8-20-2015

Wavelet-based Feature Extraction Methodology for Pattern Classification in Engineering Applications

Andre A. Silva
University of Connecticut - Storrs, aasilva80@gmail.com

Wavelet-based Feature Extraction Methodology for Pattern Classification in Engineering Applications

Andre A. Silva

B.S., University of Connecticut, 2012

A Thesis

Submitted in Partial Fullfilment of the

Requirements for the Degree of

Master of Science

at the

University of Connecticut

2015

APPROVAL PAGE

Master of Science Thesis

Wavelet-based Feature Extraction Methodology for Pattern Classification in Engineering Applications

Presented by

Andre A. Silva, B.S.

Major Advisor _____
Dr. Shalabh Gupta

Associate Advisor _____
Dr. Ali M. Bazzi

Associate Advisor _____
Dr. Krishna R. Pattipati

University of Connecticut

2015

To my mother Seida Davila,

my father Ricardo Silva,

and my brother Carlos Silva.

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation and gratitude to my advisor Dr. Shalabh Gupta for his invaluable recommendations in my research work and his constant amount of encouragement throughout my pursuit of a deeper level of knowledge. I am humbled to have worked with a very considerate professor that truly values the spirit of learning in research and further endeavors.

I would also like to express my sincere thanks to my committee members, Dr. Ali M. Bazzi and Dr. Krishna R. Pattipati, for investing their time and feedback as part of my research efforts.

I am grateful to my colleagues at the University of Connecticut for their gregariousness and support throughout my time in graduate school.

Finally, I would like to thank my parents for their love and encouragement. Thank you for telling me I can achieve anything I set my mind to.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

Wavelet-based Feature Extraction Methodology for Pattern Classification in Engineering Applications

Andre A. Silva, M.S.

University of Connecticut, 2015

Feature extraction is fundamental in the framework of pattern recognition. In classification applications where raw data is collected from scientific experiments or simulations, the resulting data may have been exposed to noise, excited by external dynamics, or affected by uncertainties not accounted for during the data acquisition process. Furthermore the data can be high dimensional, which can hinder the accuracy of decision algorithms if additional dimensions of the data make the patterns non-separable, not to mention the added computational complexity burden to train such algorithms. The objective of this research is to understand how specially designed features can improve the class separability in Euclidean space, and how adaptive the feature extraction approach can be when implemented in various pattern classification applications.

The wavelet analysis, with known benefits for simultaneous time and frequency localization, has seen recent application to feature extraction in pattern recognition problems. However, there is still a gap in understanding if the entire time-scale domain of wavelets is necessary for pattern classification. Despite the benefits of the additional dimension present in the wavelet domain, there is an added computational complexity for machine learning. Furthermore, information in the

wavelet domain that is not useful for separation of classes can in fact degrade the performance of a classifier. This thesis presents a novel filtering approach that can utilize the benefits of wavelets as well as select the most useful regions in the wavelet domain to extract features for improving the classifier performance and thereby reducing the computational complexity.

This thesis presents a wavelet-based filtering method for feature extraction. The method extracts optimal regions in the wavelet domain where the between-class separation is improved while the within-class separation is reduced. The approach is verified on four different applications, ranging from electric machine fault diagnosis to medical health detection. The experimental results demonstrate significant improvements in classification accuracies on these applications when using the optimal regions in the wavelets. The approach is also compared with existing methods used on the same applications.

NOMENCLATURE

List of Abbreviations

ECG	Electrocardiogram
PD	Parkinson's Disease
FOG	Freezing of Gait
ADL	Activities of Daily Living
KLT	Karhunen-Loève Transform
FDD	Fault Detection & Diagnosis
EDM	Electrical Discharge Machining
RUL	Remaining Useful Life
MCSA	Motor Current Signature Analysis
FFT	Fast Fourier Transform
PCA	Principal Component Analysis
FDA	Fisher Discriminant Analysis
BIH	Beth Israel Hospital
SVDB	Supraventricular Arrhythmia Database
CWRU	Case Western Reserve University
OVA	One vs. All
RAS	Rhythmic Auditory Stimulation
pmf	probability mass function

PC	Principal Component
FOC	Field-Oriented Control
CWT	Continuous Wavelet Transform
k-NN	k-Nearest Neighbor
CART	Classification and Regression Tree
RF	Random Forest
HMM	Hidden Markov Model
ANN	Artificial Neural Network
SVM	Support Vector Machine
CCR	Correct Classification Rate
SPC	Specificity
SNS	Sensitivity
FAR	False Alarm Rate
MDR	Missed-Detection Rate
GUI	Graphical User Interface

NOMENCLATURE

List of Symbols

$f(t)$	example time-series function
$\psi(t)$	mother wavelet function
S_i	i^{th} scale subset
T_j	j^{th} time shift subset
$\boldsymbol{C}^0$	set of class labels
$\boldsymbol{V}_{C_\alpha}$	set-theoretic difference of $\boldsymbol{C}^0$ and $C_\alpha \in \boldsymbol{C}^0$
$W(\cdot, \cdot)$	wavelet coefficients
$W_{i,j}$	wavelet coefficients for the i^{th} scale subset and j^{th} time shift subset
$R(\gamma)$	wavelet cells for the γ^{th} index
$R_x^\eta(\gamma)$	wavelet cells for the x^{th} class index and η^{th} observation index
$D(\gamma)$	objective function for the γ^{th} index
Θ^*	optimal cell indices
X	matrix of optimal cells
$\|\cdot\|_p$	$\mathcal{L}^p$-norm/ℓ^p-norm of a vector
$\|\cdot\|$	cardinality of a set
$\mathbb{R}$	set of real numbers
$\mathbb{N}^+$	set of natural numbers, except 0

Chapter 1

Introduction

1.1 Research Background & Motivation

The problem of finding patterns in time-series data is fundamental in scope, with origins in statistics and engineering. Recently, the field of pattern recognition has advanced rapidly, primarily due to its conceptual backbone and interdisciplinary applications in Statistical Inference, Bioinformatics, Image Recognition, Optimization, Decision Theory, and other related fields. For supervised learning, typically there are two phases that are necessary for classification. In the *training phase*, the data of interest has multiple instances of the respective patterns where information about the pattern labels are known *a priori*. This data is then utilized to train a mathematical model that can correctly assign these labels to the data with minimum error. This model, classed the *classifier*, should attempt to generalize well for different dynamics and operating conditions respective to the domain of the classification problem. In the *testing phase*, where new data arrives with unknown class labels, the goal is to correctly classify the patterns online using the trained classifier.

Feature extraction is fundamental to the framework of pattern recognition. In classification applications, where raw data is collected from scientific experiments or simulations, the resulting

data may have been exposed to noise, excited by external dynamics, or affected by uncertainties not accounted for during the data acquisition process. Furthermore, the data can be high dimensional, which can hinder the accuracy of decision algorithms if additional dimensions of the data make the patterns non-separable, not to mention the added computational complexity burden to train such algorithms. As with most designs in pattern recognition, the demand for near-perfect classification accuracy, while minimizing the false alarms and missed detections, are some of the most critical requirements in the synthesis of satisfactory solutions. In critical applications such as in defense, where an unknown target is to be classified as an enemy or a friendly, there are unavoidable trade-offs in accuracy to consider when minimizing the error in classifying friendly targets as the enemy vs. classifying enemy targets as the friendly. Other criteria include the time allowed to make the classification of the unknown pattern, as well as the amount of financial costs allowed to implement the methodology in hardware. Therefore, it is essential for feature extraction methods to enhance the separability of patterns in the raw data to improve the accuracy of classification and to reduce the computational complexities to train and implement the classifiers.

1.1.1 Wavelets for Feature Extraction

As mentioned above, Machine Learning approaches typically employ several pre-processing steps to aid in improving the classifier's ability to discriminate among the patterns in the data. As an example, these steps may include signal conditioning for denoising, input normalization such that the input values exist within the realm of the capabilities of the selected algorithms, and feature extraction for enhancement of separability of the classes. Standard techniques for feature extraction include Principal Component Analysis [1] (PCA), Fisher Discriminant Analysis [2] (FDA), Local Linear Embedding [3], and supplementary extensions present in the literature. Li et al. [4] presented

a fuzzy-based feature extraction method utilizing FDA. Huang et al. [5] utilized Slow Feature Analysis [6] to propose the Slow Feature Discriminant Analysis (SFDA), for handwritten digit recognition, which uses an objective based on temporal variations. Mo et al. [7] utilized the Kernel PCA for feature extraction in an indoor positioning environment. Jiang et al. [8] implemented the Positive Matrix Factorization (PMF) to extract the features from drugs and explosives concealed by body packing. Zabalza et al. [9] proposed the Folded-PCA approach that can improve the feature extraction efficiency and classification accuracy. Li et al. [10] further presented a semi-supervised feature extraction based on fuzzy FDA for hyperspectral image classification. Gu et al. [11] extended SFDA using uncorrelated features and global preserving projections.

The wavelet analysis, with known benefits for simultaneous time and frequency localization, has seen recent application to feature extraction in pattern recognition problems. Godfrey et al. [12] applied the continuous wavelet transform in a classification tree architecture for delirium subtyping using patient accelerometer information. Yu et al. [13] implemented their proposed cluster-based wavelet feature extraction for machine fault diagnosis. Eristi et al. [14] utilized wavelet analysis and sequential forward selection for feature extraction and selection that is used to classify power system disturbances. Avci et al. [15] designed a feature extractor utilizing wavelet energy and entropy that are used to train a neural network optimized by a genetic algorithm. Saravanan et al. [16] utilized discrete wavelets for feature extraction and selection for diagnosis of incipient gearbox faults. Local energy histograms in the wavelet domain are proposed by Dong et al. [17] for the classification of Brodatz texture images. Shankar et al. [18] applied discrete wavelets as features to a fuzzy classifier for classification of remote sensing images. Liu et al. [19] proposed the sparse coding model as a feature extraction method from wavelets for bearing fault diagnosis. Guo et al. [20] implemented a

genetic programming based feature extraction system utilizing the wavelet coefficients for epileptic classification. Seshadrinath et al. [21] presented a classification methodology for motor current signatures by applying the wavelet analysis to train a neural network based classification algorithm. Bafroui et al. [22] utilized the energy and entropy calculated from the wavelet transform to design a feature extraction method for classification of gearbox failures.

Other works in the literature explore extensions or alternates of wavelet analysis. Xian et al. [23] utilized wavelet packets and proposed a hybrid Support Vector Machines (SVM) classifier for fault diagnosis of rotating machinery failures. Lin et al. [24] improved the Empirical Mode Decomposition (EMD) method for feature extraction of vibration signals and electrical signals in the brain. Huang et al. [25] utilized their proposed non-separable wavelets with Gaussian Markov random fields for iris feature extraction. Khushaba et al. [26] proposed a feature extraction method based on fuzzy sets and mutual information for classification of driver drowsiness. This method finds the optimal wavelet decomposition tree from each data channel and combines these features together to form one net feature vector. Wang et al. [27] applied the wavelet packets in conjunction with the best basis selection for feature extraction implemented to detect epileptic conditions. Bin et al. [28] applied the EMD method and wavelet packets as the bases for a feature extractor of rotating machinery. Li et al. [29] utilized the EMD to calculate the coefficients of variation and fluctuation indices that were used to detect seizures with SVM. Deepa et al. [30] implemented the combination of the Discrete Fourier and Cosine Transform for feature extraction and Particle Swarm Optimization (PSO) for feature selection for face recognition. Chen et al. [31] proposed the overcomplete rational dilation discrete wavelet transform for feature extraction of gearbox failures. Han et al. [32] utilized wavelet packets combined with an improved version of the Fast-ICA algorithm

with the third-order Newton iteration adopted. Wang et al. [33] applied Ensemble EMD (EEMD) combined with a tunable Q-factor wavelet transform for the feature extraction of bearing faults. Poornima et al. [34] proposed a feature extraction technique based on Multiscale PCA for face recognition. Lu et al. [35] designed a feature extractor using adaptive multiwavelets for rotor fault diagnosis. Bennet et al. [36] utilized the discrete wavelets with a moving window technique for feature extraction and selection applied to cancer classification. Wang et al. [37] applied the lifting wavelet packets with the fuzzy c-means method for classification of load dynamics characteristics. Zhang et al. [38] designed a combined wavelet filter that is tuned with a quantum-based PSO for demodulation of bearings information. Yuwono et al. [39] proposed a frequency extractor using the wavelet filter banks and the wavelet kurtogram in conjunction with a comb filter to generate a feature vector of defect frequencies. These feature vectors are utilized with a swarm-based method to train the hidden states of a Hidden Markov Model (HMM) used for classification.

The above methods are useful in utilizing the 2-D wavelet domain to extract features for pattern classification. However, there is still a gap in understanding if the entire time-scale domain of wavelets is necessary for pattern classification. Despite the benefits of the additional dimension present in the wavelet domain, there exists added computational complexity for machine learning. Furthermore, information in the wavelet domain that is not useful for separation of fault classes can in fact degrade the performance of a classifier. Thus, a novel filtering approach is needed that can utilize the benefits of wavelets as well as select the most useful regions in the wavelet domain to extract features for improving the classifier performance and reducing the complexity.

1.1.2 Applications in Fault Detection & Diagnosis

One of the applications considered in this thesis involves Fault Detection and Diagnosis (FDD) of different faults within electric machines and drives systems. The induction motor, and in general electric machine and drive systems, are the de facto standard in the industry due to their consistency of speed control, cost effectiveness, and range of applications including: electric vehicles such as ships, boats, cars, underwater vessels and other applications, such as air handling systems, extruders, hoists and conveyors. Several non-linearities of the induction motor make the drive structure more complex such as the interdependence of the main flux and armature current, variable speed power inverters, and AC signal data acquisition devices. Furthermore, as with most electrical systems, the electric machines are prone to component failures. Fault detection and diagnosis (FDD) of the motor drive systems are essential for early warning of incipient failures that may help in system recovery and significant improvement of their remaining useful life (RUL).

The analysis of these systems during the presence of faults and methods for fault diagnosis have been the subject of interest in the power electronics community since the later part of the 20^{th} century [40–46]. At a course level, there are three main critical failure modes exhibited by induction motors: bearing failure modes, stator winding failure modes, and rotor bar faults. The bearing failure affects the motor wear by increasing the rotational friction of the rotor. Undesirable wear is the main cause of bearing faults, which may include: normal fatigue wear after prolonged service past the remaining useful life; overheating, due to constant operation in extreme conditions or excessive loading on the bearings; and misalignment, where a raceway ball track is not parallel to the raceway edges. The stator winding failures are the second most common type of failure in these systems, where they account for 21% of the distribution of faults

in electric motor drive systems [47]. The main reason these faults occur is due to the unforeseen breakdown of insulation between components, which may occur across one or more of the phase windings in the stator or across the phase and nearby components. In addition, rotor bar failures have received notable attention. Broken rotor bars are generally caused by stresses from electromagnetic forces or overloaded operating conditions, inadequate rotor fabrication, and rotor component wear from poor operating environments or lack of maintenance [48]. Therefore, the application of FDD is essential for early warning of these failures to aid in the recovery of motor drive systems and add significant improvement to their remaining useful life (RUL).

Given that fault diagnosis of electric machine systems is salient to the industry, there exist methodologies implemented for diagnosis of these failures. From the prior studies in electric machine diagnosis, the location of the frequency bands where the information of the fault is present have been identified. Motor Current Signature Analysis (MCSA) is a method that applies Fourier analysis to the motor transient parameters [49–51] and seeks to discriminate the faults based on their unique spectra. The main advantage of MCSA is that the practicality of implementation, since Fast Fourier Transforms (FFTs) are very efficient to generate. To further enhance the practicality, the information of the frequency locations and the automation of the diagnosis technique led to widespread acceptance in the industry. The disadvantage of MCSA is that it is highly dependent on the localization of the fault's respective frequency interval. In other words, effects of outside uncertainties can affect unique frequency representation of the faults. Fault frequency values may overlap and so it is no longer trivial approach to discriminate the faults. Measurement noise and disturbances also play a role in this effect. Furthermore, for motor transient behavior that is not always periodic, the method lacks the information to gather what frequencies occur at what point

of time. Several methods have been reported in literature for fault diagnosis of electric motor drives that deal with these issues. Schoen et al. [52] implemented an Artificial Neural Network (ANN) classifier for unsupervised, online learning of induction motor failures. Tallam et al. [53] extended the application of stator winding turn-fault detection for closed-loop induction motor drives based on ANNs. Murphey et al. [54] proposed the fault diagnostic ANN for single-switch and post-short circuit faults. Martins et al. [55] proposed a Hebbian-based ANN for unsupervised, online diagnosis of the stator faults utilizing vector current information. Liu et al. [56] proposed multi-agent systems that can automate the various tasks required for real-time health monitoring. Tran et al. [57] proposed a feature selection of current sensors based on decision trees to be implemented in a neuro-fuzzy inference system. Ghate et al. [58] proposed an optimal multi-layer perceptron ANN and later explored cascaded ANN systems for induction motor fault detection [59].

Kallesoe et al. [60] proposed an observer-based estimation of interturn short circuit faults in delta-connected induction motors. Tabbache et al. [61] implemented the Extended Kalman Filter (EKF) for residual generation of the motor parameters for sensor fault detection and post fault-tolerance. De Angelo et al. [62] generated vector residuals for stator-interturn short-circuit detection. Cheng et al. [63] proposed a fault detection and identification approach of stator-turn faults using the transfer impedance of closed-loop multiple-motor drives. Georgoulas et al. [64] applied the Principal Component Analysis (PCA) with Hidden Markov Models (HMM) for broken rotor fault diagnosis in asynchronous machines.

FDD methods utilizing the wavelet analysis [65], which has been cited as an ideal diagnostic method for all-electric ship systems [66], has seen significant presence in this community for time-scale feature extraction [67]. Mohammed et al. [68] implemented the wavelets for faults diagnosis

of permanent magnet machines using a recently validated machine model based on Finite Element Analysis (FEA). Ordaz-Moreno et al. [69] designed a broken bar detection algorithm based on discrete wavelets for FPGA implementation. Siddiqui et al. [70] implemented discrete wavelets for detection of broken rotor bar faults. Cusido et al. [71] utilized power spectral density techniques in wavelet decomposition for machine fault detection. Li et al. [72] applied wavelet-based kurtosis statistics for fault diagnosis in rolling bearings. Rajagopalan et al. [73] implemented the Zhao-Atlas-Marks distribution for nonstationary motor fault detection. Rosero et al. [74] utilized the Empirical Mode Decomposition (EMD) and Wigner-Ville distribution for short-circuit detection of permanent magnet machines. Sadeghian et al. [75] proposed an algorithm for broken rotor bar detection using wavelet packets and neural networks. Konar et al. [76] utilized wavelet analysis with Support Vector Machines (SVM) for bearing fault detection. They later implemented a Genetic Algorithm (GA) as a feature selection of wavelet information for induction motor fault diagnosis [76]. Bin et al. [28] applied the EMD method and wavelet packets as the bases for feature extraction in rotating machinery. Keskes et al. [77] implemented the Stationary Wavelet Packet Transform (SWPT) to diagnose broken rotor bar failures. Seshadrinath et al. [78, 79] proposed Dual Tree complex wavelets for interturn fault diagnosis. They also presented a classification methodology by applying the wavelet analysis for optimized Bayesian inference [21].

The first application of the thesis includes the fault diagnosis of rotor bar faults and electric phase failures for motor drives used in electric ships [80–82], and the second focuses on bearing fault diagnosis in induction motor systems [83].

1.1.3 Applications in the Medical Field

The second application the thesis explores includes detection of medical conditions of patients from symptoms. There are two particular problems that are explored in this field. Cardiac arrhythmia is a medical condition in which the heartbeat of a person is either faster or slower than a typical heartbeat. A particular type of arrhythmia known as Atrial Fibrillation affects nearly 2% of the population in the western countries as of 2014 [84]. Although some episodes of arrhythmia posed harmless effects, there are other instances where the symptoms can include the higher risk for blood clotting within the heart and higher risk of blood flow reduction delivered to the heart. A number of tests for the diagnosis of this heart condition include electrocardiogram (ECG) devices that measure the electrical activity through the heart. Expert analysis from cardiologists allows one to discriminate the periods of the nominal heart activity against the arrhythmia condition. However, due to the common case that there are more patients seeking care than doctors available, there is a need for automated heart condition diagnostic tools for timely medical care.

The behaviors of arrhythmia in the signal have been extensively studied and several approaches have been proposed for its detection. A standard approach known as ARISTOTLE [85, 86] utilizes the Karhunen-Loève Transform (KLT) to project the signal's periodic behavior to one out of the five expected signal beat patterns. Greenwald et al. [87] extended this technique by adding a post processor that learns from the inherent noise present in the ECG signals to improve the detection capabilities. Other techniques proposed additional feature parameters from the ECG beat sequence for the detection [88, 89]. Noise is a predominant issue with the hardware used to generate ECG signals, which has been a motivation for the recent

techniques using data-driven approaches. If the system has a complex structure due to various interconnections or inherent redundancy, then the data-driven approach can enable learning of the underlying behavioral patterns that stem from a system subjected to different faulty conditions. Furthermore, if there exists incomplete information of the physical phenomena, or if the system operates in the presence of uncertainties, the data-driven approaches can be useful in extracting the most relevant features of faulty behaviors. The recent literature invokes several machine learning approaches in conjunction with frequency analysis techniques [90–94], with most of the work utilizing time-frequency methods. The author in [95] utilized the data from the MIT-BIH Supraventricular Arrhythmia Database (SVDB) to propose a machine learning approach to diagnose the arrhythmia condition in patients. This data is the third application of the thesis.

Parkinson's Disease (PD) is a degenerative disorder that occurs in the central nervous system. It affects the creation of dopamine-generating cells in the region of the mid-brain, which result in the impairment of a person's motor skills. The cause of the cell destruction is currently unknown. The periods where the motor skills of a patient are impaired are commonly known as freezing conditions. Typical freezing occurs in a patient's legs, and this particular symptom is known as a Freezing of Gait (FOG). The FOG symptom causes discomfort for the patient, as it interferes with Activities of Daily Living (ADL) and ultimately impairs their quality of life [96, 97]. While current medicinal solutions cope with the symptom by reduction of the FOG episode, these results are only effective within a brief time interval. Furthermore, as the disease progresses, medicinal solutions are not as effective as in the disease's early stages.

In the initial phases of the research, the literature highlighted studies to better quantify the gait behavior [98, 99]. In [100, 101], Salarian et al. investigated the assessment of various gait

parameters in the development of PD symptom monitoring systems. Machine learning approaches became more prevalent once there was a way utilize the prior work to aid in improving patient ADL. In [102–104] the investigators found that gait patterns are dependent on auditory signals that affect patient's movement behavior, the subject of Rhythmic Auditory Stimulation (RAS). With this idea, the solution was to diagnose the FOG symptom and utilize RAS to correct the walking behavior in real-time. The authors in [105] sought to utilize the sensors placed at different body parts of the patients to identify the anomaly. They in fact proposed a thresholding-based technique using Fourier analysis to diagnose the symptom. The authors in [106] implemented an approach within a machine learning framework on this same problem. Classifiers are trained by the statistical parameters highlighting the gait behaviors. The most accurate classifier is implemented online with feedback auditory response that cues an improvement in patient walking. The Daphnet FOG [105] is the fourth application of the thesis.

In summary, there is a vast set of machine learning applications where wavelet analysis for feature extraction can be used. Furthermore the successes found in utilizing wavelet analysis for feature extraction in the aforementioned applications leave further questioning whether features can be further enhanced to improve the classification performance. The thesis explores the issues whether the entire time-scale domain of wavelets is necessary for pattern classification. Despite the additional dimension present in the wavelet domain to enhance class information, there exists added computational complexity for training pattern classifiers. Furthermore, information in the wavelet domain that is not useful for the separation of fault classes can in fact degrade the performance of a classifier. Thus, a novel filtering approach is needed that can utilize the benefits of wavelets as well as select the most useful regions in the wavelet domain to extract features for improving the

classifier performance and reducing the complexity.

1.2 Contributions of the Thesis

This thesis presents a wavelet-based filtering method for feature extraction. The contribution extracts optimal regions in the wavelet domain where the between-class separation is improved while the within-class separation is reduced. The approach is verified on the above four applications, ranging from electric machine fault diagnosis to medical health detection. The results demonstrate that significant improvements in classification accuracies are achievable while reducing the computational complexity. The approach is also compared with existing methods used on the same applications. The contributions of the thesis are as follows:

- A wavelet-based feature extraction methodology for pattern recognition is developed, with applications to the electric motor drives field and the medical field.
  - A partitioning of the wavelet domain to define the *cells*, specific subsets of the wavelet domain localized by time and scale.
  - Objective costs that measure the between-class and within-class distance of the cells, and selection of these optimal cells for feature extraction.
  - Verification of the approach in terms of increased classification accuracy and reduced computational complexity through its application to four different cases studies:
    1. Motor drives in electric ships with broken rotor faults and electrical phase failures,
    2. Bearing faults in induction motor systems,
    3. Cardiac arrhythmia in medical patients, and
    4. Freezing of gait symptoms in medical patients.

1.3 Structure of the Thesis

The thesis is organized into six chapters, including the present chapter. Chapter 2 describes the theoretical development for the proposed feature extraction technique. Wavelet analysis with a partitioning scheme is presented to construct specific regions in the wavelet domain, hereafter defined as *cells*. The optimal cells in the data are identified through the proposed objection functions where class information is better separated. Furthermore, the tools implemented for pre-processing and pattern classification are highlighted.

Chapter 3 describes the application to an experimentally validated electric drive system used in electric ships. The task is to perform fault diagnosis of the motor drive failures.

Chapter 4 describes the application of the approach to the Case Western Reserve University (CWRU) bearings data collected from an induction motor system. The task is to perform fault diagnosis of the bearing failures.

Chapter 5 describes the application of the approach to the Supraventricular Arrhythmia (SVDB) database. The task is to diagnose the arrhythmia heart condition in patients.

Chapter 6 describes the application of the approach to the Daphnet Freezing of Gait (FOG) data. The task is to diagnose the FOG symptom in patients suffering from Parkinson's Disease.

In each of the application chapters, the application is briefly introduced and the reasons for applying machine learning are stated. Discussion will follow in regards to the experimental setup for data generation. A presentation of the data visualization and how the method is realized in the wavelet domain are shown. Finally, the results of the classification as well as comparative results from other approaches are presented.

The thesis is concluded in Chapter 7, which also outlines suggestions for future research in

the related fields of study. Specifically, future research on improving the separation of classes via feature extraction techniques is discussed.

1.4 List of Publications

Journal Papers

1. **A.A. Silva**, S. Gupta, A.M. Bazzi and A. Ulatowski, "Wavelet-based Feature Extraction for Fault Diagnosis of Motor Drives in Electric Ships," *to be submitted to* IEEE Transactions on Industry Applications.

2. **A.A. Silva**, N. Najjar and S. Gupta, "Wavelet-based Feature Extraction Methodology with Pattern Recognition Applications," *In Preparation.*

Conference Papers

1. **A.A. Silva**, N. Najjar, S. Gupta, P. DOrlando and R. Walthall, "Wavelet-based Fouling Diagnosis of Heat Exchanger in the Aircraft Environmental Control System," *to appear in* SAE 2015 AeroTech Congress & Exhibition, Sep. 2015.

2. N. Najjar, J. Hare, **A.A. Silva**, S. Gupta, G. Leaper, K. Pattipati, R. Walthall and P. D'Orlando, "Heat Exchanger Fouling Diagnosis for an Aircraft Air-Conditioning System," in SAE 2013 AeroTech Congress & Exhibition, Sep. 2013, pp. 3055-3060.

3. **A.A. Silva**, A.M. Bazzi and S. Gupta, "Fault Diagnosis in Electric Drives using Machine Learning Approaches," in IEEE International Electric Machines and Drives Conference (IEMDC), May 2013, pp. 722-726.

Patents

1. **A.A. Silva**, N. Najjar, S. Gupta, P. D'Orlando and R. Walthall, "Wavelet-based Analysis for Fouling Diagnosis of an Aircraft Heat Exchanger," U.S. Patent Application 14,689,467. *Submitted on Apr. 2015.*

Chapter 2

Feature Extraction Methodology

The wavelet analysis as a tool for extraction of time-scale information is presented in this chapter. Wavelet analysis is commonly used to capture the simultaneous time and frequency characteristics of non-stationary time-series data. While the technical literature has found application of wavelets in feature extraction, the wavelet information itself has not been analyzed to determine what information is critical to construct representative features of the classes, and what information hinders the goal of class separability, and thus can be filtered out. Furthermore, selection of the mother wavelet function affects the signal decomposition of the time-series and ultimately the kinds of features extracted.

This chapter presents the contribution of the wavelet-based filtering approach as developed in this thesis. Two particular measures have been utilized for filtering that calculate the between-class and within-class distance of the regions in the wavelet and select the optimal regions for feature extraction. The classification design is then highlighted, which is divided into training and testing phases. In the training phase, the data from each application is collected for the various classes considered and their respective number of observations. They are passed through the wavelet transform for enhancing the separability of the classes. The proposed technique selects the optimal regions in the wavelet domain to use for training a pattern classifier. In the testing phase, the

trained classifier is applied to make the decisions based on data with a priori unknown class.

2.1 Wavelet Analysis

For a time-domain signal $f(t)$ in any $\mathcal{L}^2(\mathbb{R})$ space, the signal can be expanded by the use of a family of orthonormal wavelet functions, such that

$$[W_\psi f](s,\tau) = \frac{1}{\sqrt{|s|}} \int_{\mathbb{R}} f(t)\psi^* \left(\frac{t-\tau}{s}\right) dt \tag{2.1}$$

where $\psi(t)$ is the mother wavelet, $s = 1, ..., m$ and $\tau = 1, ..., n$ are the scale and translation parameters, respectively; $[W_\psi f](s,\tau)$ is the wavelet transformation of the signal $f(t)$ [107]. Wavelet analysis is an effective tool that extracts two-dimensional information from time-domain signals. The wavelet coefficients form the $m \times n$ matrix from the raw data. The magnitude-square is computed from the wavelet data for mathematical convenience.

2.1.1 Wavelet Partitioning

The wavelet matrix is partitioned into a series of sub-matrices called *cells*, as described here. Let $a \in \mathbb{N}^+$ and $b \in \mathbb{N}^+$ be the divisions of scale and translation axis of the wavelet matrix such that $m \bmod a = 0$ and $n \bmod b = 0$, respectively. Now let

$$\begin{aligned} S_i &= \left\{(i-1)\frac{m}{a} + 1, ..., i\frac{m}{a}\right\} \;\; ; \;\; \forall i = 1, ..., a \\ T_j &= \left\{(j-1)\frac{n}{b} + 1, ..., j\frac{n}{b}\right\} \;\; ; \;\; \forall j = 1, ..., b \end{aligned} \tag{2.2}$$

and let

$$W_{i,j} = [W_\psi f](S_i \times T_j) \;\; ; \;\; \forall i, j \tag{2.3}$$

Further, let an index be defined as $\gamma = (i-1)b + j; \quad \forall i = 1, ..., a$, and $j = 1, ..., b$, such that $\gamma \in \Gamma = \{1, ..., ab\}$. Then, the data matrix $R(\gamma)$ corresponding to cell with index γ is defined as

$$R(\gamma) = W_{i,j}^T \tag{2.4}$$

An illustration of the partitioning and restructuring process is found in Fig. 2.1.

2.2 Optimal Cell Selection

Once the wavelet matrix is partitioned into cells, a set of optimal cells is selected for the purpose of enhancing the classifier decisions for pattern classification. The optimal cell selection approach utilizes a metric that measures the separability between fault classes. In general, let $\boldsymbol{C}^0 = \{C_1, ..., C_N\}$ represent the set of classes. Then, for any $C_\alpha \in \boldsymbol{C}^0$, $\alpha = 1, ...N$, let us define a set $\boldsymbol{V}_{C_\alpha} = \{\boldsymbol{C}^0 \backslash C_\alpha\}, \quad |\boldsymbol{V}_{C_\alpha}| = N - 1$, which represents the classes in $\boldsymbol{C}^0$ other than the class C_α. Let $V_\beta \in \boldsymbol{V}_{C_\alpha}$ for $\beta = 1, ..., N-1$. Then, for each cell $\gamma \in \Gamma$, the separability between classes C_α

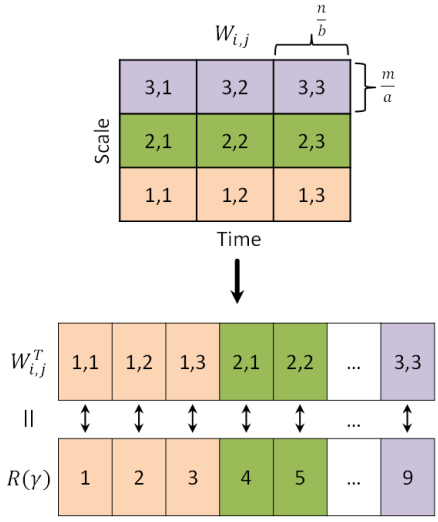

Fig. 2.1: Cell Partitioning and Restructuring

and V_β is measured by

$$\Phi_{\alpha\beta}(\gamma) = \sum_{\eta=1}^{M_\alpha} \sum_{\hat{\eta}=1}^{M_\beta} ||R^{\eta}_{C_\alpha}(\gamma) - R^{\hat{\eta}}_{V_\beta}(\gamma)||_2 \tag{2.5}$$

where η and $\hat{\eta}$ are the respective observation indices for classes C_α and V_β. M_α is the total number of observations for class C_α, and M_β is the total number of observations for class V_β.

Subsequently, the second measure seeks to find the within-class distance via computing the distance between all pairwise combinations of observations pertaining to the same class.

$$\Psi_\alpha(\gamma) = \sum_{\substack{\eta=1 \\ \eta\neq\tilde{\eta}}}^{M_\alpha} \sum_{\tilde{\eta}=1}^{M_\alpha} ||R^{\eta}_{C_\alpha}(\gamma) - R^{\tilde{\eta}}_{C_\alpha}(\gamma)||_2 \tag{2.6}$$

The two measures above are combined into a single measure as follows:

$$D(\gamma) = \frac{1}{g_1} \sum_{\alpha=1}^{N} \sum_{\substack{\beta=1 \\ V_\beta \in \boldsymbol{V}_{C_\alpha}}}^{N-1} \Phi_{\alpha\beta}(\gamma) - \frac{1}{g_2} \sum_{\alpha=1}^{N} \Psi_\alpha(\gamma) \tag{2.7}$$

An illustration of the concept is found in Fig. 2.2. The parameter g_1 counts the number of operations utilized to compute the first term, while g_2 follows for the second term. These parameters are thus used as normalization constants such that each term holds an equal weight in the objective function. From the combinatorics of the operations it is found that $g_1 = \sum_{\alpha=1}^{N} \sum_{\beta=1}^{N-1} M_\alpha M_\beta$ and

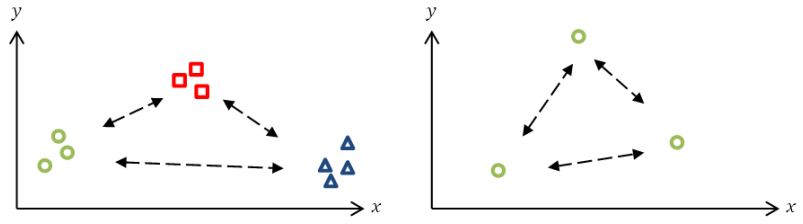

Fig. 2.2: Illustration of the Between-Class (left) and Within-Class (right) Measures

$g_2 = \sum_{\alpha=1}^{N} M_\alpha(M_\alpha - 1)$. Now, let's define indices $\theta_i, \forall i = 1, ..., ab$ where $\theta_i \in \Gamma$, such that

$$D(\theta_1) \geq D(\theta_2) \geq \cdots \geq D(\theta_{ab}) \tag{2.8}$$

Then, the optimal set of $\nu > 0$ cell locations that can enhance the class separation is defined by $\Theta^* \subseteq \Gamma$, where

$$\boxed{\Theta^* = \{\theta_1, ..., \theta_\nu\}} \tag{2.9}$$

2.3 Data Reduction and Feature Extraction

The Principal Component Analysis (PCA) is a data reduction method which transforms the data into a lower dimensional feature space [108, 109]. To apply PCA, the winning cells of each class are placed into a matrix X, where $X = [R(\theta), \forall \theta \in \Theta^*]$, such that the rows represent the length of a cell and the columns represent the number of optimal cells $\nu = |\Theta^*|$. The goal of PCA is to find a reduced feature set Y containing $q < \nu$ principal components that summarize the most useful information in X.

Using the Karhunen-Loève (KL) algorithm, the uncorrelated features, called Principal Components (PCs) or score vectors, are inferred from the data matrix. PCA is an optimal feature extraction method as the PC's computed are based on the maximization of the data matrix variance. These PC's capture the most useful information in the classes from the original data matrix, as can be seen by the formation of separable clusters in the feature space. The KL algorithm is described as follows:

1. The covariance matrix Σ of size $(\nu \times \nu)$ of X is computed and the left eigenvalues $\{\lambda_i\}$ and the corresponding eigenvectors $\{\boldsymbol{e}_i\}$ are obtained for $i = 1, 2, ..., \nu$.

2. The eigenvalues are sorted; then, the $q < \nu$ largest dominating eigenvalues are selected.

3. Using the q eigenvectors that correspond to the largest eigenvalues, a $(\nu \times q)$ transformation matrix $\mathcal{T}$ is constructed. Consequently, the original data set X is transformed into the feature set Y, using Eq. 2.10 below.

$$Y = X \times \mathcal{T} \tag{2.10}$$

2.4 Pattern Classification

In general, pattern classification algorithms can learn from the data in the form of parametric methods or non-parametric methods. For parametric designs, a decision boundary is constructed around the different class regions. Non-parametric methods seek to associate posterior probabilities that a feature is assigned to a particular class. Furthermore, most classification algorithms are based on statistical methods. For example, a classifier learns the existence of a fault by observing most frequent patterns that a fault can exhibit. After much consideration of the standard classifiers, the k-Nearest Neighbor (k-NN) algorithm was used for classification in this thesis. This algorithm acts as a majority vote classifier, a data point is assigned to a particular class according to the majority class of its k nearest neighbors [108].

The performance of the classifier was evaluated using the Cross-Validation (CV) algorithm. In general, the CV algorithm makes an certain arbitrary number of partitions of the sensor data such that a portion of these partitions are used for training and the rest are used for testing the classifier. A random test partition is then applied to a classifier to produce the predicted class of the data. The results are tabulated into a confusion matrix, which keeps tallies of the outcomes of the predicted class vs. the actual class of the data. The K-fold cross-validation method was applied

in this thesis. There are different performance measures available from the confusion matrix. One of major importance is *Correct Classification Rate* (CCR). The CCR is defined as the sum of the diagonal elements of the confusion matrix divided by the total number of elements and denotes the percentage of correct classification from all test points used throughout the iterations of the CV process. In addition, the False Alarm Rate (FAR), Missed-Detection Rate (MDR), Specificity (SPC), and Sensitivity (SNS) are also computed. Furthermore, for comparing any two classifier methodologies, the McNmemar's Test of significance [110] is applied.

Chapter 3

Case Study 1: Fault Diagnosis of Motor Drives in Electric Ships

3.1 Problem Statement

The induction motor, and in general electric machine and drive systems, are the de facto standard in the industry due to their consistency of speed control, cost effectiveness, and range of applications including: electric vehicles such as ships, boats, cars, underwater vessels and other applications such as air handling systems, extruders, hoists and conveyors. As with most electrical systems, the electric machines are prone to component failures. Fault detection and diagnosis (FDD) of the motor drive systems are essential for early warning of incipient failures that may help in system recovery and significant improvement of their remaining useful life (RUL). The analysis of these systems during the presence of faults and methods for fault diagnosis have been the subject of interest in the power electronics community since the later part of the 20^{th} century [40–46]. Short-circuit winding failures are some of the most common types of faults, with a 21% of the distribution of faults in electric motor drive systems [47]. The main reason these faults occur is due to the unforeseen breakdown of insulation between components, which may occur across one or more of the phase windings in the stator or across the phase and nearby components. In addition, rotor bar failures have received notable attention. Broken bars are generally caused by stresses

from electromagnetic forces or overloaded operation conditions, inadequate rotor fabrication, and rotor component wear from poor operating environments or lack of maintenance [48]. The broken rotor faults as well as the short-circuit faults are studied in this thesis.

3.2 Model Specifications & Data Generation

A simulation model of the electric motor drive developed and experimentally validated in [80, 81] is used for this research [82], as shown in Fig. 3.1. The system under study is a three-phase induction motor drive operating under indirect Field-Oriented Control (FOC). The motor drive includes a 2300V inverter fed from a 3500V dc source and connected to a 2300V/500A, four-pole, 2250 HP induction machine. Typical motors of this rating are used for driving voluminous marine vessels, such as cargo ships, cruise ships, and other watercraft.

A trapezoidal profile, shown in Fig. 3.2 is implemented as the drive schedule with three

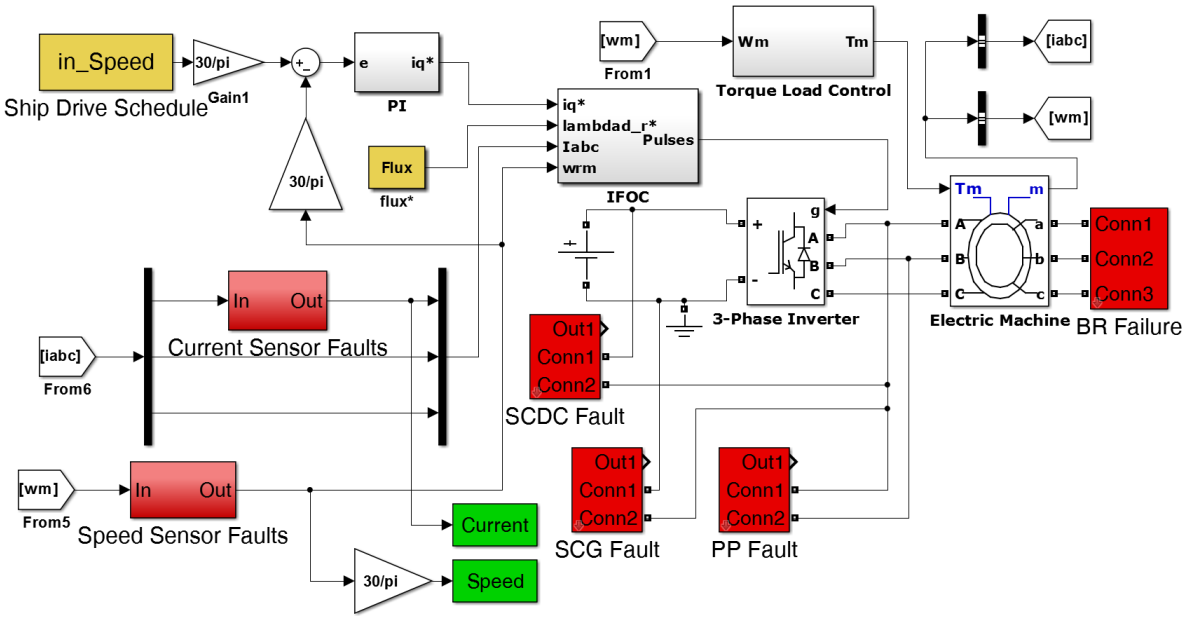

Fig. 3.1: Block Diagram of Electric Motor Drive System w/ Fault Locations

regions: the upward ramp, steady state, and downward ramp, which resemble the start-up, cruising speed, and de-acceleration of the ship, respectively. Sixteen trapezoidal profiles were simulated to study the variation of the fault signatures with respect to motor speed, ranging from maximum peak value of 1750 RPM to 1562.5 RPM in steps of 12.5 RPM. These simulations model 150 seconds of operation under the trapezoidal profile. For this study, the input flux is set to 0.4 V·sec and the torque command is proportional to a quadratic load. The system outputs of measured speed and current are collected after each simulation run.

The critical faults of the drive system are summarized in Table 3.1 and their locations are shown in Fig. 3.1 with red marks. Broken Rotor (BR) faults are simulated by increasing the resistance of the squirrel-cage rotor. The BR fault models about 9-10% loss of torque compared to the Nominal (NOM) condition. This was determined by collecting the torque values at steady state for the BR and NOM conditions. Short circuit (SC) faults included in the analysis are the SC from phase-A to ground (SCG) and phase-to-phase fault (PP) inside the machine. These are simulated by utilizing ideal switches.

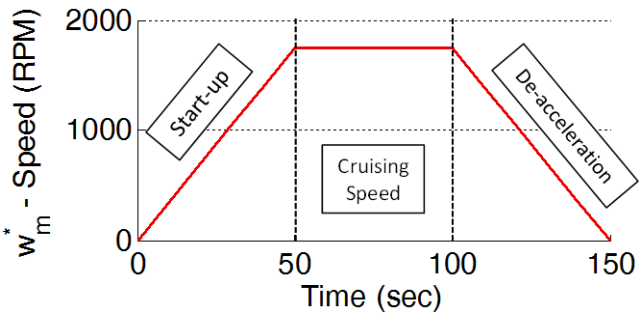

Fig. 3.2: Regions of the Driving Profile for the 1750 peak RPM

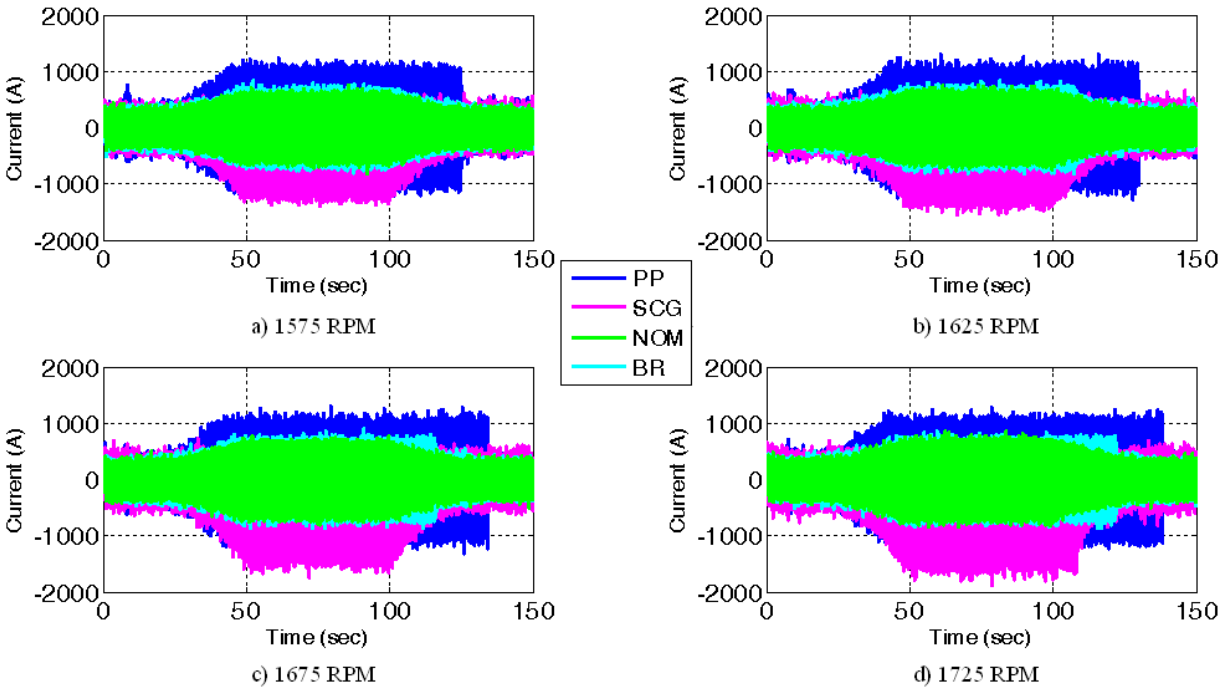

Fig. 3.3: Stator current time-series data for different classes from 1575 through 1725 peak RPM

Table 3.1: Critical Motor Faults

System Component	Fault Name	Fault Description
Induction Motor	BR: Broken Rotor Bar	simulates broken rotor
Power Electronics	PP: Phase-to-Phase	simulates phase-to-phase fault
	SCG: Short Circuit to GND	simulates short to GND

Table 3.2: Data Generation Specifications

Classes	NOM, BR, PP, SCG
Peak Drive Schedules (RPM)	[1562.5:12.5:1750]
Fault Injection Time	0 seconds
Sampling Rate	1kHz
Time range	0-150 seconds
Inputs	Speed command (RPM), Rotor flux (V·sec)
Outputs	Speed (RPM), Current (A)

To model the effects of real-world measurements, Additive White Gaussian Noise (AWGN) is injected to the sensor data, resulting in 20 dB SNR. Four unique instances of these noisy sensor data are generated that serve as additional observations of the sensor data for a particular speed and vehicle health condition. Figure 3.3 depicts the time-series data of the nominal condition and the studied failures for data collected at the 1575, 1625, 1675, and 1725 peak RPM values. In total 256 data sets were generated from the combination of data collected for four classes, sixteen speed profiles, and four instances of AWGN to the sensor data. The specifications for data generation are found in Table 3.2.

Figures 3.4a)-d) depict the wavelet transformed data of the nominal condition and the studied failures for the 1575 peak RPM trapezoidal profile. In this work the 50-55 second interval of the wavelet data, where the speed command is constant, is utilized. Figures 3.5a)-d) illustrate the cell selection operation for the case of the 1575 peak RPM input.

The ν number of cells of length $\frac{n}{b}$ are grouped to form a $(\frac{n}{b} \times \nu)$ data matrix which is the input to the PCA. The resulting feature set is a $(\frac{n}{b} \times q)$ feature matrix, where $q = 12$ dimensions explain the maximum amount of variance in the data matrix. Table 3.3 summarizes the parameters of the motor drive data collected as well as the operation for feature extraction and training.

3.3 Results & Discussion

For this work, K-fold with $K = 10$ is used, where 9 data sets are used for training, and 1 is left for testing. The cross-validation is performed as explained above, where every decision is recorded into the confusion matrix in Table. 3.4. In the ideal case the prediction should match the actual, thus a confusion matrix with large tallies in the diagonal indicate an accurate classifier. The

Table 3.3: Training Data Specifications

Specifications\Database	**Motor Drive System**
Sampling Rate	1kHz
Sensors	Speed, Current Sensors
Classes	NOM, BR, PP, SCG
Observations/Class	{64, 64, 64, 64}
Samples $= (n)$	5000
Mother Wavelet	Meyer
Cell Length $= (\frac{n}{b}; a = m)$	250 $(b = 20)$
Optimal Cells $(= \nu)$	15
Feature Points/Class	{16000, 16000, 16000, 16000}

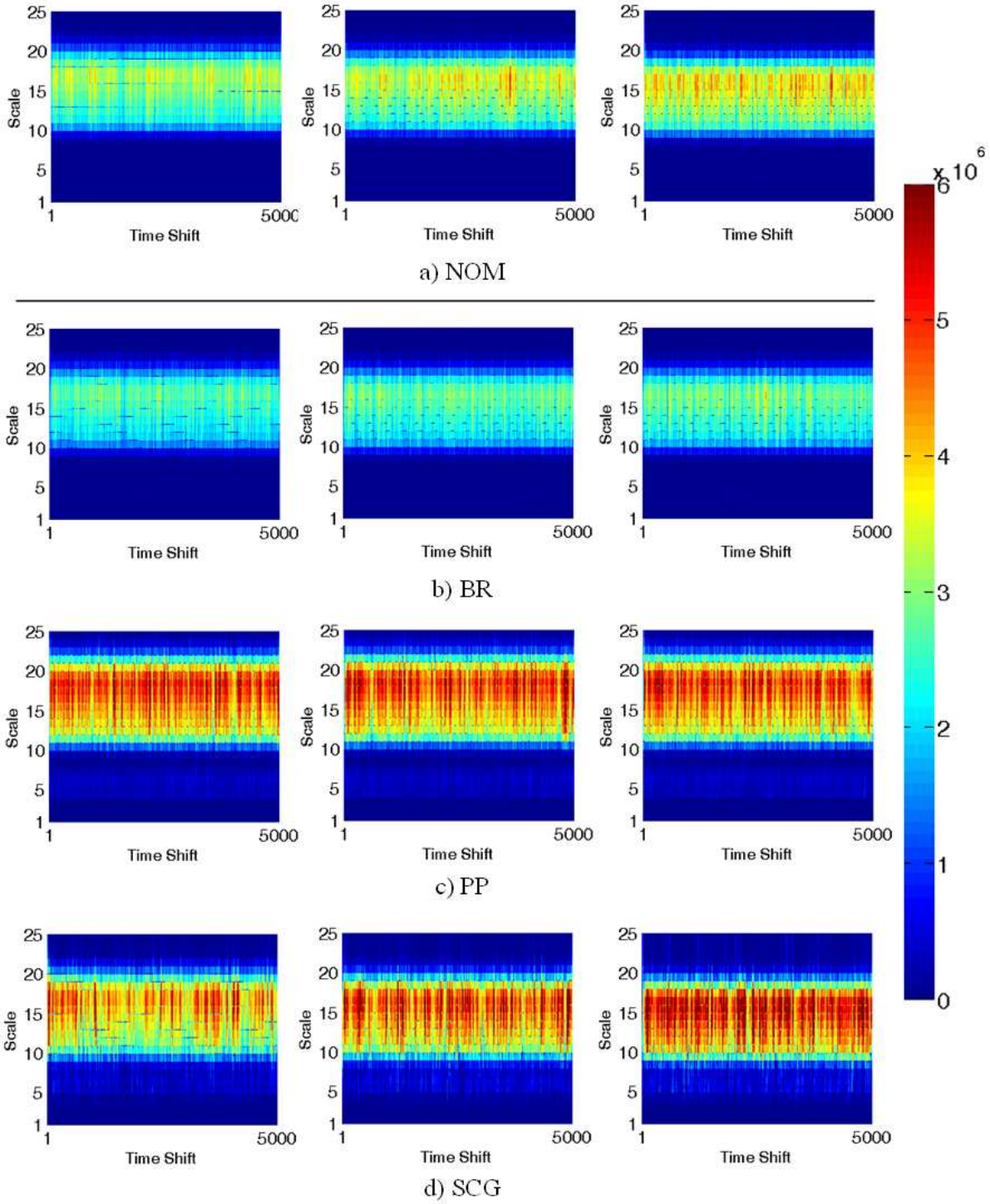

Fig. 3.4: Wavelet analysis for observations of the a)Nominal and b-d)Motor drive faults. Each row displays three specific observations of the same class.

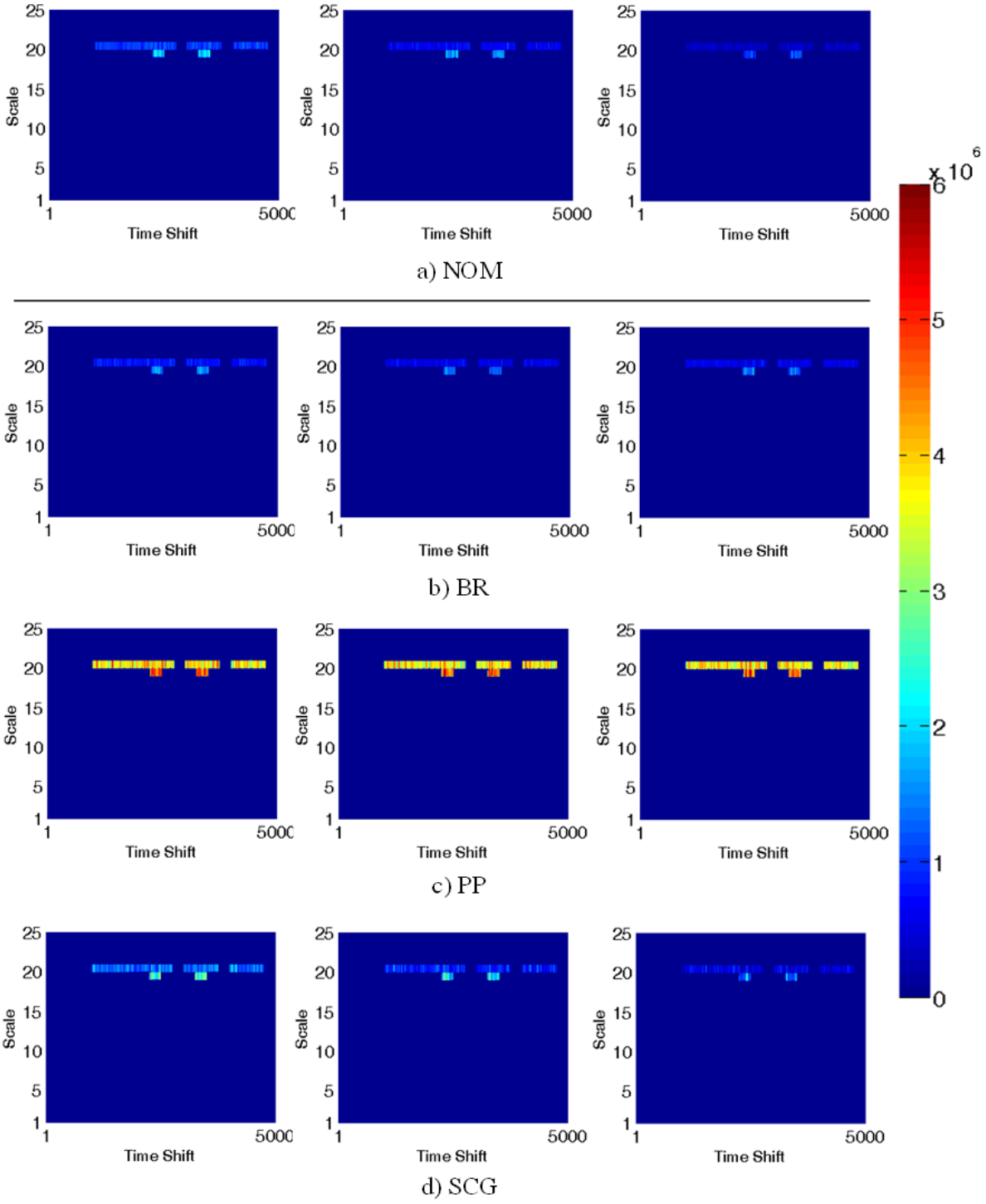

Fig. 3.5: Cell Selection of Fig. 3.4 data, $\nu = 15$ cells extracted

Table 3.4: Confusion Matrix Result

		Classifier Output			
	System Conditions	**NOM**	**BR**	**PP**	**SCG**
Actual	**NOM**	15952	19	0	29
	BR	40	15935	0	25
	PP	0	0	16000	0
	SCG	938	333	0	14729

$k = 3$, CCR = 97.84 %

Table 3.5: Classification Results of Tested Approaches

Feature Extractor	**CCR (%)**	**FAR (%)**	**MDR (%)**	**SPC (%)**	**SNS (%)**	**Time (sec)**
PCA	83.45	18.91	4.62	81.09	95.38	0.455
CWT + PCA	86.86	20.65	6.52	79.35	93.48	6.033
CWT + Optimal Cells + PCA	97.84	0.30	2.04	99.70	97.96	0.249

CCR is computed by taking the trace of the matrix and dividing by the sum of all entries. The proposed architecture attained a CCR value of 97.84% with the k-NN approach with $k = 3$, which indicates high accuracy of the design. After experimentation with the standard mother wavelets, the Meyer wavelet provided the best accuracy after applying the methodology. The machine learning framework generates a diagnosis in real-time quite fast at the expense of the computational burden of training the classifier. However, the training phase is usually performed off-line, which does not hinder the fast response of the decision. In simulations, the testing of one observation of time-series data, which is the scenario that takes place in real-time, only took 0.249 seconds to yield a decision. The simulation environment is a personal computer running Windows 7 Enterprise SP1 64-bit, Intel(R) Core(TM) i5-2400 CPU @ 3.1 GHz, and 16 GB RAM.

Additional comparative analyses were performed to compare the approach to more standard methods. The cross-validation and classifier algorithms were kept consistent to the proposed approach to compare the efficiency and the accuracy of using the approach as the feature extraction. In addition to the performance metrics from the confusion matrix, the McNemar's test of significance [110] was calculated to compare the approach with the standards techniques. In the first comparison, the time-series of the stator current were applied with PCA to extract the features of the classes. This took 0.481 seconds to make a decision. However, 83.45% of a CCR was achieved, which interprets that some features in the time-domain can capture the information, but not sufficiently. The P-value was 6.15×10^{-5}, which was computed by performing a two-tailed hypothesis test with the McNemar statistic and a 0.05 significance level. Second, the wavelet coefficients were computed and applied with PCA to generate the features of the data. This resulted in a higher accuracy of 86.86% CCR. It however, was the

slowest to achieve a decision with 6.033 seconds of time consumed. Because the wavelets compute time-scale information from the data, the computation time increased significantly when applying PCA in the wavelet domain. In the proposed approach, computation is much faster due to the removal of the non-informative cells for classifier training. Furthermore, this confirms that while the wavelet domain can enhance the class separability, there are regions in the wavelets that can hinder separation and thus reduce accuracy when training a classifier with these regions. The P-value here was 7.70×10^{-3}. The results are summarized in Table 3.5, which also included the False Alarm Rate (FAR) and Missed-Detection Rate (MDR).

For an end-user operating the vehicle, it is of interest to have a dashboard display of fault diagnosis in the electric drive. This has been incorporated into a GUI tool for the purposes of receiving such information on-line, such that the driver has the ability to make decisions on the proper action items necessary when a fault is present in the system. The GUI tool is presented in Fig. 3.6. This tool can simulate the data collection for one commanded speed and one fault per simulation. The user has the ability to select what sensor to visualize in the plot. The resulting sensor is then fed to the proposed algorithm that applies the feature selection and the classification to yield the fault diagnosis decision of the data. The diagnosis is reported by a dial object in the GUI that moves to mark the corresponding faulty condition.

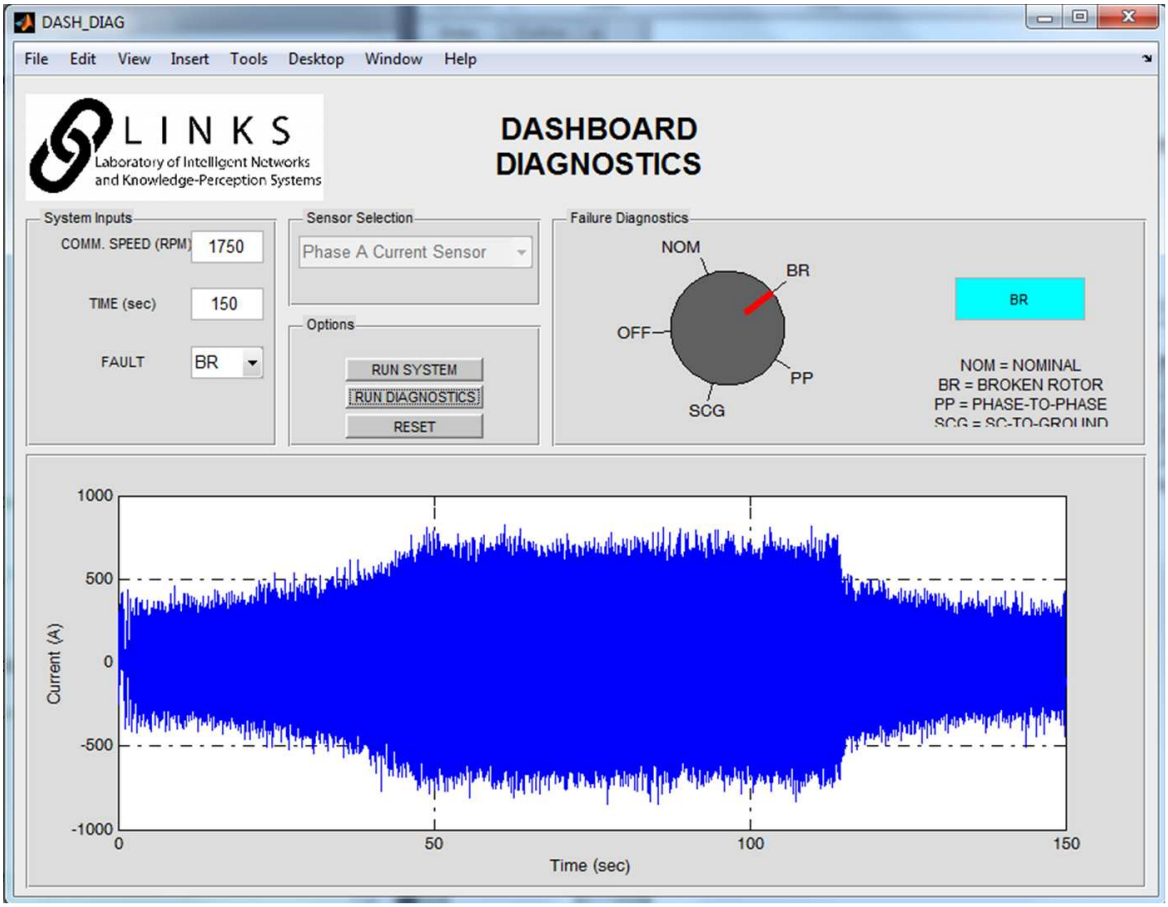

Fig. 3.6: Dashboard Diagnostics GUI

Chapter 4

Case Study 2: Bearing Fault Diagnosis in Induction Motors

4.1 Problem Statement

The flexible application of induction motors and its known benefits of high reliability and low costs make them the workhorse of various industries. Induction motors and relative components are in constant operation and it is critical that their performance be sustained even throughout the most severe operating conditions. These systems typically exhibit bearing failures, which affect the motor wear by increasing the rotational friction of the rotor. As such the needs to diagnose these kinds of failures in the early stage of fault evolution may lead to maximizing the motor remaining useful life and minimizing the amount of maintenance costs. This thesis analyzes the data pertaining to three bearing failures and the nominal condition. This was provided by the authors from Case Western Reserve University (CWRU) [83].

4.2 System Description & Data Generation

The data includes measurements from three accelerometers placed throughout the induction motor system, see Fig. 4.1 for reference. The three bearing failure modes under study are the inner-race fault (IRF), outer-race fault (ORF), and the ball fault (BF). For the experiment, the

system included a 2 hp induction motor, the bearings in the drive end and fan end of the motor, accelerometers at the base, fan end, and drive end, and a dynamometer acting as a load. The bearings failures were introduced via the Electrical Discharge Machining (EDM) method.

Accelerometer data were collected for the 3 types of faults and the nominal condition. The researchers of the original work [83] considered various fault severities and load conditions, which provides for good generalization of classification problem to these conditions. Due to the fast sampling rate and sufficient data points in the steady state, additional observations were added to this study to garner sufficient statistics in the classification step. In the end there were 40 total observations for the nominal condition, 160 observations for the IRF and ORF, and 280 observations for the ball fault. The data from the drive end sensor is shown in Fig. 4.2.

4.3 Data Specifications

Table 4.1 summarizes the parameters of the CWRU data collected as well as the operation for feature extraction and training. For the sake of computational speed, a decimation step was applied to the CWRU Bearings data to reduce the sampling rate and the total number of samples. As part of the wavelet transform, 64 scales were uniformly selected from the interval $[4, 256]$. After experimentation with the standard mother wavelets, the "Gaussian2" wavelet provided the best accuracy after applying the methodology. For the purposes of utilizing the data reduction with the winning cells, each scale information was partitioned independently to each cell, ie., $a = m$. The cell is then a vector of length $\frac{n}{b}$, with mb cells per data set. The operations of Eq's. 2.7, 2.9 are applied to determine the ν winning cells, and these cells are placed into a $(\frac{n}{b} \times \nu)$ matrix. PCA finds the $q < \nu$ principal components and applies the transformation matrix on the data X to

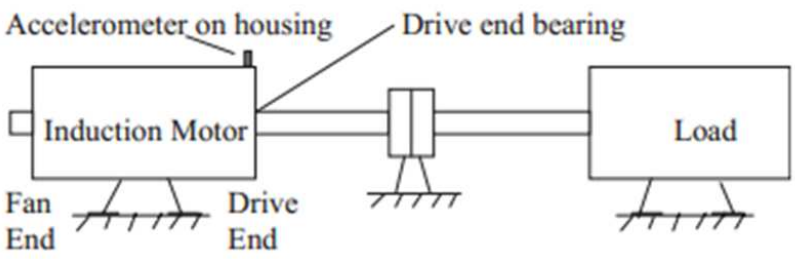

Fig. 4.1: CWRU Bearings Experimental Setup

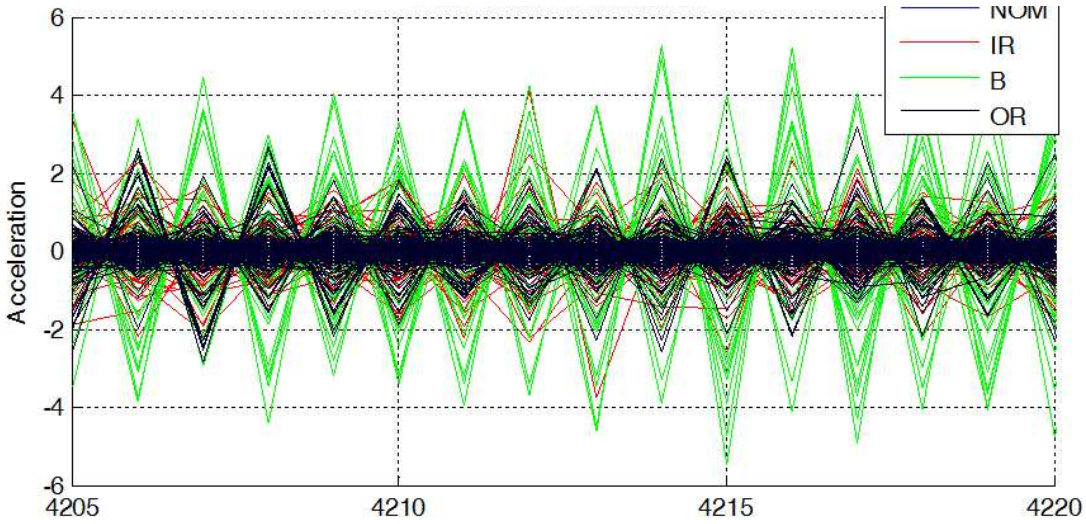

Fig. 4.2: CWRU: Drive End Accelerometer for the four classes

achieve the lower dimensional data of size $(\frac{n}{b} \times q)$.

4.4 Results & Discussion

For this work, the k-NN classification algorithm was implemented. Through various cross-validation iterations, it was found that $k = 3$ is the optimal parameter to classify the different conditions. Figure 4.3 shows the wavelet visualization and Fig. 4.4 zooms in the scale interval to show the candidate regions where the methods determined optimal cells. Each row of the figure displays three specific observations of the same class. Figure 4.5 shows the results of the optimal cell locations in the wavelet data.

From Fig. 4.5, it can be seen as the realization of what cell locations can provide for the best class separability. In the CWRU Bearings data, due to the large sample length, there is some visualization difficulty. However, given these cell size constraints, the algorithm nevertheless found the optimal locations in the wavelet data, where in this case the upper portion of the wavelet provided the highest class separability.

Table 4.2 displays the classification results. The Correct Classification Rate (CCR) was found to be 97.28% with a $k = 3$ k-NN classifier. The False Alarm Rate (FAR), Missed-Detection Rate (MDR), Specificity (SPC), and Sensitivity (SNS) are computed where necessary to provide comparative results with standard methodologies, as found in Table 4.3. The cross-validation and classifier algorithms were kept consistent with the proposed approach to compare the efficiency and the accuracy of using the approach for feature extraction.

In addition to the performance metrics extracted from the confusion matrix, the McNemar's test of significance [110] was calculated to compare the approach with the standards techniques.

Table 4.1: Training Data Specifications

Specifications\Database	**CWRU Bearings**
Sampling Rate	12 kHz [6 kHz]
Sensors	Accel. (3)
Classes	NOM, IRF, ORF, BF
Observations/Class	{40, 160, 160, 280}
Samples $= (n)$	12000 [6000]
Mother Wavelet	Gaussian2
Cell Length $= (\frac{n}{b}; a = m)$	50 $(b = 120)$
Optimal Cells $(= \nu)$	30
Feature Points/Class	{2000, 8000, 8000, 14000}

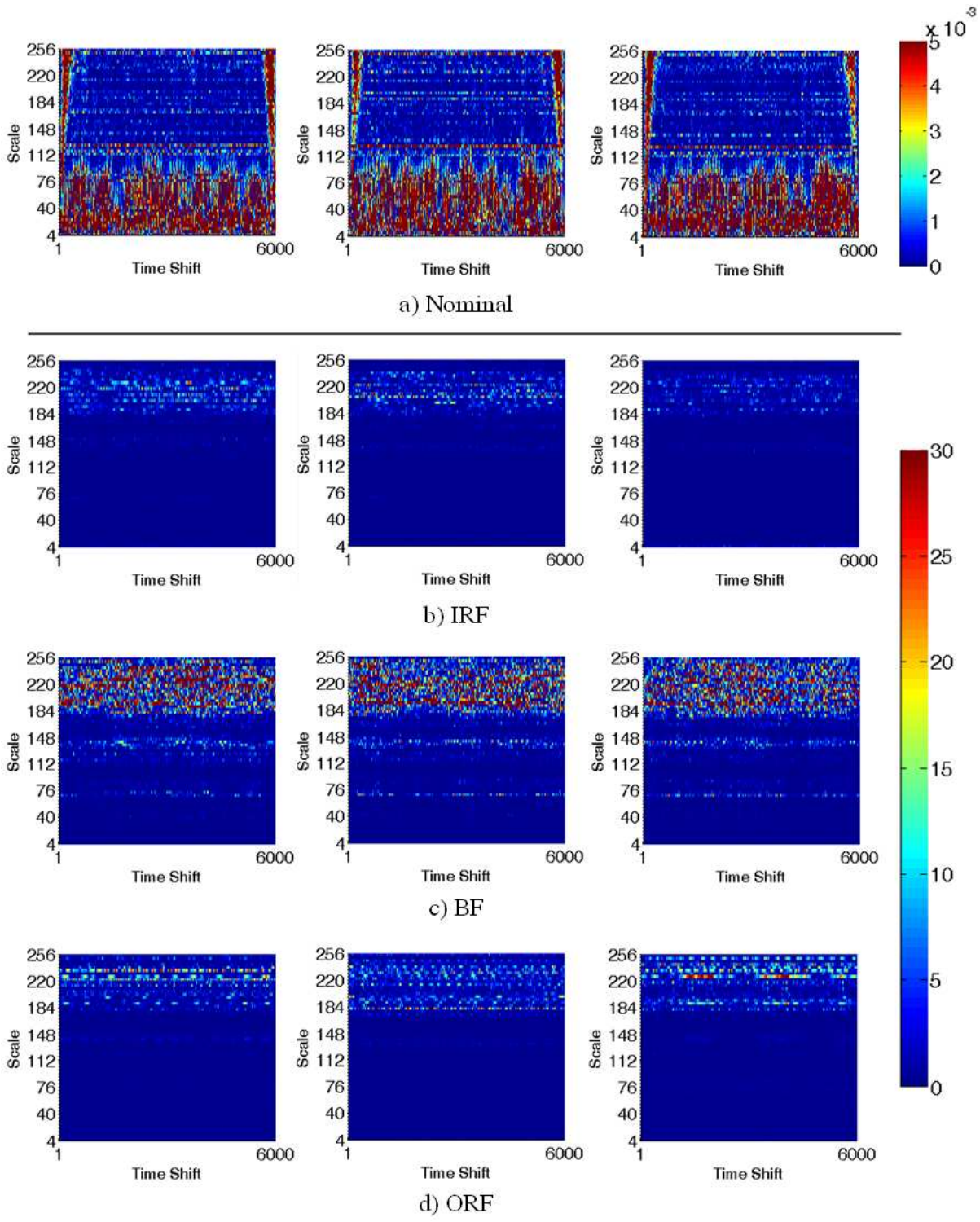

Fig. 4.3: Wavelet analysis for the various observations of the a)Nominal and b-d)Bearing faults. Each row displays three specific observations of the same class.

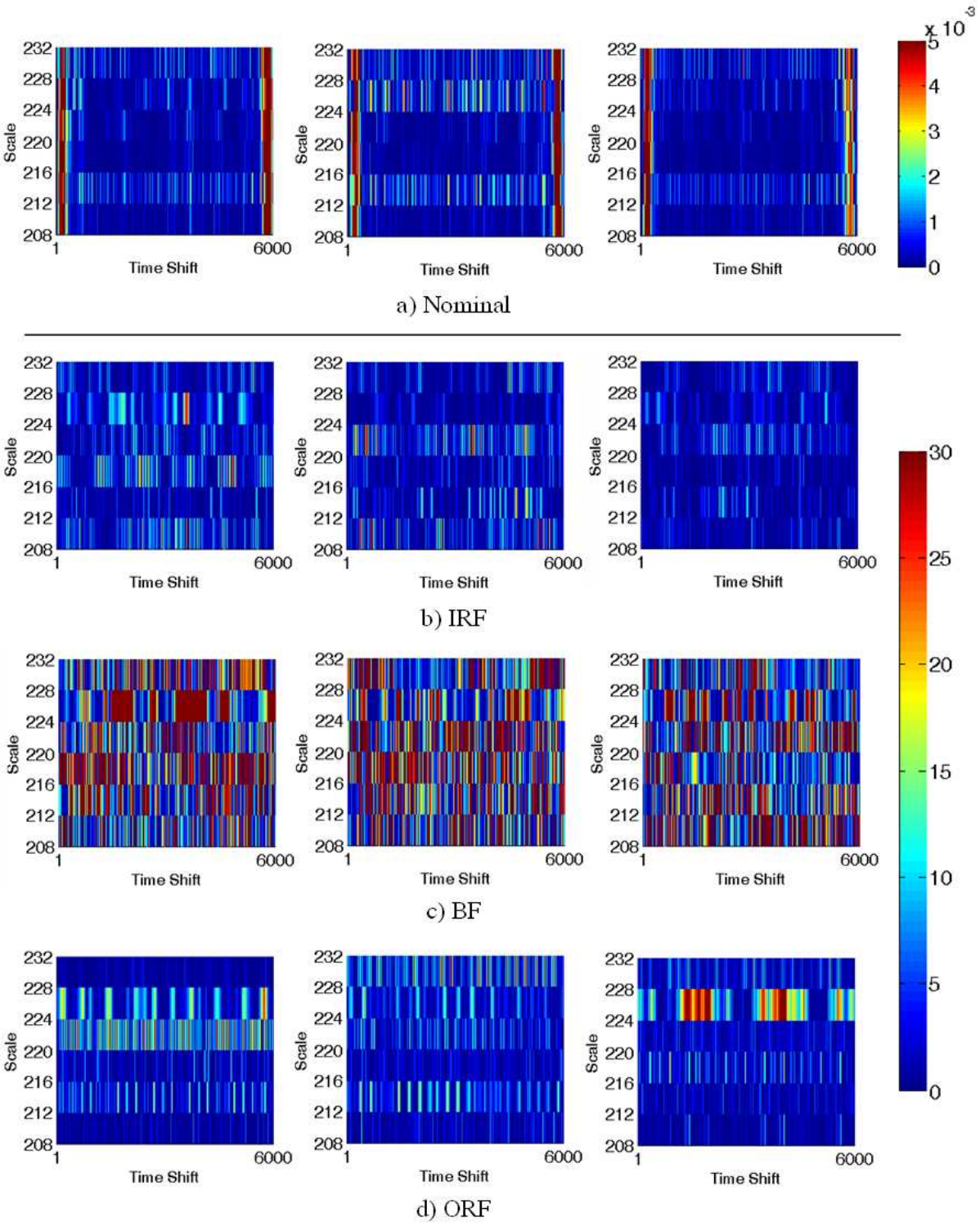

Fig. 4.4: Zoomed-in plots of Fig. 4.3 data

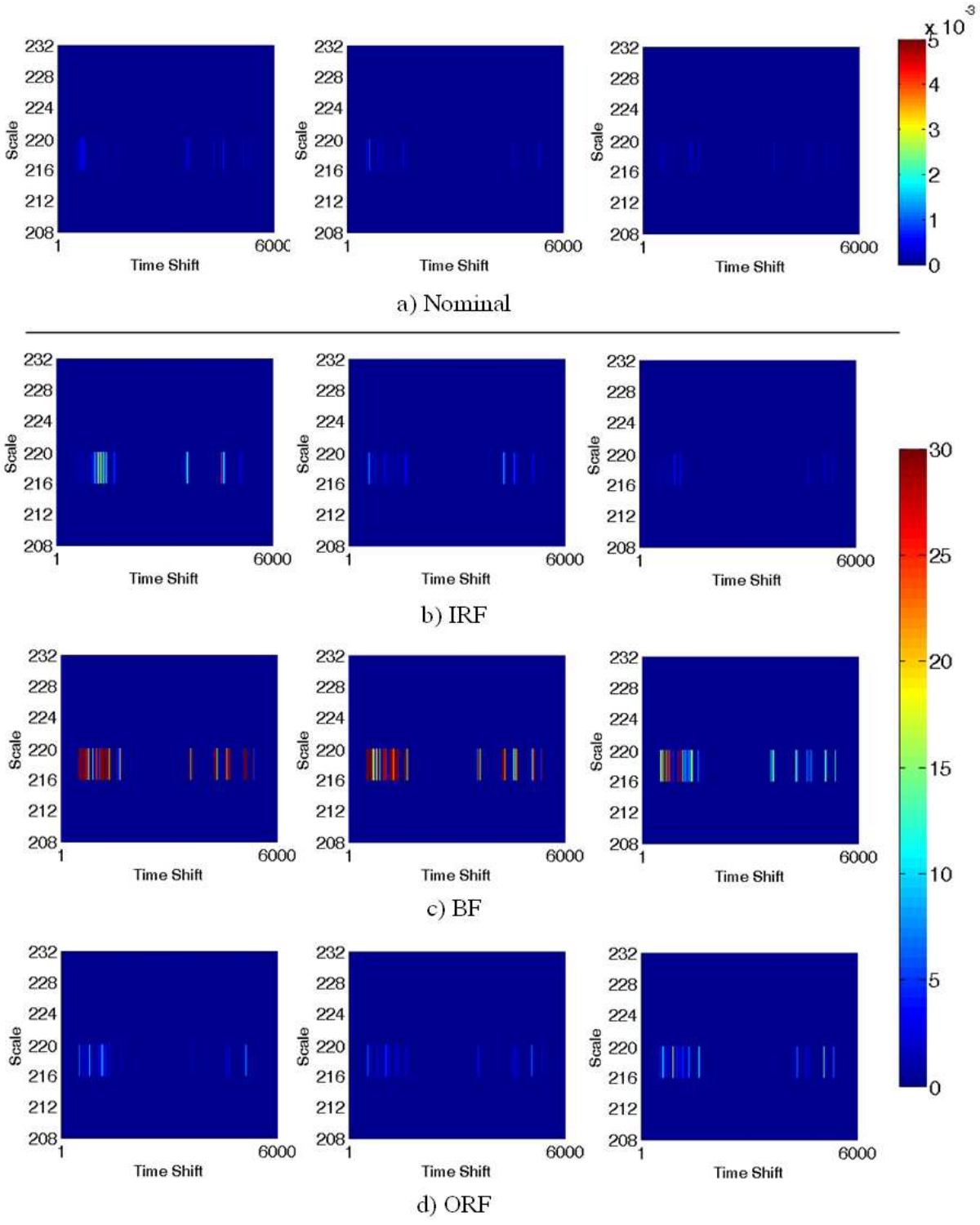

Fig. 4.5: CWRU Cell Selection of Fig. 4.3 data, $\nu = 30$ cells extracted

Table 4.2: Confusion Matrix Result

Classifier Output

Actual: System Conditions	NOM	IRF	BF	ORF
NOM	2000	0	0	0
IRF	0	7584	123	293
BF	0	13	7942	45
ORF	13	190	193	13604

$k = 3$, CCR = 97.28 %

Table 4.3: Classification Results of Tested Approaches

Feature Extractor	**CCR** (%)	**FAR** (%)	**MDR** (%)	**SPC** (%)	**SNS** (%)	**Time** (sec)
CWT + PCA	69.61	8.87	1.00	99.00	91.13	3.222
CWT + Optimal Cells + PCA	97.28	0.00	0.04	100.00	99.96	0.010

Table 4.4: Comparative Results with Approaches in Literature

Author	**Feature Extractor**	**Classifier**	**CCR** (%)	**SPC** (%)	**SNS** (%)
Yuwono et al. [39]	Filtering + PSO	HMM	95.08	92.61	96.31
Proposed	CWT + Optimal Cells + PCA	k-NN	96.87	95.23	99.22

In the comparison, the wavelet coefficients were computed and applied with PCA to generate the features of the data. For the bearings data this resulted in an accuracy of 69.61% CCR with decision-making time of 3.222 seconds. Because the wavelets compute time-scale information from the data, the computation time increased significantly when applying PCA in the wavelet domain. In the proposed approach, computation is much faster due to the removal of the non-informative cells for classifier training. Furthermore, this confirms that while the wavelet domain can enhance the class separability, there are regions in the wavelets that can hinder separation and thus reduce accuracy when training a classifier with these regions. The P-value computed was found to be 2.85×10^{-23}, which was determined by performing a two-tailed hypothesis test with the McNemar statistic and 0.05 significance level.

In addition to accuracy, the time to reach a decision in the testing phase was recorded. The approach was simulated in a computer environment comprising of: Windows 7 Enterprise SP1 64-bit operating system, Intel(R) Core(TM) i5-2400 CPU @ 3.1 GHz, and 16 GB RAM. Overall, the decision-making process was achieved fast enough with the bearings data, although having the fastest sampling rate and tying the variable frequency drive system with the largest class space. In simulations, the testing of one observation of time-series data, which is the scenario that takes place in real-time, only took 0.010 seconds. Nevertheless, these results provide reasonable feedback to what can be expected when applying the algorithm in hardware.

The CWRU Bearings data have undergone extensive analysis from multiple authors over the last decade. The authors that initially proposed the study collected bearings information while considering different input conditions pertaining to various motor speeds and loads. Based on data availability and research interests, the literature contains the work of authors that have utilized

particular subsets of the data found on the CWRU university page [83] to tests their respective FDD solutions. In a recent paper utilizing all of the faulty conditions, Yuwono et al. [39] collected the harmonics of each of the different failures and clustered these harmonics with a swarm-based method. The clusters are further used to train an HMM for classification of only the failures. Utilizing this approach with the studied subset of bearings data, 95.08% CCR was obtained with. The proposed approach netted a 96.94% CCR after using the same reduced data set and the same sensor information applied in this thesis. These calculations are found in Table 4.4.

Chapter 5

Case Study 3: Heart Condition Detection in Medical Patients

5.1 Problem Statement

Cardiac arrhythmia is a medical condition in which the heartbeat of a person is either faster or slower than a typical heartbeat. A particular type of arrhythmia known as Atrial Fibrillation affects nearly 2% of the population in the western countries as of 2014 [84]. Although some episodes of arrhythmia posed harmless effects, there are other instances where the symptoms can include the higher risk for blood clotting within the heart and higher risk of blood flow reduction delivered to the heart. A number of tests for the diagnosis of this heart condition can include electrocardiogram (ECG) devices that measure the electrical activity through the heart. Such a scenario allows for expert cardiologists to discriminate the periods of normal activity and the arrhythmia condition. However, since there are more patients seeking care than doctors available, there is a need for automated heart condition diagnostics such that medical care time is better utilized. This thesis analyzes the MIT-BIH Supraventricular Arrhythmia Database (SVDB) [111, 112] by the authors from Carnegie Mellon University [95].

5.2 System Specification & Data Generation

Figure. 5.1 shows the placements of a typical ECG device on a patient. In this work, the data includes measurements from two electrodes, chest lead V1 and a modified version of Limb Lead II. These measurements are collected from a single patient for a period of time, and the various observations of the nominal and anomalous conditions are overlayed as shown in Fig. 5.2.

5.3 Data Specifications

Table 5.1 summarizes the parameters of the SVDB data collected as well as the operation for feature extraction and training. As part of the wavelet transform, 64 scales were uniformly selected from the interval $[4, 256]$. After experimentation with the standard mother wavelets, the "Gaussian2" wavelet provided the best accuracy after applying the methodology. For the purposes of utilizing the data reduction with the winning cells, each scale information was partitioned independently to each cell, ie., $a = m$. The cell is then a vector of length $\frac{n}{b}$, with mb cells per data set. The operations of Eq's. 2.7, 2.9 are applied to determine the ν winning cells, and these cells are placed into a $(\frac{n}{b} \times \nu)$ matrix. PCA finds the $q < \nu$ principal components and applies the transformation matrix on the data X to achieve the data of size $(\frac{n}{b} \times q)$.

5.4 Results & Discussion

For this work, the k-NN classification algorithm was implemented. Through various cross-validation iterations, it was found $k = 3$ the optimal parameter to classify the nominal and arrhythmia conditions. Figure 5.3 shows the wavelet data plots for this case study. Each column of the figure displays three specific observations of the same class. Figure 5.4 shows the results of

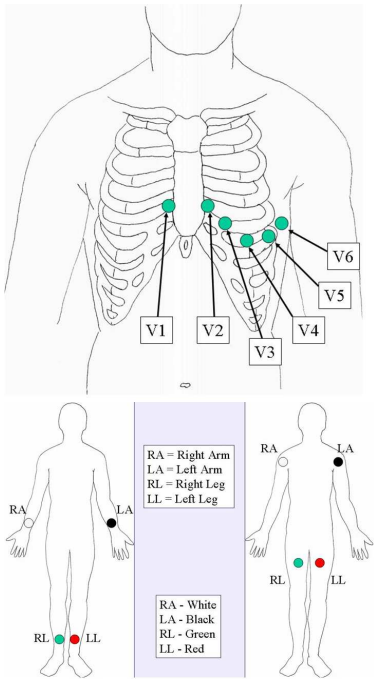

Fig. 5.1: SVDB: Electrode Placement

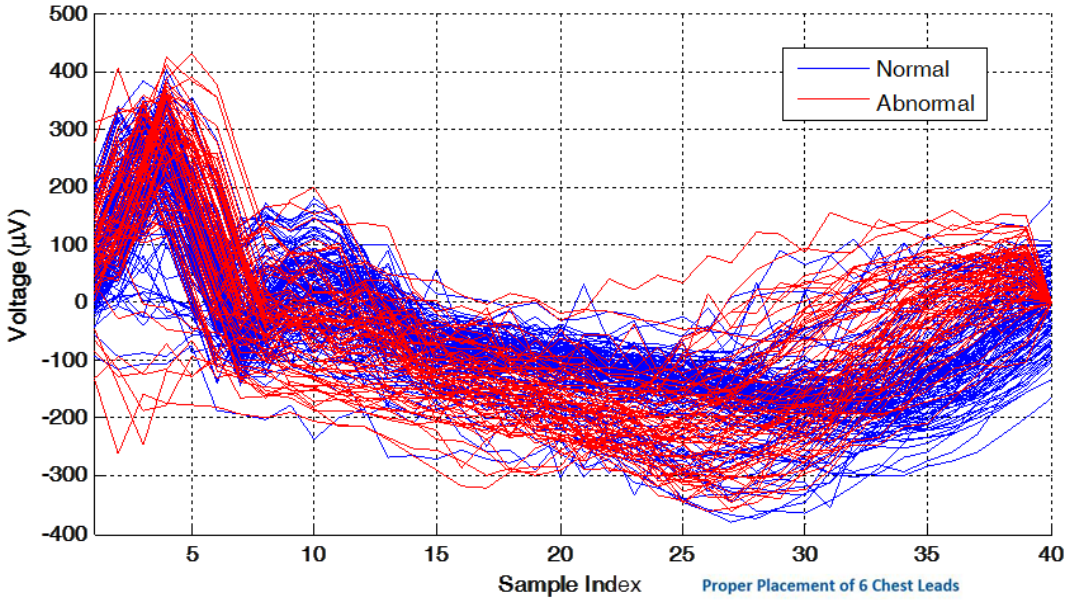

Fig. 5.2: SVDB: Lead II Voltage for the two classes

Table 5.1: Training Data Specifications

Specifications\Database	**MIT-BIH SVDB**
Sampling Rate	128 Hz
Sensors	Electrode Leads (2)
Classes	NOM, Arrhythmia
Observations/Class	{133, 67}
Samples $= (n)$	40
Mother Wavelet	Gaussian2
Cell Length $= (\frac{n}{b}; a = m)$	40 $(b = 1)$
Optimal Cells $(= \nu)$	31
Feature Points/Class	{5320, 2680}

the optimal cell locations in the wavelet data.

From Fig. 5.4, it can be seen what cell locations provide for the best class separability. In the SVDB data, the length of the cell was chosen as the length of the complete time-series. This was selected by method of trial and error, where the length and a sufficient amount of winning cells played a role in the input matrix to the PCA. Given these cell size constraints, the algorithm nevertheless found the optimal locations in the wavelet data, where in this case the upper portion of the wavelet provided the highest class separability.

Table 5.2 displays the classification results. The Correct Classification Rate (CCR) was found to be 96.27% with a $k = 3$ k-NN classifier. The False Alarm Rate (FAR), Missed-Detection Rate (MDR), Specificity (SPC), and Sensitivity (SNS) are computed where necessary to provide comparative results with standard methodologies, as found in Table 5.3. The cross-validation and classifier algorithms were kept consistent with the proposed approach to compare the efficiency and the accuracy of using the approach for feature extraction.

In addition to the performance metrics extracted from the confusion matrix, the McNemar's test of significance [110] was calculated to compare the approach with the standards techniques. In the comparison, the wavelet coefficients were computed and applied with PCA to generate the features of the data. For the SVDB data, this resulted in an accuracy of 97.40% CCR with decision-making time of 0.072 seconds. This interprets that while the complete domain has informative features for classification, there exists features that are more informative than others, as the proposed method utilized about half of the total number of cells in the wavelet to achieve a just as high of an accuracy. Furthermore, the selection of a subset of cells is also beneficial to the testing time achieved. This is the importance for selectively filtering out this

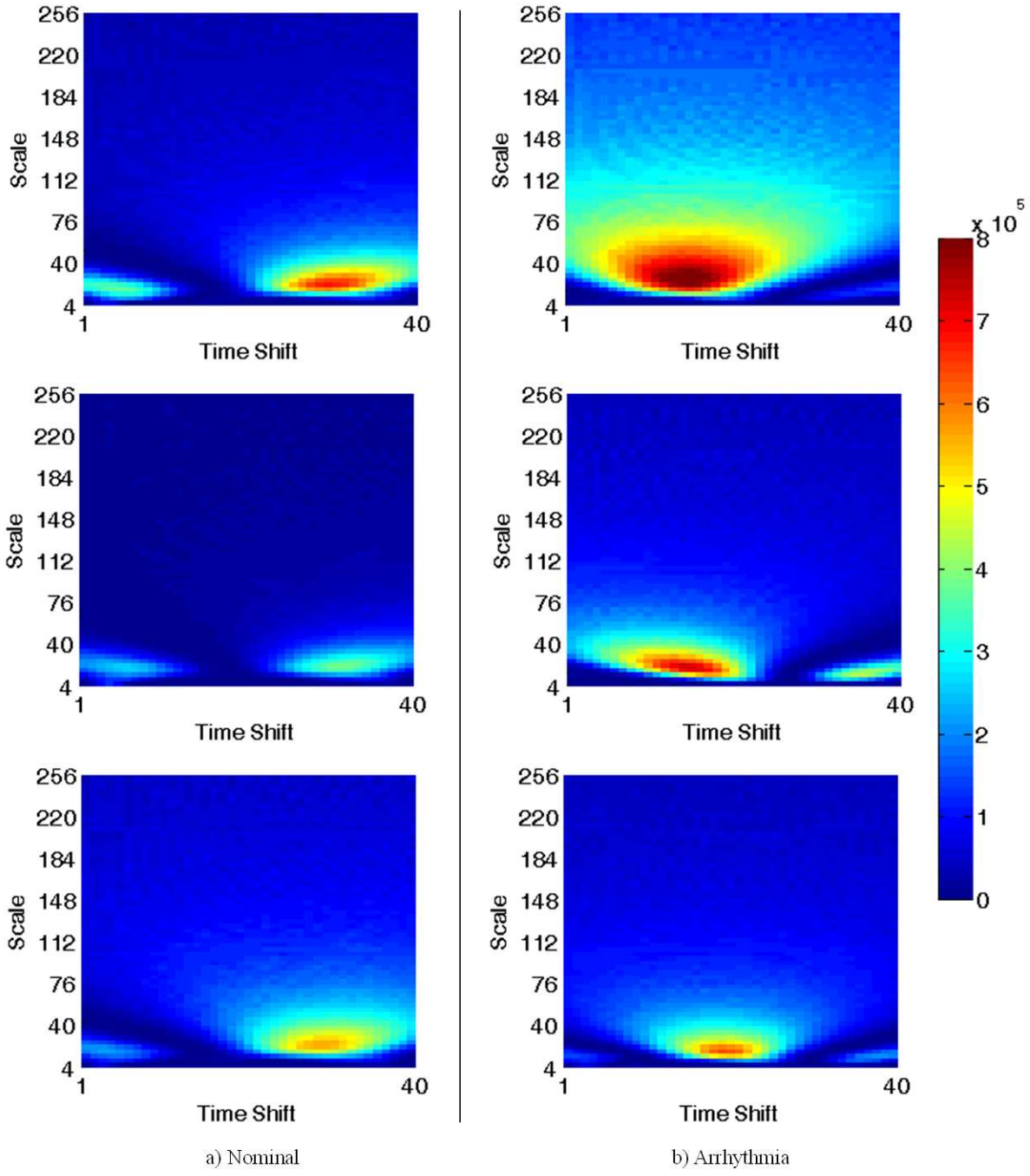

Fig. 5.3: Wavelet analysis for the various observations of the a)Nominal and b)Arrhythmia classes

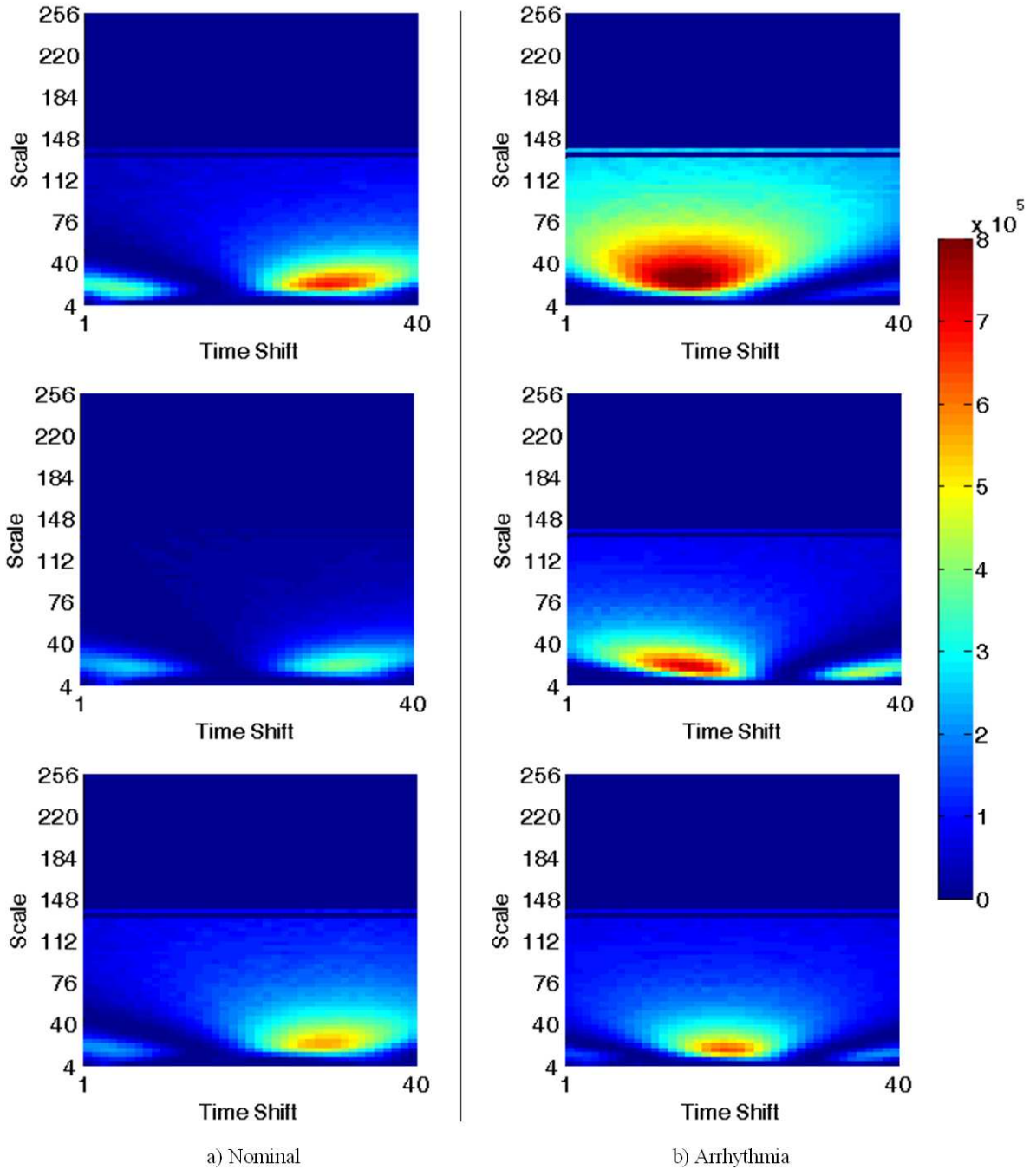

Fig. 5.4: SVDB Cell Selection of Fig. 5.3 data, $\nu = 31$ cells extracted

Table 5.2: Confusion Matrix Result

Actual: Patient Conditions	Classifier Output: NOMINAL	Classifier Output: ARRYHTHMIA
NOMINAL	5256	64
ARRYHTHMIA	234	2446

$k = 3$, CCR = 96.27 %

Table 5.3: Classification Results of Tested Approaches

Feature Extractor	CCR (%)	FAR (%)	MDR (%)	SPC (%)	SNS (%)	Time (sec)
CWT + PCA	97.40	0.64	6.49	99.36	93.51	0.072
CWT + Optimal Cells + PCA	96.27	1.20	8.73	98.80	91.27	0.024

Table 5.4: Comparative Results with Approaches in Literature

Author	Feature Extractor	Classifier	SPC (%)	SNS (%)
Olszewski [95]	FFT	CART	90.30	79.70
Olszewski [95]	Structural Features	CART	91.30	90.80
Proposed	CWT + Optimal Cells + PCA	k-NN	98.80	91.27

information to preserve maximum class separability and maximizing testing speed. The P-value was found to be 0.462, which was computed by performing a two-tailed hypothesis test with the McNemar statistic and 0.05 significance level.

In addition to accuracy, the time to reach a decision in the testing phase was recorded. The approach was simulated in a computer environment comprising of: Windows 7 Enterprise SP1 64-bit operating system, Intel(R) Core(TM) i5-2400 CPU @ 3.1 GHz, and 16 GB RAM. Overall, the testing time of one observation of time-series data was achieved at one of the faster rates with the SVDB data set, a time of 0.024 seconds, which would be indicative of the relatively slower sampling rate and tied with the smallest class space to train out of all the other applications.

Other comparative results were found from the original work performed by the authors that generated the data as shown in Table 5.4. In the SVDB data, a method based on Structural Features was proposed, which was paired with a Classification and Regression Tree (CART) classifier [95]. They also compared their results with frequency analysis of the lead data and the CART classifier. The former yielded the SPC of 91.30% and SNS of 90.80%. Utilizing the proposed method, the SPC was 98.80% and the SNS was 91.27%.

The tuning parameter of the number of cells, mentioned in Table 5.1, adds a significant contribution to the separability of classes in the feature space when using the proposed method. To achieve a high CCR with an efficient number of these cells a plot of the CCR with respect to the this tuning parameter is presented in Fig. 5.5. Starting with a minimum number of ν cells in the parameter search, the optimal number of cells for the thesis was selected such that the CCR is maximized while minimizing these number of cells. These selected cells are the most informative regions in the wavelet data as they achieve a high CCR without utilizing the complete wavelet data.

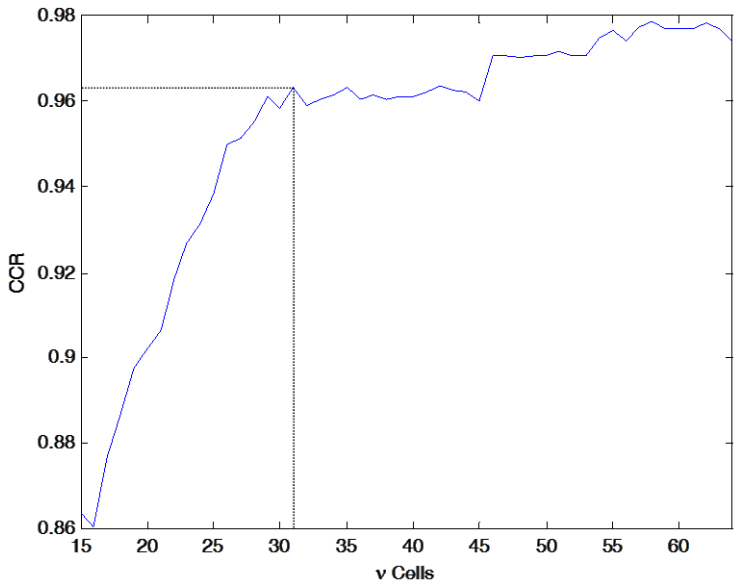

Fig. 5.5: Achieved CCR utilizing the top ν cells

Chapter 6

Case Study 4: Freezing of Gait Detection in Medical Patients

6.1 Problem Statement

Parkinson's Disease (PD) is a degenerative disorder that occurs in the central nervous system. It affects the creation of dopamine-generating cells in the region of the mid-brain, which result in the impairment of a person's motor skills. The cause of the cell destruction is currently unknown. The periods where the motor skills of a patient are impaired are commonly known as freezing conditions. Typical freezing occurs in a patient's legs, and this particular symptom is known as a Freezing of Gait (FOG). The FOG symptom causes discomfort for the patient, as it interferes with Activities of Daily Living (ADL) and ultimately impairs their quality of life [96, 97]. While current medicinal solutions cope with the symptom by reduction of the FOG episode, these results are only effective within a brief time interval. Furthermore, as the disease progresses, medicinal solutions are not as effective as in the disease's early stages. This thesis analyzes the Daphnet Freezing of Gait (FOG) data set [105] submitted to the UCI Machine Learning Repository for classification of the nominal and PD symptom.

6.2 System Description & Data Generation

The data includes measurements from three 3-axis accelerometers placed throughout the patient's body, specifically in the patient's hip, thigh, and ankle, as shown in Fig. 6.1. As part of the data generation process in the experiment, patients were asked to performed three types of walking tasks:

1. Walking in a straight line, a 180° turn and repeat.
2. Random walking in a hall area, including periods of stoppage and 360° turns.
3. Walking to simulate patient ADL (e.g., entering/leaving rooms, going to the kitchen for a cup of water and returning to the starting room).

Various observations of the FOG anomaly are exhibited during the each patient experiment. Once all accelerometer data for each patient and each walking task are collected, the data for each class is combined into one data matrix. This amounted to a matrix size of 40 rows (the sample length) and 315 columns (total number of observations) for the nominal class and (40×265) data matrix for the FOG class. The data from the ankle sensor is shown in Fig. 6.2.

6.3 Data Specifications

Table 6.1 summarizes the parameters of the Daphnet FOG data collected as well as the operation for feature extraction and training. As part of the wavelet transform, 64 scales were uniformly selected from the interval $[4, 256]$. After experimentation with the standard mother wavelets, the "Coiflet1" wavelet provided the best accuracy after applying the methodology. For the purposes of utilizing the data reduction with the winning cells, each scale information was

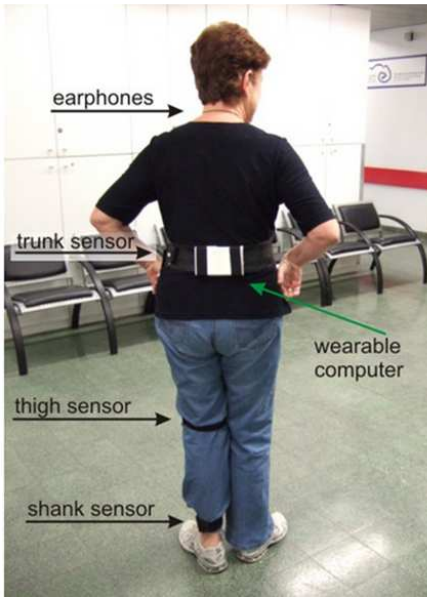

Fig. 6.1: FOG: Patient Sensor Placement

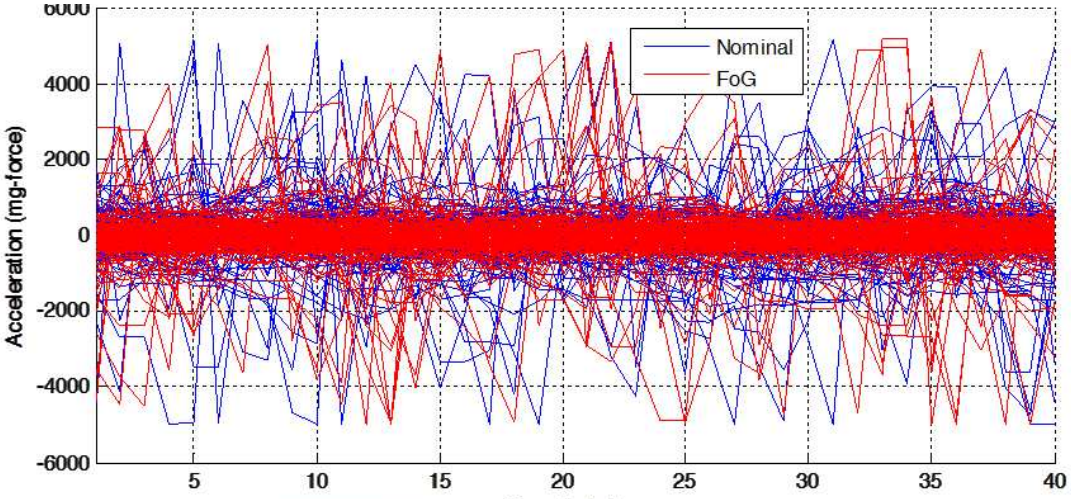

Fig. 6.2: FOG: Ankle Accelerometer for the two classes

partitioned independently to each cell, ie., $a = m$. The cell is then a vector of length $\frac{n}{b}$, with mb cells per data set. The operations of Eq's. 2.7, 2.9 are applied to determine the ν winning cells, and these cells are placed into a $(\frac{n}{b} \times \nu)$ matrix. PCA finds the $q < \nu$ principal components and applies the transformation matrix on the data X to achieve the feature data of size $(\frac{n}{b} \times q)$.

Table 6.1: Training Data Specifications

Specifications\Database	**Daphnet FOG**
Sampling Rate	64 Hz
Sensors	{X,Y,Z}-Accel. (3)
Classes	NOM, FOG
Observations/Class	{315, 265}
Samples $= (n)$	40
Mother Wavelet	Coiflet1
Cell Length $= (\frac{n}{b}; a = m)$	20 $(b = 2)$
Optimal Cells $(= \nu)$	49
Feature Points/Class	{6300, 5300}

6.4 Results & Discussion

For this work, the k-NN classification algorithm was implemented. Through various cross-validation iterations, it was found $k = 3$ the optimal parameter to classify between the nominal condition and the FOG symptom. Figure 6.3 shows the wavelet data plots for this case study. Each column of the figure displays three specific observations of the same class. Figure 6.4 shows the results of the optimal cell locations in the wavelet data.

From Fig. 6.4, it can be seen what the cell locations can provide for the best class separability. In the FOG data, the cell length was that of half of the time-series length, and as such generally had a more spread of optimal cell locations. This was selected by method of trial and error, where the length of the cell and number of winning cells played a role in the input matrix to the PCA. Given these constraints, the algorithm nevertheless found the optimal locations in the wavelet data.

Table 6.2 displays the classification results. The Correct Classification Rate (CCR) was found to be 95.30% with a $k = 3$ k-NN classifier. The False Alarm Rate (FAR), Missed-Detection Rate (MDR), Specificity (SPC), and Sensitivity (SNS) are computed where necessary to provide comparative results with standard methodologies, as found in Table 6.3. The cross-validation and classifier algorithms were kept consistent with the proposed approach to compare the efficiency and the accuracy of using the approach for feature extraction.

In addition to the performance metrics extracted from the confusion matrix, the McNemar's test of significance [110] was calculated to compare the approach with the standards techniques. In the comparison, the wavelet coefficients were computed and applied with PCA to generate the features of the data. For the FOG data this resulted in an accuracy of 95.5% CCR with decision-making time of 0.096 seconds. This interprets that while the complete domain has

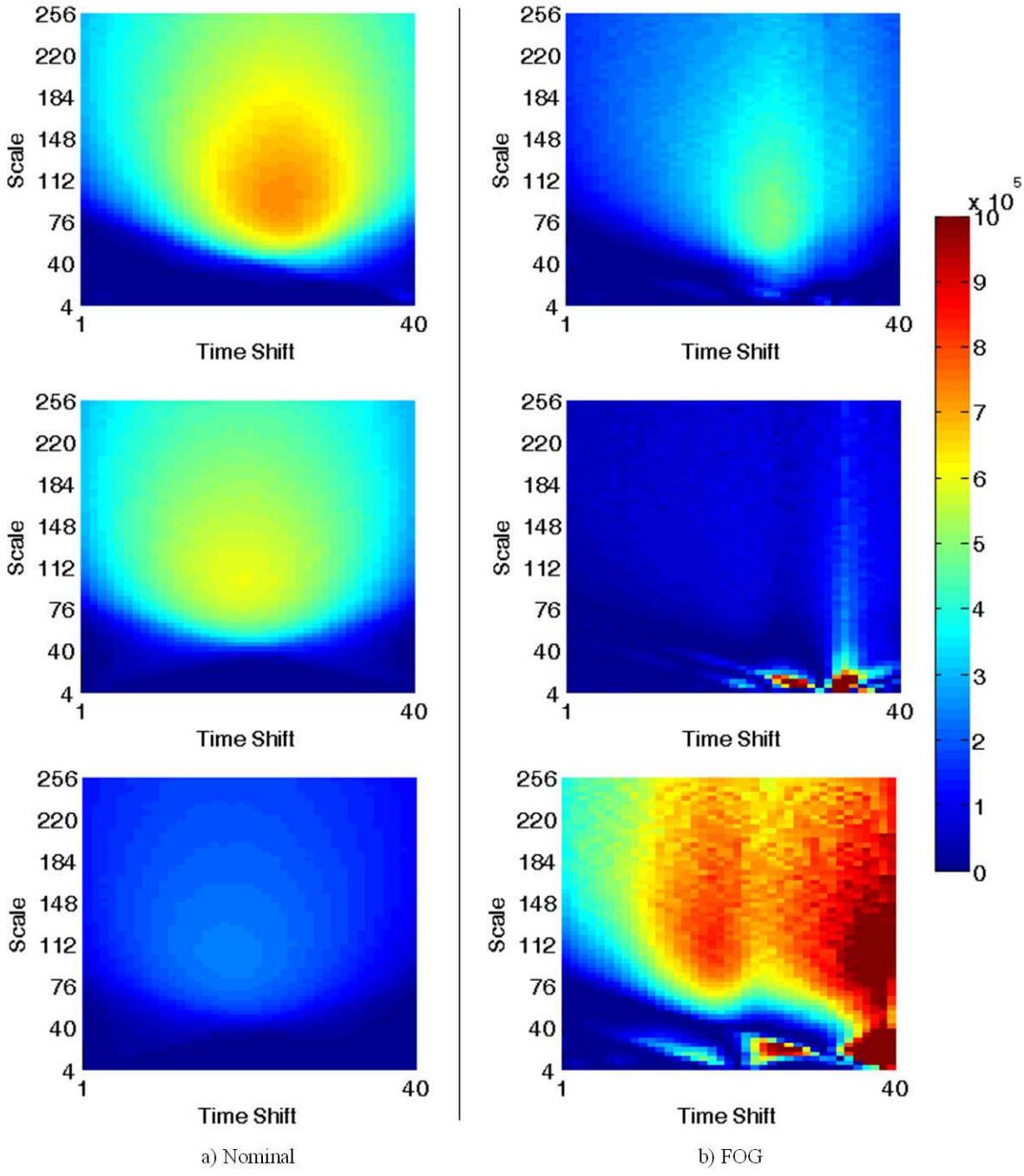

Fig. 6.3: Wavelet analysis for the various observations of the a)Nominal and b)FOG classes

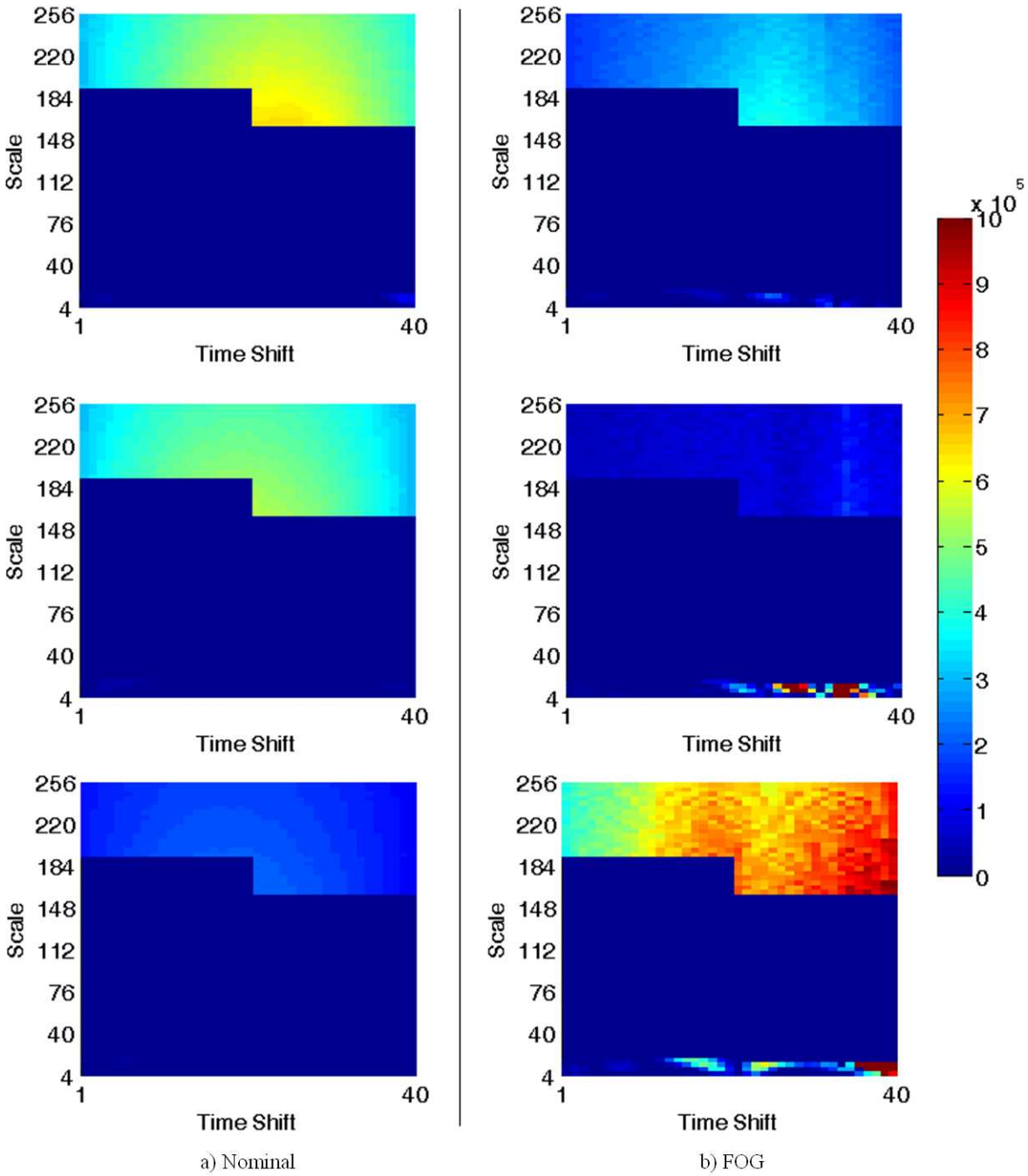

Fig. 6.4: SVDB Cell Selection of Fig. 6.3 data, $\nu = 49$ cells extracted

Table 6.2: Confusion Matrix Result

Classifier Output

Actual: Patient Conditions	NOM	FOG
NOM	6057	243
FOG	302	4998

$k = 3$, CCR = 95.30 %

Table 6.3: Classification Results of Tested Approaches

Feature Extractor	CCR (%)	FAR (%)	MDR (%)	SPC (%)	SNS (%)	Time (sec)
CWT + PCA	95.50	3.57	5.60	94.40	96.43	0.096
CWT + Optimal Cells + PCA	95.30	3.86	5.70	96.14	94.30	0.020

Table 6.4: Comparative Results with Approaches in Literature

Author	Feature Extractor	Classifier	SPC (%)	SNS (%)
Bachlin et al. [105]	FFT + Statistical Features	Thresholding	81.60	73.10
Mazilu et al. [106]	Time-domain Features	RF	95.38	66.25
Proposed	CWT + Optimal Cells + PCA	k-NN	96.14	94.30

informative features for classification, there exists features that are more informative than others, as the proposed method utilized about half of the total number of cells in the wavelet to achieve a just as high of an accuracy. Furthermore, the selection of a subset of cells is also beneficial to the testing time achieved. This is the importance for selectively filtering out this information to preserve maximum class separability and maximizing testing speed. The P-value was found to be 1.25, which was computed by performing a two-tailed hypothesis test with the McNemar statistic and 0.05 significance level.

In addition to accuracy, the time to reach a decision in the testing phase was recorded. The approach was simulated in a computer environment comprising of: Windows 7 Enterprise SP1 64-bit operating system, Intel(R) Core(TM) i5-2400 CPU @ 3.1 GHz, and 16 GB RAM. Overall, the testing time of one observation of time-series data was achieved at a faster rate with the FOG data set, a time of 0.020 seconds, which would be indicative of the relative slowest sampling rate and smaller class space to train out of all the other applications.

Other comparative results were found from the original work performed by the authors that generated the data as shown in Table 6.4. In the Daphnet FOG data, a threshold-based classification rule was proposed [105]. The measurements were transformed to the frequency domain where the energy of the signal at different frequency bands constructed the decision function. Other collaborators proposed a methodology within a machine learning framework, where the Random Forest (RF) classifier was trained with time-domain features [106]. In both scenarios, tests were performed to determine the accuracy of the methods when training with the data for each patient (patient-dependent), and for training with the whole data (patient-independent). The comparative results highlighted are for the patient-independent case and are consistent with our approach. The

latter study yielded the SPC of 95.38% and SNS of 66.25%. Utilizing the proposed method, the SPC was 96.14% and the SNS was 94.30%.

The tuning parameter of the number of cells, mentioned in Table 6.1, adds a significant contribution to the separability of classes in the feature space when using the proposed method. To achieve a high CCR with an efficient number of these cells a plot of the CCR with respect to the this tuning parameter is presented in Fig. 6.5. Starting with a minimum number of ν cells in the parameter search, the optimal number of cells for the thesis was selected such that the CCR is maximized while minimizing these number of cells. These selected cells are the most informative regions in the wavelet data as they achieve a high CCR without utilizing the complete wavelet data.

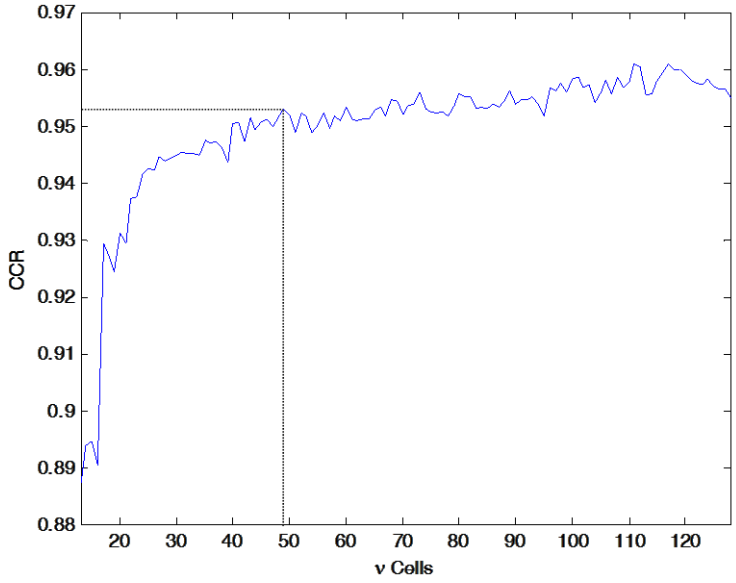

Fig. 6.5: Achieved CCR utilizing the top ν cells

Chapter 7

Discussion, Conclusion & Future Work

The thesis presented a methodology for feature extraction utilizing a wavelet-based approach. In the wavelet domain, the optimal cells that maximize the between-class distance and minimize the within-class distance are selected to enhance the class separability to improve the performance of classification algorithms. Four different problems with applications in pattern recognition were tested with this approach and a consistent degree of high accuracy was found, with results achieving $> 95\%$ CCR across the four applications. These results were compared to the works of the original authors and other standard approaches. Also, the method has been found to perform the classification in the testing phase within fractions of a second. From these results, the approach confirms significant improvements on the prior works in the field. The table summary of the data specifications for all the applications part of thesis is found in Table 7.1. The classification results table for all the applications part of thesis and their respective comparison with standard approaches are found in Table 7.2. The comparison table of the classification results of this thesis with results from other authors using the same applications is found in Table 7.3.

The future work for the motor drive in electric ships involves the online implementation of the classifier using an experimental test-bed that is currently being designed. Additionally, the

Table 7.1: Summary of Training Data Specifications

Specifications\Database	**Motor Drive System**	**CWRU Bearings**	**MIT-BIH SVDB**	**Daphnet FOG**
Sampling Rate	1kHz	12 kHz [6 kHz]	128 Hz	64 Hz
Sensors	Speed, Current Sensors	Accel. (3)	Electrode Leads (2)	{X,Y,Z}-Accel. (3)
Classes	NOM, BR, PP, SCG	NOM, IRF, ORF, BF	NOM, Arrhythmia	NOM, FOG
Observations/Class	{64, 64, 64, 64}	{40, 160, 160, 280}	{133, 67}	{315, 265}
Samples $= (n)$	5000	12000 [6000]	40	40
Mother Wavelet	Meyer	Gaussian2	Gaussian2	Coiflet1
Cell Length $= (\frac{n}{b}; a = m)$	250 $(b = 20)$	50 $(b = 120)$	40 $(b = 1)$	20 $(b = 2)$
Optimal Cells $(= \nu)$	15	30	31	49
Feature Points/Class	{16000, 16000, 16000, 16000}	{2000, 8000, 8000, 14000}	{5320, 2680}	{6300, 5300}

Table 7.2: Summary of Classification Results

Data Set (Classes)	**Solution**	**Feature Extractor**	**CCR (%)**	**FAR (%)**	**MDR (%)**	**SPC (%)**	**SNS (%)**	**Time (sec)**
Motor Drive System (4)	Comparative	PCA	83.45	18.91	4.62	81.09	95.38	0.455
	Comparative	CWT + PCA	86.86	20.65	6.52	79.35	93.48	6.033
	Proposed	CWT + Optimal Cells + PCA	97.84	0.30	2.04	99.70	97.96	0.249
CWRU Bearings (4)	Comparative	CWT + PCA	69.61	8.87	1.00	99.00	91.13	3.222
	Proposed	CWT + Optimal Cells + PCA	97.28	0.00	0.04	100.00	99.96	0.010
MIT-BIH SVDB (2)	Comparative	CWT + PCA	97.40	0.64	6.49	99.36	93.51	0.072
	Proposed	CWT + Optimal Cells + PCA	96.27	1.20	8.73	98.80	91.27	0.024
Daphnet FOG (2)	Comparative	CWT + PCA	95.50	3.57	5.60	94.40	96.43	0.096
	Proposed	CWT + Optimal Cells + PCA	95.30	3.86	5.70	96.14	94.30	0.020

Table 7.3: Summary of Comparative Results from Literature

Data Set (Classes)	**Author**	**Feature Extractor**	**Classifier**	**SPC (%)**	**SNS (%)**
CWRU Bearings (4)	Yuwono et al. [39]	Filtering + PSO	HMM	92.61	96.31
	Proposed	CWT + Optimal Cells + PCA	k-NN	95.23	99.22
MIT-BIH SVDB (2)	Olszewski [95]	FFT	CART	90.30	79.70
	Olszewski [95]	Structural Features	CART	91.30	90.80
	Proposed	CWT + Optimal Cells + PCA	k-NN	98.80	91.27
Daphnet FOG (2)	Bachlin et al. [105]	FFT + Statistical Features	Thresholding	81.60	73.10
	Mazilu et al. [106]	Time-domain Features	RF	95.38	66.25
	Proposed	CWT + Optimal Cells + PCA	k-NN	96.14	94.30

study of the proposed method onto automobile drive schedules is the next direction of this work, where the driving input targets stop-and-go traffic dynamics. The fault diagnosis problem will also account for a larger set of the electric machine fault universe. A GUI has been built from the initial success of the research and can be an initial platform for onboard diagnostics of these types of failures in electric vehicles.

The future work in terms of the theoretical contributions are to explore additional formulations of objective functions based on measures in probability and information theory that select features that further improve the classification performance of the presented approach. A combination of selecting different wavelet bases and combining the resulting optimal cells per basis is being considered. Another direction capable from this work is that of applying multi-sensor fusion on these data to consider further improvements on the classification accuracy.

References

[1] L. Chiang, E. Russell, and R. Braatz, "Fault diagnosis in chemical processes using fisher discriminant analysis, discriminant partial least squares, and principal component analysis," *Chemometrics and Intelligent Laboratory Systems*, vol. 50, no. 2, pp. 243–252, Mar. 2000.

[2] L. Chiang, M. Kotanchek, and A. Kordon, "Fault diagnosis based on fisher discriminant analysis and support vector machines," *Computers and Chemical Engineering*, vol. 28, no. 8, pp. 1389–1401, July 2004.

[3] S. Roweis and L. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, Dec. 2000.

[4] C. Li, B. Kuo, and C. Lin, "Lda-based clustering algorithm and its application to an unsupervised feature extraction," *IEEE Transactions on Fuzzy Systems*, vol. 19, no. 1, pp. 152–163, Feb. 2011.

[5] Y. Huang, J. Zhao *et al.*, "Slow feature discriminant analysis and its application on handwritten digit recognition," in *International Joint Conference on Neural Networks, 2009*, June 2009, pp. 1294–1297.

[6] L. Wiskott and T. Sejnowski, "Slow feature analysis: unsupervised learning of invariances," *Neural Computation*, vol. 14, no. 4, pp. 715–770, Apr. 2002.

[7] Y. Mo, Z. Zhang *et al.*, "Random forest based coarse locating and kpca feature extraction for indoor positioning system," *Mathematical Problems in Engineering*, vol. 2014, no. 1, pp. 1–8, Oct. 2014.

[8] Y. Jiang and P. Liu, "Feature extraction for identification of drug and explosive concealed by body packing based on positive matrix factorization," *Measurement*, vol. 47, no. 1, pp. 193–199, Jan. 2014.

[9] J. Zabalza, J. Ren *et al.*, "Novel folded-pca for improved feature extraction and data reduction with hyperspectral imaging and sar in remote sensing," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 93, no. 1, pp. 112–122, July 2014.

[10] C. Li, H. Ho *et al.*, "A semi-supervised feature extraction based on supervised and fuzzy-based linear discriminant analysis for hyperspectral image classification," *Applied Mathematics and Information Sciences*, vol. 9, no. 1, pp. 81–87, Feb. 2015.

[11] X. Gu, C. Liu *et al.*, "Uncorrelated slow feature discriminant analysis using globality preserving projections for feature extraction," *Neuralcomputing*, May 2015.

[12] A. Godfrey, R. Conway, M. Leonard *et al.*, "A continuous wavelet transform and classification method for delirium motoric subtyping," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 17, no. 3, pp. 298–307, June 2009.

[13] G. Yu, C. Li, and S. Kamarthi, "Machine fault diagnosis using a cluster-based wavelet feature extraction and probabilistic neural networks," *The International Journal of Advanced Manufacturing Technology*, vol. 42, no. 1-2, pp. 145–151, May 2009.

[14] H. Eristi, A. Ucar, and Y. Demir, "Wavelet-based feature extraction and selection for classification of power system disturbances using support vector machines," *Electric Power Systems Research*, vol. 80, no. 1, pp. 743–752, July 2010.

[15] E. Avci, A. Sengur, and D. Hanbay, "An optimum feature extraction method for texture classification," *Expert Systems with Applications*, vol. 36, no. 3, pp. 6036–6043, Apr. 2009.

[16] N. Saravanan and K. Ramachandran, "Incipient gear box fault diagnosis using discrete wavelet transform (dwt) for feature extraction and classification using artificial neural network (ann)," *Expert Systems with Applications*, vol. 37, no. 6, pp. 4168–4181, June 2010.

[17] Y. Dong and J. Ma, "Wavelet-based image texture classification using local energy histograms," *IEEE Signal Processing Letters*, vol. 18, no. 4, pp. 247–250, Feb. 2011.

[18] B. Shankar, S. Meher, and A. Ghosh, "Wavelet-fuzzy hybridization: Feature-extraction and land-cover classification of remote sensing images," *Applied Soft Computing*, vol. 11, no. 3, pp. 2999–3011, Apr. 2011.

[19] H. Liu, C. Liu, and Y. Huang, "Adaptive feature extraction using sparse coding for machinery fault diagnosis," *Mechanical Systems and Signal Processing*, vol. 25, no. 2, pp. 558–574, Feb. 2011.

[20] L. Guo, D. Rivero *et al.*, "Automatic feature extraction using genetic programming: An application to epileptic eeg classification," *Expert Systems with Applications*, vol. 38, no. 8, pp. 10 425–10 436, Aug. 2011.

[21] J. Seshadrinath, B. Singh, and B. Panigrahi, "Incipient interturn fault diagnosis in induction machines using an analytic wavelet-based optimized bayesian inference," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 5, pp. 990–1001, May 2014.

[22] H. Bafroui, A. Ohadi *et al.*, "Application of wavelet energy and shannon entropy for feature extraction in gearbox fault detection under varying speed conditions," *Neurocomputing*, vol. 133, no. 1, pp. 437–445, June 2014.

[23] G. Xian and B. Zeng, "An intelligent fault diagnosis method based on wavelet packer analysis and hybrid support vector machines," *Expert Systems with Applications*, vol. 36, no. 10, pp. 12 131–12 136, Dec. 2009.

[24] L. Lin and J. Hongbing, "Signal feature extraction based on an improved emd method," *Measurement*, vol. 42, no. 5, pp. 796–803, June 2009.

[25] J. Huang, X. You *et al.*, "Rotation invariant iris feature extraction using gaussian markov random fields with non-separable wavelet," *Neurocomputing*, vol. 73, no. 4-6, pp. 883–894, Jan. 2010.

[26] R. Khushaba, S. Kodagoda *et al.*, "Driver drowsiness classification using fuzzy wavelet-packet-based feature-extraction algorithm," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 1, pp. 121–131, Sept. 2010.

[27] D. Wang, D. Miao, and C. Xie, "Best basis-based wavelet packet entropy feature extraction and hierarchical eeg classification for epileptic detection," *Expert Systems with Applications*, vol. 38, no. 11, pp. 14 314–14 320, Oct. 2011.

[28] G. Bin, J. Gao *et al.*, "Early fault diagnosis of rotating machinery based on wavelet packetsempirical mode decomposition feature extraction and neural network," *Mechanical Systems and Signal Processing*, vol. 27, no. 1, pp. 696–711, Feb. 2012.

[29] S. Li, Z. W. *et al.*, "Feature extraction and recognition of ictal eeg using emd and svm," *Computers in Biology and Medicine*, vol. 43, no. 7, pp. 807–816, Aug. 2013.

[30] G. Deepa and R. e. a. Keerthi, "Face recognition using spectrum-based feature extraction," *Applied Soft Computing*, vol. 12, no. 9, pp. 2913–2923, Sept. 2012.

[31] B. Chen, Z. Zhang *et al.*, "Fault feature extraction of gearbox by using overcomplete rational dilation discrete wavelet transform on signals measured from vibration sensors," *Mechanical Systems and Signal Processing*, vol. 33, no. 1, pp. 275–298, Nov. 2012.

[32] L. Han, C. Li *et al.*, "Feature extraction method of bearing ae signal based on improved fast-ica and wavelet packet energy," *Mechanical Systems and Signal Processing*, vol. 62-63, no. 1, pp. 91–99, Oct. 2015.

[33] H. Wang, J. Chen *et al.*, "Feature extraction of rolling bearings early weak fault based on eemd and tunable q-factor wavelet transform," *Mechanical Systems and Signal Processing*, vol. 48, no. 1-2, pp. 103–119, Oct. 2014.

[34] K. Poornima, A. Danti *et al.*, "Novel feature extraction for face recognition using multiscale principal component analysis," *International Journal of Computer Science and Engineering Communications*, vol. 2, no. 1, pp. 121–128, Feb. 2014.

[35] N. Lu, Z. Xiao *et al.*, "Feature extraction using adaptive multiwavelets and synthetic detection index for rotor fault diagnosis of rotating machinery," *Mechanical Systems and Signal Processing*, vol. 52-53, no. 1, pp. 393–415, Feb. 2015.

[36] J. Bennet, C. Ganaprakasam, and K. Arputharaj, "A discrete wavelet based feature extraction and hybrid classification technique for microarray data analysis," *The Scientific World Journal*, vol. 2014, no. 1, pp. 1–9, July 2014.

[37] Z. Wang, S. Bian *et al.*, "Feature extraction and classification of load dynamic characteristics based on lifting wavelet packet transform in power system load modeling," *International Journal of Electrical Power and Energy Systems*, vol. 62, no. 1, pp. 353–363, Nov. 2014.

[38] S. Zhang, Y. Zhang, and J. Zhu, "Rolling element-bearing feature extraction based on combined wavelets and quantum-behaved particle swarm optimization," *Journal of Mechanical Science and Technology*, vol. 29, no. 2, pp. 605–610, Feb. 2015.

[39] M. Yuwono, Y. Qin, J. Zhou, Y. Guo, B. Celler, and S. Su, "Automatic bearing fault diagnosis using particle swarm clustering and hidden markov model," *Engineering Applications of Artificial Intelligence*, p. To appear, Mar. 2015.

[40] R. Kryter and H. Haynes, "Condition monitoring of machinery using motor current signature analyis," in *Seventh Power Plant Dynamics, Control and Testing Symposium*, May 1989.

[41] R. Schoen and T. Habetler, "Effects of time-varying loads on rotor fault detection in induction machines," *IEEE Transactions on Industry Applications*, vol. 31, no. 4, pp. 900–906, Aug. 1995.

[42] R. Schoen, T. Habetler, F. Kamran, and R. Bartheld, "Motor bearing damage detection using stator current monitoring," *IEEE Transactions on Industry Applications*, vol. 31, no. 6, pp. 1274–1279, Dec. 1995.

[43] F. Filippetti, G. Franceschini, C. Tassoni, and P. Vas, "Recent developments of induction motor drives fault diagnosis using ai techniques," *IEEE Transactions on Industrial Electronics*, vol. 47, no. 5, pp. 994–1004, Oct. 2000.

[44] A. Siddique, G. Yadava, and B. Singh, "A review of stator fault monitoring techniques of induction motors," *IEEE Transactions on Energy Conversion*, vol. 20, no. 1, pp. 106–114, Mar. 2005.

[45] R. Tallam, S. Lee *et al.*, "A survey of methods for detection of stator-related faults in induction machines," *IEEE Transactions on Industry Application*, vol. 43, no. 4, pp. 920–933, Aug. 2007.

[46] A. Gandhi, T. Corrigan, and L. Parsa, "Recent advances in modeling and online detection of stator interturn faults in electrical motors," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 5, pp. 1564–1575, May 2011.

[47] A. Bonnett and C. Yung, "Increased efficiency versus increased reliability," *IEEE Industry Applications Magazine*, vol. 14, no. 1, pp. 29–36, Feb. 2008.

[48] A. Bonnett and G. Soukup, "Cause and analysis of stator and rotor failures in three-phase squirrel-cage induction motors," *IEEE Transactions on Industry Applications*, vol. 28, no. 4, pp. 921–937, Aug. 1992.

[49] M. Benbouzid, M. Vieira, and C. Theys, "Induction motors' faults detection and localization using stator current advanced signal processing techniques," *IEEE Transactions on Power Electronics*, vol. 14, no. 1, pp. 14–22, Jan. 1999.

[50] M. Benbouzid, H. Nejjari, R. Beguenane, and M. Vieira, "Induction motor asymmetrical faults detection using advanced signal processing techniques," *IEEE Transactions on Energy Conversion*, vol. 14, no. 2, pp. 147–152, Jun. 1999.

[51] M. Benbouzid, "A review of induction motors signature analysis as a medium for faults detection," *IEEE Transactions on Industrial Electronics*, vol. 47, no. 5, pp. 984–993, Oct. 2000.

[52] R. Schoen, B. Lin, T. Habetler, J. Schlag, and S. Farag, "An unsupervised, on-line system for induction motor fault detection using stator current monitoring," *IEEE Transactions on Industry Applications*, vol. 31, no. 6, pp. 1280–1286, Dec. 1995.

[53] R. Tallam, T. Habetler, and R. Harley, "Stator winding turn-fault detection for closed-loop induction motor drives," *IEEE Transactions on Industry Applications*, vol. 39, no. 3, pp. 720–724, May 2003.

[54] Y. Murphey, M. Masrur, Z. Chen, and B. Zhang, "Model-based fault diagnosis in electric drives using machine learning," *IEEE/ASME Transactions on Mechatronics*, vol. 11, no. 3, pp. 290–303, Jun. 2006.

[55] J. Martins, V. Pires, and A. Pires, "Unsupervised neural-network-based algorithm for an on-line diagnosis of three-phase induction motor stator fault," *IEEE Transactions on Industrial Electronics*, vol. 54, no. 1, pp. 259–264, Feb. 2007.

[56] L. Liu, K. Logan *et al.*, "Fault detection, diagnostics, and prognostics: Software agent solutions," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 4, pp. 1613–1622, July 2007.

[57] V. Tran, B. Yang, M. Oh, and A. Tan, "Fault diagnosis of induction motor based on decision trees and adaptive neuro-fuzzy inference," *Expert Systems with Applications*, vol. 36, no. 2, pp. 1840–1849, Mar. 2009.

[58] V. Ghate and S. Dudul, "Optimal mlp neural network classifier for fault detection of three phase induction motor," *Expert Systems with Applications*, vol. 37, no. 4, pp. 3468–3481, Apr. 2010.

[59] ——, "Cascade neural-network-based fault classifier for three-phase induction motor," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 5, pp. 1555–1563, May 2011.

[60] C. Kallesoe, R. Izadi-Zamanabadi *et al.*, "Observer-based estimation of stator-winding faults in delta-connected induction motors: A linear matrix inequality approach," *IEEE Transactions on Industry Applications*, vol. 43, no. 4, pp. 1022–1031, July 2007.

[61] B. Tabbache, M. Benbouzid *et al.*, "Dsp-based sensor fault detection and post fault-tolerant control of an induction motor-based electric vehicle," *International Journal of Vehicular Technology*, vol. 2012, no. 1, pp. 1–7, Nov. 2012.

[62] C. De Angelo, G. Bossio *et al.*, "Online model-based stator-fault detection and identification in induction motors," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 11, pp. 4671–4680, Nov. 2009.

[63] S. Cheng, P. Zhang, and T. Habetler, "An impedance identification approach to sensitive detection and location of stator turn-to-turn faults in a closed-loop multiple-motor drive," *IEEE Transactions on Industrial Electronics*, vol. 58, no. 5, pp. 1545–1554, May 2011.

[64] G. Georgoulas, M. Mustafa *et al.*, "Principal component analysis of the start-up transient and hidden markov modeling for broken rotor bar fault diagnosis in asynchronous machines," *Expert Systems with Applications*, vol. 40, no. 17, pp. 7024–7033, Dec. 2003.

[65] K. Gaeid, H. Ping *et al.*, "Survey of wavelet fault diagnosis and tolerant of induction machines with case study," *International Review of Electrical Engineering*, vol. 27, no. 3, pp. 4437–4456, June 2012.

[66] K. Logan, "Intelligent diagnostic requirements of future all-electric ship integrated power system," *IEEE Transactions on Industry Applications*, vol. 43, no. 1, pp. 139–149, Jan. 2007.

[67] X. Jin, S. Gupta *et al.*, "Wavelet-based feature extraction using probabilistic finite state automata for pattern classification," *Pattern Recognition*, vol. 44, no. 11, pp. 1343–1356, July 2011.

[68] O. Mohammed, Z. Liu *et al.*, "Internal short circuit fault diagnosis for pm machines using fe-based phase variable model and wavelets analysis," *IEEE Transactions on Magnetics*, vol. 43, no. 4, pp. 1729–1732, Apr. 2007.

[69] A. Ordaz-Moreno, R. Romero-Troncoso *et al.*, "Automatic online diagnosis algorithm for broken-bar detection on induction motors based on discrete wavelet transform for fpga implementation," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 5, pp. 2193–2202, May 2008.

[70] K. Siddiqui and V. Giri, "Broken rotor bar fault detection in induction motors using wavelet transform," in *2012 International Conference on Computing, Electronics and Electrical Technologies (ICCEET)*, Mar. 2012, pp. 1–6.

[71] J. Cusido, L. Romeral *et al.*, "Fault detection in induction machines using power spectral density in wavelet decomposition," *IEEE Transactions on Industrial Electronics*, vol. 55, no. 2, pp. 633–643, Feb. 2008.

[72] F. Li, G. Meng *et al.*, "Wavelet transform-based higher-order statistics for fault diagnosis in rolling element bearings," *Journal of Vibration and Control*, vol. 14, no. 11, pp. 1691–1709, Nov. 2008.

[73] S. Rajagopalan, J. Restrepo *et al.*, "Nonstationary motor fault detection using recent quadratic timefrequency representations," *IEEE Transactions on Industry Applications*, vol. 44, no. 3, pp. 735–744, June 2008.

[74] J. Rosero, L. Romeral *et al.*, "Short-circuit detection by means of empirical mode decomposition and wignerville distribution for pmsm running under dynamic condition," *IEEE Transactions on Industrial Electronics*, vol. 56, no. 11, pp. 4534–4547, Nov. 2009.

[75] A. Sadeghian, Y. Zhongming, and B. Wu, "Online detection of broken rotor bars in induction motors by wavelet packet decomposition and artificial neural networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 58, no. 7, pp. 2253–2263, Feb. 2009.

[76] P. Konar and P. Chattopadhyay, "Bearing fault detection of induction motor using wavelet and support vector machines (svms)," *Applied Soft Computing*, vol. 11, no. 6, pp. 4203–4211, Sept. 2011.

[77] H. Keskes, A. Braham, and Z. Lachiri, "Broken rotor bar diagnosis in induction machines through stationary wavelet packet transform and multiclass wavelet svm," *Electric Power Systems Research*, vol. 97, no. 1, pp. 151–157, Apr. 2013.

[78] J. Seshadrinath, B. Singh, and B. Panigrahi, "Investigation of vibration signatures for multiple fault diagnosis in variable frequency drives using complex wavelets," *IEEE Transactions on Power Electronics*, vol. 29, no. 2, pp. 936–945, Apr. 2013.

[79] ——, "Incipient turn fault detection and condition monitoring of induction machine using analytical wavelet transform," *IEEE Transactions on Industry Applications*, vol. 50, no. 3, pp. 2235–2242, Sept. 2013.

[80] A. Bazzi, A. Dominguez-Garcia, and P. Krein, "Markov reliability modeling for induction motor drives under field-oriented control," *IEEE Transactions on Power Electronics*, vol. 27, no. 2, pp. 534–546, Feb. 2012.

[81] A. Bazzi and P. Krein, "Utilization of median filters in power electronics: Traction drive applications," in *2013 28th Annual Conference on Applied Power Electronics Conference and Exposition (APEC)*, Mar. 2013, pp. 3055–3060.

[82] A. Silva, S. Gupta, and A. Bazzi, "Fault diagnosis in electric drives using machine learning approaches," in *IEEE International Electric Machines and Drives Conference (IEMDC)*, May 2013, pp. 722–726.

[83] "Case western reserve university bearing data center website," http://csegroups.case.edu/bearingdatacenter/pages/welcome-case-western-reserve-university-bearing-data-center-website, accessed: 2015-04-08.

[84] M. Zoni-Berisso, F. Lercari, T. Carazza *et al.*, "Epidemiology of atrial fibrillation: European perspective," *Clinical Epidemiology*, vol. 6, no. 1, pp. 213–220, June 2014.

[85] N. Thakor, "From holter monitors to automatic defibrillators: Developments in ambulatory arrhythmia monitoring," *IEEE Transactions of Biomedical Engineering*, vol. 31, no. 12, pp. 770–778, Dec. 1984.

[86] S. Artis, G. Moody, and R. Mark, "Algorithms for improved detection of supraventricular arrhythmias," in *Proceedings. Computers in Cardiology 1990*, Sep. 1990, pp. 507–510.

[87] S. Greenwald, R. Patil, and R. Mark, "Improved detection and classification of arrhythmias in noise-corrupted electrocardiograms using contextual information," in *Proceedings. Computers in Cardiology 1990*, Sep. 1990, pp. 461–464.

[88] N. Thakor, "Applications of adaptive filtering to ecg analysis: noise cancellation and arrhythmia detection," *IEEE Transactions of Biomedical Engineering*, vol. 38, no. 8, pp. 785–794, Aug. 1991.

[89] S. Chen, P. Clarkson, and Q. Fan, "A robust sequential detection algorithm for cardiac arrhythmia classification," *IEEE Transactions of Biomedical Engineering*, vol. 43, no. 11, pp. 1120–1124, Nov. 1996.

[90] L. Khadra, "Detection of life-threatening cardiac arrhythmias using the wavelet transformation," *Medical and Biological Engineering and Computing*, vol. 35, no. 6, pp. 626–632, Nov. 1997.

[91] K. Minami, H. Nakajima, and T. Toyoshima, "Real-time discrimination of ventricular tachyarrhythmia with fourier-transform neural network," *IEEE Transactions of Biomedical Engineering*, vol. 46, no. 2, pp. 179–185, Feb. 1999.

[92] R. Ceylan and Y. Ozbay, "Comparison of fcm, pca and wt techniques for classification ecg arrhythmias using artificial neural network," *Expert Systems with Applications*, vol. 33, no. 2, pp. 286–295, Aug. 2007.

[93] B. Weng, J. Wang *et al.*, "Atrial fibrillation detection using stationary wavelet transform analysis," in *30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2008. EMBS 2008.*, Aug. 2008, pp. 1128–1131.

[94] S. Saminu, N. Ozkurt, and I. Karaye, "Wavelet feature extraction for ecg beat classification," in *2014 IEEE 6th International Conference on Adaptive Science and Technology (ICAST)*, Oct. 2014, pp. 1–6.

[95] R. T. Olszewski, "Generalized feature extraction for structural pattern recognition in time-series data," Ph.D. dissertation, Carnegie Mellon University, 2001.

[96] H. Ringendahl and T. Sierla, "The freezing phenomenon in parkinson's disease," *Fortschr Neurol Psychiatr*, vol. 65, no. 10, pp. 435–445, Oct. 1997.

[97] N. Giladi, "Freezing of gait: Clinical overview," *Advances in Neurology*, vol. 87, pp. 191–197, 2001.

[98] J. Ghika, A. Wiegner *et al.*, "Portable system for quantifying motor abnormalities in parkinson's disease," *IEEE Transactions on Biomedical Engineering*, vol. 40, no. 3, pp. 276–283, Mar. 1993.

[99] O. Sofuwa, A. Nieuwboer *et al.*, "Quantitative gait analysis in parkinsons disease: Comparison with a healthy control group," *Archives of Physical Medicine and Rehabilitation*, vol. 86, no. 5, pp. 1007–1013, May 2005.

[100] A. Salarian, H. Russmann *et al.*, "Gait assessment in parkinson's disease: toward an ambulatory system for long-term monitoring," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 8, pp. 1434–1443, Aug. 2004.

[101] ——, "Quantification of tremor and bradykinesia in parkinson's disease using a novel ambulatory monitoring system," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 2, pp. 313–322, Feb. 2007.

[102] M. Thaut, G. McIntosh *et al.*, "Rhythmic auditory stimulation in gait training for parkinson's disease patients," *Movement Disorders*, vol. 11, no. 2, pp. 193–200, Mar. 1996.

[103] G. McIntosh, S. Brown *et al.*, "Rhythmic auditory-motor facilitation of gait patterns in patients with parkinson's disease," *J Neurol Neurosurg Psychiatry*, vol. 62, no. 1, pp. 22–26, Aug. 1996.

[104] J. Hausdorff, J. Lowenthal *et al.*, "Rhythmic auditory stimulation modulates gait variability in parkinson's disease," *The European Journal of Neuroscience*, vol. 26, no. 8, pp. 2369–2375, Oct. 2007.

[105] M. Bachlin, M. Plotnik, D. Roggen *et al.*, "Wearable assistant for parkinsons disease patients with the freezing of gait symptom," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 436–446, Nov. 2009.

[106] S. Mazilu, M. Hardegger, Z. Zhu *et al.*, "Online detection of freezing of gait with smartphones and machine learning techniques," in *2012 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*, May 2012, pp. 123–130.

[107] S. Mallat, *A wavelet tour of signal processing: the sparse way.* Burlington, MA: Academic Press, 2008.

[108] C. Bishop, *Pattern Recognition and Machine Learning.* New York, NY: Springer, 2006.

[109] N. Najjar, J. Hare, P. D'Orlando, G. Leaper *et al.*, "Heat exchanger fouling diagnosis for an aircraft air-conditioning system," in *SAE Technical Paper 2013-01-2250*, Sep. 2013, pp. 3055–3060.

[110] W. Conover, *Practical Nonparametric Statistics.* New York, NY: John Wiley & Sons, Inc., 1999.

[111] S. D. Greenwald, "Improved detection and classification of arrhythmias in noise-corrupted electrocardiograms using contextual information," Ph.D. dissertation, Massachusetts Institute of Technology, 1990.

[112] A. Goldberger, L. Amaral, L. Glass *et al.*, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. 215–220, May 2000.