Ph.D. Thesis

IMPLEMENTATION OF AN UNDERWATER TARGET CLASSIFIER USING HIGHER ORDER SPECTRAL FEATURES

Submitted to

THE COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY

in partial fulfilment of the requirement for the award of the degree of

Doctor of Philosophy

by

MOHANKUMAR K

Under the guidance of

Prof. (Dr.) P. R. S. Pillai

DEPARTMENT OF ELECTRONICS
COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY
COCHIN - 682 022, INDIA

AUGUST 2015

IMPLEMENTATION OF AN UNDERWATER TARGET CLASSIFIER USING HIGHER ORDER SPECTRAL FEATURES

Ph.D. Thesis in the field of Ocean Electronics

Author

Mohankumar K
Research Scholar
Department of Electronics
Cochin University of Science And Technology
Cochin - 682 022, India
e-mail: kmohankumar@gmail.com

Research Advisor

Dr. P. R. S. Pillai
UGC Emeritus Fellow
Department of Electronics
Cochin University of Science And Technology
Cochin - 682 022, India
e-mail : prspillai@cusat.ac.in

August 2015

Dedicated to... ... ...

My Mother

COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY

DEPARTMENT OF ELECTRONICS

Thrikkakara, Cochin - 682 022

CERTIFICATE

This is to certify that this thesis entitled, ***Implementation of an Underwater Target Classifier using Higher Order Spectral Features*** is a bonafide record of the research work carried out by Mr. Mohankumar K. under my supervision in the Department of Electronics, Cochin University of Science and Technology. The results presented in this thesis or parts of it have not been presented for any other degree(s). Also certified that thesis is adequate and complete for the award of the Ph.D. Degree.

Cochin - 22
10-08-2015

Prof. (Dr.) P.R.S. Pillai
Supervising Guide

DECLARATION

I hereby declare that the work presented in this thesis entitled ***Implementation of an Underwater Target Classifier using Higher Order Spectral Features*** is a bonafide record of the research work carried out by me under the supervision of Dr. P.R.S. Pillai, UGC Emeritus Fellow, in the Department of Electronics, Cochin University of Science and Technology. The result presented in this thesis or parts of it have not been presented for any other degree(s).

Cochin - 22 **MOHANKUMAR K**

10 Aug, 2015

COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY

DEPARTMENT OF ELECTRONICS

Thrikkakara, Cochin - 682 022

CERTIFICATE

This is to certify that this thesis entitled, ***Implementation of an Underwater Target Classifier using Higher Order Spectral Features*** has been modified to effect all the relevant corrections suggested by the Doctoral Committee and the audience during the Pre-synopsis Seminar

Cochin - 22
10-08-2015

Prof. (Dr.) P.R.S. Pillai
Supervising Guide

Acknowledgements

I would like to express my deepest sense of gratitude to my research guide, **Prof. (Dr.) P.R.S. Pillai**, UGC Emeritus Fellow, Department of Electronics, Cochin University of Science and Technology for his excellent guidance and incessant encouragement. It has been a great pleasure and privilege to work under him and he was always there when I needed help.

I am much grateful to **Dr. Supriya M. H.**, Associate Prof. & Head, Department of Electronics, Cochin University of Science and Technology, for the whole hearted support and constant encouragement.

Sincere thanks are due to **Prof. (Dr.) P. Mohanan**, **Prof. (Dr.) K. Vasudevan** and **Prof. (Dr.) C. K. Aanandan**, Department of Electronics, Cochin University of Science and Technology, for providing adequate help and fruitful suggestions.

I take this opportunity to express my sincere thanks to **Dr. Tessamma Thomas**, **Dr. James Kurian**, Department of Electronics, for the constant encouragement rendered to me.

I would like to give a special word of thanks to the research fellows and project staff in the **Centre for Ocean Electronics (CUCENTOL)**, Mr. Jinto George, Mr. Prajas John, Ms. C. Prabha, Mr. Binesh T., Mr. Prasanth P.P., Mr. Sujith Kumar S. Pai, Mrs. Sabna N., Mr. Suraj Kamal, Mr. Satheesh Chandran, Shameer K. Mohammed and Sherin Ben.

I thankfully bear in mind the sincere co-operation and support I received from the **library and administrative staff** of the Department.

I also take this opportunity to thank all the **M.Tech. and M.Sc.** students, Department of Electronics, Cochin University of Science and Technology who contributed and helped me in completing my thesis.

My heartfelt gratitude goes to Mr. Subhash Chandra Bose M. R. and Mrs. Reshmi S for their constant support and encouragement.

It is beyond words to express my gratitude to my mother **Mrs. Pankajakshy K.** for her constant prayers and support in connection with preparation of my thesis.

A word of mention is deserved by my cousins Ammu and Ajitha, who were a constant source of motivation and energy for me. I also thank Mr. Sajeevu and Mr. Somashekharan, who always lended me an emotional support to overcome difficult situations.

I am grateful to all of my friends and neighbours at my native Malamakkavu, for their immense love and constant encouragement.

I thankfully acknowledge the sincere co-operation and support I received from the staff and management of Grey Technolab India Pvt. Ltd., Cochin. I am sure I could not have completed this blessed task without their support and cooperation.

MOHANKUMAR K

10^{th} August 2015.

ABSTRACT

The detection and classification of underwater targets has gained considerable research interest due to its strategic as well as commercial importance. As the operator assisted classification turns out to be tedious and time consuming, there is a need and requirement for the intelligent classifiers that could analyze the received noise and identify the targets. The process of automated classification involves the extraction of source specific features from the noise emanated from the targets, followed by the application of a classification algorithm, based on some form of pattern matching techniques.

The ocean, as a propagation medium, is full of interfering noise sources which include biological, natural as well as man-made sources. This makes the extraction of robust features for such classification a challenging problem, as the signals emitted by the target are corrupted by the ambient noise. There has been many feature extraction techniques, which are primarily based on the second order power spectral analysis. Such traditional techniques can fail to provide acceptable confidence levels during classification, especially when there are deviations from Gaussianity and linearity.

The proposed work envisages the implementation of an underwater target classifier making use of the source specific features extracted using Higher Order Spectral analysis, especially the Bispectrum. The higher order spectrum (HOS) has many attractive properties which, if utilized properly, can make it a potential candidate for the extraction of features from the underwater noise signals. Different types of bispectral features have been extracted from the target waveforms and a suitable feature selection technique is applied to generate the feature vector.

Several classification frameworks have been implemented by the combination of different feature selection criterion and classification algorithms, *viz.*, the k-Nearest Neighbour (k-NN), Artificial Neural Network (ANN) and Support Vector Machines (SVM) and their performances investigated.

Contents

List of Figures

List of Tables

Abbreviations

AIB	Axially Integrated Bispectrum
ANN	Artificial Neural Network
BGTCC	Bispectral Gammatone Cepstral Coefficient
BMFCC	Bispectral Mel Frequency Cepstral Coefficient
CGF	Cumulant Generating Function
CIB	Circularly Integrated Bispectrum
DCT	Discrete Cosine Transform
DFT	Discrete Fourier Transform
ERB	Equivalent Rectangular Bandwidth
FFT	Fast Fourier Transform
GTCC	Gammatone Cepstral Coefficient
HMM	Hidden Markov Model
HOS	Higher Order Spectra
JMI	Joint Mutual Information
k-NN	k-Nearest Neighbour
LSA	Least Squares Approach
MFCC	Mel Frequency Cepstral Coefficient
MGF	Moment Generating Function
MI	Mutual Information
MLP	Multi Layer Perceptron
QPC	Quadratic Phase Coupling
RBF	Radial Basis Function
RIB	Radially Integrated Bispectrum
SNR	Signal to Noise Ratio
SVM	Support Vector Machine

Chapter 1

Introduction

This chapter provides an introduction to the area of research and starts with an overview of the sonars which are the primary means for acquiring the target signals. Sonars play a key role in the ocean research, helps in remotely detecting, locating and classifying objects underwater. The chapter also touches upon active and passive sonar systems as well as the various noise sources in the ocean and their spectral characteristics. The need and requirement of a target classifier for identifying the noise sources in the ocean along with the underlying architecture and the principle of operation have been discussed. Various techniques for feature extraction, feature selection and different classification algorithms have also been briefly introduced, together with an outline of the proposed work.

1.1 Background

The oceans have always been a fascination and mystery to the mankind. It has even frightened him to the extent that many old civilizations practised worshipping it. The bewilderness, fear and fascination slowly paved the way for inquisitiveness and exploration and the trend increased as the oceans served as a source of rich resources. The exploration became a necessity when oceans became a pathway for commercial and military sea faring. A good navy became essential for empire expansion plans and to secure own boundaries. Consequently, the demand for newer vessels with improved capabilities increased and the research in this area became indispensable.

With the advent of more and more ships and vessels, the question *who is around me* or *who is hiding there* gained importance. Such a knowledge is significant from strategic, social as well as commercial viewpoints. Long back as 1490, Davinci had observed a way to detect ships travelling at a distance by carefully listening the radiated sound through a tube dipped in the ocean waters. Even though the works of Isaac Newton and Rayleigh provided the mathematical analysis and fundamental theories describing generation, propagation and reception of sound in the 18^{th} century itself, the first electrically driven underwater transceiver was operational only in 1914, just before the First World War. With the outset of the World War I, acoustic sensing devices like SC tubes were developed to detect submarines and ships. With World War II, the detection and identification of the type and nature of the vessels became even more necessary and the studies in this direction led to the development of the modern SONAR (Sound Navigation and Ranging) systems.

It has been found that the sound travels with less attenuation in water compared to other forms of radiated energies like electromagnetic waves, and the sonars utilise this property of sound for navigation and ranging. Though, the process of detection of even remote targets is relatively easy by listening to the target signals, the identification of the corresponding sources remains elusive and specially trained operators are generally employed for identifying the targets. Modern high-speed digital electronics and the improvements in the fields of signal processing have influenced the sonar design to incorporate automated software-based algorithms to aid the sonar operators in the process of such decision making. With the advent of research in the field of Artificial Intelligence and Machine Learning, techniques for automatic target classification have also been developed, that could help the sonar operators in detecting and identifying various targets. However, the variety and diversity of the targets along with the variability and complexity of the ocean environment always posed challenges to such attempts.

1.2 The Sonar

In general, the Sonar refer to the devices or methods that use underwater sound propagation to explore the presence, position and properties of objects in the sea. A basic block diagram of the Sonar receiving system is as shown in Fig. 1.1, which depicts the sensor array, data acquisition system, beam forming, detection processing and finally the output display.

The sensor array converts the acoustic signals to electrical signals, which is digitised by the use of a proper data acquisition system. Beam

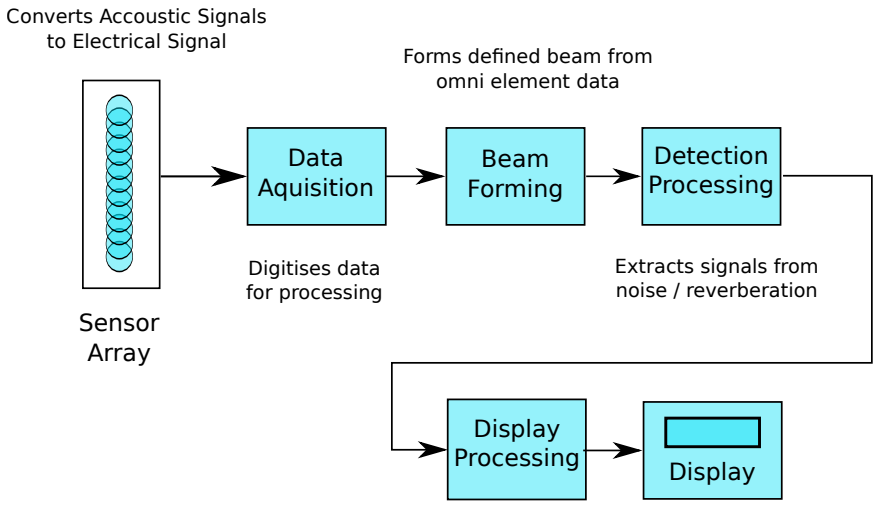


Figure 1.1: Basic SONAR receiver block diagram

forming can be used for spatial filtering, to selectively receive the signals from a particular direction. The received signals are processed for the detection of any targets and are finally presented to the sonar operators for decision making.

1.2.1 Types of Sonars

Generally, sonars are of two types, the *passive sonar* that listens to the incoming sound and the *active sonar* that illuminates the ambience with acoustic energy and observes the back scatterings. Active sonars transmit the acoustic energy in the form of a pulse, often called a *ping* and then listen for its reflection, or the *echo*. The transmission of acoustic energy is achieved using projectors and the reception is carried out using hydrophones and the received echoes form the basis for detection and classification of the targets. With active sonars, one can measure the distance and bearing of the targets, more easily and accurately.

Passive sonars are listening sonars and just listen to the acoustic signals radiated by the targets. There is only one way transmission, from target to the sonar, and the sonar system comprises of only hydrophones as transducers, for the reception of the radiated signals generated by the targets. A typical passive sonar system has been illustrated in Fig. 1.2. Unlike active sonars, no signals are sent out. The signals generated by the target, after retrieving from the hydrophone array are analysed to detect, classify and locate the target.

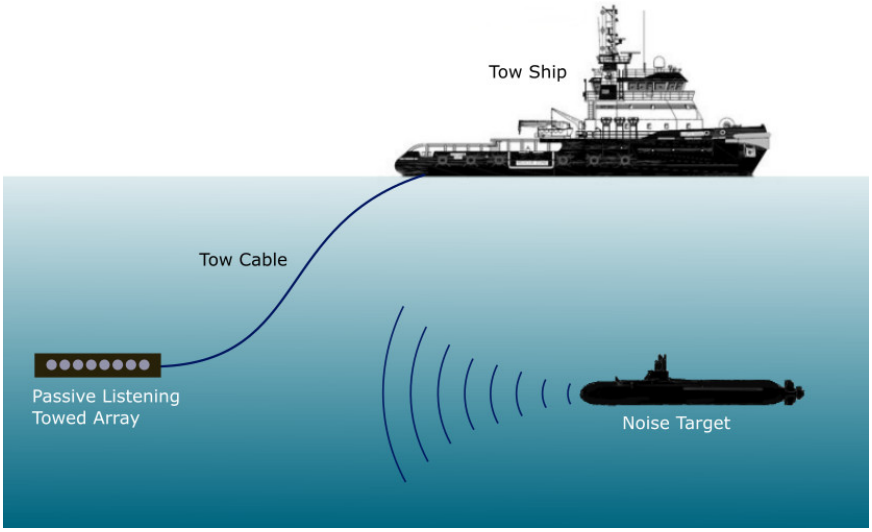


Figure 1.2: A Typical Passive Sonar System

The characteristics of the target of interest generally determine whether to use active or passive sonar systems. Since, only listening for the incoming signals generated by the targets is involved, the passive sonar system has the advantage that it does not expose itself by transmitting any form of energy and in this sense, it is more ideal for target classification purposes compared to active sonars, especially for strategic applications. Moreover, unlike active sonars, the passive sonars does not interfere with the marine ecological balance with its

high energy acoustic emissions. Although in World War II, active sonars were mainly used, with the advent of more noisy submarines and surface vessels, passive sonar became the preferred choice for early detection and warning applications. This thesis deals with underwater passive sonar target classification using Higher Order Spectral Features.

1.2.2 Sonar Equations

The working of the sonar and its efficiency is determined by various factors. The sonar equation takes into account the effects of such factors of the medium, target and the sonar equipment and establishes working relationships between them. The most common practical uses of sonar equations include the design of the sonar system with the required characteristics and the prediction of performance of the sonar equipment of known or existing design.

The sonar equations are formed from the basic equality between the *desired* (signal) and *undesired* (background) portions of the received signal. For just detecting a target using sonars, one requires that:

$$SignalLevel = BackgroundLevel$$

This equality can be expanded in terms of the sonar parameters which can be grouped into three sets depending on whether they are related to the equipment, medium or the target.

- *Equipment Related Parameters*: Projector Source Level (SL), Self Noise Level (NL), Receiving Directivity Index (DI) and Detection Threshold (DT)

- *Medium Related Parameters*: Transmission Loss (TL), Reverberation Level (RL) and Ambient Noise Level (NL)
- *Target Related Parameters*: Target Strength (TS) and Target Source Level (SL)

Consider a transducer producing a signal with a *source level* of SL decibels at a unit distance. Let the radiated sound reaches a target of *target strength* TS, and get reflected back so as to reach the radiating transducer again. As the signal travels through the medium, let its level be reduced by the *transmission loss*, TL and thus the echo level at the hydrophone terminals becomes $SL - 2TL + TS$. If the transducer has a directivity index of DI and if the *background level* is NL, the relative noise power at the terminals of the hydrophone will be $NL - DI$. Hence, at the hydrophone terminals, the signal-to-noise ratio (SNR) is $SL - 2TL + TS - (NL - DI)$. The detection of the target will occur, when this input SNR is above a certain detection threshold, DT, satisfying certain probability criteria. Thus, if the target is present, just at the point of detection, the SNR will be equal to the detection threshold, and the active-sonar equation becomes

$$SL - 2TL + TS - (NL - DI) = DT \tag{1.1}$$

For a reverberation background, the term $(NL - DI)$ is replaced by an equivalent plane-wave *reverberation level* RL observed at the hydrophone terminals. The active-sonar equation then becomes $SL - 2TL + TS = RL + DT$.

In the passive case, the target itself produces the signal so that the parameter target strength becomes irrelevant. As only one way transmission is involved, the signal undergoes attenuation due to

transmission loss only once. Thus, the passive sonar equation can be obtained as:

$$SL - TL - (NL - DI) = DT \tag{1.2}$$

1.3 Functions of Sonar

In the context of target recognition, the ultimate goal of the sonar is to extract the relevant parameters of the radiated acoustic space-time field, so that any or all of the functions *viz.*, detection, localisation, estimation and classification of the target can be achieved.

1.3.1 Detection

In sonar terminology, the term detection refers to the process of determining the presence of a target in the noise background, by analysing the received noise signals. Generally, the targets are detected when the acoustic field generated by them produce localised peaks which stand out against the relatively smooth background noise. The decision of the presence of a source or target is made when the measured signal statistics exceeds a certain predetermined threshold. This decision making may be carried out either automatically, or by a trained human operator.

1.3.2 Localisation

Upon detecting the target, the position of the targets can be determined, in order to make some tactical decisions, and the process is termed as localisation. Generally, localisation involves the estimation

of range and bearing of the targets. The bearing of the target can be deduced using beam forming technique, in the detection stage itself. For passive sonars, the approximate range can be determined by time delay estimates obtained from the received signals at spatially separated points. However, active sonars employ echo ranging, in order to achieve better accuracy. The localisation becomes difficult in the case of manoeuvring targets and in such cases, techniques like Kalman filtering can be applied to estimate the track of the target.

1.3.3 Estimation

The process of estimation involves the computation of statistics used to facilitate target detection, classification and the determination of relevant target parameters. One of the most widely used estimation technique is the spectral estimation which is employed to determine the tonal components in a received waveform, and may be used for both the detection and classification functions. Some other commonly estimated parameters include direction of arrival (DOA), direction of departure (DOD) and the target strength.

1.3.4 Classification

One of the most notable objectives of a sonar system is to interpret the information contained in the received signal in order to identify the signal generating mechanism (also known as the target). The detection and classification problems generally addressed by the sonar systems is achieved by comparing the level of certain statistics, with the assumed or estimated statistics where these statistics are related to the extracted features of the received acoustic space time field. As an example, the

information on the target dynamics can be used to identify a school of fish from a freighter or submarine. Though the classification can be done manually, it becomes a tedious task as the number of targets increases which necessitates the use of automated systems.

1.4 Underwater Acoustic Propagation Effects

The detection, localisation and classification performance of sonar depends on the propagation effects induced by the acoustic channel and various environmental factors. For the active sonars, the performance is also determined by the transmitting subsystems. In the case of passive sonars, the nature and characteristics of the radiated noise can be a factor that influences its performance. The following sections briefly examine the channel effects and other factors that can influence the sonar performance.

1.4.1 Channel Effects

1.4.1.1 Reverberation

When active sonar is used, the sound returned to the receiver generally consists of returns from many other sources besides the target of interest and constitute a dominant source of noise and is termed as reverberation. The sources of reverberation include the surface, the bottom, and the volume of water. The marine life, bubbles, and other inhomogeneities in the water contribute to the volume reverberation. The reverberation phenomenon in the ocean is analogically similar to the scattering of light from the cars headlights in fog or mist. A

high-intensity pencil beam will penetrate the fog and as the main headlights are less directional will result in white-out, where the returned reverberation dominates. Hence, to reduce the effect of reverberation in the ocean, the active sonar needs to transmit in narrow beams.

1.4.1.2 Doppler Effect

Doppler shift refers to the change in the received frequency when there is a relative motion between the source and the receiver. This effect can be an influencing factor in underwater noise propagation. When the source and the receiver recede, the received frequency is decreased and when they are nearing each other the apparent frequency increases. This shift can be appreciable especially when the sonar and the source are on the opposite course.

1.4.1.3 Multipath Rayleigh Fading

The underwater acoustic propagation channels are affected by the multipath effects contributed by the ocean surface, ocean bottom, reflectors along with scatterers having random heterogeneity. This multipath effect results in the formation of additional sub-eigenpath components which override the contribution of the dominant component severely. The interaction of the acoustic signals travelling along multiple paths can induce fluctuations in the received signal's amplitude and phase, a phenomenon known as multipath fading, thereby affecting the SNR of the signal significantly.

1.4.2 Environmental Factors

The propagation of the sound through the ocean water is also affected by various environmental factors, which can in turn affect the performance of the sonar operation. The speed of sound varies in water with density which in turn depends on temperature, pressure, dissolved molecules and salinity. Temperature of the ocean waters varies with depth and for depth ranges between 100 and 300 feet, there is often the formation of thermoclines which divides the warmer surface water layer from the cold bottom waters. This will lead to inaccuracies in sonar predictions as a sound originating on one side of the thermocline tends to be bent or refracted off the thermocline.

In sonar, since the wave propagation speed is a time-varying function of depth and range, with significant dependencies on geographic location and season of the year, the estimations turn out to be cumbersome due to complex refractive phenomenon, especially when the propagating energy interacts with the sea surface or bottom. Motions of the water mass, sea surface, sonar platforms and the targets lead to a wide variety of channel dispersions in time, frequency and angle.

1.5 Noise Sources in the Ocean

The general background noise prevailing in the ocean due to the collective contribution of all the oceanic noise sources is called the ambient noise and is characterised by a broad frequency range. As obvious, the level and nature of ambient noise is highly variable and depends on a number of factors including the climate, wind speed, presence of aquatic organisms, etc. As the ambient noise prevails as a

backdrop for the signals of interest, it can considerably affect the performance of the sonar, and in turn the underwater target classifier.

Noise sources in the ocean, which contribute to the ambient noise, can be either natural or manmade (anthropogenic). Natural sources include marine organisms that produce sound and natural processes such as earthquakes, wind-driven waves, rainfall etc. Anthropogenic noise is generated by a variety of human initiated activities, including shipping, oil and gas exploration, military operations, etc. With the increase in these activities, anthropogenic noises are becoming more pervasive and more intense, increasing the overall level of oceanic ambient noise.

1.5.1 Natural Sources of Ambient Noise

The natural sources of ambient noise can be broadly classified into the following categories:

- Biological sources
- Hydrodynamic sources
- Seismic sources
- Thermal agitations
- Cracking of ice

1.5.1.1 Biological Sources

Oceans contain the greatest diversity of life on Earth. Among these, a large variety of marine organisms like crustaceans, mammals and fishes contribute to the ambient noise backdrop. Noise from such sources exhibits a wide frequency spectrum from 10 Hz to 100 kHz. The individual sounds are repetitive and of short duration. The

individual sounds may merge during choruses that result when a large number of animals are calling, and can prevail substantially over a region for a period of time.

1.5.1.2 Hydrodynamic Sources

There exists a wide variety of hydrodynamic processes in the ocean which can be considered as noise sources, where the noises are being generated due to various physical phenomena like movement of water due to winds, tides, currents etc. Major hydrodynamic sources include surface waves, turbulence, bubbles and water droplets. Surface waves originate due to the wind action on the ocean surface and have a dominant influence on the noise spectrum, especially towards the low frequency region. Turbulence is caused by the movement of layers of water streaming over each other with varying speeds and generally occurs in regions near to coastal areas, straits and harbours. The noise generated by turbulence can have a frequency range of 1 to 100 Hz. The bubbles are yet another source of hydrodynamic noise and the noise is generated during the generation, oscillation, joining and splitting of bubbles. Water droplets can also generate noise as they drop on to the sea surface. While the natural frequency of the noise generated due to the bubbles is found to be inversely proportional to the size of the bubbles, that generated due to the droplets depend on the kinetic energy of the droplets.

1.5.1.3 Seismic Sources

Various tectonic as well as volcanic activities can also contribute to the ambient noise in the sea. Seismic waves that result from such

disturbances in the earths crust can propagate substantial amount of energy to the sea even if the sources of such disturbances are far away from the sea. The spectral characteristics of such noises depend on the magnitude and range of the seismic activity, the propagation path, etc. It has been observed that, in general, the spectral peaks due to the seismic activities occur between 2 and 20 Hz, when the disturbances are waterborne, and in some cases, it can go up to 100 Hz.

1.5.1.4 Thermal Agitations

The effects of thermal agitations of the medium itself can generate a minimum noise level for that medium, even in the absence of any other noise sources. The thermal agitations contribute a minimum ambient noise level especially at the upper frequency limits, around 20 to 30 kHz, of the ambient noise data.

1.5.1.5 Cracking of Ice

Shifting and breaking of ice, along with the interaction of the pack of ice and floes with air and water can contribute considerably towards the ambient noise, especially in the polar regions. The noise originating from these phenomenon exhibit seasonal variations and covers a wide range of frequencies.

1.5.1.6 Other Sources

Other natural phenomena like precipitation and rain can also contribute to the ambient noise. While precipitation generally contributes to noise at a frequency above 500 Hz, heavy rains can raise

the noise level about 30 dB in the 5 to 10 kHz range of the ambient noise spectrum.

1.5.2 Manmade Sources of Ambient Noise

The major manmade noise sources include

- Shipping
- Oil and gas exploration
- Military operations

1.5.2.1 Shipping

At low frequencies, around 500 Hz, shipping is one of the major noise sources in the ocean. Ship generated noise can be attributed to machinery noise, propeller noise and hydrodynamic noise. The effective detectable range for shipping noise can be as high as 1000 miles or even more.

1.5.2.2 Oil and Gas Exploration

The oil and gas industry also contribute to ocean ambient noise through operations related to exploration and production, such as pipe laying, drilling and platform operations.

1.5.2.3 Military Operations

The R & D operations in naval forces across the world are ever increasing and considerably contribute to short term changes in the

ambient noise levels. Such noise generations which include explosions, active sonars, test firing of weapons etc., can even disturb the ocean environment significantly. High-frequency acoustic pulses emitted by active sonars and their echoes can also contribute to ocean noise. Submarines, explosives used in military tests or exercises and firing of torpedoes can also be significant sources of undesirable noise.

1.6 Passive Sonar Underwater Target Classifier

Automatic Target Classifiers are such systems, where machine learning has been employed to categorise the targets. Such systems have been widely used in conjunction with radars due to the advantages provided by the active RF ranging technology. However, in underwater perspective, the application of automated target classification systems has found to be even more challenging, considering the complexity of the channel, the ambient environment and the nature of targets.

Conventionally, the task of such passive sonar target classification is carried out by human experts by analysing the audio visual representations of the retrieved and processed signals. The operators need the skill to identify the type and origin of the noises of interest. However, as the number of signals presented to the operator increases, the identification and classification also turn out to be difficult. There can also be errors in interpreting the results due to the lack of expertise. Together with the extreme alertness required for a prolonged period, operators may feel fatigue and even boredom due to the monotonous nature of the monitoring task, resulting in the decline of detection performance. Due to these confounding factors, there is a

growing need for automatic classification of underwater targets.

Underwater target classification systems have wide applications in both strategic and commercial applications. Classifying and locating military targets like submarines and torpedoes are of predominant importance, in taking appropriate counter measures. Such classification techniques can also be extended to identify commercially important marine biological species.

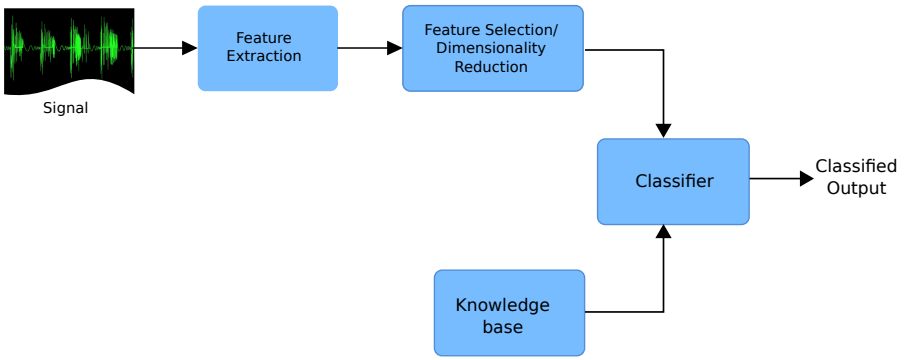


Figure 1.3: General block diagram of underwater target classifier

The general block diagram of an automated underwater target classifier has been depicted in Fig. 1.3. The classifier is fed with the pre-processed signals received from the passive sonar system. The acoustic signals received from the sonar front-end are processed to extract source specific classification clues, called features which may further be processed to get an optimal subset and stored in a knowledge base. During classification, the features derived from the signals are compared with those stored in the knowledge base using some pattern matching algorithms in order to find the closest match.

1.6.1 Feature Extraction

Features are individual measurable parameters and generally represent some distinctive aspect, quality or characteristics of a phenomenon being observed. Such characteristics can be used to identify the phenomenon and often a combination of different features may be required to yield satisfactory results and this collection can be represented as feature vector.

Since the features contain information that describes the unique characteristics of an object to differentiate it from the others, extracting informative, discriminating and independent features is an important task that can considerably affect the efficiency and efficacy of the classification process.

The extracted features should be informative and non-redundant so as to facilitate subsequent learning and generalisation procedures for classification. Feature selection algorithms select a subset from the extracted features to concisely represent the targets without redundancy. The elimination of trivial features ensures dimensionality reduction, thereby reducing the processing, computational as well as storage overheads. The relevant features of known targets identified after the feature selection process, along with their corresponding target labels, are generally compiled and stored as a knowledge base.

1.6.2 Classification Algorithms

Once a knowledge base is created, the features from the incoming test signals are compared with the template features stored in the

knowledge base to identify the targets. The term classification algorithm generally refers to the mathematical or statistical function that does such comparisons and maps the input data to a category. There have been many classification algorithms developed so far and they are broadly categorised into two, the unsupervised and supervised.

In unsupervised systems, such as the k-means clustering, the system is presented with unlabelled training samples and the system needs to find some hidden structure in the unlabelled data. Unsupervised classification, at many instances, can be regarded as a clustering problem which solely is based on the correlations in the natural or latent patterns in the observed data. For a classifier, the particular objective is to find the discriminative patterns that effectively separate the individual classes. However, in a complex environment such as the oceanic environment in which the target signatures are dynamically affected by various complex factors, the clustering techniques may not give well discriminative clusters as there are no explicit mechanisms to drive the clustering algorithms to produce independent clusters of individual targets.

Unlike the unsupervised classifiers, the supervised classifiers have a training stage where the classifier is trained with labelled data, so that the classifier knows the features corresponding to each of the targets. The supervised learning algorithm will analyse the labelled training data and produces an inferred function, such that it can generalise the training data and hence it can correctly determine the class labels for the unseen instances.

Examples of some of the supervised classifiers, which have been used in this thesis include k-Nearest Neighbour (k-NN) Classifier,

Artificial Neural Network (ANN) Classifier and Support Vector Machines (SVMs).

1.7 The Proposed Work

As has already been discussed, feature extraction is one of the important objectives of the sonar processing systems so as to interpret and identify the received signals unambiguously. Traditionally, power spectral analysis and its variants have been used as the technique for feature extraction for such systems. However, being a linear method, and most of the complex signals like underwater noise being generated by nonlinear mechanisms, the use of power spectral analysis turned out to be inadequate. Nonlinear methods may be used in such cases, in order to gain a more complete understanding of the underlying signal dynamics. This thesis, thus explores the feasibility of making use of the Higher Order Spectral analysis, especially the bispectral analysis as a feature extraction technique for the noise sources in the ocean.

The bispectrum, which is based on the third order cumulant sequence of a signal, can play a key role in characterising the nonlinearities of the underlying signal generating mechanisms, especially those containing quadratic nonlinearities. Also, as bispectrum and all higher order spectra for Gaussian process are identically zero, it suppresses the effect of additive white Gaussian noise, while preserving the magnitude and phase information of the original signal, making it a suitable candidate for the feature extraction.

As it is obvious, a robust feature vector should capture the most

invariant characteristics of the underlying signal, even in the presence of noise as well as in situations where the signal undergoes arbitrary scaling and translations, that are quite common in underwater acoustic channels. As an attempt to achieve this, the work also explores the possibilities of incorporating other higher order features like various types of integrated bispectra, due to its attractive properties of low dimensionality, scaling and translation invariance among others.

There have been many attempts to extract the acoustic features leveraging the perceptual attributes of the human auditory system. Many filter banks have been devised mimicking the auditory response of the human ear to different spectral bands. Generally, cepstral coefficients are derived from the spectrum of the signal after the application of such filter banks and are used as the feature vector. This work explores two such filter banks, namely the Mel and the Gammatone filters and bispectral analysis has been applied to enhance the cepstral coefficients derived using them.

Selection of an optimal and non-redundant subset of features is an important step in handling the classification problem, as it can improve the overall performance of the classifier. Various techniques for feature selection and dimensionality reduction have also been explored. The robustness of the generated feature vectors was validated using three different classification algorithms namely, k-Nearest Neighbour, Neural Networks and Support Vector Machines. It has been found that the higher order feature set can yield acceptable classification accuracy and can be used for underwater target classification. In summary, the main contributions of the thesis include:

- Feature vector generation using Higher Order Spectral Analysis for

complex underwater signals

- A novel HOS feature named Bispectral Gammatone Cepstral Coefficients has been proposed and validated
- Realisation of an underwater target classifier using HOS features

1.8 Outline of the thesis

Remaining parts of the thesis, after this introductory chapter, has been arranged as follows: The second chapter provides a comprehensive review of literature related to the area of research. It covers a review of the past works and basic theoretical concepts.

Third chapter covers the methodology adopted for the realisation of the proposed underwater target classifier with HOS features. The fourth chapter provides a detailed description of the Higher Order features and their extraction techniques. A novel HOS feature based on Gammatone filter bank has been proposed and detailed in this chapter.

The fifth chapter furnishes a comprehensive description of the various classification techniques adopted in this thesis. The performance achieved with various classification frameworks has also been presented.

Finally, the sixth chapter gives an overview of the main achievements of the work, proposes areas where future research can be pursued and draws conclusions. This is followed by the list of references and publications brought out as a part of this research work.

1.9 Summary

The need and relevance of intelligent and automated underwater target classifiers have been discussed in this chapter. The basics of sonar, different types and functions of sonars, etc., have been discussed along with the importance of feature extraction techniques from the received sonar signals for the purpose of target classification. The various components and functionalities associated with the target classifiers have also been briefly described followed by a discussion of the proposed work.

Chapter 2

Review of Past Work

A comprehensive review of literature related to the area of research has been presented in this chapter. Various research works reported in the areas of target specific feature extraction and classification have been covered. Conventional feature extraction techniques including spectral estimation and cepstral analysis together with the works and results published in the area of Higher Order Spectral analysis have been consolidated and presented. An in-depth coverage of various higher order techniques like bispectrum, trispectrum, bicepstrum, etc., along with their theoretical framework and applications are also highlighted in this chapter. The works of various researchers on biologically inspired feature extraction techniques have also been consolidated. Various methodologies for feature selection and dimensionality reduction have been reviewed along with different types of classifiers reported in the open literature.

2.1 Background

For decades, trained professionals are being employed for the detection and classification of underwater targets by listening to the sonar data. Development of intelligent systems for classifying the noise sources, in order to assist the sonar operators is one of the hot topics in the field of sonar signal processing. However, the complexity of the underwater environment, makes the task of classification a challenging problem.

The classification process, in general, can be decomposed into estimation of features or signatures of the source from a set of received signals by applying suitable feature extraction techniques, followed by the application of a pattern recognition algorithm on to the estimated source signatures in order to arrive at the class labels.

So far, a number of techniques have been developed and reported in the open literature for extracting features from the emanated sound, and majority of them centres around the classical power spectral analysis. The spectral analysis and its variants like cepstral analysis have gained considerable attention from the research community. However, being a linear method, power spectral analysis cannot fully characterise the nonlinear signal generating mechanisms. This, coupled with the complex nature of the radiated noise signals, which actually are composites of diverse sources, some of which are totally stochastic and nonlinear, calls for the use of non-Gaussian and nonlinear feature extraction techniques, like Higher Order Spectral analysis. The Higher Order Spectrum (HOS) has many attractive properties which, if utilised properly, can make it a potential candidate for the extraction

of features from the underwater noise signals. Upon extracting the features, only a relevant set of the features are selected and fed to a classifier, for the purpose of classification. Thus, a review of existing literatures which throws light on the properties and nature of the underwater noise, various feature extraction techniques, HOS and its applications and different types of classifiers as well as their working principles have been carried out during the design and development of the proposed classifier. The following sections briefly discuss the outcomes of the literature survey.

The review paper by Gordon M. Wenz [1] explores the various aspects of the four basic kinds of underwater noise namely, radiated noise, self-noise, ambient noise and reverberation noise. The basic challenges of research for each type of noise along with the objectives, areas of effort and accomplishments have been explored. The review also addresses the major problems related to noise, such as the noise measurements as well as noise reduction and prevention techniques.

Pieng et al. describe the development of an ambient noise database in the frequency range of 11 - 8300 Hz in [2] using the data collected from the Singapore straits and the surrounding waters. The collected data indicate the presence of high levels of noise due to shipping in the frequency range below 1 kHz and snapping shrimp above 2 kHz. The spectral features of the observed data are also compared with the spectral curves reported in open literature.

A discussion of the trends and the effects of merchant shipping across the world, considering the changes in number, size and propulsion technology has been presented in [3]. The paper observes that there has been an increase in world shipping that results in an

increase in low-frequency ambient noise at an average rate of about 1/2 dB per year. In another work, Potter and Delory [4] have also studied and confirmed the rise in the background noise due to shipping, especially in the northern hemisphere.

The mechanism of generation of Very Low Frequency (VLF) ambient ocean noise has been considered by Nichols in [5]. The trials carried out for 40 days at three deep water sites suggest that the breaking waves are likely to be a source of very low frequency (1 to 20Hz) ambient noise. A discussion of the low-to-mid frequency (LMF) noise characteristics, and the probable mechanisms of their generations have been discussed in Carey [6], along with the effects of micro bubbles and the breaking of waves.

The changes in ambient noise levels with wind speed, has been analysed from the data obtained from moored buoy near Alaska, and is reported in [7]. It has been found that the measured ambient noise level at 900 Hz lies well below the Knudsen curve for a wind speed of about 5 knot. The noise level gradually approaches the Knudsen curve with an increase in the wind speed from 5-10 kn, and above 10 kn, the measured ambient noise level matches the Knudsen curve.

In [8], Brocket et al. discusses the statistical analysis of ambient acoustic noise for detecting the presence of non-linearity and non-Gaussianity, based on bispectral analysis. It has been found that when the time series is considered for the order of a minute, the ambient noise generation, due to distant shipping and wind effects, can be approximated as linear and Gaussian, but for shorter periods of the order of seconds, it is fairly nonlinear and non-Gaussian. Another statistical modelling attempt carried out by Michel Bouvet et al. [9],

shows that the snapping shrimp noise appears to be non-stationary and the background noise is very close to the Gaussian. The study carried out by Webster in [10], also examines the non-Gaussian characteristics of the ambient noise and finds that the performance of the signal processing algorithms designed for Gaussian environment can degrade considerably in a non-Gaussian environment.

2.2 Conventional Feature Sets

Robust features to represent the targets are one of the key requirements for having a dependable target classification system that can be relied upon in different environments. Features can be extracted using time domain or frequency domain techniques. The acoustic features have been well defined during the course of the years, the prominent one being the frequency domain analysis, where the frequency components are extracted and represented in various forms.

Features in time domain are usually quick and easy to implement, as these features do not need any transformations. Various time domain features such as short time energy, short time magnitude, zero crossing rate, autocorrelation, slope sign changes as well as waveform length have found applications in medical [11] and speech signal processing [12], especially when low noise signals are involved.

The performance of the time domain analysis however degrades in the presence of noise as well as when time varying parameters are involved [13]. Therefore, time domain features are generally combined with spectral features to yield better classification results [14, 15].

The complex acoustic ambience of the ocean with several noise

sources and dynamic channel effects makes the task of underwater target classification extremely challenging. In such a scenario, most of the target classifiers rely on frequency domain algorithms for feature extraction [16, 17].

2.2.1 Spectral Estimation

Marple, in [18] presents a summary of several spectral estimation methods - both parametric and non-parametric. Though emphasis is given to parametric time series modelling, non-parametric techniques like classical spectral estimation, autoregressive, ARMA, Prony, Maximum likelihood, Pisarenko and MUSIC are also discussed. Another tutorial paper [19] by Kay and Marple also provide in-depth treatment of many modern spectral estimation techniques.

In [20], Shin et al., present techniques to improve the detection performance of passive emissions from quiet sources in littoral waters, focusing on the full spectrum of the target signature. Various noise emissions corrupted with ambient noise are analysed and the results are discussed.

In [21], the acoustic line spectrum has been used for the detection of torpedo. The spectrum generated by torpedo's blade motion has been analysed to enable its detection by a submarine sonar system. Classification of marine vessels based on the features which are directly extracted from the Power Spectral Density of their acoustic radiated noise has been described in [22]. A bank of 71 files of real radiated ship noise data has been used for the performance evaluation.

The authors, in [23], review various feature extraction techniques

and classification algorithms related to the concept of music information retrieval. A number of different spectral based features like spectral centroid, spectral roll off etc., have been discussed. A discussion on classifiers, feature combination and classifier fusion has also been presented.

2.2.2 Miscellaneous

In [24], a chaotic feature set is proposed for classifying ships from their radiated noise, by making use of the nonlinear regularities in the emitted signals, using chaotic features. The results show that some classes that cannot be classified well by using power spectra can be satisfactorily classified using the proposed chaotic feature. Another study on similar lines to augment existing feature extraction methods has been carried out in [25], based on Fractal analysis. The proposed method includes fractal Brownian motion (FBM) based analysis, fractal dimension analysis along with wavelet analysis. Analysis on six different classes of ships shows that the fractal approaches are effective and can be used to augment traditional features like line and average spectra.

Locke et al., in [26], present a short-time method based on the fractional Fourier transform for detecting frequency modulated narrow-band signals. Such methods can be used for detecting cetacean vocalisations such as dolphin whistles. The proposed method also overcomes some of the disadvantages of the existing fractional Fourier analysis by applying appropriate correction factors.

2.3 Biologically Inspired Feature Set

2.3.1 MFCC

In [27], Childers et al., present a tutorial review of the cepstrum concepts focusing on data processing. The power, complex, and phase cepstra are discussed and shown to be easily related to one another. General problems associated with phase unwrapping, linear phase components, spectrum notching, aliasing, oversampling and extending the data sequence with zeroes are discussed. Concepts of the log spectrum, complex cepstrum and the effects of various forms of liftering the cepstrum are presented.

Molau et al. [28], present a method to compute the Mel-frequency cepstral coefficients directly from the power spectrum of a speech signal. The authors have observed that the omission of filter bank in signal analysis does not affect the word error rate. It simplifies the front end of the speech recognisers by merging subsequent signal analysis steps into a single one and also avoids possible interpolation as well as discretisation problems and results in compact implementation.

A technique incorporating power cepstrum and complex cepstrum techniques for decomposing a composite signal of unknown multiple wavelets overlapping in time has been presented in [29]. The proposed technique makes use of the property of the power cepstrum for efficiently recognising wavelet arrival times and amplitudes and the property of the complex cepstrum in estimating the form of the basic wavelet and its distorted echoes.

Underwater acoustic signal recognition using Mel-Frequency

Cepstral Coefficients (MFCC) and Linear Predictive Coding derived Cepstral Coefficients (LPCCC) have been discussed in [30]. The features extracted frame by frame and using a Bag of Acoustic Words (BoAW) have been considered in this study. Garcia et al. [31], present the development of an automatic recognition system for infant cry using acoustic characteristics obtained from the mel-frequency cepstrum technique and a feed forward neural network. The objective is to classify two types of infant cries, *viz.*, the normal cry and the pathological cry from deaf babies.

In [32], Molla and Hirose present a study on the effectiveness of mel-frequency cepstrum coefficients (MFCCs) and some of their statistical distribution properties such as skewness, kurtosis, standard deviation etc., as the features for text dependent speaker identification. The result shows that the first MFCC degrades the identification competence as it contains more information about speech than the speaker and the statistical distribution parameters enhance the training speed of the neural network.

In [33], the authors discuss the use of Weighted Filter Bank Analysis (WFBA) to increase the discriminating ability of Mel-frequency cepstral coefficients (MFCCs). Two WFBA schemes which differ in the computation of weighting functions, one based on a fuzzy membership function and other based on log energy of each critical band, have been investigated. The experiments for recognition of continuous Mandarin telephone speech show that, by properly adjusting the fuzzy factor, the WFBA scheme based on fuzzy membership function has higher capability in enhancing the discriminating ability of cepstral features.

2.3.2 Gammatone Filter Bank

Performance of the Conventional speaker recognition systems degrades under noisy conditions. Zhao et al. [34], examine Gammatone Frequency Cepstral Coefficient (GFCC), which is based on an auditory periphery model, and show that this feature captures speaker characteristics and performs substantially better than conventional speaker features under noisy conditions. Their findings indicate that GFCC features out-perform conventional MFCC features under noisy conditions.

In their recent work, Valero and Alias [35] have examined the robustness of Gammatone Cepstral Coefficients (GTCC), for non-speech audio classification. They have evaluated the performance over two different corpora of 4 hours each, containing general sounds and audio scenes, respectively. According to them, logarithmic energy compaction in the GTCC computation process yields better performance.

In [36], the author presents a technique which relies on the Gammatone filter bank to a speaker recognition system. The simulation studies over two datasets show that compared to the standard Mel frequency cepstral coefficients and the perceptual linear prediction analysis front ends, the proposed auditory based front-end yielded higher recognition rate.

The work by Schluter et al. [37], investigate the performance of Gammatone features on a large vocabulary speech recognition task. Since the results were competitive, experiments were also carried out by combining Gammatone features with a number of other

state-of-the-art acoustic features and found that a relative performance in word error rate improved by a factor of about 12% compared to the best performing single feature system.

2.4 HOS Based Feature Sets

Although, traditionally, feature extraction techniques mostly rely on power spectral analysis or its variants, there have been a surge of interest in techniques based on Higher Order Spectral Analysis. This section discusses some of the salient works reported in the field of HOS, along with the concepts of moments and cumulants that form the basis for higher order spectra.

The probability density function (PDF) of a random signal $x(k)$ gives an insight into the distribution of the amplitudes of $x(k)$. The shape of the PDF can be characterised by a set of measures called *moments*, which can be defined in terms of the Moment Generating Function (MGF) [38].

Let $w = [w_1, w_2, \ldots, w_n]^T$ and $x = [x(k), x(k+\tau_1), \ldots, x(k+\tau_{n-1})]^T$. Then, the MGF can be defined as [39]:

$$\phi(w) = E[exp(jw^T x)] \tag{2.1}$$

Mathematically, $m_n(\tau_1, \tau_2, \ldots . \tau_{n-1})$, the n^{th}-order moment of $x(k)$ can be defined as the coefficients of Tylor's expansion of the Moment Generating Function. Equivalently, the n^{th}-order moment, m_n, can also be defined as the expected value of the process $x(n)$ multiplied by $(n-1)$ lagged versions of itself [39]. Thus,

$$\begin{aligned} m_1 &= E[x(k)] \\ m_2 &= E[x(k)x(k+\tau)] \\ m_3 &= E[x(k)x(k+\tau_1)x(k+\tau_2)] \\ m_4 &= E[x(k)x(k+\tau_1)x(k+\tau_2)x(k+\tau_3)] \end{aligned} \tag{2.2}$$

In general, the n^{th} order moment is given by,

$$m_n(\tau_1, \tau_2, \cdots, \tau_{n-1}) = E[x(k)x(k+\tau_1)\cdots x(k+\tau_{n-1})] \tag{2.3}$$

The first moment, the *mean*, gives a measure of the location of the PDF, while the second moment, the *variance*, gives a measure of the spread of the PDF. The third and fourth order moments are called *skewness* and *kurtosis*, which give a measure of the asymmetry and sharpness of the PDF, respectively.

The Cumulant Generating Function (CGF) is defined as the natural logarithm of MGF.

$$\kappa(w) = \ln E[exp(jw^T x)] \tag{2.4}$$

and the coefficients of the Taylor's expansion of the CGF are termed as the *cumulants* [38].

From Eq. (2.1) and (2.4), it is clear that the moments and cumulants

are closely related.

$$\begin{aligned}
c_1(\tau) &= m_1(\tau) \\
c_2(\tau) &= m_2(\tau) - (m_1)^2 \\
c_3(\tau_1,\tau_2) &= m_3(\tau_1,\tau_2) - m_1[m_2(\tau_1) + m_2(\tau_2) + m_2(\tau_2-\tau_1)] + 2(m_1)^3 \\
c_4(\tau_1,\tau_2,\tau_3) &= m_4(\tau_1,\tau_2,\tau_3) - m_2(\tau_1)m_2(\tau_3-\tau_2) - m_2(\tau_2)m_2(\tau_3-\tau_1) - \\
&\quad m_2(\tau_3)m_2(\tau_2-\tau_1) - m_1[m_3(\tau_2-\tau_1,\tau_3-\tau_1) + \\
&\quad m_3(\tau_2,\tau_3) + m_3(\tau_1,\tau_2)] + (m_1)^2[m_2(\tau_1) + m_2(\tau_2) + m_2(\tau_3) + \\
&\quad m_2(\tau_3-\tau_1) + m_2(\tau_3-\tau_2) + m_2(\tau_2-\tau_1)] - 6(m_1)^4
\end{aligned} \tag{2.5}$$

If $x(k)$ is a zero mean process, $m_1(\tau) = 0$, then the second and third order cumulants are identical to second and third order moments. One of the notable features of the cumulants is that, for a Gaussian process, all cumulants of order greater than two are identically zero and this helps in distinguishing a non-Gaussian process from a Gaussian one. Also, cumulants are symmetric in their arguments, that is one can interchange the arguments of the cumulant without changing the value of the cumulant. That is, $c_n(\tau_1,\tau_2,\ldots,\tau_{n-1}) = c_n(\tau_{i1},\tau_{i2},\ldots,\tau_{i(n-1)})$ where $(i1,\ldots,i(n-1))$ is a permutation of $(i,\ldots,n-1)$. For example, third order cumulants exhibit the following symmetries:

$$\begin{aligned}
c_3(\tau_1,\tau_2) &= c_3(\tau_2,\tau_1) = c_3(-\tau_2,\tau_1-\tau_2) \\
&= c_3(-\tau_1,\tau_2-\tau_1) = c_3(\tau_2-\tau_1,-\tau_1) \\
&= c_3(\tau_1-\tau_2,-\tau_2)
\end{aligned} \tag{2.6}$$

Detailed descriptions of the properties of the cumulants can be found in [38, 40].

The higher order spectra is defined in terms of higher order

cumulant sequence. The n^{th} order spectrum is defined to be the $(n-1)$-dimensional Fourier transform of the n^{th} order cumulant sequence. The bispectrum is obtained by taking the two dimensional Fourier transform of the third order cumulant. The general motivations behind the use of the bispectrum include deviations from normality, phase estimation, and the detection and characterisation of nonlinear mechanisms that generate the time series [40].

Bispectrum can be estimated using parametric and non-parametric methods. In [41] a parametric method for bispectrum estimation of an autoregressive (AR) model driven by non-Gaussian white noise is presented. According to the authors, the parametric method would provide bispectrum estimates of higher fidelity and resolution.

The authors, in [42], present a general approach for evaluating the principal domain of the discrete bispectrum of a stationary, band limited random signal. The basic statistical issues of testing for non-zero bispectral structure are reviewed and a test has been proposed which can be used to provide inferences about the joint density of the signal with relevant prior knowledge.

Hinich et al. in [43], examine the ability of the bispectrum to detect a non-Gaussian time series when that time series is corrupted by non-Gaussian (or Gaussian) noise. The effect of detection performance for the variation of different parameters such as sample size and average SNR has been investigated and it has been found that the bispectrum performs well compared to other energy detectors in detecting non-Gaussian signals.

The bicoherence, which is a normalised form of bispectrum, has some attractive properties. Bispectrum normalisation scheme suggested by

Kim and Powers has been examined in [44]. The advantage of the Kim and Powers method is that it produces a result guaranteed to be bounded by zero and one, and may be easily interpretable as the fraction of signal energy due to quadratic coupling. However, Hinich and Wolinsky, in this paper show that wrong decisions can be obtained by blindly relying on the Kim and Powers normalisation, as this normalisation depends on the resolution bandwidth of the sample bispectrum.

A nonparametric bicoherence estimation technique which is resistant to transient contamination has been suggested in [45]. The authors suggest a stepwise outlier rejection algorithm based on the assumption that the transients often have different probability distributions when compared to the original signal.

Richardson and Hodgkiss in [46], discuss the estimation of the bispectrum and bicoherence of underwater acoustic signals. It is demonstrated how the bispectrum estimate can be used to detect non-Gaussianity, nonlinearity and harmonic coupling. Actual data taken from a freely drifting swallow float and also from an element of a moored acoustic array are used for the analysis.

The bispectral analysis have also been utilised in fields like fault detection of rotating machinery [47], characterisation of failure mechanisms of the pipes for the oil industry [48], identification of damages in corroded and non-corroded galvanized steels [49]. Apart from these, it has also been found useful in music information retrieval [50, 51], and for analysing bio medical signals [39].

The work by McLaughlin et al. [52], describe the use of normalised bispectrum to detect the presence of quadratic phase coupling arising as a result of nonlinear mechanisms, for passive sonar data. Two

techniques, one based on a segment and average approach and other based on single record phase coupling detector have been analysed.

An algorithm for extracting features from radiated noise of underwater targets using bispectrum has been presented in [53]. Features were extracted after bispectrum estimation on three target signals and low-dimension feature vectors were obtained.

The extracted features were passed into the radial basis function (RBF) neural network classifier. The results confirm the usability and robustness of the proposed feature set.

In [54], the authors investigate the potential of the bispectrum as a robust feature for speaker identification, in varying noise conditions. The results were compared against cepstrum features. While the cepstrum gave the best results using clean data, its performance falls off sharply in varying conditions with cross-conditions being the most difficult for the cepstrum. The bispectrum features from the principal region were used for generating the feature set. By averaging and smoothing, the bispectrum did very well for additive Gaussian noise.

The wavelet bispectrum and bicoherence, which provide the instantaneous bi-amplitude and biphase information has been used to form a set of 23 features for the nonlinear analysis of wheezes [55]. The proposed feature set was evaluated on a dataset of wheezes, acquired from adult patients with diagnosed asthma and Chronic Obstructive Pulmonary Disease (COPD), and found to yield satisfactory results.

Most bispectrum based signal reconstruction techniques rely on estimated values on the bispectral plane, which are not asymptotically unbiased and perform poorly with high noise levels. Sundaramoorthy

et al. [56], suggest two magnitude reconstruction approaches that use bispectrum samples drawn only from regions, where the estimates are asymptotically unbiased, in the bispectral plane. The performance of the two proposed approaches, namely Closed-Form Approach and Least Squares Approach (LSA), has been verified using simulation studies.

In [57], the authors present analytic performance evaluation of the complex cepstrum and bicepstrum (i.e., cepstrum of the bispectrum) methods. Explicit expressions of the bias and variance of cepstrum parameters are derived.

The authors, in another work [58], have extensively evaluated the analytic performance of the complex cepstrum, power cepstrum, bicepstrum and power bicepstrum by providing approximate expressions of the bias and variance of the cepstrum parameters due to the presence of Gaussian noise with moderate signal-to-noise ratios. Monte Carlo simulation results suggest that the improved performance of the bicepstrum is a function of SNR, number of records and samples per record.

In [59], the authors suggest a novel parametric estimation technique for bicepstrum (cepstrum of the bispectrum) making use of both second and third order statistics and have successfully applied it to time delay estimation and nonminimum phase system identification. The performance of this method is compared to that of the second-order least-squares methods, for different noise conditions and lengths of data and found to provide larger performance gain when the noise sources are Gaussian and spatially correlated with unknown correlation function.

2.5 Feature Selection Techniques

Many algorithms and techniques have been devised for feature selection, to find a subset of features that would maximise the relevancy while minimising the redundancy. Feature selection techniques can be broadly grouped into three, *viz.*, wrapper, embedded and filter approaches. The wrapper and embedded approaches are classifier dependent, while the filter approach is classifier independent [60].

Wrapper methods consider the subset selection as a search problem, where different combinations are prepared and each candidate subset is evaluated using the training/validation accuracy of a particular classifier. However, this may lead to considerable computational overheads, and may produce subsets that are overly specific to the classifier used. Embedded approaches overcome this disadvantage of wrapper approaches by incorporating the knowledge about the specific structure of the classifier. Even though, this makes the embedded approaches far less computationally intensive, the learning part and the feature selection part become inseparable, making it more dependant on the classifier model assumptions.

In contrast, the filter techniques assess the relevance of the features by examining only the intrinsic properties of the data, independent of the classification algorithm [61]. Filter methods define a heuristic scoring criterion by evaluating the statistics of the data and use this score to determine the relevance of the features. The filters have advantage in terms of computational speed and can be easily scaled to very high-dimensional datasets.

Feature selection methods can also be divided as univariate or multivariate [62]. Univariate methods use univariate statistics, where each feature is considered separately ignoring the feature dependencies, to rank the features. Multivariate feature selection methods employ multivariate statistics for feature ranking and incorporate feature dependencies to some degree and can preserve relevant features that might be discarded by univariate methods.

2.5.1 Fisher Ratio

The Fisher Ratio is a well known univariate method, that comes under the filter approach to rank the features. The Fisher Ratio depends on the interclass difference and the intraclass spread or variance, and is defined as the ratio of the interclass difference to the intraclass spread [63]. Let the mean and variance of the l^{th} feature of the classes C_i and C_j are denoted by $\mu_{i,l}$, $\mu_{j,l}$, $\sigma^2_{i,l}$ and $\sigma^2_{j,l}$ respectively. Then the Fisher Ratio of the l^{th} feature is defined as,

$$\lambda_{i,j,l} = \frac{(\mu_{i,l} - \mu_{j,l})^2}{(\sigma^2_{i,l} + \sigma^2_{j,l})} \tag{2.7}$$

This criterion measures the difference of two means normalised by the sum of variances. The definition given in Eq. (2.7) is for a two class problem. It can be further extended to multiclass problems, where there are C classes. In such cases, the average class separability measure λ_l of the l^{th} feature [64] is given by,

$$\lambda_l = \frac{\sum_{i=1}^{C} \sum_{j=1}^{C} \lambda_{i,j,l}}{C(C-1)} \quad i \neq j \tag{2.8}$$

A high value of λ_l indicates less intraclass spread and more interclass

difference, and hence represents a strong discriminative feature. Feature selection can be accomplished by retaining all the features having the λ_l value above a predetermined threshold.

2.5.2 Joint Mutual Information

Joint Mutual Information (JMI) is a multivariate feature selection algorithm and can be used to select a relevant and non redundant subset of features from the raw feature set.

The Mutual Information MI [65] of two random variables X and Y, measures the amount of information shared by X and Y and can be expressed as,

$$MI(X;Y) = H(X) - H(X \mid \mathtt{Y}) = H(Y) - H(Y \mid \mathtt{X}) \tag{2.9}$$

where $H(X)$ is a measure of the *a priori uncertainty* of X and $H(X \mid \mathtt{Y})$ measures the *conditional a posteriori uncertainty* of X after Y is observed. $MI(X;Y)$ quantify the reduction in the uncertainty of X if Y has been observed. Thus, a large value of MI signifies high correlation between the two variables and a zero value indicates that the variables are independent.

The two advantages of MI that make it unique among the other dependency measures include capacity to measure any kind of relationship between the variables, as it makes no assumptions about the nature of the relationship between the variables and its invariance under space transformations such as translations, rotations, as well as any transformation that preserve the order of the original elements of the variables [66].

The concept of mutual information can be easily expanded to include more than two random variables to get the Joint Mutual Information (JMI) between the variables $(X_1, X_2, \cdots, X_N)$ and Y as given by,

$$JMI(X_1, X_2, \cdots, X_N; Y) = \sum_{i=1}^{N} MI(X_i; Y \mid X_{i-1}, X_{i-2}, \cdots, X_1) \tag{2.10}$$

JMI provides a measure of the decrease in the uncertainty of Y, on the basis of the information provided by the feature vector $X_1, , X_2, \cdots, X_N$.

Even though MI, as a feature selection criterion, selects subsets of the informative features, the selected features can contain redundancy. In real world applications, it can be expected that some of the features may be dependent on each other, and in such cases, extending the concept of MI, to Joint MI avoids the redundancy in the feature space and provides an optimal subset that contains not only the most relevant but also the least redundant features.

The Mutual Information (MI) measure is suitable for assessing the "information content" of features in complex classification tasks, where other methods based on linear relations (like the correlation) are prone to mistakes. The applicability of the mutual information criterion to evaluate a set of features and to select an informative subset from this feature set has been investigated in [67]. An algorithm is proposed that is based on a greedy selection of the features and ranking them according to their MI with respect to the class discounted by a term that takes the mutual dependencies into account. Also, some examples are presented from different classification areas where the method is satisfactory.

In [68], Guyon presents a tutorial overview of variable selection and feature selection algorithms and covers a wide range of aspects like

feature construction, feature ranking, multivariate feature selection etc. Subset selection methods including wrapper methods and embedded methods are also introduced.

In [69], the authors illustrate the importance of feature selection in combining features from different data models, and demonstrate the potential difficulties in performing feature selection in small sample size situations, due to the curse of dimensionality.

The authors in [70], investigates the selection of good features based on mutual information criterion and a new method called minimal-redundancy-maximal-relevance criterion (mRMR) has been devised. A two-stage feature selection algorithm by combining the mRMR and more sophisticated feature selectors has been proposed. Extensive experimental comparison of the proposed algorithm has been carried out with three different classifiers and four different data sets. The results confirm that mRMR leads to promising improvements on feature selection and classification accuracy.

2.6 Existing Classification Systems

The following sections briefly review the theory behind the working of the 3 classifiers, namely the k-Nearest Neighbour, Artificial Neural Networks and Support Vector Machines that have been used in this thesis. Some of the applications of these classifiers that have been reported in the open literature is discussed. A brief review of other existing classification systems like Hidden Markov Model has also been provided.

2.6.1 k-Nearest Neighbour Classifier

The k-Nearest Neighbour (k-NN) technique is a simple and intuitively appealing method for classification which works on the premise that two instances are less similar if they are far apart in the instance space, as defined by some distance function. Two nearly situated instances are generally more similar and may belong to the same class [71]. That is, the classification of unknown instances can be achieved by relating them to the known instances based on some distance or similarity function like the Euclidean distance. A description of some of the popular similarity measures, which are being widely used, have been discussed in [72].

Let each instance be represented with a feature vector with n features. ie., each instance can be represented as a point in a n dimensional space. Let there be M instances in the training set and d_i be the distance computed between the unknown instance and the i^{th} instance, where $i = 1, \cdots, M$. The k-Nearest Neighbour classifier selects k instances having smaller d_i values [73]. The classification rule is to assign the unknown instance to the majority category label of the selected k- nearest training instances. The best choice of k depends upon the data and can be selected by various heuristic techniques and usually chosen to be odd, so as to avoid any ties. A demonstration of how the choice of k can affect the classifier performance is illustrated in Fig. 2.1. When $k = 3$ the test instance represented by the star symbol is labelled as circle, while a choice of $k = 7$, leads to its classification as a triangle.

Though, a larger value of k generally reduce the effect of noise on

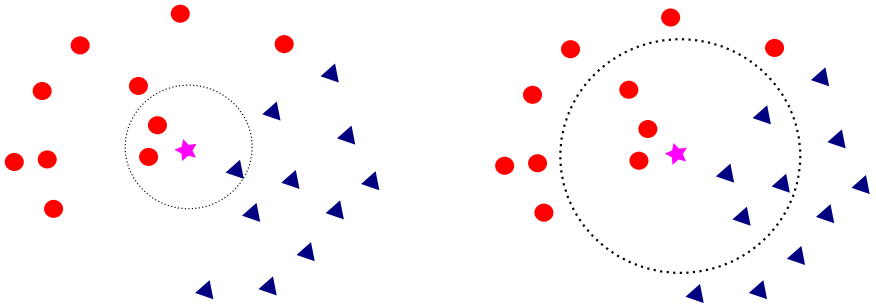


Figure 2.1: How the choice of k can affect the classification

the classification, it does make boundaries between classes less distinct. It can be shown that the k-Nearest Neighbour rule becomes the Bayes optimal decision rule as k becomes infinity [74].

An application of the k-NN classifier for the detection and classification vehicles from their acoustic signals has been described in [75] using MFCC as feature set. The paper also compares the result with a neural network classifier. EEG brainwave behaviour due to RF exposure has been studied using k-NN classifier in [76]. k-NN is used to classify the different groups for the exposure sessions and also to prove that there are significant difference in the EEG signals due to the RF radiation.

2.6.2 Neural Network Classifiers

An Artificial Neural Network is an information processing system that have been developed as generalisations of mathematical models of human cognition. The basic processing elements of ANN are called *artificial neurons*, or simply the neurons which mimics, to a certain extent, the behaviour of biological neurons in the brain. These neurons

are interconnected to form a network as in the case of the biological nervous system [77].

Each neuron receives signals from other neurons through weighted and directed links. Typically, these connection weights that modulate the effect of the associated input signals approximately model the effects of the biological synapses. Each neuron is also characterised by a transfer function that mimics the nonlinear characteristics exhibited by biological neurons. The output y_i of the i^{th} neuron is generated [78], by computing the weighted sum of its n input signals, x_j, $j = 1, 2, \cdots, n$ and applying the transfer function $f_i(\cdot)$, as shown in Fig. 2.2.

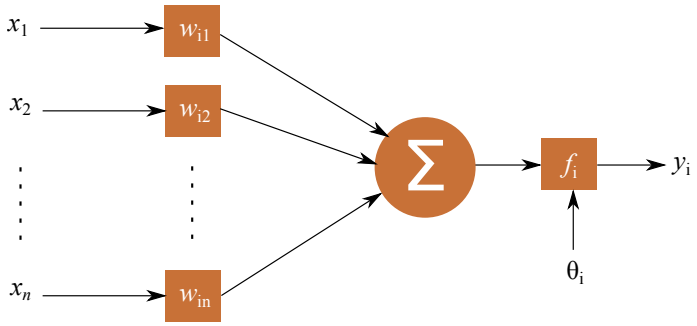


Figure 2.2: Artificial Neuron

That is,

$$y_i = f_i\left(\sum_{j=1}^{n} w_{ij}x_j - \theta_i\right) \tag{2.11}$$

where, w_{ij} is the connection weight between nodes i and j and θ_i is the threshold (or bias) of the node.

The architecture of ANN refers to the pattern of connection between the neurons. ANNs can be viewed as weighted directed graphs in which artificial neurons are nodes and directed edges with weights

serve as the connections between the neurons [77]. ANNs can be divided into feed-forward and recurrent classes based on the connection pattern. In a recurrent network, there exists loops because of feedback connections and loops are absent in the case of feed-forward networks. In general, feed-forward networks produce only one set of response values for a given input and the response is independent of the previous network state. Thus feed-forward networks are static and memory-less, compared to recurrent networks, which are dynamic systems. Examples of feed-forward networks include single layer perceptron, multi-layer perceptron (MLP) and Radial Basis Function Nets, where as Kohonen self-organising maps (SOMs), Competitive and Hopfield networks are examples of recurrent networks. Among these, MLPs are one of the most common NN architecture used for classification tasks.

Typically, MLPs are feed forward nets with an input layer (consisting of the input variables), zero or more hidden (intermediate) layers and an output layer, as shown in Fig. 2.3.

Generally, the neurons in a layer share same characteristics like the transfer functions. One can choose the number of neurons in each layer, number of hidden layers and the transfer functions of each layer. The complete network therefore represents a very complex set of interdependencies which may incorporate any degree of nonlinearity, allowing very general functions to be modelled in many areas including pattern classification, pattern association, regression, clustering and constrained optimisation.

The learning capability of an artificial neuron is achieved by adjusting the connection weights iteratively in accordance with some learning algorithm, so that the trained network can perform certain

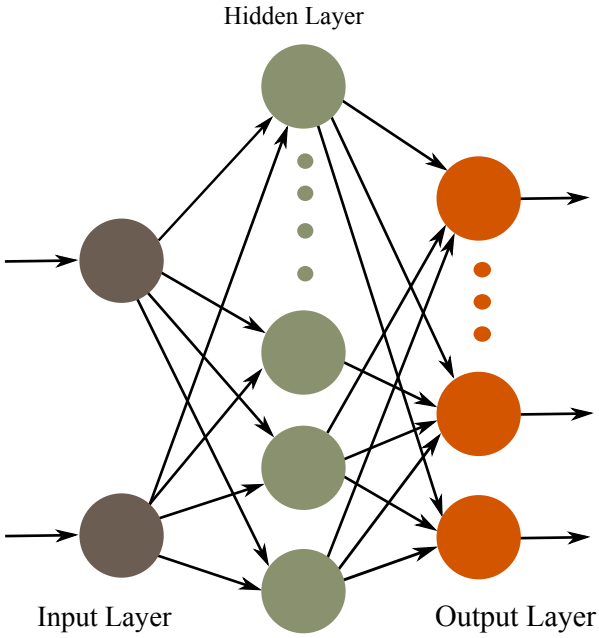


Figure 2.3: Multilayer Perceptron Architecture

tasks. Learning in ANNs can roughly be divided into supervised, and unsupervised. Supervised learning is based on direct comparison between the actual output of an ANN and the desired output based on some error function, such as the total mean square error, between the actual and the desired outputs summed over all available data. The connection weights are iteratively adjusted in order to minimise the error, using a gradient descent-based optimisation algorithms such as back propagation (BP) [78]. Unsupervised learning utilises the correlations among the input data, and is used in situations where no information on the correct output is available.

The design of an intelligent system for classifying marine vessels using Probabilistic Neural Network (PNN), based on passive sonar listening has been addressed in [79]. An Autoregressive model with appropriate

order and coefficients has been developed for the acoustic radiated noise of ships and has been used as the feature vector. The results of evaluating the performance of the proposed system indicate that this method is successful in classifying vessels into three separate classes - heavy ships, medium ships, and boats.

The performance of ANN-based helicopter sound detection systems has been evaluated in [80]. Linear Prediction Coefficients (LPCs) along with other forms of the LPC parameters such as reflection coefficients, cepstrum coefficients, and Line Spectral Pairs (LSPs) have been used as feature vectors for the training and testing of the ANN detectors. The analysis shows that the performance of the ANN detectors can be improved if the wavelet transform is applied for de-noising the signal prior to the feature extraction stage, or if the detection system is trained using a combination of clean as well as noisy signals.

A neural network classifier for classifying underwater mines and mine-like targets from the acoustic backscattered signals has been proposed in [63]. The proposed subband-based classification system consists of a feature extractor using wavelet packets in conjunction with linear predictive coding (LPC), a feature selection scheme, and a back propagation neural-network classifier.

In [81], the authors describe object recognition from underwater images using a hierarchical neural tree system, in which each level is composed by a neural-based classifier. Such neural tree architecture has the advantage of not requiring any a priori information about the network structure. Here the input image is divided into small regions and a neural tree is used to classify each region into different object classes.

2.6.3 SVM Classifiers

Introduced in 1995 by Vladimir Vapnik and colleagues, the Support Vector Machines [82] has become one of the most popular supervised learning techniques which can be used for classification and regression analysis. Initially developed for solving two-group classification problems, it was later extended to handle multiclass scenarios.

The SVMs conceptually implements the idea that "if there exists two input vectors that cannot be separated linearly, they can be separated by non-linearly mapping them into a higher dimension feature space, where a linear decision surface can be constructed." This helps in separating the data, which was otherwise inseparable [83].

Consider the two classes, indicated by circles and triangles in Fig. 2.4(a). Since the classes are linearly separable, there can be many lines or hyperplanes [84], that can define a boundary that separates the two classes as in Fig. 2.4(a). However, one can define a hyperplane that maximises the distance between the hyperplane and the nearest data point of each class and is termed as the optimal hyperplane. Thus, the optimal hyperplane maximises the margin, as shown in Fig. 2.4(b).

Consider two linearly separable classes, with a set of training vectors $(\mathbf{x}_1, y_1), \cdots, (\mathbf{x}_l, y_l)$, where $\mathbf{x}_i \in \mathbb{R}$ and $y_i \in \{-1, +1\}$. A linear function of the form

$$f(\mathbf{x}) = \mathbf{w}^T\mathbf{x} + b \tag{2.12}$$

can be used to correctly separate the classes represented by the training

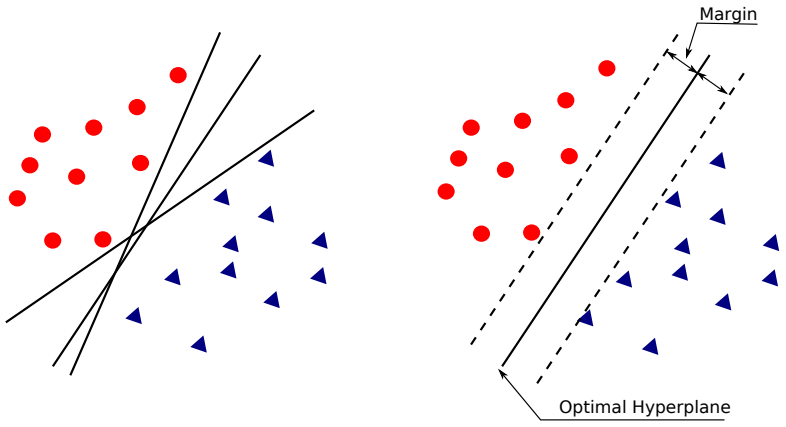


Figure 2.4: Illustration of a simple linear classification scenario and optimal hyperplane

vectors. ie., for every training vector $\mathbf{x}_i$,

$$\begin{aligned} f(\mathbf{x}_i) &\geq 0 \quad for \quad y_i = +1 \\ &< 0 \quad for \quad y_i = -1 \end{aligned} \tag{2.13}$$

Specifically, the two classes are separated by the hyperplane $f(\mathbf{x}) = \mathbf{w}^T\mathbf{x}+b = 0$. However, there may exist, many such hyperplanes that can separate the two classes and one needs to find out the optimal hyperplane which maximises the separating margin between the classes. Solution to this optimisation problem can be obtained by minimising the following cost function [85]:

$$J(\mathbf{w}) = \tfrac{1}{2}\,\mathbf{w}^T\mathbf{w} = \tfrac{1}{2}\,\|\mathbf{w}\|^2 \tag{2.14}$$

where the separability constraints are given by,

$$\begin{aligned} \mathbf{w^T x_i + b} &\geq +1, \quad for \quad y_i = +1 \\ \mathbf{w^T x_i + b} &\leq -1, \quad for \quad y_i = -1 \end{aligned} \tag{2.15}$$

Equation (2.15) can be written compactly as:

$$y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1; \quad i = 1, 2, \ldots, l \tag{2.16}$$

In many cases, the classes may not be completely separable by the hyperplanes, and in order to address this scenario, Cortes and Vapnik [82] introduced the concept of slack variables ξ_i, that provides a measure of the misclassification error. Using the slack variable, the separability constraints in Eq. (2.16) can be relaxed as follows:

$$y_i(\mathbf{w}^T\mathbf{x}_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0; \; i = 1, 2, \ldots, l. \tag{2.17}$$

Defining a penalty function, $F(\xi) = \sum_{i=1}^{l} \xi_i$, and a regularisation parameter C, the cost function in Eq.(2.14) can be modified as:

$$J(\mathbf{w}, \xi) = \|\mathbf{w}\|^2 + CF(\xi) \tag{2.18}$$

where $C > 0$ is a constant that sets the relative importance of maximising the margin and minimising the amount of slack.

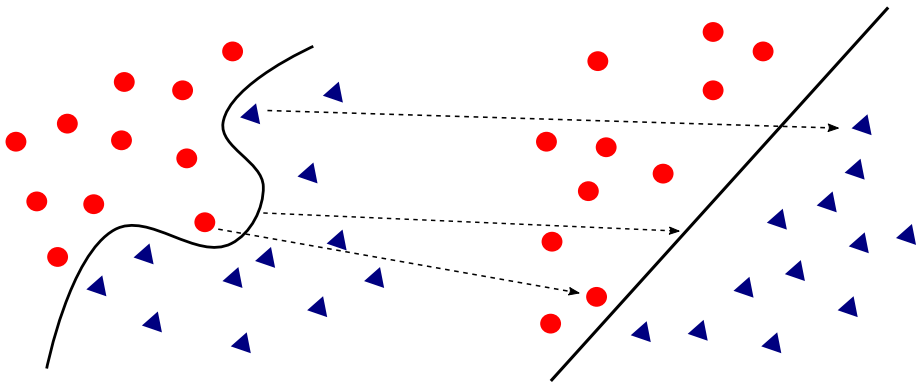


Figure 2.5: Nonlinear to Linear Mapping

The concept of linear SVM can be extended to handle nonlinear

classification scenarios by a technique termed as kernel mapping [86]. In the nonlinear case, the samples in the input space cannot be separated by any linear hyperplanes as shown in Fig. 2.5. A nonlinear operator $\phi(\cdot)$ is applied to map the input pattern into a higher dimensional feature space, where the classes are linearly separable. However, in practice, the mapping function is not directly employed, instead a kernel function $K(\cdot,\cdot)$ which implicitly defines $\phi(\cdot)$ is made use of.

Consider the case of mapping an n-dimensional feature space into an m-dimensional feature space, using the mapping function $\phi(\cdot)$ such that, $x \rightarrow \phi(x)$, $x \in X$, $\phi(x) \in H$, where X is the original n-dimensional feature space and H is the new m-dimension feature space. x is an arbitrary vector in X and $\phi(x)$ the corresponding vector in H.

The kernel function $K(x,y)$, for all $x,y \in X$ can be defined as $K(x,y) = \langle\phi(x),\phi(y)\rangle$, where $\langle\phi(x),\phi(y)\rangle$ denotes the inner product of $\phi(x)$ and $\phi(y)$.

A number of kernels that can be used along with Support Vector Machines have been proposed by various researchers. However the choice of the kernel function highly depends on the nature of the data and the parameters to be modelled. The most popular kernel functions are: linear, polynomial and Radial Basis Function. A polynomial kernel function helps to model the feature space up to the order of the polynomial whereas a Radial Basis Function (RBF) allows to model hyperspheres while the linear kernel allows only the modelling of hyperplanes.

The RBF kernel function can be represented as [71],

$$K(x,y) = \exp\left(-\gamma\|x-y\|^2\right) \tag{2.19}$$

where the classifier parameter γ plays a major role in the performance of the kernel, and can be adjusted optimally in order to yield better classifier performance.

The concept of kernels as a way of generating nonlinear boundaries has been discussed in [87]. Practical implementation of SVMs for real life problems, including the effect of the SVM kernel parameters and their effective selection have also been illustrated. A discussion on SVM training algorithms and the use of SVMs for unbalanced data have also been given. The accuracy of an SVM can severely degrade if the data are not normalised, and as such, various normalisation techniques have been reviewed and analysed.

An approach for automatically tuning the kernel parameters of an SVM has been discussed in [88]. This proposed approach is based on the possibility of computing the gradient of various bounds on the generalisation error with respect to these parameters. Techniques are also proposed for smoothing these bounds without affecting their accuracy. These smoothed gradients allow the gradient descent to search the kernel parameter space, improving the performance and reducing the complexity of the solution, making it possible to design highly complex and tunable kernels.

In [89], an improved grid-search algorithm is proposed that can choose the optimal parameters for an SVM. The algorithm makes use of a big variable step size to search for the parameter values in a large range and optimal parameters are obtained. Then small steps are used to search near the optimal parameters to fine tune the results. Analysis shows that the improved grid-search algorithm can reduce the SVM classifier's computational complexity effectively and improve its

performance as well as classification accuracy.

An analysis of classification algorithms of underwater targets from the backscattered acoustic signals has been considered in [90]. Several different classification algorithms Viz. K-nearest neighbour (K-NN), neural networks, probabilistic neural networks (PNNs), Support Vector Machines (SVMs), multivariate Gaussian classifier are tested and benchmarked not only for their performance, but also to gain an insight to the properties of the feature space using a wideband 80-kHz acoustic backscattered data set. The performance of the PNN and multivariate Gaussian were not as good as other classifiers and the SVM classifier has shown to give good classification results.

The SVM classifier has been found to be used in a variety of applications including the detection of micro-calcification clusters in digital mammograms [85], text classifications [91] and content-based image retrieval systems [92].

2.6.4 Other Important Classifiers

An expert system is a system that intends to act or behave generally like a human expert. Expert systems are basically designed to solve very complex and rigid problems not by following the procedure of a developer as is the case in conventional programming but by reasoning about the knowledge in the same manner as an expert does. In [93], the analysis of the expert system attributes along with the general architecture is discussed. The paper also describes the engineering processes in expert system development and the possible techniques that can be applied to expert system development.

Lourens and Coetzer in [94], discuss the detection of the mechanical features like propeller shaft speed, number of propeller blades and type of propulsion by analysing the underwater acoustic noise. Classification of the ships into different classes is carried out using a small expert system. The development of an autonomous sonar classification expert system for AUVs is investigated in [95] by Brutzman, et al. The use of Geometric analysis techniques and an expert system for heuristic reasoning has been examined in this paper. Classification of sonar contacts is performed by comparing the attributes of detected objects with predetermined attributes of known objects of interest.

Hidden Markov Models have been widely used for different classification scenarios, especially for applications involving speech processing. In [96], Rabiner presents an in-depth tutorial overview of Hidden Markov Models and its applications. The paper discusses Discrete Markov Models and its extension to Hidden Markov Models. Basic elements of HMM are explained in great detail. Various types of HMMs, comparison of HMMs and optimisation criteria are also discussed. The paper concludes by discussing the implementation of a real world speech recognition system using HMM.

Automatic frequency line tracking (FLT) is a crucial function for sonar operators for detecting, classifying or tracking moving targets of interest. In [97], the authors propose an automatic tracking technique for a frequency line with HMM using Viterbi and Forward-Backward algorithms. The notion of probabilistic integration of the spectral power has also been applied to the classical hidden Markov model (HMM) algorithms in order to achieve acceptable performance levels.

An integrated hybrid hidden Markov model and neural network

(HMM/NN) classifier has been proposed in [98], for sonar signal classification. In the proposed classifier, a left-to-right HMM module is employed and the system has been validated using sonar biological signals.

Fuzzy rule-based classifiers are a popular counterpart of fuzzy control systems. There are numerous practical designs for such classifiers, like neuro-fuzzy models, genetic algorithm based fuzzy systems, etc. In [99], Kuncheva summarises the architectural as well as theoretical aspects of fuzzy if-then classifiers in a view to understand them better. A rigorous discussion about the fuzzy classification, exact match of the classification boundaries and universal approximations with fuzzy classifiers are discussed along with the equivalence between fuzzy rule-based and non-fuzzy classifiers.

In [100], the authors present a fuzzy logic based system to detect and classify a series of underwater acoustic transients. The proposed classifier is based on the fuzzy logic theory which can incorporate human knowledge into the process of inference. Two classes of real life underwater acoustic transients have been used to evaluate the performance of the proposed classifier.

Pal et al. [101], in their work, attempt to build a fuzzy version of the multilayer perceptron using the gradient-descent back propagation algorithm, by incorporating concepts from fuzzy sets at various stages. The proposed fuzzy neural network model is capable of handling input features presented in quantitative and also in linguistic form. The effectiveness of the proposed model is demonstrated using a speech recognition problem where the classes have ill-defined, fuzzy boundaries. A comparison is made with the standard Bayes classifier

and the conventional multilayer perceptron, and the performance of the proposed model is found to be better.

2.7 Summary

This chapter presents a review of pertinent open literatures available in the topic covered by the thesis. A detailed survey on the higher order spectral analysis and their applicability in various domains has been included. Various feature selection algorithms, classification of targets using statistical and expert system classifiers, Hidden Markov Model classifiers, Neural Network and Support Vector Machine classifiers, etc. have been reviewed for their adaptability in the proposed classifier design.

Chapter 3

Methodology

This chapter addresses the methodology adopted for the realisation of the proposed target classifier, which primarily involves the extraction of higher order feature vectors from the emanated target signals and the feature selection procedures to identify an optimal feature subset. The feature extraction steps mainly concentrate on Higher Order Spectral analysis and involve bispectral and bicepstral analysis. A knowledge base containing the target specific features has been developed from the estimated feature set, and serves to provide the required classification clues for the classifier. The classifier compares the features of the unknown signal with the ones in the knowledge base and checks for a suitable match, based on which the system performs the decision making process.

3.1 Background

The underwater ambient noise characteristics are highly complex and varying. One of the reasons for this could be the richness in the number and diversity of the noise generating sources. The biodiversity catered by the ocean is remarkable and as such, many of the aquatic organisms generate a variety of noises. These include the crackling sound of shrimps to cetacean vocalisations. Many of the environmental phenomena like the wind, waves, rain, precipitation, cracking of ice and even earthquakes also contribute to the ambient noise. The biological and environmental sources are generally grouped under natural noise sources.

Apart from these, considerable noise levels are being contributed by human activity. Such noise include shipping noise, noise generated due to exploration, military research activities, etc. The manmade noise levels have been increasing considerably, and in many cases, it even reaches a level that disturbs the natural ecology of the ocean itself.

Thus it is apparent that the underwater acoustic environment presented to the sonar system comprises of a complex mixture of noises, which include the biological noise combined with environmental and manmade noises. Considering the large operational range of such systems, the effective acoustic field will also encompass sources from a variety of locations. Thus, the computation of various statistics for interpreting the passive sonar signals for achieving target classification becomes a challenging task.

3.2 Model of the proposed Target classifier

A simplified block diagram of the proposed classifier is shown in Fig. 3.1. The acoustic target signals are preprocessed and relevant features are extracted from these signals. Techniques based on HOS analysis are being used for the feature extraction. The feature extraction is followed by dimensionality reduction. A knowledge base is formed from the known target signatures and has been used to train the classifier. During the testing phase, the classifier is presented with the extracted and selected features of the unknown signal to predict the class label. The following sections outline the extraction of higher order features, the generation of the feature vector and the different classification algorithms employed in this work.

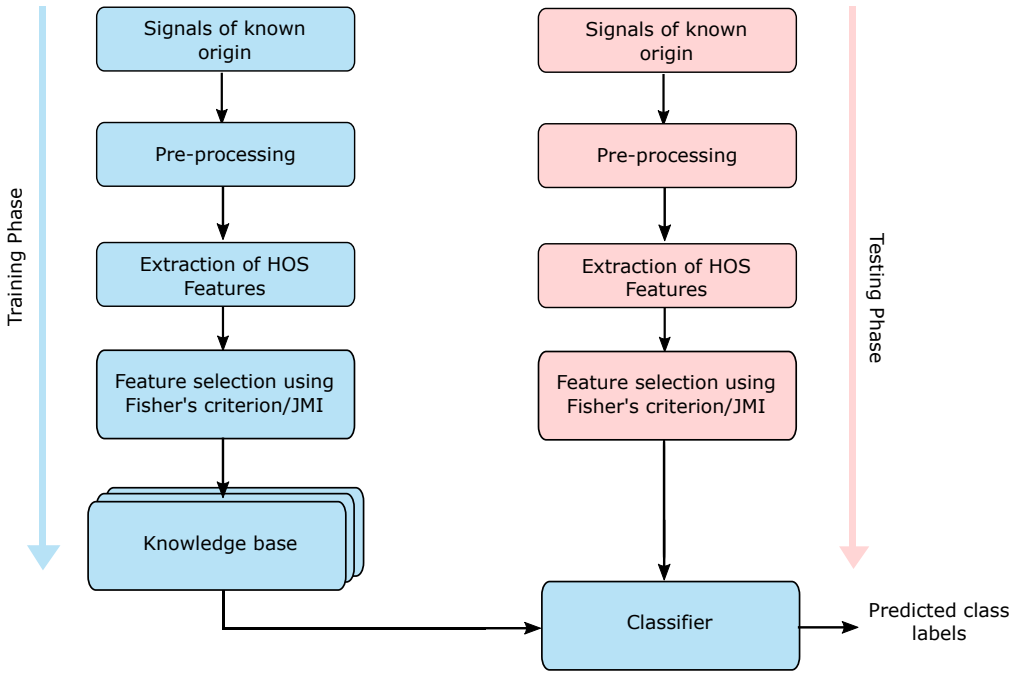


Figure 3.1: Block diagram of the proposed classifier

3.3 Feature Extraction

Extraction of robust features from noisy acoustic signals emanating from underwater noise sources is one of the challenging steps in the process of classification. As has been outlined in chapter 2, Most of the existing feature extraction techniques rely on the conventional power spectral estimation methods and its variants. However, being a linear method, the use of power spectral analysis may fail to provide information pertaining to the deviations from linearity and Gaussianity of stochastic processes, which is quite often the case in underwater scenario. This calls for the use of alternative methods like higher order spectral analysis, in order to gain a more complete understanding of the underlying signal dynamics.

Defined in terms of higher order cumulants, the Higher Order Spectra (HOS) or polyspectra, offers a novel set of techniques, algorithms and methodologies for the study and analysis of signals, especially those originating from nonlinear systems. The study of HOS has been dominated by work on the third order polyspectra, or the bispectrum and to some extent the fourth order polyspectra, also known as the trispectrum. The motivations behind the use of HOS analysis include its ability to suppress Gaussian noise, to reconstruct the phase and the magnitude response of the signals or systems and to detect and characterise nonlinear interactions [58]. Due to these attractive properties, Higher Order Signal processing, especially the bispectrum has been used in many applications, including pattern recognition [102], biomedical signal processing [39], underwater signal processing [53], etc.

Since bispectrum, which is a higher order spectra of order three, is identically zero for many noise processes, especially the Gaussian process, it can be used to suppress the effect of additive white Gaussian noise [40]. Bispectrum also allows the detection and characterisation of second-order nonlinear interactions between frequency components, known as the Quadratic Phase Coupling (QPC). Signal reconstruction or system identification can also be achieved in bispectrum domain due to its ability to preserve both magnitude and phase information, even in the presence of additive Gaussian noise. Thus, the limitations of the second order methods and the advantages of the higher order methods motivate the use of higher order spectral analysis for feature extraction. A consolidated view of the various higher order features that have been combined to form the feature vector proposed in this work is depicted in Fig. 3.2.

After framing and windowing of the source signal, bispectral and bicepstral estimation have been carried out to extract different higher order feature sets. The bispectral matrix resulting from the bispectral estimation has been utilised to generate bicoherence, integrated bispectra and biologically derived features. The bispectral matrix contains information regarding the quadratic nonlinear interaction between the frequency components of the source signal. However, it is found that, the variance of the bispectrum at the bifrequency is proportional to the product of the power of the signals [8] at these frequencies.

The bicoherence, which is a normalised form of the bispectrum is generally employed in order to make the bispectrum independent of the energy content at the bifrequencies. The diagonal elements of the bicoherence matrix which represent the self coupling components in the

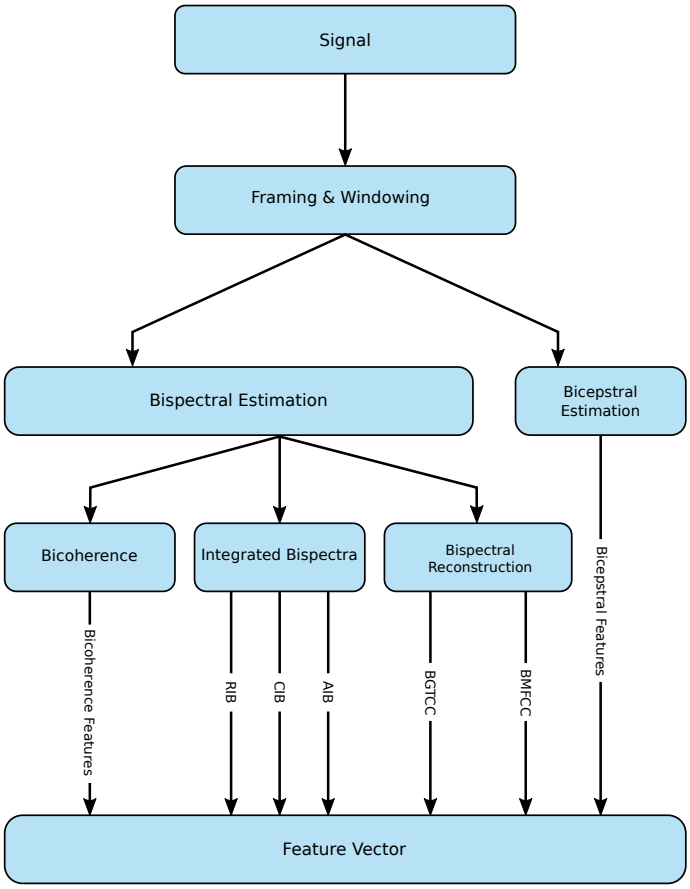


Figure 3.2: Consolidated view of feature vector generation

signal, together with the anti-diagonal elements have been considered for the feature vector generation.

The information in the two dimensional bispectral matrix can be transformed into a one dimensional vector using different types of integrated bispectra, namely the Axially Integrated Bispectra (AIB), Radially Integrated Bispectra (RIB) and Circularly Integrated Bispectra (CIB). Such integrated bispectra also possess some attractive properties like translational and rotational invariance and have been

included in the feature vector.

The human auditory system is excellent in processing and perceiving acoustic signals and many attempts have been made to model and mimic both its functional as well as structural properties, over the past few decades. As a result, many perceptually motivated filter banks have been derived by modelling the auditory response of the human ear to different spectral bands with nonlinear spacing. Such filter banks can be utilised in order to extract robust features from the target signals. Two such conventional cepstral feature extraction approaches, namely the Mel Filter Cepstral Coefficients (MFCC) and Gammatone Cepstral Coefficients (GTCC) have been studied further, for the purpose of deriving the Bispectral MFCC and Bispectral GTCC and have been incorporated into the feature vector.

The bicepstrum has also proven to be useful in a number of applications in different domains, varying from fault detection to the analysis of power generation systems. The bicepstrum can be computed by taking the logarithm of the bispectrum and inverse transforming the log-spectrum. The definition of the cepstrum in terms of higher order cumulants enables us to retrieve the cepstrum of the source wavelet analytically, conserving the phase information. Hence, the bicepstral coefficients have also been incorporated into the feature vector.

A detailed theoretical description and the extraction algorithms of the proposed higher order features have been presented in Chapter 4. The higher order feature vector thus generated has been optimised using feature selection techniques to achieve efficient classification performance.

3.4 Feature Selection

Generally, for a robust classification framework, the features will be extracted using different techniques such that the resultant feature vector encompasses the various defining attributes of the targets. Some of these features may be completely irrelevant or redundant and the feature selection process eliminates such features and chooses an optimal and relevant subset of features using some decision criterion [65]. Formally, the problem of feature selection can be stated as follows: given a set of D features, select a subset of m features that leads to the smallest classification error. Thus, feature selection removes irrelevant or redundant features and reduces the data dimensionality.

In this work, two feature selection approaches namely Fishers Criterion and Joint Mutual Information (JMI), which are detailed in Chapter 2 have been employed.

3.5 Knowledge base Compilation

Feature extraction followed by feature selection lead to the formation of the target feature vector which forms the input to the classifier. The feature vectors of known signals thus formed are appropriately labelled to create a knowledge base for the classifier.

The knowledge base used in the proposed classifier consists of HOS features extracted from the signals of 30 different acoustic targets of natural and anthropogenic origin. The natural sources include recordings of biological sources like vocalisations of different types of whales, as

well as recordings of environmental noises like rain and cracking of ice. Recordings of noise emanating from manmade sources like ships and boats have also been included in the knowledge base. Some of these raw noise data for the compilation of the knowledge base have been collected during the scheduled cruises off Cochin and off Mangalore. Other recordings have been collected from various open source sound databases available in the Internet.

3.6 Classifiers

The classifier compares the labelled signals in the knowledge base against the feature vector of the unknown signal to perform the decision making process. Three classifiers namely the k-Nearest Neighbour (k-NN), Artificial Neural Networks (ANNs) and the Support Vector Machine (SVM) Classifier have been employed and their performance has been analysed in this thesis.

A k-Nearest Neighbour Classifier is a relatively simple yet efficient classifier, where each sample is labelled according to the majority of its k nearest neighbours, based on some form of distance measures. The performance of the classifier using Euclidean, city block and correlation distance measures have been analysed in this work. The value of k has also been varied to analyse its effect on the classifier performance.

Artificial Neural Networks (ANNs) are inspired by the biological nervous system and can be described as a properly weighted directed graph, in which each node performs a transfer function [78]. The nodes are generally organised in layers, and a typical ANN will have an input layer, an output layer and one or more hidden layers in between them.

In the training phase, the compiled knowledge base of the targets are applied as input to the network and the network adjusts its weights for capturing the feature values of the targets. The number of hidden layers, the number of neurons in each layer and the transfer function of the layers have been varied to obtain optimal classifier performance.

In many practical classification scenarios, the classes are non-separable linearly, in the feature space. The Support Vector Machine (SVM) Classifier relies on the fact that it is possible to transform the data into a higher dimensional space where the classes are linearly separable, by employing kernel functions [82]. For any kernel, there will be some parameters that would effectively determine the properties and efficiency of the classifier involved. The optimal values of these parameters need to be determined using some kind of model selection or parameter search, in order to ensure the best possible performance of the classifier. In this thesis, an RBF Kernel has been used and Cross-validation as well as Grid-search approaches have been employed for identifying the optimal values for the kernel parameters.

3.7 Summary

The methodology adopted for the development of an intelligent target classifier has been discussed in this chapter. An overview of the feature vector generation using various higher order features including bicoherence, integrated bispectra and biologically derived feature sets have been presented. The knowledge base formation for the proposed target classifier is provided along with a description of the various types of classifiers employed in this work.

Chapter 4

Higher Order Feature Sets

A common concern of feature extraction techniques is to obtain the features that are fairly invariant to irrelevant aspects of the data, as far as the classification is concerned. This chapter examines the Higher Order Spectral analysis techniques like bispectral and bicepstral analysis to obtain the robust feature vectors from the underwater noise data. The bispectrum is a third-order spectrum that has a number of properties which make it a valuable signal processing tool. This chapter also describes the various integrated bispectra for estimating the features from the noise data. Axial, radial and circular integration in the bifrequency plane has been considered. Conventional techniques like Mel Frequency Cepstral Coefficients and Gammatone Cepstral Coefficients, which are biologically motivated, have also been considered for the purpose of generating their bispectral counterparts.

4.1 Background

One of the main objectives of signal processing is to extract the relevant information from the obtained signal, that is generally corrupted by noise. Such information, in the context of classification is termed as the characteristic features and the process of feature extraction involves the application of relevant techniques to generate a set of parameters that can uniquely characterise the signal generating mechanism. The techniques to be employed for this purpose depend on the nature of the signal and the noise. For example, the noise and the signal can be Gaussian or non-Gaussian, and the underlying generating mechanism can be linear or nonlinear. Till recently, such analysis were carried out by assuming and approximating the signals as Gaussian, originating from linear processes. One of the reasons for such an approximation is that almost all the classical signal analysis tools are centered around the power spectral analysis, which is based on the second order statistics that cannot characterise non-linearity and non-Gaussianity. In power spectral estimation, the frequency components of the signal under consideration along with their power distribution is estimated and such an analysis is sufficient to fully describe the Gaussian signals. However, many practical situations call for the extraction of information regarding deviations from Gaussianity and linearity, and in such circumstances, one may have to look for techniques that can provide capabilities beyond the power spectral analysis.

Recently, a lot of research is being carried out in higher order statistics and as a result, analysis techniques with higher order spectra, also known as polyspectra, which is defined in terms of higher order

statistics, have been developed and are extensively used in places where the second order statistics fail. The most commonly used higher order tools are based on the third and fourth order cumulants, and are called the bispectrum and trispectrum, respectively.

4.2 Higher Order Spectra or Polyspectra

Higher Order Spectra can be defined in terms of either cumulants (cumulant spectra) or moments (moment spectra).

Assuming that the n^{th} order cumulant sequence of a random signal $x(k)$, as defined in section 2.4, is absolutely summable, the n^{th} order **cumulant spectrum** of $x(k)$, $S_n(\omega_1, \omega_2, \ldots, \omega_{n-1})$ exists, and is defined to be the $(n-1)$-dimensional Fourier transform of the n^{th} order cumulant sequence. In an analogous manner, the n^{th} order **moment spectrum** is defined as the $(n-1)$-dimensional Fourier transform of the corresponding moment sequence.

In general, cumulant spectra is more widely used in processing random signals when compared to moment spectra. This wide usability of cumulant spectra can be attributed to the following properties that the moment spectra do not share:

1. Additivity: The cumulants of the sum of two independent random processes equals the sum of the cumulants of the process
2. Cumulant spectra of order > 2 are zero if the underlying process in Gaussian
3. Cumulants quantify the degree of statistical dependence of time series
4. Cumulants of higher-order white noise are multidimensional

impulses and the corresponding cumulant spectra are flat.

Let $c_n(\tau_1, \tau_2, \cdots, \tau_{n-1})$ be the n^{th} order cumulant and is absolutely summable. Then the general formula for the n^{th} order (cumulant) spectrum is given by

$$S_n(\omega_1, \omega_2, \cdots, \omega_{n-1}) = \sum_{\tau_1=-\infty}^{\infty} \cdots \sum_{\tau_{n-}1=-\infty}^{\infty} c_n(\tau_1, \tau_2, \cdots, \tau_{n-1}) exp\left[-j \sum_{i=1}^{n-1} \omega_i \tau_i\right] \tag{4.1}$$

In general, $S_n(\omega_1, \omega_2, \cdots, \omega_{n-1})$ is complex, and has both magnitude and phase. The power spectrum, bispectrum and trispectrum are the

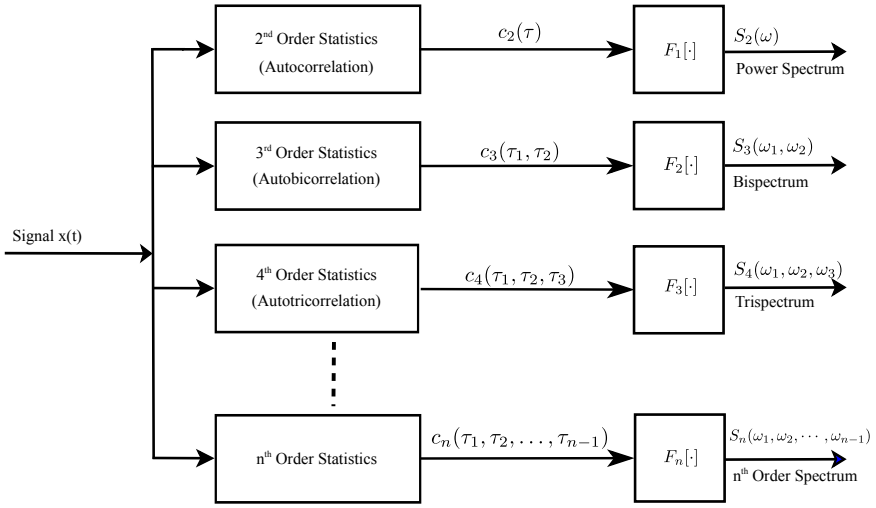


Figure 4.1: Computation of Polyspectra

special cases of n^{th} order cumulant spectra for $n = 2, 3, 4$ respectively. Thus, using Eq. (4.1):

Power Spectrum ($n = 2$)

$$S_2(\omega) = \sum_{\tau=-\infty}^{\infty} c_2(\tau) exp\left[-j\omega\tau\right] \tag{4.2}$$

Bispectrum $(n = 3)$

$$S_3(\omega_1, \omega_2) = B(\omega_1, \omega_2) = \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} c_3(\tau_1, \tau_2) exp\left[-j(\omega_1\tau_1 + \omega_2\tau_2)\right] \tag{4.3}$$

Trispectrum $(n = 4)$

$$S_4(\omega_1, \omega_2, \omega_3) = \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} \sum_{\tau_3=-\infty}^{\infty} c_4(\tau_1, \tau_2, \tau_3) exp[-j(\omega_1\tau_1+\omega_2\tau_2+\omega_3\tau_3)] \tag{4.4}$$

A diagram illustrating the computation of the polyspectra of different orders is given in Fig. 4.1.

4.3 The Bispectrum

The bispectrum is defined as the two dimensional Fourier transform of the third order cumulant. It is the second member in the family of polyspectra, the first member being the power spectrum.

4.3.1 Principal Domain

From the symmetry of the third order cumulant as indicated in Eq. (2.6), the $(\tau_1 - \tau_2)$ plane can be divided into six regions. By knowing the cumulants in any one of these regions, the cumulants in the other five regions can be easily calculated. It follows from this symmetry property of the third order cumulant that the third order cumulant spectrum of the bispectrum is also sixfold symmetric. That

is,

$$\begin{aligned} S_3(\omega_1,\omega_2) &= S_3(\omega_2,\omega_1) = S_3^*(-\omega_2,-\omega_1) \\ &= S_3^*(-\omega_1,-\omega_2) = S_3(-\omega_1-\omega_2,\omega_2) \\ &= S_3(\omega_1,-\omega_1-\omega_2) = S_3(-\omega_1-\omega_2,\omega_1) \\ &= S_3(\omega_2,-\omega_1-\omega_2) \end{aligned} \tag{4.5}$$

Thus, by knowing the bispectrum in the triangular region as illustrated in Fig. 4.2, the bispectrum of the other regions can be easily computed [103]. This region represented by $\omega_2 \geq 0$, $\omega_2 \geq \omega_1$, $\omega_1 + \omega_2 \leq \pi$, is generally called the principal region or the non-redundant region.

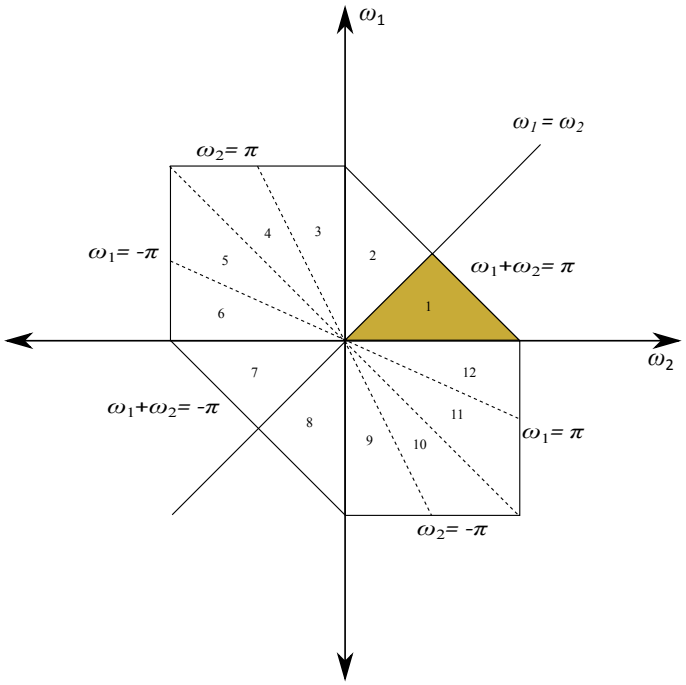


Figure 4.2: Illustration of symmetry in bispectral plane

4.3.2 Properties

- Bispectrum $B(\omega_1, \omega_2)$ is generally complex. That is it has both magnitude and phase.

$$B(\omega_1, \omega_2) = |B(\omega_1, \omega_2)|\, exp(j\phi(\omega_1, \omega_2)) \tag{4.6}$$

- $B(\omega_1, \omega_2)$ is doubly periodic with a period 2π

$$B(\omega_1, \omega_2) = B(\omega_1 + 2\pi, \omega_2 + 2\pi) \tag{4.7}$$

- Bispectrum has 12 regions of symmetry as shown in Figure 4.2
- Gaussian Processes: For a stationary zero-mean Gaussian process $x(k)$, the third order cumulant sequence $c_3(\tau_1, \tau_2) = 0$ for all (τ_1, τ_2) and therefore its bispectrum $B(\omega_1, \omega_2)$ is also identically zero.
- Linear Phase Shifts: For a process $x(k)$ let the power spectrum and bispectrum be represented as $P_x(\omega)$ and $B_x(\omega_1, \omega_2)$. Let $y(k) = x(k-N)$ be a shifted version of $x(k)$, where N is a constant integer and has the power spectrum $P_y(\omega)$ and bispectrum $B_y(\omega_1, \omega_2)$. Since the second and third-order moments suppress linear phase information, we have, $P_x(\omega) = P_y(\omega)$ and $B_x(\omega_1, \omega_2) = B_y(\omega_1, \omega_2)$. It may be noted that while the power spectrum suppresses *all* phase information, the bispectrum does not.

4.3.3 Estimation

In digital signal processing, the computations require discrete or sampled data and practically, the data available will be of finite length.

Just as power spectrum estimation, the estimation of the bispectrum from finite length data can be achieved either using the conventional (non-parametric) methods or by the parametric methods. In non-parametric method, there are two classes of estimators-direct and indirect [40]. The direct class of estimators compute the bispectrum directly from the data. The time domain data is first transformed into the frequency domain by Fourier transform, and from this the bispectrum is estimated. In the indirect method, cumulant sequence is estimated first, from the time series, and then the 2D Fourier transform is computed to estimate the bispectrum. It may be noted that the conventional estimators make no assumptions about the model of the process.

Parametric estimators are based on Auto Regressive (AR), Moving Average (MA) or Auto Regressive Moving Average (ARMA) models. These estimators assume an underlying model for the process under consideration and the parameters of the corresponding model are computed leading to the estimation of the bispectrum.

While the parametric techniques can provide better estimates of the bispectrum when there is an understanding about the underlying model, the conventional estimators have the advantages of simplicity and ease of implementation. Also, the conventional estimators can provide good estimates with sufficiently long data records.

4.3.3.1 Direct Method

For estimating the bispectrum using the direct method [40], consider a series comprising of N samples, $\{x(1), x(2), \ldots, x(N)\}$.

- Divide the series into K segments or records, each of length M. Let $x_i^{'}(k)$, represent the i^{th} record. $i = 1, 2, \ldots, K$ and $k = 1, \ldots, M$.
- The mean μ_i of the i^{th} segment is computed and subtracted from each sample of that segment.

$$x_i(k) = x_i^{'}(k) - \mu_i \tag{4.8}$$

- The M-point Discrete Fourier Transform, $X_i(k)$ of each segment is computed, i.e.,

$$X_i(k) = \sum_{n=0}^{M-1} x_i(n) e^{-j\frac{2\pi}{M}nk}, \quad k = 0, 1, \ldots, M-1, \quad i = 1, 2, ..., k. \tag{4.9}$$

- The bispectrum of each segment is obtained as

$$B_i(k_1, k_2) = \frac{1}{M} X_i(k_1) X_i(k_2) X_i^*(k_1 + k_2), \quad i = 1, ..., K. \tag{4.10}$$

- Finally, the bispectrum is computed as the mean of all $B_i(k_1, k_2)$

$$B(\omega_1, \omega_2) = \frac{1}{K} \sum_{i=1}^{K} B_i(\omega_1, \omega_2), \quad \omega_i = \frac{2\pi}{M} k_i, \quad i = 1, 2. \tag{4.11}$$

It may be noted that the symmetry properties of the bispectrum can be used to reduce the computational complexity. $B_i(k_1, k_2)$ need to be computed only in the triangular region $0 \leq k_2 \leq k_1$, $k_1 + k_2 < M/2$, and the values of other regions can be easily deduced from this.

A two-dimensional rectangular smoothing window of size $(M_3 \times M_3)$ can be applied to the computed bispectrum, in order to

reduce the variance of the estimate.

$$\tilde{B}_i(k_1,k_2) = \frac{1}{M_3^2} \sum_{n_1=-M_3/2}^{M_3/2-1} \sum_{n_2=-M_3/2}^{M_3/2-1} B_i(k_1+n_1, k_2+n_2) \tag{4.12}$$

The final bandwidth of this bispectrum estimate is $\Delta = M_3/M$, which is the spacing between frequency samples in the bispectrum domain. For large N, the direct method produce asymptotically unbiased and consistent bispectrum estimates, with real and imaginary part variances, given by:

$$\begin{aligned} var\left(Re\left[\hat{B}(\omega_1,\omega_2)\right]\right) &= var\left(Im\left[\hat{B}(\omega_1,\omega_2)\right]\right) \\ &= \tfrac{1}{\Delta^2 N} X(\omega_1)X(\omega_2)X(\omega_1+\omega_2) \\ &= \tfrac{M}{KM_3^2} X(\omega_1)X(\omega_2)X(\omega_1+\omega_2) \end{aligned} \tag{4.13}$$

From the above expressions, it is clear that the variance of the bispectrum estimate can be reduced by increasing the number of records or the size of the frequency smoothing window (M_3).

4.3.3.2 Indirect Method

Consider a series comprising of N samples, $\{x(1), x(2), \ldots, x(N)\}$.

- Divide the series into K segments or records, each of length M. Let $x_i^{'}(k)$, represent the i^{th} record. $i = 1, 2, \ldots, K$ and $k = 1, \ldots, M$.
- The mean μ_i of the i^{th} segment is computed and subtracted from each sample of that segment.

$$x_i(k) = x_i^{'}(k) - \mu_i \tag{4.14}$$

- Estimate the moments of each segment

$$m_3^i(\tau_1, \tau_2) = \frac{1}{M} \sum_{l=l_1}^{l_2} x^i(l) x^i(l+\tau_1) x^i(l+\tau_2) \tag{4.15}$$

where,

$$l_1 = max(0, -\tau_1, -\tau_2), \quad l_2 = min(M-1, M-2)$$
$$|\tau_1| < L, \quad |\tau_2| < L, \quad i = 1, 2, \cdots K$$

Note that the third-order moments and cumulants are identical, as each segment has zero mean: $c_3(\tau_1, \tau_2) = m_3(\tau_1, \tau_2)$.

- Obtain the average cumulants as:

$$c_3(\tau_1, \tau_2) = \frac{1}{K} \sum_{i=1}^{K} m_3^i(\tau_1, \tau_2) \tag{4.16}$$

- Compute the third-order bispectrum estimate by taking the double Fourier transform:

$$B(\omega_1, \omega_2) = \sum_{\tau_1=-L}^{L} \sum_{\tau_2=-L}^{L} c_3(\tau_1, \tau_2) \exp^{-j(\omega_1\tau_1 + \omega_2\tau_2)} \omega(\tau_1, \tau_2) \tag{4.17}$$

where $L < M - 1$. Also note that a two-dimensional window $\omega(\tau_1, \tau_2)$ can be introduced to smooth out edge effects, as described in the case of direct method.

4.3.4 Quadratic Phase Coupling

There are situations in which the interactions between two harmonic components cause contribution to the power at their sum and/or difference frequencies. As an illustration, consider two sinusoids

$x_1(n) = cos(2\pi f_1 n + \phi_1)$ and $x_2(n) = cos(2\pi f_2 n + \phi_2)$ passed through a nonlinear system, such that the output signal $y(n)$ is $x_1(n) + x_2(n) + [x_1(n) + x_2(n)]^2$.

That is,

$$\begin{aligned} y(n) = cos(2\pi f_1 n + \phi_1) + cos(2\pi f_2 n + \phi_2) + \\ cos^2(2\pi f_1 n + \phi_1) + cos^2(2\pi f_2 n + \phi_2) + \\ 2cos(2\pi f_1 n + \phi_1) cos(2\pi f_2 n + \phi_2) \end{aligned} \tag{4.18}$$

Or,

$$\begin{aligned} y(n) = cos(2\pi f_1 n + \phi_1) + cos(2\pi f_2 n + \phi_2) + 1 + \\ \frac{1}{2} cos(4\pi f_1 n + 2\phi_1) + \frac{1}{2} cos(4\pi f_2 n + 2\phi_2) + \\ cos[2\pi(f_1 + f_2)n + (\phi_1 + \phi_2)] + \\ cos(2\pi(f_1 - f_2)n + (\phi_1 - \phi_2)) \end{aligned} \tag{4.19}$$

As given in the above equation, the output signal $y(n)$ contains the components with frequencies $2f_1, 2f_2, (f_1 + f_2)$ and $(f_1 - f_2)$ along with the original components f_1 and f_2. The component of $y(n)$ also exhibits certain phase relations as shown:

$$\begin{aligned} 2f_1 &\longleftrightarrow 2\phi_1 \\ 2f_2 &\longleftrightarrow 2\phi_2 \\ f_1 + f_2 &\longleftrightarrow \phi_1 + \phi_2 \\ f_1 - f_2 &\longleftrightarrow \phi_1 - \phi_2 \end{aligned} \tag{4.20}$$

and such phase relationships arising out of quadratic non-linearity in the signal generating mechanism is called the Quadratic Phase Coupling (QPC) . It may be noted that, in situations where there is only frequency

coupling and no phase coupling between the components, no QPC exists.

Consider two signals $x_1(k)$ and $x_2(k)$ such that,

$$\begin{aligned} x_1(k) &= \cos(\lambda_1 k + \theta_1) + \cos(\lambda_2 k + \theta_2) + \cos(\lambda_3 k + \theta_3') \\ x_2(k) &= \cos(\lambda_1 k + \theta_1) + \cos(\lambda_2 k + \theta_2) + \cos(\lambda_3 k + \theta_3) \end{aligned} \tag{4.21}$$

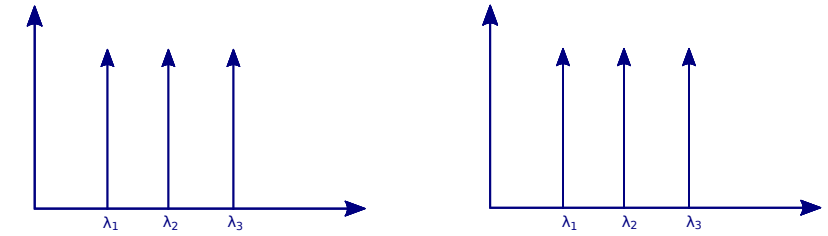
where $\lambda_3 = \lambda_1 + \lambda_2$, $\theta_3 = \theta_1 + \theta_2$ and $\theta_3' \neq \theta_1 + \theta_2$. $\lambda_1, \lambda_2, \lambda_3$ are said to be harmonically related and $\theta_1, \theta_2, \theta_3, \theta_3'$ are independent random variables uniformly distributed between $[0, 2\pi]$. λ_3, in the case of $x_1(k)$, is an independent harmonic component because θ_3' is an independent random-phase variable. However, λ_3 of $x_2(k)$ is a result of phase coupling between λ_1 and λ_2. The autocorrelation sequences of the two signals are given by

$$c_2^{x_1}(\tau_1) = c_2^{x_2}(\tau_1) = \frac{1}{2}\{\cos(\lambda_1\tau_1) + \cos(\lambda_2\tau_2) + \cos(\lambda_3\tau_3)\} \tag{4.22}$$

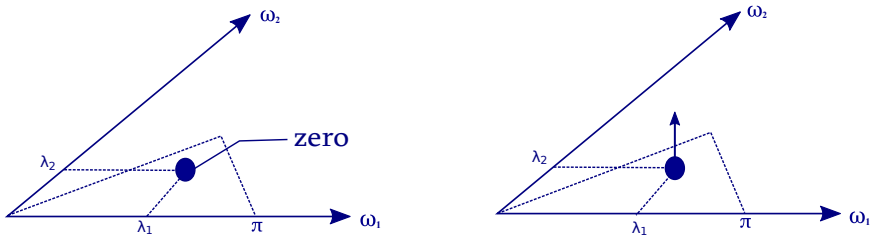
Thus it is apparent that, $\{x_1(k)\}$ and $\{x_2(k)\}$ have identical power spectra with peaks at λ_1, λ_2 and $\lambda_3 = \lambda_1 + \lambda_2$. This is has been illustrated in Fig. 4.3(a). The third order cumulants of the signals can be obtained as

$$\begin{aligned} c_3^{x_1}(\tau_1, \tau_2) &= 0 \\ c_3^{x_1}(\tau_1, \tau_2) &= \tfrac{1}{4}\{cos(\lambda_2\tau_1 + \lambda_1\tau_2) + cos(\lambda_3\tau_1 - \lambda_1\tau_2)\} \\ &\quad + cos(\lambda_1\tau_1 + \lambda_2\tau_2) + cos(\lambda_3\tau_1 - \lambda_2\tau_2) \\ &\quad + cos(\lambda_1\tau_1 - \lambda_3\tau_2) + cos(\lambda_2\tau_1 - \lambda_3\tau_2) \end{aligned} \tag{4.23}$$

Thus, it is clear that the bispectrum of $\{x_1(k)\}$ is identically zero while the bispectrum of $\{x_2(k)\}$ is not. The bispectrum of $\{x_2(k)\}$ exhibits peak at (λ_1, λ_2) (if $\lambda_1 \geq \lambda_2$), in the triangular region $\omega_2 \geq 0$, $\omega_1 \geq \omega_2$, $\omega_1 + \omega_2 \leq \pi$. This has been illustrated in Fig. 4.3(b)



(a) Power spectrum of $x_1(k)$ and $x_2(k)$



(b) Magnitude Bispectrum of $x_1(k)$ and $x_2(k)$

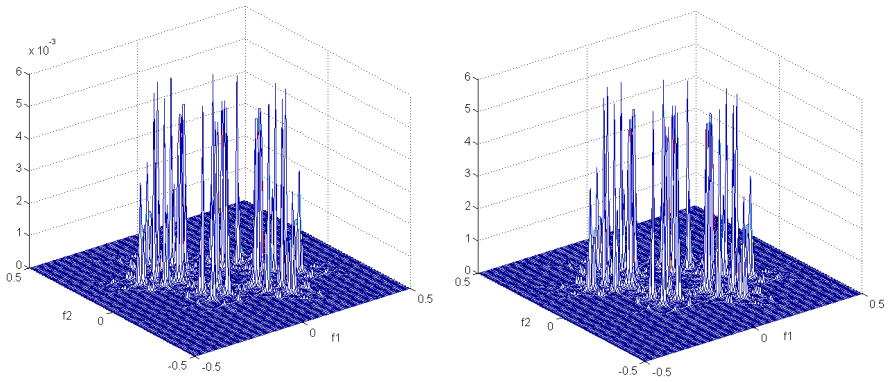
Figure 4.3: Illustration of Quadratic Phase Coupling

Thus, the power spectrum fails to discriminate $\{x_1(k)\}$ from $\{x_2(k)\}$ as it suppresses phase relations of harmonic components. However, the bispectrum does preserve quadratic phase relations and therefore becomes useful for extracting them quantitatively.

4.3.5 Normalisation

In the case of bispectrum, it is found that, at the bifrequency (f_1, f_2), the complex variance is proportional to the product of the power of the signals at the frequencies f_1, f_2 and $(f_1 + f_2)$ [44]. i.e.,

$$var[B(f_1, f_2)] \propto P(f_1)P(f_2)P(f_1 + f_2) \tag{4.24}$$



(a) Signal amplified by 10 (b) Signal amplified by 100

Figure 4.4: Bispectral plots illustrating the energy dependency of bispectrum

Fig. 4.4(a) and (b) shows the bispectral plots of a signal amplified by a factor of 10 and 100 respectively.

It can be seen that as the amplitude of the signal gets increased, the magnitude of the bispectral plots also get increased on a large scale.

Thus, in order to make the bispectrum independent of the energy content at the bifrequencies, a normalised measure, referred to as the *bicoherence* can be used. Bicoherence, which is a normalised form of the bispectrum, can be defined as

$$bic(f_1, f_2) = \frac{|B(f_1, f_2)|}{\sqrt{P(f_1)P(f_2)P(f_1 + f_2)}} \tag{4.25}$$

Since the bicoherence is independent of the energy or amplitude of the signal, it can be used as a convenient test statistic for the detection of non-Gaussian, non-linear and coupled processes. Fig. 4.5 shows the contour and mesh plots of the bicoherence of a typical noise waveform. The same signal has been amplified by a factor of 10 and 100 and the

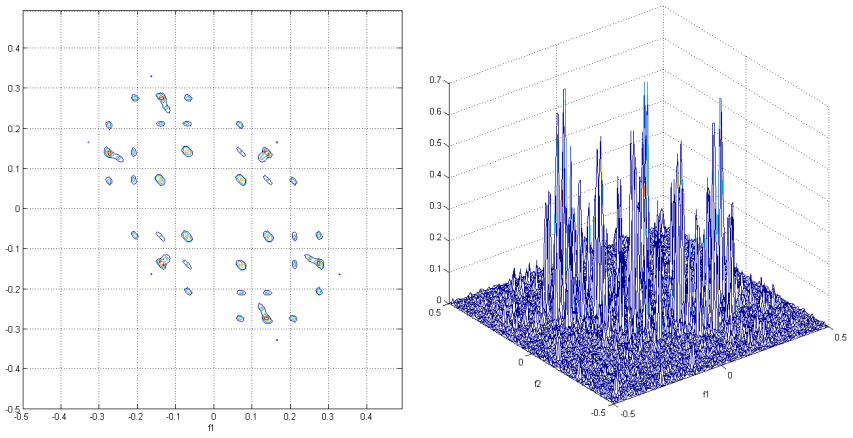
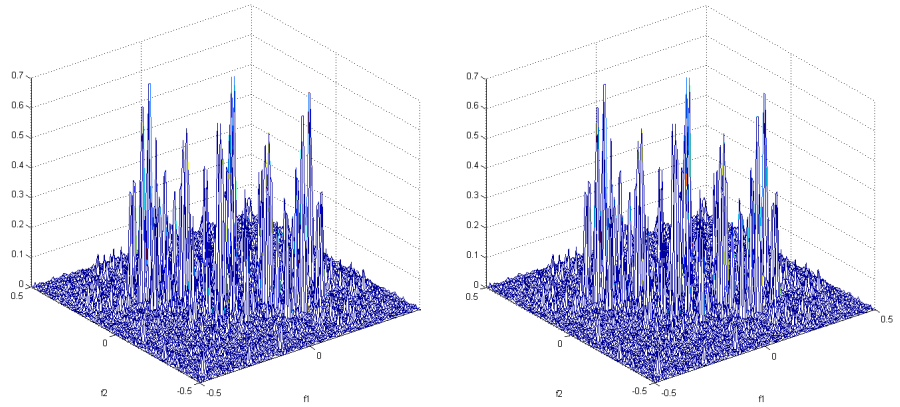


Figure 4.5: Bicoherence of a typical signal - The contour and mesh plot

corresponding bicoherence plots are presented in Fig. 4.6(a) and (b). From the figures, it is clear that the bicoherence is independent of the energy content or the amplitude of the signal.

4.4 Bispectral Analysis of Underwater Noise Signals

A detailed analysis of the various noise sources present in the marine environment using bispectrum has been carried out, from a target classification perspective. The signals from various sources, both of manmade as well as biological origin, have been analysed. The target data waveforms have been segmented into records and the bispectrum of the records are computed using direct method, as outlined in the previous sections. In direct method, the FFT of the target waveform is computed using a suitable FFT of length N. Generally, the value of N is chosen to be a power of 2, such as 64, 128, 256 or 512. Figure 4.7 shows the mesh and contour plots of the



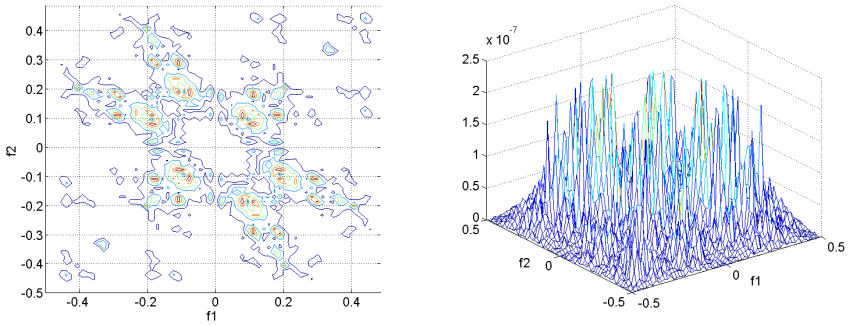
(a) Signal amplified by 10 (b) Signal amplified by 100

Figure 4.6: Bicoherence plots of amplified signals

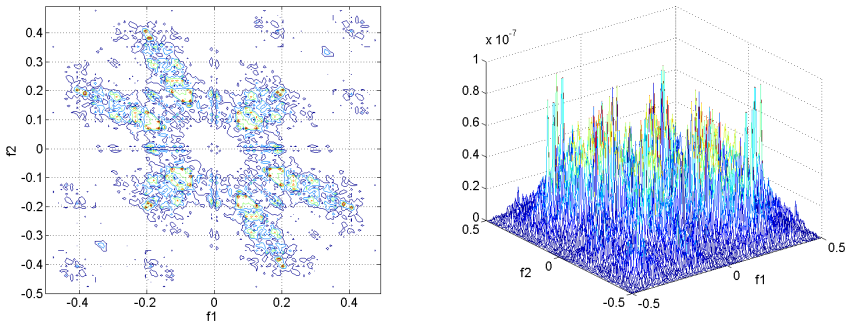
bispectrum of a typical signal, generated using different values of N.

The plots generated with $N = 64$, has much broader peaks, as seen from the mesh plot, due to poor the frequency resolution. The plots generated with 128 point FFT have moderate peaks and those generated with 256 point FFT have narrow peaks. Thus, it may be desirable to use an FFT size, as large as possible, since it would increase the frequency resolution. This is illustrated in Fig. 4.7(c). However, increasing the DFT size, can conflict with the requirement of having a large number of segments K, as given in Eq. (4.13) which is an essential requirement for obtaining reliable estimates with low variance and bias. Thus, this thesis has chosen 128 point FFT, which strikes a balance in both the resolution and number of records, to compute the bispectral matrix.

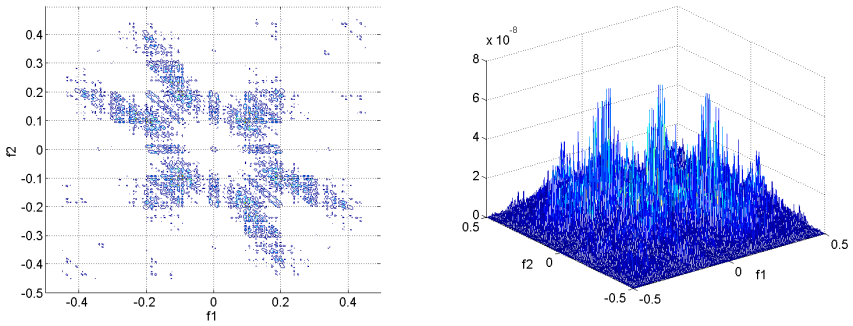
As has already been discussed, the bispectrum depends on the amplitude of the signal under consideration. Thus, its normalised form, the bicoherence, has been used for the analysis of the noise waveforms,



(a) Bispectral Plot with 64-Point FFT



(b) Bispectral Plot with 128-Point FFT



(c) Bispectral Plot with 256-Point FFT

Figure 4.7: Bispectral plots of a target computed with different FFT bin sizes

for the detection of non-Gaussian, nonlinear and coupled processes. The mesh and contour plots of the bicoherence computed for two different targets, a ship and a whale, have been shown in Figs. 4.8 and 4.9.

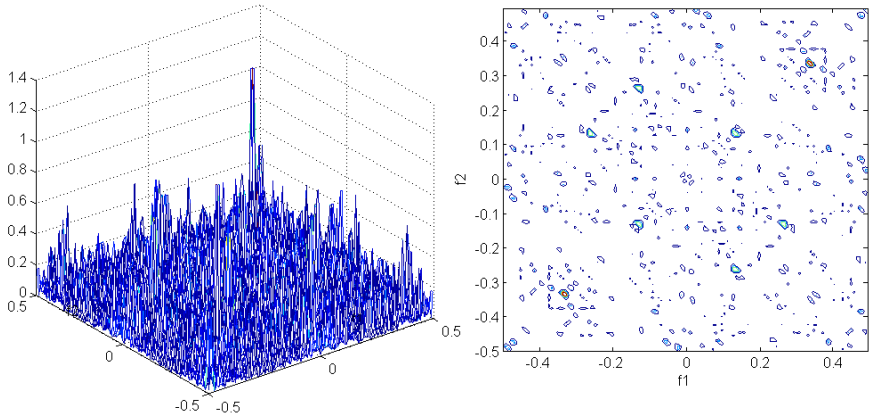


Figure 4.8: Bicoherence plot for Ship

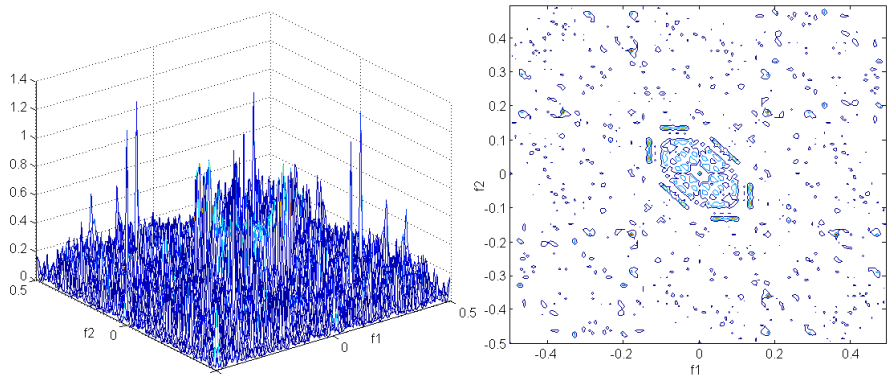


Figure 4.9: Bicoherence plot for Whale

These plots reveal the presence of many peaks which represents Quadratic Phase Couplings (QPC) as described in section 4.3.4. A peak corresponding to a bifrequency (f_1, f_2), reveals that QPC exists

between the constituent frequency components f_1 and f_2 of the signal, which generally arises due to the quadratic nonlinearity in the signal generating mechanism. In order to remove the peaks due to less important couplings, some thresholding can be applied. Figures 4.10 (a) and (b) respectively show the contour plot of the previous two targets, filtering out the peaks which are less than 30% of the maximum peak value.

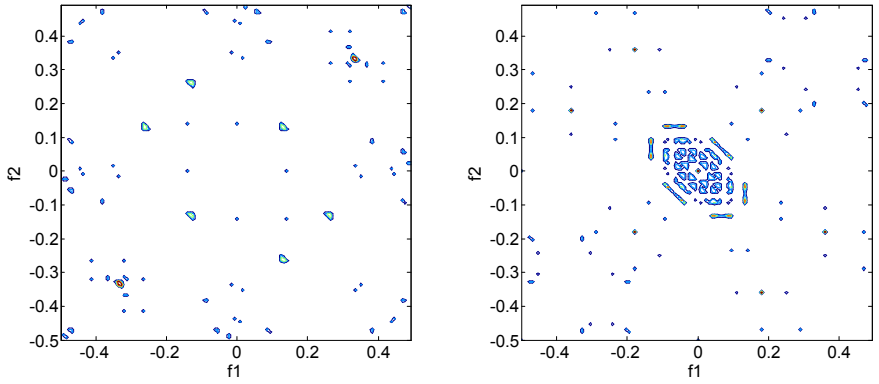


Figure 4.10: Contour plot of (a)Ship and (b)Whale after thresholding

4.5 Integrated Bispectra

Although the bispectra have all the advantages of cumulants/polyspectra in feature extraction, their direct use has certain implications. Being an $N \times N$ matrix, with N generally equal to 128, the feature vector generated by the two-dimensional (2-D) matrix can be huge, even after thresholding. This can consume both computational power and time during the classification process, and can also leads to overfitting during the training stage of the classifier. Thus, in order to efficiently use bispectra as feature vectors for

identification and classification problems, integrated bispectral methods may be investigated [102], [104].

4.5.1 Axially Integrated Bispectrum (AIB)

The Axially Integrated Bispectrum (AIB) is obtained by integrating the bispectra along paths parallel to the ω_1 or ω_2 axes in bifrequency plane and thus retains the scale characteristics of the signal [105].

$$\begin{aligned} AIB(\omega) &= \tfrac{1}{2\pi}\int_{-\infty}^{\infty} B(\omega_1,\omega_2)d\omega_2 \\ &= \tfrac{1}{2\pi}\int_{-\infty}^{\infty} B(\omega_1,\omega_2)d\omega_1 \end{aligned} \tag{4.26}$$

Consider a zero mean stationary random process $x(t)$ and let $B(\omega_1,\omega_2)$ be its bispectrum. Using Eq. (4.3), the third order cumulant $c_3(\tau_1,\tau_2)$ can be expressed as,

$$c_3(\tau_1,\tau_2) = \frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} B(\omega_1,\omega_2)\times exp\{j(\omega_1\tau_1+\omega_2\tau_2)\}d\omega_1 d\omega_2 \tag{4.27}$$

Let $\tilde{y}(t) = x^2(t) - E\left\{x^2(t)\right\}$. The cross correlation $r_{\tilde{y}x}(\tau 1)$ between $\tilde{y}(t)$ and $x(t)$ is defined to be:

$$r_{\tilde{y}x}(\tau_1) = E\left\{\tilde{y}(t)x(t+\tau_1)\right\} = E\left\{x^2(t)x(t+\tau_1)\right\} = c_{3x}(0,\tau_1) \tag{4.28}$$

and its cross spectrum is given by:

$$S_{\tilde{y}x}(\omega) = \sum_{-\infty}^{\infty} c_{3x}(0,\tau_1)\exp\{-j\omega\tau_1\} \tag{4.29}$$

with

$$c_{3x}(0,\tau_1) = \frac{1}{2\pi}\int_{-\pi}^{\pi} S_{\tilde{y}x}(\omega)\exp\{-j\omega\tau_1\}\,d\omega \tag{4.30}$$

Comparing Eq. (4.27) and (4.30),

$$S_{\tilde{y}x}(\omega) = \frac{1}{2\pi}\int_{-\pi}^{\pi} B_x(\omega,\omega_2)d\omega_2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} B_x(\omega_1,\omega)d\omega_1 \tag{4.31}$$

Thus the axially integrated bispectrum is a function of a single frequency variable, and can be considered as a cross spectrum between the signal $x(t)$ and its square [106].

Though the AIB contains less phase information when compared to the bispectra, the estimation variance of the AIB is much less, equivalent to that of power spectrum.

4.5.2 Radially Integrated Bispectrum (RIB)

Chandran and Elgar [102], first proposed the use of radially integrated bispectra in pattern recognition and demonstrated its applicability. The RIB is obtained by integrating the bispectrum along radial lines passing through the origin in bifrequency space, as shown in Fig. 4.11. The integrated bispectra can be defined as:

$$RIB(a) = \int_{0^+}^{1/(1+a)} B(f_1, af_1)df_1 \tag{4.32}$$

In the discrete domain, let the bispectrum $B(k,l)$ be computed with an N-point FFT, with $0 \le l \le k \le k+l \le (N/2-1)$. The integration in

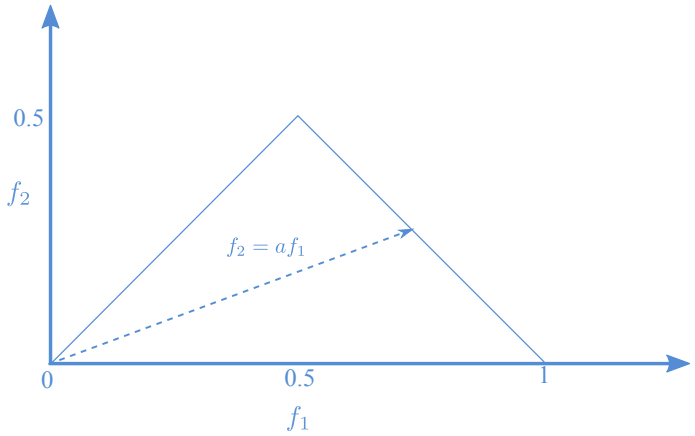


Figure 4.11: RIB Computation - Integrating the bispectrum along the dashed line with slope=a.

Eq. (4.32) can be approximated by,

$$RIB(a) = \sum_{k=1}^{\lfloor(\frac{N}{2}-1)/(1+a)\rfloor} B(k, ak) \tag{4.33}$$

for $0 < a \leq 1$. The bispectrum is interpolated for this summation by

$$B(k, ak) = pB(k, \lceil ak \rceil) + (1-p)B(k, \lfloor ak \rfloor) \tag{4.34}$$

where $p = ak - \lfloor ak \rfloor, and \lfloor x \rfloor$ represents the largest integer contained in x and $\lceil x \rceil$ represents the smallest integer containing x.

4.5.3 Circularly Integrated Bispectrum (CIB)

As the name suggests, for the Circularly Integrated Bispectrum, the integral paths are a set of concentric circles with the origin as the centre

The CIB can be defined as [107]:

$$CIB(a) = \int B_p(a, \theta) d\theta \tag{4.35}$$

where $B_p(a, \theta)$ is the polar representation of $B(\omega_1, \omega_2)$.

4.5.4 Feature Extraction using Integrated Bispectra

The Integrated Bispectra (IB) represent the two dimensional bispectral matrix into a vector of one dimension. The bispectral matrix is computed from the target waveforms and the matrix is integrated axially, radially and circularly to get the respective integrated bispectra. In the discrete domain, the integration has to be approximated with summation.

The AIB is obtained by summing the values in each column of the bispectral matrix. Since a 128 point DFT has been used for the computation of the bispectrum, the AIB vector will contain 128 values. RIB has been computed using the first quadrant of the bispectral matrix. The point where $\omega_1 = \omega_2 = 0$ is considered to be the origin and a line with a slope $m = \tan(\theta)$ is assumed to pass over the bispectral plane. All the values in the matrix through which the line passes is summed up to get the RIB corresponding to that slope. In this case, θ was varied from 0° to 90°.

Considering the point $\omega_1 = \omega_2 = 0$ as the centre, the bispectral values are summed over a circular path of radius r, to obtain the CIB corresponding to that radius. In this work, CIB values for radius ranging from 1 to 64, have been computed to generate the feature set, 64 being the maximum number of values on either side of the center point of the

128×128 bispectral matrix.

The IBs for three targets namely pinniped, hump and a submarine have been generated and plotted in Figs. 4.12, 4.13 and 4.14 respectively, for illustrating the discriminative nature of these features for different types of targets. Each plot shows the respective Integrated Bispectra computed for 20 different records of that particular target.

The Axially Integrated Bispectra (AIB) shown in the top plots of Figs. 4.12 through 4.14 have been obtained by summing all the frequency components along the column of the bispectral matrix. Considering the symmetry of the bispectral plots as described in section 4.3.2, these plots are symmetric about the middle feature index and only half of the values may be considered for the inclusion in the feature vector. Since AIB represent the sum of the bispectral values corresponding to all the bifrequency pairs (f_1, f_2), keeping f_1 as constant, the height of the peak in the AIB plot is indicative of the strength of the coupling where that particular frequency is involved.

By integrating the bispectral matrix axially, radially and circularly, AIB, RIB and CIB attempt to represent the information contained in the bispectral matrix using a much reduced set of values. The interclass variability of these features has been illustrated in Fig. 4.15, using the averaged AIB, RIB and CIB computed from 20 different records each of pinniped, humpback whale and submarine. Also, the averaged IBs of three different boats, which are considered as different targets, but from the same class/type have been plotted in Fig. 4.16. These plots reveal the variabilities among the target feature values that can be used to characterise these individual targets.

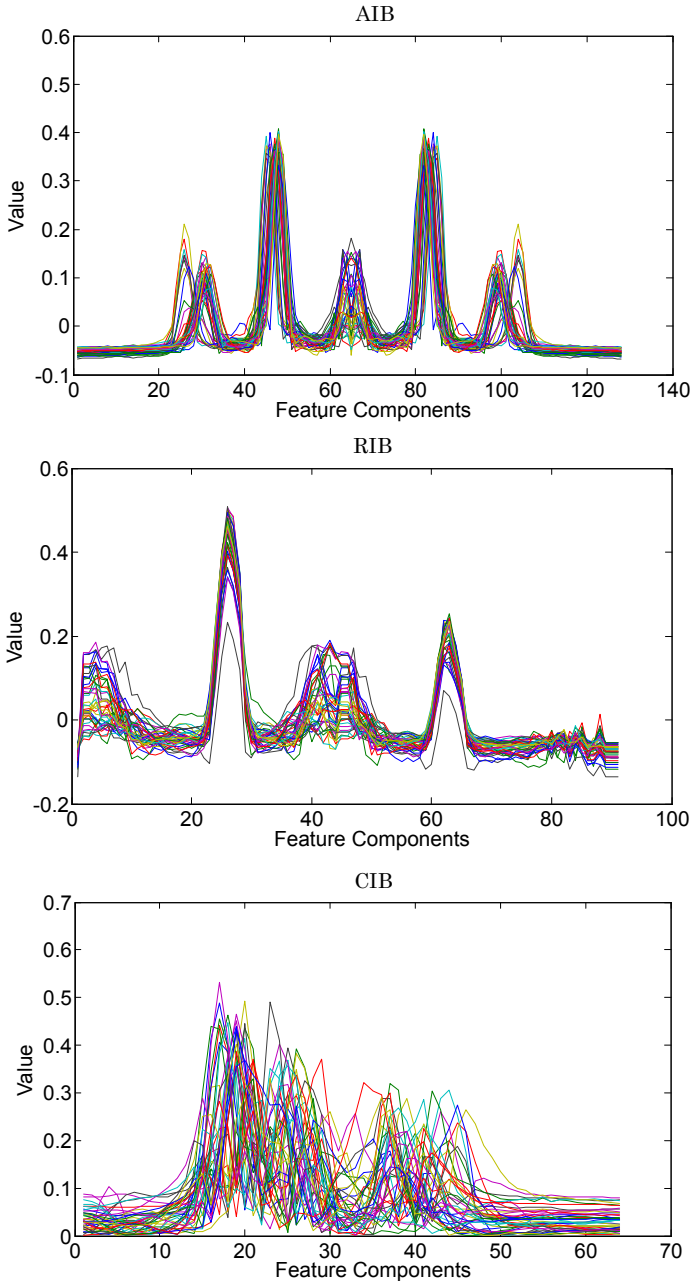


Figure 4.12: Various Integrated Bispectra of Pinniped

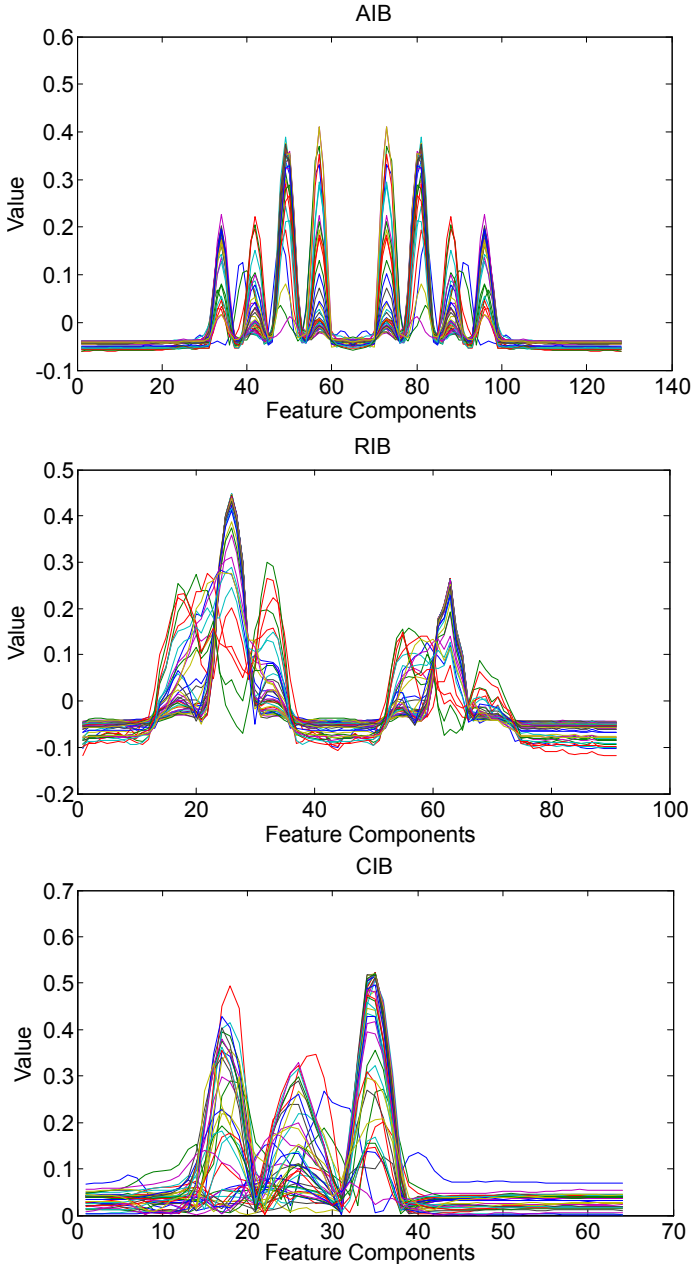


Figure 4.13: Various Integrated Bispectra of Hump

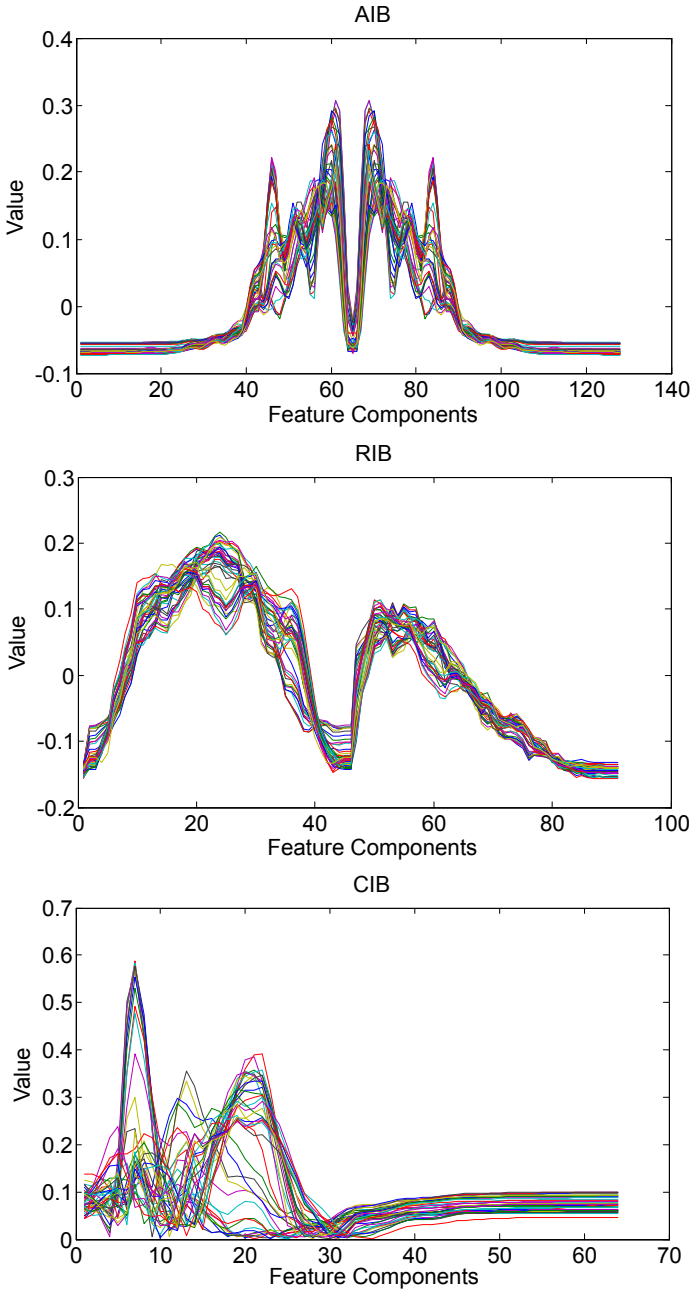


Figure 4.14: Various Integrated Bispectra of a Submarine

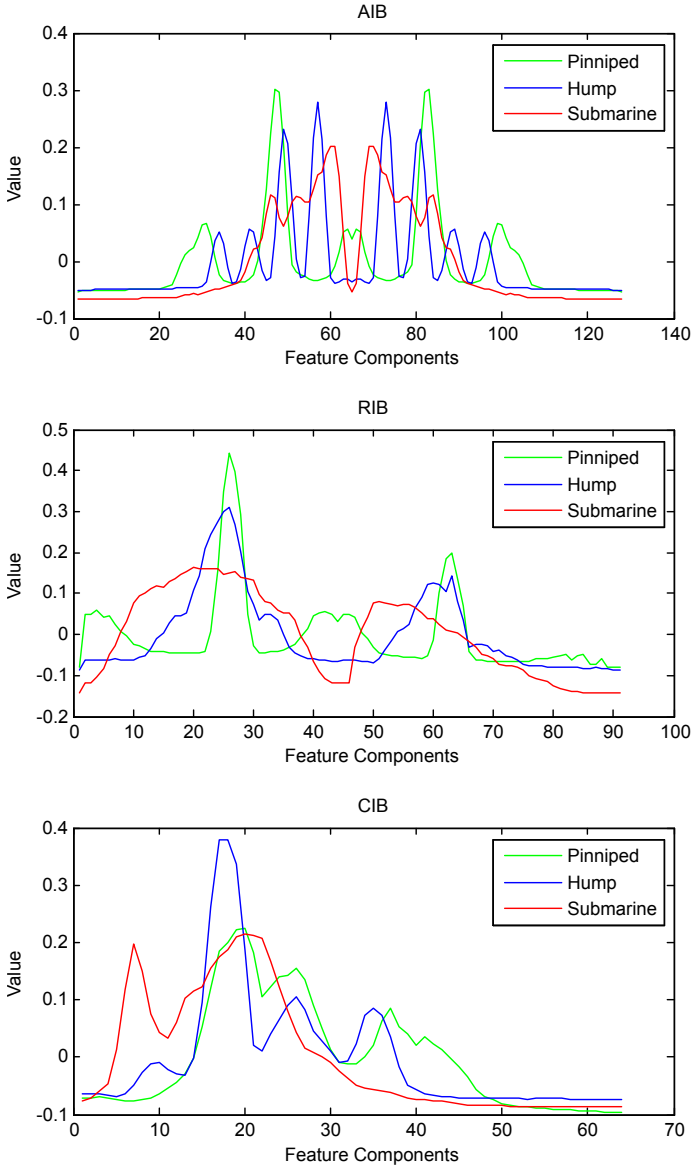


Figure 4.15: Averaged Integrated Bispectra of Three Targets

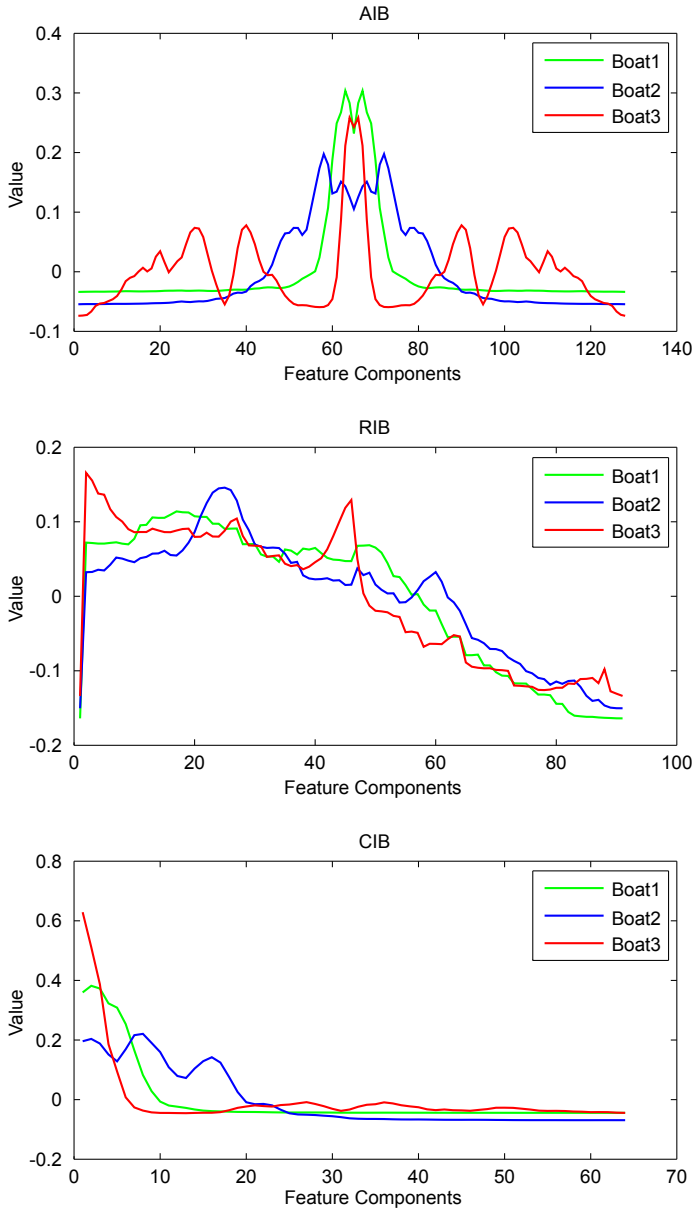


Figure 4.16: Averaged Integrated Bispectra of Different Boats

4.6 Homomorphic Transforms

Homomorphic transform literally means *same structure*. Signals combined in a nonlinear way (multiplication, convolution, etc.), cannot be separated by linear filtering. Homomorphic techniques attempt to separate signals combined in a nonlinear way by transforming the problem to a linear domain [108]. That is, the problem is converted to the same structure as a linear system. Homomorphic decomposition techniques like the cepstral analysis can be employed for feature extraction by separating the convolved signal components and has been widely used especially in acoustic signal processing.

4.6.1 Cepstrum

Cepstral analysis is designed to separate convolved signal components by transforming the composite signal into a domain where components are additive, using logarithmic transformation. The cepstrum is defined as the inverse Fourier transform of the log magnitude Fourier spectrum of the signal and is said to be in the quefrency domain, an anagram of frequency [27].

Consider a composite signal $s(t)$ generated by the convolution of the two components, $x(t)$ and $y(t)$. By taking the Fourier transform, the convolution in the time domain is transformed into multiplication in the frequency domain. That is,

$$\begin{aligned} DFT\left[s(t)\right] &= DFT\left[x(t) \otimes y(t)\right] \\ S(f) &= X(f) \times Y(f) \end{aligned}$$

Considering only the magnitude of the spectrum and taking the logarithms,

$$ln\,|S(f)| = ln\,|X(f)| + ln\,|Y(f)|$$

Thus, the nonlinear combination of the components in time has been transformed into a linear combination of log-magnitude components in the frequency domain. The components $x(t)$ and $y(t)$ can be separated by applying the inverse Fourier transform, or equivalently the discrete cosine transform for real signals. However, it should be noted that the phase information from the original signal has been lost, as a result of the magnitude operation on the complex spectra.

4.6.2 Bicepstrum

Bicepstrum is defined as the inverse 2D Z-transform of log bispectrum [40]. ie,

$$b_h(n,m) = Z_2^{-1}(\log\,(S_3(\omega_1,\omega_2)) \tag{4.36}$$

Defining $\log\,(S_3(\omega_1,\omega_2))$ as $B_h(\omega_1,\omega_2)$ and taking partial differentiation,

$$\frac{\partial B_h(\omega_1,\omega_2)}{\partial\omega_1} = \frac{1}{S_3(\omega_1,\omega_2)}\frac{\partial S_3(\omega_1,\omega_2)}{\partial\omega_1}$$

Re-arranging,

$$S_3(\omega_1,\omega_2)\frac{\partial B_h(\omega_1,\omega_2)}{\partial\omega_1}\omega_1 = \frac{\partial S_3(\omega_1,\omega_2)}{\partial\omega_1}\omega_1 \tag{4.37}$$

From Eq. (4.37) it can be shown that [40], the bicepstrum $b_h(n,m)$ and the third order moment sequence $m_3(n,m)$ can be related by the linear

convolution formula

$$-n{\cdot}m_3(n,m) = m_3(n,m) * (-mb_h(n,m)) \tag{4.38}$$

A method for computing the cepstrum coefficients that does not require phase unwrapping can be derived from Eq. (4.38) based on two-dimensional Fast Fourier Transform operations as:

$$n{\cdot}b_h(n,m) = F_2^{-1}\left\{\frac{F_2(n{\cdot}m_3(n,m))}{F_2(m_3(n,m))}\right\} \tag{4.39}$$

where $F_2(*)$ represents the 2-D Fourier transform.

The bicepstrum can be applied to deconvolve both deterministic and stochastic signals. Bicepstrum contains phase information and it can be applied to obtain the system impulse response and the inverse filter. Another advantage of bicepstrum is that it less sensitive to the presence of additive Gaussian noise [59], [109]

4.6.3 Estimation of Bicepstrum

The steps for bicepstral estimation can be summarised as:

1. Segment the waveform data into K frames or records of M samples each.
2. Mean adjust each record by subtracting the mean and apply a suitable window.
3. Estimate the third order moment sequence $m_3^i(n,m)$ for each record using Eq. (4.15).

4. Compute the average moment sequence.

$$m_3(n,m) = \frac{1}{K}\sum_{i=1}^{K} m_3^i(n,m)$$

5. Compute the 2-D FFT of $n{\cdot}m_3(n,m)$ and $m_3(n,m)$
6. Compute the bicepstrum coefficients by applying inverse 2D Fourier transform as in Eq. (4.39).

A block diagram illustrating the estimation procedures of the bicepstrum is given in Fig. 4.17.

4.6.4 Bicepstral Features of Underwater Targets

The bicepstrum of the target waveforms were computed from their third order cumulants (moments) as discussed in the previous section. Figures 4.18 and 4.19 show the mesh and contour plots of the estimated bicepstrum for two records of a boat noise. Plots of two more targets, an onboard motor and a snapping shrimp have also been shown in Figs. 4.20 and 4.21.

4.7 Perceptually Motivated Higher Order Features

Over the time, the physiology and engineering of the auditory system has been a focus of intensive research and this knowledge has paved the way for the creation of models, which at least to some extent, tries to describe and mimic the human hearing mechanisms. Mathematical and computational models are helpful to carry out quantitative as well

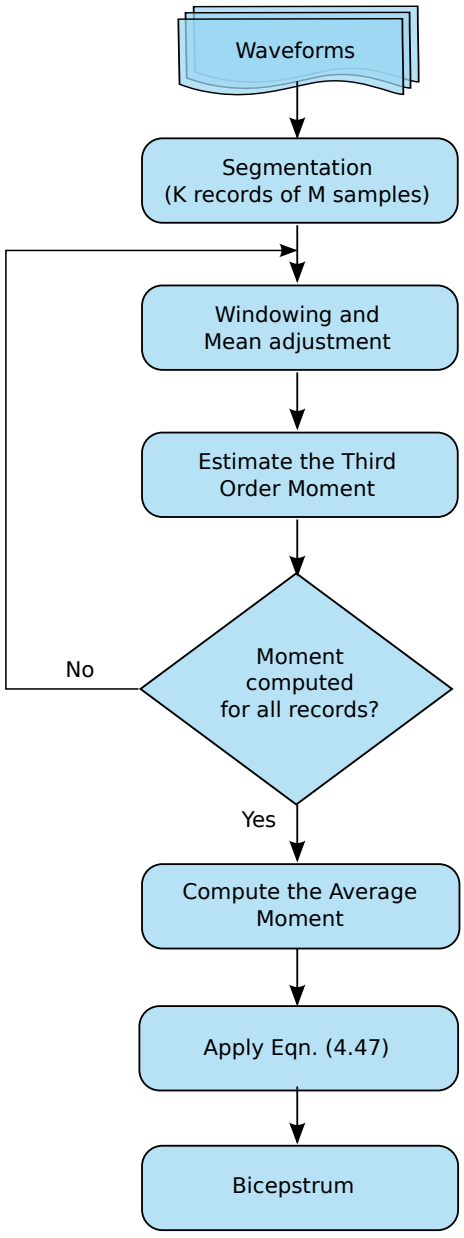


Figure 4.17: Computation of Bicepstrum

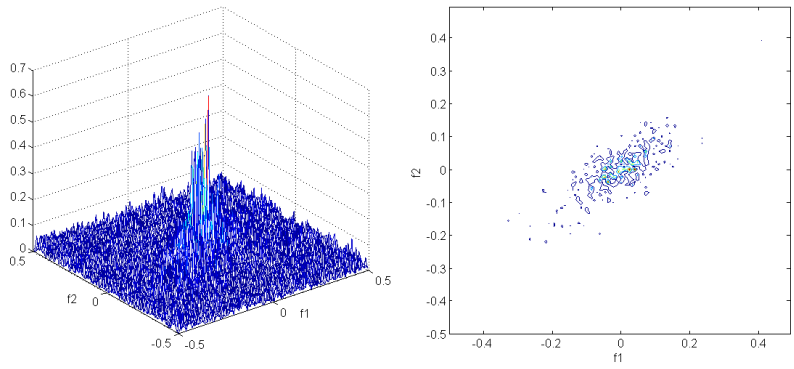


Figure 4.18: Bicepstrum of a record of a Boat Waveform

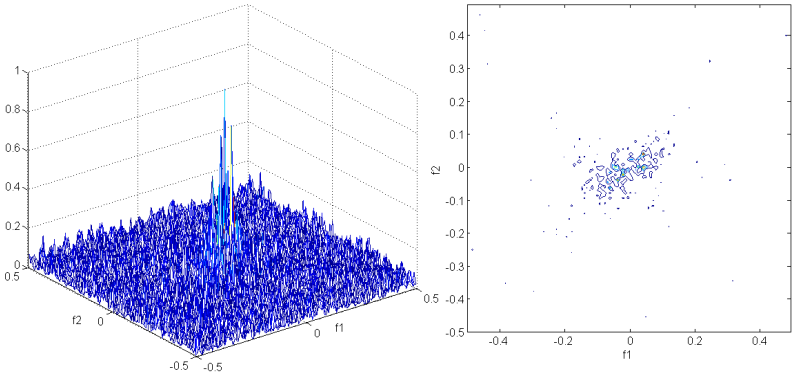


Figure 4.19: Bicepstrum of another record of a Boat Waveform

as qualitative simulations of the original system and aids in providing insights into the system. The development of such auditory models might be used for several purposes, such as understanding the mechanisms of sound perception, tracking the problems of hearing, etc.

4.7.1 Hearing Physiology

Basically, the peripheral auditory system has three main parts:

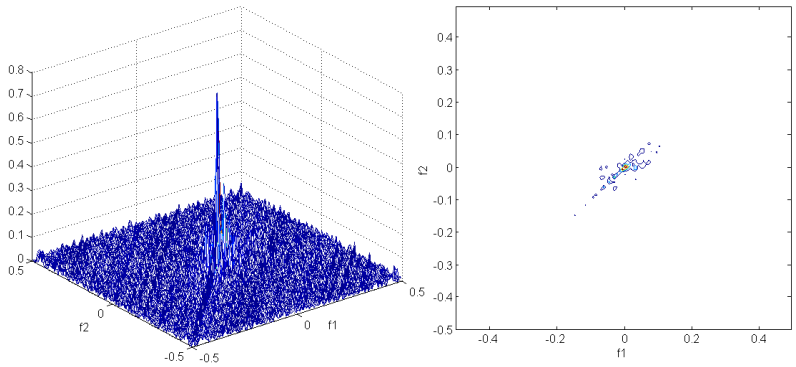


Figure 4.20: Bicepstrum of a record of an onboard motor Waveform

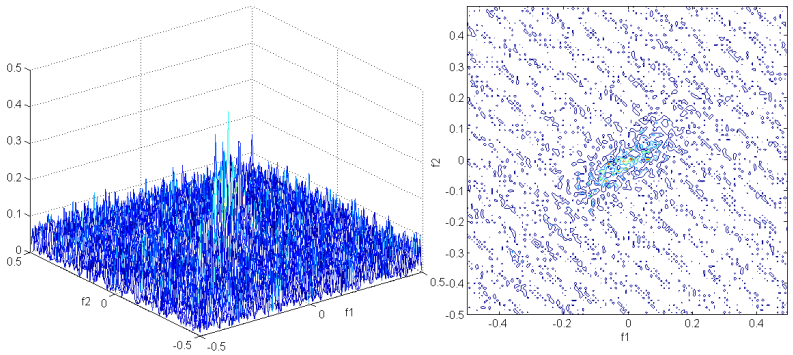


Figure 4.21: Bicepstrum of another record of a Snapping Shrimp Waveform

- The outer,
- Middle and
- Inner ear

The outer ear is the visible portion of the ear and includes the pinna (auricle), the ear canal and the eardrum. The auricle is responsible for the collection of sounds and directing it to the ear canal which acts as a quarter-wavelength resonator and enhances spectral components to which the human ear is more sensitive. The main component of the

middle ear is the ossicle chain and is made of three tiny bones incus, malleus and stapes. The primary function of the middle ear is to efficiently transfer acoustic energy in air to fluidmembrane waves located in the inner ear. The inner ear is located just behind the oval window, which is connected to the stapes footplate and comprises of the vestibular apparatus, the cochlea and the auditory nerve terminations. The cochlea is regarded as having remarkable frequency analysis capabilities and most of the perceptually motivated auditory filter banks are designed modelling the properties of the cochlea.

4.7.1.1 The Cochlea as a Filter Bank

The cochlea is a fluid filled structure and is divided by two membranes *viz.*, the Reissner's membrane and the basilar membrane. The incident acoustic energy is propagated to the basilar membrane by the pressure difference generated by the movements of the ossicle. The membrane possesses some peculiar properties, as illustrated in Fig. 4.22. At its basal end, the membrane is narrow and stiff and responds best to high frequencies, while at its apical end it is wider and more compliant, responding best to low frequencies [110]. This helps the cochlea to separate the incoming vibrations into overlapping frequency bands. Mechanical motion of the basilar membrane leads to displacements of the inner hair cells, stereocilia which in turn stimulate the generation of action potentials in the neurons of the auditory nerve. These action potentials are propagated into the auditory brainstem.

The signal processing function of the basilar membrane and the surrounding structures is to filter the incoming broadband sound. There are approximately 3,000 hair cells which connect to the auditory

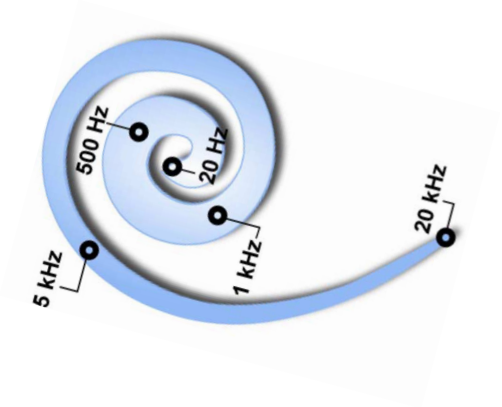


Figure 4.22: Frequency response of the basilar membrane

nerve which acts as 3,000 narrowband channels [35]. Thus, in order to reproduce the filtering characteristics of the auditory system, one has to simulate a 3,000 channel filter bank, which is computationally cumbersome. A usual methodology is to use a single filter bank channel to approximate and model a local group of fibres. Thus, in an auditory model, a single filter models a local group of fibres and the auditory filter bank often consists of many such filters, whose centre frequencies are distributed according to different frequency scales characterising the auditory response of the cochlea.

4.7.1.2 Frequency Scales

Several perceptually motivated scales to approximate the non-uniform tiling of the filters in the auditory filter banks have been proposed, with the Bark, Mel and ERB scales being the most widely used ones.

The Mel scale was a classical approach, to derive from psychoacoustic experiments, a perceptual measure of pitch. It was based on how the

subjects divided a series of simple tones into what sounded like equal intervals. The Mel scale and the frequency f in Hertz is related by [28]

$$Mel(f) = 2595log_{10}(1 + f/700) \tag{4.40}$$

Thus, it can be shown that the Mel-frequency scale is linear below 1 kHz, and logarithmic at higher frequencies.

The Bark scale, based on the critical bandwidth, was designed as an improvement on the Mel scale, which was based on the human perception about equal pitch differences whereas the Bark scale was based on the interaction of frequency and loudness [111]. For human ear, two pure tones with frequencies very close together will sound no louder than a single tone. However, if the spacing between the frequencies is increased, at a certain point the two tones begin sounding louder to subjects, and this point is called the critical bandwidth. The formula for the Bark scale [112] is given by

$$Bark(f) = (26.81 * f)/(1960 + f) - 0.53 \tag{4.41}$$

The Equivalent Rectangular Bandwidth (ERB) of the auditory filter is assumed to be closely related to the critical bandwidth, but is more smoothly behaved than the Bark scale and is measured using the notched-noise method [113]. The notch-noise method has often been employed in the analysis of auditory frequency selectivity and involves the determination of the detection threshold for a sinusoid, centred around a spectral notch of a noise, as a function of the width of the notch. The ERB scale $ERBS(f)$, can be defined as the number of

equivalent rectangular bandwidths below the given frequency f [113].

$$ERBS(f) = 21.4log_{10}(0.00437f + 1) \tag{4.42}$$

A plot illustrating the comparison of the three frequency scales is given in Fig. 4.23.

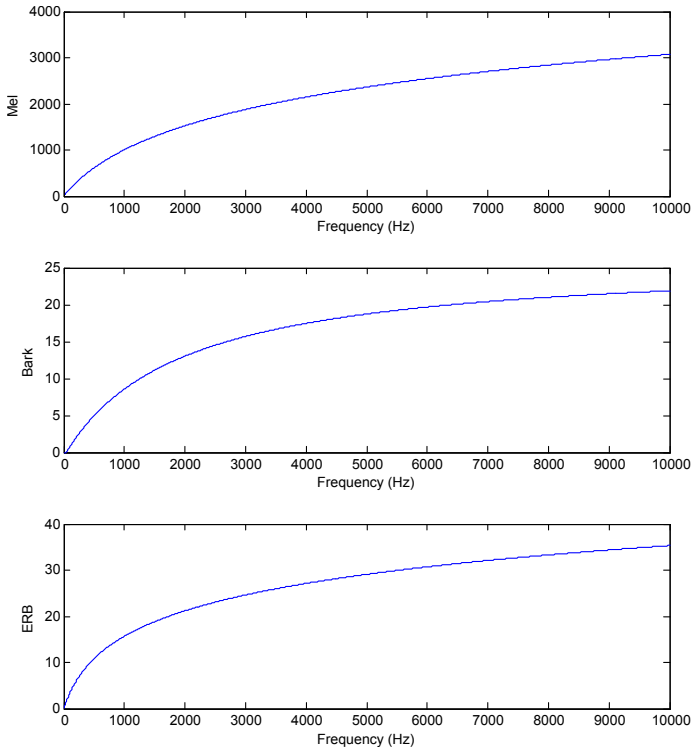


Figure 4.23: Comparison of Mel, Bark and ERB scales

4.7.2 Cepstral Analysis using Auditory Filter Banks

Cepstral analysis can detect repeated patterns in a spectrum, that are difficult or impossible to observe in the various spectral analysis techniques, making it useful for a wide variety of signal processing

applications. The cepstral analysis can be further extended by the application of various auditory filter banks during the computational procedure, in order to warp the spectral bands according to the perceptual scale of the human auditory system. Two of the most popular types of auditory filter banks, the more traditional Mel filter bank and the recently proposed Gammatone filter bank can be employed to extract the corresponding cepstral coefficients namely the Mel Frequency Cepstral Coefficients (MFCC) and Gammatone Cepstral Coefficients (GTCC).

4.7.2.1 Mel Frequency Cepstral Coefficients

The Mel filter bank consists of a set of triangular shaped band-pass filters, which are equidistant on the Mel scale as shown in Fig. 4.24. The Mel filters have narrow bandwidth at low frequencies and get wider as the frequency increases, in accordance with the Mel scale. The spectrum

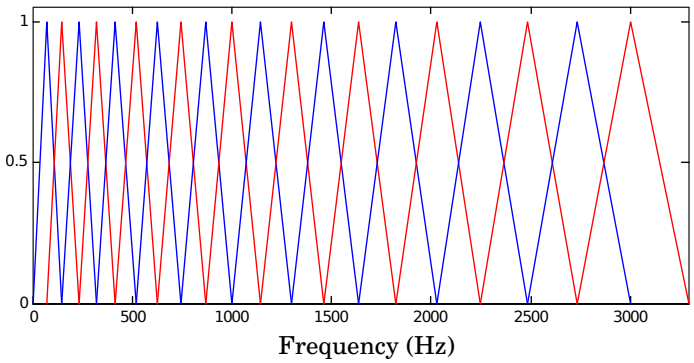


Figure 4.24: Mel scale filter bank

of the signal is transformed such that it is emphasised at Mel intervals using the Mel filter bank. The cepstrum of this transformed spectrum in turn yields Mel Frequency Cepstral Coefficients (MFCC).

4.7.2.2 Gammatone Cepstral Coefficients

Gammatone filter banks are non-uniform overlapping bandpass filters, similar to the Mel filter bank, which models the response of the auditory system more closely. The impulse response of the filter can be described as sinusoid, with an amplitude envelope modulated by a gamma distribution function as illustrated in Fig. 4.25. It can be described mathematically as [37]:

$$g(t) = Kt^{n-1}e^{-2\pi Bt}\cos(2\pi f_c t + \phi) \tag{4.43}$$

where, K is the output gain, B is the duration of the impulse response, n is the order of the filter and f_c is the centre frequency of the filter.

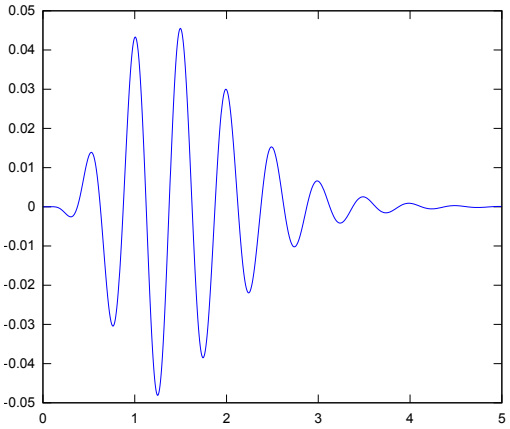


Figure 4.25: A Gammatone impulse response

Equation (4.43) provides a close approximation to experimentally-derived auditory nerve fibre impulse responses, as measured by de Boer and de Jongh using a reverse-correlation technique [114].

Also, the fourth-order gammatone filter provides a good

approximation to the 'rounded-exponential' models of human auditory filter response. Hence, the gammatone filter is in good agreement with the empirically derived perceptual estimates of the auditory frequency selectivity. The distribution of the centre frequencies of the filters in

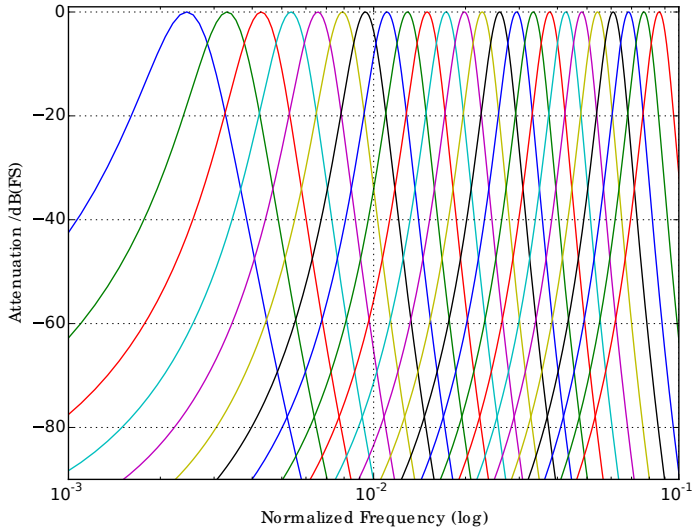


Figure 4.26: A Typical Gammatone Filter Response

the gammatone filter bank follows the Equivalent Rectangular Bandwidth (ERB) scale. Even though several physiologically motivated formulas have been derived for the ERB scale, the model suggested by Glasberg and Moore has been adopted in this work, [35], [36] according to which,

$$ERB(f_c) = 24.7(4.37f_c/1000 + 1) \tag{4.44}$$

where f_c is the centre frequency. The bandwidth of each filter in the Gammatone filter bank is determined according to the critical band (CB) corresponding to its centre frequency. The bandwidth related factor B has been computed using the equation suggested by Patterson [115] and

the value of B is given by,

$$B = 1.019 * 24.7(4.37 f_c/1000 + 1) \tag{4.45}$$

The Gammatone Cepstral coefficients (GTCC) are computed by applying the gammatone filter bank to the spectrum of the signal, followed by the application of the logarithmic and Discrete Cosine Transformations.

4.7.3 Cepstral Features from Bispectrum

MFCCs and GTCCs have been widely used in various feature extraction scenarios due to its low computational complexity and fairly acceptable performances [30, 31, 37, 116]. However, being derived from the second order power spectrum, the performance of MFCC and GTCC depend on the Signal to Noise Ratio (SNR) and hence the performance generally degrades in the presence of noise [117, 118].

Since the bispectrum suppress the additive white Gaussian noise, while preserving the magnitude and phase information of the original signal, it can be used to compute a clean estimate of the magnitude spectrum of the noisy signal. Thus, it is intuitive to use the estimated spectral magnitudes to compute the cepstral coefficients. The MFFC extraction procedure can be extended to generate the Bispectral Mel Frequency Cepstral Coefficients by employing the bispectrally reconstructed spectral components [117]. This thesis also considers the extension of the Gammatone Cepstral Coefficients to generate the Bispectral Gammatone Cepstral coefficients (BGTCC) by employing bispectrally reconstructed spectral components in the GTCC

extraction procedure.

4.7.3.1 Bispectral Reconstruction of Signals

Consider the signal $x(n)$ and its noise corrupted observation $y(n)$, such that $y(n) = x(n) + w(n)$, where $w(n)$ is additive coloured stationary Gaussian noise. Using the fact that Gaussian processes have identically zero bispectra, Bispectral Reconstruction attempts to filter out the additive Gaussian noise from $y(n)$ by first computing the sample bispectrum and then estimating the magnitude and phase of the Fourier transform of $x(n)$. Many techniques have been proposed for the recovery of Fourier amplitude from bispectra, such as recursive reconstruction techniques, closed form approach and Least Squares Approach [56]. This work concentrates on the Least Squares Approach (LSA), as it takes only the bispectral values of the principal region. This method also has the advantage that, it excludes the bispectral values on the axis ($\omega_2 = 0$), as on the axis, the power spectrum of the additive noise contributes to the asymptotic bias [56].

4.7.3.2 Least Squares Approach (LSA)

Consider the Bispectrum $B(k, l)$ of a signal $x(n)$ of length N. For convenience, let N be a multiple of 4, so that both $N/2$ and $N/4$ are integers. Then from the definition of bispectrum [119],

$$B(k, l) = X(k)X(l)X^*(k + l) \tag{4.46}$$

where $X(k)$ represents the DFT of the signal $x(n)$. If we consider only the Bispectral values at the principal region, then, $k = 1, 2, ..., N/4$ and

$l = k, k+1, ..., N/2 - k.$

In order to express Eq. (4.46) as a linear combination, one may rewrite the equation as:

$$\widetilde{B}(k,l) = \widetilde{X}(k) + \widetilde{X}(l) + \widetilde{X}(k+l) \tag{4.47}$$

where,

$$\widetilde{X}(k) = \ln(|X(k)|) \tag{4.48}$$

$$\widetilde{B}(k,l) = \ln(|B(k,l)|) \tag{4.49}$$

Equation (4.47) may be represented in a matrix form as

$$\widetilde{B} = A\widetilde{X} \tag{4.50}$$

Where,

$$\widetilde{B} = [\widetilde{B}(1,1), \widetilde{B}(1,2), \widetilde{B}(1,N/2-1), \widetilde{B}(2,2), \widetilde{B}(2,3) \cdots \widetilde{B}(N/4,N/4)]^T \tag{4.51}$$

which is a vector of dimension $N^2/16$, representing all bispectral points in the principal domain, excluding the frequency axes.

$$\widetilde{X} = [\widetilde{X}(1), \widetilde{X}(2), \cdots \widetilde{X}(N/2)]^T \tag{4.52}$$

is an $N/2$ dimensional vector representing the DFT values. The coefficient matrix A has a dimension of $(N^2/16) \times (N/2)$ and is given

by,

$$A = \begin{bmatrix} 2 & 1 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 \\ 1 & 0 & 1 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 2 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & 1 \end{bmatrix} \tag{4.53}$$

Equation (4.50) represents a set of equations and the solution for $\widetilde{X}$ can be found out in a least square sense as,

$$\widetilde{X} = (A^T A)^{-1} A^T \widetilde{B} \tag{4.54}$$

Considering Eq. (4.48), we can recover the magnitudes of the DFTs as

$$|X_{LSA}(k)| = exp(\widetilde{X}(k)) \tag{4.55}$$

where, $k = 1, 2, ..., N/2$

4.8 Bispectral MFCC

4.8.1 Computation of BMFCC

The Bispectral Mel Frequency Cepstral Coefficients (BMFCC) are generated by first computing the bispectrum of the processed waveform records using the direct method as outlined in Section 4.3.3.1. The magnitude of the DFT, $|X_{LSA}[k]|$, is then computed from the Bispectral values using Least Squares Approach. The power spectrum is then computed as:

$$P_{LSA}[k] = |X_{LSA}[k]|^2 \tag{4.56}$$

For a Mel-frequency filter bank comprising of P filters with individual transfer functions $H_j[k]$, $j = 1, 2, \cdots, P$, the energy m_j in each band is given by

$$m_j = \sum_{k=0}^{N-1} |X_{LSA}[k]|^2 . H_j[k]) \tag{4.57}$$

The Bispectral Mel-frequency cepstrum is the discrete cosine transform of the logarithm of the P filter outputs and is represented as

$$C_i = \sqrt{\frac{2}{N}} \sum_{j=1}^{P} \log(m_j) \cos\left(\frac{\pi i j}{P}\right) \tag{4.58}$$

where C_i is the i^{th} BMFCC coefficient.

The block diagram illustrating the steps of the BMFCC extraction is given in Fig. 4.27.

4.8.2 BMFCC Features

The bispectrum of the processed waveform records are estimated using the direct method and the magnitude of the DFT is computed from the Bispectral values using Least Squares Approach, as elucidated in Eq. (4.56). A mel-filter bank has been generated with 40 filters, with the centre frequencies equally distributed on the Mel scale and has been applied to the computed power spectrum, followed by log and discrete cosine transformations to obtain the BMFC coefficients.

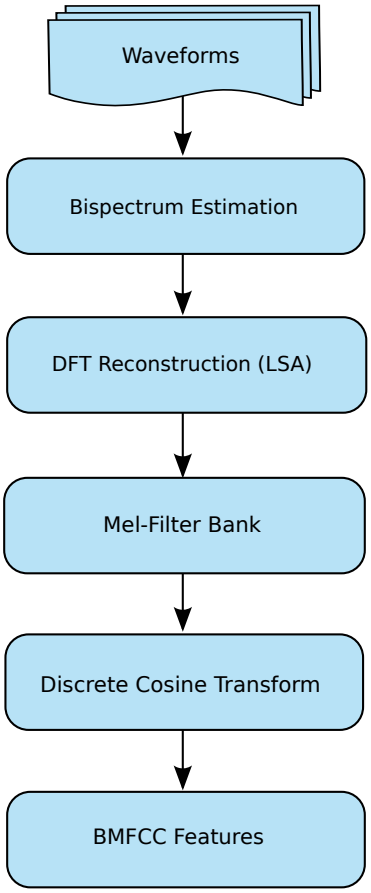


Figure 4.27: BMFCC Feature Extraction

Being a cepstral feature, the higher numbered BMFC coefficients contain less characteristic information and most of the details occur near the origin. Thus only 19 coefficients are considered for the feature vector generation. The computed BMFCCs for some of the underwater targets namely ship1, boat3, finwhale and whale1 are furnished in Fig. 4.28. The plots contain 19 BMFCCs computed from each of the 20 different records of these targets. These plots illustrate the the similarity in the extracted coefficients among the different records of the same target. The variability of these coefficients across different

targets has been illustrated in Fig. 4.29, using the averaged BMFC coefficients computed from 20 records of each of these four targets.

Figure 4.30 presents the BMFC coefficients of three types of boats, averaged over 20 records of each target. The plot shows the variability of the feature set among different targets of the same class. It can be seen that while the general trend of the plots is similar, there is sufficient variation in the corresponding coefficient values among different targets. The values of the corresponding BMFC coefficients computed for the three different boats are also furnished in Table 4.1, to reveal the variabilities in the feature values.

Table 4.1: BMFC Coefficients of Different Boats

Boat1	**Boat2**	**Boat3**
-0.7456	-0.8724	-0.7501
-0.1634	-0.3167	-0.3361
-0.2833	-0.0072	-0.1155
0.1004	0.0816	-0.1955
-0.1909	-0.0329	-0.0831
0.1494	0.1071	-0.0675
-0.0351	0.0725	0.1020
0.0791	0.1059	0.0122
0.0525	0.0838	0.1999
0.1280	0.0692	0.1133
0.0355	0.1072	0.1687
0.2238	0.1516	0.2280
0.1358	0.1382	0.1589
0.2122	0.0984	0.1805
0.0642	0.0839	0.1559
0.1416	0.0833	0.1096
0.0719	0.0444	0.1026
0.0121	0.0094	0.0199
0.0116	-0.0074	-0.0037

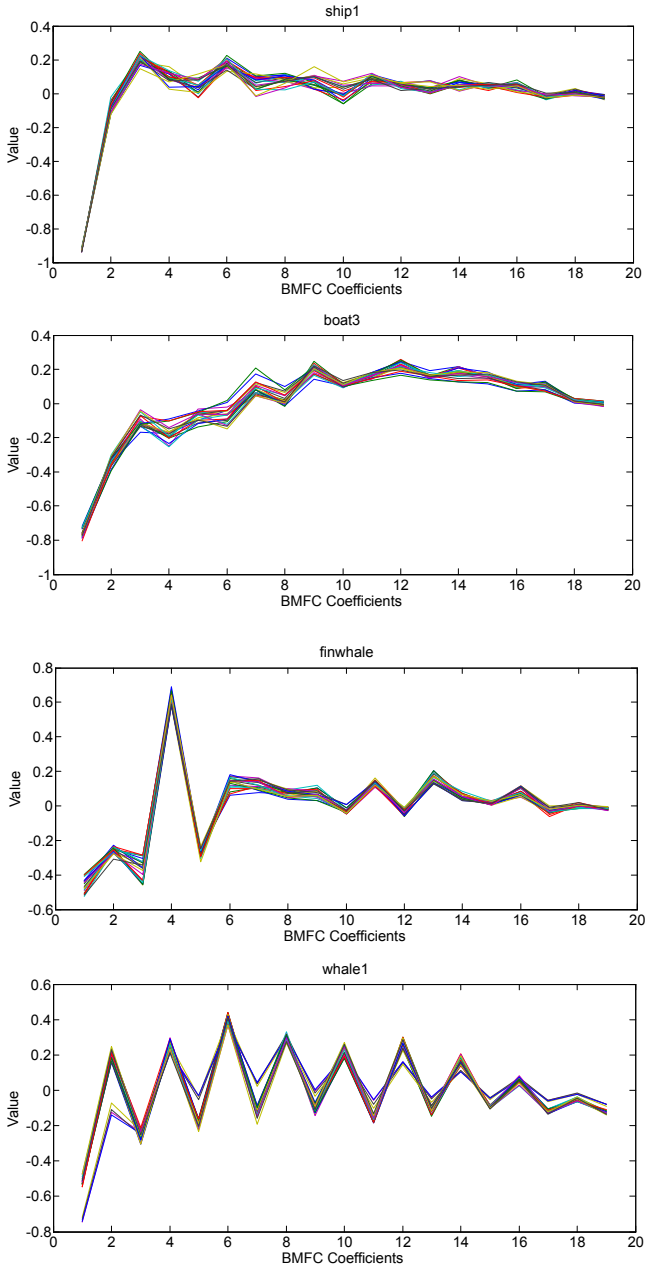


Figure 4.28: BMFCC Plots of Four Underwater Targets

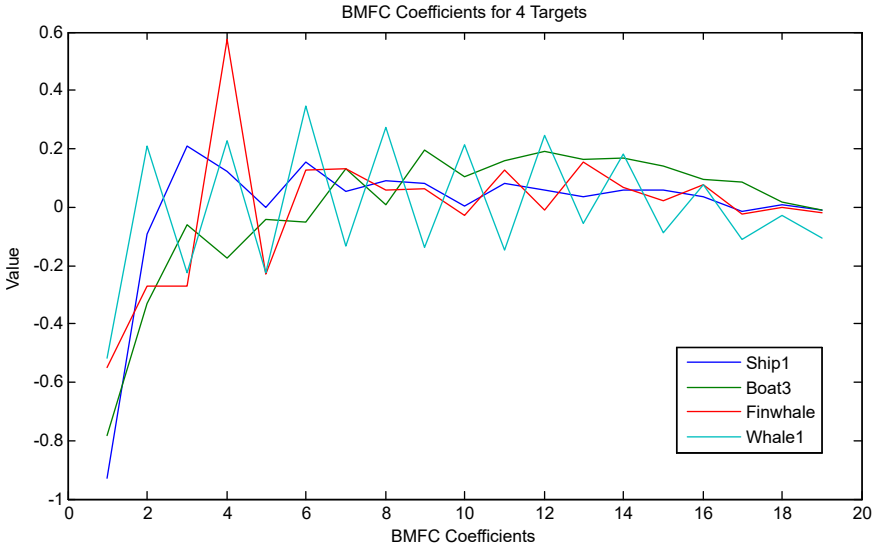


Figure 4.29: Averaged BMFCC Plots of Four Underwater Targets

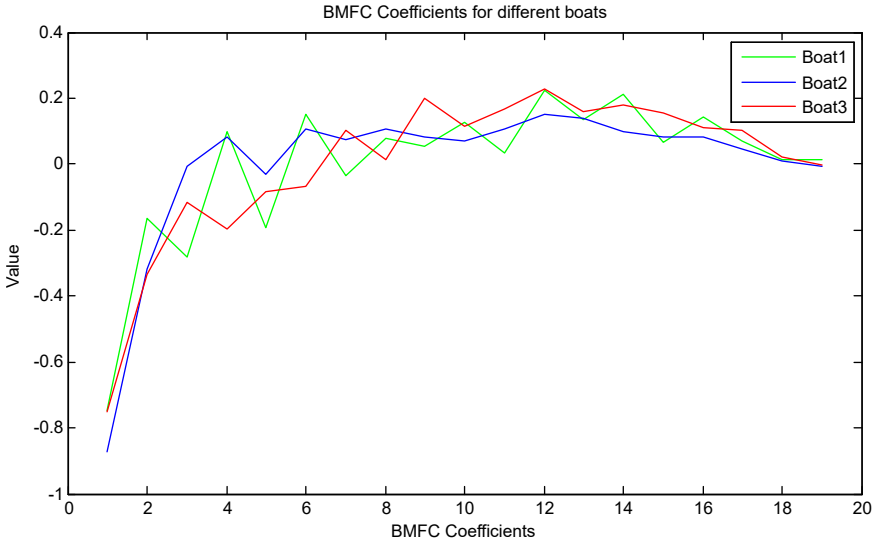


Figure 4.30: Averaged BMFCC Plots of Different Boats

4.9 Bispectral GTCC

4.9.1 Computation of BGTCC

A gammatone filter bank, with M channels has been implemented using Eq. (4.43), with suitable filter order and bandwidth. The magnitude of the Fourier components of the signal has been computed from the bispectral estimate using Least Squares Approach, and the power spectrum is generated.

The filter bank generated is applied to the spectrum of the signal, emphasising the perceptually meaningful frequency components. The filter output of the m^{th} Gammatone filter, X_m has been computed using the equation,

$$X_m = \sum_{k=0}^{\frac{N}{2}-1} P_{LSA}[k].H_m[k] \tag{4.59}$$

where, $P_LSA[k]$ represents magnitude of the N-point frequency spectrum and $H_m[k]$ is the magnitude of the frequency response of the m^{th} Gammatone filter . Also, $1 \leq m \leq M$.

Now the i^{th} Bispectral Gammatone Cepstral Coefficients has been computed as,

$$BGTCC_i = \sqrt{\frac{2}{M}} \sum_{m=1}^{M} \log(X_m) \cos\left(\frac{i\pi}{2M}(2m+1)\right) \tag{4.60}$$

The algorithm for the computation of BGTCC is furnished below.

Algorithm 1 Computation of BGTCC

```
1:  procedure BGTCC
2:      N ← No. of Samples in a Record
3:      R ← Number of Records
4:      Generate R Records of N samples
5:      i ← 1
6:  loop:
7:      Compute the bispectrum of i^th Record
8:      Extract Principal region of the bispectrum matrix
9:      Compute the magnitudes of the DFT using LSA
10:     Generate the Gammatone filter bank
11:     Apply the filter bank to the computed DFT
12:     Compute logarithm of the filter bank output
13:     Perform DCT to generate the BGTCC
14:     i ← i + 1
15:     if i <= R then
16:         goto loop.
17:     else
18:         close;
```

4.9.2 BGTCC Features

The preprocessed noise data waveforms are segmented into records and the bispectrum of each record is estimated using direct method, with 128 point DFT. From the bispectral estimate, the Fourier amplitudes are reconstructed using the Least Squares Approach. A fourth order Gammatone filter bank has been generated with 50 filters, with the centre frequencies equally distributed on the ERB scale and the filter

bank has been applied on to the computed DFT, followed by logarithmic as well as discrete cosine transformations. The BGTC Coefficients are computed using Eq. (4.60). The first coefficient has been discarded as in the case of MFCC computation procedure. The remaining 19 coefficients form the BGTCC feature set.

The BGTC Coefficients of the noise generated by typical manmade noise sources and some marine species are illustrated in Fig. 4.31. The manmade sources include ship noise and boat noise, while the biological noise include that of a fin whale and whale1. Each plot has 19 BGTC Coefficients computed for 20 different records of a particular target. Comparing this with the BMFCC plots in Fig. 4.28, it can be observed that the BGTCC plots are little more bundled together, indicating the similarity of the features with in a class. Thus the intra class variance for the BGTCC can be considered less when compared to BMFC coefficients.

The averaged BGTC coefficients of these four targets, over 20 records each, are plotted in figure Fig. 4.32 to illustrate the inter class variabilities of the coefficient values. The averaged values of the coefficients are also furnished in Table 4.2. Further, the BGTC coefficients of boat1, boat2 and boat3, which are different targets of the 'boat' class, have been plotted in figure Fig. 4.33 and the corresponding values are also furnished in Table 4.3. As can be seen from the table, the coefficient values among different boats vary, even though they are from the same class.

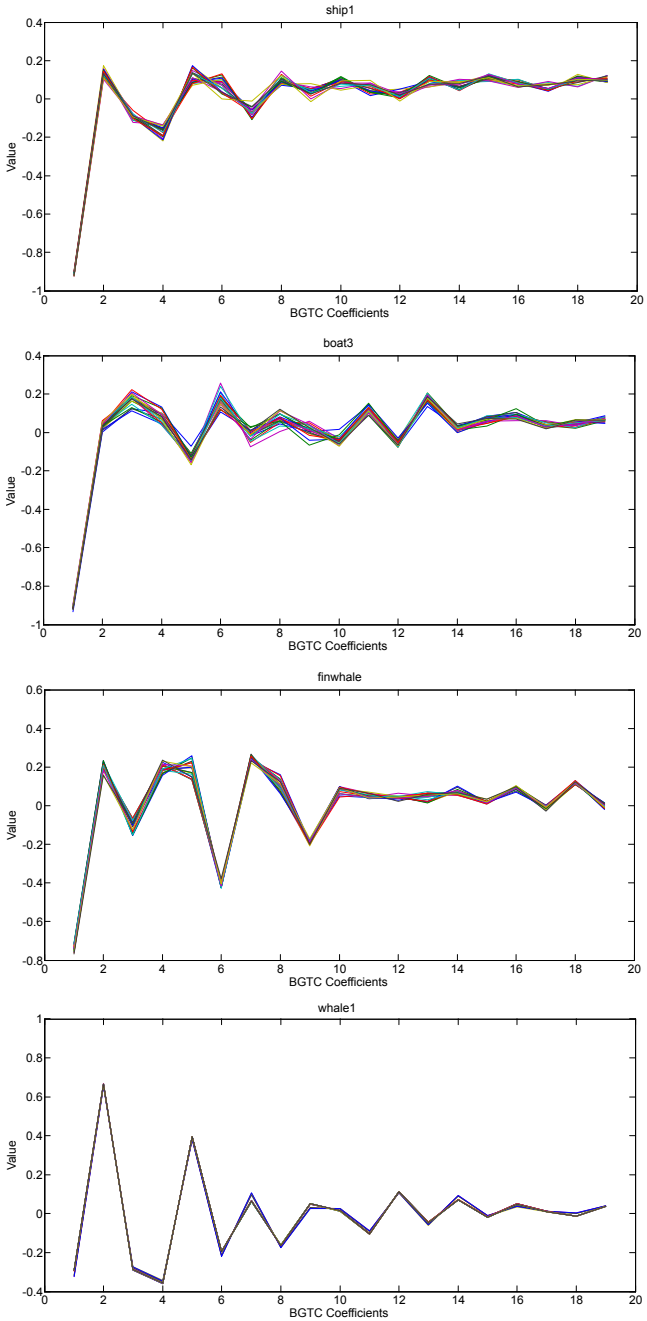


Figure 4.31: BGTCC Plots of Four Underwater Targets

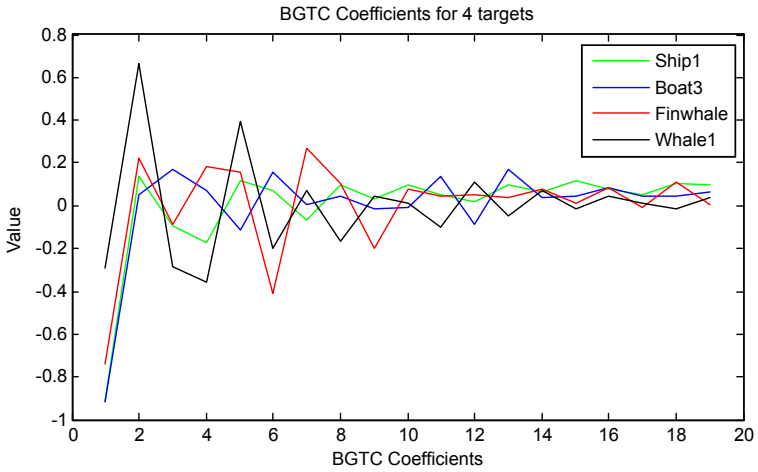


Figure 4.32: Averaged BGTC Coefficients of 4 Different Targets

Table 4.2: BGTC Coefficients of 4 Different Targets

Ship1	**Boat3**	**Finwhale**	**Whale1**
-0.9131	-0.9153	-0.7398	-0.2925
0.1356	0.0524	0.2209	0.6656
-0.0925	0.1730	-0.0836	-0.2860
-0.1696	0.0735	0.1844	-0.3555
0.1201	-0.1159	0.1546	0.3949
0.0694	0.1600	-0.4111	-0.1971
-0.0655	0.0059	0.2710	0.0687
0.0984	0.0461	0.1037	-0.1627
0.0337	-0.0141	-0.1994	0.0482
0.0976	-0.0098	0.0756	0.0146
0.0514	0.1362	0.0453	-0.1016
0.0160	-0.0842	0.0487	0.1117
0.0964	0.1703	0.0394	-0.0479
0.0679	0.0404	0.0768	0.0732
0.1153	0.0433	0.0147	-0.0172
0.0804	0.0851	0.0865	0.0469
0.0540	0.0443	-0.0072	0.0108
0.1063	0.0449	0.1133	-0.0110
0.0981	0.0639	0.0064	0.0370

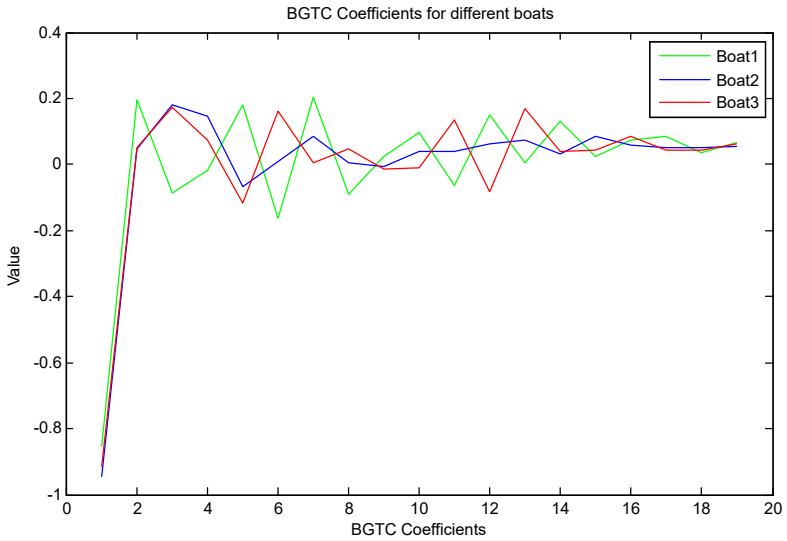


Figure 4.33: Averaged BGTC Coefficients of Different Boats

Table 4.3: BGTC Coefficients of Different Boats

Boat1	**Boat2**	**Boat3**
-0.8543	-0.9437	-0.9153
0.1958	0.0457	0.0524
-0.0862	0.1789	0.1730
-0.0175	0.1476	0.0735
0.1817	-0.0657	-0.1159
-0.1629	0.0074	0.1600
0.2033	0.0846	0.0059
-0.0903	0.0052	0.0461
0.0250	-0.0080	-0.0141
0.0975	0.0414	-0.0098
-0.0630	0.0386	0.1362
0.1505	0.0619	-0.0842
0.0044	0.0723	0.1703
0.1325	0.0331	0.0404
0.0224	0.0839	0.0433
0.0753	0.0596	0.0851
0.0855	0.0519	0.0443
0.0360	0.0515	0.0449
0.0643	0.0539	0.0639

4.10 Summary

This chapter presents the techniques and procedures involved in extracting the various target specific features required for the proposed underwater target classifier. The chapter begins with an introduction to the concepts of Higher Order Spectral analysis, mainly concentrating on the bispectrum. The chapter presents bispectral analysis and various forms of integrated bispectra and related feature extraction techniques. This chapter also touches upon the concepts of bicepstral analysis, which belongs to an area of homomorphic signal processing. The Mel Frequency Cepstral Coefficients and Gammatone Cepstral Coefficients, which are biologically inspired feature extraction techniques have been considered and extended to the bispectral scenario to derive the Bispectral MFCC and Bispectral GTCC.

Chapter 5

The Target Classifier

This chapter describes the various classification schemes implemented for identifying the noise sources in the ocean using the higher order feature set extracted from the noise emissions of the targets of interest. A description of the knowledge base used to evaluate the performance of the classifiers is also included. The realisation of k-Nearest Neighbour (k-NN), Artificial Neural Network (ANN) and Support Vector Machine (SVM) classifiers is studied in detail. The chapter also discusses the various steps involved in the implementation of different classification frameworks, along with their performances. Each classification framework is characterized by the feature selection criterion and the classification algorithm.

5.1 Background

It can be fairly assumed that the noise generated by a source carry certain characteristic information of the source, and by employing proper techniques, this information can be extracted, finally leading to the classification of the noise source. In the case of an underwater target classifier, the noise emanated from different sources are received using the hydrophone arrays which are typically a part of the passive sonar system. The received signals are further processed using suitable feature extraction techniques for extracting the target specific features.

There have been many feature extraction techniques, which are primarily based on the second-order spectral analysis. Such traditional techniques can fail to provide acceptable confidence levels during classifications, especially when there are deviations from Gaussianity and linearity. Thus, there is a surge of interest in making use of the concepts of Higher Order Spectral analysis. Once the features are extracted, the latent information in the feature set needs to be properly interpreted to identify the target and can be achieved by making use of suitable pattern matching techniques.

This chapter discusses the implementation of the underwater target classifier with the proposed higher-order feature set. k-Nearest Neighbour (k-NN), Artificial Neural Network (ANN) and Support Vector Machine (SVM) are the three techniques implemented to realise the underwater classifier. Two feature selection algorithms, which can select the appropriate features from the original feature set has been used to aid better classification capability.

A supervised classification procedure can be divided into two phases,

namely the training phase and validation or test phase. During the training, the classifier is presented with the features already collected and stored in the knowledge base, with their class labels. At this stage, the classifier gets adapted itself and captures the information contained in the features, so that it can classify the unknown signals presented to it, during the validation phase. In the validation phase also, the features are extracted from the test signal, processed and is finally fed to the classifier, which will give the class label of the test signal based on certain pattern matching criterion. The k-NN classifier slightly differs from this concept, as there is no specific learning stage, but the test signal is compared with the templates stored in the knowledge base at the time of validation to get the class label.

5.2 Creation of the Knowledge Base

5.2.1 The Noise Sources

The knowledge base contains relevant parameters extracted from the noise data of different types of targets of biological and anthropogenic nature. The noise data collected from the different targets are processed, and the characteristic features are extracted to create the knowledge base. It is a well-known fact that, the robustness of the knowledge base is one of the basic requirements for achieving good classifier performance. A description of the targets and their characteristics that have been used for the creation of knowledge base in this work is presented in the following section. Some of the anthropogenic noises have been recorded during scheduled cruises, while the other data have been downloaded from open source repositories in the Internet..

5.2.1.1 Anthropogenic sources

(a) Ship

The knowledge base contains features extracted from the noise emanations of three ships. One of the ship recordings was carried out during the cruise #321 of the Fishery Oceanographic Research Vessel Sagar Sampada, in the south-eastern Indian Ocean. Recording from a large commercial ship cruising at approximately 18 kilometres per hour at a distance of about 3 km away from the hydrophone deployment site was collected. Another recording collected from the same site was that of an oil tanker having a speed of about 22 kilometres per hour.

Another two recordings of the ship noise were carried out along the shipping channel, 5 km off-shore the Cochin transhipment terminal. The hydrophones were deployed from a boat with its engine and echo sounder turned off, to avoid the self-noise. The recordings constituted that of a merchant vessel cruising at approximately 25 kilometres per hour, about 1.5 km away from the measurement location and that of a hopper dredger spotted at a distance of about half a kilometre, moving past the recording site at 11 kilometres per hour. Apart from these, the recording of a ship Klaxon is also included in the knowledge base.

(b) Boats

Signals from 3 boats and an outboard motor were recorded. One of the boat recordings was that of a ferry, a vessel that is used to carry passengers, vehicles as well as cargo, across the Cochin backwaters. Another noise recorded was that of a tugboat, which is a type of vessel that manoeuvres other vessels by pushing or pulling them. As the propellers of a typical tugboat are recessed to avoid the damage in case

of grounding, they produce less near-surface sound compared to a ship, due to its reduced cavitation noise. Trawling is a method of fishing that involves pulling a fishing net through the water behind one or more boats, and the boats used for this purpose are called trawlers. The recording of a 120 HP trawler boat was carried out on its way back to the shore.

Outboard Motors are one of the most common methods of propelling small boats and watercrafts that are used in coastal areas. The signal from a 50 HP outboard motor boat was recorded. The recordings of the noise from a small Zodiac 35 HP engine has also been included in the knowledge base. One can hear the engine starting, going into the gears and the noise generated due to propulsion. Underwater propeller cavitation noise is composed of tonal blade rate noise and high-frequency broadband noise. The cavitation noise generated by a motor-driven propulsion mechanism has also been included in the knowledge base.

(c) Torpedo

Torpedoes are weapons designed with an explosive warhead, that can be self-propelled underwater and explode on reaching a target. The recording of the sound of a live torpedo shot obtained from the Internet is also considered. After the firing, the closing sound of the hatch followed by the sound of the propulsion through the water is recorded in the signal. The propulsion sound is mainly due to the propellers, propulsion machinery and cavitation at the nose. The duration of this propulsion sound depends on the distance to the target. A large explosion can be heard towards the end, that corresponds to the detonation

(d) Submarines

Completely underwater submerged operations characterize Submarines, one of the most high-tech vessels that can perform stealth operations, reconnaissance and rescue missions. The design of the submarines constitutes a cylindrical body with hemispherical ends and with pressure hulls that are resistant to deep water pressure. There are also ballast tanks, propeller and propulsion compartment, engine and hydrodynamic control fins. There are several mechanisms for noise radiation from the submarine which includes power plant equipment, propellers, the vibration of the hull and flow noise (hydrodynamic noise). The submarine noise available in the Internet has also been included in the database.

Time series waveforms of two typical man-made targets viz., a boat and a merchant vessel are depicted in Fig. 5.1.

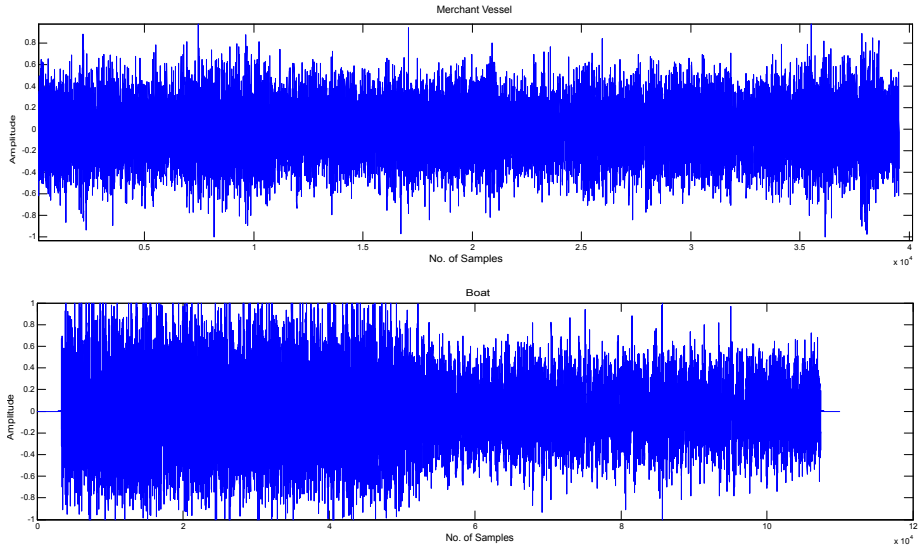


Figure 5.1: Typical noise signals of a boat and merchant vessel

5.2.1.2 Natural sources

(a) Beluga whales

Beluga whales are highly vocal and rely on sound to communicate, navigate, and locate breathing holes. They employ echolocation for their navigation. Belugas are known as 'canaries of the sea' due to the wide range of sound they produce which include sounds like clicks, moos, whistles and clangs, which span a broad range of frequencies. Social sounds, including calls, whistles and buzzes are also common. The sounds recorded are mostly in the frequency range of 0.1 to 12 kHz.

(b) Toadfish

The toadfish, with its name derived due to the button-like markings, is known for its vociferous night-time mating call. During the summer mating season, the male toadfish vibrate the muscles of their swim bladder to invite the females to their nests. These calls are known as *boatwhistles* and may range between 250 to 650ms in duration with a fundamental frequency of about 180 Hz. Males also emit a single, short duration pulse or 'grunt' of short duration of about 100 ms, which are emitted almost exclusively during the *boatwhistles* of a conspecific male which can jam the signal.

(c) Damselfish

Damselfishes are renowned marine chatterboxes and use their swim bladders to produce sound. They generate 'pops' and 'chirps' during courtship and also to exhibit aggression while defending their territory. Males produce pulsed sounds during the courtship behaviour which are

known as the *signal jump*, in which they rise in the water column and then rapidly swim downward while producing a pulsed sound. Pops contain only one or two pulses and are commonly made towards heterospecifics than conspecifics. Aggressive chirps consist of 3 to 11 pulses and are usually made towards conspecifics. The pulse rate of aggressive chirps has been found to be faster than signal jump sounds. The number of pulses produced during the pops and chirps is species-specific, with peak frequencies varying between 100 Hz to 4 kHz.

(d) Fin whale

The fin whale, the second largest mammal, reaching lengths upto 24 m produce high-intensity calls with a frequency range of 1528Hz. The vocalisations of blue and fin whales are the lowest-frequency sounds made by any animal. Each sound lasts one to two seconds, and various sound combinations occur in patterned sequences lasting 7 to 15 minutes each.

(e) Searobin

Sea robins, also known as gurnards are a family of bottom feeding fish with armoured bony heads and two dorsal fins. Their pectoral fins are fan-shaped, which, when swimming, open and close like a bird's wings in flight. They have a "drumming muscle" that makes sounds by beating against the swim bladder. The sounds have a mean duration of about 60 ms with a fundamental frequency of approximately 200 Hz with harmonics at approximately 400 Hz and 600 Hz.

(f) Minke whale

The minke whale, the smallest of the baleen whales, is found

globally in tropical, temperate, and polar waters. Their vocalisations comprise of very low frequency, long wavelength repetitive pulses or groans. Vocalizations of the Minke from the Great Barrier Reef area are popularly known as "star wars call. They span a wide frequency range (50 Hz to 9.4 kHz) and are composed of distinct and stereotypically repeated units with both amplitude and frequency-modulated components.

(g) Pinnipeds

Pinnipeds, commonly known as seals, have a natural amphibian history resulting in some fascinating sensory capabilities. They emit and receive sounds both in air and water. Pinnipeds produce vocalisations in a variety of social contexts, notably reproduction and aggression. The effective bandwidth of pinniped vocalisations ranges from 10 Hz to 15 kHz.

(h) Scad Fish

Scad fish refers to any of several species of fishes in the family Carangidae. Scad fishes have forked tails and reach lengths of about 25 to 30 cm. They produce sound by grinding of the pharyngeal teeth. Enormous schools may collectively produce low-frequency hydrodynamic sounds while swimming. Narrowband spectral analysis of their echoes has revealed significant structures in the frequency range from 200 Hz to 5 kHz.

(i) Grunt Fish

Grunt fish is a member of the family Haemulidae. They are usually found along the shores in warm and tropical waters. They produce pig-like grunts with their pharyngeal teeth using a process known as

stridulation.

(j) Sculpin

The sculpin is a catadromous fish with no swim bladder. They have a pair of sonic muscles that produce sounds by rapidly contracting the muscles. The sound produced by sculpin is composed of a series of pulses lasting for 40-60 ms, with the peak frequency at approximately 100 Hz.

(k) Croaker

The croaker is an important commercial fish, which "croak" by vibrating their swim bladders with special sonic muscles as part of their spawning or fright response. Compared to other Sciaenids, where only the males possess sonic muscles, both the male and female croakers have sonic muscles.

(l) Sperm whale

Sperm whales produce clicking and creaking sounds for echolocation purposes. Clicking sounds are used to gather information about acoustically reflective features, and creaking sounds are used while foraging. They are mostly silent while on the surface. The clicks of male sperm whale contain higher frequency components with energy ranging up to 12 kHz and with distinctive peaks at 400 Hz and 2 kHz.

(m) Humpback Whale

Humpback whales are medium-sized baleen whales which have shown to produce complex vocalisations. They produce structured series of vocalisations, termed as 'song', as mating calls as well as 'social sounds' while on their low latitude wintering grounds. Songs are repeated, continuous rhythmic sound patterns that can reach upto 30

minutes duration. The overall frequency range of these songs has been estimated to be 20-1900 Hz. They also produce unpatterned sounds associated with their social and feeding behaviour. Certain vocalisations also appear to be coordinated for organised feeding activity which are found to span the frequency range 40 Hz 1250 Hz.

(n) Gray Whale

Gray whale is a type of baleen whale that migrates between feeding and breeding grounds annually. The common sounds generated by the gray whale are moans, clicks, knocks and metallic bonging sounds occurring as pulses in quick succession with frequencies ranging from 20 Hz to 3 kHz.

(n) Snapping Shrimp

Snapping shrimps are found in shallow tropical and subtropical waters wherever rock, coral or other materials on the bottom provides interstices on which they can thrive. They produce sound by the snapping of their claws. The sound produced by a shrimp colony can be so loud that the sonars may miss other proximate targets.

(o) Other Sources

Rain falling on the ocean can be a source of underwater noise, which can considerably increase the ambient noise levels. The knowledge base also includes light rain recorded just below the surface. The sound is mainly generated due to the impact of the raindrop hitting the ocean surface combined with the sound that gets radiated from bubbles trapped underwater during the splash. Another natural source incorporated into the knowledge base is the sound generated during the cracking of ice. The noise is generated when the ice cracks, when waves, icebergs or wind

pushes against it.

Some of the typical target signals of biological origin have been depicted in Fig. 5.2.

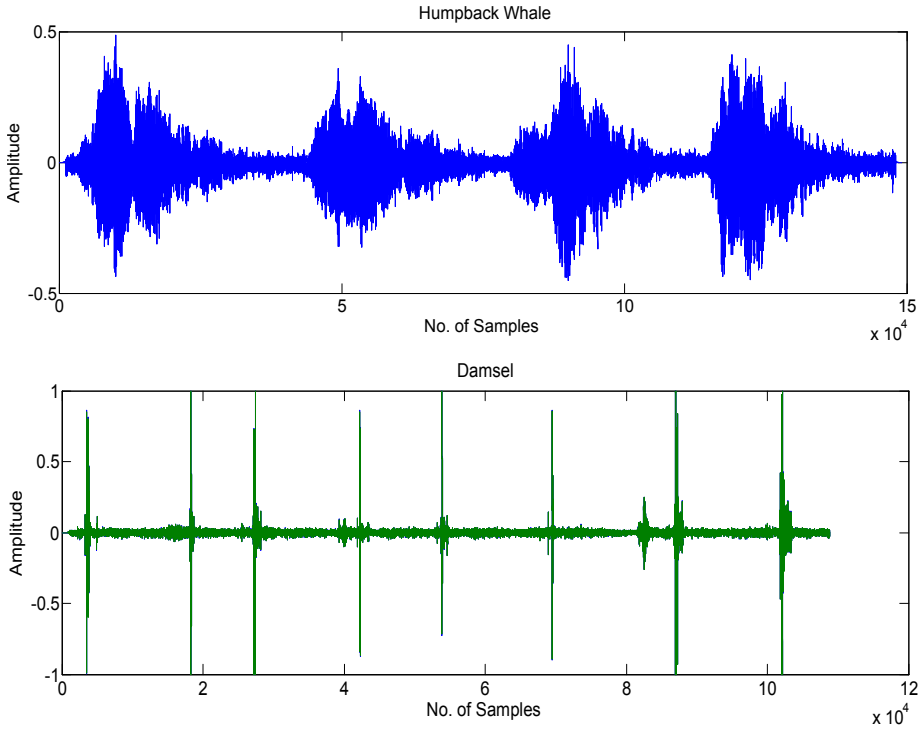


Figure 5.2: Typical marine biological noise signals

5.2.2 Generation of the Feature Vector

For generating the feature vector, the noise data waveforms of various targets are processed using the procedures of higher order spectral analysis as described in the previous chapter. Each preprocessed target waveform, sampled at 22.050 kHz, is segmented into fixed size records, and each record is analysed to extract the features using different procedures. The feature sets so obtained are

concatenated to get the feature vector for that record. The feature vectors of all the records of a particular target constitute the raw Target Feature Vector.

In the underwater scenario, the dimensionality of the original measurement space created by the observation vectors will be massive. However, for each target/sources emitting acoustic radiation, there will be certain associated attributes or characteristic features that are unique to the individual class of targets. The feature selection algorithms are designed to rank the features according to some criteria and select the most relevant features from the feature sets to form an optimal feature vector, which would be dimensionally smaller than the original feature vector.

Based on the belief that there exist some low dimensional qualitative or representative features in the measurement vector, which characterise the nature of the actual target, the feature selection stage tries to estimate this low-dimensional subspace to form the final Target Feature Vector, which contains the most salient features with reduced redundancy. The rest of the features are assumed to be trivial, often considered as noise and are discarded, thereby reducing the dimensionality of the measurement space and the computational complexity, significantly.

Since each feature used as part of a classification procedure can increase the cost and execution time of a recognition system, there is a natural tendency to use smaller feature sets. At the same time, there is a potentially opposing need to include a sufficient set of features to achieve reliable recognition rates under difficult conditions. The number of features selected by the feature selection stage has been

varied and its effect on the classification accuracy has been analysed. It is found that a final Target Feature Vector with 32 features could give optimal performance. The variation of the success rates for a typical SVM classifier with the number of features selected is as shown in Fig. 5.3. The classification performance was poor with less number of features and started increasing as the number of features increased, reaching a maximum value when 32 features was considered, beyond which the performance started declining.

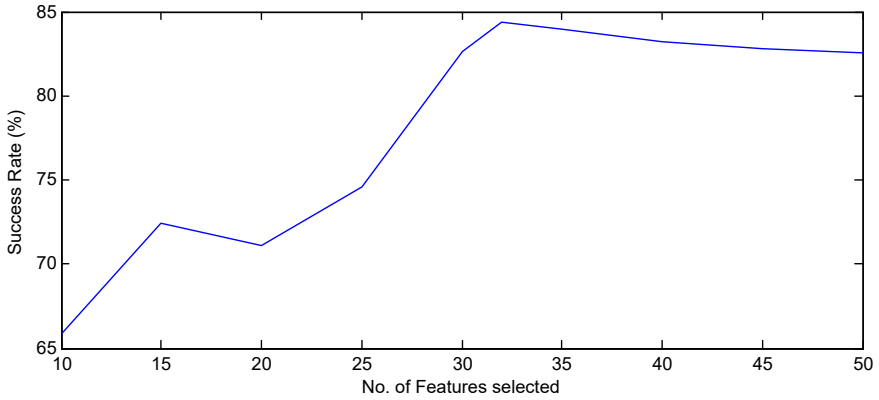


Figure 5.3: Variation of success rate with number of features

A flowchart illustrating the complete procedure for generating the Target Feature Vector and creation of the knowledge base is furnished in Fig. 5.4.

The final Target Feature Vector after the feature selection process for all the targets, averaged over 20 records each, has been presented in Fig. 5.5 through 5.7. The features being significantly different for different underwater targets, can characterise the corresponding targets.

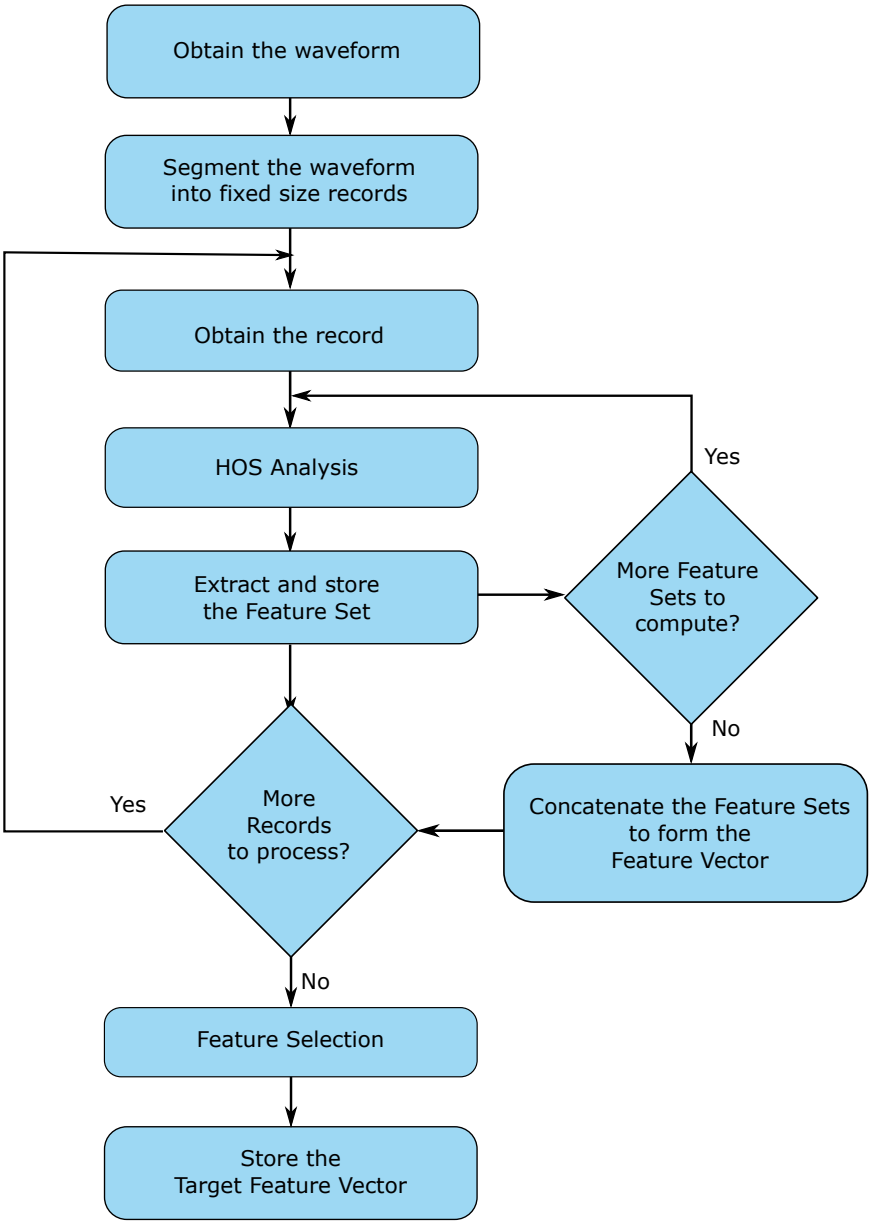


Figure 5.4: Flowchart for Target Feature Vector generation

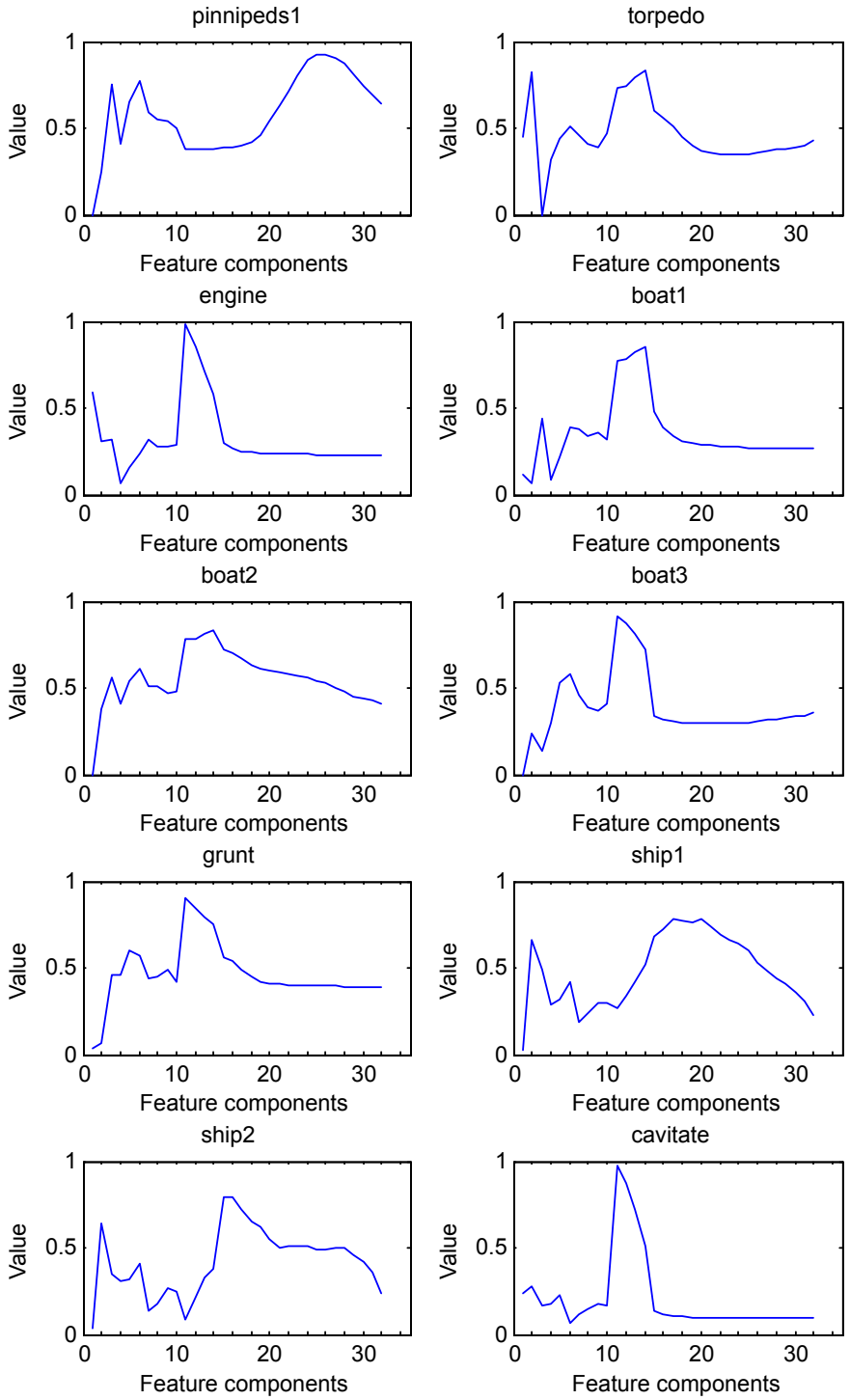


Figure 5.5: Target Feature Vectors for Targets 1 to 10

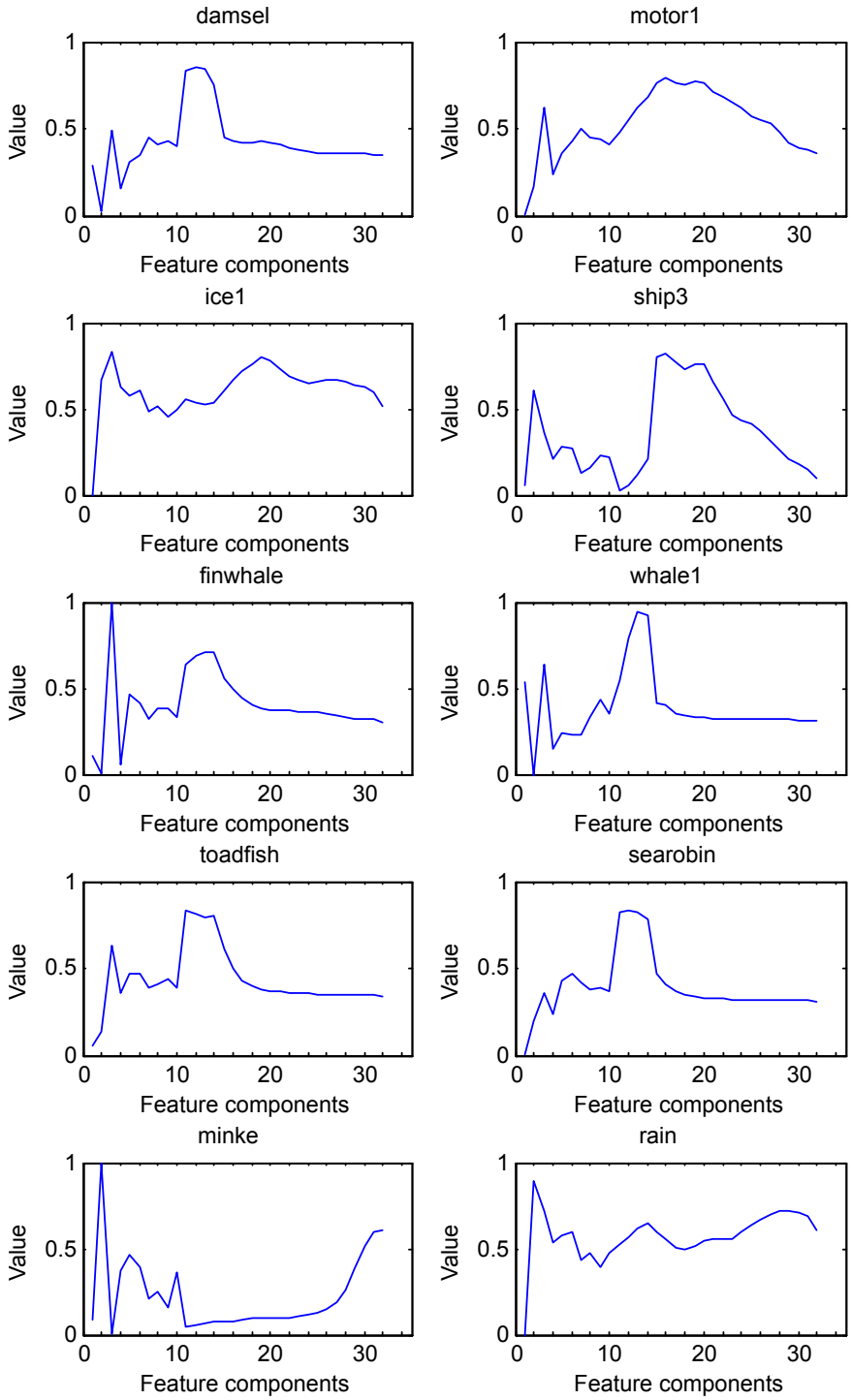


Figure 5.6: Target Feature Vectors for Targets 11 to 20

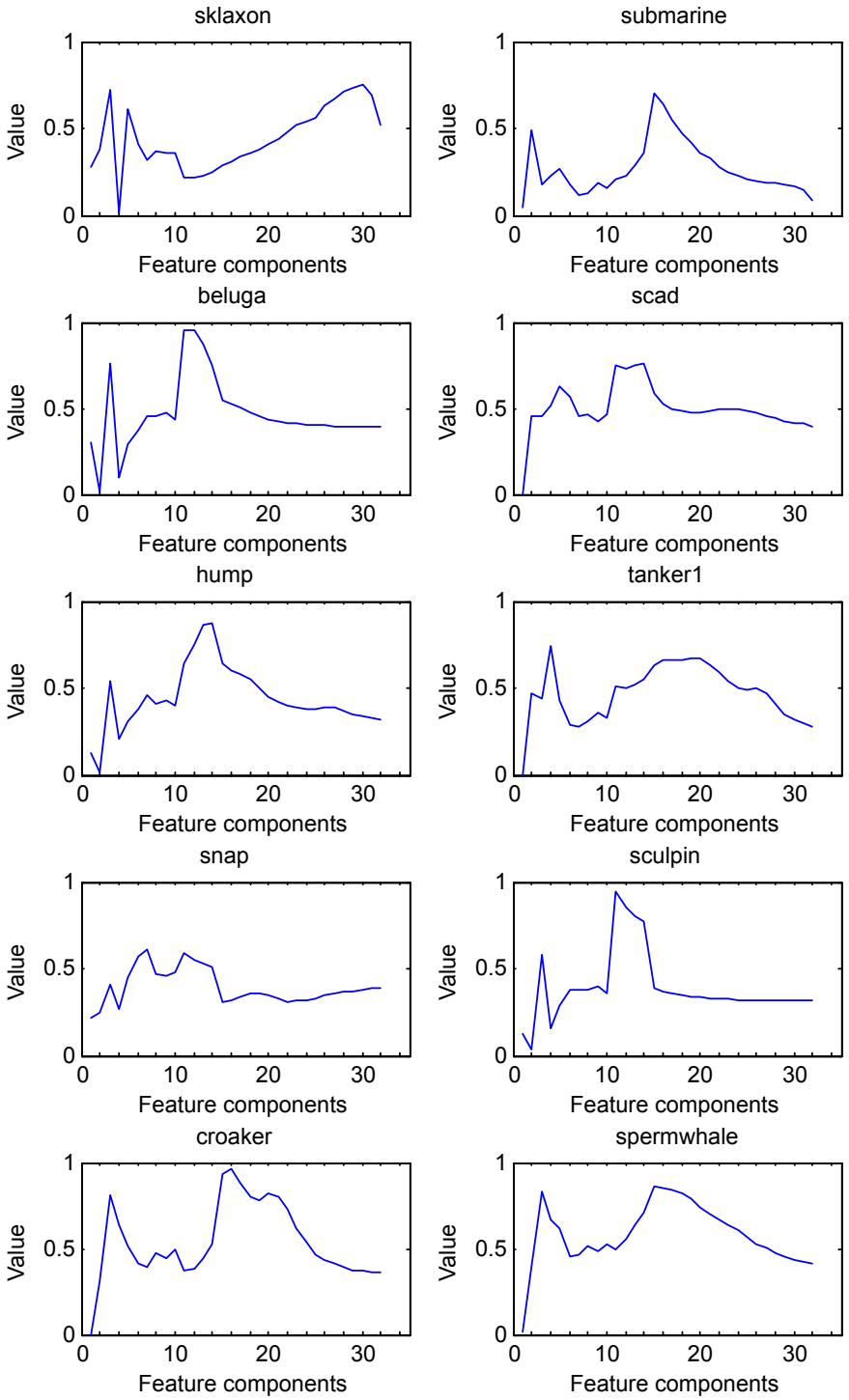


Figure 5.7: Target Feature Vectors for Targets 21 to 30

5.2.3 Implementation of the Prototype classifier

Upon completion of the feature selection procedures, the optimal target feature subset obtained will serve as the knowledge base. The target classifiers are trained and evaluated with the help of the resulting knowledge base. The following sections describe the different classifiers, their parameter selection and validation.

5.3 k-NN based Underwater Target Classifier

An underwater target classifier, which can identify the targets from their noise emanations has been implemented using the k-Nearest Neighbour algorithm.

An underwater target classifier, which can identify the targets from their noise emanations has been implemented using the k-Nearest Neighbour algorithm. The main advantage of k-NN is its simplicity, and its highly intuitive nature makes it easy to understand and implement. However, being an instance-based classifier, there is no explicit training stage, and as a result, the distance of each query instance to all training samples needs to be computed at the time of classification, which increases the computation time, especially if the training set is large. Also, the classification accuracy depends on the choice of the distance measure and the value of the parameter k. The effect of the value of k on classification is illustrated in Fig. 2.1.

The other parameter that needs to be optimally chosen is the distance measure. While one of the commonly used distance measures is the Euclidean distance, the use of other metrics such as city block or

Manhattan, cosine, correlation distance, etc., are also in practice. If x and y are the N dimensional feature vectors representing two instances, then the Euclidean, city block and correlation distance measures can be defined as:

$$\begin{aligned} \textit{Euclidean}: \quad & d_e(x,y) &=& \sqrt{\textstyle\sum_{i=1}^{N}(x_i-y_i)^2} \\ \textit{City block}: \quad & d_m(x,y) &=& \textstyle\sum_{i=1}^{N}|x_i-y_i| \\ \textit{Correlation}: \quad & d_c(x,y) &=& \frac{\sum xy-\frac{\sum x\sum y}{N}}{\sqrt{(\sum x^2-(\sum x)^2/N)(\sum y^2-(\sum y)^2/N)}} \end{aligned}$$

The cosine metric is a similarity measure defined in terms of the cosine of the angle between the two vectors x and y and is defined as,

$$\textit{cosine similarity} = \cos(\theta) = \frac{x \cdot y}{\|x\|\,\|y\|}$$

The simulation was carried out to determine these parameters with the available data in the knowledge base. The variation of the success rates of the classifier with different values of k is depicted in Fig. 5.8. It has been found that the classifier has optimal performance for a value of $k = 7$. Similarly, by varying the distance functions, simulations were carried out. It has been observed that for the k-NN classifier, the Manhattan (city block) distance measure has a slight performance advantage when compared to other distance measures, and the outcome of this study is illustrated in Fig. 5.9.

The performance of the optimal k-NN based classifier has been evaluated with noise waveforms of 30 underwater targets and the success rate of classification of each target has been estimated. For performance evaluation, the whole data set can be randomly partitioned into test and training sets. The partitioning has been carried out in such a way that the training set has 100 records of each

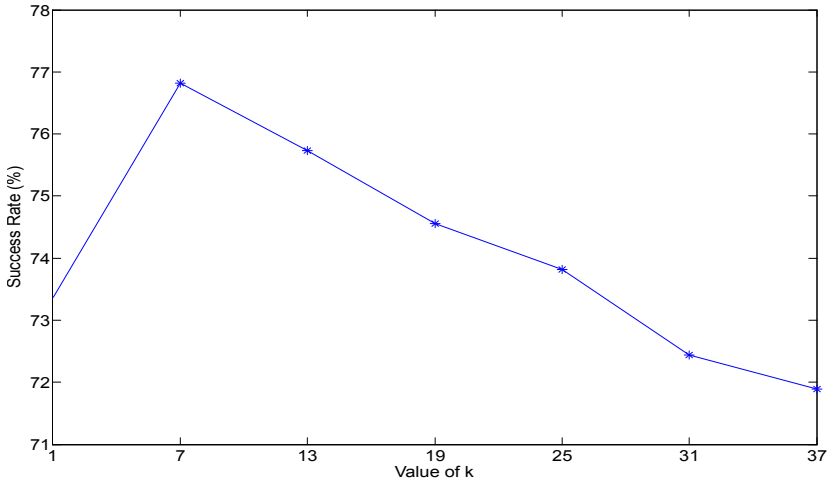


Figure 5.8: Typical variation of success rates for the k-NN classifier for different values of k

target. In the case of the k-NN classifier, the training set will constitute the template vectors, against which the test vectors are compared. The test set is used to evaluate the performance of the classifier. The success rate for each target has been computed as the percentage of the ratio of the correctly classified records to the total number of records tested. The outcome of the performance results is summarised in Table 5.1.

The table shows the success rates obtained for the individual targets on the application of the Fisher and JMI algorithms. It is seen the target damselfish exhibited the lowest success rate of 66.03% and 63.51% respectively for Fisher and JMI feature selection. Average success rates of 78.84% and 80.12% have been obtained respectively when the Fisher and JMI feature selection criterion has been applied. The normalized confusion matrix, rounded to 2 decimal points, obtained with the JMI feature selection has been furnished in Fig. 5.10. The labels T1, T2, . . . , T30 represent the 30 different targets mentioned in Table 5.1, in the

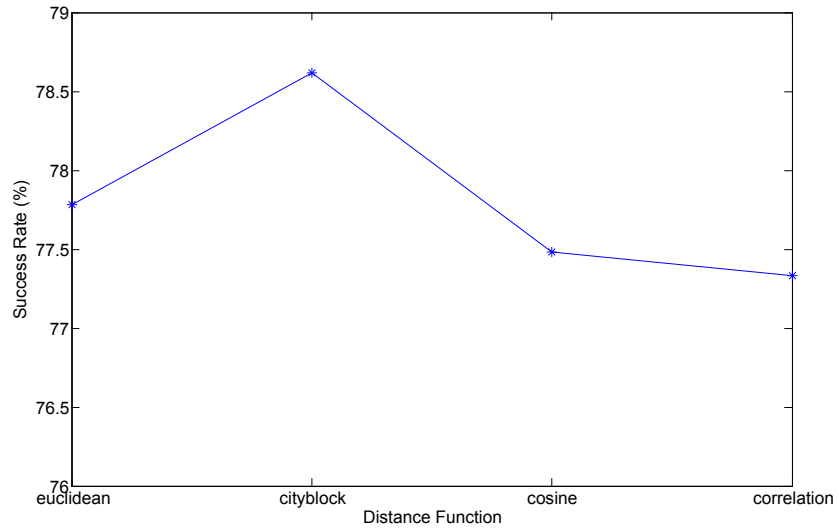


Figure 5.9: Variation of success rates for the k-NN classifier for different distance functions

same order. As seen from the confusion matrix, the lowest success rate corresponds to the target 'damsel', which has a probability of almost 11% to be misclassified as 'beluga'.

5.4 ANN based Underwater Target Classifier

Design and implementation of Artificial Neural Network based target classifier have been successful. As the ANN architecture has many parameters, it is important that one has to find out a proper set of network parameters for its optimum performance. The various parameters that were analysed include the number of hidden layers and the transfer functions for each layer. The number of neurons in the hidden layers was also varied to get the optimal performance. Simulation studies were carried out by varying the relevant parameters to obtain an optimal combination, for making the classification process

Table 5.1: Success Rates of Individual Targets for the k-NN Classifier

Targets	**No. records used for validation**	**Success Rate (%)**	
		Fisher	**JMI**
pinnipeds1	326	83.21	82.54
torpedo	325	81.93	81.65
engine	324	84.44	83.38
boat1	327	77.86	82.03
boat2	324	83.21	82.54
boat3	345	80.58	80.00
grunt	242	78.64	78.17
ship1	220	69.82	75.64
ship2	217	82.32	82.41
cavitate	500	78.65	82.79
damsel	320	66.03	63.51
motor1	341	81.03	82.79
ice1	451	77.36	74.99
ship3	251	81.54	76.47
finwhale	296	78.65	82.98
whale1	324	80.38	78.45
toadfish	184	79.84	78.29
searobin	271	69.91	75.14
minke	336	83.55	82.79
rain	362	73.13	81.42
sklaxon	165	79.61	82.79
submarine	232	82.76	82.08
beluga	411	79.61	81.39
scad	285	82.02	81.19
hump	462	75.58	79.21
tanker1	354	83.01	81.74
snap	326	76.85	82.03
sculpin	336	83.51	79.46
croaker	369	73.82	83.43
spermwhale	500	76.38	82.38
Average Success Rate		*78.84*	*80.12*

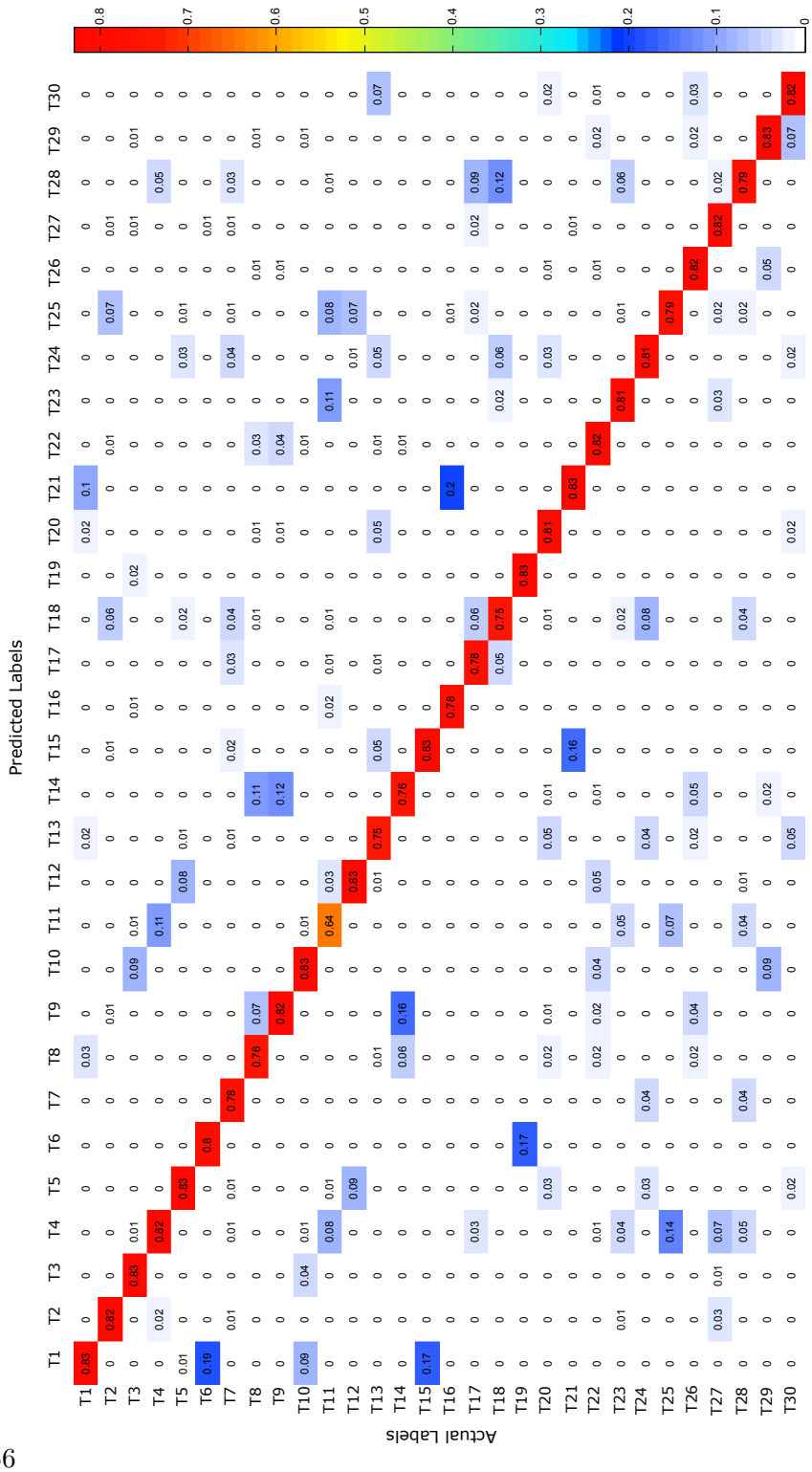


Figure 5.10: The 30 × 30 confusion matrix for the k-NN classifier

more efficient.

From the studies it has been observed that a network with two hidden layers is found to yield satisfactory performance. As the number of hidden layers is increased, the learning process is found to become inefficient, resulting in unacceptable generalization of the network. Simulations were carried out with *linear*, *tan* and *log-sigmoid* transfer functions. Acceptable performance was observed, when the log-sigmoid transfer function was used for all the layers in the network.

The simulation results for ANN based underwater target classifier is as presented in Table 5.2 and shows the success rates for each target with Fisher and JMI feature selection criterion. While the classifier yielded, an average success rate of 83.30% with the JMI feature selection technique, Fisher feature selection criterion resulted in only 81.93%. The normalized confusion matrix corresponding to the JMI feature selection scheme is depicted in Fig. 5.11.

5.5 SVM based Underwater Target Classifier

A Support Vector Machine classifier with an RBF kernel has been implemented and trained with the feature vectors extracted using the higher order spectral analysis. In line with the earlier classifier systems, here also the Target Feature Vectors were partitioned randomly into test and training sets, and for each target, 100 training records were used for training the SVM classifier. The SVM classifier was trained with the training set, and the performance validation was carried out with the test set to obtain the success rates of each target. The procedure was repeated ten times, with different randomly selected test sets, and the

Table 5.2: Success Rates of Individual Targets for the ANN Classifier

Targets	**No. records used for validation**	**Success Rate (%)**	
		Fisher	**JMI**
pinnipeds1	326	84.28	85.60
torpedo	325	83.54	87.97
engine	324	84.14	86.08
boat1	327	79.36	85.01
boat2	324	84.65	83.19
boat3	345	84.29	85.10
grunt	242	80.14	80.48
ship1	220	81.91	77.87
ship2	217	83.30	85.35
cavitate	500	84.33	85.78
damsel	320	68.13	65.63
motor1	341	82.37	85.74
ice1	451	78.34	78.06
ship3	251	84.65	86.07
finwhale	296	84.41	86.02
whale1	324	81.41	81.11
toadfish	184	80.42	80.47
searobin	271	71.13	77.78
minke	336	84.05	85.67
rain	362	84.29	84.06
sklaxon	165	84.00	85.01
submarine	232	84.07	84.46
beluga	411	81.15	84.61
scad	285	82.68	83.68
hump	462	78.02	81.86
tanker1	354	83.87	84.78
snap	326	82.28	85.24
sculpin	336	84.77	84.89
croaker	369	84.08	86.05
spermwhale	500	83.73	85.35
Average Success Rate		*81.93*	*83.30*

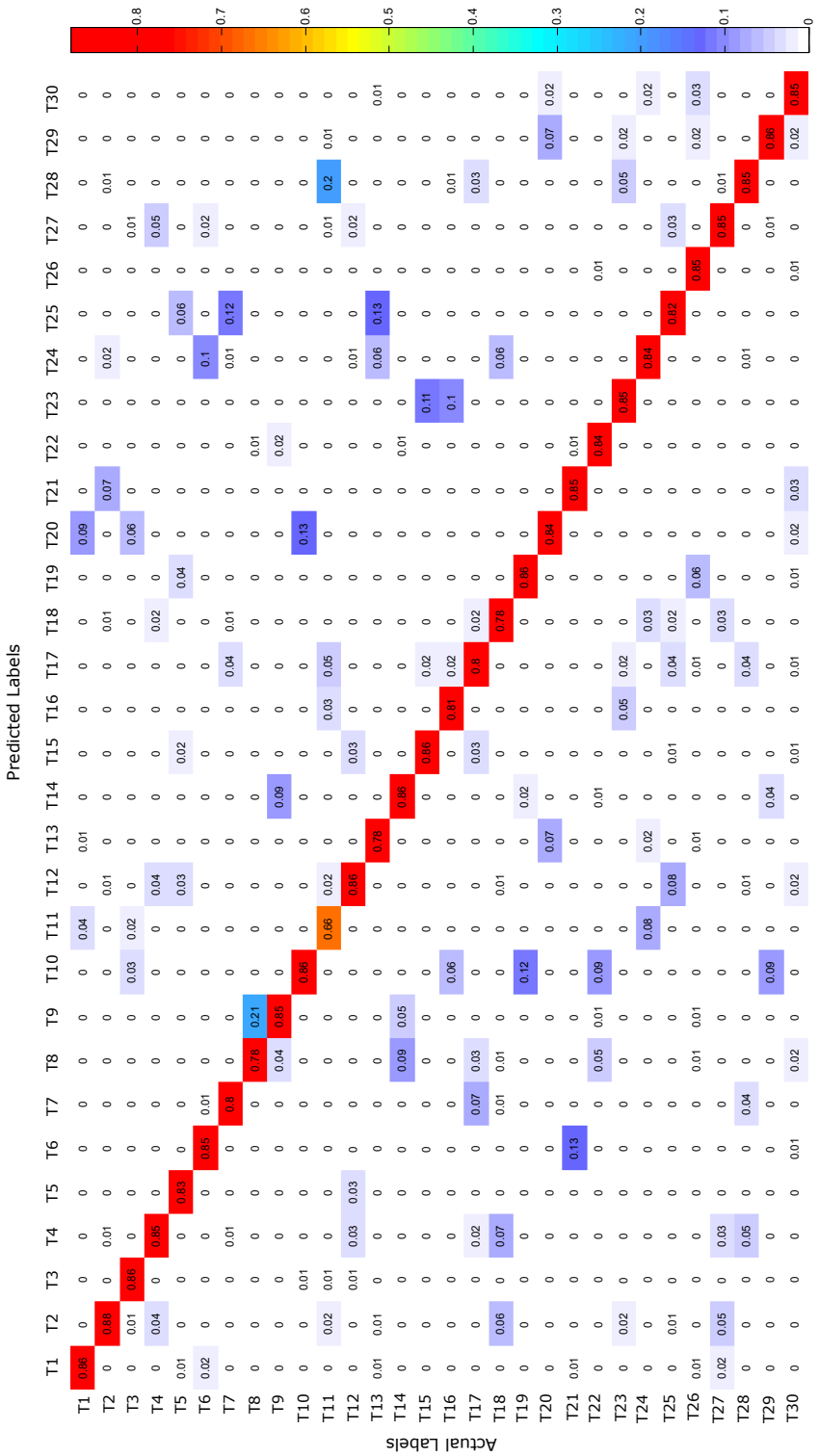


Figure 5.11: The 30 × 30 confusion matrix for the ANN classifier

average success rates have been computed, in order get a statistically reliable result. The success rates for each target has been computed as the percentage of the ratio of the correctly classified records to the total number of records tested.

The choice of the classifier parameter γ, can considerably affect the classifier performance as discussed in section 2.6.3. Simulations are carried out to estimate the optimal value of the classifier parameter γ, using a method similar to the grid search algorithm [89]. The average success rates of the SVM classifiers corresponding to different values of γ have been determined, keeping all other conditions unchanged. The obtained results depicting the variations of the success rates of the classifier for different values of γ are shown in Fig. 5.12. It is found that, for the HOS feature sets used, the maximum performance for the classifier has been obtained when the value of $\gamma = 0.1$.

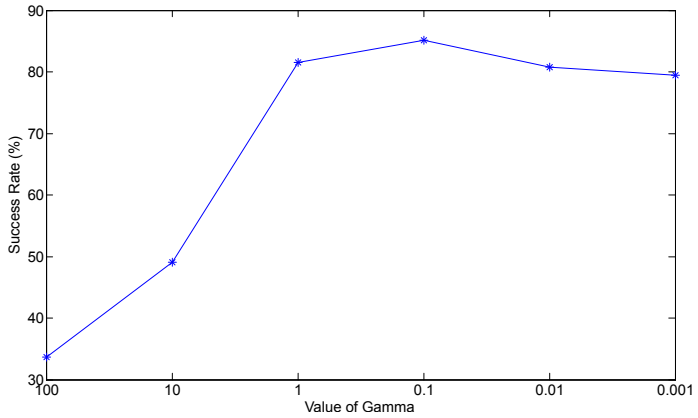


Figure 5.12: Variation of the success rate for different values of γ for SVM classifier

The success rates of each target, for the SVM classifier with an optimal value of γ has been furnished in Table 5.3, which gives the number of records used for validation and the corresponding success

rates in percentage for the different classes of individual targets. The success rates have been estimated by employing the two different feature selection techniques, namely Fisher's criterion and JMI. It can be seen that the classifier performs with acceptable success rates for the proposed HOS feature sets. An average success rate of 84.18% and 84.49% has been obtained with the Fisher criterion and JMI respectively. It is observed that the average success rate obtained for the JMI feature selection technique is slightly higher when compared to that of the Fisher's criterion.

The normalized confusion matrix corresponding to the JMI feature selection scheme is as shown in Fig. 5.13. The classification accuracy of each target, along with the false positives and false negatives are displayed. As seen from the confusion matrix, the success rate corresponding to the target 'damsel', which was low in the case of k-NN and ANN classifiers, has been considerably improved in the case of SVM classifier. T10 ('cavitate') with a score of 0.89 has the highest probability of correct classification while the targets T11 ('damsel') and T22 ('submarine') have the lowest probability score of about 0.77 correct classification.

5.5.1 System performance under noise conditions

Studies have been carried out for validating the classification capability of the proposed system with HOS feature set, under Gaussian as well as ambient noise conditions.

Table 5.3: Success Rates of Individual Targets for the SVM Classifier

Targets	**No. records used for validation**	**Success Rate (%)**	
		Fisher	**JMI**
pinnipeds1	326	87.46	87.19
torpedo	325	84.88	83.91
engine	324	88.00	87.60
boat1	327	83.56	86.44
boat2	324	87.32	87.32
boat3	345	86.60	84.17
grunt	242	84.18	82.54
ship1	220	80.40	79.60
ship2	217	79.89	80.49
cavitate	500	88.00	89.11
damsel	320	75.21	76.53
motor1	341	81.80	81.29
ice1	451	83.80	82.34
ship3	251	84.32	84.15
finwhale	296	88.00	85.30
whale1	324	81.75	82.02
toadfish	184	79.39	85.78
searobin	271	77.12	86.02
minke	336	87.87	87.87
rain	362	86.30	84.96
sklaxon	165	87.79	87.76
submarine	232	77.38	77.38
beluga	411	81.58	83.07
scad	285	85.84	86.31
hump	462	82.10	82.19
tanker1	354	85.64	83.65
snap	326	87.60	86.56
sculpin	336	88.05	88.14
croaker	369	85.85	87.64
spermwhale	500	87.82	87.65
Average Success Rate		*84.18*	*84.49*

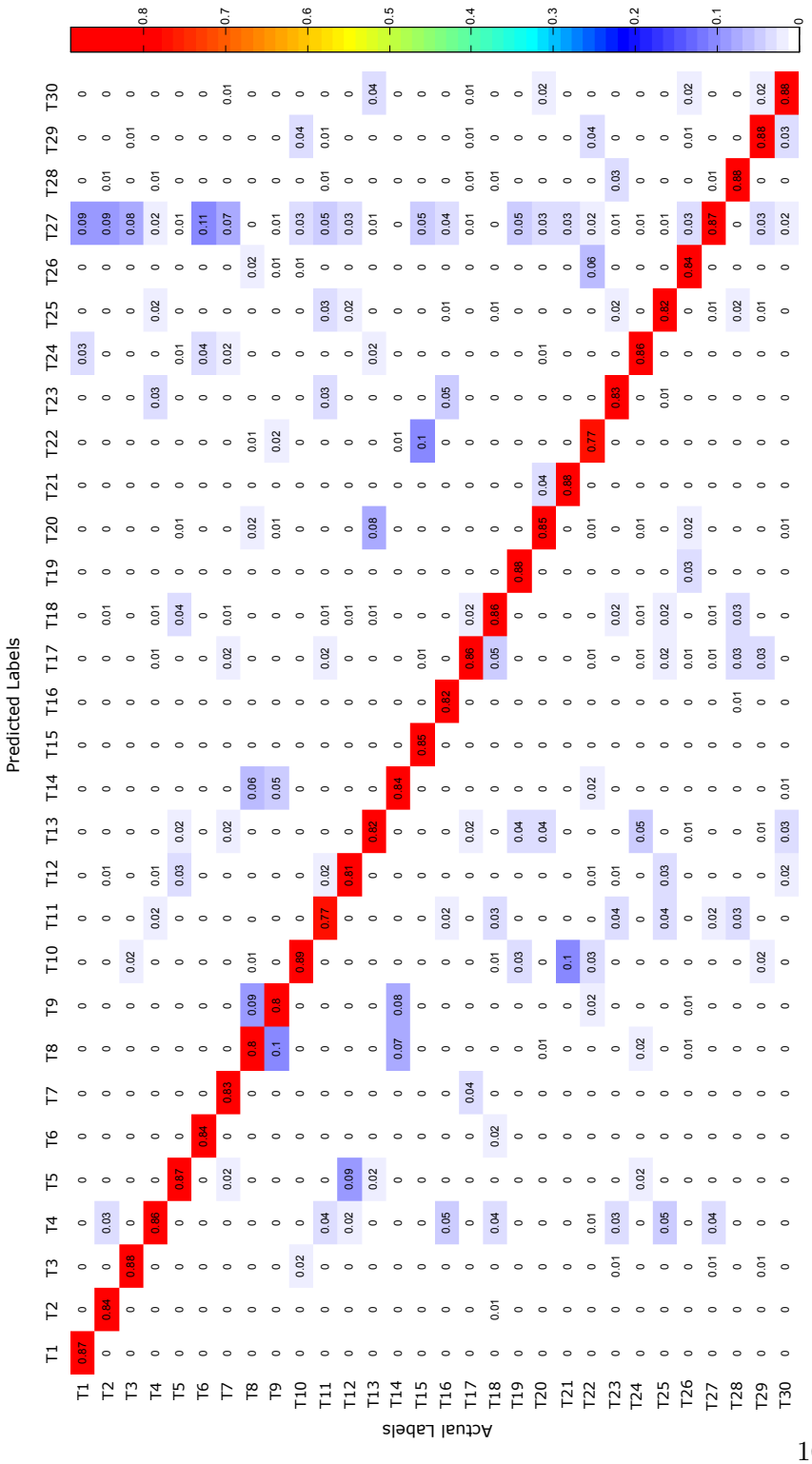


Figure 5.13: The 30 × 30 confusion matrix for the SVM classifier

5.5.1.1 Gaussian Noise

The SVM classifier trained with features extracted from normal target signals was tested with target waveforms corrupted with additive Gaussian noise for evaluation of the performance of the system. Simulated white Gaussian noise was added to the signal waveforms to simulate different SNRs and features were extracted from these signals and applied to the classifier for testing. The average success rates of the classifier for 20, 25 and 30 dB noise have been estimated and is presented in Fig. 5.14. The plot shows the average success rates for Fisher and JMI feature selection criterion, with additive Gaussian noise.

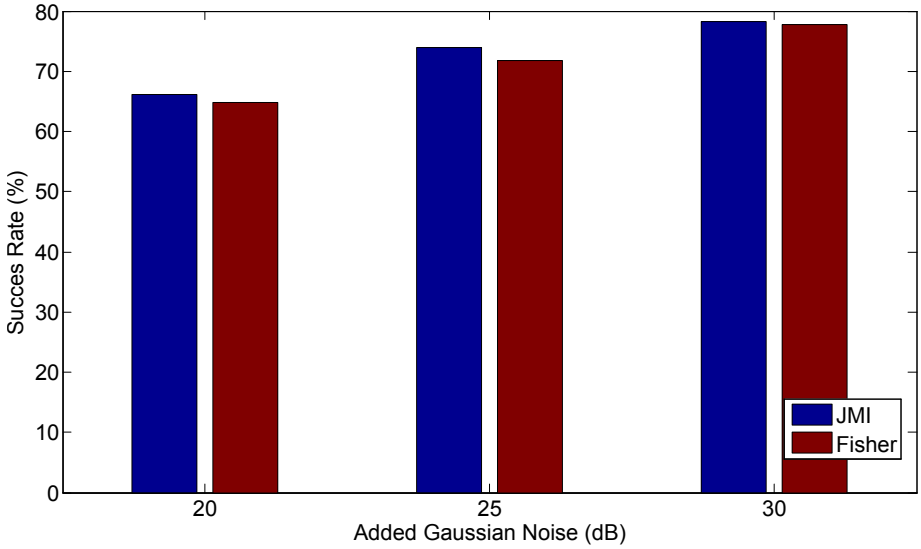


Figure 5.14: Average Success rates of the SVM classifier for the validation set with added white Gaussian noise of 20, 25 and 30 dB

It has been found that, the classifier had an average success rates of 66.21%, 74.02% and 78.34% for JMI, while the success rates with the Fisher feature selection criterion applied was found to be 64.77%, 71.81%

and 77.69%, respectively under 20, 25 and 30 dB white Gaussian noise.

5.5.1.2 Additive Ambient Noise

The performance of the classifier has also been validated under additive ambient noise conditions. A pre-recorded ambient noise data has been added to each target signal in the time domain for validating the effect of ambient noise. The added ambient noise data was recorded during the cruise #321 of the Fishery Oceanographic Research Vessel Sagar Sampada.

The SVM classifier trained with a clean feature set extracted from the normal target records has been used for the evaluation. Feature set extracted from the noise corrupted signals of each target was used in the testing phase. With the validation set corrupted by ambient noise, the classifier yielded an average success rate of 79.12% with the Fisher criterion and 80.20% with JMI.

5.5.2 Rayleigh Fading Compensation

The interaction of signals travelling along different paths in the underwater channel induces multi-path Rayleigh fading, and it affects the SNR of the signal significantly. The received signals, thus require fading compensation to ensure classification accuracy.

The probability density function of the Rayleigh distribution is defined as

$$f(x:\sigma) = \left(\frac{x}{\sigma^2}\right) exp(-\frac{x^2}{2\sigma^2}) \tag{5.1}$$

for $\sigma > 0$ and $x \in [0, \infty)$. The mean and variance of a Rayleigh random

variable can be expressed as $\mu(X) = \sigma\sqrt{(\pi/s)}$ and $var(X) = 0.5(4 - \pi)\sigma^2$.

The underwater signal before applying to the classifier, is applied to the compensation system, which describe the Rayleigh fading characteristics. Accordingly, a prototype compensation system as described in [120] has been incorporated to the proposed underwater target classifier for the estimation of Rayleigh fading effects, thereby providing the required compensation. The mathematical model that has been used for this purpose is the Clarke's Model, expressed as:

$$g(t) = \sqrt{\frac{2}{N}} \sum_{n=1}^{N} exp[j(\omega_d t \cos \alpha_n + \phi_n)] \tag{5.2}$$

where α_n represents the direction of the incoming wave, ϕ_n is the initial phase associated with the n^{th} propagation path and ω_d is the maximum angular Doppler frequency occurring when $\alpha_n = 0$. This mathematical model is assumed to provide a close approximation of the actual fading effects.

The prototype compensation system consists of a Quadrature Amplitude Modulation (QAM), transmission, demodulation and compensation stages. A pilot signal is emitted along with the target signal, and both get multiplied by the coefficients of the channel, generated by Eq. (5.2), to introduce Rayleigh fading to the signals.

The use of pilot signal helps in providing Rayleigh fading compensation. The receiver extracts the fading information of the channel concerning the pilot signal, which in turn is used in estimating fading channel coefficients. Interpolation techniques utilizing transformations can be applied for channel estimation, thereby

Table 5.4: Success Rates of SVM classifier under Rayleigh Fading

Training set	**Test set**	**Success Rate (%)**	
		Fisher	**JMI**
Without Rayleigh Fading	Without Rayleigh Fading	84.18	84.49
Without Rayleigh Fading	With Rayleigh Fading, non compensated	51.42	54.3
With Rayleigh Fading	With Rayleigh Fading, non compensated	79.45	79.7
Without Rayleigh Fading	With Rayleigh Fading, but compensated	72.12	73.56

providing Rayleigh fading channel compensation to the affected signals. Out of the various interpolation schemes available, the linear interpolation scheme as suggested in [121], has been utilized in the prototype compensation system. The success rates of the SVM classifier under Rayleigh fading scenario have been presented in Table 5.4.

With JMI feature selection, the classifier yielded a success rate of 84.49%, without considering Rayleigh fading, while it was 84.18% for the Fisher criterion. When the classifier was tested with a dataset subjected to Rayleigh fading without channel compensation, the success rate was considerably decreased to 54.3% and 51.42% of JMI and Fisher respectively. When the channel compensation was applied to the test set, a success rate of 73.56% and 72.12% have been obtained respectively with JMI and Fisher criterion, even when the classifier was trained with a clean dataset. On further analysis, it was also found that when the classifier was trained with the dataset subjected to Rayleigh fading, the success rate improved to 79.7% and 79.45%.

5.6 Summary

Prototype classifier systems for classifying the underwater noise sources using higher order spectral features extracted from the noise emissions has been described. Various classification algorithms *viz.*, the k-Nearest Neighbour, Artificial Neural Networks and Support Vector Machines have been considered for implementing the classifier. From the performance validation studies, it has been seen that all the three types of classifier systems are found to have comparable and acceptable success rates within $80\pm5\%$, using the proposed higher order spectral feature vectors with the application of Fisher and JMI feature selection criterion. The performance of the higher order feature set was also analysed with Gaussian and ambient noise as well as under Rayleigh fading conditions.

Chapter 6

Conclusions

This thesis addresses one of the emerging topics in Sonar Signal Processing, *viz.*, the implementation of a target classifier for noise sources in the ocean. The main challenge includes the diversities of the noise sources and the variabilities of the ambient noise. In the work reported in this thesis, target classifiers based on the Higher Order Spectral feature sets have been studied. A number of Higher Order Spectral features and classification schemes have been examined. It has been shown that by optimally selecting the features, all the three classifier frameworks, *viz.*, the k-NN, ANN and SVM can give acceptable levels of classification performance within 80±5%. This chapter also presents the salient highlights of the work and the inferences gathered along with the scope and directions for future research work in this area.

6.1 Salient Highlights of the Work

The classification of noise sources in the ocean has become one of the main tasks in modern underwater acoustic signal processing. The variability, diversity and abundance of such noise sources, together with the composite nature of the general ambient noise background make this task a complex one. The work reported in the thesis entitled *Implementation of an Underwater Target Classifier using Higher Order Spectral Features* addresses one of the emerging topics of digital signal processing, *viz.*, the Higher Order Spectral approach, for the analysis and classification of the noise sources in the ocean. The following are the salient highlights of the thesis.

6.1.1 Requirement for an Automated Intelligent Target Classifier

The conventional human operated target classification systems suffer from many disadvantages due to human errors and hence the need and requirement for an intelligent and automated classifier has been brought out in the thesis, considering the complex nature of the ocean environment. The introductory chapter briefly discusses the various functionalities and suggested methodology for the realization of the proposed classifier along with a discussion of various noise sources in the ocean. The use of higher order spectral analysis, especially the bispectral analysis as a feature extraction technique and its advantages have also been highlighted in this chapter.

6.1.2 Compilation of the State-of-the-art Literature

The development of a target classifier involves the extraction of the characteristic features of the noise sources, selection of the relevant features and application of the classification algorithms. As a prelude to the development of the classifier, a compilation and review of the state-of-the-art literature has been carried out. A special emphasis has been given to the higher order spectral analysis domain, which is an emerging topic in the field of digital signal processing.

6.1.3 Classifier utilizing the HOS Features

A methodology for the realization of the underwater target classifier has been proposed. HOS analysis of the signals emanating from the various noise sources in the ocean has been carried out to extract various feature sets. As the dimensionality of the resulting feature vector is high, suitable feature selection algorithms are to be applied for selecting only a minimal and non-redundant subset of the features. The optimal feature subset so obtained has been used for the classification process, using k-NN, ANN and SVM classifiers.

6.1.4 Feature Vector Generation using HOS Analysis

The robustness of the feature set is an important aspect in determining the performance of the classifier. Various Higher Order Spectral techniques have been adopted for the feature extraction. The target feature vector includes bicoherence, bicepstral and various integrated bispectral features. The procedures for the conventional

perceptually motivated feature extraction techniques like MFCC have been applied on to the bispectrum to generate the Bispectral MFCC features. A set of new features called the *Bispectral Gammatone Cepstral Coefficients (BGTCC)* have also been proposed, by applying the conventional Gammatone Cepstral Coefficient extraction procedures on to the bispectrum.

6.1.5 Implementation of Underwater Target Classifier

Once the feature vector has been extracted and the knowledge base is created, the classifier needs to be trained with the knowledge base and the parameters of the classifier need to be tuned to get an acceptable classification performance. Different classifiers *viz.*, the k-Nearest Neighbour, Artificial Neural Network and the Support Vector Machine have been studied and the classifier performances were evaluated along with the Fisher and JMI feature selection criterion. The various parameters of the classifier have been fine tuned in order to get an acceptable performance validation results. The repeatability of the results have been ascertained by using randomly selected test records and the reproducibility of the results were found to be acceptable within the limits of permissible errors.

6.2 Future Scope for Research

The work presented in this thesis has a significant role to play in view of its practical applications in underwater classification scenario. This work also provides substantial scope for further research towards improving the overall system performance, especially considering the

relatively unexplored Higher Order Spectral analysis domain. Some of the suggested areas for future work in this area are enlisted below.

6.2.1 Hardware Implementation

The proposed classifier works on a simulated environment. The algorithms can be implemented in high-end Digital Signal Processors, and a hardware version of the classifier can be developed for the use in real-time practical systems. However, development of such a system needs huge investments as it requires costly hardware subsystem components like the hydrophone arrays, receiving subsystem electronics, pre-processing modules and decision making subsystems.

6.2.2 Collection of Field Data

As obvious, the performance of the classifier depends on the vastness of the knowledge base, and in order to make the system more reliable and robust, the knowledge base needs to be updated with the feature vectors for all the classes and types of the targets under consideration. Thus, the augmentation and updation of the knowledge base with more field data, from the environment where the system is designated to operate, is an important requirement. This also enables the validation of the classifier performance with realistic field data. Also, there must be provisions for real time feature learning by which the system can learn by itself the new classes and types of the targets.

6.2.3 Multiple Targets Scenario

The current scope of the work focuses the classification of a prominent target, in the noise background. If multiple targets are present, one may have to separate the signals of each individual targets before the process of classification, using reliable source separation or unmixing algorithms. This would enable the classifier to handle multiple targets, which may be inevitable in some of the practical scenarios.

6.2.4 Augmentation of Feature Vectors

The performance of the classifier can be improved by augmenting the feature vector used, with other feature sets that can have some other characteristics not represented by the existing feature sets. A possible candidate for this from the HOS would be the trispectrum, which is based on the fourth order statistics. However, as the spectral order increases, the computational complexity and storage requirement also increases, and the development of efficient algorithms to tackle such situations is also an important area of further research. Incorporation of more complex auditory models to extract perceptually motivated features may also be worked up on.

6.3 Summary

In this chapter, an attempt has been made to bring out the salient highlights of the work carried out for the implementation of an underwater target classifier, using Higher Order Spectral features along

with the general inferences gathered. A discussion on the scope and directions for future research works in this area has also been presented.

References

[1] G. M. Wenz, "Review of Underwater Acoustics Research: Noise," *The Journal of the Acoustical Society of America*, vol. 51, no. 3B, pp. 1010–1024, 1972.

[2] T. Pieng, K. Beng, P. Venugopalan, M. Chitre, and J. Potter, "Development of a shallow water ambient noise database," in *Proceedings of the 2004 International Symposium on Underwater Technology*, pp. 169–173, IEEE, 2004.

[3] D. Ross, "Ship Sources of Ambient Noise," *IEEE Journal of Oceanic Engineering*, vol. 30, pp. 257–261, Apr. 2005.

[4] J. Potter and E. Delory, "Noise sources in the sea and the impact for those who live there," in *Proceeding of the Conference on Acoustics and Vibration Asia'98*, 1998.

[5] R. H. Nichols, "Infrasonic ocean noise sources: Wind versus waves," *The Journal of the Acoustical Society of America*, vol. 82, no. 4, pp. 1395–1402, 1987.

[6] W. Carey, "Oceanic low frequency ambient noise," in *OCEANS 2000 MTS/IEEE Conference and Exhibition. Conference Proceedings*, vol. 1, pp. 453–458, Ieee, 2000.

[7] D. Hollinberger and D. Bruder, "Ambient noise data logger buoy," *IEEE Journal of Oceanic Engineering*, vol. 15, no. 4, pp. 286–291, 1990.

[8] P. L. Brockett, M. J. Hinich, and G. Wilson, "Nonlinear and non-Gaussian ocean noise," *The Journal of the Acoustical Society of America*, vol. 82, no. 4, pp. 1386–1394, 1987.

[9] M. Bouvet and S. C. Schwartz, "Underwater noises: Statistical modeling, detection, and normalization," *The Journal of the Acoustical Society of America*, vol. 83, no. 3, pp. 1023–1033, 1988.

[10] R. Webster, "Ambient noise statistics," *IEEE Transactions on Signal Processing*, vol. 41, pp. 2249–2253, June 1993.

[11] P. Geethanjali, Y. K. Mohan, and J. Sen, "Time domain Feature extraction and classification of EEG data for Brain Computer Interface," in *2012 9th International Conference on Fuzzy Systems and Knowledge Discovery*, no. FSKD 2012, pp. 1136–1139, IEEE, may 2012.

[12] M. Jalil, F. A. Butt, and A. Malik, "Short-time energy, magnitude, zero crossing rate and autocorrelation measurement for discriminating voiced and unvoiced segments of speech signals," in *2013 The International Conference on Technological Advances in Electrical, Electronics and Computer Engineering (TAEECE)*, no. m, pp. 208–212, IEEE, may 2013.

[13] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Feature reduction and selection for EMG signal classification," *Expert Systems with Applications*, vol. 39, no. 8, pp. 7420–7431, 2012.

[14] G. Tzanetakis and P. Cook, "Musical genre classification of audio signals," *IEEE Transactions on Speech and Audio Processing*, vol. 10, pp. 293–302, jul 2002.

[15] S. Chu, S. Narayanan, and C.-C. J. Kuo, "Environmental Sound Recognition With TimeFrequency Audio Features," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 17, pp. 1142–1158, aug 2009.

[16] a. Kundu, G. Chen, and C. Persons, "Transient sonar signal classification using hidden Markov models and neural nets," *IEEE Journal of Oceanic Engineering*, vol. 19, no. 1, pp. 87–99, 1994.

[17] C. Kang, X. Zhang, A. Zhang, and H. Lin, "Underwater acoustic targets classification using welch spectrum estimation and neural networks," in *Advances in Neural Networks – ISNN 2004* (F.-L. Yin, J. Wang, and C. Guo, eds.), (Berlin, Heidelberg), pp. 930–935, Springer Berlin Heidelberg, 2004.

[18] S. Marple, "A tutorial overview of modern spectral estimation," in *International Conference on Acoustics, Speech, and Signal Processing*, pp. 2152–2157, IEEE, 1989.

[19] S. Kay and S. Marple, "Spectrum analysis–A modern perspective," *Proceedings of the IEEE*, vol. 69, no. 11, pp. 1380–1419, 1981.

[20] F. Shin and D. Kil, "Full spectrum signal processing," in *Conference Proceedings. OCEANS '95 MTS/IEEE*, vol. 1, pp. 397–403, IEEE, 1995.

[21] M. J. Hinich, "Detecting a Hidden Periodic Signal When its Period is Unknown," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 30, no. 5, pp. 747–750, 1982.

[22] M. Farrokhrooz and M. Karimi, "Marine Vessels Acoustic Radiated Noise Classification in Passive Sonar Using Probabilistic Neural Network and Spectral Features," *Intelligent Automation & Soft Computing*, vol. 17, pp. 369–383, jan 2011.

[23] Z. Fu, G. Lu, K. M. Ting, and D. Zhang, "A Survey of Audio-Based Music Classification and Annotation," *IEEE Transactions on Multimedia*, vol. 13, pp. 303–319, Apr. 2011.

[24] S. Yang and Z. Li, "Classification of ship-radiated signals via chaotic features," 2003.

[25] S. Yang, Z. Li, and X. Wang, "Ship recognition via its radiated sound: The fractal based approaches," *The Journal of the Acoustical Society of America*, vol. 112, no. 1, pp. 172–177, 2002.

[26] J. Locke and P. R. White, "The performance of methods based on the fractional Fourier transform for detecting marine mammal vocalizations.," *The Journal of the Acoustical Society of America*, vol. 130, pp. 1974–84, Oct. 2011.

[27] D. Childers, D. Skinner, and R. Kemerait, "The cepstrum: A guide to processing," *Proceedings of the IEEE*, vol. 65, no. 10, pp. 1428–1443, 1977.

[28] S. Molau, M. Pitz, R. Schluter, and H. Ney, "Computing Mel-frequency cepstral coefficients on the power spectrum," in *2001 IEEE International Conference on Acoustics, Speech, and Signal*

Processing. Proceedings (Cat. No.01CH37221), vol. 1, pp. 73–76, IEEE.

[29] R. Kemerait and D. Childers, "Signal detection and extraction by cepstrum techniques," *IEEE Transactions on Information Theory*, vol. 18, pp. 745–759, Nov. 1972.

[30] M. Kucukbayrak, O. Gunes, and N. Arica, "Underwater Acoustic Signal Recognition Methods," *Journal of Naval Science and Engineering*, vol. 5, no. 3, pp. 64–78, 2009.

[31] J. Garcia and C. Reyes Garcia, "Mel-frequency cepstrum coefficients extraction from infant cry for classification of normal and pathological cry with feed-forward neural networks," in *Proceedings of the International Joint Conference on Neural Networks, 2003.*, vol. 4, pp. 3140–3145, IEEE, 2003.

[32] M. Molla and K. Hirose, "On the effectiveness of MFCCs and their statistical distribution properties in speaker identification," in *2004 IEEE Symposium on Virtual Environments, Human-Computer Interfaces and Measurement Systems, 2004. (VCIMS).*, vol. 00, pp. 136–141, IEEE, 2004.

[33] W.-w. Hung, H.-c. Wang, and S. Member, "On the use of weighted filter bank analysis for the derivation of robust MFCCs," *IEEE Signal Processing Letters*, vol. 8, pp. 70–73, Mar. 2001.

[34] X. Zhao, Y. Shao, and D. Wang, "CASA-Based Robust Speaker Identification," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, pp. 1608–1616, July 2012.

[35] X. Valero and F. Alias, "Gammatone Cepstral Coefficients: Biologically Inspired Features for Non-Speech Audio Classification," *IEEE Transactions on Multimedia*, vol. 14, pp. 1684–1689, Dec. 2012.

[36] W. Abdulla, "Auditory based feature vectors for speech recognition systems," in *Advances in Communications and Software Technologies*, pp. 231–236, N. E.Mastorakis andV.V.Kluev,Eds. Greece:WSEAS Press, 2002.

[37] R. Schlüter, I. Bezrukov, H. Wagner, and H. Ney, "Gammatone features and feature combination for large vocabulary speech

recognition," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 4, pp. IV–649–IV–652, IEEE, 2007.

[38] J. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications," *Proceedings of the IEEE*, vol. 79, pp. 278–305, Mar. 1991.

[39] K. C. Chua, V. Chandran, U. R. Acharya, and C. M. Lim, "Application of higher order statistics/spectra in biomedical signals–a review," *Medical engineering & physics*, vol. 32, pp. 679–689, Sept. 2010.

[40] C. Nikias and M. Raghuveer, "Bispectrum estimation: A digital signal processing framework," *Proceedings of the IEEE*, vol. 75, no. 7, pp. 869–891, 1987.

[41] M. Raghuveer and C. Nikias, "Bispectrum estimation: A parametric approach," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, pp. 1213–1230, Oct. 1985.

[42] M. Hinich and H. Messer, "On the principal domain of the discrete bispectrum of a stationary signal," *IEEE Transactions on Signal Processing*, vol. 43, no. 9, pp. 2130–2134, 1995.

[43] M. Hinich and G. Wilson, "Detection of non-Gaussian signals in non-Gaussian noise using the bispectrum," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, pp. 1126–1131, July 1990.

[44] M. J. Hinich and M. Wolinsky, "Normalizing bispectra," *Journal of Statistical Planning and Inference*, vol. 130, pp. 405–411, Mar. 2005.

[45] J. Fackrell and S. McLaughlin, "Robust nonparametric bicoherence estimation by stepwise outlier rejection," *Electronics Letters*, vol. 36, no. 4, pp. 368–370, 2000.

[46] A. M. Richardson, "Bispectral analysis of underwater acoustic data," *The Journal of the Acoustical Society of America*, vol. 96, no. 2, pp. 828–837, 1994.

[47] L. Jiang, Y. Liu, X. Li, and S. Tang, "Using bispectral distribution as a feature for rotating machinery fault diagnosis," *Measurement*, vol. 44, pp. 1284–1292, Aug. 2011.

[48] J. J. G. de la Rosa, R. Piotrkowski, and J. E. Ruzzante, "Higher Order Statistics and Independent Component Analysis for Spectral Characterization of Acoustic Emission Signals in Steel Pipes," *IEEE Transactions on Instrumentation and Measurement*, vol. 56, pp. 2312–2321, Dec. 2007.

[49] R. Piotrkowski, E. Castro, and A. Gallego, "Wavelet power, entropy and bispectrum applied to AE signals for damage identification and evaluation of corroded galvanized steel," *Mechanical Systems and Signal Processing*, vol. 23, pp. 432–445, Feb. 2009.

[50] F. Argenti, P. Nesi, and G. Pantaleo, "Automatic Transcription of Polyphonic Music Based on the Constant-Q Bispectral Analysis," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, pp. 1610–1630, Aug. 2011.

[51] D. Brillinger and R. Irizarry, "An investigation of the second- and higher-order spectra of music," *Signal Processing*, vol. 65, pp. 161–179, Mar. 1998.

[52] S. McLaughlin, M. Levonen, R. Lennartsson, J. Robinson, and L. Persson, "Comparison of quadratic phase coupling detectors on sonar data," in *Oceans 2003. Celebrating the Past ... Teaming Toward the Future*, vol. 0, pp. 1685–1690 Vol.3, IEEE, 2003.

[53] X. Li, M. Yu, Y. Liu, and X. Xu, "Feature Extraction of Underwater Signals Based on Bispectrum Estimation," in *2011 7th International Conference on Wireless Communications, Networking and Mobile Computing*, no. 1, pp. 1–4, IEEE, Sept. 2011.

[54] S. Wenndt and S. Shamsunder, "Bispectrum features for robust speaker identification," in *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 2, pp. 1095–1098, IEEE Comput. Soc. Press, 1997.

[55] S. A. Taplidou and L. J. Hadjileontiadis, "Analysis of wheezes using wavelet higher order spectral features," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 7, pp. 1596–1610, 2010.

[56] G. Sundaramoorthy, M. Raghuveer, and S. Dianat, "Bispectral reconstruction of signals in noise: amplitude reconstruction issues," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, pp. 1297–1306, July 1990.

[57] A. Petropulu and C. L. Nikias, "Analytic Performance Evaluation of The Bicepstrum," *IEEE*, pp. 2337–2340, 1989.

[58] A. Petropulu and C. Nikias, "The complex cepstrum and bicepstrum: analytic performance evaluation in the presence of Gaussian noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, pp. 1246–1256, July 1990.

[59] C. Nikias and F. Liu, "Bicepstrum computation based on second- and third-order statistics with applications," in *International Conference on Acoustics, Speech, and Signal Processing*, vol. 6, pp. 2381–2385, IEEE, 1990.

[60] G. Brown, A. Pocock, M. Zhao, and M. Luján, "Conditional likelihood maximisation: a unifying framework for information theoretic feature selection," *The Journal of Machine Learning Research*, vol. 13, pp. 27–66, 2012.

[61] Y. Saeys, I. Inza, and P. Larrañaga, "A review of feature selection techniques in bioinformatics.," *Bioinformatics (Oxford, England)*, vol. 23, pp. 2507–17, Oct. 2007.

[62] L. Wang, Y. Lei, Y. Zeng, L. Tong, and B. Yan, "Principal feature analysis: a multivariate feature selection method for fMRI data.," *Computational and mathematical methods in medicine*, vol. 2013, pp. 1–7, Jan. 2013.

[63] M. R. Azimi-Sadjadi, D. Yao, Q. Huang, and G. J. Dobeck, "Underwater target classification using wavelet packets and neural networks," *IEEE Transactions on Neural Networks*, vol. 11, pp. 784–794, Jan. 2000.

[64] K. Z. Mao, "RBF neural network center selection based on Fisher ratio class separability measure.," *IEEE Transactions on Neural Networks*, vol. 13, pp. 1211–1217, Jan. 2002.

[65] G. D. Tourassi, E. D. Frederick, M. K. Markey, and C. E. Floyd, "Application of the mutual information criterion for feature

selection in computer-aided diagnosis," *Medical Physics*, vol. 28, no. 12, pp. 2394–2402, 2001.

[66] P. A. Estévez, M. Tesmer, C. A. Perez, and J. M. Zurada, "Normalized mutual information feature selection," *IEEE Transactions on Neural Networks*, vol. 20, no. 2, pp. 189–201, 2009.

[67] R. Battiti, "Using mutual information for selecting features in supervised neural net learning," *IEEE Transactions on Neural Networks*, vol. 5, pp. 537–550, Jan. 1994.

[68] I. Guyon and A. Elisseeff, "An Introduction to Variable and Feature Selection 1 Introduction," *Journal of Machine Learning Research*, vol. 3, pp. 1157–1182, 2003.

[69] A. Jain and D. Zongker, "Feature selection: evaluation, application, and small sample performance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 2, pp. 153–158, 1997.

[70] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy.," *IEEE transactions on pattern analysis and machine intelligence*, vol. 27, pp. 1226–38, Aug. 2005.

[71] M. Lan, C. L. Tan, J. Su, and Y. Lu, "Supervised and traditional term weighting methods for automatic text categorization.," *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, pp. 721–35, Apr. 2009.

[72] M. A. Wajeed and T. Adilakshmi, "Different similarity measures for text classification using KNN," *2011 2nd International Conference on Computer and Communication Technology, ICCCT-2011*, pp. 41–45, 2011.

[73] K. Odajima and A. P. Pawlovsky, "A detailed description of the use of the kNN method for breast cancer diagnosis," *Proceedings - 2014 7th International Conference on BioMedical Engineering and Informatics, BMEI 2014*, vol. 1, no. Bmei, pp. 688–692, 2014.

[74] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification and Scene Analysis (2nd ed)*. John Wiley & Sons, Inc., 1995.

[75] J. George, L. Mary, and K. S. Riyas, "Vehicle Detection and Classification From Acoustic Signal Using ANN and KNN," *International Conference on Control Communication and Computing (ICCC)*, no. Iccc, pp. 436–439, 2013.

[76] R. M. Isa, I. Pasya, M. N. Taib, A. H. Jahidin, W. R. W. Omar, N. Fuad, and H. Norhazman, "EEG brainwave behaviour due to RF Exposure using kNN classification," *Proceedings - 2013 IEEE 3rd International Conference on System Engineering and Technology, ICSET 2013*, pp. 385–388, 2013.

[77] A. Jain and K. Mohiuddin, "Artificial Neural Networks: A Tutorial," *IEEE Computer*, vol. 29, pp. 31–44, Mar. 1996.

[78] X. Yao, "Evolving artificial neural networks," *Proceedings of the IEEE*, vol. 87, no. 9, pp. 1423–1447, 1999.

[79] M. Farrokhrooz and M. Karimi, "Ship noise classification using Probabilistic Neural Network and AR model coefficients," *Europe Oceans 2005*, vol. 2, pp. 1107–1110, 2005.

[80] S. Akhtar, M. Elshafei-Ahmed, and M. S. Ahmed, "Detection of helicopters using neural nets," *IEEE Transactions on Instrumentation and Measurement*, vol. 50, pp. 749–756, June 2001.

[81] G. L. Foresti, S. Member, and S. Gentili, "A Hierarchical Classification System for Object Recognition in Underwater Environments," *IEEE Journal of Oceanic Engineering*, vol. 27, no. 1, pp. 66–78, 2002.

[82] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, pp. 273–297, Sept. 1995.

[83] W. S. Noble, "What is a support vector machine," *Nature Biotechnology*, vol. 24, no. 12, pp. 1565–1567, 2006.

[84] C. J. C. Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," *Data Mining Knowledge Discovery*, vol. 2, pp. 121–167, 1998.

[85] I. El-Naqa, Y. Yang, M. N. Wernick, N. P. Galatsanos, and R. M. Nishikawa, "A support vector machine approach for detection

of microcalcifications.," *IEEE transactions on medical imaging*, vol. 21, pp. 1552–63, Dec. 2002.

[86] E. Wold, T. Blum, D. Keislar, and J. Wheaten, "Content-based classification, search, and retrieval of audio," *IEEE Multimedia*, vol. 3, no. 3, pp. 27–36, 1996.

[87] A. Ben-Hur and J. Weston, "A user's guide to support vector machines.," *Methods in molecular biology (Clifton, N.J.)*, vol. 609, pp. 223–39, Jan. 2010.

[88] O. Chapelle, V. Vapnik, O. Bousquet, and S. Mukherjee, "Choosing multiple parameters for support vector machines," *Machine learning*, pp. 131–159, 2002.

[89] J. Li, C. Zhang, and Z. Li, "Battlefield Target Identification Based on Improved Grid-Search SVM Classifier," in *International Conference on Computational Intelligence and Software Engineering*, pp. 1–4, Ieee, Dec. 2009.

[90] D. Li, M. R. Azimi-Sadjadi, and M. Robinson, "Comparison of different classification algorithms for underwater target discrimination.," *IEEE Transactions on Neural Networks*, vol. 15, pp. 189–194, Jan. 2004.

[91] S. Tong and D. Koller, "Support Vector Machine Active Learning with Applications to Text Classification," *Journal of Machine Learning Research*, pp. 45–66, 2001.

[92] L. Zhang, F. Lin, and B. Zhang, "Support vector machine learning for image retrieval," in *Proceedings 2001 International Conference on Image Processing*, vol. 2, pp. 721–724, IEEE, 2001.

[93] M. Agarwal and S. Goel, "Expert system and it's requirement engineering process," in *International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2014)*, pp. 1–4, IEEE, May 2014.

[94] J. Lourens and M. Coetzer, "Detection of mechanical ship features from underwater acoustic sound," in *ICASSP '87. IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 12, pp. 1700–1703, 1987.

[95] D. Brutzman, M. Compton, and Y. Kanayama, "Autonomous Sonar Classification Using Expert Systems," in *OCEANS 92 Proceedings on Mastering the Oceans Through Technology*, vol. 2, pp. 554–559, IEEE, 1992.

[96] L. R. Rabiner, "Tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.

[97] S. Paris and C. Jauffret, "Frequency line tracking using HMM-based schemes," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, pp. 439–449, Apr. 2003.

[98] A. Kundu and G. Chen, "An integrated hybrid neural network and hidden Markov model classifier for sonar signals," *IEEE Transactions on Signal Processing*, vol. 45, no. 10, pp. 2566–2570, 1997.

[99] L. I. Kuncheva, "How good are fuzzy If-Then classifiers?," *IEEE transactions on systems, man, and cybernetics. Part B, Cybernetics : a publication of the IEEE Systems, Man, and Cybernetics Society*, vol. 30, pp. 501–509, Jan. 2000.

[100] T.-T. Leung and P. White, "A fuzzy logic based underwater acoustic transient classifier," in *1996 IEEE Digital Signal Processing Workshop Proceedings*, pp. 494–497, IEEE, 1996.

[101] S. K. Pal and S. Mitra, "Multilayer perceptron, fuzzy sets, and classification.," *IEEE Transactions on Neural Networks*, vol. 3, pp. 683–97, Jan. 1992.

[102] V. Chandran and S. Elgar, "Pattern Recognition Using Invariants Defined From Higher Order Spectra - One Dimensional Inputs," *IEEE Transactions on Signal Processing*, vol. 41, pp. 205–212, Jan. 1993.

[103] M. J. Hinich, "Bispectrum of ship-radiated noise," *The Journal of the Acoustical Society of America*, vol. 85, no. 4, pp. 1512–1557, 1989.

[104] X. Zhang, Y. Shi, and Z. Bao, "A new feature vector using selected bispectra for signal classification with application in radar target recognition," *IEEE Transactions on Signal Processing*, vol. 49, no. 9, pp. 1875–1885, 2001.

[105] J. Tugnait, "Detection of non-Gaussian signals using integrated polyspectrum," *Signal Processing, IEEE Transactions on*, vol. 42, no. 11, pp. 3137–3149, 1994.

[106] J. Ramirez, J. M. Gorriz, and J. C. Segura, "Statistical voice activity detection based on integrated bispectrum likelihood ratio tests for robust speech recognition," *The Journal of the Acoustical Society of America*, vol. 121, no. 5, p. 2946, 2007.

[107] X. Liao and Z. Bao, "Circularly integrated bispectra: Novel shift invariant features for high-resolution radar target recognition," *Electronics Letters*, vol. 34, no. 19, pp. 1879–1880, 1998.

[108] S. W. Smith, *The scientist and engineer's guide to digital signal processing*. California Technical Publishing, San Diego, CA, USA, Feb. 1997.

[109] A. Iturrospe, D. Dornfeld, V. Atxa, and J. Manuel Abete, "Bicepstrum based blind identification of the acoustic emission (AE) signal in precision turning," *Mechanical Systems and Signal Processing*, vol. 19, pp. 447–466, May 2005.

[110] P. T. Bhatti and J. H. McClellan, "A Cochlear Implant Signal Processing Lab: Exploration of a Problem-Based Learning Exercise," *IEEE Transactions on Education*, vol. 54, pp. 628–636, Nov. 2011.

[111] E. R. Thomas, *Sociophonetics: An Introduction*. Palgrave Macmillan, Houndmills, UK, 2011.

[112] H. Traunmuller, "Analytical expressions for the tonotopic sensory scale," *The Journal of the Acoustical Society of America*, vol. 88, no. 1, p. 97, 1990.

[113] J. Smith and J. Abel, "Bark and ERB bilinear transforms," *IEEE Transactions on Speech and Audio Processing*, vol. 7, no. 6, pp. 697–708, 1999.

[114] D. L. Wang and G. J. Brown, "Separation of speech from interfering sounds based on oscillatory correlation.," *IEEE Transactions on Neural Networks*, vol. 10, pp. 684–97, Jan. 1999.

[115] R. Patterson, K. Robinson, J. Holdsworth, D. McKeown, C. Zhang, and M. Allerhand, "Complex sounds and auditory

images," in *Auditory physiology and perception, Proc. 9th International Symposium on Hearing* (Y. Cazals, L. Demany, and K. Horner, eds.), no. 1992, pp. 429– 446, 1992.

[116] A. B. Nielsen, S. Sigurdsson, L. K. Hansen, and J. Arenas-García, "On the relevance of spectral features for instrument classification," in *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, vol. 2, pp. II–485–II–488, IEEE, 2007.

[117] P. K. Ajmera, N. S. Nehe, D. V. Jadhav, and R. S. Holambe, "Robust feature extraction from spectrum estimated using bispectrum for speaker recognition," *International Journal of Speech Technology*, vol. 15, pp. 433–440, June 2012.

[118] V. Mitra, H. Franco, M. Graciarena, and D. Vergyri, "Medium-duration modulation cepstral feature for robust speech recognition," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1749–1753, IEEE, May 2014.

[119] C. Nikias and J. Mendel, "Signal processing with higher-order spectra," *IEEE Signal Processing Magazine*, vol. 10, pp. 10–37, July 1993.

[120] T. Binesh, M. H. Supriya, and P. R. Saseendran Pillai, "An Efficient HMM Underwater Signal Classifier with Enhanced Fading Channel Performance," *Journal of Circuits, Systems and Computers*, vol. 23, p. 1450121, oct 2014.

[121] E. Okamoto, Huan-Bang Li, and T. Ikegami, "Rayleigh fading compensation for QAM by using FFT," in *Proceedings of PIMRC '96 - 7th International Symposium on Personal, Indoor, and Mobile Communications*, vol. 3, pp. 1079–1082, IEEE.

Publications brought out in the field of research

1. Mohankumar K., Supriya M.H., and P.R.Saseendran Pillai, "Extraction of Bispectral features using Principal Component Analysis," in *Proc. SYMPOL-2007 (International Symposium on Ocean Electronics)*, Allied Publishers Pvt. Ltd., Chennai, pp. 42 - 47, 2007.
2. Mohankumar K., Supriya M.H., and P.R. Saseendran Pillai, "Implementation of a Bispectral Target Classifier with Neural Networks," Presented in NSA (National Symposium on Acoustics), 2009.
3. Mohankumar K., Supriya M. H., and P. R. Saseendran Pillai "Implementation of a Neural Network Classifier for Noise Sources in the Ocean," Proc. SYMPOL-2009/ IEEE Xplore (DOI: 10.1109/SYMPOL.2009.5664142), pp. 72 - 78.
4. Mohankumar K., Supriya M. H., and P. R. Saseendran Pillai, "Implementation of a Neural Network Based Bicepstral Classifier For Marine Noise Sources," Proc. SYMPOL-2011/ IEEE Xplore (DOI: 10.1109/ SYMPOL.2011. 6170514), pp. 159 - 165.
5. Mohankumar K., Supriya M.H., and P.R. Saseendran Pillai, "Implementation of a Bicepstral Target Classifier with Support Vector Machines for Noise Sources in the Ocean," J. Pure Appl. Ultrason., 36(1), pp. 3 - 7, 2014.
   (This paper has secured the Dr. S. Parthasarathy Memorial Award for the Best Paper Published in the Journal of Pure and Applied Ultrasonics in the year 2014)
6. Binesh T, Mohankumar K., Supriya M. H., and P. R. Saseendran Pillai, "Underwater Target Classifier using A Modified Transform Based Feature Set," Int. J. Electron Commun Eng. (IJECE), 2(3), pp. 21 - 32, 2013.
7. Mohankumar K., Supriya M.H., and P.R. Saseendran Pillai, "Implementation of an Intelligent Target Classifier with Bicoherence Feature Set," Int. J. of Innov. Res. Sci Eng Technol., 3(11), pp. 17551 - 17555, 2014.
8. Mohankumar K., Supriya M.H, P.R. Saseendran Pillai, "Higher Order Feature Set For Underwater Noise Classification," Sig. Process.: Int. J. (SPIJ), 8 (5), pp. 88 - 97, 2014.
9. Mohankumar K., Supriya M.H., and P.R. Saseendran Pillai,

"Bispectrally Extractable Features for Underwater Target Classification," Presented in WOSC 2015 (World Ocean Science Congress 2015).

10. Mohankumar K., Supriya M.H., and P.R. Saseendran Pillai, "Bispectral Gammatone Cepstral Coefficient Based Neural Network Classifier," in 2015 IEEE Underwater Technology (UT), pp. 1-5, IEEE, Feb. 2015.
    (This paper was adjudicated as the third Best Student Poster in the IEEE Underwater Technology 2015 (UT-15))

Other Publications

1. Supriya M.H., Mohankumar K., and P.R. Saseendran Pillai, "Bispectral Estimation for Detecting Phase Coupled Noises in the Ocean," J. Acoust. Soc. In, Vol 33, pp. 551 - 556, 2005.
2. Mohankumar K., Supriya M.H., and P.R. Saseendran Pillai, "Detection of Nonlinear Noise Sources in the Ocean Using Bicoherence," Proc. SYMPOL-2005 (National Symposium on Ocean Electronics), Allied Publishers Pvt. Ltd., New Delhi, pp. 207 - 213, 2005.

Subject Index

T