CHAPTER 2

REVIEW OF PAST WORK

This chapter is devoted to the review of the research work reported in open literature in the areas of underwater noise, spectral estimation, cepstral analysis, bispectral estimation, target classification, etc. Classification of targets centred around the statistical classifiers, expert system classifiers, neural network classifiers, etc. have been reported and the results of comparison of the different methods have also been consolidated by various researchers. The target classification is also achieved through energy detection as well as spectral modelling. The functional and operational requirements of the classification systems in use, along with the various target specific features as well as the feature selection criteria required to be adopted for realizing the various state-of-the-art classifiers such as the statistical classifiers, expert system classifiers, neural network classifiers, fuzzy classifiers, sonar signal processor based classifiers, etc. are highlighted in this chapter. This chapter also covers the recent trends in the classifier implementation based on the Hidden Markov Model.

2.1 Introduction

In modern Sonar systems, the dry-end comprising of the receiver and the post processing modules performs underwater target detection,
Chapter 2 Review of past work

estimation, localisation, tracking and classification functions. The detection and estimation procedures in sonar involve the computation of various statistics for improving the overall performance of target detection, localisation and classification capabilities of the end system, taking into consideration all the undesirable propagation effects.

Upon judiciously selecting the characteristic features of a target, the relevant features are combined to form a feature vector. In general, such a feature vector forms the input to the classifier. Reports on classifiers based on statistical as well as expert system concepts are available in open literature. Though the results of such activities are reported, most of the papers do not touch upon the implementation details and other related issues to aid the researchers for the practical realisation of such a system. In the area of sonar signal processor based classifiers, many research works have been reported, both in the field of active as well as passive sonars. Some studies even indicate the potential use of broadband sonar as a tool for species and size classification of fish and other marine species.

Yet another major contribution in the field of classification is through the use of neural networks. There are literatures which give an overview of the practical and potential application of neural network models, viz. Hopefield Net, Multilayer Perceptron and Self Organizing Feature maps. Classification schemes implemented using fuzzy logic principles as well as algorithms typically used in pattern recognition systems have also been reported. Of these, the most widely used classifiers are the Neural Network (Multi-layer Perceptron), K-Nearest Neighbours, Gaussian Mixture Model and the Bayes' Classifier.
2.2 Characteristic Signatures of Typical Ocean Noise

Bertilone and Killeen [1], investigated the noise from snapping shrimp, which often dominates the ambient underwater noise environment of warm, shallow waters, at frequencies ranging from 1 kHz to 200 kHz. It is found that the conventional sonars perform poorly in this highly impulsive noise environment, and there is a potential for significant enhancements to detection performance using detectors that are tuned to the non-Gaussian noise. The paper reports the general statistics of band pass snapping shrimp noise data collected from the Timor Sea and investigates the performance of several generalized energy detectors (GEDs) for passive band pass detection of characteristic random processes in the noise data.

A review on the various aspects of radiated noise, self noise and ambient noise has been carried out by Gordon M. Wenz [2]. The review addresses the objectives, accomplishments and basic challenges of underwater noise research. The review also highlights the major problems such as the noise measurements, noise reduction and prevention.

Carey [3] discusses the low-to-mid frequency (LMF) noise characteristics, mechanisms and computations of basin noise based on the breaking wave as a random source of sound. The ambient noise is seen to be attributed as due to the bubble spray and splash produced by wind action and breaking waves at higher frequencies. However, the cause of the low-to-mid-frequency noise (10-500 Hz) remained a puzzle, since large bubbles required to produce low frequency noise are not found at sufficient depths. The LMF ambient noise measurements were often dominated by emissions from ships and industrial activity, making observations of local noise generating mechanisms difficult. Wave-wave interaction, wave-turbulence interaction and the pressure fluctuations due to the turbulent boundary layer...
Chapter 2 Review of past work

above the sea surface were all examined as LMF mechanisms but were found to radiate insufficient sound levels. Breaking waves were found to generate microbubble plumes and clouds, while radiating low frequency sound. Recently, cumulative oscillations of these compact microbubble clouds were shown to be responsible for generating LMF noise. Results reported by Williams et al. in [4] seem to confirm the theory that the Knudsen region of the ambient noise is produced by breaking waves even at low sea states. It also shows that the main mechanism which produces the noise is the free oscillations of the bubbles. It is concluded that the main source of energy is the radial flow around the proto-bubble at the moment it breaks away to form the actual bubble.

Tan Soo Pieng, et al. [5] describe the collection of ambient noise data and the structured compilation of the collected information into a useful database. The data collected spans a frequency range of 11 Hz to 8300 Hz. The data has been indexed and stored in a database and accessed through a GUI. A brief summary of the data collected is presented in terms of the power spectral density variations. The observed data is compared with the ‘classic’ curves reported in open literature.

Potter and Delory [6] studied the ambient noise levels from shipping and other human activities in Northern and Southern Hemisphere sites and concluded that shipping appears to have raised the background noise significantly throughout the Northern Hemisphere. Available evidence and moderate extrapolation of known features of marine mammal hearing leads to the conclusion that the total noise levels are likely to adversely affect several species.

Pflug, et al. [7], investigated the stationarity and Gaussianity of ambient shipping noise. To identify the time periods of non-stationarity in
the noise, up to fourth moments are analyzed and the summing up of the investigations indicate that the third order moments deviate from Gaussianity more than the fourth order moments. It has been found that, while shipping noise at the deeper waters appear to be somewhat non-Gaussian during certain time periods, the shallow depth data appears Gaussian.

A theoretical model for the prediction of ambient noise levels due to cumulative oscillations of air bubbles under breaking wind waves has been presented by Pavlo Tkalich and Eng Soon Chan [8]. The model uses a budget of the energy flux from the breaking waves to quantify the acoustic power radiation by a bubble cloud and derives good estimates of the magnitude, slope, and frequency range of the noise spectra using the wind speed or height of the breaking waves. In this model, it has been assumed that the wind is the source of the Knudsen spectra only through the mediation of breaking waves, which are themselves the sources of the sound radiating bubbles.

A recent work by Wales et al. [9] present an evaluation of the classical model for determining an ensemble of the broadband source spectra of the sound generated by individual ships and propose an alternate model to overcome the deficiencies in the classical model proposed by Ross. The alternate model proposed here represents the individual ship spectra by a modified rational spectrum where the poles and zeros are restricted to the real axis and the exponents of the terms are not restricted to integer values. An evaluation of this model on the source spectra ensemble indicates that the rms errors are significantly less than those obtained with the model where the frequency dependence is represented by a single baseline spectrum. Furthermore, at high frequencies (400 to 1200 Hz), a single-term rational spectrum model is sufficient to describe the frequency
Chapter 2 Review of past work

dependence and, at the low frequencies (30 to 400 Hz), there is only a modest reduction in the rms error for a higher order model. Finally, a joint probability density on the two parameters of the single term model based on the measured histograms of these parameters is proposed. This probability density provides a mechanism for generating an ensemble of ship spectra.

A model has been proposed by Gray and Greeley [10] for the acoustic source strength of blade rate line tonal produced by merchant vessels. These source strengths are based on observed cavitation time histories of merchant vessels and limitations imposed by considerations of propeller design procedures and ship vibration criteria. Relationships are presented for the expected value of the blade rate source strength for ships of different lengths, expressed both as a monopole source strength located at a known depth below a free surface and as a dipole source strength that describes the pressure radiated to the farfield. These relationships are based on a small sample of merchant ship characteristics and are exercised for the estimated population of ships at sea. This calculation yields a statistical description of the distribution of source level and frequency of propeller blade rate acoustic energy for the fleet of single-screw merchant vessels.

Arveson and Vendittis [11] present the results of extensive measurements made on the radiated noise of a bulk cargo ship, powered by a direct-drive low-speed diesel engine, which is a representative design for many modern merchant ships. The radiated noise data show high-level tonal frequencies from the ship's service diesel generator, main engine firing rate, and blade rate harmonics due to propeller cavitation. Radiated noise directionality measurements indicate that the radiation is generally dipole in form at lower frequencies, as expected. There are some departures from this pattern that may indicate hull interactions. Blade rate source level agrees reasonably well with a model of fundamental blade rate radiation
previously reported by Gray et al., but agreement in blade rate harmonics is not as good.

Observations obtained from a buoy moored near Alaska have been presented by Hollinberger et al. in [12]. The buoy recorded the omni directional ambient noise sound pressure level and wind speed. The analysis shows that for a wind speed of about 5 knot, the measured ambient noise level at 900 Hz lies well below the Knudsen curve for open ocean wind generated noise. As the wind speed increases from 5-10 kn, the measured noise level approaches Knudsen curve, and above 10 kn, the measured ambient noise level matches the Knudsen curve.

Michel Bouve et al. [13] present a study of statistical modelling of underwater noise using a Gaussian- Gaussian Mixture. Three underwater noise samples are studied with emphasis on noise PDF modeling. The results show that the snapping shrimp noise appears to be non-stationary. The background noise is very close to Gaussian while the merchant ship noise seems to be adequately described by a Gaussian-Gaussian Mixture.

The results of an experiment carried out to investigate the relative importance of wind and waves as noise generators are given by Nichols in [14]. Trials were carried out for 40 days to measure wind speed, wave height and noise spectrum levels at three deep water sites. The results suggest that breaking waves are likely to be a source of VLF ambient noise.

The non-Gaussian characteristics of ambient noise have been examined by Webster in [15]. Signal processing algorithms optimized for Gaussian noise may degrade significantly in a non-Gaussian noise environment. A generic distribution suitable for modelling non-Gaussian ambient noise has also been suggested.
Chapter 2 Review of past work

Huynh [16] demonstrates underwater mammal sound classification using a novel application of wavelet time–frequency decomposition and feature extraction using a Bienenstock, Cooper, and Munro (BCM) unsupervised network. Different feature extraction methods and different wavelet representations have been discussed. The system achieves outstanding classification performance even when tested with mammal sounds recorded at very different locations, from those used for training. The results suggest that nonlinear feature extraction from wavelet representations outperforms different linear choices of basis functions.

Certain fish sound recordings and marine mammal vocalizations are available in the archives of certain web sites [17], [18]. Information on some marine life forms like whales is given in [19]. This site also includes few sound files and spectrograms of various sound sources.

2.3 Classes of features

The identification and selection of features play a crucial role in the realisation of the classifier with acceptable success rates. A wide range of features extracted from spectral, cepstral, and bispectral methods have been used for implementing various types of classifiers such as statistical classifiers, expert system classifiers, etc..

2.3.1 Spectral Features

The methods and procedures that have been suggested for extracting some of the vital spectrally decomposable features are reported in [20] – [40] and are briefly discussed below.

In [20], Chun Ru Wan et al. analysed the statistical property of the power spectrum observations and developed a novel tonal detector by optimally integrating the spectral inferences. The optimal detectors are
derived by using the method of maximum likelihood hypothesis test. The results from simulations and open ocean trial data have shown that the proposed detectors have a promising role in detecting tonals.

Marple, in [21] presents a summary of several modern spectral estimation methods. Most of the methods are explained in the context of parametric time series modelling. Non parametric techniques discussed include classical spectral estimation, autoregressive, ARMA, Prony, Maximum likelihood, Pisarenko and MUSIC. The paper also throws some light on current spectral estimation research trends.

Shin, F.B. [22], suggests methods 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 analyzed and the results are presented. A classify before detect algorithm is used, which takes advantage of the microstructures present in the aquatics signature for improved performance.

A detailed tutorial on power spectral estimation, periodograms, random signals, fundamental principles of estimation theory, various procedures for power spectral density estimations, etc. are discussed in [23].

In [24], Hinich proposes a method for detecting an unknown periodic signal in additive noise. The period is unknown, but the amplitudes of the fundamental and the first \((M - 1)\) harmonics are known to be nonzero. One application of such a method is the detection of a torpedo by a submarine sonar system from the observed acoustic line spectrum generated by the torpedo's blade motion.
Chapter 2 Review of past work

In the work reported by Cremona et al. [25], an approach of Frequency Modulated Continuous Wave (FMCW) radar spectrum analysis with the Auto Regressive parametric estimation is described. The proposed method yields additional information compared to power spectra and a more detailed signal characterization.

A method for estimating signal harmonics in the spectrum is presented by Eftestol in [26]. The method's potential for discriminating between cardiac rhythms organised to different degrees is studied using features based on the signal harmonic frequencies and corresponding amplitudes in a classification system. This study demonstrates that the proposed method for estimating signal harmonics in the spectrum has potential for discriminating between ECGs with different levels of organization.

Ricardo S. Zebulum et al.[27] in their work investigate the application of Artificial Neural Network in speech recognition. The performance of a neural network based recognition system when using different spectral analysis models has been compared and different sets of coefficients, such as Autocorrelation and Mel cepstrum, have been extracted. A hybrid system, combining the two different sets of coefficients outperforming the other models, has also been described.

A comparison of different spectrum estimation techniques applied to nonstationary signals has been carried out by Massino Aletto et al.[28]. The comparison examines applications in which spectral analysis is applied to nonstationary signals. Different spectral analysis algorithms were tested in order to compare their behaviour in detecting defined harmonic frequencies. The results showed that the chirp-Z transform outperforms
other techniques especially when a restricted frequency band has to be analyzed.

Friedlander and Porat [29] present an iterative frequency domain technique, based on minimizing the error between the smoothed sample power spectrum and a spectral model, for estimating AR-plus-noise and autoregressive moving average (ARMA) parameters. The estimation error of this proposed technique, with less computational requirements than maximum likelihood estimators, is found to be close to Cramer-Rao bound, especially for long data records.

A new ARMA model has been proposed by Talkhan et al.[30] for the power spectral density function of noisy random ergodic zero mean discrete time signals in which, the residual power not covered by the AR polynomial is represented by a limited order MA polynomial. The residual power which is still not represented by the added limited order MA polynomial has been minimized. The proposed technique which is computationally efficient has been validated and is found to consume less storage space.

Qi Tian et al. [31] made an attempt to enhance the performance of Split Spectrum Processing (SSP) in detecting multiple targets which exhibit different spectral characteristics. An iterative procedure that combines group delay moving entropy and SSP is proposed, whereby the multiple targets are identified one at a time. It has been established that the proposed group delay moving entropy technique can be used to select the optimal frequency region for SSP, when detecting multiple targets. The dominant target is subsequently eliminated using time domain windows, which improves the detection of the remaining weaker targets.
Chapter 2 Review of past work

A high-resolution spectral estimation algorithm based on the approximation to maximum likelihood criterion, has been proposed by Luzin et al. [32]. Here, the structure of estimation filter comprises of two-channels, of which the first channel corresponds to a measuring channel like one of Capon's algorithm, while the second channel is used for taking into account and compensating for the non-coherent white noise components. The proposed two beam method allows the design of power spectrum estimating devices having rather simple structure for real time implementations.

A new method for simultaneously estimating a number of power spectra has been suggested in [33]. It is required that a prior estimate of each spectrum is available and new information is obtained in the form of values of the autocorrelation function of their sum. The method is compared with minimum cross-entropy spectral analysis and some basic mathematical properties are discussed. Three numerical examples are included, two based on synthetic spectra, and one based on actual speech data.

The impact of the Fast Fourier transform on the spectrum of time series analysis is discussed in [34]. It is shown that the computationally fastest way to calculate the mean lagged products is to begin by calculating all Fourier coefficients with a Fast Fourier transform and then to fast-Fourier-retransform a sequence made up of $a_k^2 + b_k^2$ (where $a_k + ib_k$ are the complex Fourier coefficients). The paper also discusses the raw and modified Fourier periodograms, bandwidth versus stability aspects, and the aims and computational approaches to complex demodulation.

Arun [35] presents principal components algorithms for the problem of fitting an ARMA model to a given segment of a sample sequence of a
discrete-time stochastic process, and uses the model to estimate the power spectrum of the process. To reduce the effects of finite word length errors, the authors have suggested balanced state-space parameterization of the ARMA model, instead of the more popular difference equation parameterization. Model identification is formulated as a problem of selecting a partial state to approximately span an apparently large dimensional information interface between the past and the future of the process. Different criteria are used to measure the quality of the approximation, and it leads to various Singular Value Decomposition based principal components algorithms for the problem.

Existence of an exact relationship between the maximum likelihood method (MLM) and autoregressive (AR) signal modeling in multidimensional power spectral estimation has been investigated by Dowla and Lim [36]. For one-dimensional uniformly sampled autocorrelation functions, Burg has shown a relationship between the maximum entropy method and MLM spectral estimates. In this paper the authors have shown a similar relationship between the MLM and AR spectral estimates for m-Dimensional signals sampled uniformly or nonuniformly.

Peretto et al. [37] describes power spectrum analysis and a periodic signal estimation whose bandwidth is not limited by the mean sampling time. The procedure relies on the evaluation of the input signal autocorrelation function in different delayed time instants, located at either equispaced or random time instants. A recursive random sampling process in the time domain was used in order to avoid any bandwidth limitation due to the sampling strategy in the evaluation of each autocorrelation function. The signal power spectrum as well as its period can finally be estimated, if the approximate value of the fundamental frequency is known.
Chapter 2 Review of past work

Amin [38] suggests a simple method that provides an access to a set of sliding power spectra with different characteristics. The method is based on the use of cascading-form realisation of the infinite impulse response filter. The filter is employed as a time window to the data samples or their lagged products depending on whether the frequency or the power spectrum is of interest. The choice of the filter’s poles and zeros, as well as their cascading order, allows a direct access to a wide range of power-spectrum estimators, each with different trade-offs between temporal and spectral resolutions.

The acoustic spectrum of a transiting aircraft, when received by a hydrophone located beneath the sea surface, changes with time due to the Doppler Effect. The traditional method for analysing signals, whose frequency content changes with time is the short-time Fourier transform that selects only a short segment of the signal for spectral analysis at any one time. The short-time Fourier transform requires the frequency content of the signal to be stationary during the analysis window; otherwise the frequency information will be smeared by the transformation. Recently, joint time-frequency distributions, which highlight the temporal localisation of a signal’s spectral components, have been used to analyse nonstationary signals whose spectra are time dependent. In this paper Ferguson [39], applied the short-time Fourier transform and the Wigner-Ville time-frequency distribution to a time-series data from the hydrophone so that the instantaneous frequency of the propeller blade rate of a aircraft can be estimated at short time intervals during the aircraft’s transit over the hydrophone. The variation with time of the estimates of the Doppler-shifted blade rate is then compared with the corresponding temporal variation predicted using a model that assumes the sound propagation from the
airborne acoustic source to the subsurface receiver through two distinct isospeed media separated by a plane boundary.

Omologo and Svaizer [40] report on the use of cross power spectrum phase (CSP) analysis as an accurate time delay estimation (TDE) technique. It is used in a microphone array system for the location of acoustic events in noisy and reverberant environments. A corresponding coherence measure (CM) and its graphical representation are introduced to show TDE accuracy. Using a two-microphone pair array, experiments show less than 10 cm average location error in a 6 m x 6 m area.

2.3.2 Cepstral Features

Automatic genre classification of music in audio format has gained significant importance as, in addition to automatically structuring large music collections, such classification can be used as a way to evaluate features for describing musical content [41]. A comparison of the automatic results with human genre classifications on the same dataset has been done. The results show that, although there is room for improvement, genre classification is inherently subjective and therefore perfect results can not be expected from either automatic or human classification. The experiments also showed that features derived from an auditory model have similar performance with features based on mel frequency cepstral coefficients (MFCC).

Holmes et al. [42] describe the use of subword units based on allophones with an allophone-dependent model structure, to improve subword HMM (hidden Markov model) recognition performance when using vocabulary-independent training. The new system is an extension of an approach based on sub-triphone units called phonicles. The original
Chapter 2 Review of past work

system modelled major phonetic context effects, but did not take account of context effects wider than one immediately adjacent phone or the differences in duration and spectral complexity which exists between different types of phoneme. The recognition system has therefore been extended so that phoneme transcriptions are first converted to allophone transcriptions. Each allophone is then transformed to a sequence of one or more allophonicles, where different allophonicles can have different numbers of states and one allophonicle may be shared across allophones. Using a Mel cepstrum front end, isolated-word speaker-dependent recognition experiments on six application vocabularies have shown extremely good recognition performance for allophonicle models, with an average error rate of only 0.3%.

A new algorithm of extracting MFCC for speech recognition which reduces the computation power by 53% compared to the conventional algorithm and has an accuracy of 92.93% has been suggested by Han et al. [43]. By a reduction of only 1.5%, the number of logic gates required to implement the new algorithm is about half of the MFCC algorithm, which makes the new algorithm very efficient for hardware implementation.

A method for warping the frequency axis of cepstral coefficients in a way analogous to the pre-processing performed by the human ear has been described by Merwe and Preez [44]. The computation is a two-step procedure in which the bilinear transform is used to represent the LPC coefficients on a warped frequency scale. A warping constant determines the degree of transformation. This results in an ARMA representation of the filter transfer function. The second step determines recursively the cepstral coefficients corresponding to this ARMA transfer function.
Imai[45] reports a new technique for cepstral analysis-synthesis on the mel frequency scale, the log spectrum on the mel frequency scale is considered to be an effective representation of the spectral envelope of speech. This analysis-synthesis system uses the mel log spectrum approximation (MLSA) filter which was devised for the cepstral synthesis on the mel frequency scale. The filter coefficients are easily obtained through a simple linear transform from the mel cepstrum defined as the Fourier cosine coefficients of the mel log spectral envelope of speech. The MLSA filter has low coefficient sensitivity and good coefficient quantization characteristics. The spectral distortion caused by the interpolation of the filter parameters of two successive frames is small and as such, the data rate of this system is very low. The same quality speech is synthesized at 60-70% of the data rates in the conventional cepstral vocoder or the LPC vocoder.

New techniques for automatic speaker verification using telephone speech based on a set of functions of time obtained from acoustic analysis of a fixed, sentence-long utterance has been suggested by Furui [46]. Cepstral coefficients are extracted by means of LPC analysis successively throughout an utterance to form time functions. The time functions are expanded by orthogonal polynomial representations and, after a feature selection procedure, brought into time registration with stored reference functions to calculate the overall distance. This is accomplished by a new time warping method using a dynamic programming technique. A decision is made to accept or reject an identity claim, based on the overall distance. Reference functions and decision thresholds are updated for each customer. Results of the experiment indicate that verification error rate of one percent or less can be obtained even if the reference and test utterances are subjected to different transmission conditions.
Chapter 2 Review of past work

Concentrating mainly on the signal processing and physical models behind the algorithms used to classify ships by making use of the radiated noise, the physical model for cavitation is expanded to include the losses by acoustical radiation and the heat transfer from the vapor to the fluid by Lourens [47]. Out of the five algorithms developed for estimating the propeller speed, the performance of the three most promising ones are judged with respect to the ratio of the expected value to the variance of the estimator. A complete Bayes hypothesis test on second-order autoregressive power density spectrum poles are then described for determining the kind of propulsion a vessel uses. The nature of gearbox noise is described and the cepstrum is proposed as a technique to detect this kind of noise.

Xiong et al.[48] present a comparison of 6 methods for classification of sports audio. For the feature extraction, the two choices of MPEG-7 audio features and Mel-scale frequency cepstrum coefficients (MFCC) are considered, while for the classification the two choices of maximum likelihood hidden Markov models (ML-HMM) and entropic prior HMM (EP-HMM) are considered. EP-HMM, in turn, have two variations, viz. with and without trimming of the model parameters. Thus there exist 6 possible methods, each of which corresponds to a combination. The results show that all the combinations achieve classification accuracy of around 90% with the best and the second best being MPEG-7 features with EP-HMM and MFCC with ML-HMM.

Eronen [49] has compared several features with regard to recognition performance in a musical instrument recognition system. Both mel-frequency and linear prediction cepstral and delta cepstral coefficients were calculated. Linear prediction analysis was carried out both on a uniform and a warped frequency scale, and reflection coefficients were also 40
Realisation of a Target Classifier for Noise Sources in the Ocean

used as features. The data base consisted of 5286 acoustic and synthetic solo tones from 29 different Western orchestral instruments, out of which 16 instruments were included in the test set. The best performance for solo tone recognition, 35% for individual instruments and 77% for families, was obtained with a feature set consisting of two sets of mel-frequency cepstral coefficients and a subset of the other analysed features. The confusions made by the system were analysed and compared to results reported in a human perception experiment.

A vowel system identification using phonological typologies is discussed in [50]. The phonological study of vowel systems shows that a typology of the languages may be issued from the description of their vowel system. The vocalic space has been modelled as a Gaussian mixture and two algorithms, viz. LBG and LBG-Rissanen algorithms have been used to estimate it. The success rate of about 75% has been achieved.

Molau et al. [51] presents a method to derive Mel-frequency cepstral coefficients directly from the power spectrum of a speech signal. Omission of filter bank in signal analysis does not affect the word error rate and it simplifies the speech recognizers front end by merging subsequent signal analysis steps into a single one. It avoids possible interpolation and discretization problems and results in compact implementation. The frequency warping schemes like vocal tract normalization can be integrated easily without additional computational efforts.

Molau et al.[52] describes a technique called histogram normalization that aims at normalizing feature space distributions at different stages in the signal analysis front-end, namely the log-compressed filter bank vectors, cepstral coefficients, and LDA (local density approximation) transformed acoustic vectors. Best results are obtained at
Chapter 2 Review of past work

the filter bank, and in most cases there is a minor additional gain when normalization is applied sequentially at different stages. It is shown that histogram normalization performs best if applied both in training and recognition, and that smoothing the target histogram obtained on the training data is also helpful.

Tyagi et al. [53] presents new dynamic features derived from the modulation spectrum of the cepstral trajectories of the speech signal. Cepstral trajectories are projected over the basis of sines and cosines yielding the cepstral modulation frequency response of the speech signal. It has been shown that the different sines and cosines basis vectors select different modulation frequencies, whereas the frequency responses of the delta and the double delta filters are only centred over 15 Hz. Therefore, projecting cepstral trajectories over the basis of sines and cosines yield a more complementary and discriminative range of features. In this work, the cepstrum reconstructed from the lower cepstral modulation frequency components is used as the static feature. In experiments, it is shown that these new dynamic features yield a significant increase in the speech recognition performance in various noise conditions when compared directly to the standard temporal derivative features.

Garcia et al. [54] presents the development of an automatic recognition system for infant cry, with the objective to classify two types of cry, viz. normal and pathological cry from deaf babies. Acoustic characteristics obtained from the mel-frequency cepstrum technique and a feed forward neural network that was trained with several learning methods were used for this classifier and this resulted in a better scaled conjugate gradient algorithm.
Speech and music discrimination has gained much popularity in recent years for efficient coding and automatic retrieval of multimedia sources and Automated Speech Recognition (ASR). Mubarak et al.[55] present two novel features that can be concatenated with Mel frequency cepstral coefficients, viz. the Delta Cepstral Energy (DCE) and Power Spectrum Deviation (PSDev). Employing a Gaussian mixture model for classification as a back-end to the system, a significant improvement in the error rate was found using these features.

Black et al.[56] describe an algorithm for detecting and estimating pitch in acoustic audio signals using the generalized spectrum (GS). A performance evaluation of a GS-based and two classical, autocorrelation- and cepstrum-based, pitch determination algorithms have been conducted on a set of Wavetable synthesized musical signals. The experiment performs the tasks of pitch detection and estimation. Pitch estimation performance is presented in terms of gross pitch errors and mean-squared fine pitch error. The pitch detection performance is evaluated by a receiver operating characteristic analysis of the detection statistics. Results demonstrate that the GS-based estimator generally performs worse than the autocorrelation and cepstrum-based methods. However, the GS-based method performed consistently better for the detection problem, especially at low signal-to-noise ratio levels.

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 is presented in [57] by Molla and Hirose. Multi-layer neural network with back propagation learning algorithm is used here as the classification tool. The MFCCs representing the speaker characteristics of a speech segment are computed by nonlinear filter bank
Chapter 2 Review of past work

analysis and discrete cosine transform. The speaker identification efficiency and the convergence speed of the neural network are investigated for different combinations of the proposed features. The result shows that the first MFCC degrades the identification competence and statistical distribution parameters enhance the training speed of the neural network.

Automatic knowledge extraction from music signals is a key component for most music organization and music information retrieval systems. Nielsen et al. in [58], consider the problem of instrument modelling and classification from the rough audio data. Two different models on the spectral characterization of musical instruments have been considered. The first assumes a constant envelope of the spectrum (i.e., independent from the pitch), whereas the second assumes a constant relation among the amplitude of the harmonics. The first model is related to the Mel frequency cepstrum coefficients (MFCCs), while the second to harmonic representation (HR). Experiments on a large database of real instrument recordings show that the first model offers a more satisfactory characterization and therefore MFCCs should be preferred to HR for instrument modelling/classification.

Hung et al. [59], discusses the use of Weighted Filter Bank Analysis (WFBA) to increase the discriminating ability of Mel Frequency Cepstral Coefficients (MFCCs). The WFBA emphasizes the peak structure of the log filter bank energies (LFBEs) obtained from filter bank analysis while attenuating the components with lower energy in a simple, direct, and effective way. Experimental results for recognition of continuous Mandarin telephone speech indicate that the WFBA-based cepstral features are more robust than those derived by employing the standard filter bank analysis and some widely used cepstral liftering and frequency filtering schemes both in channel-distorted and noisy conditions.
2.3.3 Bispectral Features

Higher Order Spectral analysis as well as bispectrum concepts have been described in [60] – [76].

An important task in underwater passive sonar signal processing is the determination of target signatures based on the narrow-band signal content in the received signal. To achieve good classification performance it is important to be able to separate the different sources (e.g. engine, hull and drive) present in the signature, and to determine the distinct frequency coupling pattern of each of these sources. Lennartsson, et al. [60], attempt to achieve these using bispectral techniques. It was found that the harmonics that propagated through water are engine related at low speeds and drive related at high speeds. The hull vibrations are only present at very low speeds. Moreover, it is found that the normalized bispectrum measures could provide additional coupling information not visible in the standard bispectrum.

Raghuveer and Nikias in [61] provide a detailed study of parametric estimation of bispectrum. Power spectrum estimation essentially contains the same information as the autocorrelation and hence provides a complete statistical description of a process only if it is Gaussian. In cases where the process is non-Gaussian or is generated by nonlinear mechanisms, higher order spectra defined in terms of higher order moments or cumulants provide additional information which cannot be obtained from the power spectrum. This paper concentrates on the third-order spectrum or the bispectrum. The bispectrum of a third-order stationary process can be defined as the double Fourier transform of its third moment sequence. It has the important property of being identically zero for a zero-mean Gaussian process and can thus be used to measure deviations from normality.
Chapter 2 Review of past work

Raghuveer and Nikias in [62], also describes the parametric method for bispectrum estimation which is based on a non-Gaussian white noise driven autoregressive (AR) model. A detailed overview of bispectrum and parametric methods are presented. The paper proposes a parametric approach to bispectrum estimation based on AR modelling of time series. The definition and properties of a parametric bispectrum estimator in the general ARMA case are also presented.

Regazzoni, et al. in [63] compare the spectral and bispectral analysis techniques and investigate the acoustical underwater communication problem in the low frequency range, up to 1 kHz, where the shipping noise is dominant and expected to be non-Gaussian. Classical detector performance is found to degrade in the presence of non-Gaussianity. The results indicate that for detection and identification in non-Gaussian environment, HOS based approaches are capable of providing more robust and efficient results.

Another paper which tries to compare the effectiveness of classical spectral analysis as well as modern techniques is given in [64]. The objective of the paper is to ascertain whether passive sonar signals can be classified on the basis of higher order statistics of their time series and whether higher order statistics can have any additional classification information that is not present in the power spectrum. The paper describes the limitations of the conventional higher order spectra (HOS) and defines new higher order spectra called Phase Only Spectra (POS). The studies reveal that conventional HOS could provide no more information than is present in the Power spectra. Higher order analysis should use POS to extract additional information.
Realisation of a Target Classifier for Noise Sources in the Ocean

Bispectrum estimation is found useful in detecting quadratic phase coupling among sinusoids in noise, the phase measurements for non-Gaussian processes, system identification, etc. When the given data records are short, all known methods perform poorly in terms of resolution. Raghuveer and Nikias in [65] presents a method called Constrained Third Order Mean (CTOM), which can perform very well in detecting quadratic phase coupling when the data records are short. The method proposes the estimation of the parameters of an autoregressive (AR) model driven by non-Gaussian White Noise (NGWN) by setting the sample mean of the Third Order Recursion error process to zero.

Papadopoulos, et al. [66], suggest a method for transient signal reconstruction using bispectral estimation techniques. The proposed method is capable of reconstructing the transient signal in environments, where the noise is coloured Gaussian with unknown autocorrelation function. The method could out perform the conventional Prony’s method whereas the existing methods could not perform well in the presence of significant additive noises.

In [67], Garth and Bresler re-examines many statistical tests for stationary time series, formalizing the consistency requirements for the component HOS estimators. The paper also proposes a new F-test statistic. Studies are carried out on the detrimental factors in Hinich’s test and modify this test for coloured scenarios.

In [68], Grassia, et al. consider the statistical characterization of non-Gaussian noise, with a particular reference to shipping underwater noise. The bispectrum of sample data are analyzed using both parametric as well as non-parametric methods to obtain useful phase instantaneous information applicable to classification and characterization.
Chapter 2 Review of past work

Nikias and Raghueer [69] describe the bispectrum estimation in a digital signal processing framework. Definitions and properties of bispectrum is presented. Conventional Power Spectral estimation and its limitations are discussed. General reasons behind the use of bispectrum in signal processing are addressed. Both conventional and parametric models of bispectrum estimations are explained. The paper also briefs various applications of the bispectrum like detection of Quadratic phase coupling, deconvolution, etc.

Hinich, et al. [70], describe the bispectrum estimation of a ship radiated noise received using a towed array. The result shows that there exist frequency dependent bispectral components in the ships radiated noise. The ambient noise does not contain any significant bispectral components. Since the existence of a nonzero bispectrum indicates the presence of non-linear components in the noise source, it is estimated that the ship generated noise contains non-linear components. The paper suggests a means for differentiating between ship noise and at least some other forms of ambient noise sources using bispectral analysis. Another significant point to be noted is that the data used was in the narrow band (of bandwidth 130 Hz) and the cavitation noise, which is expected to be a major contributor of the shipping noise, was out of band.

The bispectrum and bicoherence estimates of underwater acoustic signals have been studied by Richardson and Hodgkiss [71]. The properties of bispectrum and bicoherence are described. Bispectra of data collected from a freely drifting swallow float is estimated. The results show how the bispectrum can be used to detect non-Gaussianity, nonlinearity and harmonic coupling. Special stress is given in determining whether the spectral lines are harmonically related.
Mendel in [72], gives an in-depth treatment on higher order statistics. The paper collects some of the most useful theoretical results, making them readily accessible. Various fields of applications of higher order statistics are described. The paper covers various definitions and properties related to higher order statistics and also discuss various results.

Quazi [73] suggests an attempt to utilize the basic quantities of the information theory, viz. entropy and mutual information for detection and localization of underwater sources. The entropy of a process having a finite number of sample points is maximum when the received process consists of noise alone and decreases when a correlated signal is present. The paper analyses both active and passive sonar signals and compares them with the results of traditional techniques.

Martin [74], presents a detection statistics, which exploits features in the three dimensional response of the non-stationary bispectrum for an assumed class of transient signals. The results are presented relative to the performance of a conventional power spectrum detector and a detection statistics based on the spectral correlation. The paper also discusses the merits of bispectral detectors relative to other transient detection methods.

Frazer and Boashash [75] have demonstrated the application of the Wigner-Ville time frequency distribution, the bispectrum, the time varying bispectrum and Gerr's third order Wigner distribution, to underwater acoustic data. Use of higher order spectral analysis improves time, frequency and time-frequency analysis methods and provides the analyst with important additional information.

Roy et al. [76] in their paper present a novel feature and its estimation method, for the classification of marine vessels using passive
Chapter 2 Review of past work

A classification feature, namely the coefficient of quadratic nonlinearity (CQNL), which quantifies the quadratic coupling that exists between a pair of running machinery on the target platform, has been proposed. The classification is based on the premise that the degree of nonlinear coupling between the various machineries on a target platform is unique for a particular class of targets and depends on the types of machinery and their placements. The estimation of the CQNL feature uses higher-order spectral analysis in conjunction with a method analogous to matched field processing. The CQNL feature provides unique information about the target platform that is not present in spectrum based features. The performance of the algorithm is good even in low SNR conditions.

2.4 Classifiers

2.4.1 Statistical Classifiers

Statistical classification uses statistical procedure for classification, in which individual items are placed into groups based on quantitative information on one or more characteristics inherently associated with the items (referred to as traits, variables, characters, etc.) and based on a training set of previously labelled items.

In [77], Rajagopal et al. describe the classification of marine vessels using Passive Sonar methods. The signals from various surrounding sources are sensed by a receiver. The data is processed and transformed to obtain the input to the classifier, which combines these with any other stored information to make a decision. The paper proposes a scheme for a general classifier and describes the practical constraints on each block of the classifier. Three well known classification techniques, viz. statistical,
expert system and neural network are compared and finally combined to give a hybrid classifier.

Shapo and Bethel [78], in their paper introduces *Cell Probability Density Function* (CPDF), a new statistical detecting and tracking algorithm suitable for imaging arrays. The input to the algorithm is the 20 array of intensity levels in all beams as a function of time. CPDF is a three-step algorithm, involving pre-processing, detection, and tracking/bearing estimation. It is found that CPDF has been very successful in detecting and tracking targets on broadband data collected by SONAR arrays, and has excelled in especially challenging scenarios with high bearing rates and multiple crossing targets.

### 2.4.2 Expert System Classifiers

Rajagopal *et al.* [79] describes an expert system approach for sonar target classification. It deals with passive listening for classification of underwater targets.

There are basically two different techniques of classification, *viz.* the statistical approach which makes use of classical pattern recognition methods and the expert system approach. The paper also identifies and discusses dominant sources of noise such as propeller cavitation noise, blade-rate tonals, piston-slap tonals, gear noise, injector noise and low frequency radiation from the hulls. The structure of the expert system consists of three major parts, *viz.* a knowledge base which deals with rules of inference and facts, a database consisting of the facts made available to the system by the programmer at any given time and the inference engine that guides the reasoning process through the knowledge base by attempting to match the facts in the data base to the rule conditions.
Chapter 2 Review of past work

Algorithms to detect and classify various parameters of ships by analyzing the underwater acoustic noise are presented in [80]. Lourens and Coetzer in this work, discuss the detection of the mechanical features like propeller shaft speed, number of propeller blades and type of propulsion. Classification of the ships into different classes is carried out using a small expert system. The propeller shaft speed and number of blades can be determined by two approaches. The first method consists of computing the Discrete Fourier Transform of the short time power of band pass filtered underwater acoustic noise. The second method determines a two dimensional discrete Fourier transform to extract the required information.

The development of an autonomous sonar classification expert system for AUVs is investigated in [81] 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.

Adnet and Martin [82] presents the first results of a study of an expert system dedicated to spectral analysis. Spectral analysis methods have been put together with a unification principle stemming from filter bank analysis and the strategy has been applied to generate a Knowledge based system.

2.4.3 Neural Network Classifiers

The various neural network based classifiers reported in open literature are described below.

Jae-Byung Jung, et al. [83] discusses the results from a set of experiments conducted on the target classification capabilities of broadband
sonar using targets of differing sizes and materials like Styrofoam balls and hollow plastic bodies. The experiments were carried out by analyzing the spectral components of echoes from the targets using neural networks. The studies indicate the potential use of broadband sonar as a tool for species and size classification of fish and other marine targets.

Several different classification algorithms are tested and benchmarked in [84], by Donghui Li, et al. not only for their performance but also to gain insight into the properties of the feature space. Results of a wideband 80-kHz acoustic backscattered data set collected for six different objects are presented in terms of the receiver operating characteristics (ROC) and robustness of the classifiers with respect to reverberation. Classification methods like Multivariate Gaussian Classifier, Evidential K-Nearest Neighbour Classifier, support vector machines, etc. are considered.

In [85], Purnell, et al. present the implementation of a classifier which discriminates between ships based on the radar back scatter from the targets. Three different classification methods, viz. correlation filters, peak extraction with a feed forward neural network classifier and a feed forward neural network using raw radar data were used. It was concluded that the feed forward neural network method that uses raw radar data showed better performance.

De Yao, et al. in [86] propose a classification system which consists of several subsystems including pre-processing, sub-band decomposition using wavelet packets, linear predictive coding, feature selection and neural network classifier. A multi-aspect fusion system is introduced to further improve the classification accuracy. The classification performance of the overall system is demonstrated and benchmarked on two different acoustic backscattered data sets with 40 and 80-kHz bandwidth. A comprehensive
Chapter 2 Review of past work

study has been presented to compare the classification performance using these data sets in terms of the receiver operating curves, error locations, generalization and robustness on a large set of noisy data. Additionally, the importance of different frequency bands for the wideband 80 kHz data is also investigated. For the wideband data, a sub-band fusion mechanism is introduced, which offers very promising result.

[87] – [90] discuss the introductory papers which throw some light on the underlying principles and methodologies of realizing neural network based classifiers [87] outlines the biological neuron structure and its similarity to the artificial neural network (ANN) alongwith a historical overview of ANNs. A brief review of the various components of the network and various training / learning functions is also presented.

Schoonees [88], gives an introduction to artificial neural networks and provides an overview of their potential application in signal processing. A brief survey of three network models, viz. Hopefield Net, Multilayer perceptron and Self Organizing Feature maps is also described.

Richard P. Lippmann [89] in his review paper presents the concepts of artificial neural networks in detail. He reviews six important neural net models that can be used for pattern classification. A comparison of neural nets and traditional classifiers are also presented. Descriptions of Hopefield, Hamming, Single Layer perceptron, Multi layer perceptron, Kohonen’s Self Organizing Feature maps, etc. are presented.

Back propagation algorithms are extensively used in almost all applications involving neural networks including pattern as well as target recognition. Paul and Byrne [90] present an efficient learning algorithm for the back propagation networks. Two conditions for reducing the learning
iterations are deduced, without affecting the memory retention or generalization capabilities of the network.

Abdel Allim, *et al.* [91] describe a neural network system which can recognize different types of sonar signals. The work compares the parameters that affect the shape of the echoes returned from different underwater targets like submarines, mines, etc., using fourteen echo signals from three different types of military targets.

A discussion of the real-time digital signal processor based hierarchical neural network classifier, capable of classifying both analog and digital modulation signals is presented in [92]. A high performance DSP processor, viz. the TMS320C6701, has been made use for implementing the different kinds of classifiers, including a hierarchical neural network classifier. A total of 31 statistical signal features are extracted and used to classify 11 modulation signals corrupted by white noise.

Martinez Madrid, *et al.* [93], describe a target classification system which uses the measured Doppler signature to excite a neural network. The paper describes the use of Multilayer perceptron based neural network and its training using back propagation algorithm. The paper also points out the advantages of using neural networks, like fault tolerance, learning capabilities, etc.

Eapen, A [94], proposes the use of a neural network for detecting underwater targets in the presence of random noise. Here a neural network is made to adapt to the signal output of a hydrophone. Then the changes triggered by the presence of targets will be detected with the complex classification space of the neural network. Neural networks offer powerful
collective computational capability for designing special purpose hardware, which can implement automatic detection of targets in real time. The ability to learn is the key property of ANNs. Modern learning procedures fall into two broad categories, viz. supervised methods, which require a teacher to specify the desired outputs and unsupervised procedures, which construct internal models that capture regularities in input signals. The work presented in this paper uses a variant of the back propagation rule, which is one of the most widely used algorithms for multilayer perceptron-like networks, called the Modified selective update back propagation algorithm.

A comparison of the relative performance of a number of classical classification methods with the neural network is performed by Patel, et al. [95]. Feature data extracted from infrared images are used for the comparison. Classical classification techniques, viz. k-Nearest neighbour, Euclidean distance, weighted Euclidean distance and Mahalanobis distance are described and tested. The neural network used was a simple Multilayer perceptron network trained with error back propagation coupled with a gradient descent algorithm. The network consisted of an input layer, a hidden layer and an output layer. A sigmoid function was used as the activation function. The studies show that the neural network and k-nearest neighbor methods could outperform all the other classical techniques. Considering the adaptability and the computational efficiency of Neural Networks, the MLP method is shown to have a distinct advantage.

Chin-Hsing Chen, et al. [96] describe the results of four kinds of neural network classifiers that have been used for the classification of underwater passive sonar signals radiated by ships. Classification process is divided into two stages, viz. pre-processing feature extraction stage and classification stage. In the pre-processing stage, Two-Pass Split- Windows (TPSW) algorithm is used to extract tonal features from the average power
Realisation of a Target Classifier for Noise Sources in the Ocean

spectral density of the input data. In the classification stage, four kinds of static neural network classifiers are used to evaluate the classification results:

- The probabilistic based classifier-Probabilistic Neural Network (PNN)
- The hyperplane based classifier-Multilayer Perceptron (MLP)
- The kernel based classifier- Adaptive Kernel Classifier (AKC) and
- The exemplar based classifier- Learning Vector Quantization (LVQ).

The data were collected from fishing boats, which were classified into three groups. From each boat, three types of signals (at low speed, at medium speed and at high speed) were recorded. It has been experimentally established that the exemplar classifiers-LVQ have the most efficient learning.

Azimi-Sadjadi, et al. [97] describe a new sub band based classification scheme developed for classifying underwater mines and mine like objects from the acoustic backscattered signals. The 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. Multi aspect fusion was performed to obtain great improvement in the classification performance of the system.

Roth, M.W in his review paper [98] on Automatic Target Recognition (ATR), highlights the use of neural network technology in the field of ATR. The paper describes ATR sensor development and Multi
sensor fusion, various issues related to ATR, feature extraction procedures and the scope of neural networks in the field of ATR and its advantages.

Overman and Louri, [99], investigates the design of a neural network architecture that can take noisy serial pixel data as an input and report the detected position of the target with respect to the sensors field of view. The neural network target detection architecture presented is based on Multilayer perceptron neural receiver. The paper describes how a neural network can be implemented as an optimum likelihood ratio receiver and discusses back propagation training process. Preparations in setting up neural net detection architecture for Monte-Carlo simulations against various noise types are also discussed.

Solinsky and Nash [100], in their paper attempt to describe the applications of neural network in sonar. Various neural network classifiers operating on the DARPA Phase I data set has been analyzed using classical decision theory. An important element of the assessment is to include a ground truth of events in the data set. A trained human operator produces such ground truths based on aural analysis of the data.

The use of hybrid neural approaches for passive sonar recognition and analysis using both unsupervised and supervised network topologies are investigated by Howell and Wood [101]. The results presented demonstrate the ability of the network to classify biological, man made, and geological noise sources. The capabilities of the networks to identify the complex vocalizations of several fish and marine mammalian species are also described. Basic structure, processor requirements, training and operational methodologies as well as application to autonomous observation are described. For training the network, Self Organising Map(SOM) - Kohonen maps are used, since it is efficient for unsupervised
learning. It also contains a review of various types of source files (.wav, MP3 etc) and the issues arising out of over sampling data in these types of files.

Hallinan and Jackway [102], describes a novel feature selection algorithm which utilizes a genetic algorithm to simultaneously optimize a feature subset and the weights for a three-layer feed-forward neural network classifier. It has been shown that this method needs only fewer input features and simpler neural network architecture. The results indicate that tailoring a neural network classifier to a specific subset of features has the potential to build a classifier with low classification error and relatively low computational overheads.

The design and evaluation of a comprehensive classifier for short duration oceanic signals obtained from passive sonar is described in [103]. The paper highlights the importance of selecting appropriate feature vectors for efficient classification. Wavelet based feature extractors are examined. A number of static neural network classifiers are evaluated and compared with traditional statistical techniques. The paper highlights the fact that each algorithm is designed to handle only a few set of problems and may have many limitations and a synergistic approach can lead to better results. It is found that a judicious combination of several classifiers will yield higher accuracy, since it can overcome the limitations of a single type of network and the system was tested with DARPA data set.

Yanning Zhang, et al. in [104], discuss a local adaptive neural network based classifier to classify ship noises. Combining wavelet theory with neural network to form adaptive wavelet neural network has the advantage of feature automatic compression, extraction and classification from signal. The neural network consists of input layer, adaptive wavelet
Chapter 2 Review of past work

feature extraction layer, hidden layer and output layer. A modified Back propagation algorithm is used to train the network.

Adams, et al. [105], present the statistical properties of underwater acoustic ambient noise fields obtained by analyzing the acoustic output of a single hydrophone in the 75-200 Hz band. The analysis demonstrates the non-stationarity of the noise power in 0.78 Hz bands. In addition, the correlation between the noise power fluctuations in 0.78 Hz bands is shown as a function of time and frequency separation of the bands. The fluctuations appear to consist of a slow broadband power increase and a smaller amplitude fluctuation process which has small correlation across frequency and time.

John R. Potter [106], establishes that a feed forward ANN is very effective in self training the system to recognize the end notes of bowhead whale songs. A three layer feed forward network is used for testing.

2.4.4 Fuzzy Classifiers

Argenti, et al. [107], address the problem of detecting ships in SAR images in a fully automated way. A classification scheme implemented using fuzzy logic principles is also discussed in this paper.

Amo et al. [108], suggests an algorithm for terrain matching, leveraging an existing fuzzy clustering algorithm, and modifying it to its supervised version for apply the algorithm for georegistration as well as pattern recognition. Georegistration is the process of adjusting one drawing or image to the geographic location of a "known good" reference drawing, image, surface or map. The terrain matching algorithm will be based on fuzzy set theory as a very accurate method to represent the imprecision of the real world, and presented as a multicriteria decision making problem.
The energy emitted and reflected by the Earth’s surface has to be recorded by relatively complex remote sensing devices that have spatial, spectral and geometrical resolution.

2.4.5 Sonar Signal Processor based Classifiers

Gaunaurd, in [109], describes the bistatic and mostly monostatic techniques that are useful for target classification by means of active sonar. The echoes returned by any submerged elastic body contain features caused by the poles of the scattering amplitude of the problem. These poles are studied and it was shown that they naturally split into two large sets from which one can separately extract shape or material composition information. The composition information seems easier to determine than details about the shape. Together these sets of poles unambiguously characterize any scatterer.

A filter structure has been proposed for target signal enhancement in reverberation limited environment by Kim et al. [110]. The proposed structure consists of an adaptive filter and a non-adaptive filter. The input signal is filtered by the non-adaptive filter whose coefficients are obtained from the adaptive filter working with the delayed signal. The investigations were carried out on the data from an active sonar system for target signal enhancement problem and the results have shown that the proposed method can yield fairly acceptable performance in a time varying channel and is robust to target cancellation effects.

Dwyer [111], discusses the processing technology that enables the classification of target echoes from very wide bandwidth transmitted signals, which can reduce the complexity of classifiers. Implementation of two types of sequential classifiers is discussed. One of the sequential
Chapter 2 Review of past work

classifier used the spectrum of the data while the other used the fourth order cumulant spectrum of the data.

A new method for target localization and classification has been discussed in [112]. Firing of pulses from two closely spaced transmitters, with a time separation of the order of 1 ms is the basic tool for this approach. Maximum Likelihood Estimation (MLE) method is used for the implementation.

A tutorial illustrating some interactive demos, tips and tricks of Digital Signal Processing are described in the website [113].

2.4.6 Recent Trends

A classification algorithm using Hidden Markov Models (HMMs) is presented in [114]. Recognition of three classes of targets such as personnel, tracked and wheeled vehicles can be carried out using this algorithm. The procedure described is based on target Doppler signatures. While conventional Doppler based methods consider the Doppler signature to be stationary, the suggested method utilizes the time-varying nature of Doppler signature as well for efficient classification. One of the advantages of this technique is that the classifier requires only a modest amount of data for training.

Lourens [115] considers an improved cavitation model and proposes algorithms for determining propeller speed and number of blades. The new model takes into account the acoustic losses and heat losses. Gear box identification is also addressed in this paper. A complete Bayes hypothesis test is also described for determining the kind of propulsion the vessel uses.
Paul Chestnut, Helen Landsman and Robert W. Floyd, [116], present a study carried out on an active sonar target recognition system. The data were obtained from 16 targets, submerged in a salt water pool. The frequency responses from the echoes were analyzed and the information is extracted by energy detection in a bank of filters and Spectral modelling. The classification techniques use methods in conventional pattern recognition.

Chan, et al. [117] present a new bearings-only method of detecting and tracking low signal-to-noise ratio (SNR) wideband targets on a constant course and velocity trajectory. A track-before-detect strategy based on matched velocity filtering is adopted using spatial images constructed from a sequence of power bearing map (PBM) estimates accumulated during tracking. To lower the threshold SNR for detection, a discrete bank of matched velocity filters integrates the PBM images over a range of hypothesized trajectories. The distribution of the matched filter output is derived based on a single point target in diffuse noise conditions. Receiver operating characteristic curves show a definite detection gain under low SNR conditions for matched velocity filtering over detection from a single PBM.

Attempts by Chen Xiangdong and Wang Zheng [118], throw light on a non-linear signal processing technique called the Similar Sequence Repeatability to analyze the ship radiated noise and indicate its use for the acoustic target recognition. Here, the local similarity of the ship radiated noise data is studied based on the temporal behaviour of the time series itself. According to the local similarity property of the time domain acoustic signal, a phase space is constructed. From this, repeatability parameters are calculated and repeatability (RPT) curves are drawn, from which the entropy information is extracted. The RPT curves and the
Chapter 2 Review of past work

entropy can effectively represent the time domain features of the ship radiated noise and aids in underwater target recognition and classification.

Guo Guirong, et al. [119] discuss the problem of ship recognition and a suitable method is proposed. The paper focuses on algorithms for target recognition from radar video returns. The method mainly uses the Fourier transform and Mellin transform for feature extraction. Fourier Transform is invariant to shifting. This property provides a useful means for extracting features insensitive to the time delay of radar returns, that is, the target range. The Mellin transform produces a set of features invariant to scaling changes. This implies that the features extracted from the MT are insensitive to the aspect angle of the radar.

Pezeshki, et al. in [120], suggest a feature extraction method for underwater target classification that exploits the linear dependence between two sonar returns. Canonical co-ordinate decomposition is applied to resolve two consecutive acoustic backscattered signals into their dominant canonical coordinates. The hypothesis behind this feature extraction method is that for certain aspect separations linear dependence of coherence between the sonar returns reveals common target/non-target attributes, whereas linear independence reveals bottom reverberation features.

A work that discusses the nonlinear regularities in ship-radiated signals is presented by Su Yang and Zhishun Li [121]. They propose a chaotic feature for classification. Certain classes that cannot be classified effectively by using spectra can be satisfactorily classified using the chaotic features. Experimental results show that this feature is effective and outperforms the spectrum feature in identifying some classes. It can augment current solutions by providing complementary information.
Fractal based approaches for the recognition of ships from the radiated noise is being considered in [122]. The methods proposed by Yang, et al. include fractal Brownian motion based analysis, fractal dimension analysis and wavelet analysis, to augment existing feature extraction methods that are based on spectral analysis. The results show that fractal approaches are effective and when used to augment two traditional features, line and average spectra, fractal approaches lead to better classification results. This implies that fractal approaches can capture some information not detected by traditional approaches alone.

A sequential decision feedback approach for target classification of underwater mine-like objects in a changing environment is described by Azimi-Sadjadi et al.[123]. An adaptive target classification system developed using the decisions of multiple aspects of an object through a tapped delay line mechanism to impact the final decision of the current aspect is also discussed here. This system minimizes the error of the classifier while it maps the new feature vector to a familiar feature space for the classifier. The test results presented are obtained on a wideband acoustic backscattered data set collected using four different objects with 1 degree of aspect separation for two different bottom (smooth and rough) conditions.

Paul Gaunard et al.[124] discuss the automatic classification of environmental noise sources from their acoustic signatures, recorded at the microphone of a noise monitoring system (NMS), using Hidden Markov Models. The performance of the proposed system, which is based on a time frequency analysis of the noise signal, was evaluated experimentally for the classification of five types of noise events, viz. car, truck, moped, aircraft and train. The HMM based approach is found to outperform human
Chapter 2 Review of past work

listeners as well as previously proposed classifiers based on the average spectrum of noise event with success rates as high as 95%.

Transformation of a segment of acoustic signal, by processing into a vectorial representation such as the spectrum, can permit the identification of the constituent phonemes within spoken speech according to Grant [125]. A comparison with stored replicas of the data segments using techniques such as dynamic time warping or hidden Markov modelling then permits a speech recognition operation to be accomplished. These signal processor intensive transform and graph-search-based pattern matching techniques are reviewed and currently achievable recognition accuracies are reported.

Existing concepts of Walsh power spectra for wide sense stationary stochastic processes are restricted to the case of auto power spectra because they are based on real Walsh functions. Blaesser [126], describe a Walsh power spectrum, which is based on a system of complex Walsh functions for wide-sense stationary stochastic processes and the concepts have been extended to auto and cross power spectra as well.

2.5 Summary

An attempt has been made in this chapter to present a state-of-the-art literature in the topic covered by the thesis highlighting the characteristic signatures of typical ocean noise as well as the classes of features that have been considered for realizing the various types of classifiers as reported in open literature. The literature survey has also brought out the operational features of various classifiers such as statistical, expert system, neural network and fuzzy k-NN classifiers.