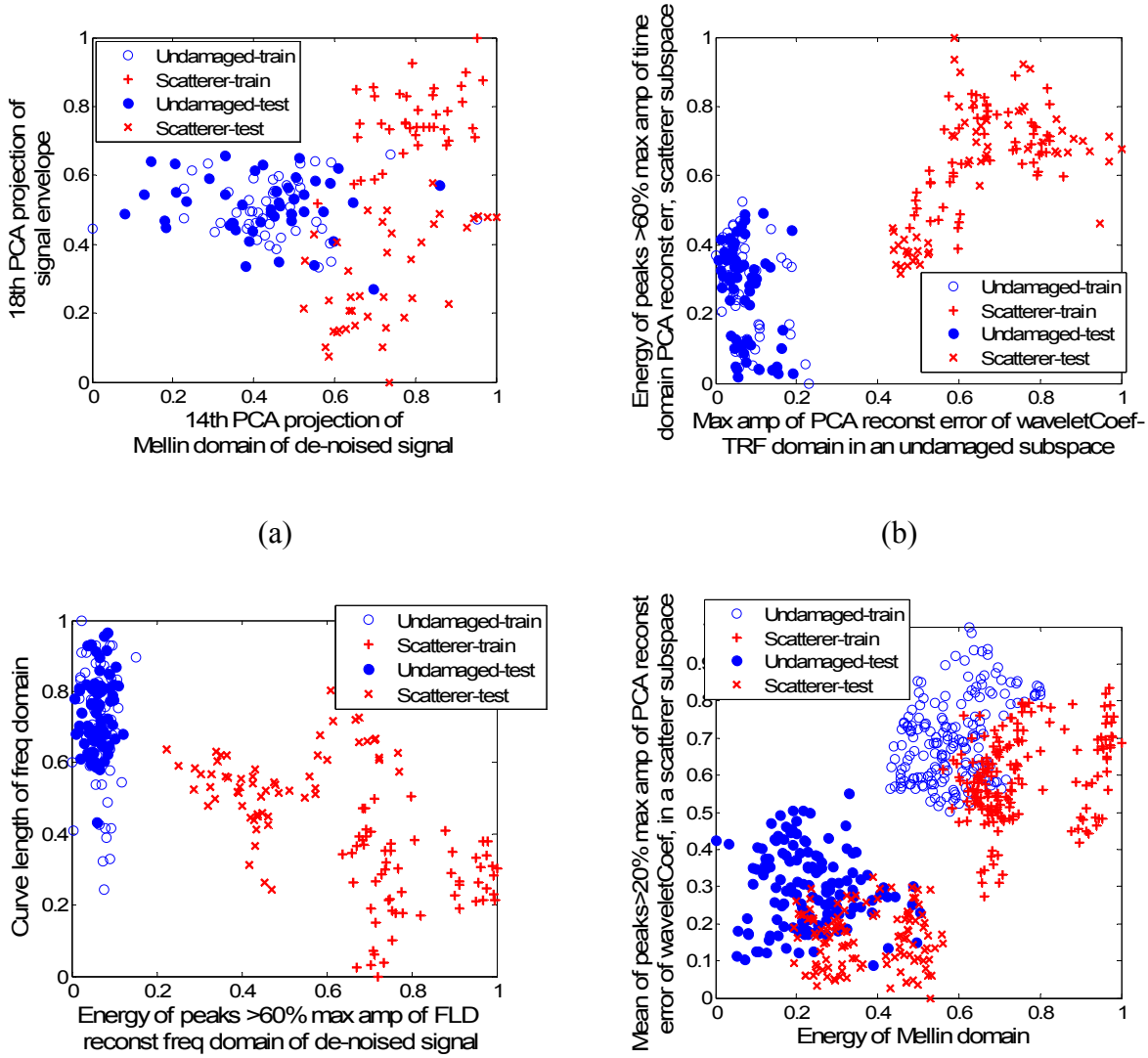
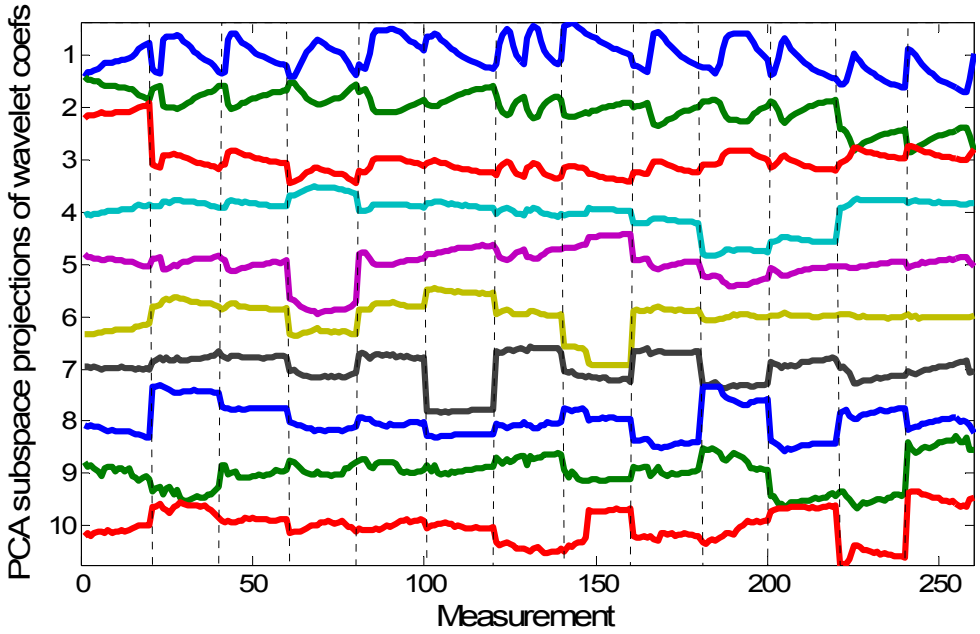
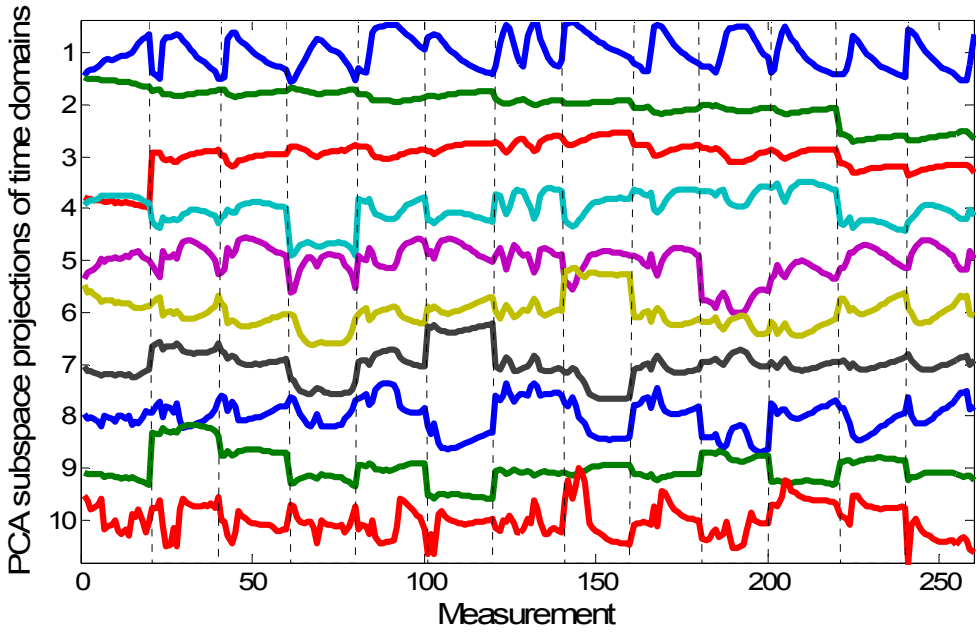


Figure 5.12 Examples of a normalized feature space defined by the first two features selected in the systematic testing on: (a) Collection I, (b) Collection II, (c) Collection III, and (d) combination of the three collections.



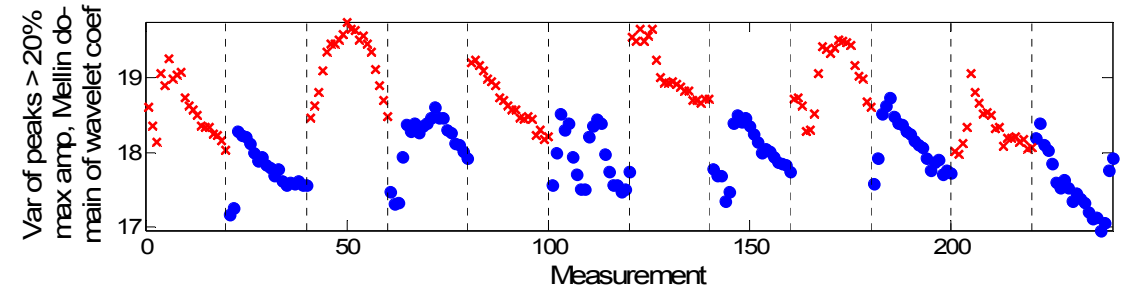
(a)



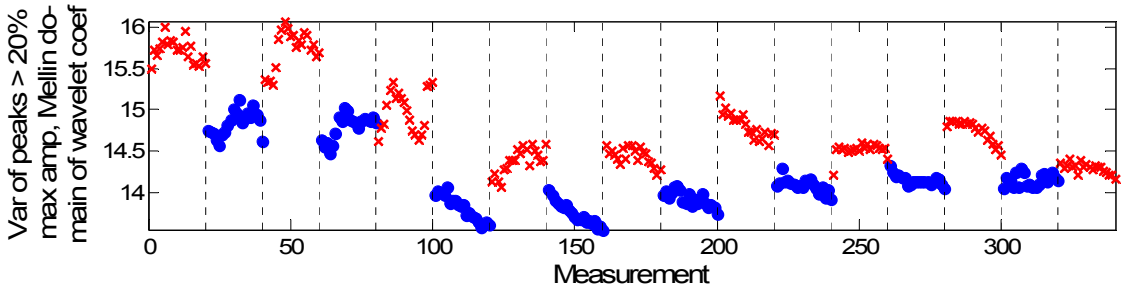
(b)

Figure 5.13 Projections of (a) the wavelet coefficients and (b) the time domain signals in a PCA subspace (Collection I data).

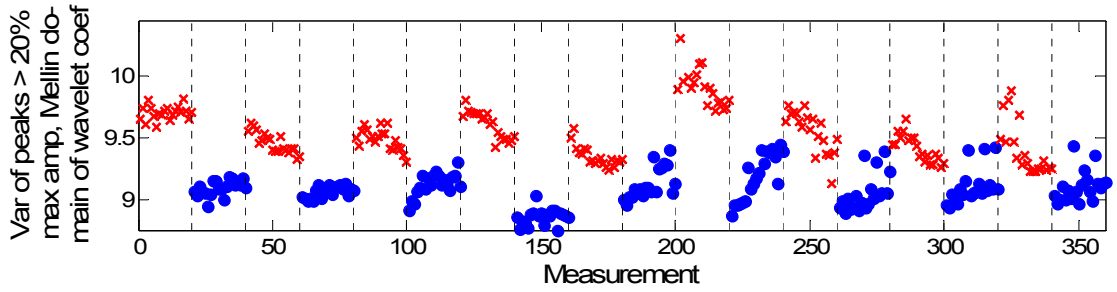
As described in the data-driven framework, the feature selection process can recommend good features as change detectors for monitoring purposes. Figure 5.14 plots all the measurements of the three collections using one selected baseline-free feature: the sample variance of peak values greater than 20% of the maximum amplitude in the Mellin domain of wavelet coefficients. Unlike the energy norm of the differential signals (Figure 5.3), this Mellin-wavelet feature is less sensitive to the environmental variations and introduces a noticeable departure from the undamaged instances (dots) to the scatterer instances (crosses). Wavelet processing is advantageous to reveal local discontinuities of a signal; the scale domain created by the Mellin transform is invariant to scaling changes in the signal. The hypothesis to explain the effectiveness of the Mellin-wavelet feature is that the pressure variations in the pipe alter the ultrasonic wave velocity, presented as scaling effects in the signal; the Mellin-wavelet domains create features that allow for more accurate change detection in the presence of those effects. To verify this hypothesis, slices of 20μs time domain signals are taken to inspect the detailed changes caused by pressure variations and by the mass scatterer. An undamaged record under 110 PSI (Figure 5.15a, solid line) and a scatterer record under 0 PSI (Figure 5.15b, solid line) are compared to a baseline signal under 0 PSI (Figure 5.15, dashed line). It can be observed that both the pressure increment and the mass scatterer produce changes in the wave amplitude, while the pressure variation introduces additional scaling changes in the signal.



(a)

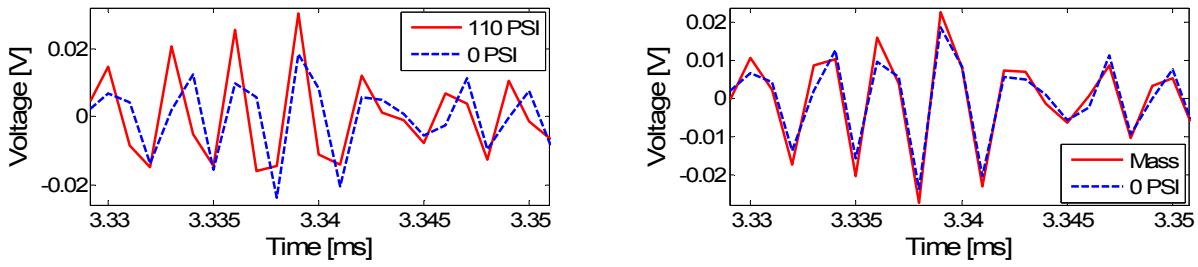


(b)



(c)

Figure 5.14 Sample variance of peak values greater than 20% of the maximum amplitude in the Mellin domain of wavelet coefficients: (a) Collection I, (b) Collection II, and (c) Collection III.



(a) (b)

Figure 5.15 Signal variation due to (a) pressure change, and (b) the mass scatterer.

5.6 Pattern Recognition for Scatterer Detection and Localization

To detect and localize damage in a pipe, four classification approaches were utilized: adaptive boosting (AdaBoost), modified AdaBoost (M-AdaBoost), support vector machines with a linear kernel function (linear-SVM), and support vector machines with a radial basis function as kernel (RBF-SVM).

This section first introduces the boosting and SVM algorithms and the machine learning evaluation metrics, and then reports and discusses the results of scatterer detection and localization in a pressurized pipe based on experimental investigations.

5.6.1 Adaptive Boosting Algorithms

Boosting is an ensemble approach that linearly superposes a number of weighted "weak" binary classifiers to generate a final "strong" classifier with an improved classification performance. The weak learners, also known as base classifiers, are usually simple and moderately inaccurate, as long as their error rate is less than 0.5. Each weak classifier focuses on the instances that are misclassified by the previous classifier. The final classifier is a sequential

combination of all the weak classifiers, where the ones with higher accuracy are assigned more weight (Schapire 2003). Adaptive boosting (AdaBoost) is perhaps the most commonly used boosting algorithm at present. Developed by Freud and Schapire in 1995, AdaBoost iteratively learns and adjusts the weight of each weak classifier according to the weighted error rate computed after every iteration.

Suppose that we have an input training set consisting of N data points, $(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \cdots, (\mathbf{x}_N, y_N)$, where $\mathbf{x}_i$ is an M-dimensional vector representing M features of the i-th data point, and y_i is the label of this instance (+1 or -1). In the AdaBoost implementation (Schapire 2003), we want to devise T weak classifiers $h_1, h_2, \cdots, h_T$, where $h_t \in \{-1, 1\}$ corresponds to the T selected features, as well as to determine their weights $\alpha_1, \alpha_2, \cdots, \alpha_T$. We first initialize the distribution over the training set as:

$$D_1(\mathbf{x}_i) = 1/N \tag{5-15}$$

We then exhaustively search all the possible weak classifiers over all the features and all the possible thresholds (in our study, we define the weak classifier as a threshold function) to find the one with the least weighted error ε_t, given by

$$\varepsilon_t = \sum \left\{ D_t(\mathbf{x}_i) \cdot \tfrac{1}{2} \left[1 - y_i h_t(\mathbf{x}_i) \right] \right\} \tag{5-16}$$

The error rate is used to compute the weight α_t of the t-th weak classifier h_t:

$$\alpha_t = \tfrac{1}{2} \ln \left[(1 - \varepsilon_t) / \varepsilon_t \right] \tag{5-17}$$

The lower the error, the higher the weight of the classifier; α_t and h_t are then used to update the distribution over the training set as:

$$D_{t+1}(\mathbf{x}_i) = D_t(\mathbf{x}_i)\exp\left[-\alpha_t y_i h_t(\mathbf{x}_i)\right] \tag{5-18}$$

After T iterations, the final classifier is obtained as below, as well as T sequentially selected features.

$$H(\mathbf{x}) = sign\left[\sum_t \alpha_t h_t(\mathbf{x})\right] \tag{5-19}$$

Traditional AdaBoost allows the weak classifiers to be selected repeatedly; this work made slight modifications to the algorithm to avoid selecting repeated weak classifiers. This method is referred to as the modified AdaBoost (M-AdaBoost). In other words, M-AdaBoost removes one dimension of the feature vector after the corresponding feature is selected to compose the weak classifier in each iteration. As a result, M-AdaBoost is more computationally efficient than AdaBoost due to gradually reducing the feature dimensionality.

5.6.2 Support Vector Machines

The SVM algorithm was first introduced by Vapnik (Cortes and Vapnik 1995) as a data classification method and has gained the popularity in the past decade. SVM, as a linear maximum margin classifier, constructs a hyperplane (decision boundary) or set of hyperplanes to maximize the separation between different classes; thus the generalization error of the classifier can be minimized in principle (Burges 1998).

The basic concept of SVM and the formulation of the maximum margin optimization are illustrated in Figure 5.16 (Burges 1998). A margin is defined in SVM to be the smallest distance between the hyperplane and any of the data points. The data points that constrain the width of the margin are called support vectors (SVs). The objective of the classifier is to determine a decision boundary to maximize the margin. Suppose that we have an input training set consisted of N data points, $(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \cdots, (\mathbf{x}_N, y_N)$, where $\mathbf{x}_i$ is an M-dimensional vector representing M features of the i-th data point, and y_i is the label of this instance (+1 or -1). We define $\mathbf{w}$ as a vector orthogonal to the decision boundary, and b as a scalar "offset" term (Burges 1998).

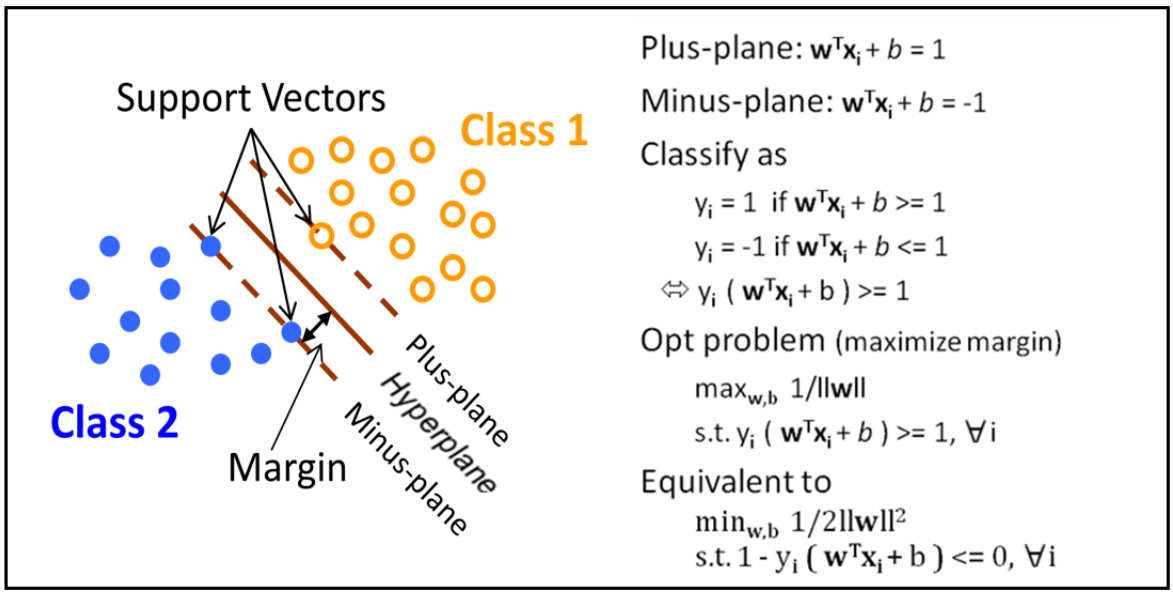


Figure 5.16 Formulation of the maximum margin optimization (Burges 1998).

A Lagrangian function is employed to form a dual representation of the maximum margin problem. The dual representation is given by,

$$\max_{\mathbf{a}} \tilde{L}(\mathbf{a}) = \sum_{i=1}^{N} a_i - \frac{1}{2}\sum_{i,j=1}^{N} a_i a_j y_i y_j \left(\mathbf{x}_i^T \mathbf{x}_j\right) \quad (5\text{-}20)$$
$$s.t.\ \ a_i \geq 0, i = 1,\ldots,N$$
$$\sum_{i=1}^{N} a_i y_i = 0$$

where a_i are Lagrange multipliers, and can be used to retrieve **w** along with the training data sets as follows,

$$\mathbf{w} = \sum_{i \in SV} a_i y_i \mathbf{x}_i \,. \quad (5\text{-}21)$$

The inner product $\mathbf{x}_i^T \mathbf{x}_j$ is known as a linear "kernel", and can be generalized to a nonlinear decision boundary, by replacing $\mathbf{x}_i^T \mathbf{x}_j$ with a nonlinear kernel function $K\left(\mathbf{x}_i, \mathbf{x}_j\right) = \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j)$ in the dual form. Without explicitly computing the function $\phi(.)$, we can transform the data points to a higher dimensional space where classification can be performed easily using a hyperplane (see Figure 5.17). SVM is probably the most well-known kernel-based algorithm that is favorable to achieve feature mapping without paying additional computational effort. In principle, there are an infinite number of kernel functions that can be applied. The most commonly used kernel functions are linear, polynomial, radial basis function, and sigmoid, as listed in Table 5-8 (Fan et al. 2005).

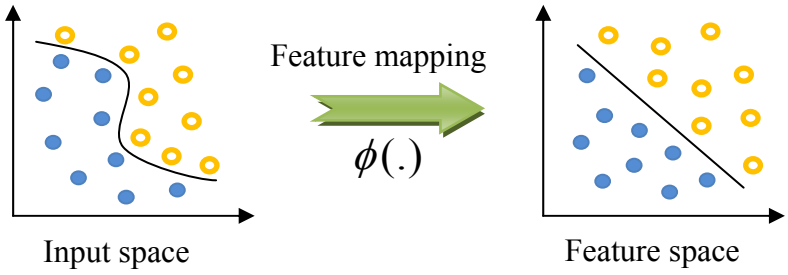


Figure 5.17 Classification becomes easier by proper feature mapping.

Table 5-8 Common types of kernel functions (Fan et al. 2005).

• Linear:	$K\left(x_i, x_j\right) = x_i^T x_j$
• Polynomial:	$K\left(x_i, x_j\right) = \left(\gamma x_i^T x_j + r\right)^d, \gamma > 0$
• Radial basis function (RBF):	$K\left(x_i, x_j\right) = \exp\left(-\gamma \left\|x_i - x_j\right\|^2\right), \gamma > 0$
• Sigmoid:	$K\left(x_i, x_j\right) = \tanh\left(\gamma x_i^T x_j + r\right)$

If we relax the hyperplane decision boundary by allowing errors in classification while assigning penalty to the misclassified instances, the algorithm becomes a soft margin SVM, as formulated below (Xing 2010),

$$\begin{aligned} &\min_{\mathbf{w},b} \frac{1}{2}\mathbf{w}^T\mathbf{w} + C\sum_i \xi_i \\ &s.t.\ \ \mathrm{y}_i\left(\mathbf{w}^T\mathbf{x}_i + b\right) \geq 1 - \xi_i, \forall i \\ &\xi_i \geq 0, \forall i \end{aligned} \tag{5-22}$$

where ξi are "slack variables", or "errors" in classification, and C is a tradeoff parameter between error and margin. The dual form of Equation (5-22) is written as,

$$\begin{aligned} &\max_{\mathbf{a}} \tilde{L}(\mathbf{a}) = \sum_{i=1}^{N} a_i - \frac{1}{2}\sum_{i,j=1}^{N} a_i a_j y_i y_j K\left(\mathbf{x}_i, \mathbf{x}_j\right) \\ &s.t.\ \ 0 \leq a_i \leq C, i = 1, \ldots, N \\ &\sum_{i=1}^{N} a_i y_i = 0 \end{aligned} \tag{5-23}$$

SVM is essentially a convex optimization problem that can be solved by quadratic programming. Currently, there are many software packages available for applying SVM. In this study, soft margin SVM with a linear kernel was applied using an online software package

LIBSVM developed by Chang and Lin (2011). The trade-off parameter *C* was determined by searching over grid points sampled at $2^{-5}, 2^{-4}, ..., 2^{15}$ for the highest accuracy in a 5-fold cross-validation within a training set. In 5-fold cross-validation, the datasets were randomly partitioned into five subsets; four subsets were used to train the SVM model, the remaining subset was used for validation of the model; such process was repeated five times, with each of the five subsets as the validation data.

5.6.3 Metrics for Machine Learning Classifiers

In this study, the reliability of the machine learning classifiers was measured by accuracy (ACC), false positive rate (FPR), false negative rate (FNR), and the area under the Receiver Operating Characteristics (ROC) curve (AUC). Higher ACC and AUC, and lower FPR and FNR indicate higher reliability of the classifier. ACC, FPR, and FNR metrics can be derived from a confusion matrix (Kohavi and Provost 1998), as shown in Table 5-9.

Table 5-9 Confusion matrix.

		Predicted	
		Damaged	**Undamaged**
Actual	**Damaged**	True positive (TP)	False negative (FN)
	Undamaged	False positive (FP)	True negative (TN)

- **ACC** is the percentage of the correct detections, including both correctly detected defects and non-defects, among all testing cases. It shows the overall correctness of the approach.

$$ACC = (TP + TN) / (TP + TN + FP + FN) \quad (5\text{-}24)$$

- **FNR** is the rate of occurrence of mistaking damage as undamaged cases. This term indicates whether the classifier tends to miss a defect that may cause a significant accident without providing an alert.

$$FNR = FN / (TP + FN) \tag{5-25}$$

- **FPR** is the percentage of false positive cases (false alarms). False alarms will result in unnecessary inspection and maintenance.

$$FPR = FP / (FP + TN) \tag{5-26}$$

FPR and FNR are dependent: decreasing one will result in increasing the other. An ROC graph depicts such a relationship by plotting the true positive rate (TPR = 1-FNR) along FPR as the decision threshold of a binary classifier varies. An ROC curve has practical values in adjusting the algorithms to determine a trade-off between the costs (FP) and benefits (TP) (Fawcett 2004). We created the ROC curves by performing a linear scan of test instances sorted in descending order by classification scores of the instances (Algorithm 2 described in (Fawcett 2004)). A higher classification score represents higher probability that an instance is a positive (the mass scatterer). The scores in the AdaBoost algorithms are defined as the weighted counts $\sum_t \alpha_t h_t(\mathbf{x})$; the scores in the SVM are the probabilistic outputs from the LIBSVM program (the probability is computed based on an algorithm described in (Platt 1999)). The closer an ROC curve is to the upper-left corner, the better the performance of a classifier is. In addition, we computed the area under the ROC curve (AUC) as another metric to evaluate the classifiers. The AUC varies between 0 and 1; the greater the area under the ROC curve, the higher the predictive

power of a classifier. The AUC has several advantageous properties compared to ACC, such as invariance to the discrimination threshold of an algorithm, and insensitivity to the changes in class distribution (Bradley 1997).

5.6.4 Results of Scatterer Detection

For each random or systematic test as formulated in Table 5-2, three feature selection methods and four machine learning classification algorithms were applied (results are summarized in Table 5-10), resulting in 12 different pairs of selector-classifier algorithms. To evaluate the performance of the 12 algorithm pairs for each data collection, the ACC and AUC of each selector-classifier pair were averaged over the six tests conducted within that data collection by permuting the roles of a basis set, a training set and a testing set.

Table 5-11 and Table 5-12 present the best algorithm pairs that achieve the highest average ACC (or AUC) within each collection, as well as the combined set of the three collections, in the random testing and the systematic testing, respectively. The results show that by choosing an appropriate selector-classifier pair, all the test datasets can be correctly classified, with 100% ACC (also 100% AUC), for any of the random tests on any of the three collections or the mixed set. Similarly, greater than 99% average ACC (100% AUC) can possibly be achieved for the single-day systematic testing, and an 84.33% average ACC (91.61% average AUC) can be obtained for the multi-day systematic testing.

However, the best algorithm pair varies under different test scenarios. For instance, Table 5-11 and Table 5-12 show that the R-filter and M-AdaBoost combination generally achieves the best detection results under most of the testing scenarios, whereas the R-filter and RBF-SVM combination provides superior performance in the multi-day systematic testing. Given different

detection or localization tasks, and numerous uncertainties in real-world operational conditions, different pipe segments, and varieties of damage, identifying a best algorithm pair may not be practical or generalizable in reality. Therefore, the average performance of a classifier regardless of the chosen feature selectors is evaluated in Figure 5.18.

Table 5-10 Summary of the applied algorithms.

		Algorithm
Feature selector	Filter	• A *t*-test filter applied relaxed correlation threshold (R-filter) • A *t*-test filter applied strict correlation threshold (S-filter)
	Wrapper	• Wrapper backward sequential selection (wrapper)
Classifier	Boosting	• Adaptive boosting (AdaBoost) • Modified AdaBoost (M-AdaBoost)
	SVM	• SVM with a linear kernel (linear-SVM) • SVM with a Gaussian radial-basis-function kernel (RBF-SVM)

Table 5-11 The best algorithm pairs of a feature selector and a classifier in the random testing for each collection and the mixed set, and the resulting average AUC and ACC.

	Algorithms applied	Results
Collection I	Any filter with M-AdaBoost; or wrapper with RBF-SVM	AUC (100%)
	Any feature selector with any boosting; or wrapper with RBF-SVM	ACC (100%)
Collection II	Any feature selector with any SVM; or any filter with M-AdaBoost	AUC (100%)
	Any feature selector with any boosting; or R-filter with linear-SVM; or wrapper with RBF-SVM	ACC (100%)
Collection III	Any filter with M-AdaBoost or any SVM	AUC (100%)
	Any feature selector with any classifier	ACC (100%)
Combined set	Any feature selector with any SVM; or R-filter with M-AdaBoost	AUC (100%)
	Any feature selector with any boosting; wrapper with any SVM	ACC (100%)

Table 5-12 The best algorithm pairs of a feature selector and a classifier in the systematic testing for each collection and the mixed set, and the resulting average AUC and ACC.

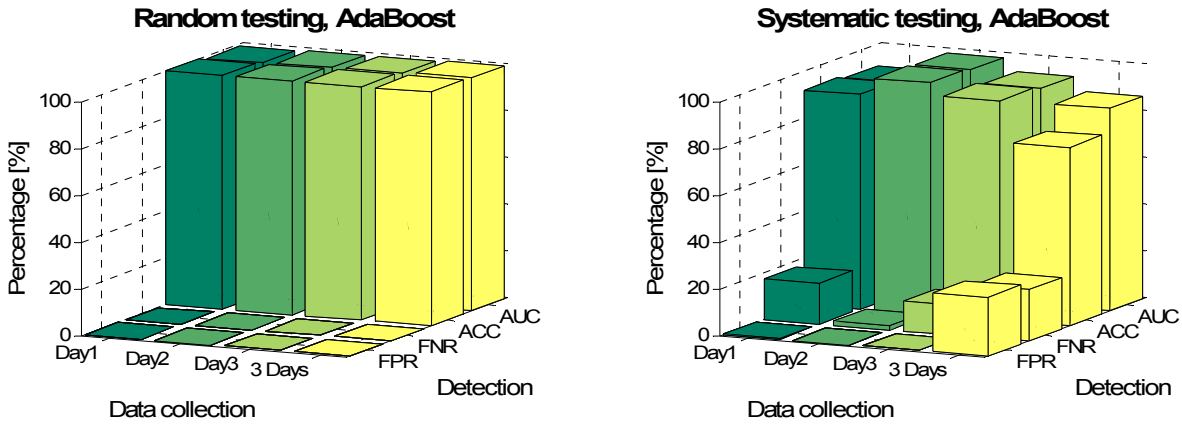
	Algorithms applied	Results
Collection I	R-filter with M-AdaBoost	AUC (100%)
	R-filter with M-AdaBoost	ACC (99.04%)
Collection II	R-filter with M-AdaBoost	AUC (100%)
	Wrapper with linear-SVM	ACC (99.44%)
Collection III	Any filter with M-AdaBoost	AUC (100%)
	R-filter with M-AdaBoost	ACC (100%)
Combined set	R-filter with RBF-SVM	AUC (91.61%)
	R-filter with RBF-SVM	ACC (84.33%)

To measure the average detectability of a classifier for a data collection in the random or systematic testing, the scatterer detection results (AUC, ACC, FPR, and FNR) are averaged over the tests with any of the three feature selectors. Therefore, the averaged results represent the reliability of the classifier irrespective of which feature selection method is used (as shown in Figure 5.18).

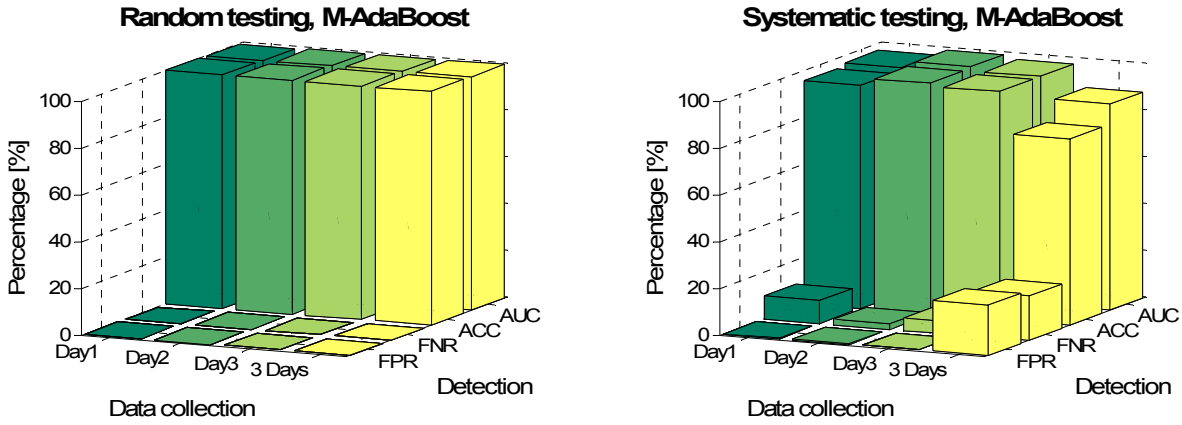
In the random testing, the boosting algorithms correctly classified all the testing instances, achieving 100% average accuracy for each of the three collections and the combined set. The linear-SVM classifier receives around 94% average accuracy for Collections I and II, and 100% and 99.97 % average accuracy for Collection III and the combined set, respectively. The RBF-SVM algorithm obtains 94% average accuracy for Collection I, and greater than 99.7% average accuracy for the other two collections and the mixed set. The systematic testing results in an overall lower average accuracy compared to that of the random testing. All four classifiers achieve around 95% average accuracy for the single-day systematic testing, while the multi-day systematic testing is much more challenging, resulting in 74.5%~79.6% accuracy. M-AdaBoost shows overall better performance in the systematic testing.

Similarly, Figure 5.19 provides the average performances of the three feature selection methods regardless of the chosen classifiers. The results represent the capability of a selector in identifying a set of effective and reliable features. In the random testing, the features selected by any of the three algorithms result in very high average accuracy (all greater than 95%, with several above 99.5%), whichever classifier is used. In the systematic testing, the features selected by the R-filter lead to the best average accuracies of 96.55%, 97.8%, 81.75% for

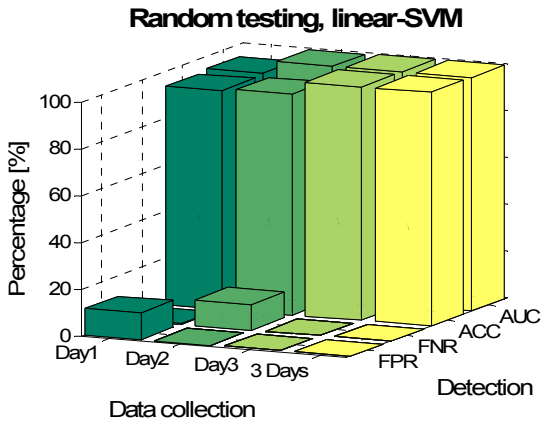
Collection I, III, and the combined set, respectively; the features selected by the wrapper method show superior results for Collection II, leading to 96.22% average accuracy.

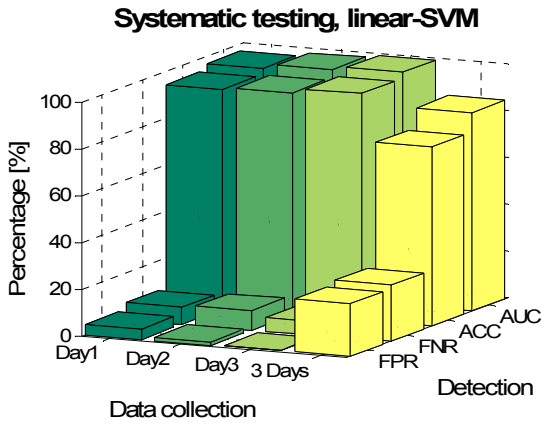


(a)

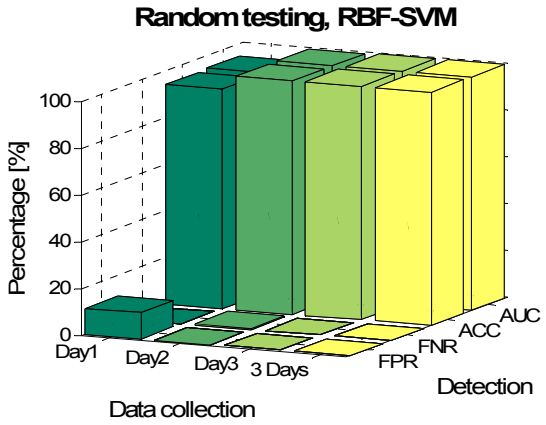


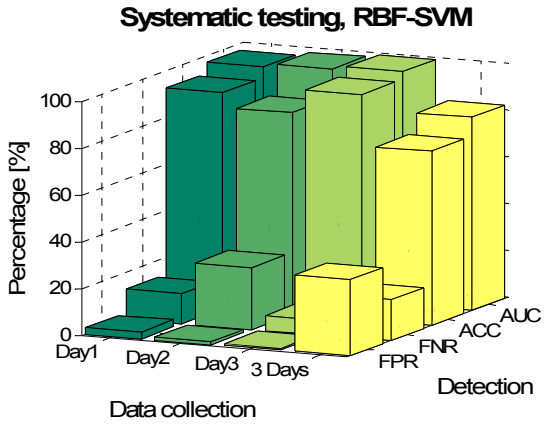
(b)





(c)

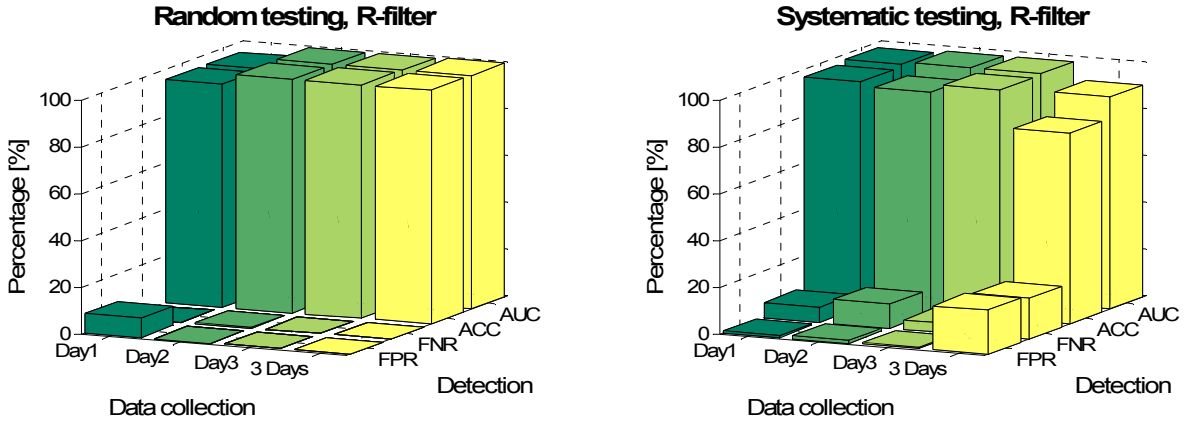




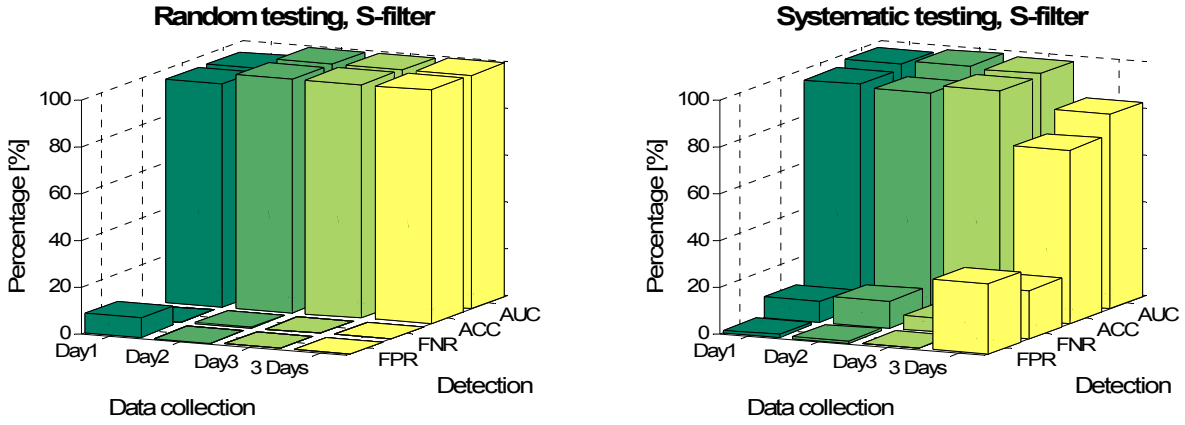
(d)

Figure 5.18 Comparison of the average performance of the classifiers for scatterer detection in the random testing (left column) and the systematic testing (right column): (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM.

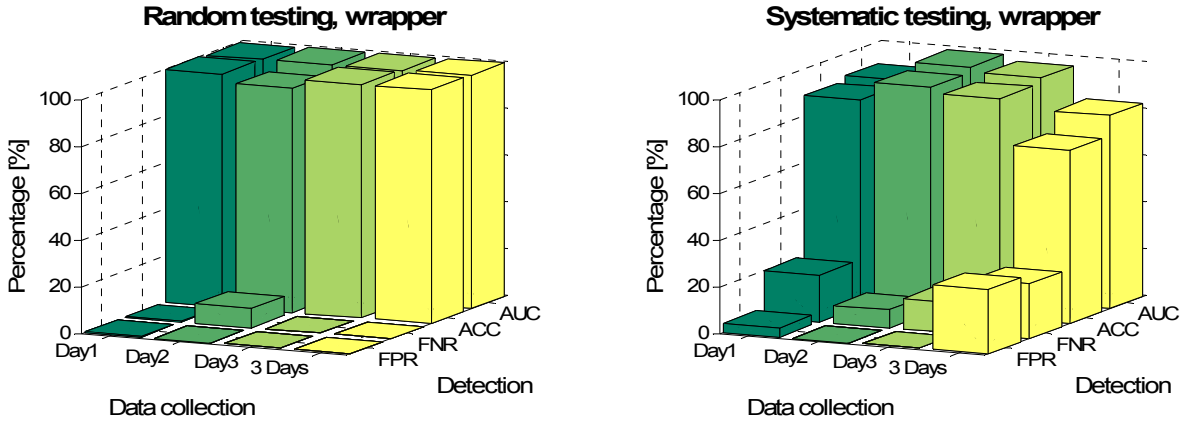
(Notation – AUC: area under the ROC curve; ACC: accuracy; FPR: false-positive rate; FNR: false-negative rate; Day1: Collection I; Day2: Collection II; Day3: Collection III; 3 Days: a combined set of the above three collections.)



(a)



(b)



(c)

Figure 5.19 Comparison of the average performance of the feature selectors in random testing (left column) and systematic testing (right column): (a) R-filter, (b) S-filter, and (c) wrapper.

5.6.5 Results of Scatterer Localization

In this work, determining the presence of a defect in a particular pipe segment is of interest. The resolution of localization can be adjusted by dividing (virtually) the pipe into a different number of zones and assigning different labels to each zone. A binary localization (to classify the presence of a scatterer in one of two possible zones, Zone 1 or Zone 3, as shown in Figure 5.1) is discussed in this dissertation. To further classify the mass positions into three zones, either a multi-class classifier or hierarchical binary classifier can be implemented.

The three data collections used in scatterer detection were also used to evaluate the data-driven method for scatterer localization. However, the data collection process was found to yield artifacts in the localization tests. The measurements of Zone 1 (with the scatterer at each of the three locations in Zone 1) were taken prior to the measurements of Zone 3. As a result, environmental and instrumental drifts over time were found to mask the signature signal characteristics produced by the scatterer in different locations. For instance, an environment-sensitive feature, the L2-norm of a differential signal, showed separation between Zone 1 and Zone 3 data in Collection II (as shown in Figure 5.3). However, such separation was mostly produced by the environmental and/or instrumental drift and seen as an overall increasing trend over time, rather than the difference caused by different scatterer locations.

To suppress such drift effects on the measurements of different zones, another set of data (Collection IV) were collected for the different mass locations in a random order. Six locations were chosen: three near the transmitter (Zone 1), and three close to the receiver (Zone 3). The Zone 1 and Zone 3 data were collected alternatively over time. For each mass position on the

pipe, 20 scatterer records were measured with varying internal air pressure ranging from 0 to 100 PSI.

For each random or systematic test (as formulated in Table 5-2), three feature selection methods and four machine learning classification algorithms were applied (as listed in Table 5-10), resulting in 12 different pairs of selector-classifier algorithms. Figure 5.20 and Figure 5.21 show the results of scatterer localization in the random testing and the systematic testing, respectively. Zone 1 instances are denoted as positives, while Zone 3 instances are denoted as negatives.

In the random testing, all the selector-classifier algorithms achieve greater than 99.5% ACC (above 97.6% AUC). In the systematic testing, applying RBF-SVM with the features selected by the wrapper method achieve the highest ACC, 95%, with a 97.5% AUC. However, AdaBoost achieves the best overall performance in the systematic localization tests, obtaining greater than 94% ACC (above 91.8% AUC) with any of the feature selectors. M-AdaBoost, linear-SVM, and RBF-SVM achieve similarly high ACC and AUC with the features selected by the wrapper method, while the localization results of these three classifiers degrade when paired up with the filter feature selectors.

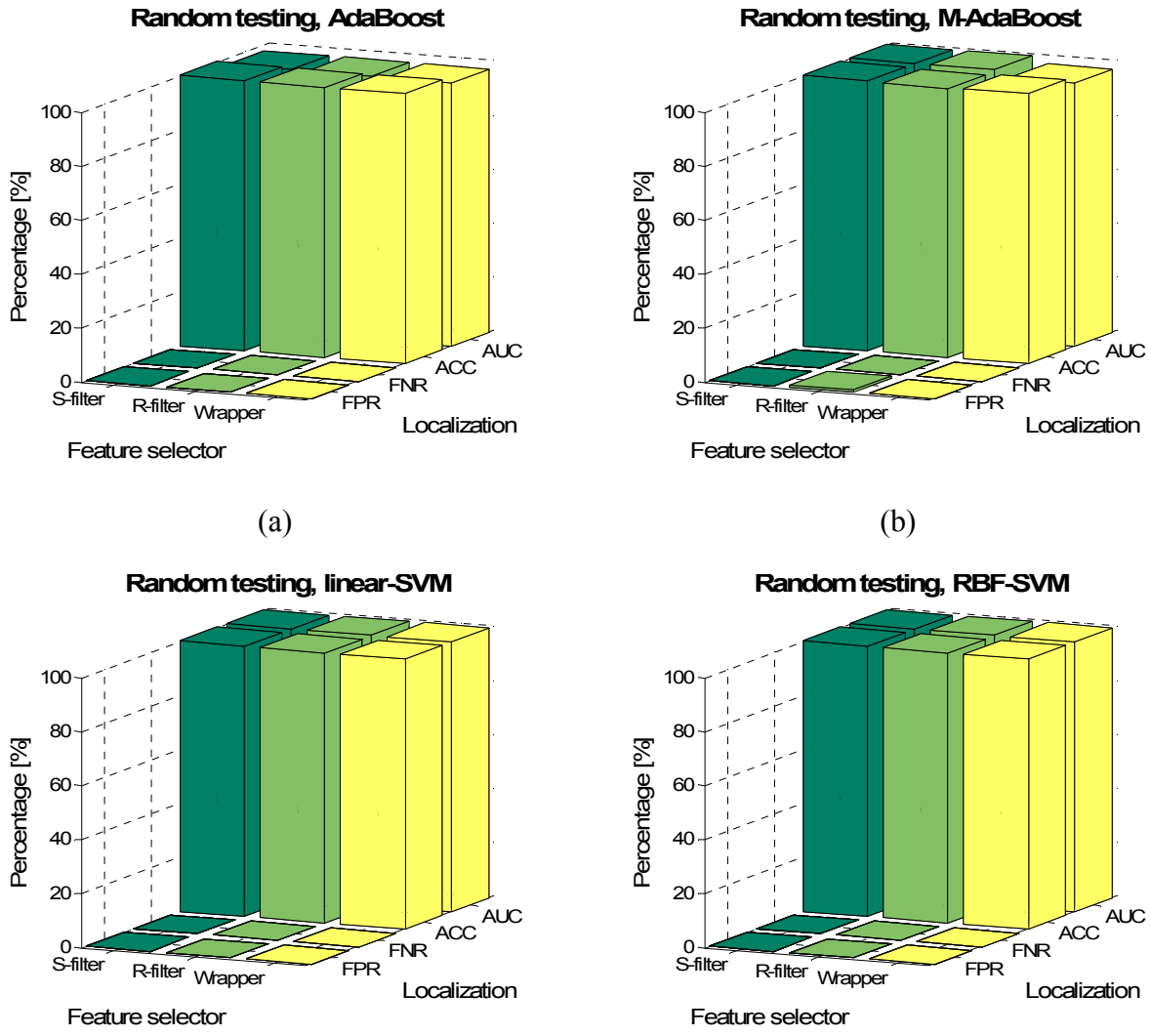


Figure 5.20 Scatterer localization results of the random testing on Collection IV, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM.

(ACC: accuracy; FPR: false-positive rate; FNR: false-negative rate; AUC: area under an ROC curve; R-filter: a relaxed-filter feature selector; S-filter: a strict-filter feature selector; and Wrapper: a wrapper sequential backward feature selector.)

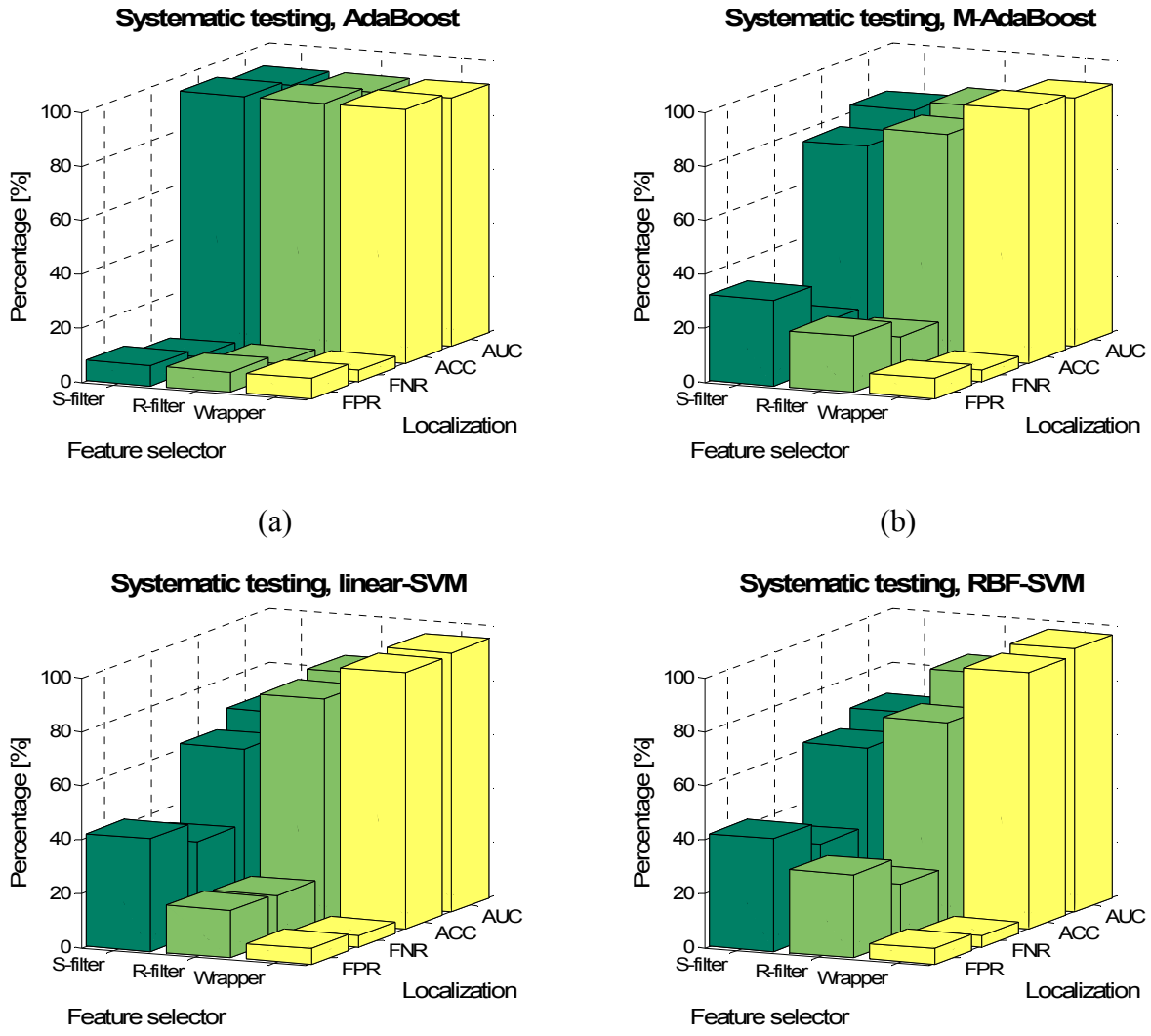


Figure 5.21 Scatterer localization results of the systematic testing on Collection IV, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM.

(ACC: accuracy; FPR: false-positive rate; FNR: false-negative rate; AUC: area under an ROC curve; R-filter: a relaxed-filter feature selector; S-filter: a strict-filter feature selector; and Wrapper: a wrapper sequential backward feature selector.)

5.7 Summary

The effectiveness and robustness of a data-driven framework for SHM were experimentally verified on a pipe under varying environmental and operational conditions. Laboratory tests were performed on a steel pipe with a mass scatterer placed at different locations and with benign variations in air pressure and ambient temperature. A total of 1500 features were extracted from the ultrasonic signals to create a feature library that might consist of effective as well as poor features. An integrated filter-wrapper feature selection approach automatically identified useful features for scatterer detection. Hybrid features constructed using combined signal processing and machine learning methods showed superior class-separability. The selected features also suggested robust change detectors for real-time monitoring purposes. Four classification methods (adaptive boosting, modified adaptive boosting, support vector machines with a linear kernel, and support vector machines with a radial basis function kernel) were investigated to determine the presence of the mass scatterer. All the classifiers achieved good results in the random testing and the systematic testing. The results show the promising applications of the integrated machine learning and signal processing methodology to pipe monitoring.

6 FIELD VALIDATIONS

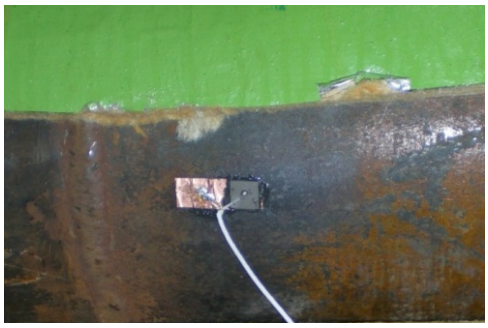
The objective of this research is to support automatic damage detection and localization in pipe monitoring. There are many different types of pipes in terms of diameters, wall thickness, ambient conditions, materials, and so on. Different types of pipes have different characteristics, which may influence the capability of the previously described approaches to detect and localize damage. If one approach can only be used under limited situations, it is not applicable for real-world deployment. As one step toward such a goal, and as a stricter test of the methods developed here, the data-driven SHM framework is studied in a series of field tests.

A guided wave ultrasonic sensing system was deployed in an in-service hot water supply system in a building, which allowed us to test the data-driven framework on a pipe under real operating conditions in a noisy environment. The following sections describe the field test environment and discuss the effects of different types of operational and environmental variations on the ultrasonic signals. During data collection, a grease-coupled mass scatterer was applied and removed, to simulate a "reversible" defect on the pipe. We demonstrate the ability of the data-driven approach to distinguish the change caused by the scatterer from change caused by ongoing variations.

6.1 Field Environment and Test Setup

The hot water supply system is located in Wean Hall, an eight-floor building on the campus of Carnegie Mellon University with a total area of roughly 290,000 sq ft and multiple mechanical spaces. The test segment is located in a three-story high mechanical room with a total area of 7,600 sq ft. That space contains mechanical components such as chillers, heat exchangers, pumps, and large-diameter pipes, with substantial mechanical and electrical noise.

(a)



(b)

(c)

Figure 6.1 (a) The pipe indicated with a dashed line is the hot water return pipe which was instrumented in our experiments, (b) PZT transducer mounted on the pipe wall, and (c) a cylinder mass with grease (or honey) attached to the pipe used to simulate damage (C. Liu et al. 2012).

Figure 6.1a shows the field environment and the pipe segment that was studied. It is a Schedule 40 steel pipe with 10 in. inner diameter and 0.365 in. wall thickness, covered by fiberglass insulation. We chose a straight section of the pipe (Figure 6.1a), because the radius bends were fabricated with bolted flanges that would not provide acoustic coupling (In other industrial applications, the bends may be welded, which would permit guided wave ultrasonics to

be employed). The pipe serves as a hot water return pipe that conveys pressurized hot water flow from the building's heat cycle back to the boiler. The heating cycle operates 24 hours a day, 7 days a week all year long. Due to the periodic pumping of hot water into the heat cycle, the hot water flow has a variable flow rate, ranging from 200 to 450 gpm, with the water temperature fluctuating from 100 to 140 ^{o}F (C. Liu et al. 2012).

We removed sections of the pipe insulation to expose the pipe surface, in order to install the transducers. The transducers used to monitor the pipes were PZT material (PSI-5A4E piezo-ceramic from Piezo System, Inc) cut into 15 mm by 15 mm wafers with thickness of 0.508 mm (Figure 6.1b). A transducer was first attached to a piece of copper foil and then mounted on the surface of the pipe by epoxy (J-B Weld 8265-S Cold Weld). The transducers were connected to a National Instrument (NI) PXI chassis for data collection. We used an NI 5421 Arbitrary Waveform Generator to excite a transmitting PZT with a sinc function with a frequency band of 200 kHz to 400 kHz. We recorded the signal from a receiving PZT with an NI 5122 Digitizer at a sampling rate of 1MHz.

Damage was simulated physically by a mass scatterer (Figure 6.1c) acoustically coupled to the pipe with vacuum grease or honey. Similar to the laboratory experiments described in Section 5.1, mass "toggling" tests were performed by applying and removing the scatterer subsequently. However, due to various access constraints, the scatterer locations were closely spaced within one pipe segment.

Table 6-1 summarizes the different experimental setup in the laboratory studies and the field tests.

Table 6-1 Comparison of the laboratory experiments and the field tests.

	Laboratory experiments	**Field tests**
Room environment	Mechanical and electrical noise significantly low	Mechanically and electrically noisy
Pipe dimensions	Inner diameter: 62 mm Wall thickness: 4 mm	Inner diameter: 10 in. (254 mm) Wall thickness: 0.365 in. (9.04 mm)
Content (fluid)	Air with varying pressure levels 0-110 PSI	Water with temperature fluctuating 100-140^{o}F, flow rate 200-450 gpm (related to pressure)
Transmitter-receiver distance	1.5 m	10ft (3.048m), 20ft (6.096m)
Mass location	6-9 spread locations for each collection	Closely spaced locations within one pipe segment due to access constraints
Test duration	1-3 hours for one collection; 4 collections months apart	10 hours to more than one day; the evaluated collections were taken about a weak or months apart
Datasets	Balanced data (approximately equal-sized undamaged instances and scatterer instances)	Mostly unbalanced datasets

6.2 Field Test Observations

Figure 6.2 shows a typical received ultrasonic signal in the field tests. Similar to the laboratory signal, we observe a multi-modal, complex wave pattern. In addition, the signal is apparently contaminated by large low-frequency noise due to mechanical vibrations that are removed using a high-pass filter.

As reported by Liu et al. (2012), we examined correlation coefficients to investigate the stationarity of the received ultrasonic signals over time, and observed an interesting structure. We calculated the correlation coefficients between each measurement signal and the first signal of the day. Figure 6.3 shows the correlation coefficients over a 20 hours period, varying from 1.0 to -0.2. Therefore, the varying correlation coefficients indicate that ambient environment and operational variations produce enormous changes in the pitch-catch signals. Furthermore, the graph of the correlation coefficients in Figure 6.3 clearly shows cyclic characteristics, in which the correlation coefficients drop dramatically and abruptly, and then climb back to near 1.0. The period of the cyclic effect varies but is most typically between 50 to 60 min. We hypothesize that the periodicity is due to the periodic pumping of hot water from the boiler into the pipe loop, which consequently changes the flow rate and temperature of the water in the pipe cyclically.

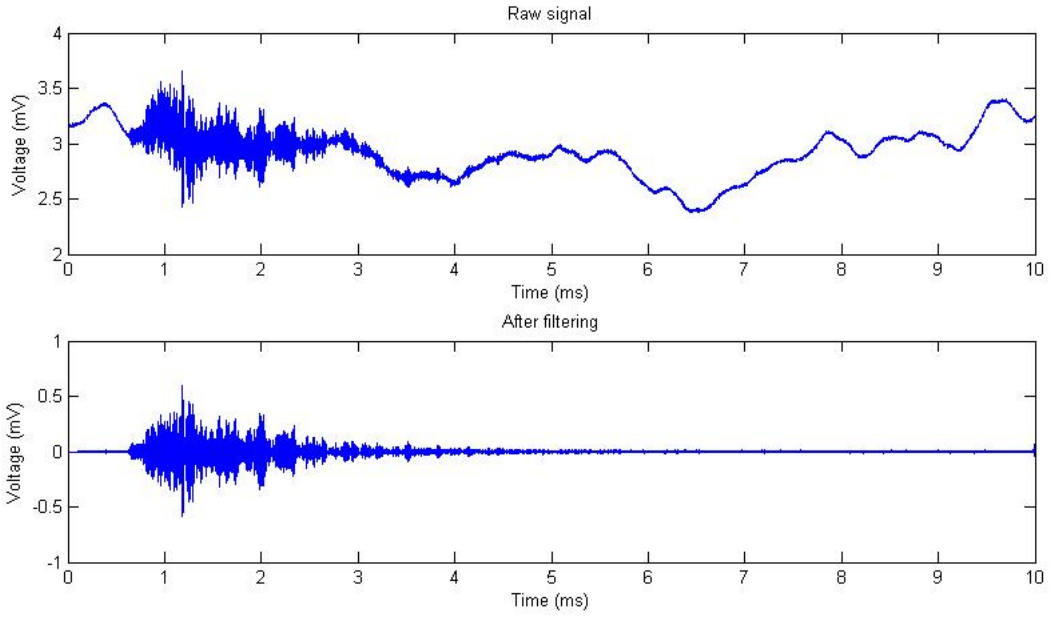


Figure 6.2 Received signal before and after applying a 125 kHz high pass filter (C. Liu et al. 2012).

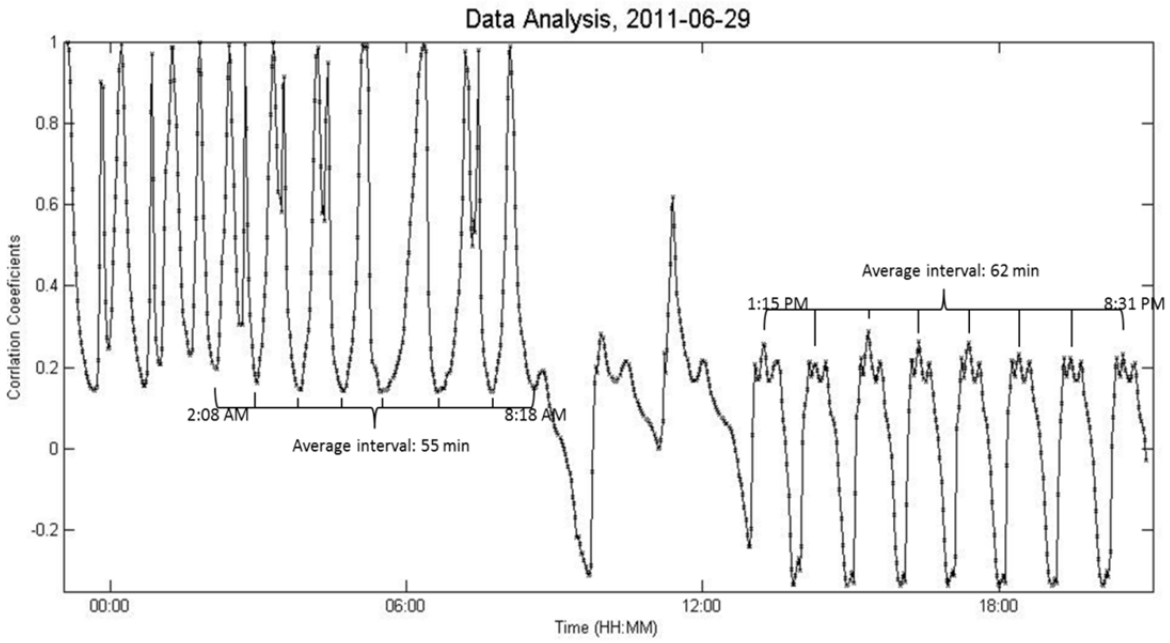


Figure 6.3 Correlation coefficients of 20-hr measurements with respect to the first measurement (C. Liu et al. 2012).

In the tests shown in Figure 6.4, a mass scatterer was applied and removed four times, respectively. However, the environmental and operational variations produce dramatic changes in the ultrasonic signals, and mask the effects of the mass scatterer. As a result, the scatterer is difficult to detect.

Another interesting observation from Figure 6.4 is that the flow rate shows similar and synchronized cyclic features with temperature, and that both align with the correlation coefficients. Note that in the hot-water pumping system the variation of the temperature results from the periodic pumping of hot-water from the boiler into the cycle, which explains why the two variations are correlated to one another. (In this real world system, we could not control the boiler and the pump for testing purposes, and therefore at this time it is not clear which of the

two changing variables was responsible for the change in signal correlation coefficients) (C. Liu et al. 2012).

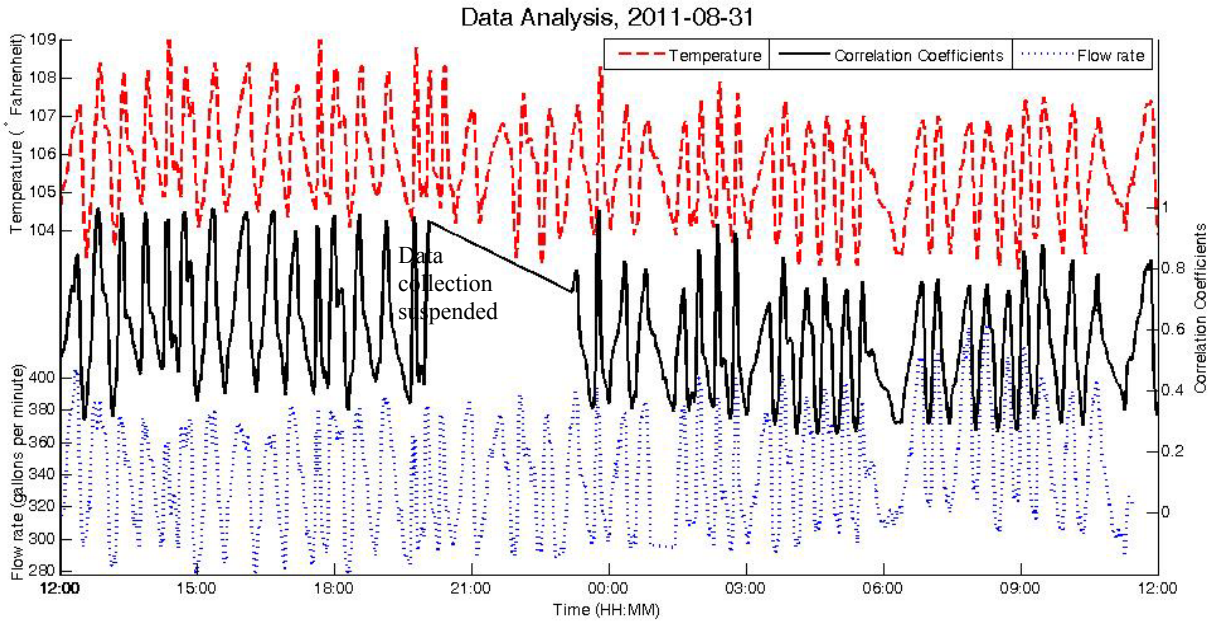


Figure 6.4 Correlation coefficients variation aligned with temperature and flow rate fluctuation (C. Liu et al. 2012).

6.3 Scatterer Detection Using Data-Driven SHM Framework

This section reports the results of scatterer detection and localization in the field tests by implementing the data-driven SHM framework developed and demonstrated in the preceding chapters. Four test scenarios are discussed in the section: the first two scenarios focus on evaluating the capability of the scatterer-detection methodology for the field data collected on different days; the third test scenario assesses the effectiveness of the data-driven framework for scatterer localization under field conditions; and the fourth scenario evaluates the generality of the data-driven methods for detecting different scatterers on the pipe.

6.3.1 Field Test 1: Detecting a Scatterer at Different Locations

6.3.1.1 Data Collection and Machine Learning Test Formulation

Field Test 1 was conducted on two different days, using the same transmitter-receiver pair, with the mass scatterer at multiple locations. The data collection details are provided in Table 6-2.

Collection F-I was acquired on 8/18/2011, from two and half cycles of scatterer toggling tests. In each cycle, the mass scatterer was applied and removed once, resulting in a half-cycle of undamaged measurements (with the absence of the scatterer on the pipe) and another half-cycle of scatterer measurements (with the mass scatterer placed at a fixed position on the pipe surface). A total of 2040 measurements were taken with about one-minute interval between each measurement, lasting 34hr 3min, consisting of 1839 undamaged measurements and 201 scatterer measurements.

Collection F-II was acquired on 8/23/2011: 418 measurements were taken without the presence of a scatterer, 1074 measurements were taken with the presence of a scatterer. The tests consisted of two toggling cycles, with a total duration of 24hr 53min.

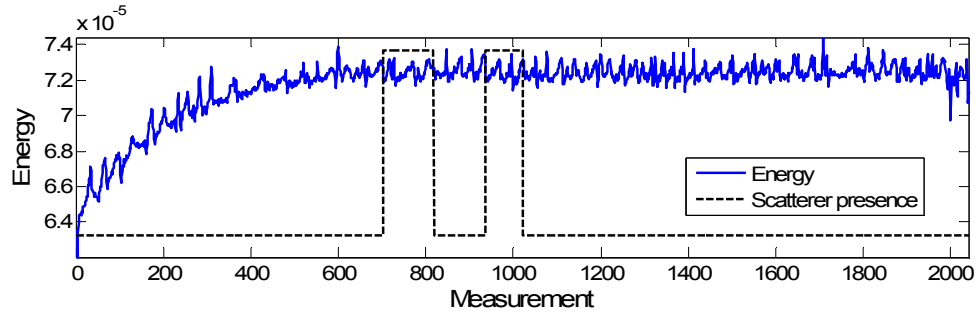
Unlike the lab experiments, the field tests were performed in long duration, dealing with unbalanced datasets (the number of the undamaged measurements is different from that of the scatterer measurements), and the changes caused by the environmental variations were more complex and dynamic.

Figure 6.5 shows the energy variations over time for the two field data collections. The energy shows cyclic patterns, with no visible indication of the introduction and removal of the

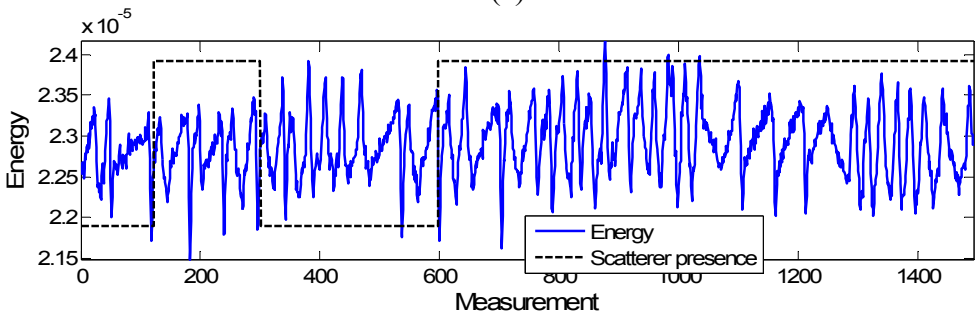
scatterer. The toggling cycles (i.e. placing and removing the scatterer) are indicated in dashed lines. Additionally, even with the same pair of transducers, the energy varies for the collections performed on two different days.

Table 6-2 Field data collection details (Field Test 1).

	Undamaged datasets	**Scatterer datasets**	**Test duration**	**Toggling cycles**	**Scatterer locations**
Collection F-I (8/18/2011)	1839	201	34hr 3min	2.5	2
Collection F-II (8/23/2011)	418	1074	24hr 53min	2	2



(a)



(b)

Figure 6.5 Energy variation over time of (a) Collection F-I, and (b) Collection F-II.

To evaluate the effectiveness of the data-driven framework for scatterer detection under the field conditions, two types of machine learning testing were formulated (Table 6-3). The random testing, similar to the lab studies (Section 5.3), was implemented on each of the collections, as well as a combined set of the two collections. Each data collection was randomly divided into three approximately equal-sized subsets, used as a basis set, a training set, and a testing set, respectively. Six random tests were performed for each collection and for the combined set, by switching the roles of the three subsets. The systematic testing was conducted differently from the lab studies, constrained by the number of scatterer locations in the tests and the number of collections with the same transducer pair (performed on different days). For systematic testing, one of the collections was used as a basis set, and the other collection was divided into a training set and a testing test based on the two different scatterer locations (cycles). The 2-fold test was performed within the second collection by switching the roles of the training and testing sets. Therefore, two systematic tests were conducted for each collection with the other collection as the basis set. A systematic test is a stricter evaluation of the predictability of a classifier; it answers two practical questions:

- Is a machine learning model, built upon the instances of the scatterer located at one location of the pipe, capable of recognizing the presence of the scatterer at different locations?
- Is a basis set consisted of datasets collected on a different day capable of generating effective baseline-dependent features for the current pipe condition assessment?

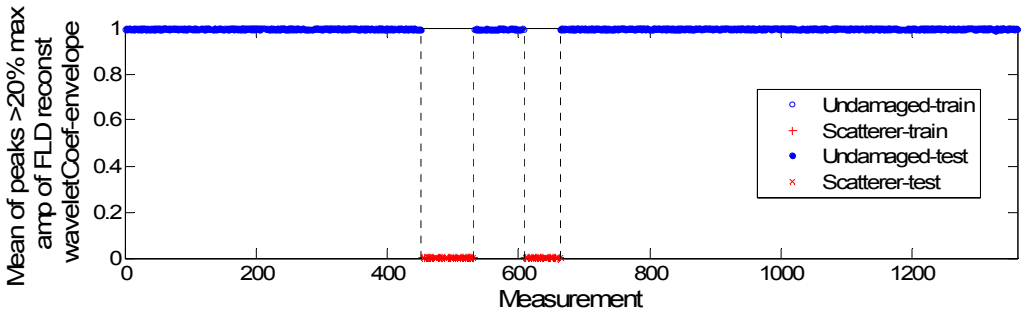
Table 6-3 Machine learning testing formulation for scatterer detection (Field Test 1).

	Single-day data	**Multi-day data**
Random testing	3 randomly divided, equal (or approximately equal) sized sets Total: 12 tests	3 randomly divided, equal (or approximately equal) sized sets Total: 6 tests
Systematic testing	One collection as a basis set, 2-fold systematic testing on the other collection (divided into training and testing sets according to different **scatterer locations**). Total: 4 tests	

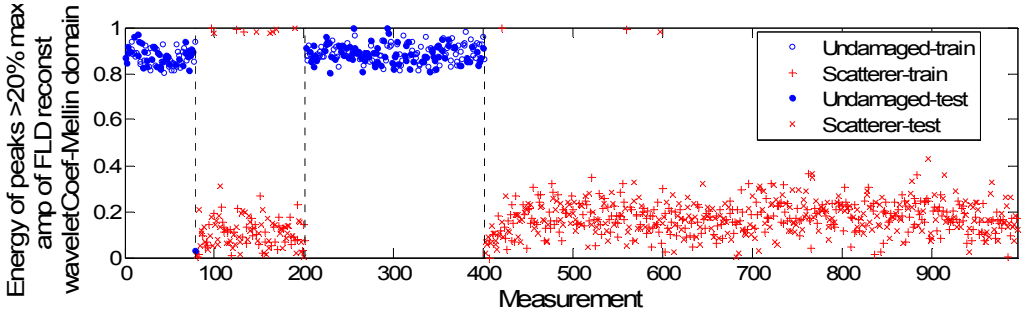
The same feature selection and classification methods studied in the lab (Table 5-10) were applied to the field tests, including three feature selectors (R-filter, S-filter, wrapper selection) and four classifiers (AdaBoost, M-AdaBoost, linear-SVM, and RBF-SVM).

6.3.1.2 Random Testing

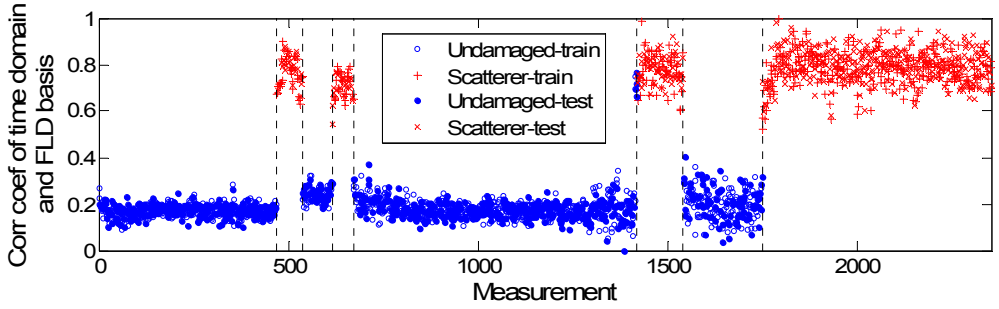
To illustrate the effectiveness of the selected features, Figure 6.6 exemplifies the features selected by the wrapper method (associated with linear-SVM) in some of the random tests performed on Collection F-I, F-II, and their combined set, respectively. The feature is normalized between 0 and 1, plotted along the test time. Each point corresponds to a signal measurement. Figure 6.6 shows that the selected features are insensitive to the cyclic environmental variations and very effective to distinguish the scatterer with large separation from the undamaged measurements. These features shown in Figure 6.6 are all FLD related features that are constructed with the aim of maximizing the discrepancy between different classes, as stated in Section 5.4.1.



(a)



(b)



(c)

Figure 6.6 Performance of the first selected feature by a wrapper method (associated with linear SVM) in (a) a random test on Collection F-I, (b) a random test on Collection F-II, and (c) a random test on the combined Collections F-I and F-II.

Four classifiers were applied to the features selected by the three selectors. Figure 6.7 shows the scatterer detection results in the random testing on Collection F-I. The classification results are expressed in terms of accuracy (ACC), false-positive rate (FPR), false-negative rate (FNR), and area under an ROC curve (AUC).

The results show that with the wrapper method used as the feature selector, all the four classifiers achieve greater than 99% ACC and AUC. When the filtering methods are used as the feature selectors, the two boosting classifiers also obtain above 99% ACC and AUC, while the SVM methods perform poorly as the number of features is increased. This is caused by the large energy variations in the first half cycle of Collection F-I. Under such a wide-range variance, the hypothesis is that only a small portion of the 1500 features preserved their discriminant power, and the wrapper method successfully identified this small set of robust features. Using a large number of features (many are possibly ineffective) degraded the performance of the SVM classifiers. However, the boosting methods were set to execute at most 100 iterations in this study, in order to save the computation time. This setting constrained the maximum number of the features fed into the classifiers. As a result, the boosting methods were less affected by the increased number of ineffective features in the feature library.

Table 6.4 summarizes the performances of the feature selection methods and the classification techniques. The combination of M-AdaBoost and R-filter performs better than all other selector-classifier algorithm pairs. Among the classification techniques, M-AdaBoost is found to be the best on average (average done over all the feature selection methods). Among feature selectors, the wrapper method is found to be the best on average (average done over all the classification methods).

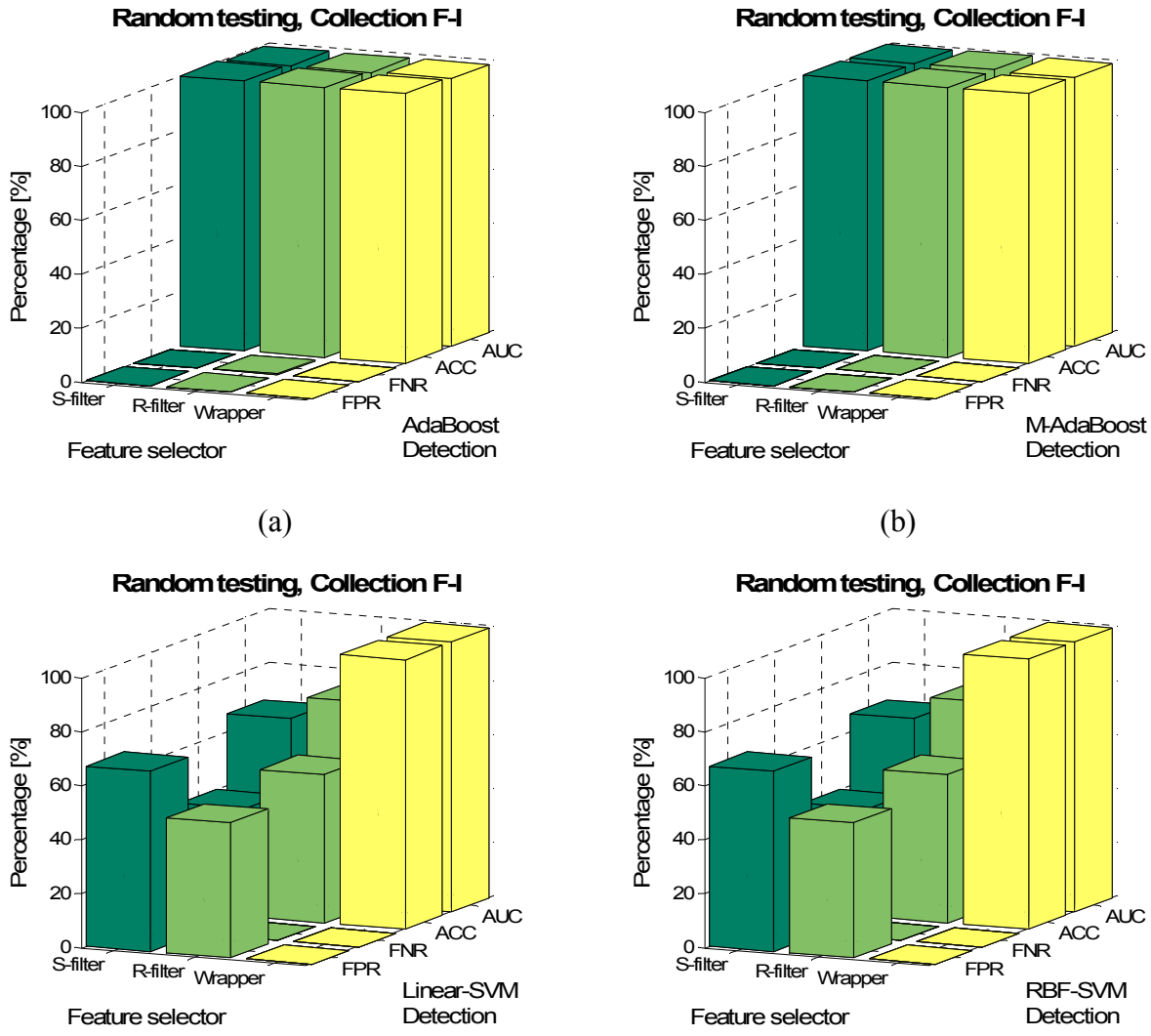


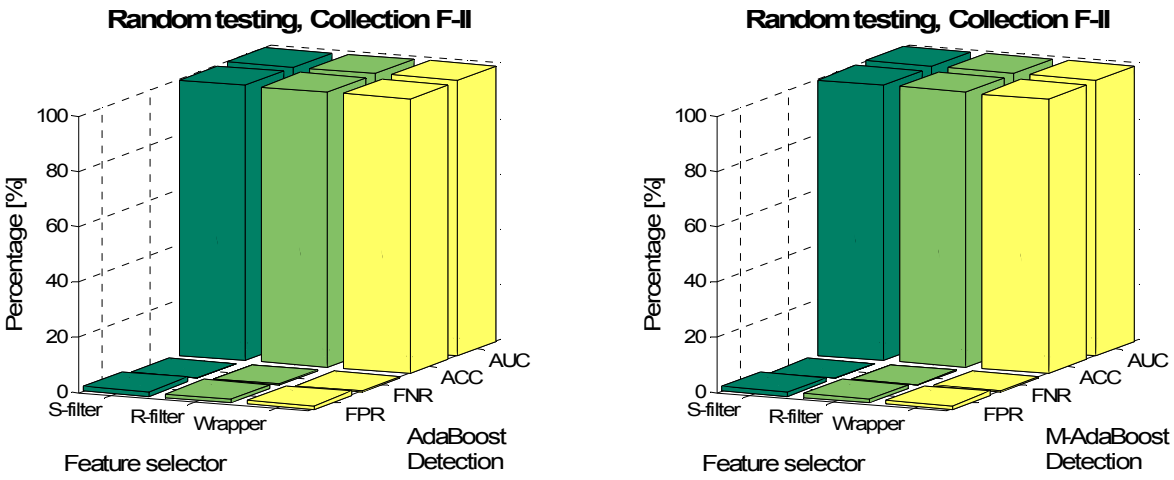
Figure 6.7 Scatterer detection results of the random testing on Collection F-I, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM.

(ACC: accuracy; FPR: false-positive rate; FNR: false-negative rate; AUC: area under an ROC curve; R-filter: a relaxed-filter feature selector; S-filter: a strict-filter feature selector; and Wrapper: a wrapper sequential backward feature selector.)

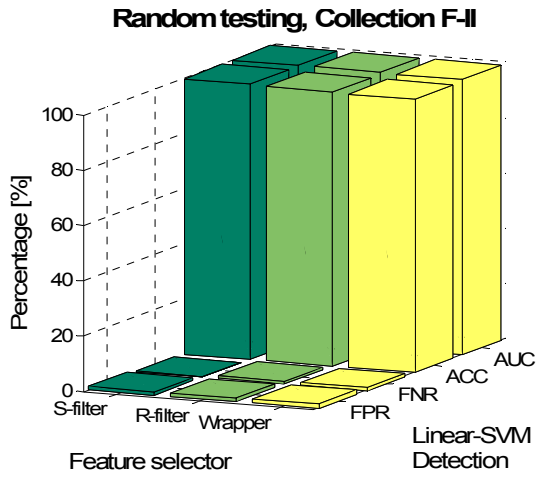
Table 6-4 Comparison of the feature selectors and the classifiers in the random testing on Collection F-I (bold font indicates the best performance).

AUC (ACC)	AdaBoost	M-AdaBoost	Linear-SVM	RBF-SVM	Average
R-filter	99.24% (99.98%)	**100%** **(100%)**	66.72% (39.9%)	66.72% (39.88%)	83.17% (69.94%)
S-filter	99.1% (99.93%)	99.9998% (100%)	75.04% (54.93%)	75.04% (54.93%)	87.3% (77.44%)
Wrapper	99.24% (99.98%)	99.61% (99.98%)	100% (99.95%)	100% (99.98%)	**99.71%** **(99.97%)**
Average	99.2% (99.96%)	**99.87%** **(99.99%)**	80.59% (64.93%)	80.59% (64.93%)	90.06% (82.45%)

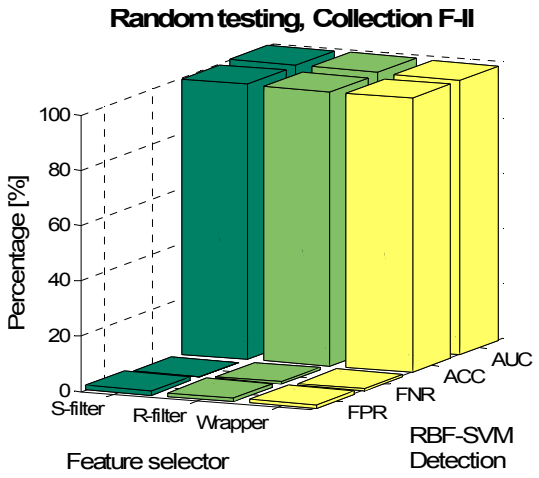
Similarly, the results of the random testing on Collection F-II are presented in Figure 6.8, and the algorithms are compared in Table 6-5. Greater than 99% ACC and AUC are received for any selector-classifier algorithm pair. Collection F-II shows better detection results than Collection F-I, because Collection F-I experienced larger energy variations over time as shown in Figure 6.5.



(a) (b)



(c)



(d)

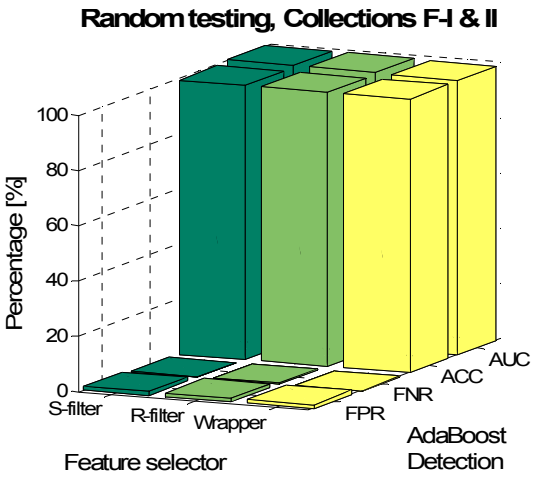
Figure 6.8 Scatterer detection results of the random testing on Collection F-II, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d. RBF-SVM.
(ACC: accuracy; FPR: false-positive rate; FNR: false-negative rate; AUC: area under an ROC curve; R-filter: a relaxed-filter feature selector; S-filter: a strict-filter feature selector; and Wrapper: a wrapper sequential backward feature selector.)

Table 6-5 Comparison of the feature selectors and the classifiers in the random testing on Collection F-II (bold font indicates the best performance).

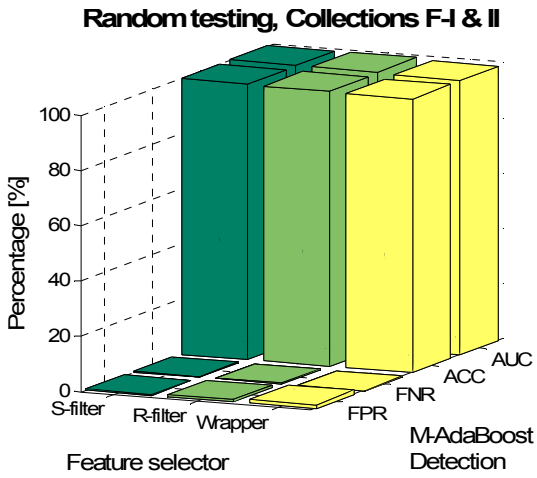
AUC(ACC)	AdaBoost	M-AdaBoost	Linear-SVM	RBF-SVM	Average
R-filter	99.59% (99.43%)	99.9% (99.5%)	99.83% **(99.53%)**	**99.92%** (99.43%)	**99.81%** **(99.47%)**
S-filter	99.66% (99.43%)	99.8% (99.46%)	99.79% (99.06%)	99.79% (99.09%)	99.76% (99.26%)
Wrapper	99.72% (99.43%)	99.65% (99.3%)	99.62% (98.83%)	99.31% (99.16%)	99.57% (99.18%)
Average	99.66% **(99.43%)**	**99.78%** (99.42%)	99.75% (99.14%)	99.67% (99.23%)	99.71% (99.3%)

For the combined set, the results are similar to the results of Collection F-I, as shown in Figure 6.9 and Table 6-6. The boosting classier receives greater than 99% ACC and AUC with any of the feature selectors. The performance of the SVM algorithms depends on the features selected. With the features selected by the wrapper method, the SVM classifiers obtain even higher ACC and AUC than the boosting classifiers, whereas result in lower ACC and AUC as the number of features is increased.

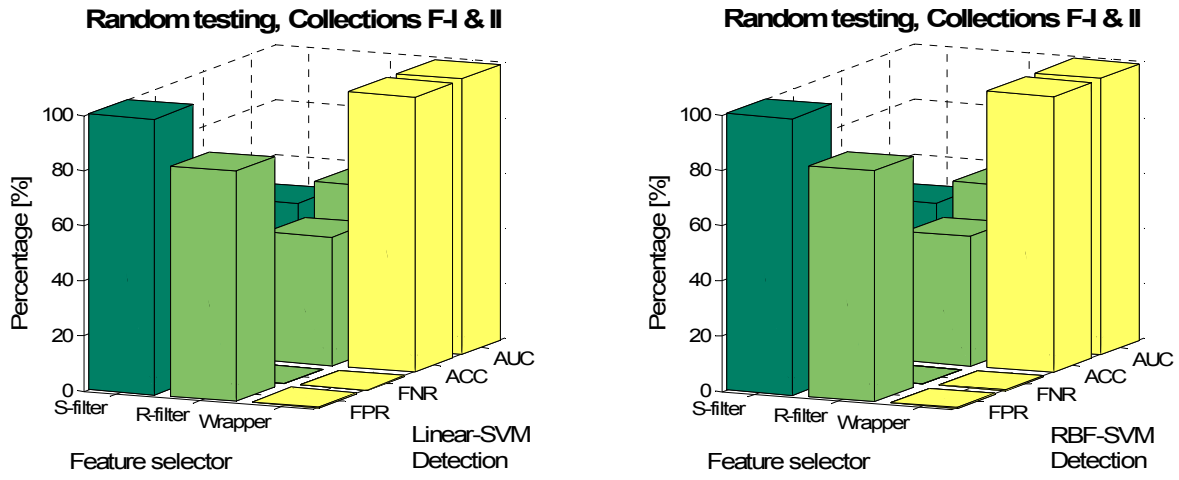
It should be noted that the classifiers in this study are designed to maximize the accuracy during the training process (either for feature selection or classification). Given the nature of unbalanced dataset in the field data collections, it is expected that by adjusting the algorithms to maximize AUC would achieve improved performance. The preceding results also infer the necessity of implementing multiple classifiers and feature selectors, so as to achieve more accurate and reliable results.



(a)



(b)



(c) (d)

Figure 6.9 Scatterer detection results of the random testing on combined Collections F-I and F-II, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM. (ACC: accuracy; FPR: false-positive rate; FNR: false-negative rate; AUC: area under an ROC curve; R-filter: a relaxed-filter feature selector; S-filter: a strict-filter feature selector; and Wrapper: a wrapper sequential backward feature selector.)

Table 6-6 Comparison of the feature selectors and the classifiers in the random testing on the combined Collections F-I and F-II (bold font indicates the best performance).

AUC (ACC)	AdaBoost	M-AdaBoost	Linear-SVM	RBF-SVM	Average
R-filter	99.95% (99.07%)	**99.98%** (99.69%)	50.07% (36.1%)	50.07% (36.1%)	75.02% (67.74%)
S-filter	99.9% (99.05%)	99.93% (99.39%)	58.39% (46.69%)	58.39% (46.72%)	79.15% (72.96%)
Wrapper	99.2% (99.04%)	99.27% (99.04%)	99.96% **(99.77%)**	99.97% **(99.77%)**	**99.6%** **(99.41%)**
Average	99.68% (99.05%)	**99.73%** **(99.37%)**	69.47% (60.85%)	69.47% (60.86%)	84.59% (80.03%)

6.3.1.3 Systematic Testing

In random testing, nearly 100% accuracy and area under the ROC curve can be received with appropriate feature selection and classification methods. This section discusses the results of the systematic testing which are more challenging but realistic as discussed in the laboratory studies (Chapter 5). Although the scatterer was closely spaced during different tests due to limited access in the field, the locations of the scatterer in the basis set, the training set, and the testing set were different positions.

Two systematic tests were conducted on each data collection. In the first systematic test on Collection F-I, Collection F-II was used as a basis set to generate baseline-dependent features, the training set consisted of the measurements taken during the first one and half toggling cycles in Collection F-I, and the remaining one cycle of collected data in Collection F-I was used as the testing set. In the second systematic test on Collection F-I, the roles of the training and testing sets were reversed. Two systematic tests on Collection F-II were formulated in a similar manner.

Figure 6.10 shows the first feature selected by the wrapper method (associated with linear-SVM) in each of the four systematic tests, respectively. Three of the features were generated by combining FLD, wavelet analysis and the Mellin transform, and the fourth were created by combining FLD, wavelet analysis and the Fourier transform. The selected features clearly present a large step as the scatterer was applied or removed from the pipe, allowing visual recognition of the master scatterer. These encouraging features can be used as robust change detectors for on-line monitoring in the field.

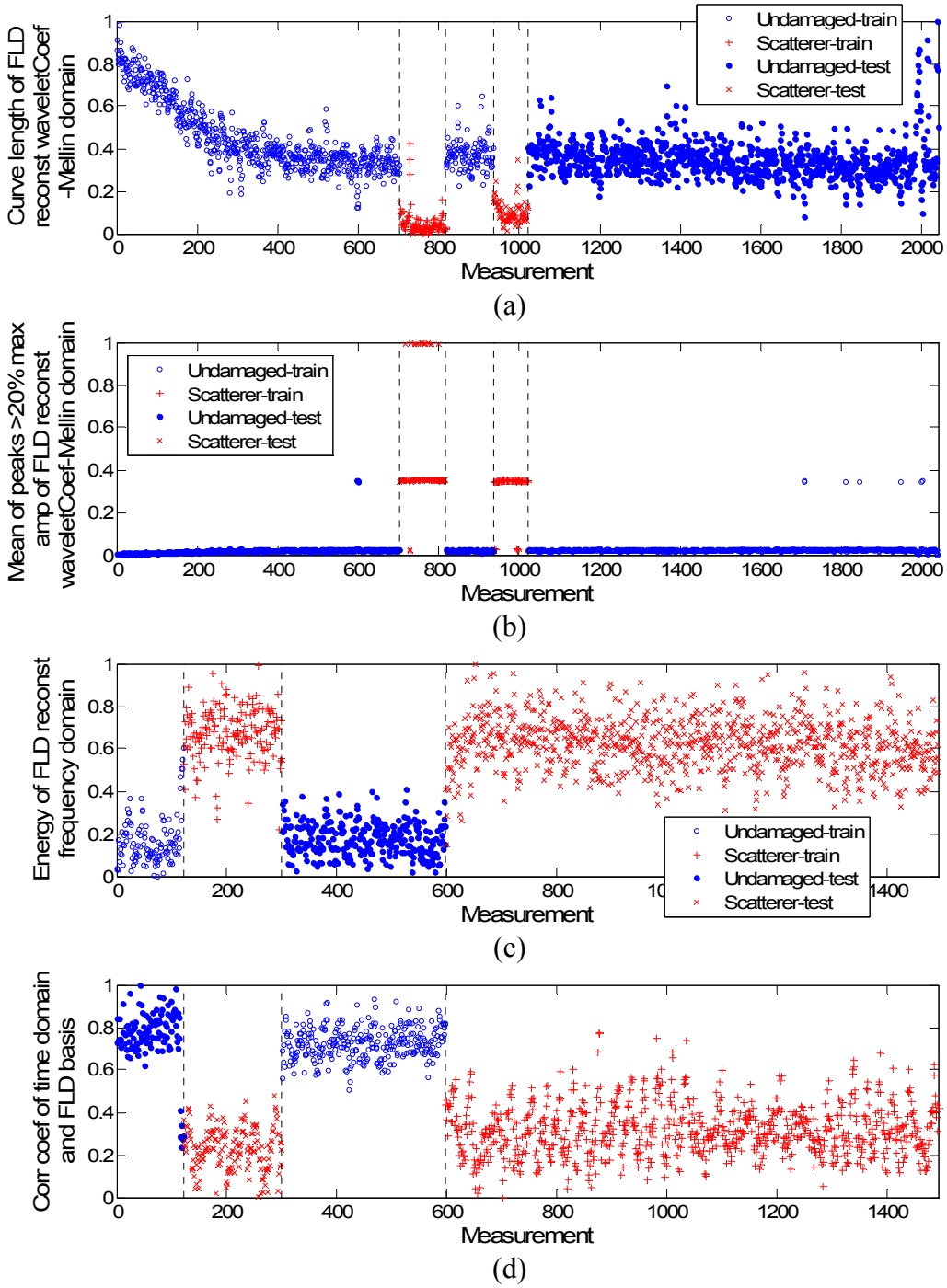
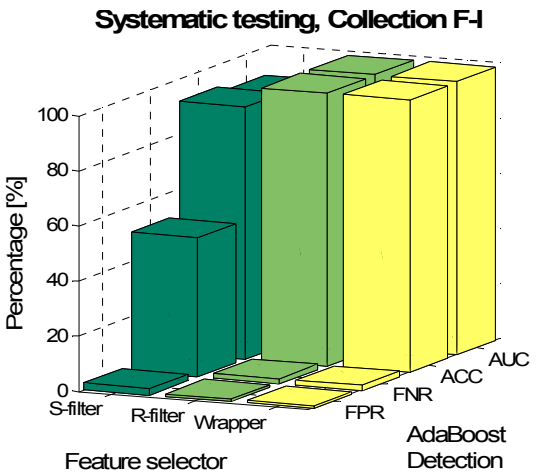


Figure 6.10 Performance of the first selected feature by wrapper method (associated with linear-SVM) in (a) the first systematic test on Collection F-I, (b) the second systematic test on Collection F-I, (c) the first systematic test on Collection F-II, and (d) the second systematic test on Collection F-II.

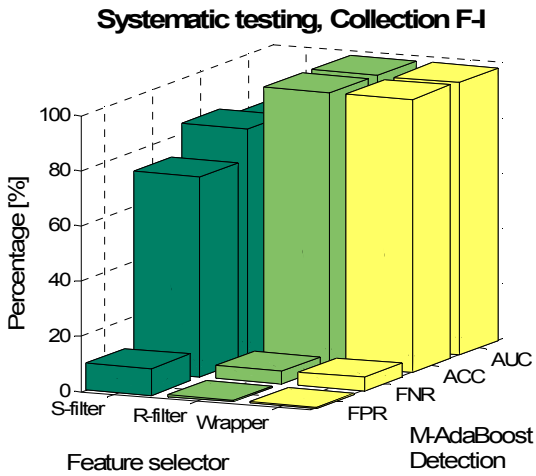
To automatically detect the scatterer, four classification methods were applied. For Collection F-I, we take the average of the classification results from the two systematic tests, and present the scatterer detection results in Figure 6.11 and summarize the algorithm comparison in Table 6-7. With a relatively small number of features selected by the S-filter or by the wrapper method, all four classifiers achieved greater than 98.9% average ACC and AUC. By using the R-filter, the classifiers achieved an average 86.06% AUC and 90.27% ACC.

The results of the systematic testing on Collection F-II are presented in Figure 6.12 and Table 6-8. The SVM classifiers outperform the boosting methods. The SVM classifiers receive higher than 99% AUC and around 95% ACC with any of the three feature selectors, while the boosting provides relatively good results with R-filter.

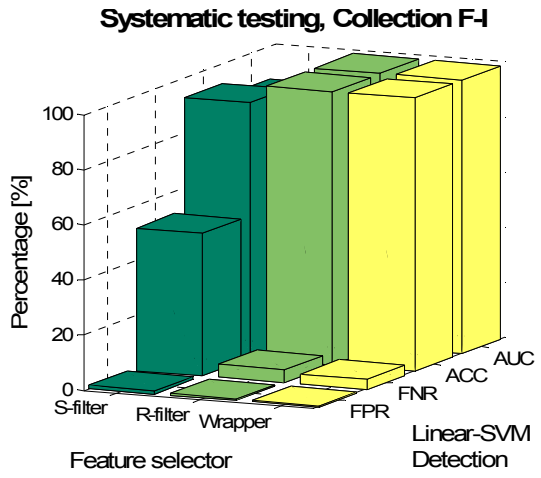
The preceding results again show that no classifier or feature selector performs better in all the scenarios, and therefore it is necessary to implement multiple selectors and classifiers as described in the data-driven framework.



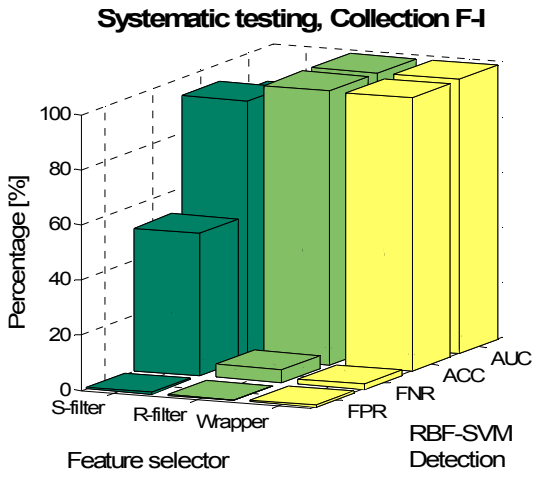
(a)



(b)



(c)



(d)

Figure 6.11 Scatterer detection results of the systematic testing on Collection F-I, with Collection F-II as a basis set, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM. (ACC: accuracy; FPR: false-positive rate; FNR: false-negative rate; AUC: area under an ROC curve; R-filter: a relaxed-filter feature selector; S-filter: a strict-filter feature selector; and Wrapper: a wrapper sequential backward feature selector.)

Table 6-7 Comparison of the feature selectors and the classifiers in the systematic testing on Collection F-I using Collection F-II as the basis set (bold font indicates the best performance).

AUC (ACC)	AdaBoost	M-AdaBoost	Linear-SVM	RBF-SVM	Average
R-filter	89.45% (91.62%)	80.07% (83.51%)	87.74% (92.69%)	86.99% (93.27%)	86.06% (90.27%)
S-filter	99.17% (98.96%)	98.97% (99.09%)	99.26% (99.09%)	99.46% **(99.41%)**	**99.21%** **(99.14%)**
Wrapper	99.17% (98.96%)	98.7% (99.09%)	99.26% (99.18%)	**99.52%** (99.18%)	99.17% (99.1%)
Average	**95.93%** (96.51%)	92.58% (93.9%)	95.42% (96.99%)	95.32% **(97.29%)**	94.81% (96.17%)

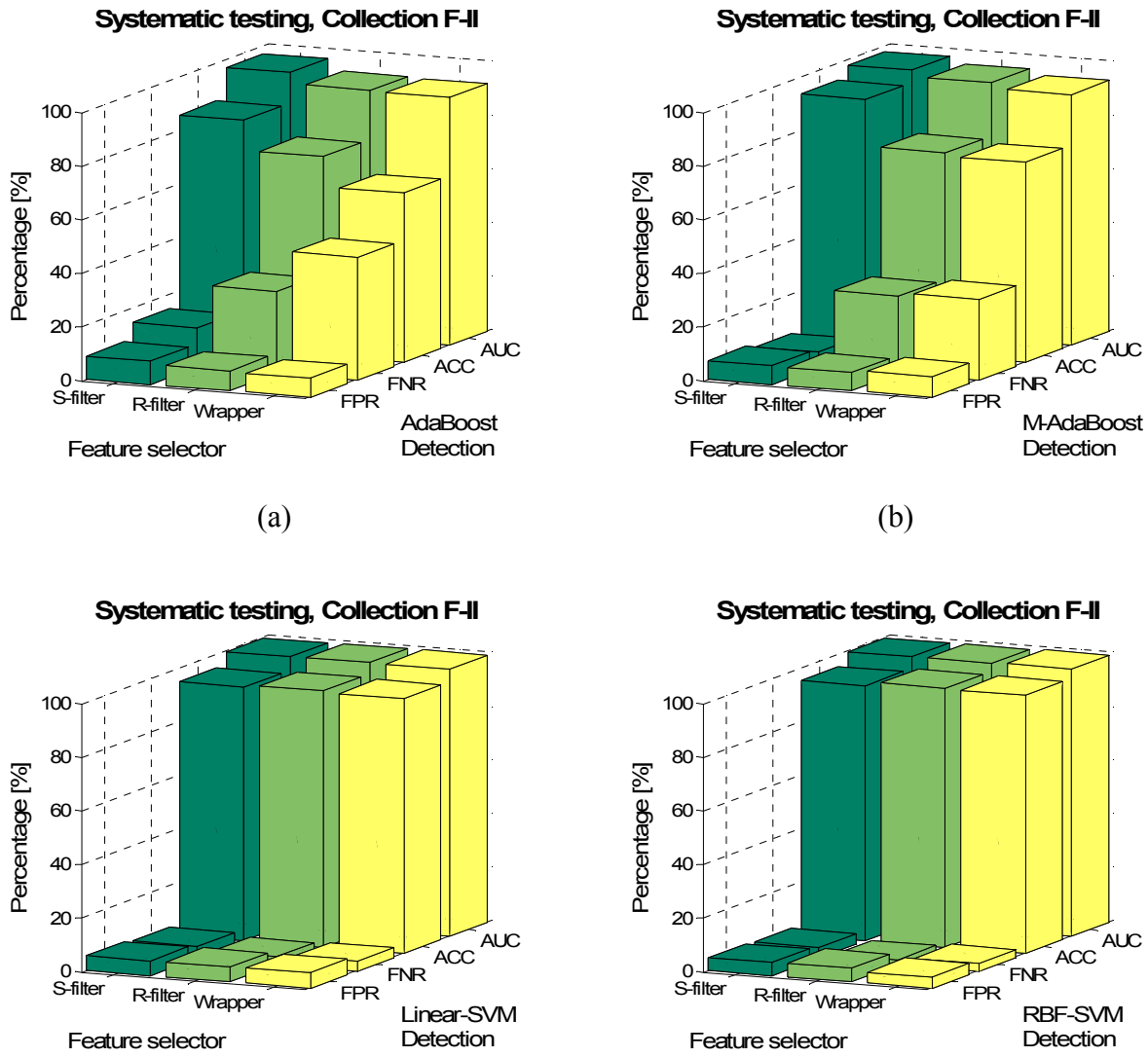


Figure 6.12 Scatterer detection results of the systematic testing on Collection F-II, with Collection F-I as a basis set, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM.

(ACC: accuracy; FPR: false-positive rate; FNR: false-negative rate; AUC: area under an ROC curve; R-filter: a relaxed-filter feature selector; S-filter: a strict-filter feature selector; and Wrapper: a wrapper sequential backward feature selector.)

Table 6-8 Comparison of the feature selectors and the classifiers in the systematic testing on Collection F-II using Collection F-I as the basis set (bold font indicates the best performance).

AUC (ACC)	AdaBoost	M-AdaBoost	Linear-SVM	RBF-SVM	Average
R-filter	96.81% (85.51%)	97.6% (93.22%)	99.32% (94.52%)	99.37% (95.06%)	**98.27% (92.08%)**
S-filter	92.6% (74.38%)	95.37% (75.73%)	99.63% (95.48%)	99.48% **(96.53%)**	96.77% (85.53%)
Wrapper	92.3% (63.18%)	93.48% (74.56%)	99.61% (95.02%)	**99.69%** (96.28%)	96.27% (82.26%)
Average	93.9% (74.36%)	95.48% (81.17%)	**99.52%** (95.01%)	99.51% **(95.96%)**	97.1% (86.62%)

6.3.2 Field Test 2: Detecting a Scatterer at One Location

6.3.2.1 Data Collection and Machine Learning Test Formulation

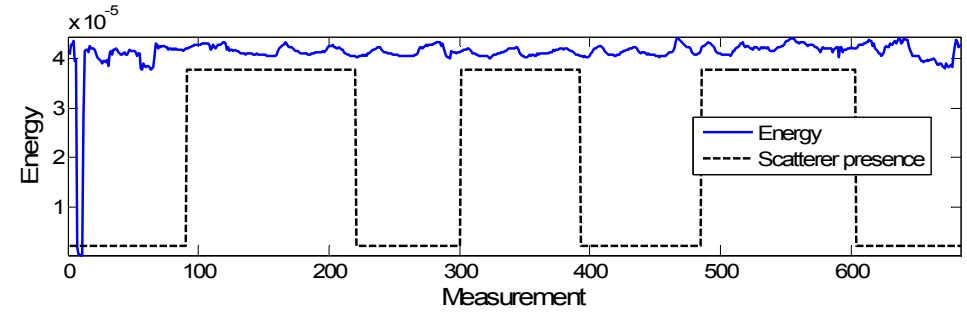
Two more collections of field test data, using a different pair of transducers, were tested using the data-driven monitoring framework. The mass scatterer was placed at a fixed location.

We provide the data collection details as listed in Table 6-9. Collection F-III was acquired on 8/24/2011: 344 measurements were taken without the presence of a scatterer, 340 measurements were taken with the presence of a scatterer. The total 684 measurements were taken with about one-minute interval between each measurement, lasting 11hr 29min. Three and half scatterer toggling cycles were performed. Collection F-IV was acquired on 8/31/2011: 891 measurements were taken without the presence of a scatterer, 553 measurements were taken with the presence of a scatterer. The tests consisted of four toggling cycles, with a total duration of 27hr 54min. The mass was placed at a fixed position for all the scatterer measurements.

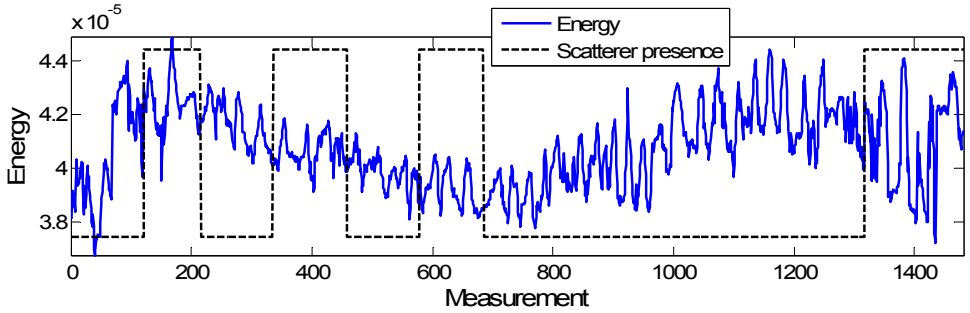
Table 6-9 Field data collection details (Field Test 2).

	Undamaged datasets	Scatterer datasets	Test duration	Toggling cycles	Scatterer locations
Collection F-III (8/24/2011)	344	340	11hr 29min	3.5	1
Collection F-IV (8/31/2011)	891	553	27hr 54min	4	1

Figure 6.13 shows the energy variations over time for the two field data collections. The toggling cycles are indicated in dashed lines. Similar to the observations in Field Test 1, the energy shows cyclic patterns, with no visible indication of the introduction and removal of the scatterer.



(a)



(b)

Figure 6.13 Energy variation over time of (a) Collection F-III, and (b) Collection F-IV.

We again used a random testing and a systematic testing to evaluate the feasibility of the data-driven methodology, as shown in Table 6-10. The random testing was formulated the same way as that done in Field Test 1, while the systematic testing was performed differently. We investigated the scatterer detectability of the data-driven methods against the energy distribution in the data. The energy of an ultrasonic response is affected by numerous factors in the monitoring system, such as any environmental fluctuations, operating condition changes, and instability of the data acquisition devices. The systematic testing in Field Test 2 focused on evaluating the reliability of the monitoring system under such variations. As a result, we divided each data collection into a training set and a testing set based on the different energy levels; one set contained half of the data collection with higher energy values, and the remaining data with lower signal energy formed the other set.

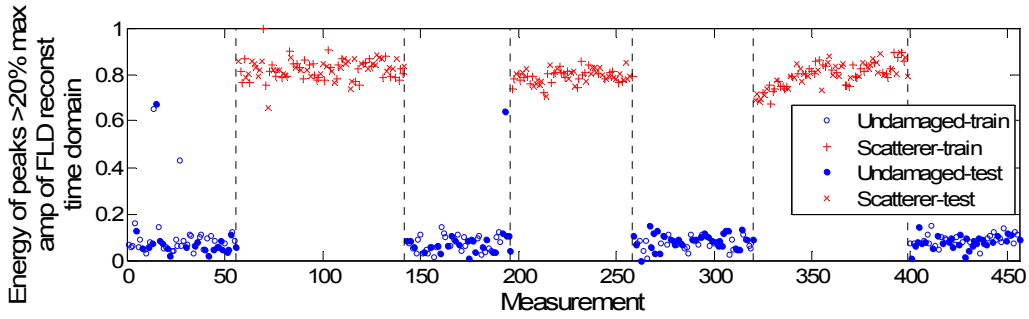
Table 6-10 Machine learning testing formulation for scatterer detection (Field Test 2).

	Single-day data	**Multi-day data**
Random testing	3 randomly divided, equal (or approximately equal) sized sets Total: 12 tests	3 randomly divided, equal (or approximately equal) sized sets Total: 6 tests
Systematic testing	One collection as a basis set, 2-fold systematic testing on the other collection (divided into training and testing sets according to different **energy variations**). Total: 4 tests	

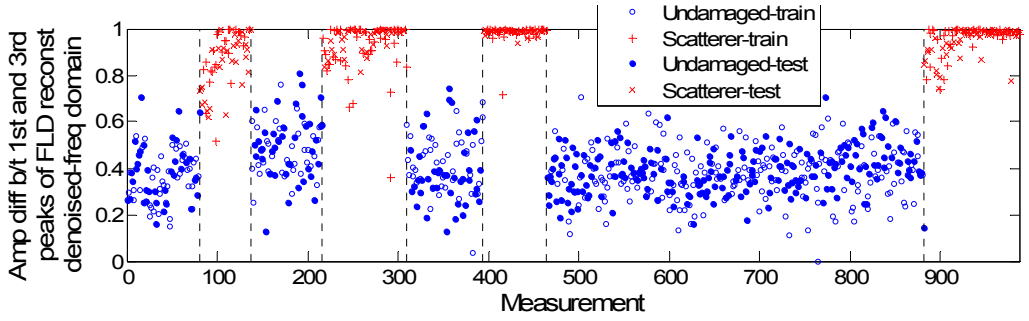
6.3.2.2 Random Testing

The three feature selection methods and four classification algorithms were applied in the single-day and multi-day random testing. Figure 6.14 exemplifies the features selected by the wrapper method (associated with linear-SVM) in some of the random tests performed on Collection F-III, F-IV, and their combined set, respectively. The feature is normalized between

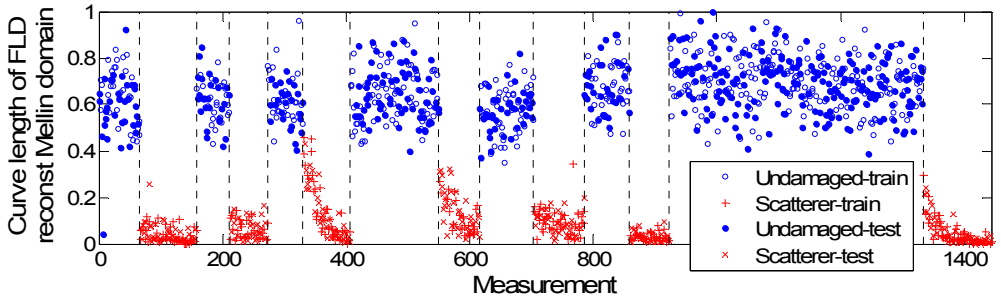
zero and one, plotted along the test time. Each point corresponds to a signal measurement. The selected features, extracted by using the FLD analysis, significantly separate the scatterer instances from the undamaged instances.



(a)



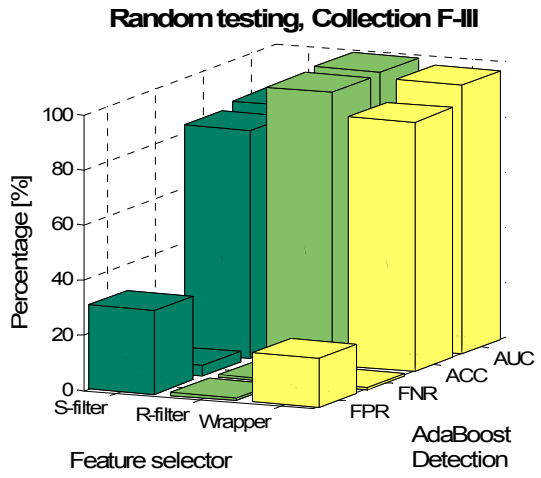
(b)



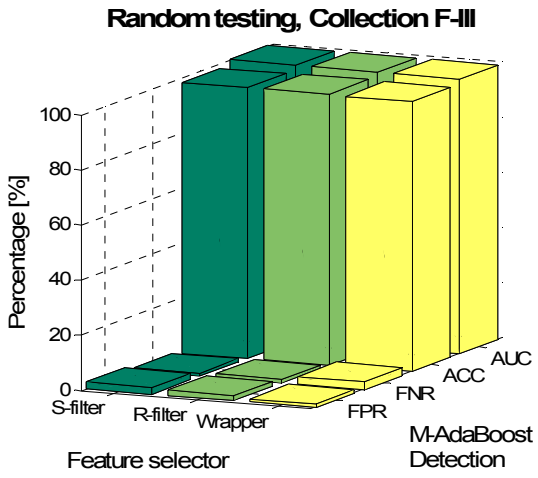
(c)

Figure 6.14 Performance of the first selected feature by a wrapper method (associated with linear-SVM) in (a) a random test on Collection F-III, (b) a random test on Collection F-IV, and (c) a random test on the combined Collections F-III and F-IV.

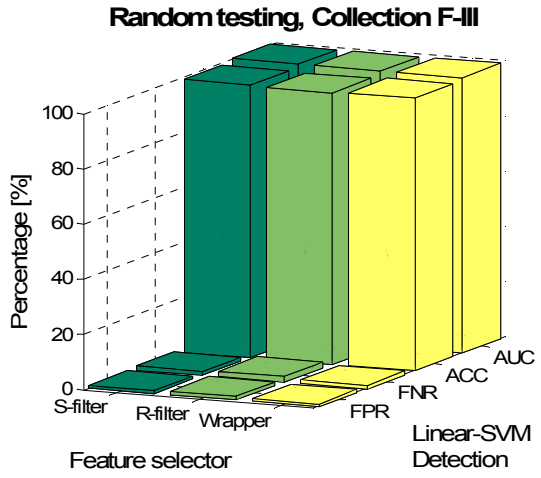
Figure 6.15 shows the scatterer detection results from the random testing on Collection F-III. M-AdaBoost, linear-SVM and RBF-SVM obtain greater than 99% AUC (>98% ACC) with any of the feature selectors. The AdaBoost method achieves the same level AUC by combining with S-filter, while receiving relatively high false-positive rates when other features selectors are used. On average, S-filter outperforms the other two feature selectors: the features selected by S-filter achieved 99.85% AUC and 98.46% ACC averaged over the four classifiers, as shown in Table 6-11.



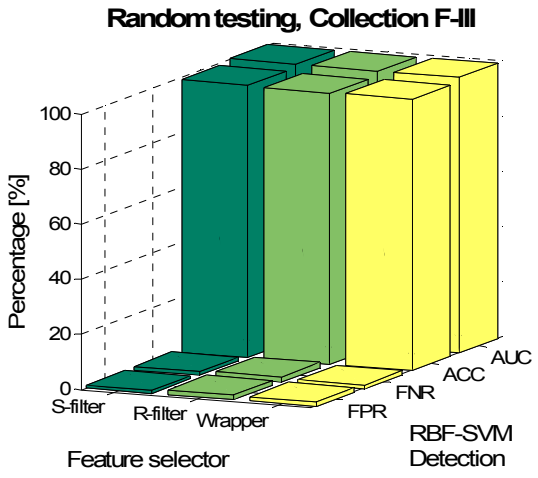
(a)



(b)



(c)



(d)

Figure 6.15 Scatterer detection results of the random testing on Collection F-III, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM.

(ACC: accuracy; FPR: false-positive rate; FNR: false-negative rate; AUC: area under an ROC curve; R-filter: a relaxed-filter feature selector; S-filter: a strict-filter feature selector; and Wrapper: a wrapper sequential backward feature selector.)

Table 6-11 Comparison of the feature selectors and the classifiers in the random testing on Collection F-III (bold font indicates the best performance).

AUC (ACC)	AdaBoost	M-AdaBoost	Linear-SVM	RBF-SVM	Average
R-filter	84.97% (82.75%)	**99.97%** (98.17%)	99.93% (98.76%)	99.9% (98.83%)	96.19% (94.63%)
S-filter	99.81% **(98.98%)**	99.88% (98.39%)	99.85% (98.32%)	99.87% (98.17%)	**99.85%** **(98.46%)**
Wrapper	97.51% (90.64%)	99.63% (98.03%)	99.82% (98.83%)	99.94% (98.54%)	99.22% (96.51%)
Average	94.1% (90.79%)	99.83% (98.2%)	99.87% **(98.64%)**	**99.9%** (98.51%)	98.42% (96.53%)

The results of the random testing on Collection F-IV are presented in Figure 6.16 and Table 6-12. A relatively small set of features selected by the wrapper methods constantly achieved an AUC greater than 99.5% by applying any of the classifiers. The boosting classification methods are robust to the feature selection process, obtaining greater than 99.5% AUC with any of the feature selectors.

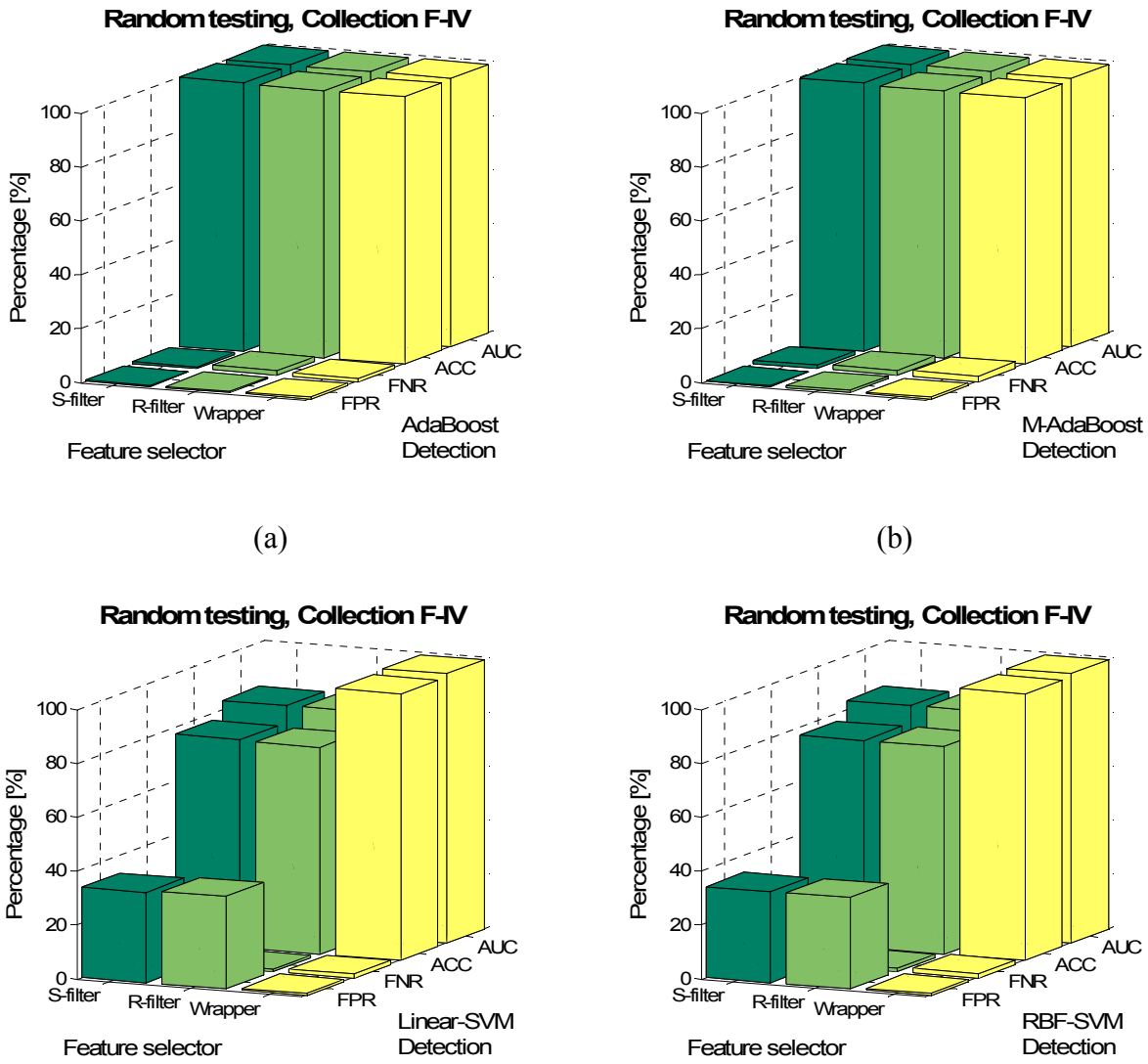


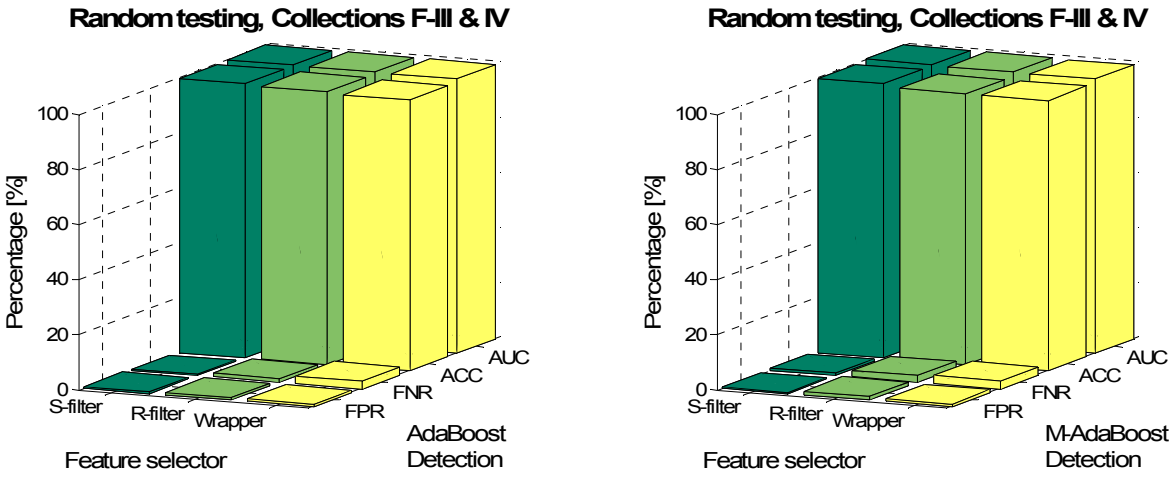
Figure 6.16 Scatterer detection results of the random testing on Collection F-IV, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM.

(ACC: accuracy; FPR: false-positive rate; FNR: false-negative rate; AUC: area under an ROC curve; R-filter: a relaxed-filter feature selector; S-filter: a strict-filter feature selector; and Wrapper: a wrapper sequential backward feature selector.)

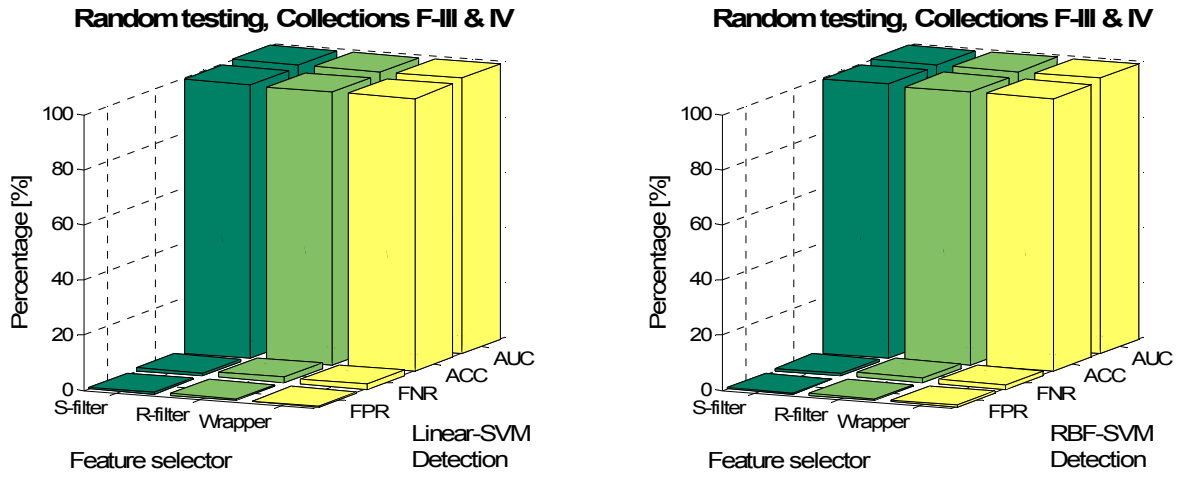
Table 6-12 Comparison of the feature selectors and the classifiers in the random testing on Collection F-IV (bold font indicates the best performance).

AUC (ACC)	AdaBoost	M-AdaBoost	Linear-SVM	RBF-SVM	Average
R-filter	99.82% **(99.6%)**	**99.99%** (99.49%)	83.36% (77.14%)	83.33% (76.87%)	91.62% (88.27%)
S-filter	99.72% (99.09%)	99.85% (98.89%)	83.36% (76.77%)	83.35% (76.9%)	91.57% (87.91%)
Wrapper	99.7% (99.16%)	99.55% (98.82%)	99.9% (99.02%)	99.94% (98.95%)	**99.77% (98.99%)**
Average	99.74% **(99.28%)**	**99.8%** (99.07%)	88.87% (84.31%)	88.87% (84.24%)	94.32% (91.72%)

Figure 6.17 and Table 6-13 provide the classification results of the multi-day random testing. Any of selector-classifier algorithm pairs achieves an average AUC above 99.7%, and ACC above 98.6%.



(a) (b)



(c) (d)

Figure 6.17 Scatterer detection results of the random testing on combined Collections F-III and F-IV, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM. (ACC: accuracy; FPR: false-positive rate; FNR: false-negative rate; AUC: area under an ROC curve; R-filter: a relaxed-filter feature selector; S-filter: a strict-filter feature selector; and Wrapper: a wrapper sequential backward feature selector.)

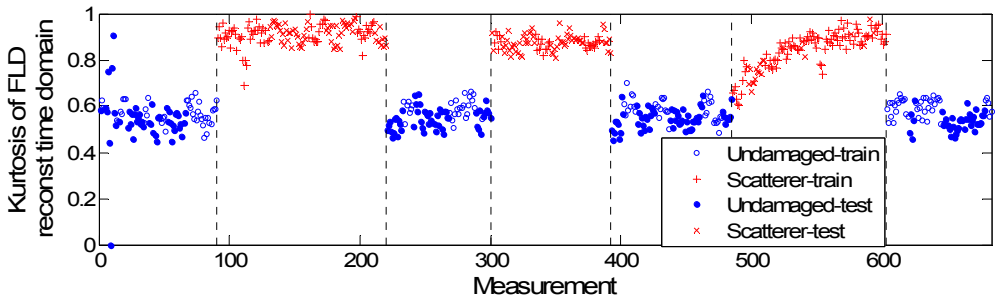
Table 6-13 Comparison of the feature selectors and the classifiers in the random testing on the combined Collections F-III and F-IV (bold font indicates the best performance).

AUC (ACC)	AdaBoost	M-AdaBoost	Linear-SVM	RBF-SVM	Average
R-filter	99.86% (99.45%)	**99.98%** (99.42%)	99.95% (99.26%)	99.98% **(99.45%)**	**99.94%** **(99.39%)**
S-filter	99.76% (98.94%)	99.71% (98.27%)	99.85% (98.87%)	99.92% (98.98%)	99.81% (98.76%)
Wrapper	99.57% (98.38%)	99.46% (98.27%)	99.9% (99.08%)	99.94% (98.91%)	99.72% (98.66%)
Average	99.73% (98.92%)	99.72% (98.65%)	99.9% (99.07%)	**99.95%** **(99.11%)**	99.82% (98.94%)

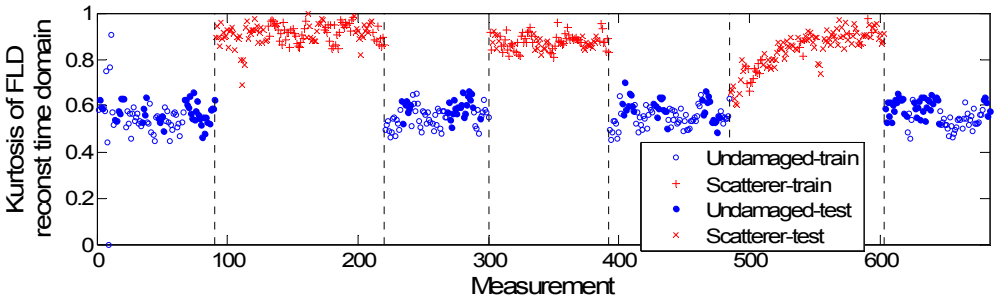
6.3.2.3 Systematic Testing

Two systematic tests were conducted on each data collection. In the first systematic test on Collection F-III, Collection F-IV was used as a basis set to generate baseline-dependent features, the training set consisted of the measurements with relatively high energy in Collection F-III, and the remaining signals with relatively low energy in Collection F-III was used as the testing set. In the second systematic test on Collection F-III, the roles of the training and testing sets were reversed. Two systematic tests on Collection F-IV were formulated in a similar manner.

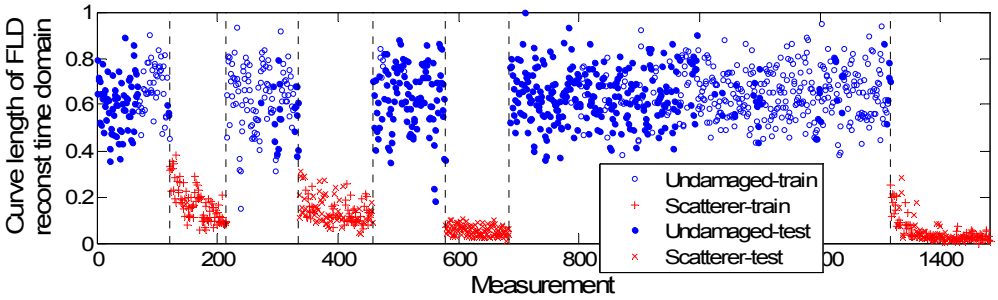
Figure 6.18 shows the first features selected by the wrapper method (associated with linear-SVM) in the four systematic tests, respectively. We observe from each collection that the same feature is found to be the most effective either by using the higher-energy data for training or by using the lower-energy data for training. This result indicates that the selected feature, again related to FLD, is independent from energy variations, and therefore provides robust scatterer detection. With such effective features, we achieved around 98.5% AUC and 92.9% ACC, averaged over all the selectors and classifiers, for Collection III (as shown in Figure 6.19 and Table 6-14) , and 99.91% average AUC and 99.43% average ACC for Collection IV (as shown in Figure 6.20 and Table 6-15).



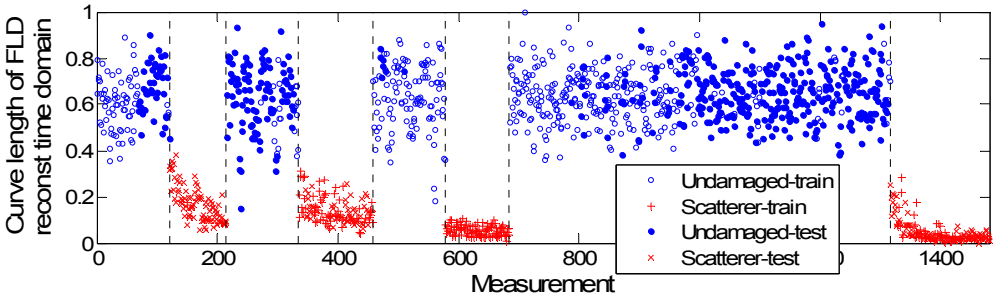
(a)



(b)



(c)



(d)

Figure 6.18 Performance of the first selected feature by a wrapper method (associated with linear-SVM) in (a) the first systematic test on Collection F-III, (b) the second systematic test on Collection F-III, (c) the first systematic test on Collection F-IV, and (d) the second systematic test on Collection F-IV.

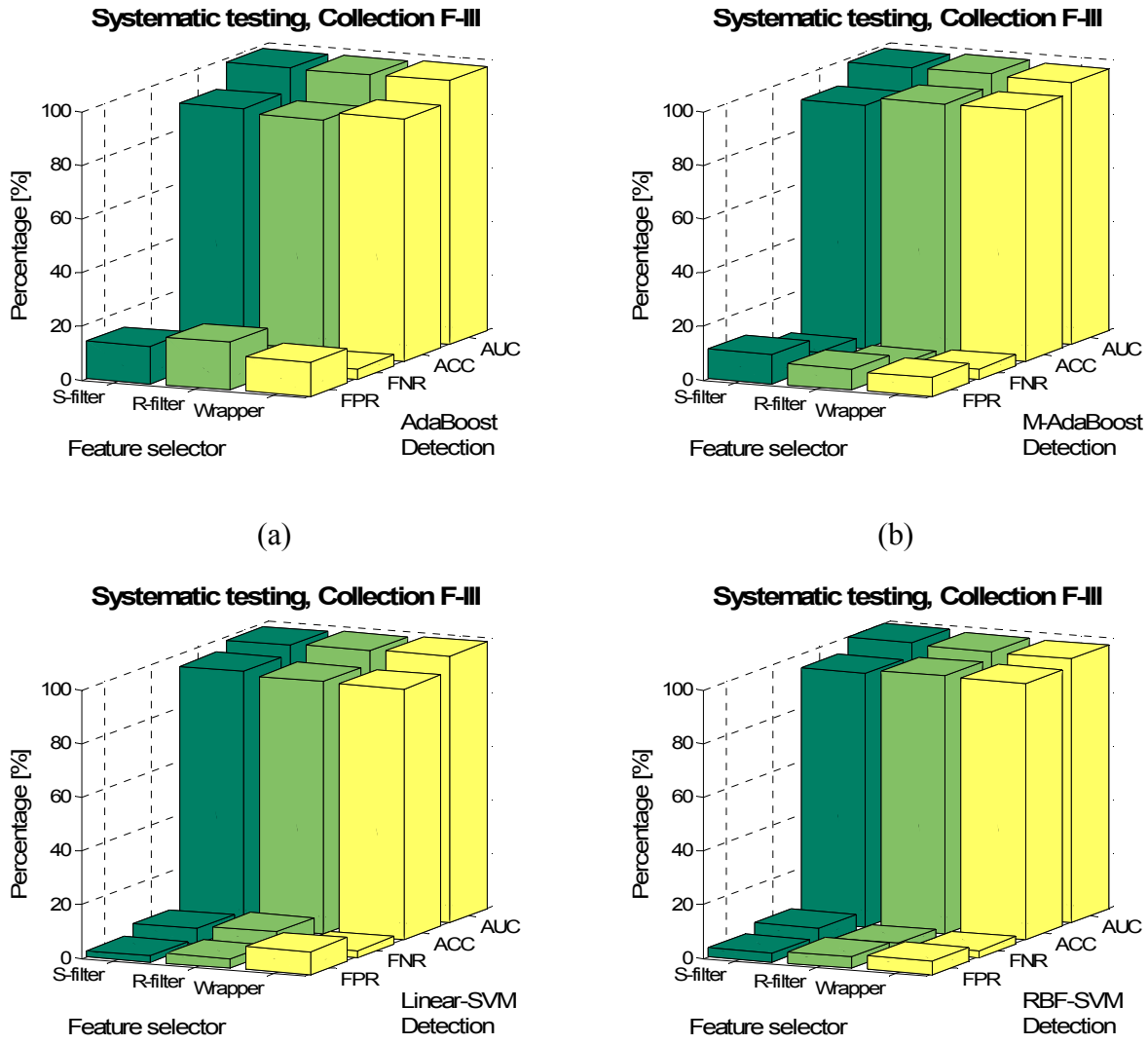
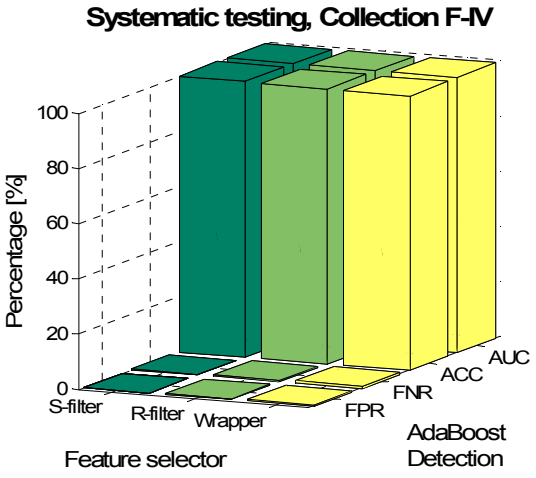


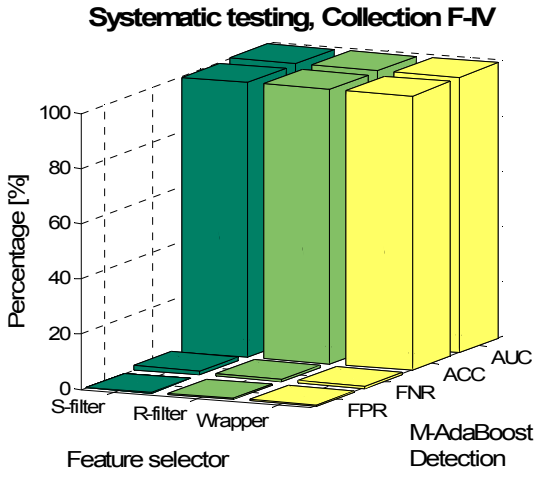
Figure 6.19 Scatterer detection results of the systematic testing on Collection III, with Collection IV as a basis set, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM. (ACC: accuracy; FPR: false-positive rate; FNR: false-negative rate; AUC: area under an ROC curve; R-filter: a relaxed-filter feature selector; S-filter: a strict-filter feature selector; and Wrapper: a wrapper sequential backward feature selector.)

Table 6-14 Comparison of the feature selectors and the classifiers in the systematic testing on Collection F-III using Collection F-IV as the basis set (bold font indicates the best performance).

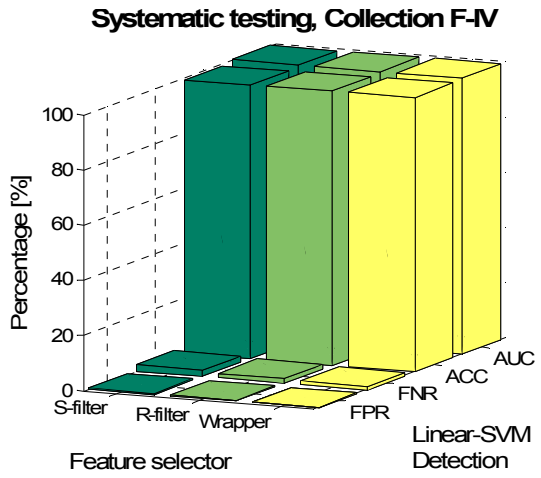
AUC (ACC)	AdaBoost	M-AdaBoost	Linear-SVM	RBF-SVM	Average
R-filter	98.2% (89.62%)	98.26% (90.79%)	98.51% (95.32%)	**99.55%** (94.59%)	**98.63%** (92.58%)
S-filter	97.99% (87.72%)	98.47% (93.71%)	98.98% (93.86%)	98.62% **(95.91%)**	98.52% (92.8%)
Wrapper	98.3% (90.5%)	97.56% (93.71%)	99.41% (93.57%)	98.24% (95.61%)	98.38% **(93.35%)**
Average	98.17% (89.28%)	98.1% (92.74%)	**98.97%** (94.25%)	98.8% **(95.34%)**	98.51% (92.91%)



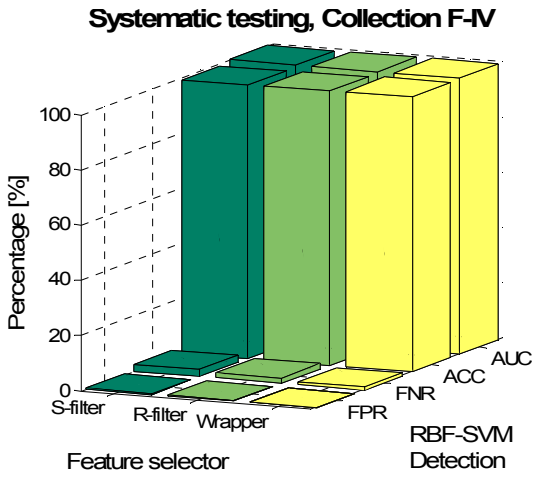
(a)



(b)



(c)



(d)

Figure 6.20 Scatterer detection results of systematic testing on Collection F-IV, with Collection F-III as a basis set, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM. (ACC: accuracy; FPR: false-positive rate; FNR: false-negative rate; AUC: area under an ROC curve; R-filter: a relaxed-filter feature selector; S-filter: a strict-filter feature selector; and Wrapper: a wrapper sequential backward feature selector.)

Table 6-15 Comparison of the feature selectors and the classifiers in the systematic testing on Collection F-IV using Collection F-III as the basis set (bold font indicates the best performance).

AUC (ACC)	AdaBoost	M-AdaBoost	Linear-SVM	RBF-SVM	Average
R-filter	99.97% **(99.93%)**	99.99% (99.49%)	99.9% (98.99%)	99.91% (98.99%)	**99.94%** (99.34%)
S-filter	99.86% (99.66%)	99.81% (99.46%)	99.92% (99.26%)	99.98% (99.33%)	99.89% (99.43%)
Wrapper	99.82% (99.46%)	99.79% (99.46%)	99.97% (99.46%)	**99.995%** (99.66%)	99.9% **(99.51%)**
Average	99.88% **(99.69%)**	99.86% (99.46%)	99.93% (99.24%)	**99.96%** (99.33%)	99.91% (99.43%)

6.3.3 Field Test 3: Localizing a Scatterer at Two Different Locations

6.3.3.1 Data Collection and Machine Learning Test Formulation

Field Test 3 focused on evaluating the capability of the data-driven methods for scatterer localization. Four cycles of scatterer toggling experiments were performed on 5/16/2012, with the mass scatterer placed and removed at two different locations 5.5 inches (139.7 mm) apart. The data collection details are provided in Table 6-16.

Specifically, the scatterer was applied at the first location within the first and third toggling cycles, and was applied at the second location within the second and fourth toggling cycles. The measurements were taken at one-minute intervals, with a total duration of 22hr 56min, including 985 signals without the presence of the scatterer, 216 signals with the scatterer at Location 1, and 171 signals with the scatterer at Location 2. The energy variation over time of Collection F-V was shown in Figure 6.21, with a dashed line indicating the toggling cycles.

To evaluate the data-driven framework for scatterer location, only the 387 records with the presence of the scatterer were used. Random testing and systematic testing were formulated as shown in Table 6-17. In random testing, Collection F-V data was randomly divided into three approximately equal-sized subsets, as a basis set, a training set, and a testing set; six random tests were performed, by switching the roles of the three subsets. In systematic testing, the basis set, the training set, and the testing set were divided according to the toggling cycles: the scatterer measurements within two toggling cycles (the first two cycles, or the third and the fourth cycles) were used as the basis set for feature extraction; a 2-fold machine learning test was performed on the remaining scatterer data. Four systematic testing were performed.

Table 6-16 Field data collection details (Field Test 3).

	Undamaged datasets	Scatterer Location 1 datasets	Scatterer Location 2 datasets	Test duration	Toggling cycles	Scatterer locations
Collection F-V (5/16/2012)	985	216	171	22hr 56min	4	2

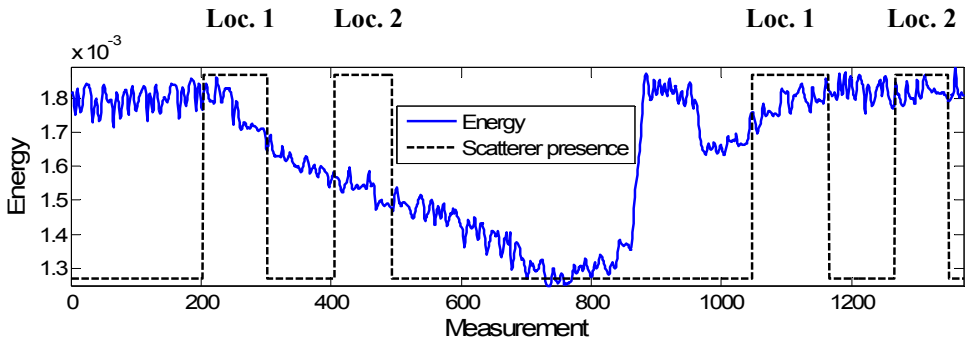


Figure 6.21 Energy variation over time of Collection F-V.

Table 6-17 Machine learning testing formulation for scatterer localization (Field Test 3).

	Single-day data
Random testing	3 randomly divided, equal (or approximately equal) sized sets Total: 6 tests
Systematic testing	Scatterer data within two toggling cycles as a basis set, 2-fold systematic testing on the remaining scatterer data (divided into training and testing sets according to different data collection time). Total: 4 tests

6.3.3.2 Testing Results

For each random or systematic test, three feature selection methods (R-filter, S-filter, and wrapper selection) and four machine learning classification algorithms (AdaBoost, M-AdaBoost, linear-SVM, and RBF-SVM) were applied, resulting in 12 different pairs of selector-classifier algorithms. Figure 6.22 and Figure 6.23 show the results of scatterer localization in the random testing and the systematic testing, respectively. Zone 1 instances are denoted as positives, while Zone 3 instances are denoted as negatives.

In the random testing, the AdaBoost and M-AdaBoost classifiers achieve 100% ACC using the features selected by any of the feature selection methods. Linear-SVM and RBF-SVM achieve 100% ACC when wrapper selection is applied, while receive high false-positive rate (many instances with the scatterer in Location 2 are misclassified as the scatterer in Location 1) when the filter feature selectors are used. This is caused by the large energy variations in the second toggling cycle of Collection F-V. The hypothesis is that only a small portion of the 1500 features preserved their discriminant power, and the wrapper method successfully identified this small set of robust features. Using a large number of features (many are possibly ineffective) degraded the performance of the SVM classifiers. However, the boosting methods were set to execute at most 100 iterations, which constrained the maximum number of the features fed into the classifiers. As a result, the boosting methods were less affected by the increased number of ineffective features in the feature library.

In the systematic testing, applying M-AdaBoost with the features selected by any of the filter methods achieve the highest ACC, greater than 97%, for scatterer localization. Other algorithms obtain around 85% ACC.

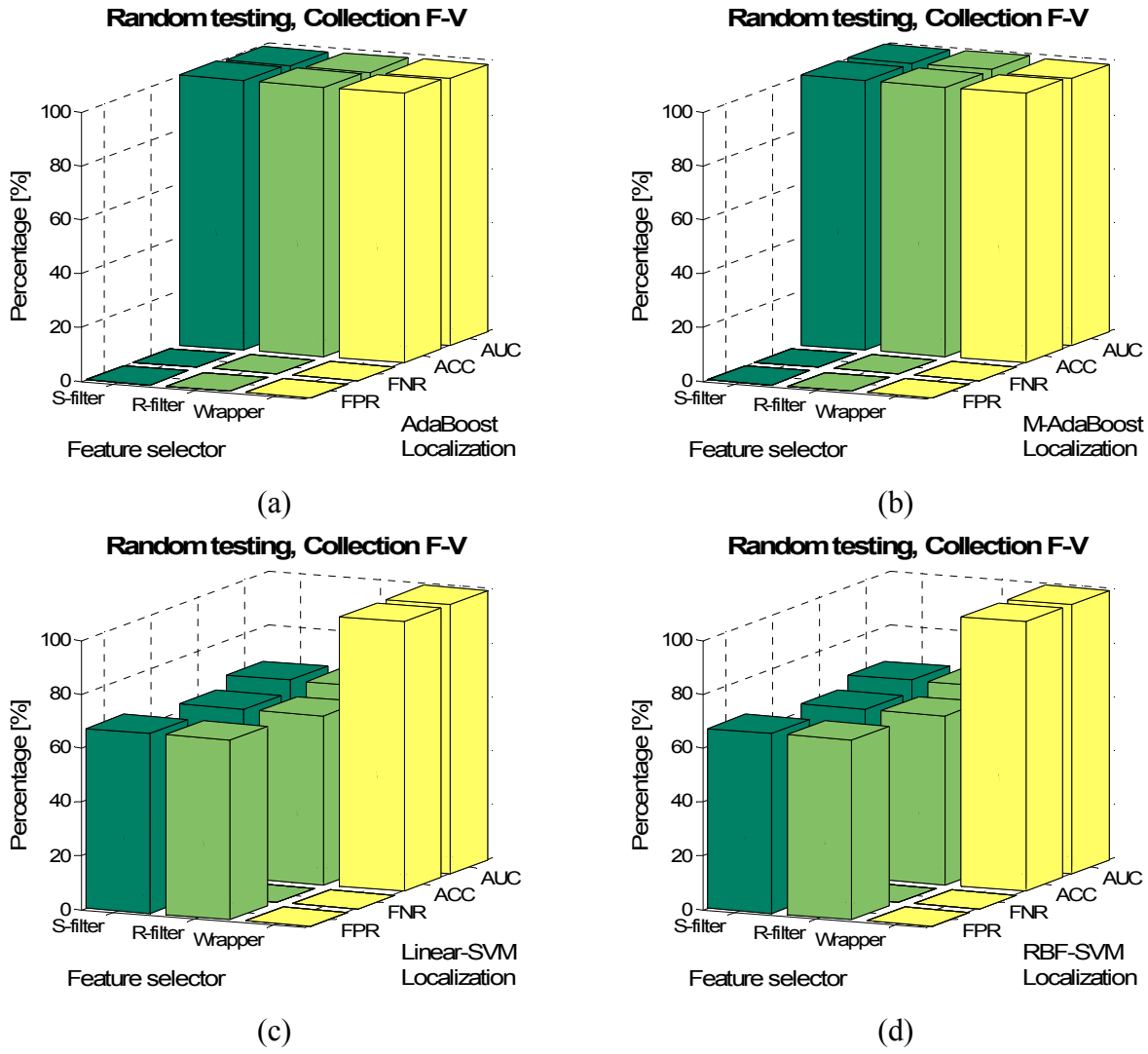


Figure 6.22 Scatterer localization results of the random testing on Collection F-V, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM.

(ACC: accuracy; FPR: false-positive rate; FNR: false-negative rate; AUC: area under an ROC curve; R-filter: a relaxed-filter feature selector; S-filter: a strict-filter feature selector; and Wrapper: a wrapper sequential backward feature selector.)

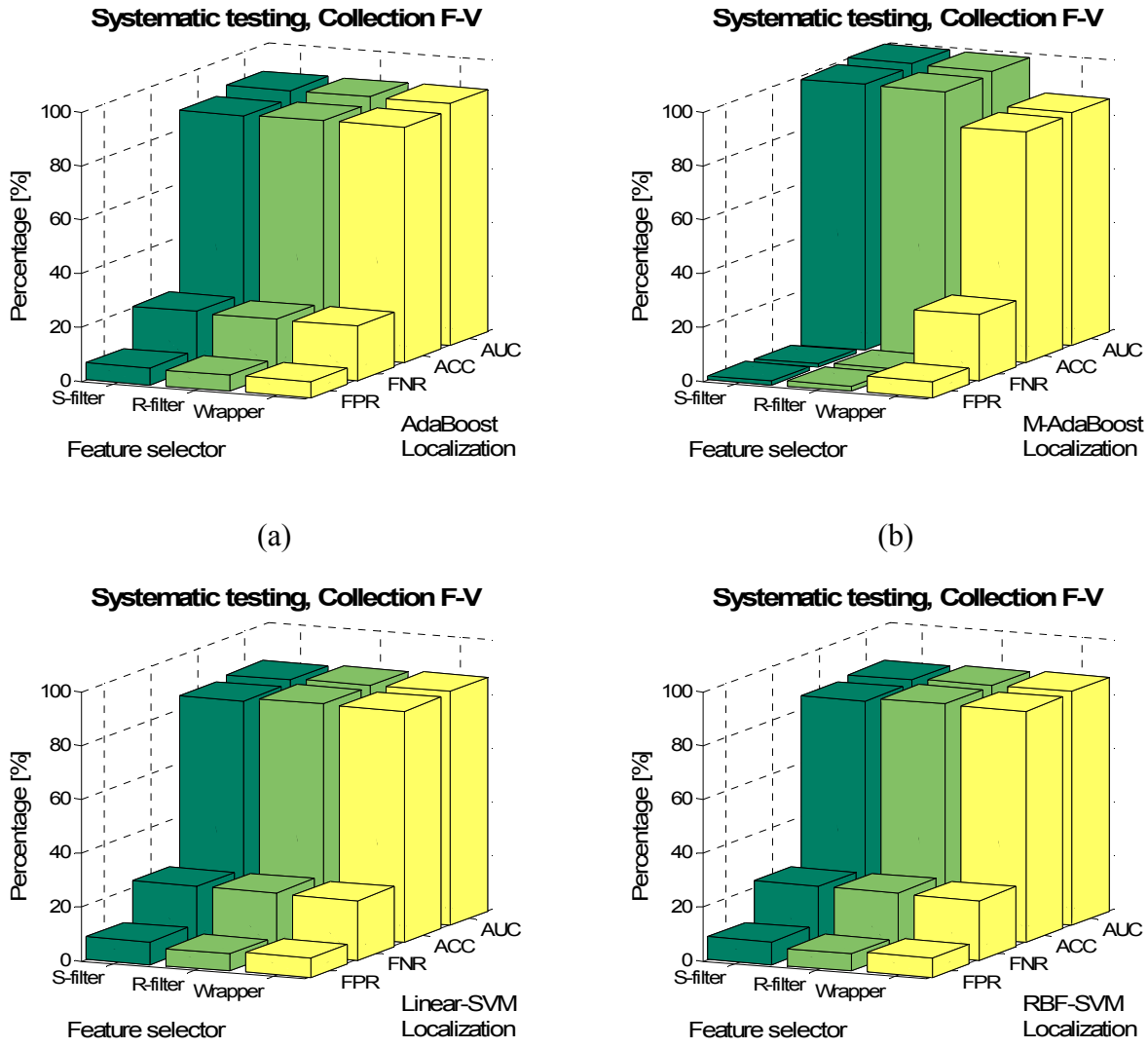


Figure 6.23 Scatterer localization results of the systematic testing on Collection F-V, by using (a) AdaBoost, (b) M-AdaBoost, (c) linear-SVM, and (d) RBF-SVM.

(ACC: accuracy; FPR: false-positive rate; FNR: false-negative rate; AUC: area under an ROC curve; R-filter: a relaxed-filter feature selector; S-filter: a strict-filter feature selector; and Wrapper: a wrapper sequential backward feature selector.)

6.3.4 Field Test 4: Detecting Different Scatterers

6.3.4.1 Data Collection and Machine Learning Test Formulation

Field Test 4 assessed the generality of the data-driven methods to different scatterers. In addition to the cylinder mass scatterer (denoted as Scatterer 1), an aluminum bar (denoted as Scatterer 2) was used to simulate a larger defect in the pipe, as shown in Figure 6.24.

Three cycles of scatterer toggling experiments were performed on 9/29/2011(Collection F-VI), with Scatterer 1 used in the first cycle, and Scatterer 2 used in the second and the third cycles. Unlike the three preceding field tests where the distance between the transmitter and receiver was 10 ft, Field Test 4 used a transducer pair with 20 ft distance. The data collection details are provided in Table 6-18. The measurements were taken at one-minute intervals, with a total duration of 24hr 1min, including 965 signals without the presence of the scatterer, 115 signals with the presence of Scatterer 1, and 351 signals with the presence of Scatterer 2. The energy variation over time of Collection F-VI was shown in Figure 6.25, with a dashed line indicating the toggling cycles.

Systematic testing was formulated as described in Table 6-19. The basis set, the training set, and the testing set were divided according to the presence of the different scatterers. More specifically, the basis set was formed by the measurements from the first toggling cycle corresponding to Scatterer 1, or by the ones from the second toggling cycle corresponding to Scatterer 2; the remaining measurements were divided into training data and testing data.

Figure 6.24 An aluminum bar (with the cylinder mass on the top) grease-coupled to the pipe to simulate larger damage (C. Liu et al. 2012).

Table 6-18 Field data collection details (Field Test 4).

	Undamaged datasets	Scatterer 1 datasets	Scatterer 2 datasets	Test duration	Toggling cycles	Scatterer locations
Collection F-VI (9/29/2011)	965	115	351	24hr 1min	3	3

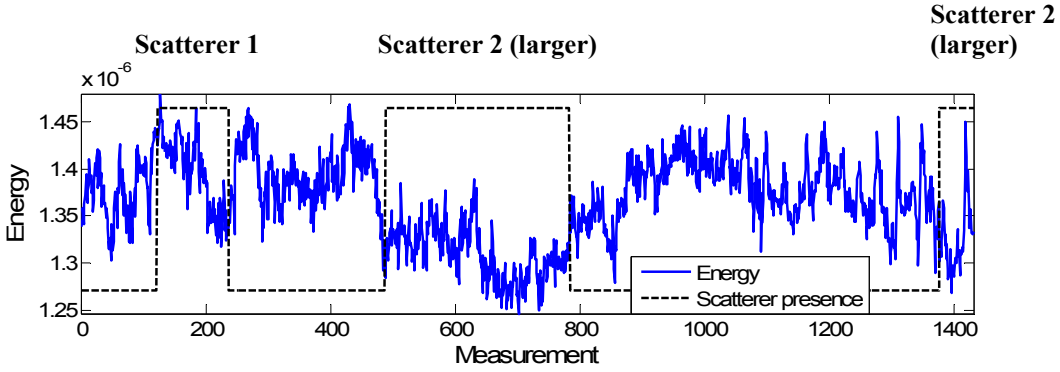


Figure 6.25 Energy variation over time of Collection F-VI.

Table 6-19 Machine learning test formulation for detection of different scatterers (Field Test 4).

	Single-day data
Systematic testing	Measurements of the first (or second) toggling cycle as a basis set, 2-fold systematic testing on the remaining data (divided into training and testing sets according to different scatterer locations). Total: 4 tests

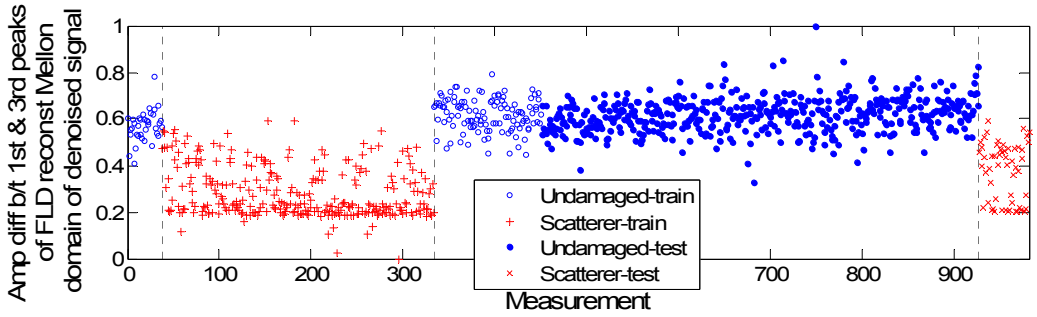
6.3.4.2 Testing Results

We examined the generality of the features to the different scatterers. Figure 6.26 shows the first feature selected by the wrapper method (associated with linear-SVM) in two systematic tests. This feature was extracted by combining FLD with wavelet processing and the Mellin transform.

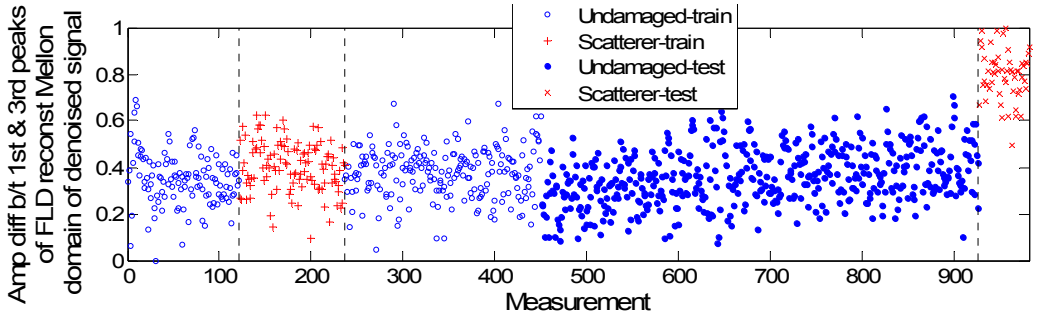
In the first systematic test, the measurements with the presence of Scatterer 1 were used in the basis set (to construct linear subspaces for feature extraction). Figure 6.26a shows that the selected feature presents a significant step change as Scatterer 2 was applied or removed from the pipe in the succeeding two toggling cycles, which allows visual recognition of the scatterer. In other words, we can use the linear subspaces constructed based on the Scatterer 1 data to reveal the pattern of the presence of a different (larger) scatterer under different operating conditions. The result also indicates that the linear-projection-based feature can be potentially used as a robust change detector for on-line monitoring in the field.

In the second systematic test, the measurements of the second toggling cycle (partial of the Scatterer 2 data) were used in the basis set. Figure 6.26b shows that the selected feature is able to detect the presence of Scatterer 2 in the third toggling cycle, but fails to distinguish Scatter 1 data from the undamaged measurements. This result is understandable, given that

Scatterer 2 is of much larger size compared to Scatterer 1. The result implies that a machine learning model learned based on the instances of a large defect might not be effective to detect the presence of a smaller defect. However, it should be noted that the distance between the transmitter and the receiver in Field Test 4 was 20ft, double of that in Field Tests 1-3. As a result, the change produced by the cylinder mass in the ultrasonic signal became much weaker and therefore more difficult to recognize.



(a)



(b)

Figure 6.26 Performance of the first selected feature by a wrapper method (associated with linear-SVM) in the systematic testing on Collection F-VI: (a) measurements of the first toggling cycle (Scatterer 1 data) as the basis set, (b) measurements of the second toggling cycle (partial of the Scatterer 2 data) as the basis set.

6.4 Summary

A guided wave ultrasonic sensing system was installed in a working hot water return pipe to investigate the feasibility of the data-driven SHM framework under real operating conditions in an environment that was electrically and mechanically noisy. A large amount of data was collected from the field instrumentation while the system operated normally. Analysis of the collected data showed significant variation in terms of energy and correlation coefficients. Tests were conducted by applying and removing a grease-coupled mass scatterer during a test period. The data-driven framework was demonstrated to effectively distinguish the change caused by the scatterer from change caused by ongoing variations. Similar to the laboratory observations, the hybrid features generated by combined signal processing and machine learning methods showed superior performance for scatterer detection.

7 CONCLUSIONS

7.1 Summary and Remarks

This work showed the effectiveness and the generality of the data-driven SHM framework to detect and localize damage (or a mass scatterer as a physical simulation of damage) in pipes in the presence of significant operational and environmental variability. Several well-known signal processing and statistical methods were employed to quantify the change in ultrasonic signals due to structural damages. However, it was observed in most cases that the change due to structural damage was masked by the changes due to benign operational and environmental variability. A data-driven framework based on pattern recognition and machine learning is then developed to differentiate between damage induced changes and changes induced by variability in the ambient conditions.

The effectiveness and robustness of a data-driven framework for SHM were experimentally demonstrated on a pipe under varying environmental and operational conditions. Laboratory tests were performed on a steel pipe with a mass scatterer placed at different locations and with benign variations in air pressure and ambient temperature. A total of 1500 features were extracted from the ultrasonic signals to create a feature library. For a specific pipe under specific circumstances, some features may be more effective than others. An integrated filter-wrapper feature selection approach automatically identifies the useful features for scatterer detection. Hybrid features constructed using combined signal processing and machine learning methods showed superior class-separability. The selected features may also suggest robust change detectors for real-time monitoring purposes – although it is surmised that the effectiveness of a given feature will depend on the structure and ambient conditions, thereby

justifying the use of machine learning techniques that will automatically pick the best features for a given circumstance instead of relying on a pre-defined feature that might be inferior under those circumstances. Four classification methods (adaptive boosting, modified adaptive boosting, support vector machines with a linear kernel, and support vector machines with a radial basis function kernel) were investigated to determine the presence of the mass scatterer, and achieved good results in the random testing and the systematic testing. Implementing multiple feature selection methods and classification algorithms were recommended for practice deployment in order to obtain a reliable decision. The results of laboratory studies show the promising applications of the integrated machine learning and signal processing methodology to pipe monitoring.

To verify the effectiveness of the data-driven SHM framework under real operating conditions in an electrically and mechanically noisy environment, a guided wave ultrasonic sensing system was installed in an in-service hot water return pipe in a building. A large amount of data was collected from the field instrumentation while the system operated normally. Analysis of the collected data showed significant variation in terms of energy and correlation coefficients. Tests were conducted by applying and removing a grease-coupled mass scatterer during a test period. The data-driven framework was found to effectively distinguish the change caused by the scatterer from the changes caused by ambient variations.

The generality of the data-driven SHM framework from different aspects is summarized in Table 7-1. The developed data-driven methods in this research were applied to different pipes, including a laboratory steel pipe specimen and segments of an in-service piping loop with much larger diameter and larger wall-thickness. The laboratory specimen was filled with pressurized

air, while the piping loop carried pressurized hot water with more dramatic change in temperature and flow rate. For both of the pipes, the data-driven framework was able to detect and localize the presence of a weak scatterer with good accuracy, although the most effective features and machine learning classifiers might vary in different test scenarios.

The data-driven methods were also generalizable to piezoelectric transducers of different size, with different longitudinal and circumferential distances between the transmitter and the receiver. Especially in the field tests, different pairs of transducers were used for data collection, including transmitters and receivers with 10 ft or 20 ft distance, 0 to180° offset along the circumferential direction; also, some transmitter-receiver pairs were located on a pipe segment with a circumferential weld in between. Moreover, the methodology developed in this work was not limited to specific frequency range: a broadband excitation was used to produce an extensive range of frequencies; the driven waveforms used in this study included sinc and Gaussian windowed sinusoidal pulses, and a chirp signal.

In addition, both the laboratory data and the field data were collected with the mass scatterer at different locations on the pipe surface. The systematic testing formulated in this dissertation showed that with the instances of the scatterer at certain locations on the pipe, a machine learning model was capable of recognizing the presence and the zone of the scatterer at other different locations. Furthermore, two different scatterers were applied to the pipe in the field tests; the result showed that with the instances of a scatterer, the methods developed in this work were also effective to detect a different scatterer of larger size.

The selected features described in this dissertation also demonstrated some degree of generality. No single “super” feature was found with good separability in all the scatterer

detection and localization tests, but the hybrid types of features (generated by the combination of signal processing and machine learning algorithms) were most frequently selected and of the highest ranks. For instance, the features extracted by applying Fisher's linear discriminant to wavelet domains were shown to be effective for scatterer detection and location in most field tests, in spite of different operating conditions (flow rate and temperature), different data-collection durations (from hours to months), different transducer confiscations (PZT locations and transmitter-receiver distance), and different scatterers (mass location and size).

Furthermore, the predictability of the machine learning classification methods were unaffected by the ratio of the undamaged and scatterer instances. The data-driven framework was first studied in the laboratory experiments, where the data of the two classes were approximately equal-size. In the field tests with unbalanced datasets, the data-driven methods still achieved good detection and localization results.

The overall comments on the data-driven methodology for damage detection and localization in pipe monitoring are concluded in Table 7-2.

Table 7-1 Generality of the data-driven SHM framework.

Category	Variable
Pipe	Diameter, wall-thickness, fluid content (pressurized air / hot-water) pipe segment with/without a weld
Operating condition	Internal air pressure, flow rate, temperature, mechanical and electrical noise
PZT transducer	Dimensions, transmitter-receiver distance
Excitation signal	Frequency, waveform
Scatterer	Location, scatterer size
Datasets	Numbers of undamaged instances and scatterer instances (balanced data / unbalanced data)

Table 7-2 Remarks on data-driven methods for pipe monitoring.

Data-Driven Approaches	
Advantages	Automated damage detection and localization with high accuracy
	Efficient for identifying feature or feature combinations that are sensitive to damage while robust to environmental and operating variations
	Possibly provides insights into the physical characteristics of the ultrasonic system
	Free of wave mode selection, no restriction to transducers or excitation signals
	Suitable when the environmental and operating conditions are unknown or difficult to measure
Challenges	Requires training of the machine learning model
	Requires validation in order for generality of the approaches
	Sometimes difficult to provide physical interpretations

7.2 Contributions and Broader Impacts

The research presented in this dissertation has practical engineering relevance as well as scientific significance. The major contribution lies in the development of the data-driven structural health monitoring framework that provides effective and robust methodology for ultrasonic monitoring of steels pipes in the presence of environmental and operational variations.

The detailed contributions of this research are threefold (Figure 7.1). First, the research presented in the thesis will make significant progress towards continuous monitoring of critical pipe systems, such as pressurized steam piping systems in power plants, and natural gas pipes especially in densely populated areas. As pipe structures age, even though replacement is deferred, continuous monitoring based on installed sensors will play a vital role in preventing catastrophic failure of pipes. On the other hand, pipe replacement will also create opportunities for embedding new sensing materials and instruments into the pipe manufacture and construction process. While current quality control of pipes is performed through scheduled inspections using relatively expensive sensor equipment and expert involvement, the emphasis of this research is to achieve automated and continuous monitoring of pipes using inexpensive PZT wafers. Another emphasis of this thesis is on the development of the software side of pipe monitoring, whereas most current research is focused on transducer design and placement on pipes.

The second contribution would be in the field of health monitoring of structures in general. A majority of research work in SHM is done on structures with simple geometries (such as plates) under controlled laboratory conditions. In contrast, the applicability of the data-driven framework developed in this dissertation has been demonstrated on an in-service hot-water pipe in a building under significant operational and environmental variations (viz. temperature, fluid

pressure and fluid flow in the pipe, etc.) as well as mechanical and electrical noise. Another unique feature of this study lies in the choice of probing ultrasonic signals. The general practice in SHM is to fine tune the excitation frequency of the probing wave and limit the number of propagating modes in order to simplify the interpretation of received ultrasonic signals by human experts. This practice leads to either increased cost of transducers or limits the applicability of the technique in complex structures or both. In this thesis, no restriction has been made on the choice of frequency or wave modes. Instead, a large pool of candidate features was considered and machine learning techniques were used to select the most effective features (in contrast to the current practice in pipeline inspection where only a small number of features quantifying the change in ultrasonic signals due to damage are picked to allow interpretation by maintenance personnel). Formulating SHM problems in the statistical pattern recognition paradigm has been investigated in the past. However, to the best of the author's knowledge, this is the first attempt to apply machine learning techniques to extremely complex ultrasonic signals in a cylindrical shell waveguide to detect and also localize damage in the presence of multiple variable operational and environmental conditions.

Thirdly, this research work broadens the scope of machine learning applications to the area of civil infrastructure monitoring and proposes a set of new algorithms, which are mainly integration of existing statistical, signal processing, and machine learning methods, suitable for SHM applications. While machine learning techniques are widely used on large datasets with distinguishable characteristics between classes, this work demonstrates the applicability of machine learning to distinguish classes with only subtle differences in characteristics that are often buried under environmental and operational variations. The impacts described above are summarized in Table 7-3.

Pipe monitoring is an integrated process of sensing, data acquisition, signal analysis and decision making, for continuous tracking of the structural functionality in a cost-effective and long-lasting manner. This research is multidisciplinary, involving comprehensive knowledge and techniques in the areas of wave propagation, signal processing, and machine learning, and exemplifies the opportunities for scientific findings and for impact on engineering practice that arise when novel technologies are introduced. The techniques developed in this research may have broader applications related to the regular inspection, maintenance, repair and management of critical infrastructures other than pipes, such as bridges, highways and buildings.

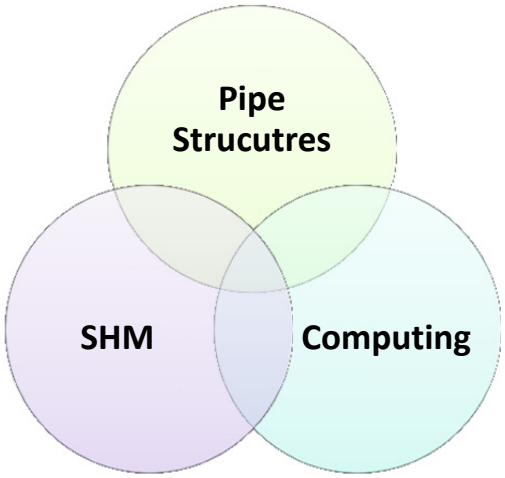


Figure 7.1 Threefold contributions of this research.

Table 7-3 Summary of the detailed impacts of this research.

Category		Majority of Existing Technologies	Technology Studied in This Research
Pipes	Solution	Inspection (schedule-based)	Monitoring (condition-based)
	Sensors	Relatively expensive and temporary devices	Permanently installed, small-size PZT sensors, low power requirements
	Technical focus	Transducer/hardware design	Data interpretation/software design
	Algorithms	Model-driven	Data-driven
Structural Health Monitoring	Structures	Plate-like simple geometries	Cylindrical shell
	Benign variations	Under constant ambient variations, or only one or two factors considered	Internal air pressure, temperature, flow rate, and combination of variations
	Excitation	Narrow band	No restrictions on frequency band
	Mode selection	Separating and selecting a small number of modes	No restrictions on wave modes
	Features	Limited number of features, and limited work on feature selection	A feature library, and automated feature selection process
	Algorithms	Limited efforts on machine learning	Comparing different machine learning algorithms
	Validation	Mostly laboratory experiments	Field tests
Computing	Application	Limited applications to infrastructure monitoring, especially to analysis of ultrasonic signals	Extending to ultrasonic infrastructure monitoring applications
	Data sets	Visually distinguishable characteristics between different classes	Complex patterns, subtle changes buried beneath significant benign changes
	Algorithms	-	Integrating statistics, signal processing, and machine learning

7.3 Future Research

Machine Learning Methods for Real Damage

In our prior laboratory experiments, we found that the degree of change caused by a mass scatterer used in this study is close to the degree of change produced by a small saw-cut (11 mm long, 1.2 mm wide, and 0.8 mm deep) on a pipe with respect to several different features, such as curve length and correlation coefficient. Further experimental work is required to validate the effectiveness of the data-driven framework on real structural defects such as cracks, corrosion and joint malfunctions.

In addition, many features generated in this study as well as the four classification methods are based on supervised learning, which means that the measurements taken on a damaged specimen are required. In real applications, measurements from damaged structures will be unavailable upfront. Future research can be done to investigate the capability of a scatterer-trained model in detecting actual damage. Ample data can be collected with a mass scatterer on a specimen to generate machine learning models (these machine learning models should be verified by creating a real damage on the pipe). Another approach would be to use one-class classification methods or novelty detection (Moya and Hush 1996), where a machine learning classifier can be constructed only based on undamaged records to recognize any anomaly in the system; the baseline-free features and the unsupervised PCA-based features generated in this study can be fed into such machine learning models.

Physics-based Insight in Data-Driven Analysis

The data-driven framework developed in this thesis uses a large number of features. However, if only a specific section of a pipe is considered, we are likely to find a smaller subset of features that are consistently effective for providing high damage detectability. By understanding the physics of the structure-sensing systems, we may manually exclude some of the features in the machine learning implementation, in order to save computation efforts. For instance, we observe that the first or first few PCA projections usually reflect the large environmental variations; therefore, these features can be removed from the feature candidate pool.

Economical Analysis and Algorithms Adjustment

The advantages of the ROC curves are not fully explored in the dissertation. The ROC curves are associated with the cost/benefit analysis and decision making. Provided the costs of false positives and false negatives, we can obtain a more cost-effective model by adjusting the decision threshold of a machine learning classifier. In practice, while devising the algorithms, we may want to place more emphasis on preventing false negatives that may lead to severe accidents. Moreover, we may consider maximizing AUC instead of ACC when devising a classification algorithm which might be more suitable for unbalanced data (Fawcett 2004).

Information Fusion

The results presented in this study are based on ultrasonic signals obtained using a single transducer-receiver channel at a time. It would be interesting to fuse the results from multiple PZT receivers. Distributed sensors can capture different characteristics of guided waves propagating along multiple channels and interacting with damage at different angles. In addition,

wrong decisions made due to a malfunctioning sensor can be largely eliminated. It would also be beneficial to utilize any available direct measurements of environmental parameters, which may eliminate certain false-positive alarms. Furthermore, methodology for decision fusion can be developed to reconcile and merge the results from the different machine learning classifiers into a final global decision.

Damage Characterization

Future research should be extended to the third stage of the SHM data-driven framework – damage characterization, in order to support decision making and cost-effective management. The feature library, the feature selection methods, and the classification algorithms developed in this dissertation should be applicable for damage characterization, while additional features could be explored based on the physical understanding of certain types of defects. The key task also includes collecting ample data sets from laboratory or field tests, to evaluate the reliability and generality of the developed data-driven system.

REFERENCES

Agrawal, A., Harley, J. B., Ying, Y., Garrett, J. H., Sohn, H., and Soibelman, L. (2009). "Preliminary Studies on the Dispersion of Signals Produced by Permanently Installed MFC Transducers for Pipeline Monitoring." *16th Workshop of the European Group for Intelligent Computing in Engineering (EG-ICE)*, Berlin, Germany.

Alleyne, D. N., and Cawley, P. (1992a). "Optimization of Lamb Wave Inspection Techniques." *NDT & E International (ISSN 0963-8695), vol. 25, no. 1, p. 11-22.*, 11–22.

Alleyne, D. N., and Cawley, P. (1992b). "The Interaction of Lamb Waves with Defects." *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, 39(3), 381-397.

Auld, B. A. (1990). *Acoustic Fields and Waves in Solids, Vol. 2*. Krieger Pub Co.

Belhumeur, P. N., Hespanha, J. P., and Kriegman, D. J. (1997). "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection." *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 19(7), 711 –720.

Bhalla, S., and Kiong Soh, C. (2003). "Structural Impedance Based Damage Diagnosis by Piezo-Transducers." *Earthquake Engineering & Structural Dynamics*, 32(12), 1897-1916.

Boller, C., Chang, F. K., and Fujino, Y. (Eds.). (2009). *Encyclopedia of Structural Health Monitoring*. John Wiley & Sons, Ltd, Chichester, UK.

Bradley, A. P. (1997). "The use of the area under the ROC curve in the evaluation of machine learning algorithms." *Pattern Recognition*, 30(7), 1145–1159.

Buck, O., Morris, W. L., and Richardson, J. M. (1978). "Acoustic Harmonic Generation at Unbonded Interfaces and Fatigue Cracks." *Applied Physics Letters*, 33(5), 371.

Burges, C. J. C. (1998). "A Tutorial on Support Vector Machines for Pattern Recognition." *Data Mining and Knowledge Discovery*, 2(2), 121–167.

Catbas, F. N., Kijewski-Correa, T., and Aktan, A. E. (2011). *Structural Identification (ST-Id) of Constructed Facilities*. American Society of Civil Engineers (ASCE) Structural Engineering Institute (SEI).

Cawley, P. (2003). "Practical Long Range Guided Wave Inspection—Managing Complexity." *AIP Conference Proceedings*, 22.

Cegla, F. B., Rohde, A., and Veidt, M. (2008). "Analytical Prediction and Experimental Measurement for Mode Conversion and Scattering of Plate Waves at Non-Symmetric Circular Blind Holes in Isotropic Plates." *Wave Motion*, 45(3), 162-177.

Chang, C. C., and Lin, C. J. (2011). "LIBSVM : A Library for Support Vector Machines." *ACM Transactions on Intelligent Systems and Technology*, 2(3), 27:1--27:27.

Cho, Y. (2000). "Estimation of Ultrasonic Guided Wave Mode Conversion in a Plate with Thickness Variation." *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, 47(3), 591-603.

Cortes, C., and Vapnik, V. (1995). "Support-Vector Networks." *Machine Learning*, 20(3), 273-297.

Crawley, E. F., and De Luis, J. (1987). "Use of Piezoelectric Actuators as Elements of Intelligent Structures." *AIAA Journal*, 25(10), 1373-1385.

Croxford, A. J., Wilcox, P. D., Drinkwater, B. W., and Konstantinidis, G. (2007). "Strategies for Guided-Wave Structural Health Monitoring." *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 463(2087), 2961 -2981.

Dasarathy, B. V. (1994). *Decision fusion*. IEEE Computer Society Press.

Demma, A., Cawley, P., Lowe, M., and Roosenbrand, A. G. (2003). "The Reflection of the Fundamental Torsional Mode from Cracks and Notches in Pipes." *The Journal of the Acoustical Society of America*, 114, 611.

Donoho, D. L. (1995). "De-noising by soft-thresholding." *Information Theory, IEEE Transactions on*, 41(3), 613 –627.

DOT. (2011). "Press Release: U.S Transportation Secretary Ray LaHood Announces Pipeline Safety Action Plan." *U.S. Department of Transportation*, <http://www.dot.gov/affairs/2011/dot4111.html> (Jun. 1, 2011).

Dutta, D. (2010). "Ultrasonic Techniques for Baseline-Free Damage Detection in Structures." *Dissertations*.

e-CFR. (2011). *Electronic Code of Federal Regulations: Title 49. Transportation*. *Electronic Code of Federal Regulations*.

EIA. (2009). "U.S. Natural Gas Pipeline Network Map." <http://www.eia.doe.gov/pub/oil_gas/natural_gas/analysis_publications/ngpipeline/ngpipelines_map.html> (May. 25, 2011).

Fan, R. E., Chen, P. H., and Lin, C. J. (2005). "Working Set Selection using Second Order Information for Training Support Vector Machines." *The Journal of Machine Learning Research*, 6, 1889–1918.

Farrar, C. R., Duffey, T. A., Doebling, S. W., and Nix, D. A. (1999). "A statistical pattern recognition paradigm for vibration-based structural health monitoring." *Structural Health Monitoring*, 764–773.

Faundez-Zanuy, M. (2005). "Data fusion in biometrics." *Aerospace and Electronic Systems Magazine, IEEE*, 20(1), 34–38.

Fawcett, T. (2004). "ROC graphs: Notes and practical considerations for researchers." *Machine Learning*, 31, 1–38.

Fink, M. (1992). "Time reversal of ultrasonic fields—Part I: Basic principles." *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, 39(5), 555–566.

GBRA. (2005). *Guadalupe-Blanco River Authority : Hazard Mitigation Final Plan - Section 14. Energy Pipeline Failures*.

Giurgiutiu, V. (2005). "Tuned Lamb wave excitation and detection with piezoelectric wafer active sensors for structural health monitoring." *Journal of Intelligent Material Systems and Structures*, 16(4), 291.

Giurgiutiu, V., and Bao, J. (2004). "Embedded-ultrasonics Structural Radar for In Situ Structural Health Monitoring of Thin-wall Structures." *Structural Health Monitoring*, 3(2), 121 – 140.

Giurgiutiu, V., and Zagrai, A. N. (2000). "Characterization of Piezoelectric Wafer Active Sensors." *Journal of Intelligent Material Systems and Structures*, 11(12), 959 –976.

Giurgiutiu, V., Zagrai, A. N., and Bao, J. (2002). "Piezoelectric Wafer Embedded Active Sensors for Aging Aircraft Structural Health Monitoring." *Structural Health Monitoring*, 1(1), 41 –61.

Goulet, J. A., Kripakaran, P., and Smith, I. F. C. (2010). "Multimodel Structural Performance Monitoring." *Journal of Structural Engineering*, 136(10), 1309.

Graff, K. F. (1975). *Wave motion in elastic solids*. Dover Pubns.

Guo, W., Garrett, J. H., de Oliveira, D. P., and Soibelman, L. (2008). "Integrating traditional inspection data and new technologies for automated/improved condition assessment for sewer maintenance." *European Group for Intelligent Computing in Engineering (EG-ICE) International Workshop 2008*, Plymouth, UK.

Guyon, I., and Elisseeff, A. (2003). "An introduction to variable and feature selection." *The Journal of Machine Learning Research*, 3, 1157–1182.

Hall, M. A. (1999). "Correlation-based feature selection for machine learning." The University of Waikato.

Harley, J. B., O'Donoughue, N., Jin, Y., and Moura, J. M. F. (2010). "Time Reversal Focusing for Pipeline Structural Health Monitoring." San Antonio, Texas, 030001–030001.

Harley, J. B., O'Donoughue, N., States, J., Ying, Y., Garrett, J. H., Jin, Y., Moura, J. M. F., Oppenheim, I. J., and Soibelman, L. (2009). "Focusing of Ultrasonic Waves in Cylindrical Shells using Time Reversal." *Proceedings of the 7th International Workshop on Structural Health Monitoring*, Stanford, CA.

Harley, J. B., Ying, Y., Moura, J. M. F., Oppenheim, I. J., Soibelman, L., and Garrett Jr., J. H. (2011). "Application of Mellin Transform Features for Robust Ultrasonic Guided Wave Structural Health Monitoring." *Review of Progress in Quantitative Nondestructive Evaluation*, Burlington, VT.

Hernández-Gómez, L. H., Durodola, J. F., Fellows, N. A., and Urriolagoitia-Calderón, G. (2005). "Locating defects using dynamic strain analysis and artificial neural networks." *Applied Mechanics and Materials, Vols*, 3–4.

Hillis, A. J., Neild, S. A., Drinkwater, B. W., and Wilcox, P. D. (2006). "Global crack detection using bispectral analysis." *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 462(2069), 1515 –1530.

Ihn, J. B., and Chang, F. K. (2004). "Detection and monitoring of hidden fatigue crack growth: I. Diagnostics." *Smart Materials and Structures*, 13, 609–620.

Ing, R. K., and Fink, M. (1998). "Time-reversed Lamb waves." *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 45(4), 1032–1043.

Jeon, B., and Landgrebe, D. A. (1999). "Decision fusion approach for multitemporal classification." *Geoscience and Remote Sensing, IEEE Transactions on*, 37(3), 1227–1233.

Jin, Y., and Moura, J. M. F. (2009). "Time-Reversal Detection Using Antenna Arrays." *IEEE Transactions on Signal Processing*, 57(4), 1396–1414.

Jin, Y., O'Donoughue, N., Moura, J. M. F., Harley, J. B., Garrett, J. H., Oppenheim, I. J., Soibelman, L., Ying, Y., and He, L. (2010). "Cognitive sensor networks for structure defect monitoring and classification using guided wave signals." San Diego, CA, USA, 76473T–76473T–12.

Kim, S. B., and Sohn, H. (2008). "Instantaneous crack detection using dual PZT transducers." *Proceedings of SPIE*, San Diego, California, USA, 693509–693509–12.

Kohavi, R., and John, G. H. (1997). "Wrappers for Feature Subset Selection." *ARTIFICIAL INTELLIGENCE*, 97(1), 273–324.

Kohavi, R., and Provost, F. (1998). "Glossary of terms." *Machine Learning*, 30(June), 271–274.

Konstantinidis, G., Drinkwater, B. W., and Wilcox, P. D. (2006). "The temperature stability of guided wave structural health monitoring systems." *Smart Materials and Structures*, 15(4), 967–976.

Lamb, H. (1917). "On waves in an elastic plate." *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character*, 93(648), 114–128.

Lee, B. C., and Staszewski, W. J. (2003). "Modelling of Lamb waves for damage detection in metallic structures: Part II. Wave interactions with damage." *Smart Materials and Structures*, 12, 815.

Li, J., and Rose, J. L. (2001a). "Excitation and propagation of non-axisymmetric guided waves in a hollow cylinder." *The Journal of the Acoustical Society of America*, 109(2), 457–64.

Li, J., and Rose, J. L. (2001b). "Implementing guided wave mode control by use of a phased transducer array." *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, 48(3), 761–768.

Liu, C., Harley, J. B., O'Donoughue, N., Ying, Y., Altschul, M. H., Garrett, J. H., Moura, J. M. F., Oppenheim, I. J., and Soibelman, L. (2012). "Ultrasonic monitoring of a pipe under operating conditions." *Proceedings of SPIE*, 8345(1), 83450B–83450B–12.

Lowe, M. J. S., Alleyne, D. N., and Cawley, P. (1998). "Defect detection in pipes using guided waves." *Ultrasonics*, 36(1-5), 147–154.

Lu, Y., and Michaels, J. E. (2008). "Numerical implementation of matching pursuit for the analysis of complex ultrasonic signals." *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 55(1), 173–182.

Lu, Y., and Michaels, J. E. (2009). "Feature Extraction and Sensor Fusion for Ultrasonic Structural Health Monitoring Under Changing Environmental Conditions." *Sensors Journal, IEEE*, 9(11), 1462–1471.

Luo, W., and Rose, J. L. (2007). "Phased array focusing with guided waves in a viscoelastic coated hollow cylinder." *The Journal of the Acoustical Society of America*, 121(4), 1945–1955.

Mallat, S. G. (1989). "A theory for multiresolution signal decomposition: the wavelet representation." *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 11(7), 674 –693.

Mallat, S. G. (1999). *A wavelet tour of signal processing*. Academic Press.

Michaels, J. E., Cobb, A. C., and Michaels, T. E. (2004). "A comparison of feature-based classifiers for ultrasonic structural health monitoring." *Proceedings of SPIE, Health Monitoring and Smart Nondestructive Evaluation of Structural and Biological Systems III*, 363–374.

Michaels, J. E., and Michaels, T. E. (2005). "Detection of structural damage from the local temporal coherence of diffuse ultrasonic signals." *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, 52(10), 1769– 1782.

Moura, J. M. F., and Jin, Y. (2007). "Detection by time reversal: single antenna." *IEEE Transactions on Signal processing*, 55(1), 187–201.

Moya, M. M., and Hush, D. R. (1996). "Network constraints and multi-objective optimization for one-class classification." *Neural Networks*, 9(3), 463–474.

Nestleroth, J. B., and Bubenik, T. A. (1999). "Magnetic flux leakage (MFL) technology for natural gas pipeline inspection." *Battelle, Report Number GRI-00/0180 to the Gas Research Institute*.

NETL. (2007). "DOE-Funded Pipeline Robot Revolutionizes Inspection Process." <http://www.netl.doe.gov/publications/press/2007/071204-Pipeline_Robot_Revolutionizes_Inspection_Process.html> (May. 30, 2011).

Nienwenhuis, J. H., Neumann, J. J., Greve, D. W., and Oppenheim, I. J. (2005). "Generation and detection of guided waves using PZT wafer transducers." *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, 52(11), 2103–2111.

Nunez, I., and Negreira, C. (2005). "Efficiency parameters in time reversal acoustics: Applications to dispersive media and multimode wave propagation." *The Journal of the Acoustical Society of America*, 117(3), 1202–1209.

O'Donoughue, N., Harley, J. B., Moura, J. M. F., Jin, Y., Oppenheim, I. J., Ying, Y., States, J., Garrett, J. H., and Soibelman, L. (2009). "Single Antenna Time Reversal of Guided Waves in Pipelines." *Proceedings of Meetings on Acoustics*, 065001–11.

Oh, C. K., and Sohn, H. (2009). "Damage diagnosis under environmental and operational variations using unsupervised support vector machine." *Journal of Sound and Vibration*, 325(1-2), 224–239.

Oppenheim, A. V., and Schafer, R. W. (1989). *Discrete-Time Signal Processing*. Prentice-Hall, Englewood Cliffs, NJ.

Parfomak, P. W. (2011). "Keeping America's Pipelines Safe and Secure: Key Issues for Congress." *Congressional Research Service (CRS) Reports*.

Park, G., Sohn, H., Farrar, C. R., and Inman, D. J. (2003). "Overview of piezoelectric impedance-based health monitoring and path forward." *Shock and Vibration Digest*, 35(6), 451–464.

Park, H. W., Sohn, H., Law, K. H., and Farrar, C. R. (2007). "Time reversal active sensing for health monitoring of a composite plate." *Journal of Sound and Vibration*, 302(1-2), 50–66.

PHMSA. (2011a). "PHMSA Stakeholder Communications: Glossary." <http://primis.phmsa.dot.gov/comm/glossary/index.htm?nocache=9390#Defect> (May. 26, 2011).

PHMSA. (2011b). “Pipeline Incidents and Mileage Reports.” <http://primis.phmsa.dot.gov/comm/reports/safety/PSI.html?nocache=6574> (May. 29, 2011).

PHMSA. (2011c). “Natural Gas Transmission, Gas Distribution, and Hazardous Liquid Pipeline Annual Mileage.” *PHMSA - US Department of Transportation Pipeline and Hazardous Materials Safety Administration*, <http://www.phmsa.dot.gov/portal/site/PHMSA/menuitem.ebdc7a8a7e39f2e55cf2031050248a0c/?vgnextoid=036b52edc3c3e110VgnVCM1000001ecb7898RCRD&vgnextchannel=3b6c03347e4d8210VgnVCM1000001ecb7898RCRD&vgnextfmt=print> (Jun. 1, 2011).

Pierce, S. G., Culshaw, B., Manson, G., Worden, K., and Staszewski, W. J. (2000). “Application of ultrasonic Lamb wave techniques to the evaluation of advanced composite structures.” *Proceedings of SPIE*, 93.

Piro, T. J. (2010). *FINAL Technology Status Assessment Report Instrumented Pipeline Initiative*. Concurrent Technologies Corporation and Carnegie Mellon University.

Platt, J. C. (1999). “Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods.” *Advances in Large-Margin Classifiers*, MIT Press, 61–74.

Posenato, D., Kripakaran, P., Inaudi, D., and Smith, I. F. C. (2010). “Methodologies for model-free data interpretation of civil engineering structures.” *Computers & Structures*, 88(7-8), 467–482.

Posenato, D., Lanata, F., Inaudi, D., and Smith, I. F. C. (2008). “Model-free data interpretation for continuous monitoring of complex structures.” *Advanced Engineering Informatics*, 22(1), 135–144.

Prada, C., and Fink, M. (1998). “Separation of interfering acoustic scattered signals using the invariants of the time-reversal operator. Application to Lamb waves characterization.” *The Journal of the Acoustical Society of America*, 104, 801.

Prada, C., Manneville, S., Spoliansky, D., and Fink, M. (1996). “Decomposition of the time reversal operator: Detection and selective focusing on two scatterers.” *Journal of the Acoustical Society of America*, 99(4), 2067–2076.

Rao, C. R. (1964). “The Use and Interpretation of Principal Component Analysis in Applied Research.” *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)*, 26(4), 329–358.

Rivola, A. (2000). “Comparison between second and higher order spectral analysis in detecting structural damages.” *Proceedings of Seventh International Conference on Recent Advances in Structural Dynamics*, 937–950.

Rizzo, P., Bartoli, I., Marzani, A., and di Scalea, F. L. (2005). “Defect Classification in Pipes by Neural Networks Using Multiple Guided Ultrasonic Wave Features Extracted After Wavelet Processing.” *Journal of Pressure Vessel Technology*, 127(3), 294–303.

Rizzo, P., Cammarata, M., Dutta, D., and Sohn, H. (2009). “An unsupervised learning algorithm for fatigue crack detection in waveguides.” *Smart Materials and Structures*, 18(2), 025016.

Rose, J. L. (1999). *Ultrasonic waves in solid media*. Cambridge University Press.

Rose, J. L., Pelts, S. P., and Quarry, M. J. (1998). "A comb transducer model for guided wave NDE." *Ultrasonics*, 36(1-5), 163–169.

Ross, A., and Jain, A. (2003). "Information fusion in biometrics." *Pattern Recognition Letters*, 24(13), 2115–2125.

Seco, F., Martiın, J. M., Jiménez, A., Pons, J. L., Calderón, L., and Ceres, R. (2002). "PCDISP: a tool for the simulation of wave propagation in cylindrical waveguides." *Proceedings of the 9th International Congress on Sound and Vibration*.

De Sena, A., and Rocchesso, D. (2007). "A Fast Mellin and Scale Transform." *EURASIP Journal on Advances in Signal Processing*, 2007, 1–10.

Sepehry, N., Shamshirsaz, M., and Bastani, A. (2011). "Experimental and Theoretical Analysis in Impedance-Based Structural Health Monitoring with Varying Temperature." *Structural Health Monitoring*, 10(6), 573–585.

Shull, P. J. (2002). *Nondestructive Evaluation: Theory, Techniques, and Applications*. CRC.

Sohn, H. (2007). "Effects of environmental and operational variability on structural health monitoring." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851), 539–560.

Sohn, H., Dutta, D., and An, Y. K. (2011). "Temperature Independent Damage Detection in Plates Using Redundant Signal Measurements." *Journal of Nondestructive Evaluation*, 30(2), 106–116.

Sohn, H., Park, H. W., Law, K. H., and Farrar, C. R. (2007a). "Damage detection in composite plates by using an enhanced time reversal method." *Journal of Aerospace Engineering*, 20, 141.

Sohn, H., Park, H. W., Law, K. H., and Farrar, C. R. (2007b). "Combination of a Time Reversal Process and a Consecutive Outlier Analysis for Baseline-free Damage Diagnosis." *Journal of Intelligent Material Systems and Structures*, 18(4), 335–346.

Sohn, H., Worden, K., and Farrar, C. R. (2002). "Statistical Damage Classification Under Changing Environmental and Operational Conditions." *Journal of Intelligent Material Systems and Structures*, 13(9), 561–574.

Su, Z., Ye, L., and Lu, Y. (2006). "Guided Lamb waves for identification of damage in composite structures: a review." *Journal of sound and vibration*, 295(3-5), 753–780.

Sun, Q. S., Zeng, S. G., Liu, Y., Heng, P. A., and Xia, D. S. (2005). "A new method of feature fusion and its application in image recognition." *Pattern Recognition*, 38(12), 2437–2448.

Taha, M. M. R., Noureldin, A., Lucero, J. L., and Baca, T. J. (2006). "Wavelet Transform for Structural Health Monitoring: A Compendium of Uses and Features." *Structural Health Monitoring*, 5(3), 267 –295.

Wang, C. H., Rose, J. T., and Chang, F. K. (2004). "A synthetic time-reversal imaging method for structural health monitoring." *Smart Materials and Structures*, 13(2), 415–423.

Weidlinger Associates Inc. (2009). *PZFlex Virtual Prototyping - Wave Propagation Analysis Software*. Weidlinger Associates Inc., Los Altos, CA.

Welch, B. L. (1947). “The Generalization of ‘student’s’ Problem When Several Different Population Varlances Are Involved.” *Biometrika*, 34(1-2), 28–35.

Worden, K., and Manson, G. (2007). “The application of machine learning to structural health monitoring.” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365(1851), 515 –537.

Worden, K., and Staszewski, W. J. (2000). “Impact Location and Quantification on a Composite Panel using Neural Networks and a Genetic Algorithm.” *Strain*, 36(2), 61–68.

Xing, E. (2010). “Machine Learning Lecture Notes: Support Vector Machines.”

Yan, F., Royer, R. L., and Rose, J. L. (2010). “Ultrasonic Guided Wave Imaging Techniques in Structural Health Monitoring.” *Journal of Intelligent Material Systems and Structures*, 21(3), 377 –384.

Yang, J., Yang, J. Y., Zhang, D., and Lu, J. (2003). “Feature fusion: parallel strategy vs. serial strategy.” *Pattern Recognition*, 36(6), 1369–1381.

Ying, Y., Garrett, J. H., Harley, J. B., Oppenheim, I. J., Shi, J., and Soibelman, L. (2012). “Damage Detection in Pipes under Changing Environmental Conditions Using Embedded Piezoelectric Transducers and Pattern Recognition Techniques.” *Journal of Pipeline Systems Engineering and Practice*, accepted 26 March 2012; posted ahead of print 29 March 2012.

Ying, Y., Harley, J. B., Garrett, J. H., Jin, Y., Moura, J. M. F., O’Donoughue, N., Oppenheim, I. J., and Soibelman, L. (2010). “Time reversal for damage detection in pipes.” *Proceedings of SPIE: Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2010*, San Diego, CA, USA, 76473S–76473S–12.

Ying, Y., Harley, J. B., Garrett, J. H., Jin, Y., Oppenheim, I. J., Shi, J., and Soibelman, L. (2011). “Applications of Machine Learning in Pipeline Monitoring.” *ASCE International Workshop on Computing in Civil Engineering*, Miami, FL.

APPENDICES

Appendix A: Comparison of Non-Destructive Evaluation Technologies

Table A-1 List of common non-destructive evaluations techniques (Shull 2002).

	Principles	Advantages	Limitations
Acoustic Emission	When a solid material is stressed above operational limits, already existing or newly born imperfections within the material may emit short bursts of energy called emission of events. Special receivers can detect these acoustic emissions. The source of emissions can be evaluated through the study of their features as strength, frequency, amplitude, and location.	Global monitoring; real-time; remote sensing; discriminate sources. eg. bridges.	Repeatability owing to each loading is difficult; suffering from attenuation; subjected to extraneous noise; contact method; multiple sensors required for flaw location; degradation must propagate during measurement.
Eddy Current Testing	Inducing electrical currents in the material being inspected and observing the interaction between those currents and the material. Eddy currents are generated by electromagnetic coils in the test probe and monitored simultaneously by measuring probe electrical impedance. The currents will be disrupted once a flow occurs.	Quick, sensitive; non-contact; no requirement for surface preparation; low-cost; adaptable to automation and in-situ examinations.	Inspecting conductors; variables must be understood and controlled; shallow-depth of penetration; lift-off effects and surface conditions; sensitive to cracks perpendicular to the interrogating surface
Liquid Penetrant Testing	An extension of visual inspection for surface-breaking flaws on non-absorbent material's surface. Basic procedure: surface clean → dye coat/penetrant → clean → developer → inspection → clean.	Portability; low-cost; sensitivity; versatility for nonporous materials; effective for production inspection.	Only detecting surface discontinuities; sensitive to temperature variation, surface condition and configuration; surface preparation required; messy; environmental and safety issues.

Magnetic Particles Inspection	Inducing a magnetic field in a ferro-magnetic material and dusting the surface with iron particles. Surface imperfections will allow the magnetic field to leak out of the material, distort the magnetic field and concentrate the iron particles near the imperfection area.	Virtually instantaneous; applied in-situ; easy to use and interpret; any size and shape of components; surface preparation is less critical; inexpensive; highly sensitive and fast.	Only effective for ferromagnetic material; limited to surface or near the surface; demagnetization may be required before/after inspection; uniformity may not be possible.
Microwave Testing	Measure properties of the electromagnetic waves scattered by or transmitted through a test particle; electromagnetic field change whenever the interact with a material structure. High frequency (from a few hundred MHz to hundred GHz; 1m-1mm) electromagnetic energy;	Non-contact; able to test moving parts or parts difficult to reach; able to test large areas; no electromagnetic interference produced due to low amount of operating power; internally inspecting dielectric objects.	Superficially inspecting conducting objects; expensive.
Optical Methods	Using light (electromagnetic energy) to discern certain material properties. Exploiting the interaction of light with matter Macroscopic/microscopic level	Non-contact; sensitive to micro-deformations; able to determine (visualize) the stress distribution.	May require a "vibration-free" environment; expensive; sophisticated.
Radiograp-hic Testing	Use of penetrating X-ray (λ 10-0.01 nm, f 30-30000PHz) to examine the interior of the object Show density changes in the film The only NDE works on all materials	Accurate; inherently pictorial, used for many shapes and sizes; sensitive assuming the discontinuity causes a reduction of cross-section thickness; widely used; volumetric method.	Safety hazard concerns; time-consuming; costly; dependent on discontinuity orientation; extensive experience and training of the personnel.
Thermography	Exploits the heat transfer within a given material, from a given material to air…; integrate infrared imaging with external heating to assess subsurface structure condition via the thermal response of the sample	Non-contact; fast; a thermal image is easy to interpret; good spatial resolution.	Only surface can be seen thermally; some materials are so thermally reflective requiring preparation with a high emissive coating; experienced thermographer may be necessary.

Ultrasound Testing	Use the transmission of high frequency sound wave (stress wave) into a material to detect imperfections within the material or changes in material properties. The perturbation travels through the object, affected by the object's geometry and properties; particle oscillation around an equilibrium point	Small discontinuities; varieties of techniques and approaches available; high-temp examination possible; thick, long or buried parts; surface and subsurface discontinuities; high-speed scanning possible; repeatability; equipment light and portable; safe; non-contact; fast; continuous monitoring.	Discontinuities oriented parallel to the beam energy may not be detected; discontinuities similar or smaller than the material's grain structure may not be detected; thin section may present resolution problems; may require the use of couplant.
Vibration Method	Exploits the global vibration of a given structure to determine or predict any anomaly. Excessive vibration of rotating machinery or shift of the natural resonance frequency can be symptoms of serious anomalies.	Global method; inexpensive; continuous monitoring.	Global method; need accurate knowledge of the structure; structural response affected by other parameters than damage only; damage sensitivity.
Visual Testing	First used; exploit the direct or remote eye-examination of a given object.	Afford the clearest view of the inspection surface; the direct view will result in the processing of the full spectrum of light wavelengths available to the eye and the brain to form an image; all visual attributes in existence are observed by the inspector.	Artificial enhancement of the image is not always possible; access restrictions may limit the view; environmental extremes hazardous to inspection personnel may limit the viewing time or visualization of the image altogether.

Appendix B: List of Relevant Publications

- "Towards Data-Driven Structural Health Monitoring: Application of Machine Learning and Signal Processing to Damage Detection." Ying, Y., Garrett Jr., J. H., Oppenheim, I. J., Soibelman, L., Harley, J. B., Shi, J., and, Jin, Y. (Accepted). *Journal of Computing in Civil Engineering*.
- "Damage Detection in Pipes under Changing Environmental Conditions Using Embedded Piezoelectric Transducers and Pattern Recognition Techniques." Ying, Y., Garrett Jr., J. H., Harley, J. B., Oppenheim, I. J., Shi, J., and, Soibelman, L. (2012, in press). *Journal of Pipeline Systems Engineering and Practice*.
- "Ultrasonic Monitoring of a Pipe under Operating Conditions." Liu, C., Harley, J. B., O'Donoughue, N., Ying, Y., Altschul, M. H., Garrett Jr., J. H., Moura, J. M. F., Oppenheim, I. J., and Soibelman, L. (2012). *Proceedings of 2012 SPIE Smart Structures/NDE Conference*, San Diego, CA.
- "Machine Learning for Pipeline Monitoring under Environmental and Operational Variations." Ying, Y., Garrett Jr., J. H., Harley, J. B., Moura, J. M. F., O'Donoughue, N., Oppenheim, I. J., Shi, J., and, Soibelman, L. (2011). *Proceedings of the 8th International Workshop on Structural Health Monitoring*, Stanford, CA.
- "Application of Mellin Transform Features for Robust Ultrasonic Guided Wave Structural Health Monitoring." Harley, J. B., Ying, Y., Garrett Jr., J. H., Soibelman, L., Oppenheim, I. J., Moura, J. M. F. (2011). *38th Annual Review of Progress in Quantitative Nondestructive Evaluation*, Burlington, VT.
- "Applications of Machine Learning in Pipeline Monitoring." Ying, Y., Harley, J. B., Garrett Jr., J. H., Jin, Y., Oppenheim, I. J., Shi, J., and, Soibelman, L. (2011). *Proceedings of 2011 ASCE International Workshop on Computing in Civil Engineering*, Miami, FL, 242-249.
- "Time Reversal for Damage Detection in Pipes." Ying, Y., Harley, J. B., Garrett Jr., J. H., Jin, Y., Moura, J. M. F., O'Donoughue, N., Oppenheim, I. J., and Soibelman, L. (2010). *Proceedings of 2010 SPIE Smart Structures/NDE Conference*, San Diego, CA, 76473S.1-12.
- "Cognitive Sensor Network for Structure Defect Monitoring and Classification Using Guided Wave Signals." Jin., Y., O'Donoughue, N., Moura, J. M. F., Harley, J. B., Garrett Jr., J. H., Oppenheim, I. J., Soibelman, L., Ying, Y., and Lin, H. (2010). *Proceedings of 2010 SPIE Smart Structures/NDE Conference*, San Diego, CA, 76473T.1-12.
- "Single Antenna Time Reversal of Guided Waves in Pipelines." O'Donoughue, N., Harley, J. B., Moura, J. M. F., Jin, Y., Oppenheim, I. J., Ying, Y., States, J., Garrett Jr., J. H., and Soibelman, L. (2009). *Proceedings of Meetings on Acoustics*, Portland, OR, 6(1), 065001-11.
- "Focusing of Ultrasonic Waves in Cylindrical Shells using Time Reversal." Harley, J. B., O'Donoughue, N., States, J., Ying, Y., Garrett Jr., J. H., Jin, Y., Moura, J. M. F., Oppenheim, I. J., and Soibelman, L. (2009). *Proceedings of the 7th International Workshop on Structural Health Monitoring*, Stanford, CA.
- "Preliminary Studies on the Dispersion of Signals Produced by Permanently Installed MFC Transducers for Pipeline Monitoring." Agrawal, A., Harley, J. B., Ying, Y., Garrett Jr., J. H., Sohn, H., and Soibelman, L. (2009). *Proceedings of the 16th Workshop of the European Group for Intelligent Computing in Engineering*, Berlin, Germany.