CLASSIFICATION OF TIME-VARYING SIGNALS FROM AN ACOUSTIC MONITORING SYSTEM USING TIME-FREQUENCY TECHNIQUES

by

Samuel P. Ebenezer

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

ARIZONA STATE UNIVERSITY

December 2001

CLASSIFICATION OF TIME-VARYING SIGNALS FROM AN ACOUSTIC MONITORING SYSTEM USING TIME-FREQUENCY TECHNIQUES

by

Samuel P. Ebenezer

has been approved

November 2001

APPROVED:

_____, Chair

_____

_____

Supervisory Committee

ACCEPTED:

_____

Department Chair

_____

Dean, Graduate College

ABSTRACT

The problem of classification plays a major role in many of our day-to-day activities in this modern digital era. Classification problems exist in a wide range of engineering and surveillance applications, including fingerprint identification, voice recognition, mobile phones and sonar systems. The emergence of digital computers has made possible the practical implementation of automatic classification which, in turn, reduces many tedious and repetitive manual classification tasks. Even though the classification problem has been well studied over the last couple of decades, it is still under the microscopic eyes of researchers from all around the world. For example, the classification of time-varying signals from a sonar system is still an emerging research topic.

This thesis developed automatic methodologies to classify time-varying warning signals from an acoustic monitoring system that may indicate the potential catastrophic structural failures of reinforced concrete structures. Since missing a single warning signal may prove costly, it is imperative to develop a classifier with high probability of correctly classifying the warning signals. Due to the time-varying nature of these signals, various time-frequency (TF) classifiers published in the literature were used. A new matching pursuit decomposition (MPD) classifier was proposed for a real acoustic monitoring system. The superior performance of this new MPD classifier was investigated and compared with existing TF classifiers for various sets of acoustic sounds, including warning signals from real-world faulty structures. The performance of the MPD classifier under different test conditions was also studied, and a combined MPD and TF classifier was proposed to speed up the processing time of the algorithm. The flexibility of the MPD, such that it can be modified for use in other types of acoustic monitoring systems, was also considered. Finally, a graphical user interface in MATLAB was also developed for nonexperts to easily handle the algorithm.

To

my parents and sisters

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor Dr. Antonia Papandreou-Suppappola for her support, motivation, and encouragement during the entire course of my research work. Without her initiation and sustained guidance, it would have been difficult completing this project. I also want to thank her for kindling my independent thought process while approaching the research problem. I would also like to thank my industrial advisor Dr. Seth B. Suppappola for his invaluable suggestions and also for the amount of time he has spent in meaningful technical discussions and meetings. It has been a great experience interacting with him and gaining various technical insights into the research problem. I would also like to thank my committee members, Dr. Darryl R. Morrell and Dr. Konstantinos Tsakalis, for their constructive feedbacks and for reviewing this thesis report. I would also like to thank my Signal Processing Lab colleague DongWon Park and other colleagues for spending their valuable time in reading this thesis and for their suggestions. It has been a cherishable experience working with all of them. I would like to thank all the people in Electrical Engineering Department.

TABLE OF CONTENTS

Page

Page

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

CHAPTER 1

INTRODUCTION

1.1. Research Objective

1.1.1. Acoustic Monitoring System. Acoustic monitoring systems are used in many applications including undersea monitoring [1], structure monitoring and whales listening for unfriendly whistles [2]. The purpose of all these monitoring systems is to listen for any warning signals that may indicate future catastrophic effects. This premonition of future events will give us some invaluable time to prevent catastrophic failures that may follow. Figure 1 shows a typical monitoring system of structures along with the warning event [3], [4], [5], [6]. For example, in real life situations the warning event could be the breaking sound of metals in steel bridges or the abnormal knocking sound in aircraft and automobile engines or the breaking sound of metals in prestressed concrete structures with reinforcing metals. There are numerous reasons why it is desirable to ascertain the condition of a structure to determine if failure is imminent. The failure of a structure may result in the loss of its use which usually implies loss of revenue. In addition to this, the replacement cost after the failure is generally more than the preventive maintenance cost. As a result, the collateral cost due to the structural failure far exceeds the cost of the structure itself. This necessitates the monitoring of structures. For example, when a prestressed concrete

structure with reinforcing metals is under distress due to corrosion of the reinforcing metal, acoustic emissions may occur which indicate the presence of distress conditions. Thus, by monitoring these warning signals, we can prevent the failure of the structure by repairing it before failure occurs and save in replacement and other collateral costs.

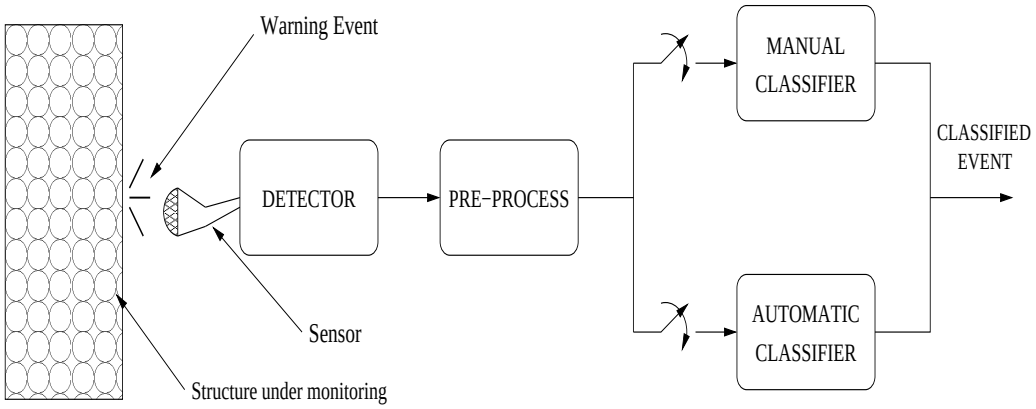

Figure 1. Typical acoustic monitoring system of structures.

1.1.2. Need for an Automatic Acoustic Classifier. Even though monitoring systems are intended to record some specific warning sounds, we cannot prevent the sensor to record all other undesirable acoustic events. For example, in a monitoring system of prestressed concrete structure with reinforcing metals, even though we are looking for the breaking noise of reinforcing metals, the monitoring system will also record other undesirable sounds like passing workers, birds, animals, machinery, automobiles or rain. Hence, it is important to classify the warning signals from all these other signals. This classification can be done manually by listening to all the sounds that are recorded by the monitoring system and using manual signal analysis combined with human judgment to identify the warning signals. However, in practice hardly one percent of recorded emissions will be of diagnostic value. Hence, manual classification becomes a tedious process. This necessitates the use of an automatic acoustic classifer [7], [8], [9].

Automatic classifiers are used in diverse fields ranging from digital communications [10], [11], [12] to surveillance applications [1]. Due to the diverse applications of classification theory, many advancements have been made over the last couple of decades through extensive research. Before the emergence of personal computers, implementation of some classification theories was not pragmatic due to intensive computational and memory requirements. However, widespread usage of computers enabled the implementation of many advanced automatic classifiers. Even though many classification theories have been developed, classification of time-varying signals whose frequency content varies with time is still a challenging task. The classification of sonar signals is one such example, as most of the recorded signals are time-varying in nature. In this thesis, we have developed an automatic classifier that classifies the time-varying acoustic emissions from a monitoring system using advanced techniques that are based on time-varying processing.

Due to the time-varying nature of the signals recorded by an acoustic monitoring system of a concrete structure, the performance of some well-established one dimensional (1-D) time domain classification methods may not be satisfactory [7], [8], [9]. The acoustic monitoring system demands that the probability of correct classification of warning signals be asymptotically equal to one. Since the main objective of an acoustic monitoring system is the premonition of catastrophic failures to be followed, missing some warning signals may prove to be very costly. Hence, we need to look for an algorithm whose probability of correct classification of warning signals is close to one. Note that we can afford to have a small level of false classifications provided that it does not increase the computational complexity of the manual analysis of classified signals.

In order to solve the classification problem in an acoustic monitoring system, we have developed some classification methods that are based on two-dimensional (2-D) trans-

formations of time and frequency, as our acoustic emissions are time-varying. Since a time-frequency representation (TFR) provides a signal's time localized frequency information, it is a powerful tool for analysing time-varying signals. As will be seen in the following chapters, the classification results of time-frequency (TF) based methods out perform the conventional 1-D time or frequency based methods. Most of our research work was concentrated on the direction of developing a classifier that is based on TF processing, and has higher performance when compared to other TF classifiers [13], [14], [15].

1.2. Thesis Organization

This thesis report is organized as follows. Chapter 2 reviews the fundamentals of TFRs and their superior quality in representing time-varying signals. In addition, we discuss a signal decomposition that is based on TF techniques and is known as matching pursuit decomposition (MPD).

In Chapter 3, we provide some background on some of the classical statistical based classifiers like the Baye's classifier and the matched filter that are based on 1-D time domain processing methods. We also discuss some existing TF based classifiers and compare their performance with conventional classifiers. The classification is performed using data samples that were obtained from a real acoustic monitoring system[1]. The TF classifiers include 2-D matched filter using the spectrogram and reassigned spectrogram. We also propose a new classifier based on the ambiguity function (AF) of the acoustic emissions.

In Chapter 4, we propose and develop a new classifier based on the MPD principle. We demonstrate the superior performance of this classifier by comparing it with the TF

[1]Our funding agency did not want to relieve its identity for commercial reasons.

classifiers discussed in Chapter 3. In this chapter, we also combine the TF classifiers with the MPD classifier to improve the efficiency of the computationally intensive MPD classifier.

In Chapter 5, we discuss different test scenarios that one may face, and the effects of using the MPD classifier with or without a TF based pre-classifier. We also discuss the graphical user interface (GUI) we developed for easy handling of the classifier algorithm by a signal processing novice. Chapter 6 concludes this thesis report with some final remarks on the different classifiers used and with future research plans.

CHAPTER 2

BACKGROUND

2.1. Time-Frequency Representations

A time domain signal can be represented in many different ways. The popular way of characterizing a signal is to represent it either in the time domain or in the frequency domain. The time domain representation of a signal provides the variations of the signal strength over time. The frequency domain representation of a signal gives the signal energy distribution across different frequencies. Oftentimes, the signals from one domain are transformed to another domain to obtain more information about the original signal that may not be easily accessible otherwise. For example, the frequency content of a signal is not often visible in the time domain. In order to obtain the frequency content of a signal, the time domain signal is transformed to the frequency domain using the Fourier transform [16], [17], [18]. However, the frequency domain representation of a signal does not provide us with the time-localized frequency content of the signal. Hence for signals whose frequency varies with time, the Fourier transform will not give us this time-varying frequency information. Similar to the Fourier transform, we can also represent the time domain signal in other domains to obtain the desirable information. The domain that represents the time-varying frequency information is called as the time-frequency domain [19], [20], [21], [22].

In order to analyse a time-varying signal effectively, the time and frequency domain characteristics should be considered jointly. These joint time-frequency representations (TFRs) characterize a given signal over the time-frequency (TF) plane by combining the time and frequency domain information to yield more revealing information about the temporal localization of a signal's spectral components. A TFR maps a 1-D time domain signal $x(t)$ into a 2-D function of time and frequency, $\mathrm{TFR}_x(t,f)$. Figures 2(a) and 2(b) show the time and frequency domain information respectively, of a linear FM chirp whose frequency varies linearly with time. As it can be seen from Figure 2(b), the frequency domain information of the linear FM chirp does not reveal the time-varying nature of the signal. Figure 2(c)[1] shows a TFR of the linear FM chirp indicating the variation of frequency over time. TFRs have been utilized to study a wide range of time-varying signals including speech, music, acoustic signals, biological signals, geophysical signals, radar and sonar signals.

2.1.1. Linear and Quadratic Time-Frequency Representations. An arbitrary TFR can be classified as linear or quadratic based on the dependence of $\mathrm{TFR}_x(t,f)$ to the signal $x(t)$ [19], [20], [21], [22].

Linear TFRs

All linear TFRs satisfy the superposition or linearity principle which states that if $x(t)$ is a linear combination of some signal components, then the TFR of $x(t)$ is the identically weighted linear combination of the TFRs of each of the signal components,

$$x(t) = ax_1(t) + bx_2(t) \Longrightarrow \mathrm{TFR}_x(t,f) = a\mathrm{TFR}_{x_1}(t,f) + b\mathrm{TFR}_{x_2}(t,f).$$

One example of a linear TFR is the short time Fourier transform (STFT) [23]. Even

[1] Note here the x-axis is time, the y-axis is frequency, and the TFR amplitude is represented by color.

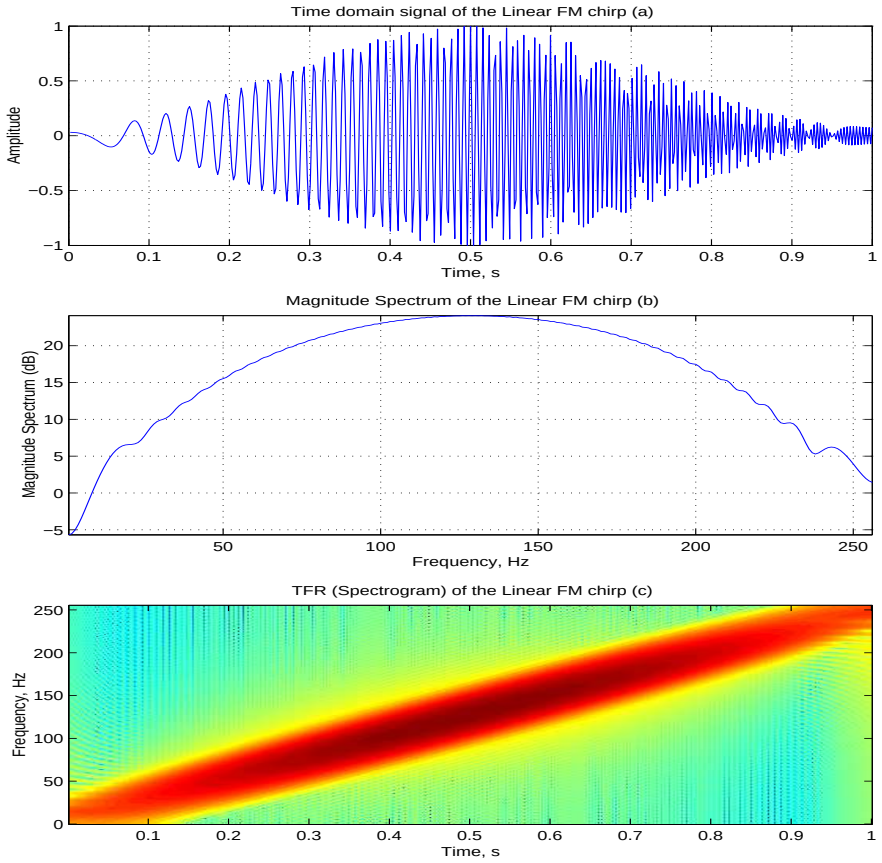

Figure 2. Time domain, frequency domain and time-frequency representations of a time-varying linear FM chirp. (a) Time domain representation, (b) Frequency domain representation, and (c) Time-frequency representations.

though the frequency domain representation in Figure 2(b) does not explicitly show the time localization of frequency components, such a time localization can be obtained by suitably applying a running window to pre-window the signal $x(t)$. The short-time spectrum of a signal $x(t)$ can be defined as

$$\mathrm{STFT}_x^{\gamma}(t_0, f) = \int \left[x(t)\gamma^*(t - t_0)\right] e^{-j2\pi f t} dt. \tag{2.1}$$

As shown in Figure 3, the STFT at time $t = t_0$ is the Fourier transform of the signal $x(t)$

multiplied by the window $\gamma(t)$ that is shifted by an amount equal to t_0.

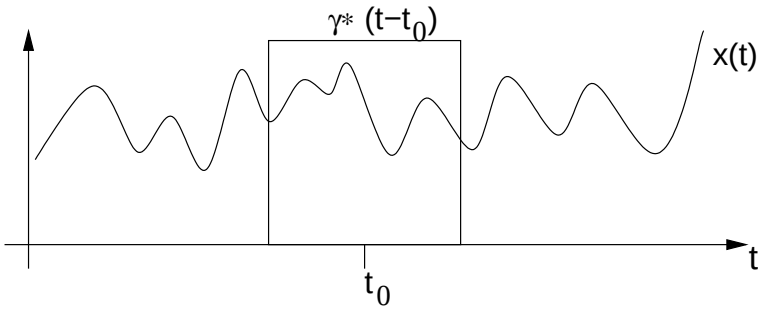

Figure 3. STFT interpretation: At time t_0, the STFT is the Fourier transform of the signal $x(t)$ multiplied by an analysis window $\gamma^*(t - t_0)$.

Quadratic TFRs

The quadratic structure of a TFR makes it possible to interpret a TFR as a TF energy distribution, since energy is also a quadratic signal representation. The linearity principle is violated by QTFR. However the QTFR satisfies the quadratic superposition principle given by,

$$x(t) = ax_1(t) + bx_2(t) \Longrightarrow$$

$$\mathrm{TFR}_x(t,f) = |a|^2\mathrm{TFR}_{x_1}(t,f) + |b|^2\mathrm{TFR}_{x_2}(t,f) + ab^*\mathrm{TFR}_{x_1,x_2}(t,f) + a^*b\mathrm{TFR}_{x_2,x_1}(t,f)$$

where $\mathrm{TFR}_{x_1}(t,f)$ is the auto TFR of the signal $x_1(t)$ and $\mathrm{TFR}_{x_1,x_2}(t,f)$ is the cross TFR of two signals $x_1(t)$ and $x_2(t)$. Thus the QTFR of a two component signal produces the auto TFR of the individual signal components called as auto terms and also the cross TFR called as cross terms that correspond to the cross component interference between the two signal components. Note that the number of interference terms grows quadratically with the number of signal components which often makes the visual analysis of the QTFR of a multicomponent signal difficult.

Many of the TFRs that are currently being used belong to the Cohen's class whose TFR member preserves the time and frequency shifts on a signal [19], [20], [21], [22]. Any Cohen's class TFR can be written as

$$\mathrm{TFR}_x(t,f) = \int\int \mathrm{AF}_x(\tau,\nu)\Phi_{TFR}(\tau,\nu)e^{j2\pi(t\nu-\tau f)}d\tau d\nu \tag{2.2}$$

where $\mathrm{A}_x(\tau,\nu)$ is called the ambiguity function (AF) of $x(t)$ and is defined as,

$$\mathrm{AF}_x(\tau,\nu) = \int x(t+\frac{\tau}{2})x^*(t-\frac{\tau}{2})e^{-j2\pi\nu t}dt.$$

Each QTFR in the Cohen's class can be interpreted as the 2-D Fourier transform of the weighted symmetric ambiguity function[2] of the analysis signal. The signal independent weighting function $\Phi_{TFR}(\tau,\nu)$ is called the kernel and it completely characterizes the TFR. Examples of some of the popular QTFRs are the Wigner distribution (WD) and the spectrogram. The WD [24], [25], [26] is defined as

$$\mathrm{WD}_x(t,f) = \int x(t+\frac{\tau}{2})x^*(t-\frac{\tau}{2})e^{-j2\pi f\tau}d\tau. \tag{2.3}$$

Equating Equations (2.2) and (2.3), we can derive the kernel of the WD which is given by

$$\Phi_{\mathrm{WD}}(\tau,\nu) = 1.$$

The square magnitude of the STFT in Equation (2.1) is called the spectrogram

$$\mathrm{SPEC}_x^{\gamma}(t,f) = |\mathrm{STFT}_x^{\gamma}(t,f)|^2.$$

The kernel function of the spectrogram [19], [22] is the AF of the analysis window $\gamma(t)$,

$$\Phi_{\mathrm{SPEC}}(\tau,\nu) = \int \gamma(t+\frac{\tau}{2})\gamma^*(t-\frac{\tau}{2})e^{-j2\pi\nu t}dt = \mathrm{AF}_{\gamma}(\tau,\nu).$$

[2]The name ambiguity function comes from the word 'ambiguous' which is widely used in RADAR signal processing. The ambiguity function is a 2-D function which is often used in RADAR and SONAR signal processing to jointly obtain the time correlation and frequency correlation of a signal.

As mentioned earlier, the spectrogram and the WD are the most widely used QTFRs. There are advantages and disadvantages when using either one of these QTFRs. The Wigner distribution has better time and frequency resolution in the TF plane when compared with the spectrogram. However, for signals with multiple components, the WD suffers from interference or cross terms that occur between any pair of these multiple components. Due to these inherent properties of the spectrogram and the WD, it is not advisable to use either one of them to analyse all kinds of signals. For example, Figure 4 shows the spectrogram and the WD of a complex sinusoidal signal at 25 Hz. As we can see from the plots, the frequency resolution of the spectrogram is poor when compared with the frequency resolution of the WD. Figure 5 shows the spectrogram and the WD of a multicomponent signal containing a complex sinusoid at 25 Hz and a non-linear FM chirp. As we can see from Figure 5 there are lot of cross terms in the WD. These noise like terms are due to interference between different frequency components of a multicomponent signals and also due to non-linear instantaneous frequency [19], [20], [21], [22]. However these visible cross terms are absent in the spectrogram. Hence, the WD is suitable only for analysing mono-component signals with linear instantaneous frequency, while for multicomponent signals, the spectrogram seems to be better when compared with the WD. As will be discussed in the following sections, the real analysis data to be classified are multicomponent signals.

2.2. Matching Pursuit Principle

As discussed in the previous section, the WD suffers from the presence of cross terms whereas the spectrogram suffers from poor time and frequency resolution. A good concentration of signal components and no misleading cross terms are necessary for an easy visual interpretation of the QTFR outcomes, as well as for accurate discrimination between

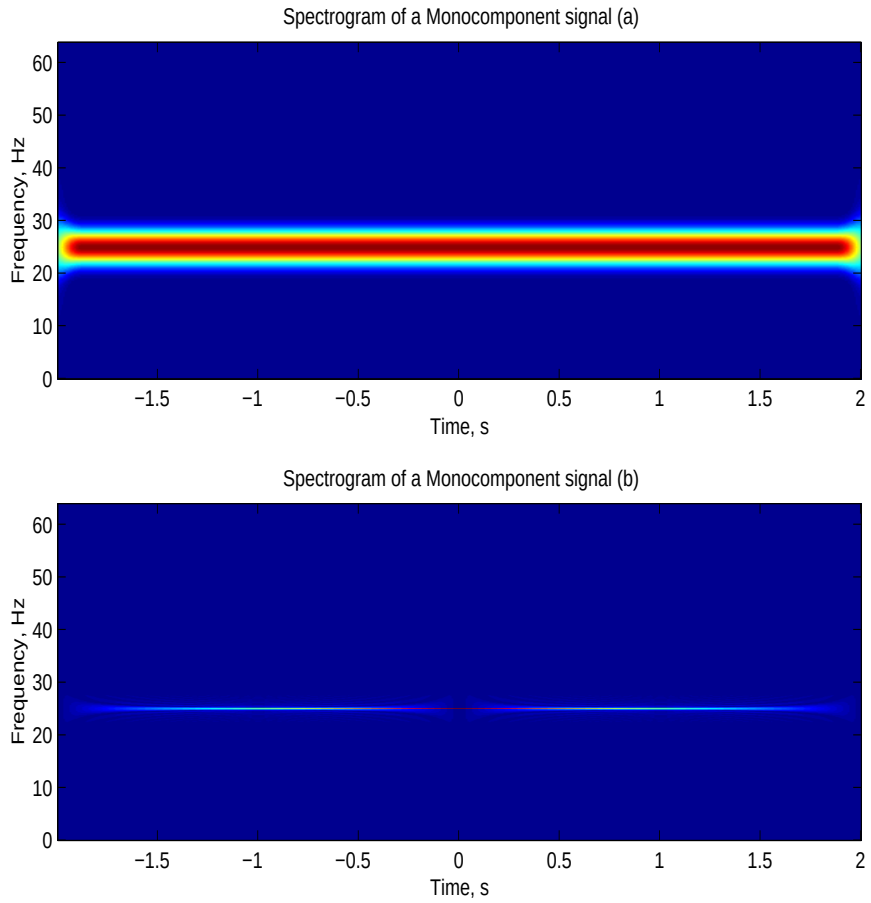

Figure 4. Comparison between the spectrogram and the Wigner distribution of mono-component signals: (a) Spectrogram, and (b) Wigner distribution of one sinusoid with frequency equal to 25Hz.

known QTFR patterns for non-stationary signal classification. Most of the recent research in time-frequency signal processing theory is in design of the kernel $\Phi(\tau, \nu)$ in Equation (2.2) such that the corresponding QTFR provides the best time and frequency resolution with reduced cross terms.

Even though many different QTFRs [27], [28], [29], [30] have been developed during the last couple of decades, most of them are optimum only for a particular class of signals. A QTFR that works perfectly for a particular class of signals may not yield the same result

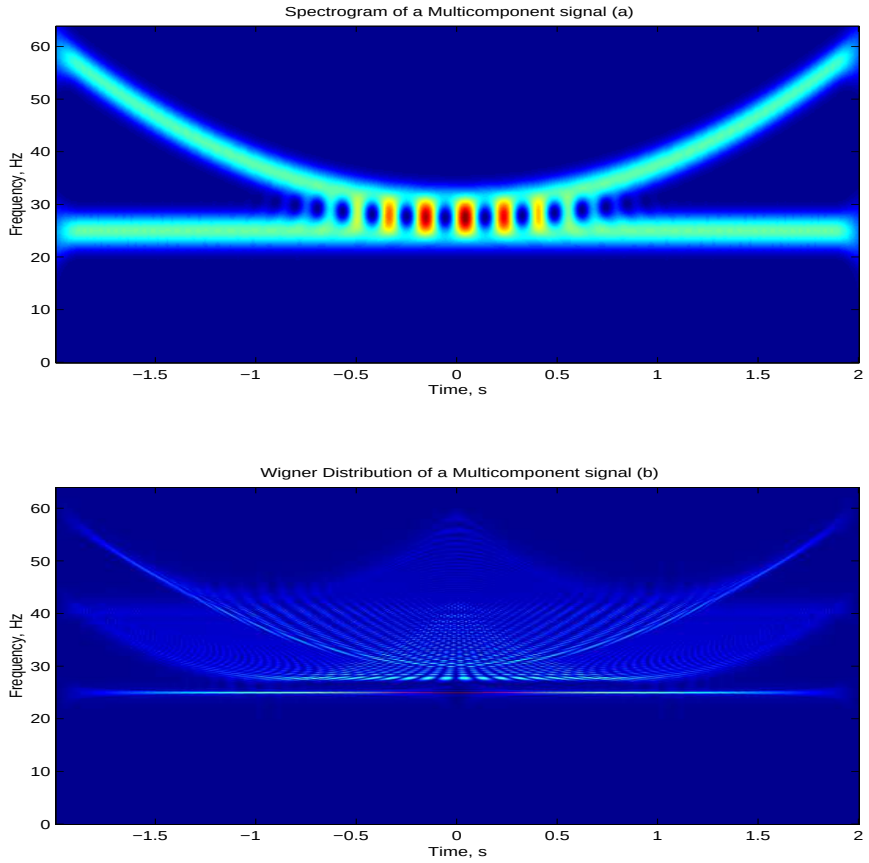

Figure 5. Comparison between the spectrogram and the Wigner distribution for multi-component signals: (a) Spectrogram, and (b) Wigner distribution of the superposition of a sinusoid at 25Hz and a non-linear FM chirp.

when used for a different class of signals. This is because different QTFRs satisfy different set of desirable signal properties. To avoid this problem, the authors in [31], [32] have developed a signal-dependent QTFR where the kernel in Equation (2.2) depends on the analysis signal. Specifically, the optimal kernel is obtained by satisfying certain constraints and performance measures such that the cross terms are reduced and the TF concentration is maximized.

The reassigned spectrogram was introduced in [33], [34] to improve the time and

frequency resolution of the spectrogram. The reassigned spectrogram is formed by moving each value of the spectrogram computed at any point (t, f) to a new, more appropriately localized point (t_{new}, f_{new}).

Other QTFRs was proposed in [35], [36], [37] where the analysis signal is first decomposed into elementary components and then the weighted sum of the Wigner distribution of the decomposed components form the signal representation. When this decomposition scheme is fitted to the analysed signal, a relevant description including fewer interference term can be obtained. Matching pursuit [35] is one such method that gives a TFR by signal decomposition. The matching pursuit decomposes any signal into a linear expansion of waveforms that belong to a redundant dictionary. When Gaussian TF atoms are used to form the dictionary, many of them are needed to decompose a signal with linear or non-linear TF characteristics. A signal's localization in time and frequency may vary widely and as a result, signal adaptive decomposition is necessary. The signal must be expanded into waveforms whose TF properties are adapted to its local structures. Thus it is important that the MPD dictionary is formed using TF atoms that are matched to the analysis signal [38], [39]. Although the MPD is an iterative non-linear algorithm, it preserves signal energy due to orthogonal expansions and thus guaranties convergence [35]. The dictionary of the MPD consists of a collection of TF atoms which are the dilated (time-scaling), translated (time-shifting) and modulated (frequency-shifting) versions of a single basic atom.

2.2.1. MPD Algorithm. Let $d(t)$ be the basic TF atom, also assume that α is the dilation or scaling factor, β is the translation factor, and κ is the modulation factor. Then, the redundant dictionary is formed by changing the factors α, β, κ. The dictionary elements $d(t; n)$ are formed as,

$$d(t;n) = \frac{1}{\sqrt{\alpha_n}}\, d\Big(\frac{t-\beta_n}{\alpha_n}\Big)\, e^{j2\pi\kappa_n t}, \quad n = 1, \ldots, N$$

where N is the total number of dictionary elements. Since all the dictionary elements are formed by dilation, translation and modulation of a single basic function, the MPD dictionary is redundant. The properties of the dictionary elements have been studied in [40]. To represent any signal $x(t)$ effectively, an appropriate subset of dictionary elements must be selected, so that $x(t)$ can be written as

$$x(t) = \sum_{k=0}^{M} a_k d(t; n(k)) \tag{2.4}$$

where $d(t;n(k))$ represents the k^{th} best matched dictionary element. We want to decompose the signal $x(t)$ as a linear expansion of signals selected from a dictionary that best matches the inner structure of the signal $x(t)$. This decomposition is done by successive approximation of $x(t)$ with orthogonal projections on dictionary elements. The signal $x(t)$ is first decomposed as[3]

$$x(t) = a^1 d^1_{bm}(t) + r^1(t), \quad a^1 = \langle\, x,\ d^1_{bm}\, \rangle$$

where $r^1(t)$ is the residual signal after approximating the signal $x(t)$ with the dictionary element $d^1_{bm}(t)$, a^1 is the projection of signal $x(t)$ onto the dictionary element $d^1_{bm}(t)$, and bm represents the best match dictionary element. Since $r^1(t)$ is orthogonal to $d^1_{bm}(t)$, we can write

$$||x(t)||^2 = |a^1|^2 + ||r^1(t)||^2.$$

To minimize $||r^1||$, the dictionary element $d^1_{bm}(t)$ must be chosen such that the parameter a^1 is maximum. The MPD is an iterative procedure that sub-decomposes the $(k-1)^{th}$ residual $r^{k-1}(t)$ by projecting it on a dictionary element that best matches it. This procedure

[3]Here, $\langle s,\ d^1_{bm}\rangle = \int s(t)\ d^{1*}_{bm}(t)dt$ and $||r^1||^2 = \langle r^1, r^1\rangle$.

is repeated each time on the subsequent residues that are obtained using the following equations.

$$r^{k-1}(t) = a^k d^k_{bm}(t) + r^k(t), \quad a^k = \langle\, r^{k-1},\ d^k_{bm}\,\rangle \tag{2.5}$$

where k is the iteration number and $d^k_{bm}(t)$ is the chosen dictionary element that best matches the residual $r^{k-1}(t)$ which is given by

$$r^k(t) = r^{k-1}(t) - a^{k-1} d^{k-1}_{bm}(t).$$

Since in Equation (2.5) the k^{th} order residue $r^k(t)$ is orthogonal to the best matched dictionary $d^k_{bm}(t)$ at the k^{th} iteration, we can write

$$||r^{k-1}(t)||^2 = |a^k|^2 + ||r^k(t)||^2. \tag{2.6}$$

From Equation (2.6), the energy conservation equation can be written as

$$||x(t)||^2 = \sum_{k=1}^{K} |a^k|^2 + ||r^K(t)||^2. \tag{2.7}$$

The energy conservation will be proved empirically in Section 4.4 by decomposing real data from an acoustic monitoring system using the MPD.

Using Equation (2.3), the WD of the decomposed signal $x(t)$ in Equation (2.4) can be written as

$$\begin{aligned}
\mathrm{WD}_x(t,f) &= \int x(t+\frac{\tau}{2}) x^*(t-\frac{\tau}{2}) e^{-j2\pi f\tau} d\tau \\
&= \int \sum_{k=1}^{K} a^k d^k_{bm}(t+\frac{\tau}{2}) \sum_{m=1}^{K} a^{m*} d^{m*}_{bm}(t-\frac{\tau}{2}) e^{-j2\pi f\tau} d\tau \\
&= \sum_{k=1}^{M} \sum_{m=1}^{K} \int a^k a^{m*} d^k_{bm}(t+\frac{\tau}{2})\ d^{m*}_{bm}(t-\frac{\tau}{2}) e^{-j2\pi f\tau} d\tau \\
&= \sum_{k=1}^{M} |a^k|^2 \int d^k_{bm}(t+\frac{\tau}{2})\ d^{k*}_{bm}(t-\frac{\tau}{2}) e^{-j2\pi f\tau} d\tau \ +
\end{aligned}$$

$$\sum_{k=1}^{K}\sum_{m=1,m\neq k}^{K} a^k a^{m*}\int d_{bm}^{k}(t+\frac{\tau}{2})\ d_{bm}^{m*}(t-\frac{\tau}{2})e^{-j2\pi f\tau}d\tau$$
$$=\sum_{k=1}^{K}|a^k|^2\mathrm{WD}_{d_{bm}^{k}}(t,f)+\sum_{k=1}^{K}\sum_{m=1,m\neq k}^{K} a^k a^{m*}\mathrm{WD}_{d_{bm}^{k}d_{bm}^{m}}(t,f).$$

The double summation in the above equation corresponds to the cross terms of the WD. Thus keeping only the first summation we obtain a modified, cross-term free QTFR, $\mathrm{MWD}_x(t,f)$, given by

$$\mathrm{MWD}_x(t,f)=\sum_{k=1}^{K}|a^k|^2\mathrm{WD}_{d_{bm}^{k}}(t,f). \tag{2.8}$$

Note that the MPD is not an approximation of the original signal $x(t)$, but a complete decomposition of the original signal. This complete decomposition of the original signal $x(t)$ is guaranteed only when the dictionary is complete [35], [40]. In fact, the authors in [35] have proved that the dictionary formed by time scaling, time shifting, and frequency shifting a basic window function is complete. The dictionary is complete when the residual $r^k(t)$ is zero as k tends to infinity, that is

$$\lim_{k\to\infty} r^k(t)=0$$

Thus the adaptive QTFR using the MPD in Equation (2.8) provides a representation with reduced cross terms unlike the WD of the entire signal $x(t)$. However, the trade off is that the MWD does not preserve all of the desirable property that the WD satisfies. Mallet and Zhang in [35] have developed and illustrated this adaptive TFR using a Gaussian basic TF atom,

$$d(t)=2^{1/4}e^{-\pi t^2}.$$

Thus, the redundant dictionary consists of time-scaled, time-shifted and frequency-shifted versions of the Gaussian atom

$$d(t;n) = \frac{2^{1/4}}{\sqrt{\alpha_n}} e^{-\pi\left(\frac{t-\beta_n}{\alpha_n}\right)^2} e^{j2\pi\kappa_n t}, \quad n = 1, \ldots, N$$

where N is the total number of transformations. Recall that α_n, β_n, and κ_n are the time-scaling, time-shift, frequency-shift parameters respectively at the n^{th} transform. Figure 6 shows TF locations of some dictionary elements formed using Gaussian atoms. The

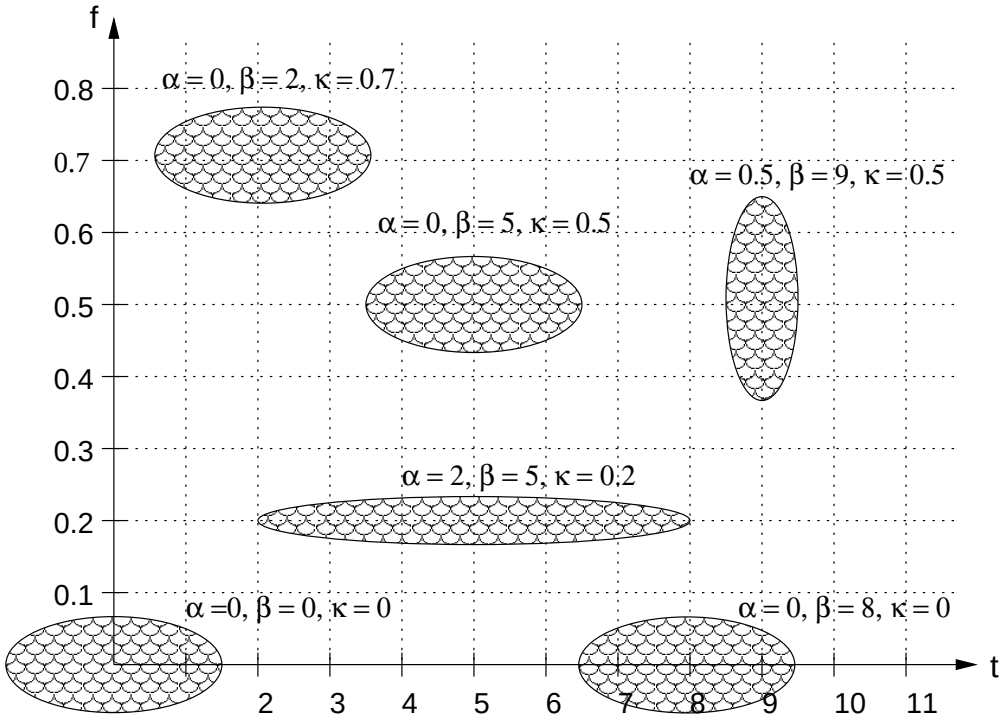

Figure 6. A basic Gaussian atom at the origin and its time-scaled by α, time-shifted by β, frequency-shifted by κ versions in the time-frequency plane.

redundant dictionary covers the entire TF plane, spanned by these time-scaled, time-shifted, and frequency-shifted Gaussian atoms.

CHAPTER 3

1-D and 2-D CLASSIFIERS

3.1. Classical Classifiers

This section explains some of the existing classical classifiers and the next section explains TF based classifiers. Signal classification is important in many process supervisory systems, diagnostic systems and fault detection systems. For example, in an undersea acoustic monitoring system, diverse kinds of signals will be heard by the hydrophones of the monitoring system. It is important for a monitoring system to correctly identify the signals which indicate the presence of an enemy submarine, from other undesirable events. Similarly, in an quadrature phase shift keying (QPSK) communication receiver [11], [12], a transmitter may send any one of the four possible, known (at the receiver) modulated signals. The main objective of the classifier/receiver is to correctly identify the transmitted signal from the four possible transmitted signals. Therefore, signal classification is a must in many monitoring and communication systems.

An intuitively appealing approach for pattern recognition is the template matching technique where in a set of templates or prototypes (one for every possible class) is required. The input unknown signal is compared with the templates of each class, and the classification is based on the preselected similarity criterion. If the input unknown signal

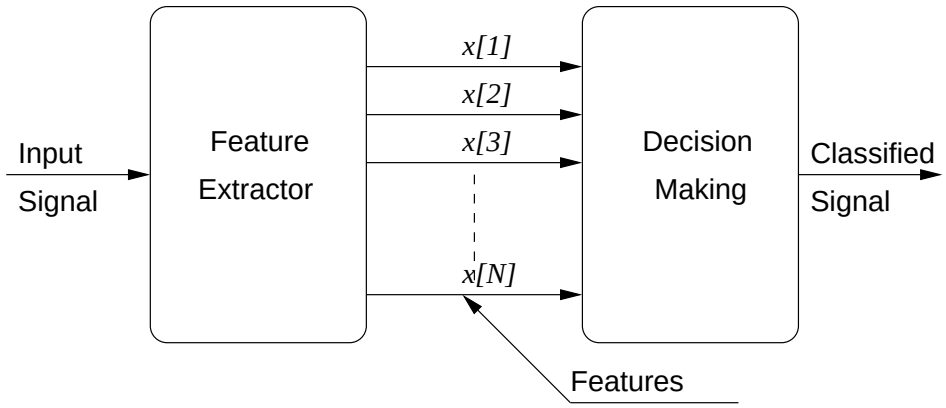

Figure 7. A typical feature based signal classifier. The first block extracts the feature set from the input signal. The classification decision is made by the second block using the extracted feature set and an appropriate similarity criteria.

matches the template of the j^{th} class better than it matches any other templates, then the input signal is classified in Class j. However, the disadvantage of the template matching approach is that it is sometimes difficult to select a good template for each class and to define a proper matching criterion. This disadvantage is especially remarkable when large variations and distortions are expected in all signals that belong to one class.

Another pattern recognition methodology is based on mathematical techniques to solve the signal classification problem [7], [8], [9]. The statistical approach is one such technique which is popular in many applications. In the statistical approach, the classification is based on a set of selected measurements called features that are extracted from the signal. These features are supposed to be less sensitive to possible signal distortions and variations compared with the other possible set of signal characteristics. The classification of a signal involves two steps. The first step is feature extraction. The criterion of feature selection is often based on the importance of the feature in characterizing the signal or the contribution of the feature to the accuracy of the signal classification. The second step is the decision making based on the extracted features and a meaningful criterion. The block diagram

in Figure 7 figuratively explains this kind of classifier. $x[1], x[2], x[3] \ldots x[N]$ in Figure 7 represents the features that are extracted from the input signal that needs to be classified. Next, we will describe some of the classical classifiers in more detail.

3.1.1. Bayes Classifier. Suppose we choose N features $x[1], x[2], x[3] \ldots x[N]$ and arrange as a vector called feature vector $\mathbf{X}$,

$$\mathbf{X} = \Big[x[1] \quad x[2] \quad x[3] \quad \ldots \quad x[N] \Big].$$

with dimensionality N. To describe the classification technique, consider that a feature $x[i], i = 1, \ldots, N$ is a random variable and that different vector features can represent different classes $C_j, j = 1, \ldots, L$. Assume that the conditional probability density function (PDF) of a feature vector[1] denoted as $p(\mathbf{X}/C_j), j = 1, \ldots L$, and the probability of occurrence of Class j are both known. On the basis of the apriori information $p(\mathbf{X}/C_j)$ and $P(C_j), j = 1, \ldots L$, the classifier performs the classification task by minimizing the probability of misclassification. Considering $L = 2$, the problem of classification can now be formulated as a statistical decision problem [7], [8], [41], [42]. Defining $\mathbf{X}$ as the feature vector of the unknown input signal $x(t)$, the decision problem can be stated as two hypothesis :

$$\mathrm{H}_1 : \mathbf{X} \text{ belongs to } C_1$$

$$\mathrm{H}_2 : \mathbf{X} \text{ belongs to } C_2.$$

If the unknown signal $x(t)$ actually belongs to Class 1, then the decision rule must be chosen such that the probability of selecting hypothesis H_1 given the feature vector $\mathbf{X}$, is greater

[1] Here, the conditional PDF refers to the PDF when the unknown signal belongs to Class j.

than the probability of selecting hypothesis H_2 given the feature vector $\mathbf{X}$, (i.e)

$$P(\mathrm{H}_1/\mathbf{X}) > P(\mathrm{H}_2/\mathbf{X}) \quad for\ (\mathbf{X}\ \epsilon\ C_1). \tag{3.1}$$

Using Bayes theorem [43], we can write,

$$P(\mathrm{H}_j/\mathbf{X}) = \frac{p(\mathbf{X}/\mathrm{H}_j)P(\mathrm{H}_j)}{p(\mathbf{X})}. \tag{3.2}$$

Substituting Equation (3.2) in Equation (3.1) we get,

$$\frac{p(\mathbf{X}/\mathrm{H}_1)}{P(\mathrm{H}_2)} > \frac{p(\mathbf{X}/\mathrm{H}_2)}{P(\mathrm{H}_1)}$$

$$l(\mathbf{X}) = \frac{p(\mathbf{X}/\mathrm{H}_1)}{p(\mathbf{X}/\mathrm{H}_2)} > \frac{P(\mathrm{H}_2)}{P(\mathrm{H}_1)}. \tag{3.3}$$

where $l(\mathbf{X})$ is called as the likelihood ratio [7], [8], [41], [42]. Hence, given the conditional density function and the probability of occurrence of different classes, the decision strategy in Equation (3.3) minimizes the probability of misclassification by maximizing the likelihood ratio.

For example, let us assume that the feature vector is a Gaussian random vector. The mean and the variance of this random vector is given by

$$\boldsymbol{\mu_j} = \Big[\mu_j[1]\ \ \mu_j[2]\ \ \mu_j[3] \ldots \mu_j[N]\Big],$$

where $\mu_j[k]$ and $\sigma_j^2[k]$ are the mean and variance of the k^{th} random variable $x[k]$ of Class j. Then the joint PDF of the feature vector $\mathbf{X}$ of Class j is

$$p(\mathbf{X}/C_j) = \frac{1}{\sqrt{(2\pi)^N |\mathrm{COV}_j|}}\ e^{\frac{-1}{2}\Big[(\mathbf{X}-\boldsymbol{\mu_j})\mathrm{COV}_j^{-1}(\mathbf{X}-\boldsymbol{\mu_j})^{\mathrm{T}}\Big]} \tag{3.4}$$

where COV_j is the covariance matrix of the feature vector, $|\mathrm{COV}_j|$ and COV_j^{-1} are the determinant and inverse of the covariance matrix of Class j respectively, and symbol T defines the transpose of a matrix. If we assume that the features $x[1], x[2], x[3] \ldots x[N]$ are

uncorrelated random variables with equal mean μ_j and variance σ_j^2, then the covariance matrix of the random vector $\mathbf{X}$ is the diagonal matrix [43],

$$\mathrm{COV}_j = \begin{pmatrix} \sigma_j^2 & 0 & \dots \\ 0 & \sigma_j^2 & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}_{N\mathrm{X}N} = \sigma_j^2 \, \mathrm{I}_{N\mathrm{X}N}. \tag{3.5}$$

Using Equation (3.5) in Equation (3.4), the joint PDF of an uncorrelated Gaussian random feature vector can be written as

$$p(\mathbf{X}/C_j) = \frac{1}{\sqrt{(2\pi)^N \sigma_j^{2N}}} \, e^{\frac{-1}{2\sigma_j^2}\sum_{k=1}^{N}(x[k]-\mu_j)^2}, \quad j = 1, 2. \tag{3.6}$$

If we further assume that $\mu_j = 0$ for both classes in Equation (3.6), the log likelihood ratio which is the log of the likelihood ratio in Equation (3.3), for the zero mean white Gaussian random vector can be written as,

$$\ln\Big(l(\mathbf{X})\Big) = \frac{1}{2}\ln\Big(\frac{\sigma_2^2}{\sigma_1^2}\Big) + \frac{1}{2}\Big(\frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2}\Big)\mathbf{X}\mathbf{X}^{\mathrm{T}}.$$

Using Equation (3.3) Class 1 will be selected if [41], [42],

$$\mathbf{X}\mathbf{X}^{\mathrm{T}} > \frac{2\sigma_1^2\sigma_2^2}{\sigma_1^2 - \sigma_2^2}\left[\ln\Big(\frac{P(\mathrm{H}_2)}{P(\mathrm{H}_1)}\Big) + \frac{1}{2}\ln\Big(\frac{\sigma_1^2}{\sigma_2^2}\Big)\right].$$

The above decision rule, derived for binary classification ($L = 2$), can be extended to L-class signal classification [41], [42].

3.1.2. Matched Filter. Let us consider the binary decision problem,

$$\mathrm{H}_1 : \mathbf{X} = \mathbf{S}_1 + \mathbf{W}, \ \mathbf{X} \text{ belongs to } C_1$$

$$\mathrm{H}_2 : \mathbf{X} = \mathbf{S}_2 + \mathbf{W}, \ \mathbf{X} \text{ belongs to } C_2$$

where $\mathbf{S}_j$ is a known deterministic vector that provides the characteristic feature set of each class. This deterministic vector will be different for different classes. $\mathbf{W}$ is a zero mean white Gaussian noise vector with covariance $\boldsymbol{\sigma^2 I}$. Thus, $\mathbf{X}$ is also white Gaussian with mean $\mathbf{S}_j$ and covariance $\boldsymbol{\sigma^2} I$ under the hypothesis H_j, $j = 1, 2$. Then, following Equation (3.6), the PDF of $\mathbf{X}$ given H_j can be written as,

$$p(\mathbf{X}/\mathrm{H}_j) = \frac{1}{\sqrt{(2\pi)^N \sigma^{2N}}}\, e^{\dfrac{-1}{2\sigma^2}\displaystyle\sum_{k=0}^{N}(x[k]-s_j[k])^2}, \quad j = 1, 2.$$

The log likelihood ratio can be expressed as

$$\ln\Big(l(\mathbf{X})\Big) = \frac{1}{2\sigma^2}\mathbf{X}(\mathbf{S}_1 - \mathbf{S}_2)^{\mathrm{T}} + \frac{1}{2\sigma^2}\Big(\mathbf{S}_2\mathbf{S}_2^{\mathrm{T}} - \mathbf{S}_1\mathbf{S}_1^{\mathrm{T}}\Big).$$

If we assume that the deterministic signals $\mathbf{S}_1$ and $\mathbf{S}_2$ have equal energy then the log likelihood ratio further simplifies to

$$\ln\Big(l(\mathbf{X})\Big) = \mathbf{X}(\mathbf{S}_1 - \mathbf{S}_2)^{\mathrm{T}} \quad and \quad \mathbf{S}_1\mathbf{S}_1^{\mathrm{T}} = \mathbf{S}_2\mathbf{S}_2^{\mathrm{T}}.$$

Then using the decision rule in Equation (3.3), the test signal will be selected in favor of Class 1 if,

$$\mathbf{X}(\mathbf{S}_1 - \mathbf{S}_2)^{\mathrm{T}} > \ln\Big(\tfrac{P(\mathrm{H}_2)}{P(\mathrm{H}_1)}\Big).$$

If we assume that the Class 1 and Class 2 signals occur with equal probability, then the decision rule becomes,

$$decide \quad H_1 \quad if \quad \mathbf{X}\mathbf{S}_1^{\mathrm{T}} > \mathbf{X}\mathbf{S}_2^{\mathrm{T}}. \tag{3.7}$$

As a result, an unknown test signal $x(t)$ will be classified in Class 1 if the inner product of the feature vector of the test signal and the known deterministic vector that characterizes Class 1 is greater than the corresponding inner product pertaining to Class 2. Since this decision rule involves matching the feature vector of an unknown test signal with the characteristic

feature vector of different possible classes, this operation is called as *matched filtering* [41], [42]. Note that the matched filter minimizes the probability of misclassification under the assumptions that

- $\mathbf{S}_1$ and $\mathbf{S}_2$ are known deterministic feature vectors;
- Additive white Gaussian noise (AWGN) with zero mean is used;
- The energy of the characteristic feature vector is the same for all classes; and
- All possible classes occur with equal probability.

The above derived decision rule is for binary classification. The matched filter developed for binary classification can be extended to L-class signal classification. This scheme is shown in the simplified block diagram in Figure 8. As will be seen in the following section, some of the classifiers used for this thesis are based on this statistical based classification method.

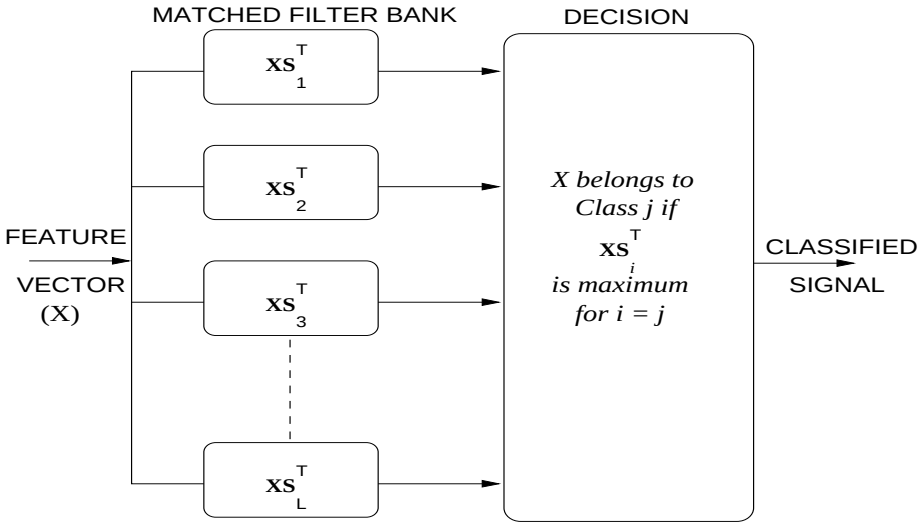

Figure 8. L-class signal classifier based on the matched filter technique. The test signal will be classified in Class j when the inner product between the template vector S_j of Class j and the feature vector of the test signal is maximum

frequency techniques. There exist many classification methods depending on the nature of the signals as we have seen in Section 3.1. In this section we will discuss about the TF based classification methods to classify acoustic events. The effectiveness of these classification methods are investigated using different acoustic signals from a real acoustic monitoring system. Figure 9 shows the spectrogram [19] of a learning signal one from each class out of eight possible classes. As seen from the figure, only Class 1 and Class 3 have high frequency components, whereas all other classes consist of low frequency signals. These kind of signals are chosen because in an acoustic monitoring system, low frequency signals like machinery, human voice or automobile noises will be quite often recorded. For this example, the empirical data indicates that Class 1 is the warning signal which usually precedes a catastrophic failure, and thus must be classified accurately.

3.2.1. TF Classifier: Spectrogram and Reassigned Spectrogram. In Section 3.1.2 we have seen that, for binary classification, if $\mathbf{S}_1$ and $\mathbf{S}_2$ are the characteristics feature vectors of two classes embedded in additive white Gaussian noise, then the matched filter operation minimizes the probability of misclassification. Any heuristic approach to the choice of feature can be used, provided it leads to a minimal misclassification rate. A common feature that we can consider is the time samples of the signal itself. Hence, the template vectors $\mathbf{X}_1$ and $\mathbf{X}_2$ of Class 1 and Class 2 respectively containing additive white Gaussian noise can be defined as

$$\overline{\mathbf{X}}_j = \Big[\overline{x}_j[1] \;\; \overline{x}_j[2] \;\; \ldots \;\; \overline{x}_j[N]\Big], \quad j = 1, 2$$

where

$$\overline{x}_j[n] = \frac{1}{K_j}\sum_{k=1}^{K_j} x_j^k[n]$$

for $n = 1, \ldots N$, denotes the sample mean of the K_j number of learning signals of Class j at discrete time n. Based on the decision rule defined in Equation (3.7), Section 3.1.2, for matched filter, a test signal $\mathbf{X}$ will be classified in Class 1 if,

$$\langle\ \mathbf{X}, \overline{\mathbf{X}}_1\ \rangle > \langle\ \mathbf{X}, \overline{\mathbf{X}}_2\ \rangle. \tag{3.8}$$

If the signals are non-stationary and the initial phase of the signal is random, then the matched filter gives poor result. In this case, other features that reveal the time-varying information may be used for classification. A standard way to represent time-varying signals is using TFRs. The authors in [44] suggested that we can use TF points as a feature to classify non-stationary signals. In order to obtain specific TF points to form a feature vector, we use the discrete time and discrete frequency versions of a TFR. Consider for example the spectrogram that was discussed in Section 2.1.1. The discrete time equivalent of spectrogram can be written as,

$$\mathbf{SPEC}(n\mathrm{T}, k\mathrm{F}) = \left| \sum_{l=0}^{N-1} x(l) h^*(l-n) e^{\frac{-j2\pi kl}{N}} \right|^2.$$

Due to the summation term involved in calculating the spectrogram, we can assume that each TF point in the TF plane is Gaussian based on the central limit theorem [43]. So using the TF points of the spectrogram as a feature vector, the authors in [44] extended the matched filter decision strategy to the TF domain. Specifically, for binary classification, one can deduce that the test signal belongs to Class 1 if[2]

$$\langle\langle\ \mathbf{SPEC}_x, \overline{\mathbf{SPEC}}_1\ \rangle\rangle > \langle\langle \mathbf{SPEC}_x, \overline{\mathbf{SPEC}}_2 \rangle\rangle$$

where $\mathbf{SPEC}_x$ is the spectrogram of a test signal $x(t)$, and $\overline{\mathbf{SPEC}}_j$ is the average spectrogram of Class j. The average spectrogram of Class j is obtained by averaging the spectrogram of all the learning signals in Class j, that is,

[2]Here, $\langle\langle \mathbf{SPEC}_x, \overline{\mathbf{SPEC}}_j \rangle\rangle$ defines the 2-D inner product $\int\int \mathbf{SPEC}_x(t,f)\ \overline{\mathbf{SPEC}}_j(t,f)\ dt\ df$.

$$\overline{\mathbf{SPEC}}_j = \frac{1}{K_j}\sum_{k=1}^{K_j} \mathbf{SPEC}_{x_j^k}, \quad j = 1, 2$$

where $\mathbf{SPEC}_{x_j^k}(t, f)$ is the spectrogram of the k^{th} learning signal in Class j and K_j is the number of learning signals in Class j. Specifically, this rule corresponds to the 2-D matched filter [41]. The inner product between the spectrogram of the test signal and the average spectrogram of the learning signals of Class 1 and Class 2 are used as the test statistic. The test signal will be assigned to the class whose 2-D correlation with the test signal is maximum. For L-class signal classification, the classification rule will be

$$x(t) \; \epsilon \; C_j \iff j = \arg\max_l \; (\; \langle\langle \mathbf{SPEC}_x, \overline{\mathbf{SPEC}}_l \rangle\rangle \;) \tag{3.9}$$

for $l = 1,\ldots,L$, where C_j is the j^{th} class.

Note that the 2-D inner product of **SPEC** and $\overline{\mathbf{SPEC}}_j$ in Equation (3.9) is not equivalent to the 1-D inner product of $\mathbf{X}$ and $\overline{\mathbf{X}}_j$ in Equation (3.8) [19], [20], [21], [22]. Instead of using TF points of spectrogram as the feature vector, we can also use the TF points of other TFRs. Since most of the acoustic events recorded by the monitoring system are multi-component, we decided not to use the popular Wigner distribution because, for a multi-component test or learning signal the WD suffers from oscillatory cross terms [19], [20], [21], [22]. There exist other TFRs that do not suffer from cross terms that correspond to smoothed versions of the WD. A highly localized reassigned spectrogram [33] [34] was also chosen for our testing to investigate whether TF localization is important in this TF classification. We have implemented this spectrogram based classifier to our data consisting of eight different acoustic signals. Each spectrogram in Figure 9 represents a learning signal from each class, and the TF variation as expected is different for different classes.

3.2.2. TF Classifier: Ambiguity Function. Recall that the ambiguity function [19] mentioned in Section 2.1.1 is defined as,

$$\mathrm{AF}_x(\tau,\nu) = e^{j\pi\tau\nu}\int x(t)\ x^*(t-\tau)e^{-j2\pi\nu t}dt. \tag{3.10}$$

The narrowband ambiguity function (AF) is the 2-D Fourier transform of the Wigner distribution. In the AF domain, all the auto terms of a multi-component signal $x(t)$ will be concentrated around the origin $(\tau,\nu)=(0,0)$, and all the cross terms will be away from the origin [19]. The definition of the auto terms and the cross terms can be found in Section 2.1.1. The AF of the different signals to be classified in our data are remarkably different. For example, Figure 10 shows the AF of a Class 1 signal and a Class 3 signal. The Class 3 signal is an 8 KHz deterministic signal that is used to check whether the detector is properly detecting the acoustic signals. The AF of each signal corresponds to an auto term, but the origin concentration differs due to the properties of each signal. Since Class 3 is a deterministic signal, its time samples are highly correlated, so we can expect an appreciable amount of correlation for nonzero time lags. Moreover, since the Class 3 signal contains most of its energy at 8 KHz there will be less frequency correlation for nonzero frequency lags. Thus it is clear from the figure that the AF of the third class signal has peaks all along the τ axis with the maximum peak at the origin whereas the AF of the first class signal has peaks only around $(0,0)$. As the magnitude of the AF called as ambiguity surface is different for different classes[3] we defined a test statistic as [13]

$$\gamma_j = \langle\langle\ |\mathbf{AF}_x(\tau,\nu)|,\ |\overline{\mathbf{AF}}_j(\tau,\nu)|\ \rangle\rangle, \quad j=1,\ldots,L$$

where

$$|\overline{\mathbf{AF}}_j(\tau,\nu)| = \tfrac{1}{K_j}\sum_{k=1}^{K_j}\ |\mathbf{AF}_{x_j^k}(\tau,\nu)|\ .$$

[3]Most of the times only the magnitude of the AF is used [19].

3.2.3. TF Classifer: Cross Ambiguity Function. Another classification method we proposed is based on the cross AF. The cross AF of the signals $x(t)$ and $y(t)$ is defined as

$$\mathrm{AF}_{xy}(\tau,\nu) = e^{j\pi\tau\nu} \int x(t)\ y^*(t-\tau)e^{-j2\pi\nu t}dt \quad . \tag{3.12}$$

When $x(t) = y(t)$, the cross AF is the auto AF in Equation (3.10). Using this basic argument and the one in Section 3.2.2, we can infer that, when $x(t)$ and $y(t)$ are highly correlated, then it is not unreasonable to expect a maximum value close to $(\tau,\nu) = (0,0)$ in the AF plane. This is because it behaves as an auto term where the maximum value of the auto AF occurs at the origin [19], and the high correlation shows that $x(t)$ and $y(t)$ are very similar.

We have made use of the above simple reasoning based on similarity measure, for signal classification. Let $x(t)$ be a test signal, and define

$$\overline{x}_j(t) = \frac{1}{K_j}\sum_{k=1}^{K_j} \quad x_j^k(t), \quad j = 1,\ldots,L \tag{3.13}$$

where $\overline{x}_j(t)$ is the time averaged learning signal of Class j, $x_j^k(t)$ denotes the k^{th} learning signal from Class j, and K_j is the number of learning signals in Class j. We define our test statistic as

$$\gamma_j = |\mathrm{AF}_{x\overline{y}_j}(0,0)|, \quad j = 1,\ldots,L \tag{3.14}$$

where $\overline{y}_j(t)$ is a time shifted version of the averaged signal $\overline{x}_j(t)$ from Class j defined in Equation (3.13). The test signal may be the time shifted version of any one of the learning signals. The time shift of the averaged prototype signal ensures that a maximum value at $\tau = 0$ in the AF plane is attained. This maximum value corresponds to the maximum possible time correlation between the test signal and the prototype signals of different classes. Note that if we substitute $(\tau,\nu) = (0,0)$, Equation (3.12) reduces to

$$\mathrm{AF}_{xy}(0,0) = \int x(t)y^*(t)\ dt.$$

Thus, Equation (3.14) reduces to the 1-D time correlation [10]. The test statistic in Equation (3.14) will be reduced to

$$\gamma_j = \max \ \big[\int x(t)\ \overline{x}_j^*(t-\tau_1)\ dt \big], \quad \text{for all possible time lags } \tau_1.$$

The test statistic in above equation is the simple time correlation between the test signal and the time averaged learning signals of all possible classes. The classification rule can then be defined as

$$x(t) \ \epsilon \ C_j \iff j = \arg\max_l \ (\ \gamma_l\ ), \quad l = 1, \ldots, L. \tag{3.15}$$

The test statistic in this case is similar to the test statistic obtained using the 1-D MF in Equation (3.8) with the template vector being the time shifted version of the time averaged learning signals. Since the test statistic in this method reduces to a simple time correlation, we will be referring this method as 1-D time correlation in later sections.

3.3. Classification Results

The classification strategies in Equations (3.9), (3.11) and (3.15) are tested using the real data obtained from an acoustic monitoring system. All the test and learning signals used were sampled at 20 KHz, and the duration of each data vector is 0.1024 seconds (2048 data samples). Eight classes of signals were used to compare the different test strategies. Figure 9 shows the spectrogram of a characteristic sample signal from each of the eight different types of signals. Table 1 shows the number of learning signals and the test signals used for each class. We have configured Class 1 to be the major class of interest and the other classes to be the undesirable events. For example, in an acoustic monitoring system, the desirable event could be a precursor warning signal that precedes catastrophic failures, and the undesirable events could be passing workers, birds, animals, machinery, automobiles

Class Number	Number of Learning Signals	Number of Test Signals
Class 1	50	342
Class 2	13	32
Class 3	10	30
Class 4	9	30
Class 5	11	14
Class 6	11	31
Class 7	13	37
Class 8	13	30

Table 1. Number of learning and test signals for each class.

or rain. As we are interested in the classification of the warning signals, we have used more number of learning signals from Class 1. By having more number of learning signals the probability of correctly classifying a Class 1 signal of varying TF structure increases. Similarly, we have used large number of test signals of Class 1, so that we can make a fair conclusion about the performance of a classifier (we cannot comment on the performance of a classifier by classifying only very few Class 1 test signal).

Table 2 provides the number of misclassified test signals for each class using different classification methods. It also shows the number of false classification (FC) events that corresponds to the number of test signals from Class 2 to Class 8 that were classified in Class 1. Table 3 provides the probability of correct classification (P_{CC}) of the test signals from each class using different methods. The methods used are (a) 1-D time-correlation obtained from cross AF (CROSS-AF) using Equation (3.15), (b) 2-D matched filtering using the spectrogram (SPEC) using Equation (3.9), (c) 2-D matched filtering using reassigned spectrogram (RSPEC) using Equation (3.9) and (d) the decision rule based on the ambiguity function (AF) in Equation (3.11).

Considering the classification results for Class 1, the performance of the 1-D time-correlation (CROSS-AF) in Table 3 is poor when compared to other methods. The poor

Class Number	CROSS AF	SPEC	RSPEC	AF
Class 1	30	13	14	13
Class 2	1	4	3	2
Class 3	0	0	0	0
Class 4	0	0	0	1
Class 5	3	0	0	3
Class 6	1	6	6	7
Class 7	9	1	1	5
Class 8	4	0	0	1
Number of FC	0	0	1	28

Table 2. Rows 2-9 provide the number of misclassified signals in each class for different methods. For example, three Class 5 signals were classified in other classes when using the CROSS AF. Row 10 provides the number of false classification of other acoustic events that were classified in Class 1.

performance of the 1-D time-correlation method can be attributed to the fact that the time-varying nature of the signals was not taken into consideration, and also due to the randomness of the initial phase of the signal. The 2-D matched filter using both the spectrogram and the reassigned spectrogram perform better than the 1-D time-correlation as shown in Table 3 demonstrating the effectiveness of the TF based methods. Even though the spectrogram and the reassigned spectrogram provide the same classification performance, the processing time using the reassigned spectrogram exceeds the corresponding one for the spectrogram by upto 10 times due to the complexity involved in calculating the reassigned spectrogram. The AF based method also works well in classifying Class 1 signals as shown in Table 3. However, the number of false classified signals in Table 2 when using the AF exceeds the corresponding number for other methods. Due to this large number of false classifications we can say that the AF based method appears to be biased towards Class 1 signals.

Even though no stringent Pcc is required for Class 2 to Class 8 signals, the TFR based classifiers classify these class of signals with acceptable performance. As the performance of

Class Number	CROSS AF	SPEC	RSPEC	AF
Class 1	0.91	0.96	0.96	0.96
Class 2	0.97	0.88	0.91	0.94
Class 3	1	1	1	1
Class 4	1	1	1	1
Class 5	0.79	1	1	0.79
Class 6	0.71	0.81	0.81	0.77
Class 7	0.89	0.97	0.97	0.87
Class 8	0.97	1	1	0.97

Table 3. Probability of correct classification in each class.

the 1-D time correlation method is poor when compared with the 2-D TF based methods, we expect that the performance of the 1-D MF will be worst, as it does not consider time shifting of the acoustic signals. Even though the performance of the 2-D matched filter in Table 3 is better than the performance of the 1-D time-correlation based method, the number of misclassification of Class 1 signals is still not acceptable for an acoustic monitoring system. Since the main objective of the acoustic monitoring system is the premonition of the structural failures, it may not be advisable to use a classifier that misses between 13 and 15 out of 342 test signals of Class 1. This necessitates the design of other classification methods with a P_{CC} of Class 1 signals closer to one and with a reduced number of false classification. In this thesis, we proposed and designed such a method as will be demonstrated in the next chapter.

CHAPTER 4

MATCHING PURSUIT BASED CLASSIFIER

4.1. Need for a New Classifier

Even though there exist well-defined mathematical techniques to solve classification problems, most of them are used for classifying stationary signals. For example, the classical 1-D classification techniques like the matched filter are optimum[1] when the signals to be classified are known and deterministic and are embedded in additive white Gaussian noise (AWGN). However, for non-stationary signals like the acoustic events in our application having random initial phase, the 1-D matched filter will not yield optimum results as we demonstrated in Chapter 3.

The TF classification techniques that were discussed in Chapter 3 are better matched to the time-varying nature of the signals. As a result, they performed better than the classical 1-D methods. However, their results need further improvements since the classification of acoustic signals from an acoustic monitoring system requires a very high probability of correct classification of the warning signals. Since the main objective of an acoustic monitoring system is the premonition of catastrophic failures to be followed, missing some warning signals may prove to be very costly. Hence, we need to design a new classification

[1] By optimum techniques we refer to the techniques that minimize the probability of misclassification.

algorithm that will provide us with the desirable P_{CC} of warning signals. The following sections describe our iterative classification algorithm that provides a misclassification rate of 1% for our real data.

4.2. TF Decomposition Based Classifier: Modified Matching Pursuit

The matching pursuit decomposition (MPD) principle that can be used to generate a localized TFR of a signal, was discussed in Section 2.2. The authors in [35] proposed this method to iteratively decompose a signal into time-scaled, time-shifted and frequency-shifted versions of a basic Gaussian atom selected from a redundant dictionary. However, when the signal to be decomposed has multiple components with different TF structures, the MPD uses many dictionary elements to decompose it, and thus many iterations are needed. The authors in [36] introduced the use of rotated Gaussian atoms with fewer waveforms as dictionary elements to analyse linear, frequency-modulated chirps more efficiently.

The authors in [38], [39] modified the MPD by using dictionary elements that are matched to the TF structure of a signal. These dictionary elements include Gaussian signals as well as signals that may have linear or non-linear phase functions such as logarithmic or power. The advantage of using a dictionary that is matched to the analysis data is that only a small number of elements will be used to decompose a signal, and hence less number of iterations will be needed.

We proposed to use the modified MPD for classifying time-varying signals. Our main objective of using the MPD is to obtain some distinctive feature set for each class during the iterative procedure of the signal decomposition algorithm and use these parameters to classify time-varying signals. For our MPD approach, we need to decompose and classify acoustic signals. As a result, we must match the dictionary elements with the acoustic

signals for a successful decomposition. Instead of using TF shifted and scaled Gaussian atoms as dictionary elements, we use TF shifted versions of the learning signals from each class to form our dictionary, as shown in Figure 11. However we must make sure that the dictionary formed by TF shifting the learning signals is complete. The authors in [45] used a wave-based dictionary consisting of wavefronts, resonances, and linear FM chirps to process scattering data. They formed the dictionary by time-shifting small number of wavefronts, resonances and linear FM chirps. They have also proved the completeness of this dictionary in [45]. The same proof applies to our dictionary also, which is basically the TF shifted versions of learning signals.

4.2.1. Design of the MPD Based Classifier. In order to design the MPD classifier, we need to select appropriate learning signals similar to other classification methods considered in Chapter 3. We start forming our dictionary by TF shifting the learning signals of each class. Looking at the representative spectrogram of the different class in Figure 9, the range of frequency shifts is limited. Since a higher range of frequency shift will increase the misclassification rate. After this initial step, the following steps demonstrate how to classify any[2] given test signal with the MPD technique.

Step 1: The test signal $x[n]$ is first correlated with all the dictionary elements.

Step 2: As our objective is to find the best matched dictionary element for that specific signal, we choose the dictionary element whose cross correlation[3] with the test signal is maximum.

Step 3: The chosen dictionary element is then subtracted from the test signal, and the

[2]Note that the test signal as well as the learning signals must be from the same acoustic system for correct classification.

[3]In order to reduce the complexity of our MPD classifier, we incorporate the time shift of the dictionary elements in the correlation computation described in Step 1.

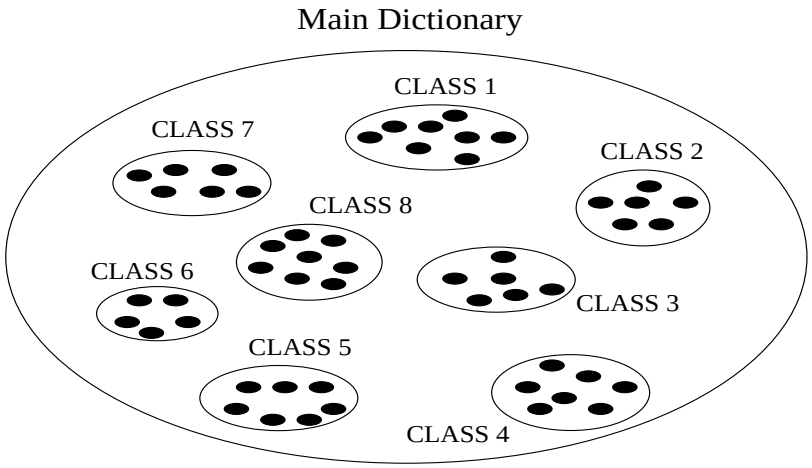

Figure 11. The dictionary elements of the modified MPD are TF shifted versions of learning signals of a specific class. Note that the learning signals of each class are represented by black ovals.

class number of the chosen dictionary element is noted. The residual signal is given by

$$x^1[n] = x^0[n] - a_j^1 d_j^1[n]$$

where $x^0[n]$ is the original test signal, $x^1[n]$ is the residual signal after the first iteration, $d_j^1[n]$ is the dictionary element that best matches the test signal $x^0[n]$ and it belongs to Class j, and a_j^1 quantifies the similarity between $x^0[n]$ and $d_j^1[n]$.

Step 4: As the MPD is iterative, we continue to decompose the residual signal from the previous iteration. At the k^{th} iteration,

$$x^k[n] = x^{k-1}[n] - a_j^k d_j^k[n], \quad k = 1, \ldots, K$$

where, $x^k[n]$ is the residual signal at the k^{th} iteration, $d_j^k[n]$ is the best matched dictionary element at the k^{th} iteration and it belongs to Class j. The correlation coefficient a_j^k at the k^{th} iteration is given by

$$a_j^k = \langle\ x^{k-1},\ d_j^k\ \rangle.$$

Step 4 is repeated for all $k = 1, \ldots, K$. The algorithm convergence criteria could be the value of the correlation coefficient or a prescribed number of iterations. Specifically when the correlation coefficient is very small compared to the threshold, then further iteration will only result in adding an error to the decomposition. A similar concept is true for a large number of iterations. In our study, we set the correlation coefficient threshold value at 0.15, and we used a maximum number of ten iterations ($K = 10$).

Step 5: When the algorithm converges, the net contribution of the correlation coefficients from each class are used as the test statistic. Specifically, the test statistic for Class j is given by

$$\gamma_j = \sum_{m=1}^{M} \mid a_j^{k_m} \mid \tag{4.1}$$

where k_m is the iteration number in which Class j was chosen during the decomposition, and M is the number of times that the Class j was chosen. Finally, the unknown test signal will be classified in Class j based on

$$x(t) \ \epsilon \ \mathrm{C}_j \iff j = \arg\max_l \ (\ \gamma_l \ ), \quad l = 1, \ldots, L. \tag{4.2}$$

As the residual signal $x^k[n]$ is uncorrelated with the best matched dictionary element $d_j^k[n]$, the MPD is an orthogonal decomposition which guaranties energy conservation. Step 5 is explained with an example that uses $L = 8$ different classes. For a given test signal, if the MPD yields the following correlation coefficients after 8 iterations, $a_1^1 = 0.9$, $a_1^2 = 0.8$, $a_1^3 = 0.7$, $a_2^4 = 0.6$, $a_3^5 = 0.5$, $a_4^6 = 0.4$, $a_5^7 = 0.3$, $a_6^8 = 0.2$, then the test statistics for each class will be $\gamma_1 = 0.9 + 0.8 + 0.7 = 2.4$, $\gamma_2 = 0.6$, $\gamma_3 = 0.5$, $\gamma_4 = 0.4$, $\gamma_5 = 0.3$, $\gamma_6 = 0.2$, $\gamma_7 = 0$, $\gamma_8 = 0$. Recall that a_1^3 is the correlation coefficient from the third iteration and the best matched dictionary element at the third iteration belongs to Class 1, and γ_1 is the test statistic for Class 1. Based on the classification rule in Equation (4.2), this test signal will

Iteration Number (k)	Residual Energy	Magnitude of Correlation Coefficient($\|a_j^k\|$)	Best Matched Class Number(j)
1	0.9192	0.2840	1
2	0.8591	0.2340	8
3	0.8089	0.2240	1
4	0.7619	0.2	8
5	0.7242	0.1941	1
6	0.6954	0.1695	1
7	0.6615	0.1840	1
8	0.6368	0.1522	8
9	0.6046	0.1540	8
10	0.5849	0.1401	1

Table 4. Matching pursuit parameters of a test signal of Class 1.

be classified in Class 1. Figure 12 shows the basic flow diagram of this classifier algorithm.

4.3. Real Data Illustration

The same set of learning and test signals that are used to evaluate the performance of the TF based classifiers will be used to evaluate the performance of the MPD based classifier. This set of learning and test signals for the eight possible classes are shown in Table 1. For illustration purposes let us investigate the different MPD parameters obtained for two specific classes of signals, namely Class 1 and Class 6. Tables 4 and 5 show the MPD parameters obtained for real data from Class 1 and Class 6, respectively. The spectrograms of these two test signals are shown in Figure 13.

The absolute values of the correlation coefficients for a Class 1 test signal after 10 iterations are

$$|a_1^1| = 0.2840,\ |a_8^2| = 0.2340,\ |a_1^3| = 0.2240,\ |a_8^4| = 0.2,\ |a_1^5| = 0.1941,$$

$$|a_1^6| = 0.1695,\ |a_1^7| = 0.1840,\ |a_8^8| = 0.1522,\ |a_8^9| = 0.1540,\ |a_1^{10}| = 0.1401.$$

The test statistics are computed using Equation (4.1) and are given by

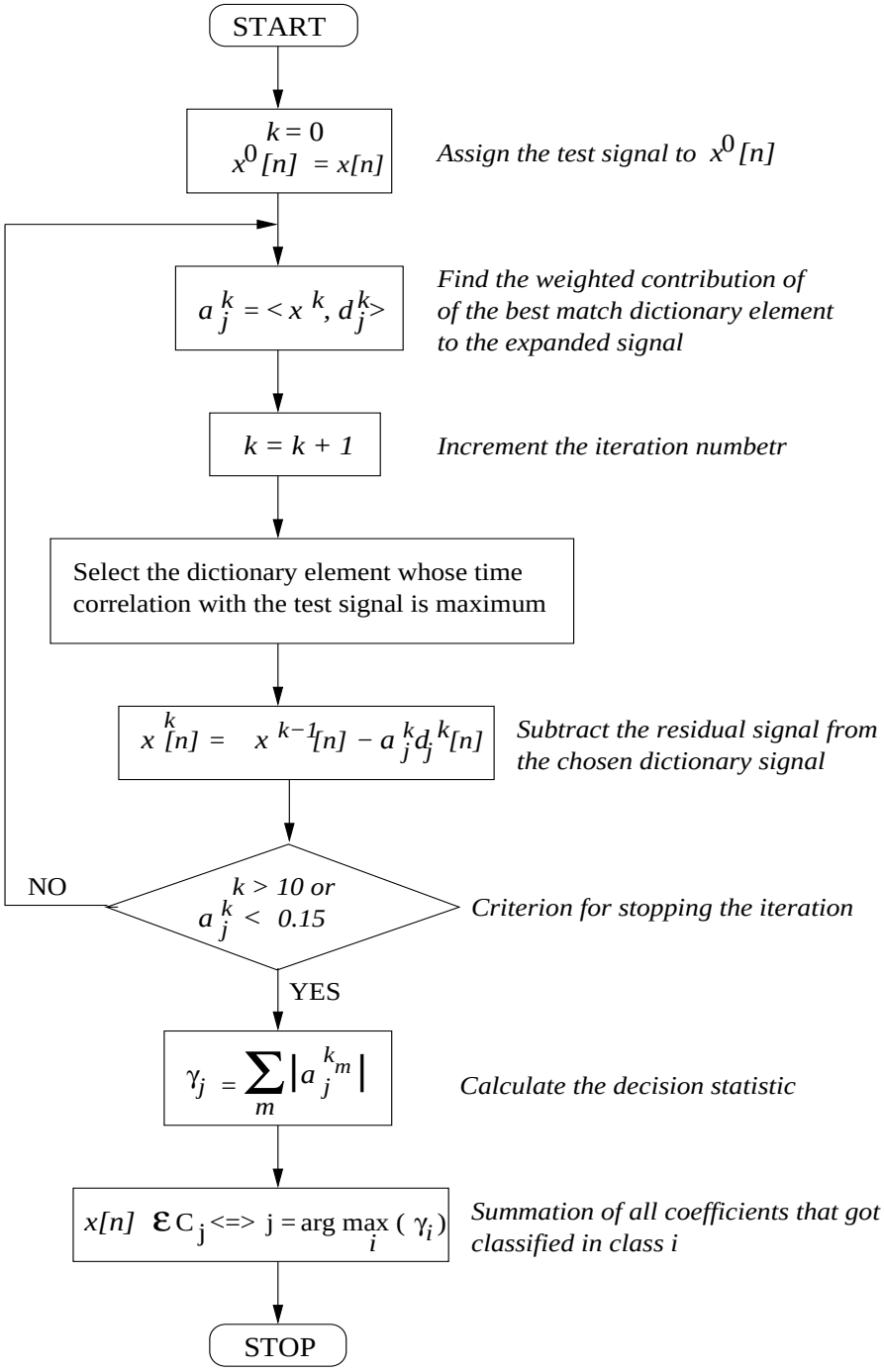

Figure 12. Implementation of the modified MPD algorithm that is explained in detail in Section 4.2.1.

Iteration Number (k)	**Residual Energy**	**Magnitude of Correlation Coefficient**$(|a_j^k|)$	**Best Matched Class Number**(j)
1	0.8425	0.3968	4
2	0.7415	0.2867	1
3	0.6861	0.2354	7
4	0.6299	0.2370	6
5	0.5612	0.2321	7
6	0.5175	0.1920	8
7	0.4890	0.1690	6
8	0.4618	0.1647	2
9	0.4372	0.1570	6
10	0.4140	0.1521	4

Table 5. Matching pursuit parameters of a test signal of Class 6.

$$\gamma_1 = 1.1921,\ \gamma_2 = 0,\ \gamma_3 = 0,\gamma_4 = 0,\ \gamma_5 = 0,\ \gamma_6 = 0,\ \gamma_7 = 0,\ \gamma_8 = 0.7402.$$

Thus, choosing the maximum test statistic,

$$\max\ [1.1921,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0.7402\ ] = 1.1921.$$

Then, choosing the maximum test statistic, the test signal will be classified in Class 1, as the maximum test statistic is γ_1.

The absolute values of the correlation coefficients for a Class 6 test signal after 10 iterations are

$$|a_4^1| = 0.3968,\ |a_1^2| = 0.2867,\ |a_7^3| = 0.2354,\ |a_6^4| = 0.2370,\ |a_7^5| = 0.2321,$$

$$|a_8^6| = 0.1920,\ |a_6^7| = 0.1690,\ |a_2^8| = 0.1647,\ |a_6^9| = 0.1570,\ |a_4^{10}| = 0.1521.$$

The test statistics are computed using Equation (4.1) and are given by

$$\gamma_1 = 0.2867,\ \gamma_2 = 0.1647,\ \gamma_3 = 0,\gamma_4 = 0.5489,$$

$$\gamma_5 = 0,\ \gamma_6 = 0.5630,\ \gamma_7 = 0.4675,\ \gamma_8 = 0.1920.$$

Thus, choosing the maximum test statistic,

$$\max\,[0.2867,\ 0.1647,\ 0,\ 0.5489,\ 0,\ 0.5630,\ 0.4675,\ 0.1920] = 0.5630.$$

Then, choosing the maximum test statistic, the test signal will be classified in Class 6, as the maximum test statistic is γ_6. Thus the MPD parameters are shown to be effective in these simple demonstrations.

As can be seen from Figure 13, the Class 1 test signal also contains some low frequency components similar to Class 8 (see Figure 9). This is the reason for the selection of some dictionary elements representing Class 8 (see Table 4). Thus the MPD parameters give us a very good feature set that clearly distinguishes one class of signals from other classes.

4.4. Energy Conservation

As shown in [35] and repeated in Section 2.2, the MPD preserves energy. Now let us empirically prove the conservation of energy using the same test signal used in the previous section. Recall that the energy conservation equation in Section 2.2, Equation (2.7) can be written as

$$||x(t)||^2 = \sum_{k=1}^{K} |a_j^k|^2 + ||r^K||^2.$$

Here $K = 10$ and a_j^k is the correlation coefficient. All the test signals and the learning signals are normalized to unit energy (i.e),

$$||x(t)||^2 = 1.$$

Substituting the corresponding values in the energy conservation equation, for a Class 1 test signal, we obtain[4](See Table 4)

[4]Note that all the correlation coefficients a_j^k obtained during the decompositon should be used irrespective of chosen class number j.

Class Number	**Number of Misclassified Events**	P_{CC}
1	3	0.99
2	2	0.94
3	1	0.97
4	0	1
5	0	1
6	5	0.84
7	3	0.92
8	0	1
Number of FC	2	

Table 6. Number of misclassified test signals and probability of correct classification for the MPD based method.

4.5. Classification Results Using the MPD

Now let us see the classification results of the MPD principle. Table 6 shows the number of misclassified test signals for each class using this method. It also provides the number of false classifications that corresponds to the number of test signals from classes other than Class 1 that are classified as Class 1, and the P_{CC} of the test signals for each class.

The test and the learning signals used in this testing are identical to the ones used in Chapter 3. Considering the classification results for Class 1 test signals, the MPD misclassified only 3 out of 342 test signals resulting in $P_{CC} = 0.99$. On the other hand, the 2-D matched filter using the spectrogram misclassified 13 out of 342 Class 1 test signals as shown in Table 2 for the same set of test signals. Thus, the MPD method appears promising for acoustic monitoring systems. Two of the three misclassified test signals of Class 1 have more dominant low frequency components along with the components of Class 1. Due to this, the MPD classifies these test signals in Class 8 which is characterized by the presence of low frequency components. There are two false classifications in the MPD

method (Table 6, last row). However, this is not a major drawback of the MPD method since a false classification does not mean that a precursor warning event is overlooked, which is the case in misclassification.

4.6. Pre-classification Prior to Matching Pursuit

Comparing Tables 6 and 2, the MPD method works best in classifying the real data from an acoustic monitoring system. However, the drawback of the MPD method is the processing time. If we compare the processing time of the 2-D matched filter using spectrogram and the MPD method, then the MPD method is slower as it is an iterative algorithm. As it computes the inner products between the test or residual signal and each elements in the dictionary at every iteration, this method is computationally intensive. The processing time increases linearly with the number of dictionary elements as well as the number of iterations. Even though it takes longer time to classify a given test signal, one cannot afford to disregard the precise classification of the MPD method. However if it becomes an issue, it is important to be able to reduce the processing time of the MPD method.

In order to achieve this, we designed a procedure to reduce the number of test signals that require classification using the MPD. Since many recorded acoustic events can easily be distinguished from Class 1 related events due to their characteristics, it is possible to classify them using other less complex and less expensive classification methods. The remaining events, reduced in number are then classified using the MPD.

4.6.1. Implementation of Pre-classifiers. The pre-classifiers were designed as follows.

- **Case 1:** Since Class 3 signals are concentrated around 8 KHz, it is easy to differentiate them from other acoustic events by checking the energy of the signal around 8 kHz. If the energy of a test signal around 8 kHz is above some pre-determined energy threshold, then TF based methods such as 2-D MF using spectrogram can be used to classify it in Class 3. Otherwise, another pre-classification method is used, as the initial classification could only determine that the test signal does not belong to Class 3. Thus the 2-D MF using spectrogram acts as a second level classifier to confirm that the test signal really belongs to Class 3.

- **Case 2:** Most of the Class 8 signals (vehicle related sounds recorded by the acoustic monitoring system) have more energy at low frequencies. Thus, they can be distinguished from the other signals by checking the signal energy at low frequencies. If a test signal is not classified in Class 3 in Case 1, then the energy of the test signal at low frequencies is compared with some pre-determined threshold. If the energy at low frequency exceeds the threshold, then 2-D MF using spectrogram is used to confirm that the test signal is from Class 8. If this is not the case, the algorithm concludes that the test signal is not a Class 8 signal, and continues with further pre-classification.

- **Case 3:** The acoustic monitoring system may record some kind of electrical pop sounds (similar to the sound from an audio amplifier when the output capacitor malfunctions). These signals are grouped under Class 4. As the electrical pops have very small duration, it is easy to differentiate them from other signals. This is done by computing the ratio of the signal energy during the first 12.5 ms to the signal energy during the next 90 ms for this particular data. As this is the third pre-classifier, the test signal cannot be from Class 3 or Class 8. However, if it is a member of Class 4, the

energy ratio exceeds some pre-determined threshold. This will be further confirmed using the 2-D MF using spectrogram. If this test signal is not classified as Class 4 in both of the above pre-classifier methods, then further classification is needed.

- **Case 4:** If the test signal is not classified in any one of the three signal classes mentioned above, then we can check whether the test signal belongs to a Class 1 related event. As mentioned earlier, the Class 1 related events are either wideband or have more energy at higher frequencies. Thus, it is possible to differentiate the Class 1 related events by calculating the ratio of the energy at higher frequencies to their energy at lower frequencies and comparing it to a threshold. If the ratio exceeds the threshold, then the test signal may have come from Class 1, which will be further confirmed by the 2-D MF using spectrogram. If this does not occur, then the next classification step is the MPD classifier.

The flow diagram of this pre-classification procedure is shown in Figure 14. Thus, the processing time is reduced by using other classifiers before the MPD procedure. Additional classes of signals may be pre-classified provided, they have some unique distinguishable features to differentiate them from other signals. Note that even if a Class 1 test signal is not pre-classified correctly in Case 4, it is still possible that it will be classified correctly by the MPD based classifer.

4.6.2. Effects of Pre-classifiers. As far as the misclassification of events is concerned, the MPD combined with the pre-classifiers yields identical results as that of the MPD without any pre-classifiers. However, the major difference is in the processing time. The MPD without pre-classifiers took 40 seconds to classify a test signal, whereas the MPD with pre-classifiers took only 12 seconds to classify the same test signal. Thus, using the

pre-classifiers we are able to reduce the processing time by about 3 times without compromising the classification results. However, this reduction of the processing time depends on

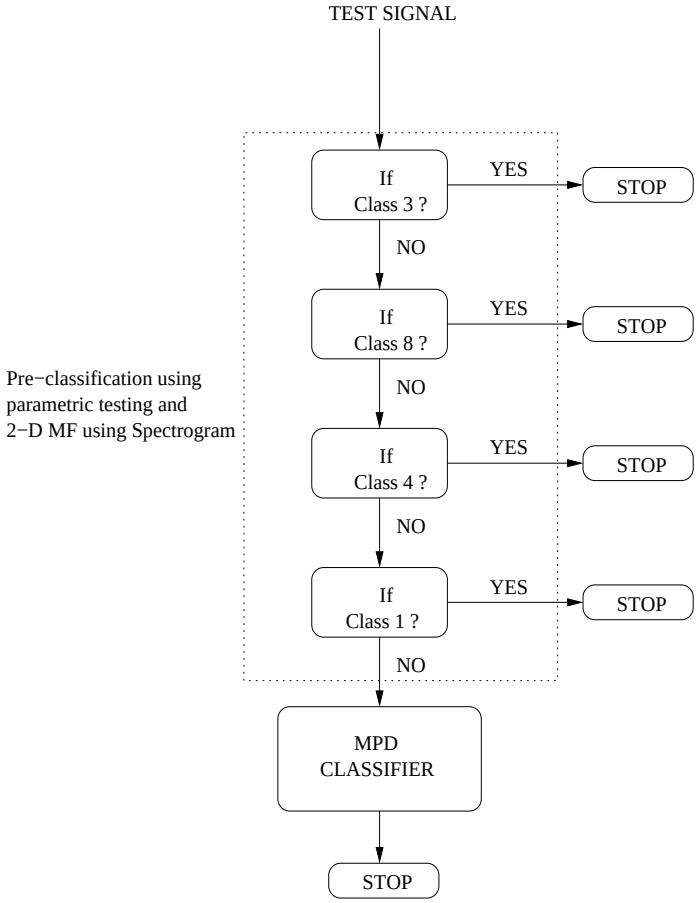

Figure 14. Flow diagram of MPD with Pre-classifier: The test signals of Class 1, 3, 4, 8 will be pre-classified and hence these signals need not go through the MPD procedure, thereby reducing the average processing time.

the probability of occurrence of signals that can be pre-classified in Class 3, 8, 4 or 1. If the probability of occurrence of these signals is high, then the processing time will be reduced considerably. For example, the reduction of processing time from 40 seconds to 12 seconds is due to the fact that 432 out of 546 test signals belonged to one of the 4 classes mentioned in Cases 1-4. Specifically, only 34 out of 432 signals went through the MPD algorithm for final classification.

CHAPTER 5

CLASSIFICATION IN DIFFERENT TEST SCENARIOS

5.1. 13-Class Classifier

In the last two chapters we have seen the 8-class classification results using TF based and MPD based methods. Now, test signals from five new additional set of classes are classified using the MPD with pre-classification method. The sample spectrograms of a signal from each of the five classes are shown in Figure 15. When classification was performed with only dictionary elements from the first eight classes, the results were not satisfactory, as the test signals from the new same class were classified among eight different classes. For example, out of 43 Class 9 test signals, 9 signals were classified in Class 1, 5 signals were classified in Class 2, 9 signals were classified in Class 4, 1 signal was classified in Class 5 and 19 signals were classified in Class 7. Note that the test signals from the previous eight classes were still classified as before in Chapter 4. Since the MPD is based on a signal decomposition principle, it is obvious that if there are no representatives from the new five classes in the MPD dictionary, the algorithm will not converge with a small residual energy. Table 7 shows the number of test signals used for the new five classes and the classification results for the test signals from the five new classes. The motive behind the above testing scenario is to investigate the performance of the MPD when there are no

Class Number	Number of Test Signals	Distribution of Classified Events							
		1	2	3	4	5	6	7	8
Class 9	43	9	5	0	9	1	0	19	0
Class 10	16	5	4	0	1	3	2	1	0
Class 11	8	0	1	0	7	0	0	0	0
Class 12	50	48	0	0	1	1	0	0	0
Class 13	21	0	1	0	0	0	7	13	0

Table 7. Number of test signals in the new classes and the number of signals classified in any one of the eight possible classes.

dictionary elements in the MPD based method is increased, which in-turn increases the processing time of this new 13-class classifier even after pre-classification. To reduce the processing time of the MPD based method, we removed the redundant learning signals of the original eight classes from the dictionary. For example, if there are two learning signals with similar TF structure, then one of them will be removed from the dictionary. The number of frequency shifts are also changed from 6 (starting at 0 Hz with increase of 50 Hz) to 5 (starting at 0 Hz with increase of 75 Hz). Table 8 shows the modified number of learning and test signals for 13 classes. Table 9 shows the classification results of the 13-class classifier using the MPD with pre-classification method.

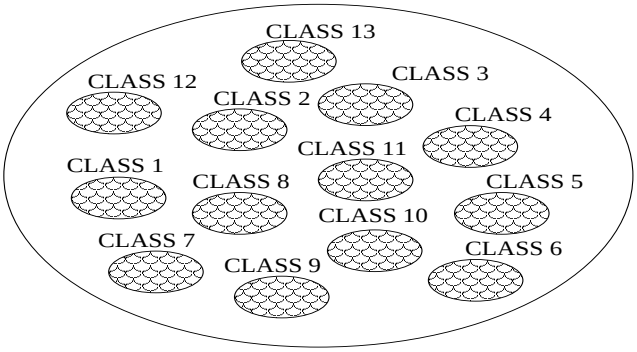

Figure 16. Dictionary of the MPD for 13-class classifier.

Class Number	Number of Learning Signals	Number of Test Signals
Class 1	46	342
Class 2	10	42
Class 3	4	45
Class 4	5	44
Class 5	8	19
Class 6	9	42
Class 7	8	54
Class 8	8	42
Class 9	12	31
Class 10	6	7
Class 11	4	4
Class 12	13	39
Class 13	9	41

Table 8. Modified number of learning and test signals in each class for a 13-class classifier using the MPD principle.

5.2. Binary Classifier

In the previous section, we have seen the classification results of a 13-class classifier using the MPD with pre-classification. Specifically, it was desirable to classify each test signal in one of the 13 specific classes. For example, the automobile noise is needed to be distinguished from birds chirping noise or from dog barking noise. But, sometimes this kind of classification is not needed in an acoustic monitoring system. Hence, consider a scenario where it is sufficient to determine whether a test signal is a warning signal or not. In this case, there are only 2 classes namely,

- Class A - Warning signals (Class 1 in the 13-class classifier), and
- Class B - All other undesirable signals like automobile noise or raining noise.

As a result, we need a binary classifier which classifies a test signal in either Class A or B. From Section 5.1, since Class 12 signals have similar characteristics as Class A signals in the

Class Number	Number of Misclassified Events	P_{cc}
Class 1	4	0.988
Class 2	2	0.95
Class 3	0	1
Class 4	1	0.977
Class 5	0	1
Class 6	5	0.88
Class 7	3	0.94
Class 8	0	1
Class 9	0	1
Class 10	2	0.71
Class 11	0	1
Class 12	8	79
Class 13	1	0.975

Table 9. Test results of the 13-class classifier using the MPD with pre-classification.

TF plane, we decided to consider Class 12 as a sub-class of Class A. Sometimes, Class 12 and Class A are so similar that it requires human judgment to make a final decision. Hence, by including Class 12 in Class A we ensure that no true warning signal is classified as an undesirable signal. Based on how similar the characteristics of some of the other classes are to Class A, we included three types of signals in Class A, *definite warning signal*, *possible warning signal*, and *may be warning signal* (that is Class 12 signals). The number of signals involved are given in Table 10.

5.2.1. Binary Classifier with only Class A Learning Signals as Dictionary Elements. The dictionary of the MPD in Section 5.1 is composed of TF shifted learning signals from 13 classes. However, in this binary classifier, the dictionary is composed of only TF shifted learning signals from Class A (including possible and may be warning signals). Table 10 lists the number of learning and test signals used for each class. The dictionary for the MPD in this case is shown in Figure 17. Since only Class A learning signals are used for this classifier, we cannot use the same decision strategy as in Equation (4.2), Section

Class Number	Number of Learning Signals	Number of Test Signals
Class A definite warning possible warning may be warning	 19 27 13	 381
Class B (non-warning signals)	0	372

Table 10. Number of learning and test signals used for the binary classifier using the MPD based method that uses only the learning signals from Class A.

4.2.1. So the following MPD parameters can be used to make a classification[1].

- Residual energy after some specific number of iterations.
- Correlation coefficient after some specific number of iterations.

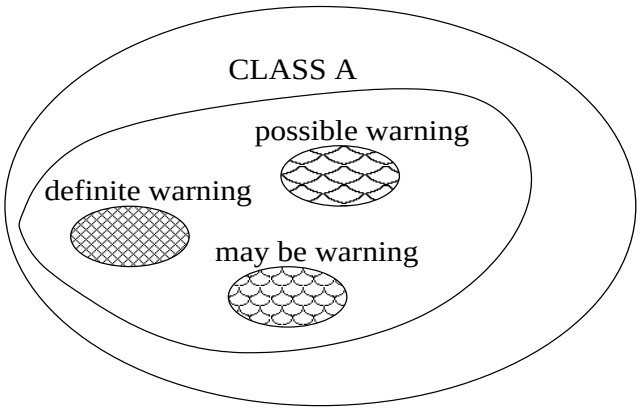

Figure 17. Dictionary of the MPD based method for a binary classifier with learning signals from Class A only.

Since there are no representative signals for Class B, we can anticipate that the residual energy of a Class B test signal after some specific number of iterations will be much higher when compared to the residual energy of a Class A test signal. In order to determine the residual energy threshold value, we used receiver operating characteristics curve (ROC[2]),

[1]Note that the classification was made by hard decision, (i.e) comparing the MPD parameter with some pre-determined threshold and a decision was made based on the comparison.

[2]ROC curves are used in detection and classification theory in order to determine the probability of detection changes as the probability of false classification increases for varying threshold values.

as shown in Figures 18 and 19 for 5 and 15 iterations, respectively.

When the correlation coefficient is used to classify a Class B test signal, we can expect a very low correlation value after some specific number of iterations, as Class B learning signals are not included in the dictionary. In contrast to this, the correlation coefficient after decomposing a Class A test signal, is expected to be high. For the classification, we require a threshold value such that if the correlation coefficient of a test signal is lower than the threshold, then the test signal is classified in Class B. As the test and learning signals are normalized to unit energy, the maximum and minimum possible correlation coefficient and residual energy are 1 and 0, respectively. As we can see from the ROC curves in

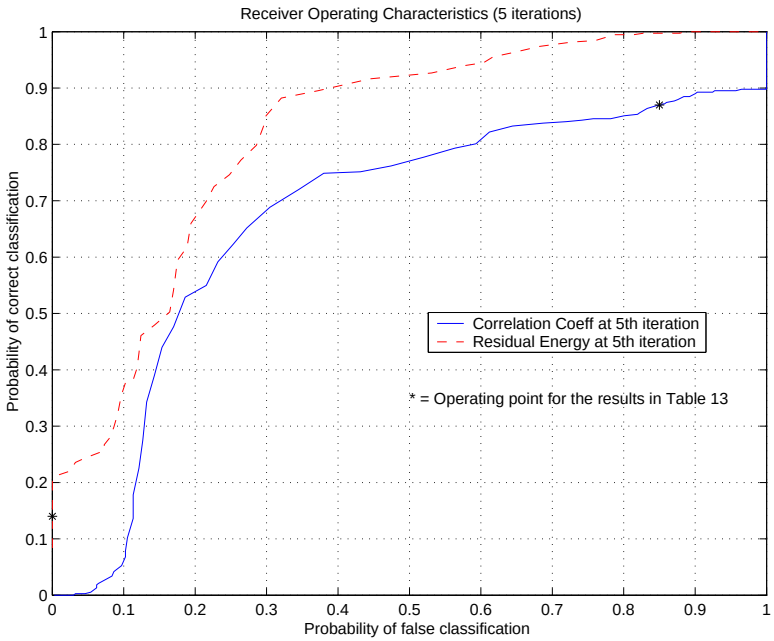

Figure 18. Receiver operating characteristic curves of the MPD based binary classifier with learning signals from Class A only. The ROC curves were obtained by varying the threshold values for the correlation coefficient and the residual energy after 5 iterations. The correlation coefficient threshold value increases as the probability of false alarm increases. The residual energy threshold value decreases as the probability of false alarm increases. The asterisk symbol shows the operating point for the results in Table 11 and 12.

Class Number	**Residual Energy Threshold = 0.3**	**Correlation Coefficient Threshold = 0.12**
Class A (warning signals)	327	49
Class B (non-warning signals)	0	318

Table 11. Number of misclassified events by using MPD binary classifier that uses only the learning signals from Class A after 5 iterations.

Class Number	**Residual Energy Threshold = 0.3**	**Correlation Coefficient Threshold = 0.12**
Class A (warning signals)	286	286
Class B (non-warning signals)	41	46

Table 12. Number of misclassified events by using MPD binary classifier that uses only the learning signals from Class A after 15 iterations.

Figures 18 and 19, for a P_{CC} equal to 0.99 the corresponding P_{FC} is close to 0.8 when the residual energy at both 5 and 15 iterations are used as decision statistics. This implies that if we want to achieve a misclassification rate of 0.01 (4 out of 381 Class A test signals) for Class A, then the misclassification rate for Class B will be 0.8 (297 out of 372 Class B test signals). The classification performance using the correlation coefficient is still worse. From these results, we can infer that none of the two classification criteria are sufficient to clearly differentiate Class A signals from Class B signals. Tables 11 and 12[3] show the misclassification rate of this classifier for 5 and 15 iterations, respectively. The correlation coefficient and the residual energy threshold values were arbitrarily chosen in the range 0 to 1 for demonstration purpose. The corresponding operating points used for these results are shown as an asterisk symbol in the ROC curves.

[3]We used 381 Class A and 372 Class B test signals for testing (see Table 10).

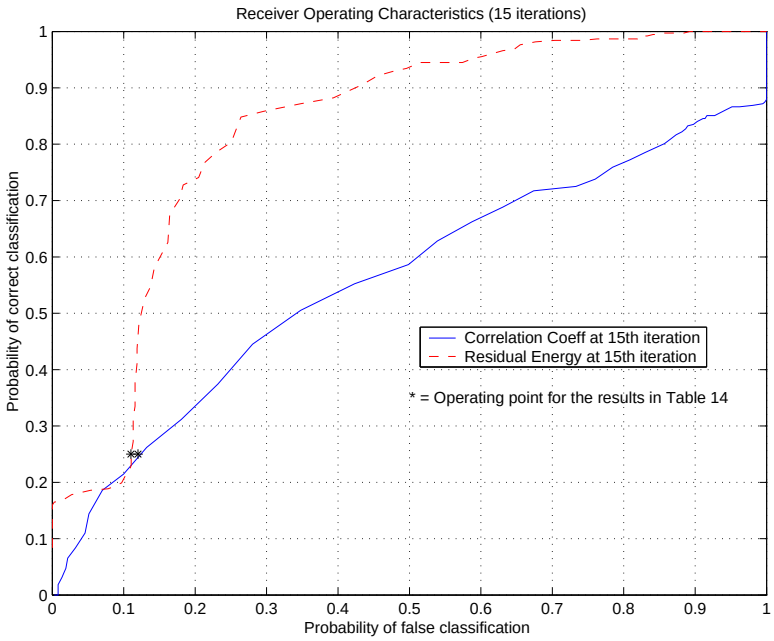

Figure 19. Receiver operating characteristic curves of the MPD based binary classifier with learning signals from Class A only. The ROC curves were obtained by varying the threshold values for the correlation coefficient and the residual energy after 15 iterations. The correlation coefficient threshold value increases as the probability of false alarm increases. The residual energy threshold value decreases as the probability of false alarm increases. The asterisk symbol shows the operating point for the results in Table 11 and 12.

Based on the results we obtained, this binary classifier yields poor classification performance even though the two MPD criteria seem reasonable. Recall that the MPD is an iterative algorithm that selects the best matched dictionary in each iteration by the projection of test or residual signal on each and every dictionary elements. This means that even though there are no representative signals from Class B, the MPD still chooses the best matched dictionary element for all kinds of test signals including signals from Class A and Class B. The time correlation between Class A learning signal and Class B test signals causes the poor performance.

Class Number	Number of Iterations = 10	Number of Iterations = 5
Class A (warning signals)	3	3
Class B (non-warning signals)	7	7

Table 13. Number of misclassified events by the MPD binary classifier that uses learning signals from both Class A and Class B.

5.2.2. Binary Classifier with both Class A and Class B Learning Signals. As we have just seen, it is very difficult to achieve a better classification performance without using some representative signals from Class B. Thus, we changed the dictionary to include some learning signals from Class B as additional dictionary elements for the MPD based method. The overall dictionary of the MPD is shown in Figure 20. The decision strategy for this binary classifier is same as that used in Equation (4.2).

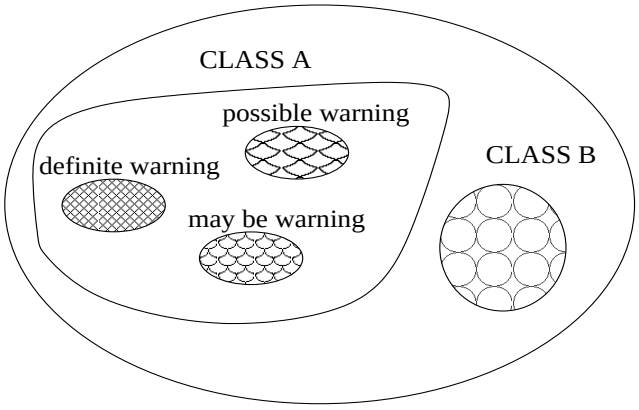

Figure 20. Dictionary of binary MPD classifier with learning signals from both classes.

Table 13 shows the classification results of this method using five and ten iterations as the stopping criterion. As we can see from the table, the MPD now performs much better than when no learning signals from Class B were used. This method misclassifies only 5 out of 381 Class A test signals, and the number of false classification is also acceptable. The average processing time using five iterations is 19 seconds/file and for ten iterations

the average processing time is 33 seconds/file. However using same number of iterations, the MPD gives the same results. As evidenced, it is important to include learning signals in the MPD dictionary from all possible classes.

5.3. Graphical User Interface

The different classifiers discussed in this thesis report involves many complex signal processing techniques. It will be difficult for a signal processing novice to use these classifier algorithms in real life applications as discussed in Chapter 1. This necessitates the need for a flexible graphical user interface (GUI) as a front end through which a novice can invoke the classifier algorithms. The entire algorithm should be designed such that the user can able to do the following operations without implicitly changing the source code of the classifier algorithms.

- The user can create a dynamic L-class classifier. The word dynamic here implies that, a user can able to create either a 3-class classifier or a 5-class classifier for example by changing some simple and manageable parameters in the GUI.
- The user can load the corresponding learning signals of any one of the L classes.
- The user can able to load the test signals that require classification in any one of the L possible classes.
- The user can create an output ASCII file so that the classifier output can be stored in that file for future reference.
- The user can select the operating sampling frequency from a set of possible sampling frequencies that may be encountered in a specific application.

- The user can execute different classifier algorithms. So the following classifier algorithms can be included as possible options for classifying a given test signal:
  - Matching pursuit
  - Matching pursuit with pre-classification
  - 2-D matched filter using the Spectrogram
  - 2-D matched filter using the Reassigned Spectrogram
  - 2-D matched filter using the Ambiguity Function
- Finally, if an acoustic engineer is familiar with signal processing concepts, the front end should allow him or her to change different classifier parameters.

Using the above mentioned operations as the guiding force, we have developed such a GUI in MATLAB as shown in Figure 21. Comprehensive details about this GUI can be found in the user reference manual for this algorithm [46]. We have developed the classifier algorithm in a generic way so that the same source code can be used to classify acoustic signals from different kinds of monitoring systems with very few minor changes.

CHAPTER 6

CONCLUSIONS

6.1. Summary

The objective of this work is the classification of acoustic monitoring signals. This classification is very important in the diagnostic testing of concrete structures. Specifically, the aim was to develop a classifier that accurately classifies warning signals of an acoustic monitoring system. As a result, we used some classical classification methods like the 1-D time correlation method and some TF based methods like the 2-D matched filter using the spectrogram, reassigned spectrogram, and ambiguity function. We then compared their results with the results of our new MPD based classifier. The MPD based classifier classified the warning signals with probability of correct classification equal to 0.99 for our test database. Since the MPD based classifier is computationally intensive, the other TF based classifiers like 2-D MF using the spectrogram are used to reduce the processing time. This is achieved by pre-classifying test signals with distinguishable features before using the MPD, which results in a computationally less intensive method [13].

Many papers proposed various methods for classifying non-stationary signals [44], [47],[48],[49]. We have used some of these methods as well as some new techniques to classify acoustic signals from an acoustic monitoring system. A method not discussed in this report

uses TFR log deviation [44] as a test statistic to classify non-stationary signals. Another test statistic is used in [47],[48],[49] which depends upon taking the inverse of the TF points of a TFR. Note that it is very difficult to use these two methods in signal classification when the signals to be classified like the acoustic events, have no common non-zero values within a band of frequencies. As a result, we did not compare our results with these methods. Moreover, the log deviation method limits the type of TFR, because only positive TFRs like the spectrogram or the reassigned spectrogram can be used.

For signals similar to those used in this report, our new MPD based method provides promising classification results. The only drawback of the MPD is the processing time. Since in an acoustic monitoring system of concrete structures, signal processing can be done off-line, the processing time may not be a major issue when compared with the high classification performance. For situations where the processing time is critical, we have introduced some pre-classifier methods before using the MPD. Thus, the MPD based classifier is well suited for classifying acoustic emissions with high degree of accuracy and a lower number of false classifications.

We have also illustrated the performance of the MPD based classifier for different test scenarios. Specifically, we have discussed two test scenarios namely,

- the binary classification when the user is only interested in the classification of warning signals, and
- the L-class classifier when the user needs to classify a test signal in one of L possible classes.

The latter scenario adds additional burden onto the MPD based classifier as it allows the classifier to not only classify a test signal as a warning signal, but also to sub-classify

a non-warning signal. Even though it appears that it is sufficient for someone to know whether a test signal is a warning signal or not, the sub-classification of non-warning signals may provide some additional information regarding the surrounding area of the concrete structure that is being monitored.

The flexibility of the MPD based classifier allows one to implement both cases by just changing the learning signal sets. The dictionary of the binary case will have two classes of learning signals while the dictionary for the multi-class case requires L classes of learning signals. Thus, the implementation of the MPD based classifier is flexible in the sense that it can be easily modified to classify signals from other acoustic monitoring systems by changing the learning signal set. However, care should be taken in selecting the range of frequency shifts of the learning signals used to form the dictionary. Since a higher range of frequency shift may increase the misclassification rate. This is due to the fact, that when a learning signal of a particular class is frequency shifted by a large amount, then it is possible that the TF structure of this frequency shifted learning signal may resemble a learning signal of another class.

We have also developed a GUI so that anybody who is unfamiliar with the complex signal processing concepts involved in the classifier algorithm can in fact use them without any difficulties. The front end is also designed in such a way that an acoustic engineer who is aware of all background signal processing concepts can play around with different parameters and optimize the algorithm for specific applications. For example, the acoustic engineer can change the spectrogram parameters for optimum results.

Even though the automatic acoustic classifiers described in this report work well, if more accurate results are needed to decide on the condition of a structure, then manual analysis will be typically performed on the test signals classified as warning signals, before

making a final decision. These automatic classifiers act as a tool to acoustic engineers to separate events which are important for further analysis, and to neglect a rather large number of signals that are of no diagnostic values. By doing this, the automatic acoustic classifiers save a lot of man power, and reduce stressful acoustic signal analysis. An acoustic engineer would rather analyse those signals that pertain to the condition of a structure than listen to all unnecessary audio signals.

6.2. Future Work

Since we have implemented this automatic acoustic classifier in Matlab, there is an extra financial burden on the end user, as the algorithm requires to have Matlab installed in computers. In order to avoid this problem, we can re-write the Matlab source code in C/C++. Another advantage of converting the source code from Matlab to C/C++ is the reduction in processing time. We know that Matlab is not computationally efficient when compared to directly running an executable file. Thus, by optimizing the source code in C/C++ we can anticipate an appreciable reduction in processing time.

As we have used the TF based and MPD based classifiers for a specific class of acoustic signals recorded by a real acoustic monitoring system of concrete structures, we can investigate the feasibility of extending these classifiers to classify signals from other applications. One example of another application is the classification of warning signals from enemy submarines in an underwater acoustic monitoring system. Also, we need to study the performance of the MPD based classifier with larger database than the one used in this report.

As seen in Section 4.6, the processing time of a MPD based classifier can be reduced by using less computationally intensive pre-classifiers. However, the performance of this

pre-classifier is increased only when the number of classes that can be pre-classified with more confidence increases. Further research can be done to include more number of classes that can be pre-classified and they do not require MPD to distinguish them from warning signals.

Even though the automatic classifiers discussed in this report work well, we can also investigate other powerful and novel classification methods based on eigen value analysis and singular value decomposition as described in [50]. Initially, we have decided not to pursue in the direction of linear analysis due to the fact that the dimension of feature vectors that will be required to classify a test signal will be large, thereby increasing the computational time.

REFERENCES

[1] J. C. Hassab, *Underwater signal and data processing.* Florida, USA: CRC Press, 1989.

[2] W. W. L. Au, *The Sonar of Dolphins.* Springer-Verlag, 1993.

[3] W. C. Michie, G. Thursby, D. Walsh, B. Culshaw, and M. Konstantaki, "Distributed sensing of physical and chemical parameters for structural monitoring," *Optical Techniques for Smart Structures and Structural Monitoring*, no. 1997/033, pp. 3/1 – 3/9, 1997.

[4] E. A. Johnson, P. G. Voulgaris, and L. A. Bergman, "Methods of system identification for monitoring slowly time-varying structural systems," in *Proceedings of the Intelligent Information Systems*, (Grand Bahama Island, Bahamas), pp. 569 – 573, 1997.

[5] R. M. Measures, "Fiber optic structural monitoring of bridges," in *Proceedings of the Instrumentation and Measurement Technology Conference*, (Ontario, Canada), pp. 600 – 602, 1997.

[6] W. Worthington, "Prestressed concrete pipe inspection and monitoring methods," in *Proceedings of the Nondestructive Evaluation of Civil Structures and Materials Conference*, (Colorado, USA), 1992.

[7] J. T. Tou and R. C. Gonzalez, *Pattern Recognition Principles.* Massachusetts, USA: Addison Wesley Publishing Company, 1974.

[8] K. S. Fu, *Digital Pattern Recognition.* New York, USA: Springer-Verlag, 1976.

[9] T. Y. Young and T. W. Calvert, *Classification, Estimation and Pattern Recognition.* New York, USA: American Elsevier Publishing Company, 1974.

[10] S. Haykin, *Communication Systems.* New York, USA: Wiley, John & Sons, Incorporated, 1994.

[11] S. Haykin, *Digital Communications.* New York, USA: Wiley, John & Sons, Incorporated, 1999.

[12] J. G. Proakis, *Digital Communications.* New York: McGraw-Hill, 2001.

[13] S. P. Varma, A. Papandreou-Suppappola, and S. B. Suppappola, "Detecting faults in structures using time-frequency techniques," in *Proceedings of the IEEE International Conference on Acoustic, Speech and Signal Processing*, (Utah, USA), pp. 3593–3596, May 2001.

[14] S. P. Ebenezer, A. Papandreou-Suppappola, and S. B. Suppappola, "Matching pursuit classification for time-varying acoustic emissions," in *Proceedings of the Thirty Fifth Asilomar Conference on Signals, Systems and Computers*, (California, USA), November 2001.

[15] S. P. Ebenezer, A. Papandreou-Suppappola, and S. B. Suppappola, "Classification of time-varying signals from an acoustic monitoring system," *to be submitted to IEEE Transactions on Signal Processing in*, 2002.

[16] S. Haykin and B. V. Veen, *Signals and Systems.* New York, USA: John Wiley and Sons Inc., 1999.

[17] A. V. Oppenheim and R. Schaffer, *Discrete-Time Signal Processing*. New Jersey, USA: Prentice-Hall, 1998.

[18] A. D. Poularikas, *The Transforms and Applications Handbook*. Florida, USA: CRC Press LLC, 1996.

[19] F. Hlawatsch and G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Processing Magazine*, vol. 9, pp. 21–67, April 1992.

[20] L. Cohen, "Time-frequency distributions - a review," in *Proc. IEEE*, vol. 77, pp. 941–981, July 1989.

[21] P. Flandrin, *Time-Frequency/Time-Scale Analysis*. California, USA: Academic Press, 1999. (Translated from French, *Temps-fréquence*. Paris: Hermès, 1993).

[22] L. Cohen, *Time-Frequency Analysis*. New Jersey, USA: Prentice-Hall, 1995.

[23] M. R. Portnoff, "Time-frequency representation of digital signals and systems based on short-time Fourier analysis," *IEEE Transactions on Acoustic, Speech, Signal Processing*, vol. 28, pp. 55–69, 1980.

[24] T. A. C. M. Classen and W. F. G. Mecklenbrauker, "The Wigner distribution - a tool for time-frequency signal analysis, Part II:Discrete time signals," *Philips Journal of Research*, vol. 35, no. 4/5, pp. 276–300, 1980.

[25] T. A. C. M. Classen and W. F. G. Mecklenbrauker, "Time-frequency analysis by means of the Wigner distribution," in *Proceedings of the IEEE International Conference on Acoustic, Speech and Signal Processing*, (Georgia, USA), pp. 69–72, 1981.

[26] M. D. F. Taylor and D. Chester, "On the Wigner distribution," in *Proceedings of*

the IEEE International Conference on Acoustic, Speech and Signal Processing, (Massachusetts, USA), April 1983.

[27] P. Flandrin, "Some features of time-frequency representations of multicomponent signals," in *Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing*, (California, USA), pp. 41B.4.1–4, March 1984.

[28] H. I. Choi and W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, pp. 862–871, June 1989.

[29] A. Papandreou-Suppappola, F. Hlawatsch, and G. F. Boudreaux-Bartels, "The hyperbolic class of quadratic time-frequency representations part I: Constant-Q warping,the hyperbolic paradigm, properties and members," *IEEE Transactions on Signal Processing*, vol. 41, pp. 3425–3444, December 1993.

[30] A. Papandreou and G. F. Boudreaux-Bartels, "Distributions for time-frequency analysis: A generalization of Choi-Williams and the Butterworth distribution," in *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, vol. 5, (California, USA), pp. 181–184, March 1992.

[31] R. G. Baraniuk and D. L. Jones, "A radially Gaussian, signal dependent time-frequency representation," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, (Canada), pp. 3181–3184, May 1991.

[32] D. L. Jones and R. G. Baraniuk, "A signal dependent time-frequency representation: Optimal kernel design," *IEEE Transactions on Signal Processing*, vol. 41, pp. 1589–1602, April 1993.

[33] F. Auger and P. Flandrin, "Improving the readability of time-frequency and time-scale representations by the reassignment method," *IEEE Transactions on Signal Processing*, vol. 43, pp. 1068–1089, May 1995.

[34] F. Auger and P. Flandrin, "The why and how of time-frequency reassignment," in *Proceedings of the International Symposium on Time-Frequency and Time-Scale Analysis*, (Philadelphia, USA), pp. 245–248, October 1994.

[35] S. G. Mallat and Z. Zhang, "Matching pursuit with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, pp. 3397–3415, December 1993.

[36] A. Bultan, "A four-parameter atomic decomposition of chirplets," *IEEE Transactions on Signal Processing*, vol. 47, pp. 731–745, March 1999.

[37] S. Qian and J. M. Morris, "Wigner distribution decomposition and cross-terms deleted representation," *Signal Processing*, vol. 47, no. 2, pp. 125–144, 1992.

[38] A. Papandreou-Suppappola and S. B. Suppappola, "Adaptive time-frequency representations for multiple structures," in *10th IEEE workshop on Statistical Signal and Array Processing*, (Pennsylvania, USA), pp. 579–583, August 2000.

[39] A. Papandreou-Suppappola and S. B. Suppappola, "Analysis and classification of time-varying signals with multiple time-frequency structures," *IEEE Signal Processing Letters*, to appear December 2001.

[40] B. Torresani, "Wavelet associated with representations of the affine weyl-heisenberg group," *Journal of Mathematics and Physics*, vol. 32, pp. 1273–1279, May 1991.

[41] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory.* New Jersey, USA: Prentice Hall, Inc., 1998.

[42] McDonough, N. Robert, and A. D. Whalen, *Detection of Signals in Noise.* Academic Press, 1995.

[43] A. Papoulis, *Probability, Random Variables, and Stochastic Process.* New York, USA: McGraw-Hill, 1991.

[44] C. D. I. Vincent and E. L. Carpentier, "Non stationary signal classification using time-frequency distributions," in *Proceedings of the International Symposium on Time-Frequency and Time-Scale Analysis*, (Philadelphia, USA), pp. 233–236, October 1994.

[45] M. R. McClure and L. Carin, "Matching pursuits with a wave-based dictionary," *IEEE Transactions on Signal Processing*, vol. 45, pp. 2912–2927, December 1997.

[46] S. P. Ebenezer, "User manual for acoustic classifier using time-frequency techniques," tech. rep., Arizona State University, Department of Electrical Engineering, Tempe, Arizona, December 2001.

[47] G. Roberts, A. M. Zoubir, and B. Boashash, "Classification of non-stationary random signals using multiple hypothesis testing," in *Proceedings of IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, (Paris, France), pp. 245–248, June 1996.

[48] G. Roberts, A. M. Zoubir, and B. Boashash, "Time-frequency discriminant analysis for non-stationary gaussian signals," in *Proceedings of the International Symposium on Signal Processing and its Applications*, (Brisbane, Australia), pp. 33–36, August 1996.

[49] G. Roberts, A. M. Zoubir, and B. Boashash, "Time-frequency classification using a multiple hypothesis test: An application to the classification of humpback whale signals,"

in *Proceedings of the International Conference on Acoustic, Speech, Signal Processing*, (Munich, Germany), pp. 563–566, April 1997.

[50] R. E. Learned, W. C. Karl, and A. S. Wilsky, "Wavelet packet based transient signal classification," in *Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*, (Victoria, Canada), pp. 109–112, October 1992.

APPENDIX A

ACRONYMS

Acronym	Meaning
1-D	One-dimensional
2-D	Two-dimensional
AF	Ambiguity function
AWGN	Additive white Gaussian noise
FC	False classification
MF	Matched filter
MPD	Matching pursuit decomposition
P_{CC}	Probability of correct classification
PDF	Probability density function
P_{FC}	Probability of false classification
QPSK	Quadrature phase shift keying
QTFR	Quadratic time-frequency representation
ROC	Receiver operating characteristics
RSPEC	Reassigned spectrogram
SPEC	Spectrogram
STFT	Short-time Fourier transform
TF	Time-frequency
TFR	Time-frequency representation
WD	Wigner distribution

Table 14. List of acronyms used in this thesis.

Table 1 provides a list of all acronyms used in this thesis.

APPENDIX B

NOTATION

Symbol	Description	Quantity
$x_j^k[n]$	j^{th} class, k^{th} learning signal, n^{th} feature. $j = 1,\ldots L$ (L classes of signals are present) $k = 1,\ldots K$ (K learning signals are included) $n = 1,\ldots N$ (N features are present)	Scalar
$\mathbf{X}$	Test signal vector	Vector
$\overline{\mathbf{X}}_j$	Template vector of j^{th} class	Vector
$p(\mathbf{X}/C_j)$	PDF of a test signal $\mathbf{X}$ given that it belongs to Class j	Vector
$\mu_j[n]$	Mean of the n^{th} feature of Class j	Scalar
$\boldsymbol{\mu}_j$	Mean vector of the template of Class j	Vector
$\sigma_j^2[n]$	Variance of the n^{th} feature of Class j	Scalar
$\sigma_j^2\mathbf{I}_{N\times N}$	Covariance matrix of the template vector of class j, whose elements are uncorrelated random variables with equal variance σ_j^2	Matrix
$\mathbf{TFR}_x$	TFR of a test signal	Matrix
$\overline{\mathbf{TFR}}_j$	j^{th} class average TFR, ‘ Average of TFR of the learning signals in Class j.	Matrix
$\mathbf{TFR}_{x_j^k}$	TFR of the k^{th} learning signal in the j^{th} class	Matrix
$\mathbf{AF}_x$	Ambiguity function of a test signal	Matrix
$\overline{\mathbf{AF}}_\mathbf{j}$	j^{th} class average ambiguity function, Average of AF of the learning signals in Class j.	Matrix
$\mathbf{AF}_{x_j^k}$	Ambiguity function of the k^{th} learning signal in the j^{th} class	Matrix

Table 15. List of some symbol notations used in this thesis.

Table 15 shows the list of notation that are adopted in the following chapters.

APPENDIX C

TERMINOLOGY

The following are some of the terms used in the thesis report.

Learning Signals :

For any classification method, prior knowledge of the different classes of signals that are to be classified is assumed known. Learning signals represent the characteristics of the different classes of signals to be classified. In other words, the learning signals in a class are the templates/prototypes of this specific class. As a result, the class of each learning signal is defined prior to classifying an unknown signal [7].

Test Signals :

Test signals are the unknown signals obtained from the sensors that require classification [7].