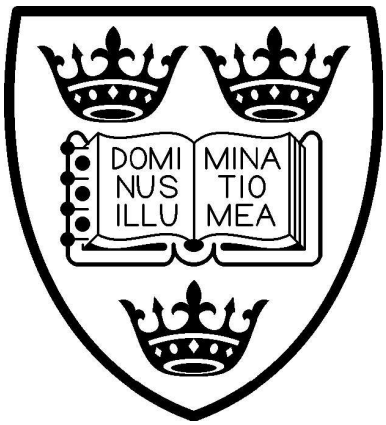


Learning from Sonar Data for the Classification of Underwater Seabeds

Louis Nicolas Atallah

Department of Engineering Science

Robotics Research Group

University of Oxford

Trinity Term 2004

Declaration

This thesis is submitted to the University of Oxford, in particular fulfillment of the requirements for the degree of Doctor of Philosophy. The thesis is entirely my own work and, except where otherwise stated, describes my own research.

Louis N. Atallah, St Cross College, Oxford

Copyright ©2004 Louis N. Atallah

All Rights Reserved

Acknowledgments

I would like to thank my supervisor Dr Penny Probert Smith for being there for every minute of this work, as well as teaching me so much about enthusiasm, dedication and perseverance.

Financially, this thesis would never have been possible without the generous support of the Karim Rida Said Foundation for three years. I would also like to thank St Cross College for offering me the Unilever Scholarship in my last year, and Dr David Kenning and Dr David Hills for offering me a lectureship in Lincoln College. Thanks also to SEA Ltd., Dr Jenny Collier (Imperial College) and Dr Nick Pace (NATO SACLANT Undersea research centre) for providing us with datasets used in this work. I'd like to thank Dr Richard Bates (St Andrews) for his great contributions to this work, in datasets and 'geophysical' points of view, and my two examiners for their valuable input.

Thanks to everyone in the robotics department who was of great help to us, especially, Prof. Michael Brady (and his comments on wavelets). Thanks also to all my colleagues in the sensors lab: David, Steve, Paul, Kostas, Russ, Emma and Ingmar for being great friends and labmates!

Thanks also to everyone in the ANC Department in Edinburgh. I have received great support from many people, including Chris Williams, David Barber and Amos Storkey as well as the lively students and postdocs.

I'd like to thank my many wonderful friends in Oxford, Edinburgh and Lebanon for being

always there, and making the last three years go so fast! (I cannot mention you all, I would need another thesis...)

This thesis is dedicated to my family (the very extended one), especially to my grandfather *Daoud* who passed away a few weeks before submitting it. During years of war and sadness, my family always insisted that education comes first, and that the success of your children is your source of happiness. To them, that meant giving with no limits...I can never repay that! The thesis is especially dedicated to my mother and father who deserve their own honorary Dphils entitled 'Parenting and self sacrifice: two very good examples'.

This one goes out to the one I love.

Learning from Sonar Data for the classification of Underwater Seabeds

Louis Nicolas Atallah

Sensors Lab-Robotics Research Group — St Cross College

Department of Engineering Science — Trinity term 2004

Abstract

The increased use of sonar surveys for both industry and leisure activities, has motivated the research for cost effective, automated processes for seabed classification. Seabed classification is essential for many fields including dredging, environmental studies, fisheries research, pipeline and cable route surveys, marine archeology and automated underwater vehicles. The advancement in both sonar technology and sonar data storage has led to large quantities of sonar data being collected per survey. The challenge, however, is to derive relevant features that can summarise these large amounts of data and provide discrimination between several seabed types present in each survey.

The main aim of this work is to classify sidescan bathymetric datasets. However, in most sidescan bathymetric surveys, only a few ground truthed areas (if any) are available. Since sidescan 'ground-truthed' areas were also provided for this work, they were used to test feature extraction, selection and classification algorithms.

Backscattering amplitude, after using bathymetyric data to correct for variations, did not provide enough discrimination between sediment classes in this work which lead to the investigation of other features. The variation of backscattering amplitude at different scales, corresponds to variations in both microbathymetry and large scale bathymetry. A method that can derive multiscale features from signals was needed, and the wavelet transform proved to be an efficient method of doing so. Wavelets are used for feature extraction in 1D sidescan and sidescan bathymetric survey data, and both the feature selection and classification stages are

automated. The method is tested on areas of known types and in general, the features selected show good correlation with sediment type, in both types of surveys.

The main disadvantage of this method, however, is that signal features are calculated per swathe (or received signal). Thus, sediment boundaries within the same swathe are not detected. To solve this problem, information present in consecutive pings of data can be used, leading to 2-D feature extraction.

Several textural classification methods are investigated for the segmentation of sidescan sonar images. The methods include 2D wavelets and Gabor filters. Effects of filter orientation, filter scale and window size are observed in both cases, and validated on given sonar images.

For sidescan bathymetric datasets, a novel method of classification using both sidescan images and depth maps is investigated. Backscattering amplitude and bathymetry images are both used for feature extraction. Features include amplitude-dependent features, textural features and bathymetric variation features. The method makes use of grab samples available in given areas of the survey for training the classifiers. Alternatively, clustering techniques are used to group the data. The results of applying the method on sidescan bathymetric surveys correlate with the grab samples available, as well as user-classified areas.

An automatic method for sidescan bathymetric classification offers a cost effective approach to classify large areas of seabed, with a fewer number of grab samples. This work sheds light on areas of feature extraction, selection and classification of sonar data.

Contents

List of Figures

List of Tables

List of Symbols and Abbreviations

Table of Symbols and Abbreviations most commonly occurring in this thesis:

θ	The grazing angle between the sonar beam and the seabed
DT	Detection Threshold of a sonar
DI	Directivity index of a sonar
SS	Spherical Spreading
AL	Absorption loss
NL	Noise level
BS	Backscattering strength
TL	One way transmission loss
A	Area insonified by the sonar
l	length of the area insonified by the sonar
w	width of the area insonified by the sonar
ψ	The mother wavelet function
ϕ	The wavelet scaling function
$Wf(u,s)$	The wavelet transform at time u and scale s
$J_{k,l}$	Distance between features k and l
BE	Backward elimination method
FS	Forward selection method
KNN	K-nearest neighbour method
PCA	Principal Component method
MLP	Multilayer Perceptron

DWT	Discrete Wavelet Transform
$GLCM$	Gray level Co-occurrence matrix method.
$S_{k,j}(x,y)$	Energy of the wavelet transform on scale k, and orientation j ($j = 1$ in 1D)
$SD_{k,j}(x,y)$	Standard deviation of the wavelet transform on scale k and orientation j
$nscale$	The number of scales in a wavelet decomposition.
$norient$	The number of orientations per scale
pw	Probablility of a window belonging to a certain class given its neighbourhood
D	Bathymetric variation feature
$(areax, areay)$	Window size
$nfeat$	Number of features
$nselectedfeat$	Number of features that a feature selection algorithm selects
$ntype$	Number of types used for classification
$M(x,y)$	Mean backscattering amplitude in window of position (x,y)
$M_s(x,y)$	Standard deviation of backscattering amplitude in window of position (x,y)

Chapter 1

Introduction

> 'There stands a tower by the Atlantic side. A grey old tower, by storm and sea-waves beat. Perch'd on a cliff, beneath it yawneth wide. A lofty cavern of yore a fit retreat. For pirates galleys; altho', now, you'll meet. Nought but the seal and wild gull; from that cave. A hundred steps do upwards lead your feet. Unto a lonely chamber! – Bold and brave. Is he who climbs that stair, all slippery from the wave.'
>
> *(poem Granuaile) Author unknown*

Seafloor classification, either in terms of physical properties or geographical provinces, plays an important role in understanding the undersea environment. However, only about 15% of the earth's oceans are mapped, and we have more detailed maps of the surface of the moon than our own seafloors. This is due to the inaccessibility of large areas of the seafloor and the attenuation of light if used to observe these areas. Since sound suffers much less attenuation, it has become the pre-eminent tool for sensing, identifying and communicating under water.

Visualisation of the seafloor is important for many activities necessary to many industries. A large commercial application is the use of seafloor surveys by oil companies. After finding where the oil fields are, safe areas have to be determined to lay pipelines leading to the oil patch. Thus, lava flows, underwater mountains and faults need to be avoided. The same applies for telephone

companies that lay fibre-optic cables underwater and want to avoid the effect of underwater land slides and fishing trawls. A precise knowledge of the environment is of great importance in mine hunting [149]. A priori contextual information of the seafloor may significantly improve detection and classification algorithms [77]. Sonar classification techniques are also of great importance to underwater archaeologists as many ancient coastal towns and cities are at present submerged under water due to eustatic and isostatic movements. Furthermore, many hundreds of wrecks are assumed to be lying on the seafloor or buried in the sediments on the seabed. For environmental as well as government agencies, observing the seafloor allows the monitoring of the environment and the observation of biological species as well as pollution and its effect on the sea. Geologists and geophysicists observe the seafloor and its layers for many reasons, for example to explain geological processes by studying morphological features, such as ridges, sediment deposits, and abyssal hills. The dredging industry could benefit greatly from detailed surveys of coastal areas that are classified into their corresponding sediment types. Automated underwater vehicles also benefit from the improvement of seabed classification techniques as that would imply an improvement in their environmental perception. These applications prove that seafloor classification is of great importance to many industries. The survey market is a rapidly growing one and the demand for fast and efficient methods to classify seafloors is increasing.

1.1 The Seabed Classification Problem

Backscattering from the seafloor involves several physical scattering processes that depend on the angle of the incidence of the sonar beam, as well as seabed sediment type and micro-bathymetry. By analysing the received backscattered signals, properties of the seabed can be deduced. However, using an accurate model of propagation and backscattering is a classical inverse problem that requires detailed knowledge of the seabed micro-bathymetry, which in most cases, is not available. This is why the approach generally used for tackling the seabed

classification problem is a multivariate statistical analysis of sonar returns rather than a model parameter fitting approach. Chapter 2 details the issues of parameter modelling, and presents several factors that affect the received sonar signals.

The approaches generally used to select statistical features from received signals can be divided into two categories:

- **Amplitude Approaches** Even in the early days of sonar, it was observed that the received backscattered amplitude differs between seabeds of different types. However, there are several other factors affecting the received signal amplitude such as depth, sonar orientation, and sonar frequency (absorption of the sonar signal in the seafloor is dependent on sonar frequency, lower frequencies are absorbed easier than higher ones [207]). When information on seabed bathymetry is available at a sufficient resolution, it can be used to correct for some of these factors and deduce the effect of seabed type on the received amplitude.

- **Spectral Approaches** By observing sonar pings plotted consecutively, texture differences can imply a difference in the type of the seabed observed. Spectral features can be observed in 1-D pings where referencing is not available, or where a real-time analysis of 1-D received signals is required. If a sonar image can be obtained, textural differences can be used to detect boundaries between different sediments.

The following section uses these two categories to present several of the techniques used in commercial systems available as well as in the literature. Some of the problems of these techniques will be discussed, leading to issues that motivated the approach investigated in this thesis.

1.1.1 Seabed Classification Techniques

This work considers swathe sonars due to their time advantage over single echo sounders and their ability to be used in obtaining a sidescan and a bathymetric image of the seafloor. Therefore, this section will focus on the commercial and research approaches to classifying seabeds

by using these sonars. On the other hand, sonars that use vertical incidence (single Echo Sounders) are widely used and were the first type developed commercially to classify seabed areas. Appendix B describes the technique used by ECHOplus and lists several commercial systems that use single echo sounders (such as Quester Tangent's QTC view and Stenmar's RoxAnn).

Methods used for Swathe Sonars (Multibeam, Sidescan and Sidescan Bathymetric Sonars)

If details of the local bathymetry of the seafloor are available, they can be used to correct for the effect of bathymetry on the received signal strength. Plotting the amplitude versus the corrected grazing angle (the angle between the sonar beam and the tangent to the seafloor) can be used efficiently to separate between sediment types. This is the approach used by several research groups([3, 148, 147, 36]). This approach, however, is usually applied to multibeam sonars with a sufficient bathymetry resolution. SESAM developed by NATO SACLANT uses this method for seabed classification of large areas, where each signal is corrected for the effect of bathymetry and depth [36, 172]. John Hughes Clark has incorporated a bottom classification tool using the NRL/Tulane scattering model, based on work by Jackson et al. [112] and Mourad and Jackson [152], into his SwathEdit Bathymetric Multibeam sonar processing software [65].

However, several sonars do not have the required bathymetric accuracy to model seafloor bathymetry accurately over small areas. The general approach used in this case is to observe textural properties of the received signals. In 1-D, features are selected from the spectral information of signals after correcting for the amplitude [179, 167, 200]. Although spectral features can be observed in 1-D signals and 2-D sonar images, there is a distinction in the literature available between these two cases. Analysing single sonar signals can be an interesting problem when it is not possible to obtain a sonar image (in some AUV's) or when features need to be extracted per ping (in a real-time approach).

When a sonar expert observes sonar images, textural differences are probably the most

important factor used to separate seabed types. To develop a machine learning approach that can separate between sonar types, features obtained from textural analysis are very widely used to classify sonar areas. Chapter 7, will explain the texture analysis problem as a general vision problem, and present several techniques that are used to classify textural images. The application of textural techniques to sonar problems will be detailed in chapter 8. Several commercial systems use textural properties for the analysis of sonar data, here is a summary of a few:

- Classiphi (Qinetiq's seabed classification and survey visualisation toolkit). Classiphi consists of two main packages - the Classiphi Toolkit and the Classiphi GIS Viewer. It provides high speed survey georeferencing and viewing of both sidescan and multibeam data. The classification algorithm used depends on sidescan texture and uses textural analysis to draw boundaries between seabed areas. The textural features considered are fractal features over user selected windows.[1]

- Seaclass by Triton Elics. This system uses an artificial neural network that is trained by selecting relevant areas of the survey by the user. Gray level Co-Occurrence matrices (detailed in chapter 7) are used to extract textural features (44 features are chosen). These features are used to train the neural network and classify the survey.[2]

- Genius by DHI Water and environment (Denmark), also selects textural features then uses a Self Organising Map for classification. [3] This system is based on the work described in the 1999 Danish Hydraulic Institute report by Babovic and Gopakumar [11].

- The Texture Mapping System by Geoacoustics Ltd. uses a supervised texture classification system. The user selects training areas in the survey. Textural features are selected from these areas and matched to other parts of the survey.[4]

[1] http://www.qinetiq.com/classiphi

[2] http://www.tritonelics.com

[3] http://www.dhi.dk

[4] http://www.geosys.co.jp

- The QTC SIDEVIEW system developed by Quester Tangent uses a large number of features selected from seabed spectral and spatial image properties. The features relating to texture include Co-occurrence matrix features, as well as FFTs, fractal features and histogram features. A feature selector is then used to select which of these features are important for classification (more than 150 features are used).

- The Amason system developed by Heriot Watt University. Among the aims of this system is the classification of sonar data by using textural information from sidescan sonar data. If other types of data are available (such as video), the system aims to extract features, then use machine learning methods for the selection of the most relevant data.More on this system can be found on http://www.ece.eps.hw.ac.uk/ amason/.

1.1.2 Some of the Problems of Seabed Classification Techniques

The previous paragraph presented a general introduction to the most commonly used approaches to seabed classification. Most of these methods attempt to overcome a similar list of problems, which will be detailed in this section.

One of the main problems is the lack of ground truth data. The only information to validate the results obtained by using a classification system, or to study the correlation of extracted features with the sediment type is usually that of a few grab samples or videos obtained from selected regions across the whole survey. Ground-truthing is usually an expensive and time consuming process, as divers are required to grab samples from different areas. These samples are analysed and features such as sediment size are studied. A classification method that can use only a few grabs as training data would be of great use to survey companies trying to reduce survey-validation costs.

A sonar expert usually examines several aspects of the data provided by sonars including spectral and spatial properties of the sonar signals as well as variations in the bathymetric graphs. However, a sonar expert would find it quite difficult to deduce similarities or differ-

ences between areas spanning tens of kilometres, due to survey size. Thus, extracting useful information from data presents a data-mining challenge, due to the size of the sonar datasets (usually several Gbytes.) A database representing a sonar survey cannot be simply considered an image database, where the main task is to find differences between areas in these images. The problem is much more complex:

-The seabed's bathymetry can vary a lot during a certain survey, leading to differences in the strength of signals received and the size of objects observed in seafloors.

-The motion of the sonar could have a great effect on the signals received, sometimes leading to artifacts in the sonar images.

-Gridding the sonar pings to their exact location is also a main issue to tackle as the image which is usually of pings received consecutively does not (due mainly to the survey track) represent the exact location of features on the seabed. (These issues will be dealt with in detail in the next chapter).

The big question is, under all these conditions, what type of sonar information can best describe different seabed sediments?

As mentioned in the previous paragraph, a model based approach is quite difficult to apply to vast areas of seabed. The parameters of the microbathymetry of the seabed are rarely known, and a general model that can fit backscattering from any type of sediment is not available (although models are available for specific seabed types under specific sonars [112]). The seabed type can vary from large areas of sand-waves of an evident bathymetric pattern, to areas covered by shells, or corals, or smooth areas of mud, or gravel. This is why statistical approaches are preferred in both research groups and commercial systems. A feature based approach is generally used. There are several problems faced :

-The type of feature to use, and which features are the most relevant considering a specific sonar system.

-The machine learning approach to deal with such a large amount of data, and in many cases (such as in the QTC sideview), the large amount of features extracted.

The problems mentioned in this paragraph highlight only a few of many problems faced in sonar classification studies. The next paragraph relates these problems to the motivation of this work and the steps taken to tackle some of these problems.

1.1.3 Direction and Themes

The availability of bathymetric data in addition to sidescan data provides extra information that can be used for seabed classification. In multibeam systems, bathymetric data of a good resolution is used to correct for the effects of seabed bathymetric changes (grazing angle) on a sonar signal. This study investigates the importance of the use of bathymetric data in sidescan bathymetric systems. In addition to using it for correcting for depth, this work investigates deriving features from bathymetry that can help in sonar classification.

The other direction that this thesis takes is that of feature selection. Is backscattering amplitude enough when the corrections due to depth are not accurate enough due to sonar resolution? and if not, can spectral features be an indication of seabed type. Inspired by the work of Reut *et al.* [179], Pace [167] and Tamsett [200], spectral features were investigated in this work for corrected 1-D signals. The approach taken is that of analysing features at different scales. Evidence of the applicability of this approach is also given.

Another issue tackled in this work is that of summarising sonar data. Studying a sonar survey in a pixel by pixel manner is almost impossible due to the size of some of these surveys (see figure 1.1 for a good example of why a pixel by pixel texture- analysis approach is very difficult to apply to the surveys used in this work). By analysing the properties of sonar signals over windows (of user defined size), an averaging approach is used. The features selected describe a window rather than describing a single point. This method is less affected by clutter, artifacts due to the sonar motion and errors in the bathymetry.

In large surveys, windows include samples from consecutive pings. Average textural properties over these windows can be an indication of seabed type. Note that the features used in this

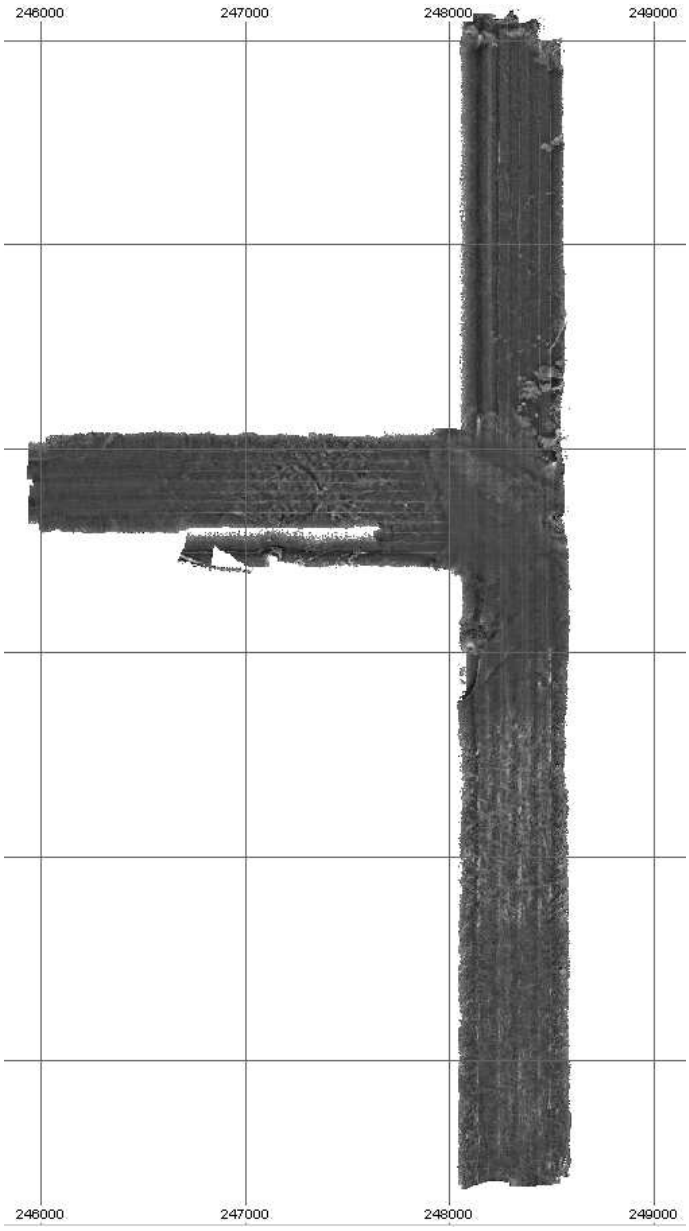


Figure 1.1: An example of displaying the sidescan amplitude for Plymouth sound, one of the surveys by SEA. Considering the whole survey as an image, or even an image database suffers from many problems. The most obvious ones in this image are the lines appearing due to sonar motion, the large influence of depth variation on sonar amplitude, and not having enough textural information if the survey is considered at such a large scale. This work attempts to correct for depth variation, as well as sonar motion. Then uses smaller portions of the survey for feature extraction.

work are derived from sonar data that is corrected for the effects of depth, sonar motion and directivity. In other words, this approach does not analyse raw sonar data as used in several "sonar-texture" studies.

The machine-learning approach to the sonar classification problem is also dealt with. After collecting features that provide separation between classes in a scatter plot, the following questions are tackled:

-What is a good way of knowing which features are relevant if we have a few training areas?

-What classifiers can be useful?

-How do the classified results compare with user observations of sonar surveys, and the properties of the sediments?

1.2 Contributions of the Thesis

This section describes the principal contributions made by this work. As far as the author is aware, this is the first study to offer an automatic method of seabed classification using sidescan-bathymetric sonar and benefiting simultaneously from both the amplitude and bathymetry properties[5]. The main contributions of this thesis are:

- The use of an acoustical model for the correction (or 'processing') of sidescan bathymetric signals. This model makes use of the provided bathymetric profile. Sparse bathymetric data is interpolated to obtain a bathymetric image providing more coverage of the seabed.

- The use of wavelets for the observation of spectral features in processed 1-D sonar signals. An automatic method is designed for feature extraction, selection and classification using 1-D wavelet properties.

[5]A UK patent application has been filed by Isis innovation Ltd. entitled "Surface texture determination method and apparatus."

- The application of the 1-D classification method to both sidescan and sidescan bathymetric datasets. The relationship between wavelet-derived features and grain size is analysed for sidescan bathymetric datasets. Neural networks are used for the classification of the whole survey where grab samples are used as training data.

- The use of information present in successive pings of 2-D sonar amplitude images and the extraction of textural features to classify different sediment classes. Unlike the 1-D approach, this method allows the observation of boundaries between sediment classes. Several textural filtering techniques such as Gabor filters and wavelets, are compared. These methods are applied to classify sidescan sonar data (artificial coral reefs, as well as sediment classes present in sidescan sonar images) as an initial step to study the effect of varying filter parameters to be used in the classification of sidescan bathymetric data. Neighbourhood information in sonar texture images (artificial coral reefs) is used for a better segmentation.

- The derivation of statistical features from depth images to observe sediment class differences, especially between rocky and smooth areas, and the selection of features that are relevant to high frequency sonar datasets.

- The combination of features for the 2-D classification of sidescan bathymetric data. The feature vector includes amplitude variation features, bathymetric variation features and textural features derived from 2-D Wavelet filtering. The most discriminating features are selected from the feature vector by using the grab samples as training data or by data clustering. These features are used to classify the surveys into sediment classes.

- Demonstration on several sidescan bathymetric datasets and comparison with other methods.

1.3 Overview of the Thesis

The first section of chapter 2 is a brief introduction to sonar and its development. The sonar types used in this work are then explained. A brief summary of the datasets used in the next chapters is given. These datasets include both sidescan data (artificial coral reef and sediment backscattered data) and sidescan bathymetric data (surveys from areas in Norway and Scotland). The factors that have an effect on sonar data are explained. These factors relate to sonar motion, the medium, the seafloor and the equipment.

Chapter 3 introduces the idea of using features over several scales. The advantages and the disadvantages of using multiscale feature selection are discussed, after proving that amplitude information on its own cannot provide efficient classification for sidescan-bathymetric datasets (even if survey and sonar parameters are corrected for). Factors that affect the features selected (such as range, look direction and beam pattern) are discussed in detail.

Chapter 4 introduces wavelet analysis as a means to observe time and scale features of signals simultaneously. The construction of the wavelet transform is summarised. Wavelet bases are then discussed, giving a few examples of wavelet bases types and wavelet bases selection criteria for various applications.

After the feature extraction stage, feature analysis plays an important role in the understanding of data. It includes feature visualisation, feature selection and classification. Chapter 5 summarises the feature analysis techniques that will be used in the next sections of the thesis.

The next part of the thesis is divided into two parts. The first part (consisting of chapter 6) uses 1-D sonar signals for signal analysis whereas the second part uses 2-D sonar signals for feature derivation.

Chapter 6 builds on the spectral feature work described by Pace [167] and Reut *et al.* [179], but uses wavelet analysis instead of Fourier transforms to extract features from sonar signals. Feature analysis methods described in chapter 5 are used for feature extraction and classification. The datasets classified by this method include a sidescan dataset and a processed

sidescan bathymetric dataset. For both methods, the process is an automatic one, using verified areas as training data.

The move to 2-D analysis starts in chapter 7 with a review of several texture classification techniques. These include gray level co-occurrence matrices, Gabor filters and 2-D wavelets, among others. The advantages of each technique are discussed and its relevance to the sonar-texture problem is emphasised. A discussion is also given on the importance of phase, multiscale resolution and orientation in texture classification problems.

2-D wavelet transforms are compared with Gabor filters in chapter 8 for the classification of two sidescan sonar datasets. One is of artificial coral reefs and the other is of areas of different sediments. This chapter serves as a guide to select filter parameters for the sidescan bathymetric datasets. In coral reef images, neighbourhood information (neighbourhood around a certain window) is used to obtain better segmentation results.

Chapter 9 introduces bathymetric features obtained from depth maps for the classification of seabed areas. These features are added to sidescan textural features and sonar amplitude features to obtain feature vectors in chapter 10. The feature selection and classification methods explained in chapter 5 are used to analyse sidescan bathymetric data. The methods are successfully applied to several surveys, using grab samples as training data when available.

Chapter 11 provides an analysis of the different results in this work, and an overview of methods used. Future applications are discussed.

1.4 Publications

The work in this thesis has resulted in four journal articles and one conference article. The work on 1-D sonar signal classification is summarised in the articles [4, 6], and in the proceedings of the sixth European Conference on Underwater Acoustics ECUA [5]. The application of this technique to the Hopvågen dataset is presented in [8]. The last chapters of this thesis present the 2-D classification technique using both sidescan and bathymetric information. This work

is summarised in [7].

Chapter 2

Sonar Types, Parameters and Datasets

Acoustical oceanography is the active or passive use of sound to study physical parameters and processes, as well as biological species and behaviours, at sea [146]. Due to the ocean floor's inaccessibility, acoustic remote-sensing with sonars, towed from research ships or installed in their hulls, presents the only feasible way of studying ocean floor processes quantitatively and over reasonably large areas. This chapter presents a general introduction to sonar, starting with a brief background, then discussing the sonar types that are used in this thesis. The sonar datasets are then described. Parameters affecting the received sonar signals are divided into many categories: parameters due to sonar motion, to the medium, the equipment and the seabed itself. The next chapter uses these parameters to develop a model for seabed correction and to discuss several of the features that can be used for seabed classification. The influence of several of the parameters discussed in this chapter on the features selected will also be dealt with in chapter 3.

2.1 Background

As early as 1490, Leonardo Da Vinci discovered that sound travels at a finite speed in water [137]. In 1635, Pierre Gassendi, a French philosopher measured the speed of sound in water.

Sir Isaac Newton published his *Mathematical principles Of Natural Philosophy* which included the first mathematical treatment of the theory of sound in 1687. Much of the theoretical work on sound in the eighteenth and nineteenth century was related to the study of pitch and quality of sound produced by musical instruments. Investigators in the field included such notables as Bernouilli, D'Alambert, Lagrange and Euler. Lord Rayleigh published his book *Theory Of Sound* [177] (1877). This monumental work established the basis for acoustic theory even as it exists today.

Fessenden developed the first high power underwater source (1912) and used it in 1914, to demonstrate echoranging on an iceberg at a distance of two miles [207]. The main stimulation for his study was the sinking of the Titanic in 1912 [32].

The initial application of acoustic technology in World War 1 was to detect and locate enemy bombers in air. But in 1915, the applications shifted underwater, with German submarine attacks on ships intensifying the effort to develop means for detecting submerged submarines. Passive systems were developed for that purpose, and had reasonable success. However, the accuracy needed for weapon delivery led to the interest in active echo ranging systems. The detection of a submarine with an active echo ranging system was demonstrated by Paul Langevin in 1917. An important understanding of the physical processes involved in sound propagation in the sea began in the 1930s. The absorption of sound by sea water was measured at ultrasonic frequencies at the Naval Research Laboratory (USA). The behaviour of the ocean as a sound transmission medium began to be understood with the invention of the bathythermograph by Spilhaus in 1937. This device permitted rapid and convenient measurement of the water temperature versus depth, from which the sound-speed depth profile could be calculated.

During the second world war an extensive program in underwater acoustics was organised by division 6 of the National Defence Research Centre (USA). In this period, other groups also carried out experiments to improve the design and use of underwater acoustic systems. At the end of the war, this knowledge was gathered in a series of reports called *Physics of Sound in the Sea*. These reports formulated the theoretical basis for describing sound propagation

in the sea, reverberation and the reflection of sound from submarines, ships and wakes in much the same form as it exists today. By the end of the war, the factors affecting the sound speed in the ocean were identified and understood. Experimental measurements of sound speed covered a wide variety of sensors, weather conditions and water depths. The progress of radar technology (the radar technology program in the Radiation Laboratory of MIT) helped in advancing underwater sonar signal processing. The problem of processing weak radar signals in the presence of noise is essentially the same as the sonar problem. From approximately 1945 to 1955 the work of Shannon [187] in the United States, and Gabor [89] and Woodward [223] in England, established the subject of information theory as a major discipline. They provided new insight into the properties of signal waveform design that affect system capability in the areas of target resolution and estimation of target range and velocity.

The 15 years immediately following the war saw the beginning and rapid development of solid-state technology. In the field of underwater acoustics, the experimental and theoretical work begun during the war was continued and expanded to cover sonic as well as ultrasonic regions. Knudsen [121], Urick [206] and others identified the various sources and characteristics of ambient noise in the ocean. Additional understanding in the causes of absorption of sound in the sea was provided by Lieberman [132] (among others) and experimentally determined absorption coefficients were available from below 100 Hz to above 1MHz [205]. The increasing availability of general purpose computers made it possible to analyse this data and to develop a statistical picture of the relative frequency of occurrence of various propagation conditions [32]. Advances in science and technique were concentrating mainly on the medium and its boundaries, but it was not until 1958 that the first open publication in the field drew attention to the fact that bottom reverberation could be presented in such a way as to be of great value in underwater geology [43] and the first application paper appeared a year later by Stride [195].

The first major scientific paper primarily based on sidescan sonar was published in 1961 [74]. By sidescan sonar we mean a sonar (looking sideways), whose images are a mapping of the strength of acoustic backscattering from the sea floor (a technical explanation of the sidescan

sonar follows in the next section). In parallel with the interest in qualitative information of the sea revealed by sidescan sonar, there was a growth of interest in numerical bathymetry (sea bottom depth profile) for scientific and commercial use, both in the deep sea and shelf waters. This stimulated much research in Swathe-Bathymetry. The idea of swathe soundings has been put forward by many researchers including Haines [96], Ritchie [180] and Tucker [204]. The standard technique of surveying the seabed was to use a ship mounted echo-sounder[1]. This only measured the depth below the survey vessel, and in order to achieve satisfactory coverage of the seabed, it is usually necessary to survey sets of parallel sounding lines [94], [95]. The number and spacing of these lines will depend upon a number of factors including the scale of the survey and its purpose. Even with closely spaced survey lines it is still not possible to guarantee full coverage of the seabed, and isolated features that lie between the lines will be missed [95].

Another approach to solve this problem was to use a transducer arrangement which provides a number of beams directed at different angles from the vertical. This was first put forward by Tucker [204], then demonstrated by Howson and Dunn [105]. Throughout the 1960s and early 1970s work was carried out on these multibeam systems, refining beam patterns, data collection techniques and generally improving their accuracy [104, 92]. For each beam the range estimate is determined by calculating the centre of mass in the echo signal contained within an appropriate reception window and exceeding a suitable threshold. However, variations in seabed scattering strength mean may cause a limitation of the accuracy of this system. This is due to the fact that the range estimate may not correspond to the centre of the beam.

An alternative to multibeam sonar that is more compact and does not suffer from this problem is the *bathymetric sidescan sonar*. This sonar measures the declination angle of a ray from the seabed by the simple procedure of measuring the phase difference of the signals in two closely spaced receivers one above the other [71] (technical details of the use of this method

[1] An echo-sounding system uses the time interval between the initiation of a sound pulse and echo returned from the bottom to determine the depth of the bottom.

known as *interferometry* will be explained in the next section). A description of such a system was first made by Denbigh [70, 69]. He used the term 'bathymetric sidescan sonar' to describe the technique, and the acronym BASS for its realization on a towfish. The only difference between this and the now more commonly adopted term of 'swathe bathymetry' is that rather than smoothing the data for display purposes, Denbigh retained almost the full resolution of the sidescan sonar in order to produce highly detailed images of sea topography.

The sidescan bathymetric sonar results in a sidescan image of swaths covering the surveyed area, and a bathymetric image of the swath series. The sidescan image is a series of backscattered amplitude profiles which are affected by the material properties of the sediments in the seafloor as well as its general profile. Thus, studying the properties of the sonar signals is a way of remote classification of seabed areas.

2.2 Sonar Types

The choice of a particular technology or type of sonar is based on the primary survey objectives such as what is to be surveyed, what resolution is needed for the survey and what spatial sampling is required. The decision is also based on the cost of each type of equipment. This section explains the sidescan and sidescan bathymetric sonars. Single echo sonars are summarised in the appendix (appendix B) since the results of classification using these sonars are compared with our results in chapters 6 and 10. However, echo sounders are not used in this work.

Another type of sonar that is widely used is the multibeam sonar. Multibeam sonars use a number of sonar-beams (usually about a 100) arranged in a fan shape, port and starboard, collecting echos from the entire swath width as the ship advances. For each beam of each ping, the system reports the sea-depth as well as the echo amplitude (or at least its maximum over time). Although the final result of the multibeam sonar is similar to that of the sidescan-bathymetric sonar, it was not used in this study, as the datasets provided were only sidescan and sidescan-bathymetric datasets. The reader can refer to [92, 68, 56, 9] for more information

on multibeam systems.

2.2.1 The Sidescan Sonar

The sidescan sonar transmits a time limited pulse of acoustic energy into the water medium. A volume of water and the sea bottom or the sea surface may be insonified. The signal is partially scattered by various irregularities in the water medium and by any changes in acoustic impedance. It is then received, displayed and analysed.

The process for producing the outgoing pulse of acoustic energy is to generate an oscillating electric field of the required frequency, duration and energy. The electrical pulse is then applied across the transducer which in most sonar systems consists of a piezoelectric material. This material expands or contracts when an oscillating voltage is applied across it, producing an oscillating pressure wave, or the acoustic pulse which is transmitted into the water medium. The acoustic pulse then propagates through the water column until it encounters something such as the seabed, sea surface or an object, some of that signal is scattered back towards the source of the acoustic signal [93].

In most circumstances the same aperture which transmitted the pulse is used to receive the scattered signal, and the acoustic pressure is converted to a voltage using the piezoelectric properties of the transducer. The system electronics then amplify, display and store the signals for post processing. In addition to the sonar signals, deployment sensors may also record the environmental data of the sidescan sonar such as roll, pitch, heave and the navigation information. To obtain a processed signal, independent of these factors, the user must correct for the errors, automatically or on-line.

Along track scanning is achieved by the physical transition of the transducer. In this direction scanning is not continuous but the object is sampled by the sequence of discrete pulse transmissions. The sidescan sonar can be towed in a vessel, or hull mounted. For technical and economic reasons, the great majority of sidescans are deployed in a towfish. The only

circumstance in which a hull mounted array would nowadays be viable would be if the ship and sonar were in common ownership and the sonar workload would justify the initial expense of the necessary work.

The transducers are normally mounted on both sides of the sonar system and the data is recorded on two separate channels to ensure maximum coverage of the sea bed. They are shaped so that they produce a beam which is very narrow in the horizontal and large in the vertical. A large vertical beam means that as much of the seabed is covered away from the fish (the vessel) as possible and a narrow horizontal beam achieves high resolution laterally along the area insonified.

The scanning process occurs in two directions. A pulse in the form of a fan-beam is sent out perpendicular to the sidescan sonar with the main beam axis at a set angle down from the horizontal, and at the same time the sidescan sonar has moved along its track and the next pulse would insonify a different area of seabed (as in figure 2.1). The angle between the beam and the tangent to the surface of the seabed will be referred to as the grazing angle θ in this study.

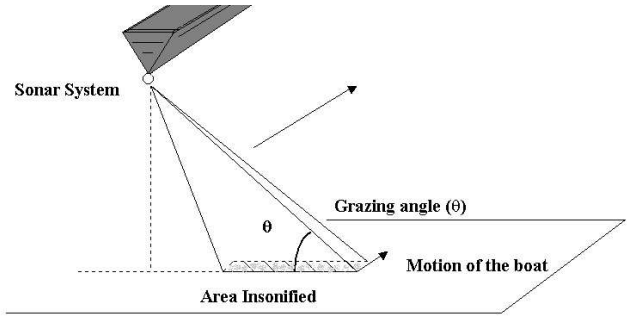


Figure 2.1: A figure showing the geometry of the sidescan sonar system.

The range resolution of a sidescan sonar is related to the bandwidth of the transmitted pulse and is of the order $\frac{c}{2B}$ where c is the velocity of sound in water and B is the signal bandwidth, leading in practical terms to a resolution varying from 10 cm at $100kHz$ to 10 m at $6kHz$ [190]. The along-track resolution is either the width of the horizontal beam on the ground or distance

travelled by the transducer during the reception interval, whichever is worse.

In designing a sidescan sonar, designers have to take many factors into consideration. One of these factors is the trade-off between the use of high frequencies and the attenuation of sound with frequency. The result is that underwater acoustic equipment is sorted into frequency groups by application. In sidescan sonar the applications range in order of decreasing frequency from local high resolution survey requirement covering swaths of 100 m or so to the 1 km suitable for offshore continental shelf waters to the deep-ocean regional applications where swaths of many kilometres are useful.

2.2.2 The Sidescan Bathymetric Sonar

In addition to the sidescan backscattered amplitude signal, this type of sonar allows the observation of a seabed's depth profile by the use of *interferometry*. The term 'interferometry' is generally used to describe swathe-sounding sonar techniques that use the phase content of the sonar signal to measure the angle of a wave front returned from a sonar target. It was inspired by the interference patterns sometimes observed in conventional sidescan records due to the Lloyd mirror effect. This effect is caused by the interference between a direct path signal and the return signal reflected off the sea surface. It was realised by Chesterman et al. [44] and by Heaton and Haslett [100] that each fringe corresponds to a certain declination angle and could therefore provide useful depth information of the sea bottom. Using a system of two receivers to measure depth was first described by Denbigh [70, 69].

The interferometric process involves a transducer consisting of a transmit array and at least two receive arrays (or staves), each array is equivalent to a normal sidescan array. The transmit and receive beams are narrow in azimuth and wide in elevation. The argument also assumes that the pulse is of a short duration, so that at only one instant a single patch (considered as a point just to make the geometry simpler) is insonified by the pulse. This patch scatters sound energy in all directions. When this scattered sound is detected back at the interferometric

transducers, the angle it makes with the transducer is measured. The range is calculated from the travel time there-and-back. The range and angle pair enable the location of the insonified patch to be known relative to the transducer.

Figure 2.2 shows a transducer pair and their position with respect to the sea bed. Figure 2.3 shows the details of a wave arriving at the transducer.

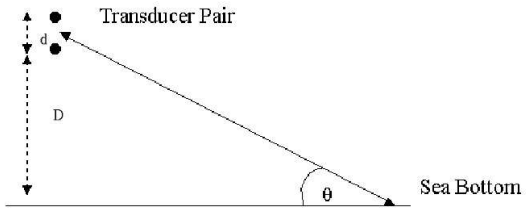


Figure 2.2: Arrangement of the transducers relative to a flat sea bed

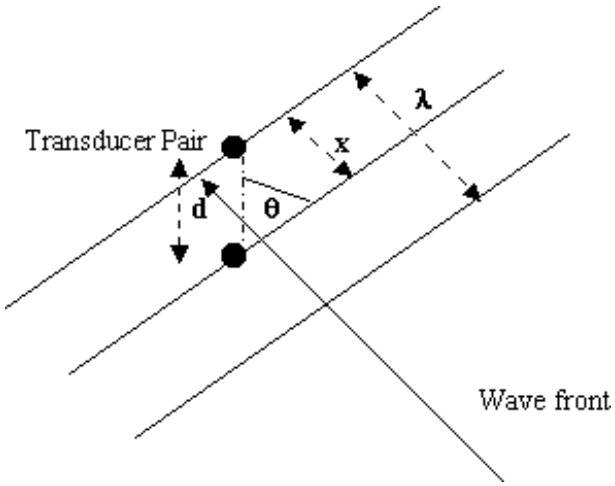


Figure 2.3: Wavefront arriving at the transducer pair

The elevation angle θ can be related to the transducer spacing d and the distance x (see figure 2.3) by the following geometrical relation

$$x = d\sin(\theta) \tag{2.1}$$

The ratio between the distance x and one wavelength λ can be related to the phase difference

ϕ between the two transducers and 2π radians:

$$\frac{x}{\lambda} = \frac{\phi}{2\pi} \tag{2.2}$$

Substituting the value of x from equation 2.2 into 2.1, the grazing angle θ can be calculated as follows:

$$\theta = \arcsin(\frac{\phi\lambda}{2\pi d}) \tag{2.3}$$

The phase angle ϕ wraps around at 2π, so that the range of angles which can be measured with a given transducer spacing is limited. Widely spaced receivers can improve final scale accuracy, but there is an error introduced due to the ambiguity problem. If the receivers are half a wavelength apart then for each value of phase difference ϕ there exists one elevation angle θ, which means there is no ambiguity problem in this case. However, the transducer spacing is increased to one wavelength λ, more than one elevation angle corresponds to the same phase difference measurement as shown in figure 2.4 [71, 72].

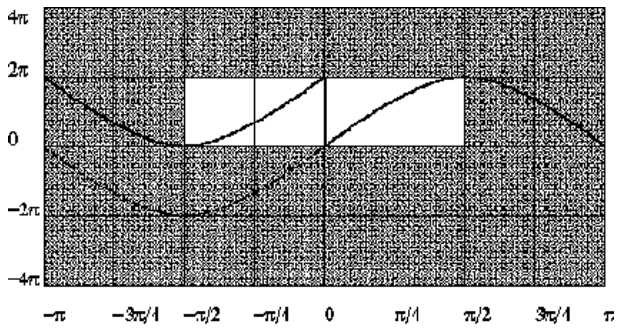


Figure 2.4: This figure presents the problem of ambiguity for a transducer spacing of one wavelength, a certain phase difference (vertical axis) corresponds to more than one elevation angle (the horizontal axis)

A technique for solving this problem is to use a second pair of receivers having a different spacing such that the combination of phase measurements constitutes a 'signature' that is unique to each declination angle. For example a single phase measurement of ϕ_1 might correspond to any of the three declination angles θ_1, θ_2 and θ_3. A second phase measurement of

ϕ_2 from a different receiver spacing might correspond to any of the three declination angles θ_4, θ_5 and θ_6 . If θ_2 and θ_5 are the same, but the other angles are different then clearly the pair of phase differences ϕ_1 and ϕ_2 correspond unambiguously to the declination angle θ_2 . This approach can be extended to overcome any ambiguity, by the use of more than 2 pairs of receivers. It is the basis for the 'vernier' technique. The name arises because if the difference in the separation of the two pairs of receivers is small, the subtraction of the two phase outputs gives a coarse but unambiguous measurement that resolves the fine but ambiguous measurement due to one pair alone. Another technique to remove ambiguities is to choose a correct starting point, then keep track of the phase, which is expected to change slowly and monotonously as successive points along the swath are collected [71].

A major difficulty with sidescan bathymetric sonars is that acoustic interference, such as sea surface backscattering, can cause multiple signals to arrive simultaneously at the receivers from different directions. In this case, the phase difference error fluctuates with time and could be very large when the signal and interference have a similar amplitude. Methods of reducing this interference include getting away from the sea surface by using a deep tow. However, this might pose a problem if large areas are to be surveyed since the sonar would be too close to the seabed, or if shallow areas are surveyed. Another alternative is to make the receivers highly insensitive in the direction of the sea surface (using amplitude shading), or to use acoustic baffles.

The use of a large acoustic baffle may be undesirable because of its effect on hydrodynamics and handling. Thus, Denbigh [72] describes the following signal processing strategies as a more practical option to minimise the effect of interference:

- Using widely spaced receivers in order to average out irregularities caused by interference. However, directional ambiguities can arise, and must be eliminated.

- Removing inconsistent measurements (from a single acoustic source) before averaging. One way of doing this is to demand that signal amplitude should be similar at each

receiver. Another way is to use many receivers and demand that the phase difference measurements of each pair are consistent with the same depth.

- Detecting measurements coming from an unwanted direction (the sea surface) and removing them before averaging.
- Rejecting depth signals corresponding to very weak received signals.
- Using a shorter acoustic pulse, which increases the number of measurements that can be averaged within a given range resolution cell.

Another problem that degrades the performance of the interference method is the variation in the speed of sound which causes a variation in the angle of elevation and thus the depth measurement. This is due to the fact that λ (the arriving wavelength) is assumed to be constant but in fact it changes with the change of speed of sound. If the speed of sound is determined accurately at the start of each survey, then this error will be minimised.

2.3 Sonar Datasets Used in This Work

The datasets used in this work were of two types: sidescan sonar datasets and sidescan-bathymetric sonar datasets. This section summarises these datasets, and presents a brief overview of the sonars used.

2.3.1 The Sidescan Datasets

Pace's Dataset

This dataset was provided for the purpose of this work by Dr. Nicholas Pace (NATO SACLANT Undersea research centre). The dataset is that of a sidescan sonar that operates at a centre frequency of 48 kHz. The sampling period is 125 μs. The sonar transmits a pulse of 1 ms duration, has a nominal bandwidth of 2 kHz, and a pulse repetition rate of 1.33 Hz.

Data were available of six sea floor types, these being: sand, mud, clay, gravel, stones and rock. Each sidescan data file was gathered at a different site and verified by grab sampling and diver inspection. For example, the gravel data files were gathered in the West English Channel at sites 1 kilometre apart, while the rock data-files were gathered in The South Baltic sea. The data file for each seabed type consists of 128 contiguous subtraces. Each subtrace is made of 1024 samples. Thus, each file corresponds to 96 m in the across track and 284 m in the along track directions [37]. Data were already processed on-line, so the traces present data which is corrected. The time varied gain was operator defined and not recorded. Further information about this data set is found in Pace's 1988 paper [167] and Tamsett's [200] paper.

The Artificial Coral Reef Dataset

This dataset was provided by Dr Jenny Collier and Stuart Humber (Imperial College). The sonar used is a 410 kHz sidescan sonar. The survey was done at 4 knots, the ping frequency is 6 Hz giving a ping spacing of 30 cm. An automatic gain control filter was used during the survey to correct for attenuation. The files received were of artificial coral reefs that were generated in the ocean. The main aim was to tell the difference between the coral reefs and the surrounding sediments. The coral reefs were of two different types as seen in figure 2.5

2.3.2 The Sidescan Bathymetric Datasets

The sidescan bathymetric datasets used in this work were provided by SEA (Ltd.). The sonar-system specifications are summarised in appendix C. The sonars used are available in two frequency versions 117 kHz and 234 kHz. Table C.1 lists some of their main properties.

As explained in appendix C, the system is provided with filters to select acceptable data (for display purposes) from the large amount of received data. However, excessive filtering results in the loss of datapoints that could be useful for signal analysis. This work considers the 'raw' sonar files and only uses the minimum amount of filtering to reserve the maximum number of

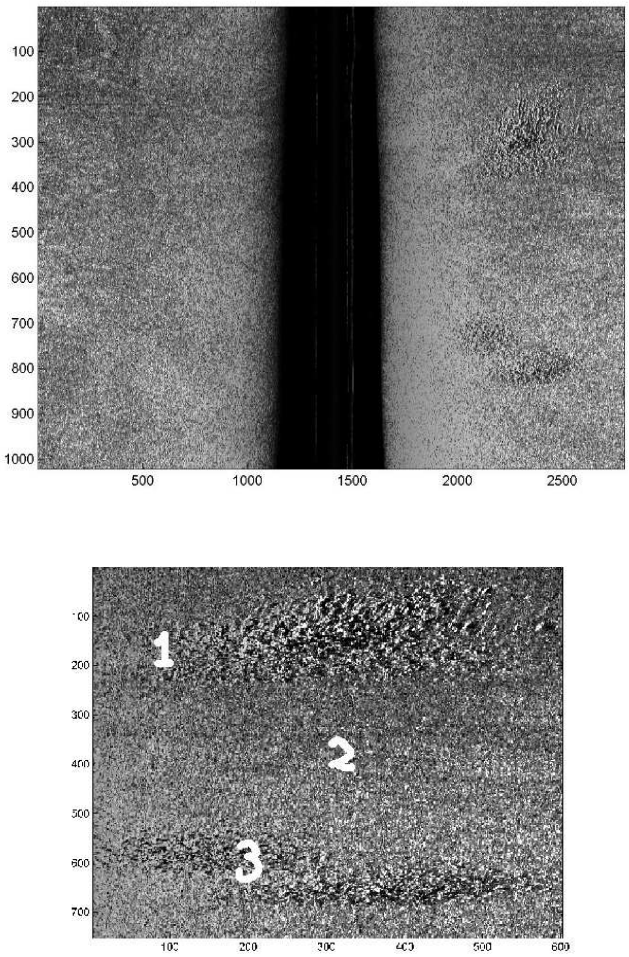


Figure 2.5: In the first figure, the total survey is presented, the artificial coral reefs appear on the right hand side. In the second, the area around the reefs is magnified and three types are considered for study: type 1, representing the first type of coral reef, type 2 which is the sediment around the reefs and type 3 which is the third type of reef.

points that could be used for data analysis. The correction and interpolation of points will be explained further in section 3.1.1.

To visualise the data received by the SEA system, the user can use the software package RTS2000 and view the received amplitude for each swathe (a one dimensional plot) or the one dimensional seabed shape (bathymetry). The user can also view a two dimensional sidescan image (grey colour coded according to the strength of the reflected signal) or a two dimensional bathymetric image which is colour coded according to depth [2]

The Hopvågen Dataset

Hopvågen bay lies approximately 80 km west of Trondheim, Norway. It has a narrow and shallow exit to the main fjord through which the bay receives an 11% turnover in water volume during each tidal cycle. The bay has a deeper northern half and a shallow southern shelf (figure 2.6). The sediments in the bay consist of fine-grained muds located in the deep northern half and a mixture of sand and silt on the shallow shelf to the south. Around the bay margins of rock outcrops and occasional boulder beds are colonised by horse mussels to a depth of approximately 8 m [15]. The bathymetric sidescan sonar of 243 kHz was provided by SEA Group Ltd.[3] and used to survey Hopvågen in a series of NE - SW transect lines. Eight lines were considered for this study (figure 2.6). Data was recorded using RTS2000. Figure 2.6 is a coloured bathymetric view of the survey, showing the lines covered by the sidescan bathymetric sonar and used for this study [15].

To provide ground truth for the classification algorithm, sediment samples were taken at 50 sites in the bay for grain size analysis in the laboratory. No samples could be obtained using a grab sampler with boulder bottom types as the boulders prevented the grab sampler from recovering material. The sediment was initially dried and then sieved at 2 mm. The remaining material was split in a riffle box to obtain $3 - 10$ grams of representative sample. This was

[2] Other programs can be used to visualise the bathymetric data such as Grid2000 and Quickgrid.

[3] www.sea.co.uk

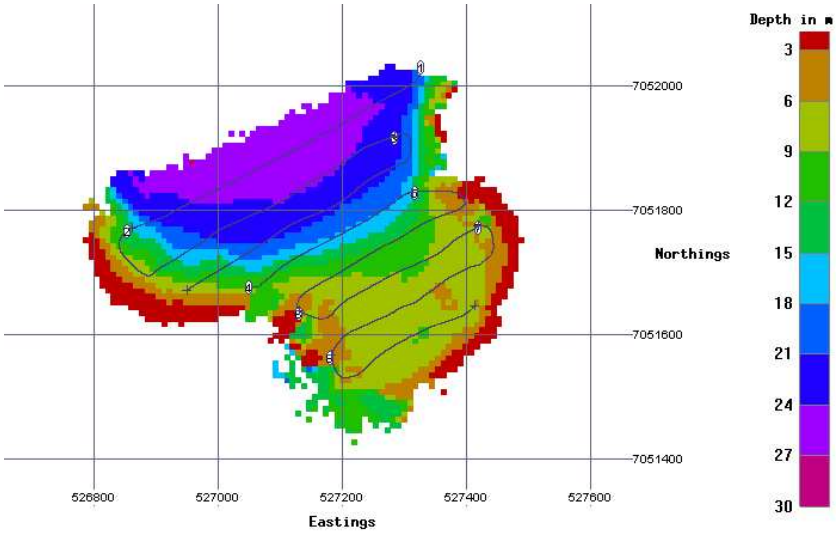


Figure 2.6: A colored bathymetric map of the Hopvågen survey, showing the lines covered by the sidescan bathymetric sonar and used for this study.

analysed using a Couter LS laser diffraction particle size analyser. Histograms of particle size were produced and the average grain size was determined [15]. Of the 50 grab samples, 22 were within the surveyed area provided for this study. The Wentworth scale [216] was used to assign a sediment class to each grab sample as given in table 2.1 Four classes were distinguished. The table also shows the number of grabs available of each type.

Sediment Type	Grain Size	Number of grabs available
1-Medium to Coarse Sand	$> 250\mu m$	8
2-Fine Sand	$125 - 250\mu m$	6
3-Very Fine Sand	$63 - 125\mu m$	2
4-Mud/Silt	$0.06 - 63\mu m$	6

Table 2.1: The number of grabs of each type of sediment in the Hopvågen survey is given in this table. The classification was done according to the Wentworth scale [216] given also in appendix A.

The Laxford Dataset

The survey in Loch Laxford was done for Scottish Natural Heritage by the Sedimentary Systems Research Unit, University of St Andrews and the Department of Biological Sciences Heriot Watt University [14]. Loch Laxford is the furthest north major sea loch on the west coast of Scotland and represents an excellent example of a large shallow inlet and bay on the west coast of mainland Scotland. The site contains a wide variety of marine habitats and communities resulting from a range of environmental conditions. These include exposed sublittoral reefs at the western loch mouth to sheltered littoral reefs and a variety of sediment habitats in the inner bay. The region is characterised by rocky coasts and high salinity, low turbidity waters. The coastal waters support a wide range of marine communities, essentially boreal in character, which reflect the range of available substrata, the variation in exposure to wave action and tidal currents, and the influence of the warm waters of the Gulf Stream. The loch has a complex fjardic shape with numerous small islands and side branches that include two subsidiary lochs. The central channel is relatively straight but the coastline is long and convoluted. The aspect of the central channel means that strong north-westerly wind and wave action can influence environmental conditions to considerable distances within the loch along this narrow zone. The main loch has two poorly defined basins and a single, broad sill across the entrance rising to a moderate depth of only 40m. The smaller outer basin reaches 67m depth whilst the rest of the loch is relatively shallow. The coastline of Loch Laxford is predominantly steep and rocky with both bedrock and boulder shores dominated by algae. Figure 2.7 is the bathymetric profile of the Loch.

Evaluation of the site included broad scale mapping at near 100% coverage using acoustical methods, high fidelity ground truth observations on point and line tracks using diver, video and remotely operated vehicle (ROV) methods, and biological and physical analyses of sediment samples collected by grabs and cores. The area was surveyed by the ECHOplus system as well as SEA's sidescan bathymetric sonar with a frequency of 117 kHz (note the difference between

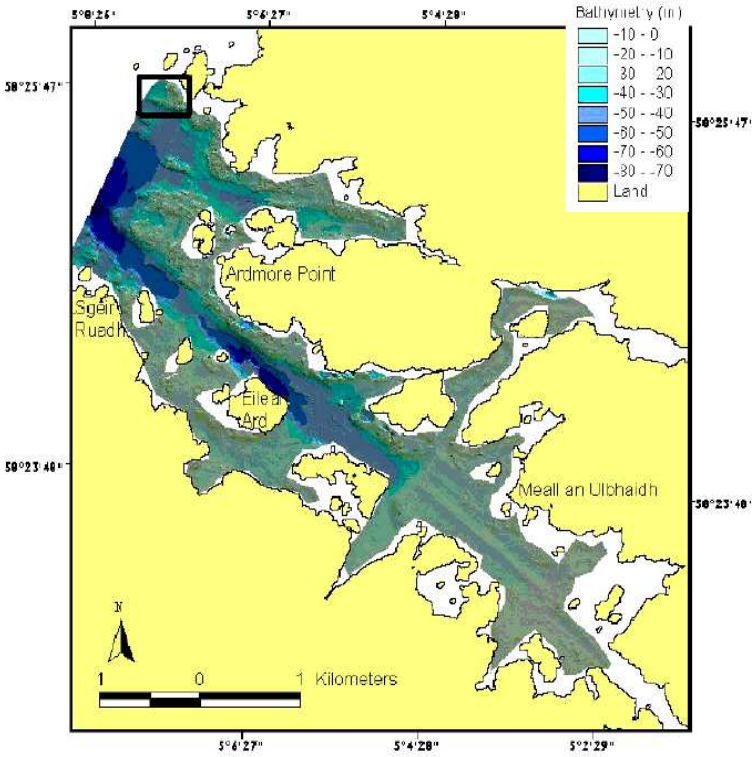


Figure 2.7: The bathymetry of Loch Laxford. The part used in this study is the selected area in the rectangle.

the two frequencies used in Laxford and Hopvågen). More details about the whole dataset can be found in [14].

This study considers only a small part of the survey which was previously studied. The part considered is a part of the outer Loch known to be varying between circalittoral rock or boulders and circalittoral sand and silty gravel . It is displayed as a rectangle in figure 2.7 and magnified in figure 2.8.

2.4 Parameters Affecting the Sonar Signal

To be able to study the effects of seabed type on the backscattered signals received by sonars, an analysis of the parameters affecting the signals is needed. An understanding of these parameters allows the correction of sonar signals for losses due to the system and the medium. The

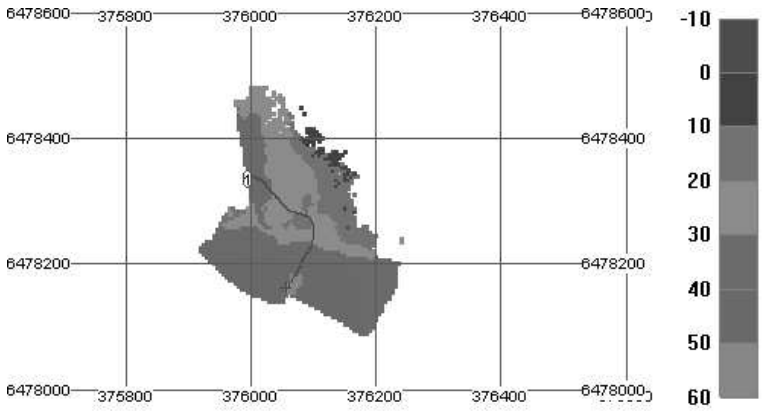


Figure 2.8: The part of Loch Laxford selected for classification in this study.

parameters affecting the absorption and backscattering of sonar signals can be divided into four categories: parameters due to sonar motion, parameters determined by the sonar equipment, parameters determined by the medium, and parameters determined by seabed type. This section discusses these parameters briefly, then derives an acoustic model for the correction of signals (the model is a standard one used in many sonar studies: [207, 193, 13]). The main aim of this work is to provide means for the correction of sidescan bathymetric sonar datasets since the other datasets used were already processed online (the sidescan datasets for example, section 2.3.1).

2.4.1 Sonar Motion

The sonar (whether sidescan or sidescan bathymetric) is usually towed from a vessel and is connected by a cable which supplies power and communication link to the data logger on board the vessel. The motion of the sonar is influenced by the motion of the boat, which is affected by winds and sea currents. The sonar's motion itself can be also affected by undersea currents, changes in sea density and eddies. The instabilities in sonar motion can be separated into two categories: translational and rotational. The translational motion describes both the sidescan position in 3-D space and its velocity. The rotational motion relates to the orientation of the

sidescan sonar and is expressed in terms of roll, pitch and yaw. This section describes some of these instabilities and discusses the calibration used to solve for some of them.

Factors Affecting Translational Sonar Motion

One of the main problems affecting the analysis and display of sonar data is the inconsistency of ship speed as the data is gathered. This leads to the variation of distance between successive pings, leading to scale distortion (either compression or elongation) parallel to the line of travel. As the towing speed gets faster, objects or sub-bottom features will appear shorter in the along track direction. They will also appear with less detail since fewer returns will be received for each object. The apparent angles of features on the seabed are also distorted by speed variations, making the determination of target orientation or sediment boundaries difficult [93].

Another problem that causes an error in sonar positioning is that of navigation latency, which is a delay between the position being measured and the information being transmitted to the sonar. The measured position is located too far back along the vessel's track. The error is along-track in direction and increases with increasing vessel speed. The scan lines from the maximum receive signal are bent forward as the acoustic axis of the receive beam pattern moves forward [17]. However, approximating the scan lines as straight lines is a valid approximation in most surveys since the maximum range of the sonar is much larger than the spacing between scan lines in the along track direction. The slower the tow speed the smaller the error of approximation. Many modern navigation systems use a predictive filter to give zero navigation latency. Other systems try to solve this problem by calibrating the system and calculating an average position offset for a given ship speed. Calibration of position offset is usually done by surveying a small area twice, with two different ship speeds, then observing the differences using a gridding program.

Heave is another problem encountered during surveys, and is defined as the rise and fall of a surface vessel or towfish in a rhythmic movement. This could lead to disjointed, jagged images on a sonar record. Towbody heave is a major cause of data distortion, particularly

when rough seas affect the stability of the surface vessel. Although, under some conditions, longer tow-cables will dampen the effect of heave seen in the data, under others it can result in a harmonic effect, increasing the heave. Heave compensators, which increase and decrease the amount of in-water cable, have been applied in the past. The concept is sound, but the systems are typically very complex and, under some conditions, will significantly add to the problem of heave.

Factors Affecting Rotational Sonar Motion

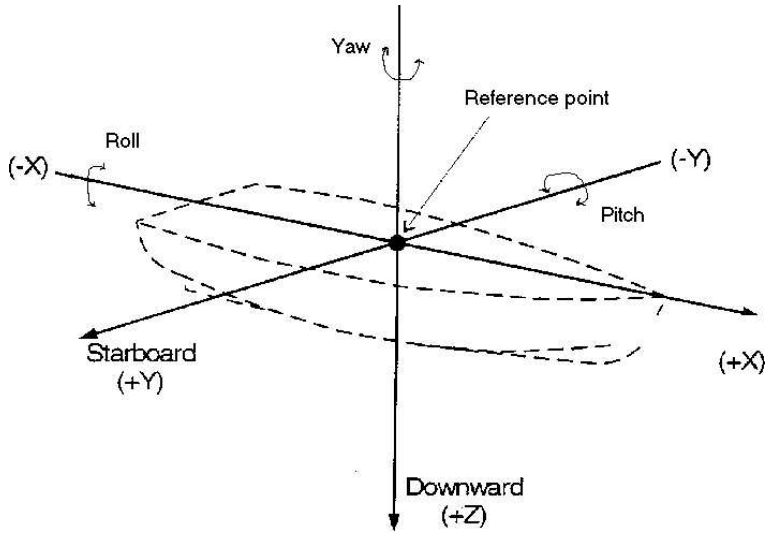


Figure 2.9: The axes as used in section 2.4.1.

Roll Roll is the rhythmic movement of a ship or tow-body about its longitudinal axis (axis X in figure 2.9), moving the beam left or right. Rolling alters the depression angle at which the acoustic axis is oriented within the vertical plane and the horizontal range of the seabed insonified by the main lobe of the beam. When working from a well designed support vessel, roll will not contribute to towfish instability to the same extent as heave in any given sea state. The main effect of rolling is that of periodic compression and stretching of the sidescan image.

Pitch Pitch is the oscillation of a ship or towed fish[4] about its transverse axis (axis Y in figure 2.9) leading to the oscillation of the sonar beam forwards and backwards within the vertical plane. It is depicted as the alternate rise and fall of the bow of a vessel proceeding through waves. Although pitch has an effect in locating the transmit footprint in a fore and aft direction, the effect is relatively slight.

Yaw Yaw is defined as the angular motion in the horizontal plane containing the projection of the longitudinal axis (X), about the mean course. It is the oscillation about the vertical axis; bow moves left and stern moves right (and vice versa). Yaw is usually due to a short period of off-course motion of the survey boat.

Correction and Calibration

When a ship is at sea, the sonar array is continuously moving with respect to the sea floor due to the ship's motion. The system must then continuously correct the sonar data for the ship's motion by comparing the orientation of the ship at any instance with some known vertical reference. Most sonars take input from separate gyroscopic or inertial navigation systems which provide these corrections in real time. These corrections are normally implemented in the sonar software, and are applied to the received sonar signals before display. Before starting a sonar survey, calibration is normally done to set sonar motion offsets, usually on a small part of the survey. Attitude sensors measure roll, pitch and yaw offsets, and in many commercial systems, these parameters are observed and used to correct for errors in the observed signals. The positioning system (usually GPS) must also be calibrated to correct for errors in positioning. We will not go into calibration details in this work since they are system specific and the reader can refer to different sonar company websites for more details on their calibration schemes.

[4]fish refers to the towed vehicle carrying the sonar.

2.4.2 Parameters Determined by the Sonar Equipment

Detection Threshold and Transmitted Source Level

The detection threshold (DT) is the measure of the intensity of the wanted signal to that of the unwanted signal. In sonar equipment, the signal to noise ratio needs to be quite significant due to the large amount of ambient sea-noise usually present in surveys (see section 2.4.3). This parameter is sonar-dependent, and is provided by the manufacturers. The transmitted source level (SL) is also sonar-dependent and usually provided by the manufacturers. It is usually defined as the intensity in dB of the transmitted pulse measured 1 metre from the transducer along the acoustic axis [93].

Directivity Index

The directivity index represents the directional properties of the transducer. It determines the amount by which the transducer array, through its beam pattern can improve the signal to noise ratio. Array directivity should reduce the noise arriving at the array from directions other than that of the signal (sea-surface for example). The directivity index of the ping depends on the angle of the transducers and the beam shape of the acoustic wave. It should also be modified to overall beam drop off along the decreasing grazing angle. (An example of directivity index correction in given later for the case of sidescan bathymetric sonars).

2.4.3 Parameters Determined by the Medium

Transmission Loss

Transmission loss may be considered to be a sum of losses due to both spreading loss (SS) and attenuation loss (AL). Spreading loss (SS) is a geometrical effect representing the regular weakening of a sound signal as it spreads outwards from the source. Attenuation loss (AL) varies linearly with range, and is expressed in decibels per unit distance.

Spreading Many models can be used to simulate the spreading of sound under water (cylindrical spreading, spherical spreading, etc.). Urick [207] considers spherical spreading to be the one that mostly approximates spreading from a sound source underwater. In the simple propagation condition of a sound source emitting in a homogeneous, unbounded, lossless medium, the power generated by the source is radiated equally in all directions. Across concentric spheres of varying radii the power is (sphere area multiplied by intensity):

$$P = 4\pi r_1^2 I_1 = 4\pi r_2^2 I_2 = ... \tag{2.4}$$

considering the first radius r_1 to be of unit length, the transmission loss to range r_2 is then:

$$SS = 10 \log \frac{I_1}{I_2} = 10 \log r_2^2 = 20 \log r_2 \tag{2.5}$$

Absorption of Sound in the Sea Absorption involves the process of conversion of acoustic energy into heat and thus represents a true loss of acoustic energy to the medium in which propagation is taking place. The logarithmic absorption coefficient when a wave travels from r_1 to r_2 can be written as:

$$\alpha = \frac{10 \log I_{1\_} 10 \log I_2}{r_{2\_} r_1} \tag{2.6}$$

Absorption in sea water is much higher than absorption in fresh water (between $5-50$ kHz it is 30 times more) [207]. Absorption is highly dependent on temperature and frequency. However, in the shallow areas considered for this work, the temperature will be considered constant. This work will follow Urick's [207] approximation of absorption loss as the following:

$$AL = \alpha r \tag{2.7}$$

where α is the absorption constant defined as [5]:

$$\alpha = kf^2 \tag{2.8}$$

[5]The absorption constant α depends on the frequency of the sonar. For the 234 kHz sonar it is 0.072 (following Urick [207]), and for the 117 kHz sonar, it is 0.027.

and k is a function of sonar frequency. The total transmission loss due to both absorption and spreading can be written as:

$$TL = SS + AL = 20 \log r + \alpha r \tag{2.9}$$

Ambient Noise

The ambient noise is the residual noise background in the absence of individual identifiable sources that may be considered the natural noise environment of the hydrophone. Measurements have been made over a large frequency range, the data shows that ambient noise has different characteristics at different frequencies, with a different spectral slope and a different behaviour with varying conditions. In deep water, a number of frequency bands in the noise spectrum can be associated with readily identifiable noise sources. In any region of the spectrum, one or more of these noise sources are apt to be dominant over the others [207]. However, in shallow water, the sources of noise are highly variable, both from time to time and from place to place. In deep water areas, some of the noise sources are: -Tides and waves: Causing hydrostatic pressure changes of relatively large amplitudes at low frequencies (usually under 10 Hz). Tidal currents produce changes in temperature and thus affect the functioning of the piezoelectric transducer.

-Seismic disturbances: Because the earth is in a constant state of seismic activity, earth unrest is a source of low frequency noise in the sea (1-10 Hz).

-Oceanic turbulence: It is in the form of irregular water currents of large or small scale; its major effect is that it gives rise to varying dynamic pressures that are picked up by a pressure-sensitive hydrophone located in the turbulent region (its also a low frequency noise).

-Shipping noise: Can exhibit variability in both space and time. The spatial variability is largely governed by the shipping routes in the oceans. Temporal variability can be induced, for example, by the seasonal activity of fishing fleets (usually in the band between 20 and 500Hz).

-Sea Surface noise (also known as 'Knudsen noise'): In the region between 500Hz and

$50kHz$. As discovered by Knudsen in 1948 [83], the noise in this part of the spectrum is correlated with wind speed. This type of noise is also depth dependent. The sound from radiating surface bubbles is modified by inactive older bubbles below the surface and by absorption, so the noise levels droop by a few dB.

-Thermal noise: It places a limit on hydrophone sensitivity at high frequencies (over $100kHz$). The spectrumlevel of this kind of acoustic background in underwater hydrophones is given by Urick [207] to rise with frequency at the rate of 6 dB/octave.

As for shallow water, the noise background at a given frequency is a mixture of three different types of noise: shipping and industrial noise, wind noise and biological noise. At a particular time, the mix of these noise sources will determine the noise level which is highly variable with time and place. For example, the noise sources in bays and harbours are a combination of the sources contributing to deep water noise (listed above), and noises due to industrial activity, marine life and tidal currents. In coastal waters, wind speed is an important source of noise levels over a wide frequency range [207]. Some sources of acoustic noise are of transient occurrence, but their effect can be quite significant. Bioacoustic noises are a good example. Their main contributors include shellfish, fish and marine mammals. Another transient source of noise is rain. The major cause of rain-generated noise is the propagation of bubbles upon water drop impact on the surface[6] [83]. The underwater ambient noise spectrum generated by rain has a unique spectral shape dominated by a broad peak at about $15kHz$. Ambient noise has been found by probability density analyses of data to have a Gaussian amplitude distribution at moderate depths [207]. This is consistent with the fact that the noise originates through many sources of random amplitude and phase. Near the sea surface, however, as with a hydrophone at a depth of a few feet below the surface, ambient noise is more spiky than Gaussian. Although it is Gaussian over short time periods, it is clearly non stationary over longer time periods because of the variability of the sources of noise.

[6]The ambient noise attributable to rainfall has been successfully used in an inverse fashion to provide estimates for oceanic precipitation [174, 158] .

Velocity Change

The velocity of sound in the sea is an oceanographic variable that determines many of the peculiarities of sound transmission in the medium. It varies with depth, seasons, geographic location and time at a fixed location. It depends on three main factors: temperature, salinity and depth. A typical sea profile can be divided into many layers:

- The surface layer: velocity is susceptible to daily and local changes in heating, cooling and wind action. Under prolonged sunny and calm conditions, the mixed layer disappears and it is replaced by water in which temperature decreases with depth.

- Seasonal thermocline: velocity decreases with depth.

- Main thermocline: slightly affected by seasonal changes, velocity decreases with depth.

- Deep isothermal layer: velocity increases with depth because of increase of pressure.

The occurrence and thickness of these layers varies with latitude, season, time of day and meteorological conditions. A shallow water profile is complicated by the effects of salinity changes caused by nearby sources of fresh water. It contains numerous gradient layers of little temporal or spatial stability.

Effects of the Sea Surface

The surface of the sea is both a reflector and scatterer of sound and has a profound effect on propagation in most applications where source and receiver lie at shallow depth. The sea surface affects underwater sound by providing a mechanism for:

1- Forward scattering and reflection loss.

2- Image interference and frequency effects.

3- Attenuation by a layer of bubbles.

4- Noise generation at higher frequencies due to surface weather.

5- Backscattering and surface reverberation.

A criterion for roughness or smoothness of the sea surface is given by the Rayleigh parameter

defined as:

$$R = kH \sin\theta \tag{2.10}$$

where: k is the wave number, H is the rms wave height (crest to trough), θ is the grazing angle. When $R \ll 1$, the surface is primarily a reflector and produces a coherent reflection at the specular angle equal to the angle of incidence. When $R \gg 1$, the surface is considered acoustically rough, acting as a scatterer and sending incoherent energy in all directions. When the surface is in motion, as is always for the surface of the sea, the vertical motion of the waves superposes itself, so to speak, upon the frequency of the sound incident upon it. So the mobile sea surface produces a frequency smearing effect on a constant frequency signal. Although it is small, it is an effect having significance for narrow band acoustic communications. In designing a sonar, the effect of the sea-surface on the incoming signals should be minimal. This is achieved by a careful selection of sonar directivity patterns and sonar beam-width.

2.4.4 Parameters Determined by the Seabed

Backscattering of Sound from the Seabed

An incident acoustic wave on the seabed can be reflected, absorbed and scattered. The relative magnitudes of these effects are dependent on seabed characteristics, seabed roughness, the frequency of the incident wave and the angle of arrival of the acoustic wave. Physically based models for acoustic scattering by the seafloor can be inverted to estimate intrinsic sea floor properties such as acoustic impedance and seafloor roughness. These classification 'clues' will presumably be only weakly dependent upon non intrinsic factors such as propagation, pulse length and transducer directivity and should, therefore, provide a means for seafloor-type recognition.

Backscattering (scattering sound back to a receiver that is at the same location of the transmitter) from the ocean bottom is generally considered to be composed of two principal components, backscatter at the water-sediment interface due to surface roughness (surface

scattering affecting the reflected signal energy) and scattering from inhomogeneities within the sediment volume (volume scattering affecting the refracted signal energy). Backscattering at the water sediment interface, particularly for sandy bottoms, is caused by surface roughness, which is due to particle size, sand ripples, dunes that are mainly hydrodynamic in origin, and features that are caused by marine organisms. Inhomogeneities within the sediment include the effects of buried rocks or shells, marine organisms or their burrows within the sediment, and layering of the subbottom [25].

Many models exist for modelling backscatter from the seafloor. These models either describe surface (roughness) and volume scattering as two different entities or provide a model that combines the effects of both at the same time. Detailed reviews of methods used to describe wave scattering from rough surfaces can be found in Ogilvy [159] and Voronovich [211]. This section is only a very brief summary of some of the methods used for the modelling of sound scattering from underwater seabeds.

Perturbation Approximation for Roughness Scattering This theory is usually used to model scattering from a slightly rough surface. The restrictions on the form of the surface make it feasible to seek a solution to the scattering problem, that is the solution for scattering from a smooth surface together with 'perturbative' terms arising from the slight surface roughness. Quantities that are functions of the surface height may be expanded as Taylor series about their value on the mean scattering surface. The theory can be first, second...order, depending on the terms in the Taylor expansion that are retained [159]. Mourad and Jackson [152], Jackson and Briggs [110], and Jackson *et al.* [111] use the perturbation theory to derive a model for surface backscattering including acoustic properties of the sediment such as the compressional velocity ratio ν, the compressional loss parameter δ and the density ratio ρ. The perturbation approximation is a reasonable approximation providing the rms surface roughness h satisfies [112]:

$$kh < 0.5 \tag{2.11}$$

where k is the acoustic wavenumber in water.

The Kirchoff-Approximation Theory for Roughness Scattering Also a widely used theory to model surface scattering, where any point on the surface is treated as a part of an indefinite plane parallel to the local surface tangent [159]. It is exact for an infinite smooth surface but is approximate for scatterers that are finite sized non planar or rough. It suffers from a few shortcomings : the theory is not self consistent and does not conserve energy, it does not consider the effects of shadowing and multiple scattering. The main condition for applying the Kirchoff theory is the following [171]: The roughness of the surface has to be sufficiently smooth in the horizontal dimension (it is applicable for any height of the surface as long as the characteristic horizontal scale of roughness exceeds the wavelength) in order to avoid shadowing effects and multiple reflections on the surface. This condition is expressed as the following:

$$2k_o r_c \cos(\theta_{inc}) >> 1 \tag{2.12}$$

where r_c is the radius of curvature at a given incidence θ_{inc} and wave number k_o. Thorsos [203] simplifies this condition saying that it is sufficient that the seafloor correlation length is longer than the wavelength λ of the incident sound wave. This hypothesis makes sense in the case of smooth seabottoms (mud, sand ,etc..), and it was experimentally proven by Pouliquen, Bergem and Pace [171].

The Composite Roughness Approximation for Roughness Scattering This approximation begins with perturbation effects and introduces corrections to account for the effect of large scale components of seafloor roughness. Thus, it overcomes many shortcomings of the Kirchoff and the perturbation theories. The large scale surface must have radii of curvature comparable to, or larger than the acoustic wavelength, and the small scale surface must have relief comparable to the wavelength [111]. Conditions for the small scale surface are the same theory conditions whereas the condition for the large scale surface arises in the derivation of

the composite roughness model through application of the Kirchoff approximation to the large scale surface [112].

Jackson *et al.* [112] use the composite roughness model for grazing angles less than 70° (grazing angle is the angle between the sound beam and the seabed), for steep grazing angles the diffraction by the large scale surface becomes important and the composite roughness model is replaced entirely by the Kirchoff model.

However, the composite method has a drawback arising from the fact that the surface roughness parameter is divided into two scales. For the continuous spectrum of roughness this parameter cannot be uniquely determined and its choice may affect the results of the calculation.

Many other models exist for scattering from rough surfaces including the small scale approximation method by Voronovich [211], which is a series expansion based on the small values of surface slopes as a small expansion parameter.

Volume Scattering Scattering by sediment inhomogeneities is recognized as a major contributor to seafloor backscattering. From an acoustic point of view, the sediment is modelled as a fluid. In an inhomogeneous fluid, scattering is caused by fluctuations in both density and sound velocity, so the backscattering is governed by the product of density and sound velocity (acoustic impedance). Jackson [111] and Ivakin [108] use the small perturbation method to model volume backscattering.

The Relationship Between Surface and Volume Scattering The two mechanisms of seabed scattering, surface roughness and volume heterogeneity, can operate at the same time, and thus methods for their separate identification are of practical importance. Ivakin [108] presents a method ('the correlation method') for separating and discriminating volume and surface components of the scattered field. This method involves measurements of spatial coherence of the field with a linear array located near the source. Ivakin [108] also proves that

the correlation scales for the volume and surface components can be significantly different, and thus spatial filtration can be used for separating these components.

The relationship between surface and volume scattering is a complex one affected by many factors, mainly the frequency used, the grazing angle and the sediment type. Experiments are on-going to test the interdependence between surface and volume scattering. In Jackson [110], a few experimental results are discussed. It is deduced that for seabeds with relatively fine sediments (grain sizes less than 0.03 mm), volume scattering is the dominant factor. Whereas for sandy seabeds (mean grain sizes of 0.13 mm), roughness (surface) scattering is the dominant factor (the frequencies used in [110] varied between $15 - 45$ kHz). Sonar frequency is also a major factor to be considered. For high frequencies (> 300 kHz), the absorption of the sound wave is minimal and the effect of volume scattering can be neglected. The nature of the seabed is of importance also. Since some seabed types such as boulders and very course sand prevent signal absorption. Thus, they minimise the effect of volume scattering.

Chapter 3

Selecting Features from Sonar Data for Seabed Classification

The previous chapters introduced the sonar classification problem, and presented an introduction to sonar properties and parameters that affect the received backscattered signals. This chapter explores the features that can summarise sonar data, and indicate variations in seabed types. The parameters that were explained in chapter 2 are used in this chapter to develop a model to correct for depth variations, sonar motion, directivity and sonar beam foot-print. This model is applied to the sidescan bathymetric data that is used in this work. To summarise large sonar datasets, features need to be derived from these corrected signals. The chapter tackles the following questions:

-Is backscattering amplitude enough for classifying sidescan-sonar data? and is an knowledge of the general bathymetry enough to correct for depth effects, and deduce a relationship between backscattering amplitude and sediment type?(section 3.1)

-Can spectral information be of use? and does a scale space approach offer any advantages? (section 3.2)

- Assuming that the features are selected across several scales, what are the parameters (relating to sonar resolution, range, look direction, and normalisation) that affect the features

observed? (section 3.2.1)

3.1 Backscattering Amplitude Features

Plotting the amplitude of corrected backscattered signals versus the grazing angle (the angle between the sonar beam and the tangent to the seafloor) can be used to classify seabed types [3, 148, 147, 36]. However, this approach is dependant on the corrections done to the received sonar signals. Details of the bathymetry should have sufficient resolution for this approach to be applicable. Due to their very good bathymetry resolution, multibeam sonars are used for the development of the SESAM system (by NATO SACLANT) for seabed classification of large areas [36, 172].

The sidescan bathymetric system used in this work was also used by Goodfellow [93] who applied the backscattering amplitude versus grazing angle approach to datasets provided by SEA (Submetrix then). One of his observations was that the resolution of the bathymetric data is not enough to apply an accurate sonar model for signal correction. Although this approach worked for several seabed types surveyed with a particular sonar frequency, he commented that it suffers from many problems. These problems include the effect of noise, sonar calibration and the most serious problem, data clipping. When amplitude data is clipped (usually due to surveying at shallow depths and not correcting for that), it is impossible to tell the difference between seabed types that seem to have the same amplitude (that of the clipping threshold). The next section develops a model that can be used for the correction of sidescan-bathymetric signals. This model is used on one of the sidescan-bathymetric datasets and the results of using backscattering amplitude versus grazing angle are discussed.

3.1.1 A Model to Correct Sidescan Bathymetric Sonar Signals

Although the SEA Submetrix RTS2000 [1] software provides filters for the correction of signals (section C), these filters result in leaving out many points of the bathymetric profile. The distance between the chosen points in RTS2000 could be tens of centimetres while the sonar wavelength is millimetric. Thus, the resolution of the resulting sea bed profile is not sufficient to model the direct effect of the sea bed geometry on the sonar signal. More data points have to be retained to attain a good resolution. Interpolation will be used between chosen points to obtain better amplitude and bathymetry profiles. The aim of this section is to use a new set of corrections on the incoming signals to correct for parameters which are due to sonar geometry, sonar type and the medium. Sonar motion is corrected for in the supplied software (RTS2000), and for each ping the roll, pitch and heave are stored and used for the display of signals.

To describe the effect of the sea bed on the sonar signal, the bottom backscattering strength (BS) is used since it provides information independent of all of the above mentioned factors. The bottom backscattering strength (in dB, and following Urick's equations [207]) is given by:

$$BS = RL - SL + 2TL - 10\log(A) - DI - NL \tag{3.1}$$

* RL is the received level in decibels

* SL is the source level in dB corresponding to 1 m along the transducer's maximum response axis

* TL is the one way transmission loss in dB

* A is the area insonified and $10\log(A)$ represents the losses due to the insonification of different areas

* DI is the directivity index depending on the source and transmit beam patterns and the transducer response for different angles of elevation.

Since the main aim is to correct the form of the received signal for all these effects, and not to exactly measure the acoustic backscatter per grazing angle, the constant factors, such

[1] RTS2000 is the software package provided by SEA for processing the sonar signals.

as the source level SL (which is supposed constant throughout the signal), and the noise levels NL are not actually subtracted from the received raw signal. The factors corrected for are the following:

-The transmission loss, as described in section 2.4.3, the transmission loss factor TL is subtracted from the received signal.

-The losses due to the beam shape and directivity of the sonar (DI).

-The losses due to the areas insonified (A).

Losses Due to Beam Shape and Directivity

The directivity index for the sidescan bathymetric sonar used in this study was provided by the manufacturers. It is dependent on the variation of grazing angle and is modified to account for both azimuthal and elevational directivity, named Rxe and Txe respectively. Rxe and Txe are given to be dependent on a constant k which is related to the grazing angle θ as follows:

$$k = \frac{\theta - \alpha + 5}{10} \tag{3.2}$$

where α is the angle between the axis of the transducer and the vertical (all angles in degrees).

The grazing angle can be easily found from the bathymetric profile of the sea bed, simply by using the depth and slant range:

$$\theta = \arcsin(\frac{\text{depth}}{\text{slant range}}) \tag{3.3}$$

The Expressions for Rxe and Txe as supplied by SEA are:

$$Rxe = 0.0929k^4 - 1.5445k^3 - 7.2351k^2 - 11.24k + 0.7203 \tag{3.4}$$

$$Txe = -5.0121k + 2.0545 \tag{3.5}$$

The directivity index (DI) should also take account of the beam directivity (DI_b) as in the following equation:

$$DI = DI_b + Rxe + Txe \tag{3.6}$$

DI_b is given by:

$$DI_b = 10 \log \frac{4\pi}{\phi\beta} \tag{3.7}$$

where ϕ is the beam elevation, and β is the beam azimuth, so

Interpolation The data received by the SEA sonar systems contains a large number of erroneous points. These points must not be included in sonar corrections. The corrections must also exclude the near-nadir and very far data points. As mentioned in appendix C, the RTS2000 system applies many types of filters on the received samples, to deduce which samples can be flagged as valid depth points. This leads to a large loss of datapoints (approximately 70% of the samples are not used). To solve this problem, the simplified depth profile (i.e. the flagged samples) is used to find the parameters (grazing angle for example), then the graph is interpolated for the whole number of samples being used. The interpolation is a simple linear one and only pings with a sufficient number of samples are used. Figure 3.1 shows the grazing angle for 55 selected points in a ping, then the grazing angle interpolated for 550 useful samples.

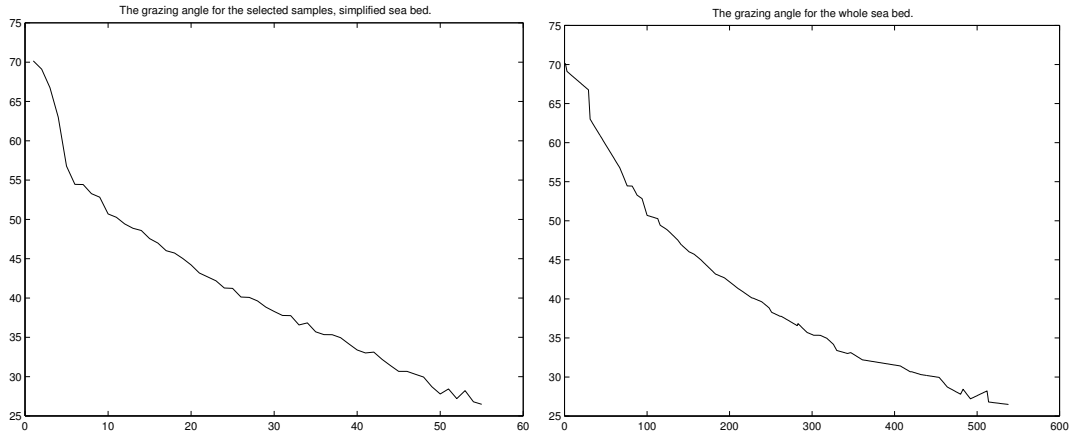


Figure 3.1: The graph on the left shows the grazing angle for the simplified sea-bed (only the flagged datapoints over the whole area), the graph on the right shows the grazing angle versus sample number for the whole region selected. The grazing angle is calculated using equation3.3.

After the calculation of the grazing angle θ, the directivity index can be calculated and

plotted versus the grazing angle, or the sample number as in figure 3.2 (Note that the grazing angle decreases with sample number starting from almost a right angle under the sonar to very small angles at large ranges. This is why the graphs in this figure seem to be 'inverted').

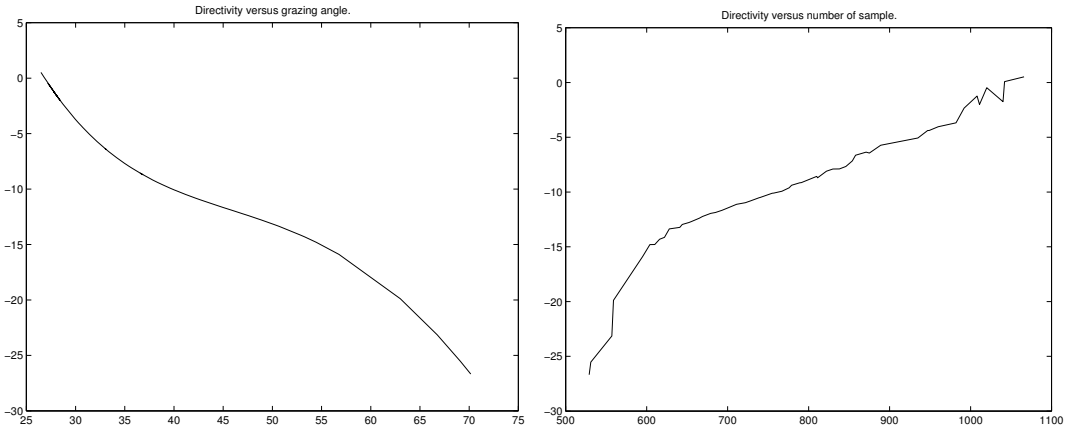


Figure 3.2: The graph on the left shows the directivity losses in dB versus grazing angle. The graph on the right is that of directivity versus sample number (for the range of samples used from this ping). The range of grazing angles selected is from about 25 to 70 degrees to avoid near nadir and far points.

Losses Due to the Insonified Areas

At two different depths, the energy reflected by the insonified areas intercepted by the same beam are different (figure 3.3). To correct for this problem, the received signal has to be corrected for the effect of the insonified area.

The area A is calculated by assuming the insonified region is a rectangle. Its width w is:

$$w = (\text{slant range}) \times \tan\beta \tag{3.8}$$

β is the beam azimuth as in figure 3.4 (given the value of 1° to -3dB points for 234 kHz for example).

The length l of the insonified rectangle is given by:

$$l = \frac{\text{pulse length}}{\cos(\theta)} \tag{3.9}$$

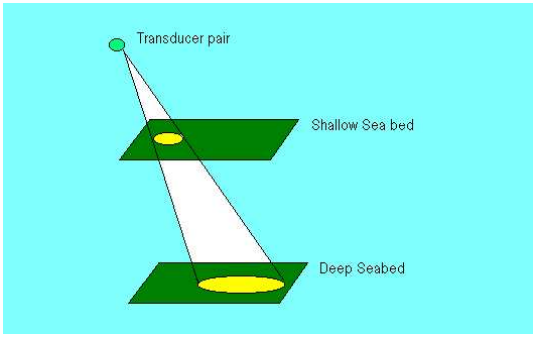


Figure 3.3: Different areas insonified by the beam at different depths.

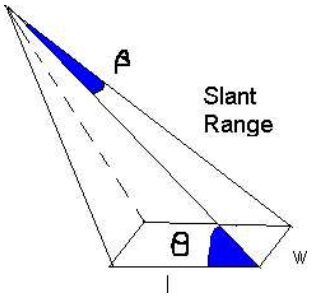


Figure 3.4: The area insonified by the beam.

where θ is the grazing angle, and the pulse length can be obtained directly from the sonar software, by using the following equation:

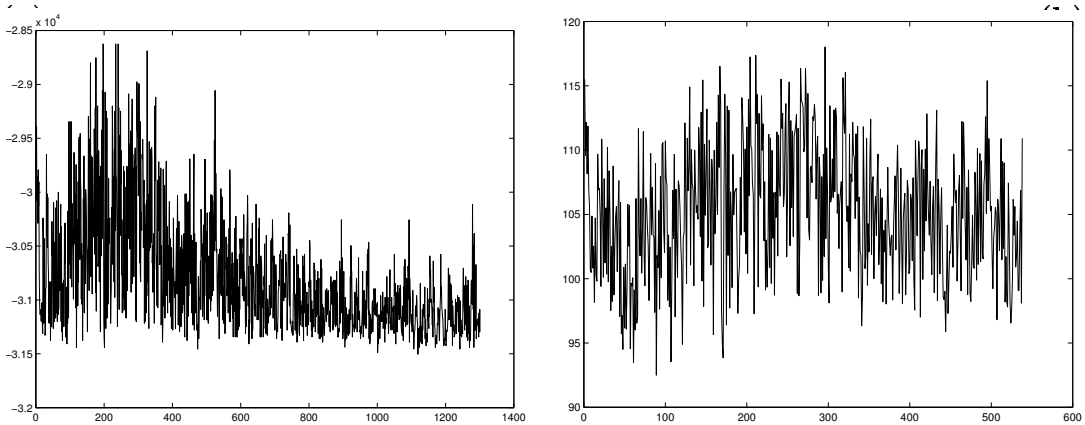
$$\text{pulse length (in meters)}=\text{sos} \times \text{pulse time(in seconds)} \tag{3.10}$$

sos is the speed of sound in water (approximately 1490 m/sec) and the pulse time is usually given by the sonar manufacturers. The area is then assumed to follow a Lambertian variation along the range, so it is multiplied by $sin^2(\theta)$ to correct for the effect of range on the area insonified.

A factor of $10\log(A)$ dB is subtracted from the received signal to correct for area losses.

Example

Figure 3.5 shows a signal from the Hopvågen dataset before and after corrections. The corrections are for area losses, directivity and transmission losses.



(a) **(b)**

Figure 3.5: Figure (a) is an unprocessed signal from an area of mud in the Hopvågen site. (b) is the selected part of the signal (omitting near nadir and far data points) corrected for area losses, directivity and transmission losses.

3.1.2 Results on SEA Data

The model developed in the previous paragraph is used to correct for each ping. The backscattered amplitude values can then be plotted versus grazing angle as in figure 3.6 (left). As noticed in this figure, the values are quite sparse. An averaging approach is used, then a curve is fitted to the resulting averaged signal (right). This is done in order to observe the values of average backscattering amplitude versus grazing angle per ping.

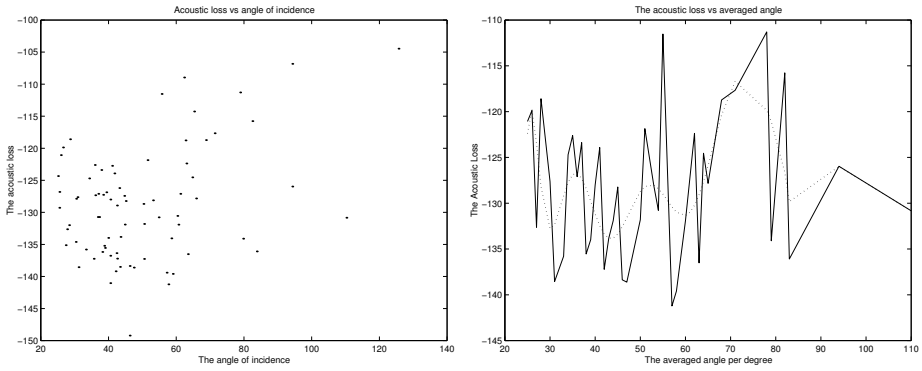


Figure 3.6: The left figure is the unaveraged acoustic loss plotted versus grazing angle, and the right figure is the acoustic loss averaged for each angle (both plotted per ping).

Using one of the supplied SEA datasets (Plymouth sound), pings are selected from known areas of four seabed types. The types are mud, sand, sandy gravel and gravel. To avoid the clipped areas of the pings, data was studied for a range of grazing angles between 30° and 60° which had less clipping than near nadir areas of the pings. The results observed are presented in the following table:

Type	**Total Ping Val.**	**30°- 60°**	**Average Val.**
Mud	-116→-132 dB	-121→-132 dB	-124.6 dB
Sandy Gravel	-125→-152 dB	-130→-145 dB	-140 dB
Gravel	-120→-160 dB	-128→-149 dB	-141dB
Sand	-130→-160 dB	-130→-162 dB	-143 dB

The second column of the table presents the range of backscattering amplitude values for

all grazing angles per ping. The third column is the range of values between 30° and 60° and the last column the average values for the whole ping.

The values for sandy gravel, gravel and sand seem to overlap for the range of angles considered as can be seen in the second and third row. Mud, however, seems to be seperable from the other classes although there is a large overlap. The results are quite similar to Goodfellow's [93] results and suggest the following: Backscattering amplitude, used on its own, is not enough for the classification of sidescan bathymetric datasets using the type of bathymetry provided.

In the sonar literature, there is ample evidence that using amplitude only can, in many cases, lead to misclassification. An example is given in Cochrane [52], where adjacent areas of rocky seafloor could not be differentiated from areas of thin sand on the basis of acoustic backscatter (i.e. grey level) alone (due to using a sonar frequency of 100 kHz which is transmitted through thin layers of sediments and reflected from the underlying rock interface). This is why additional features are needed, such as features that depend on the spatial positioning of sonar backscatter [12].

3.2 The Scale Space Approach

In the previous paragraph, backscattering amplitude alone proved to be insufficient for the classification of the sidescan-bathymetric data considered for this work, even after the correction for depth effects and sonar parameters. Looking at sonar images, a sonar expert generally uses visual textural differences as a method of telling which areas of a sonar image are different. Figure 3.7 displays four different types of seabeds surveyed by the sidescan bathymetric sonar. In this figure, the texture of (a) is of a smoother type than the gravelly area in figure (c). The textural features observed in (a) and (c) are on a much smaller scale than the textural features that can be observed in (b) and (d). To separate these images visually, we are looking at features appearing at different scales. Using a measure that can capture prooperties of different seabed types at many scales could be of use in summarising large sonar surveys, and

classifying different seabeds.

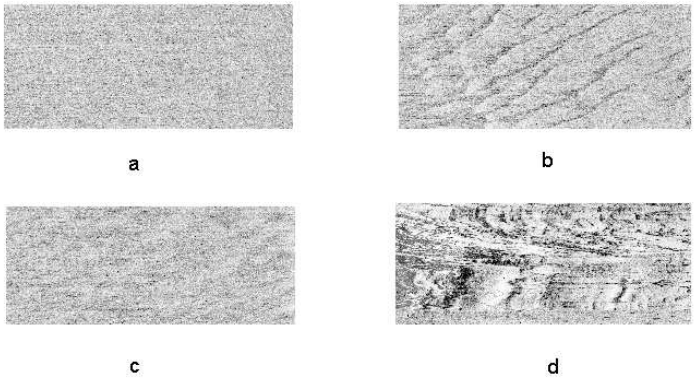


Figure 3.7: In this figure, four samples of areas of different type from one of the SEA datasets are chosen. (a) is an area of a smooth sediment (nearby grabs indicate fine sand), (b) is an area of sand waves, (c) an area of gravel, and (d) an area of coarser bathymetric structure (unknown).

In sonar literature, multiscale features for sonar classification are widely used. Pace and Gao [167] proposed a swathe seabed classification system, where features are obtained from the power spectra of 1-D swathes. Windows are selected over the power spectra to select features occurring at different frequencies. Window selection is done by observation rather than providing an automatic method for selecting window boundaries. Classification accuracies of 97% were obtained on sidescan datasets by using this method. Beck [16] and Edgecock [79] also investigate Power spectrum features. On the other hand, fractals have been successfully used for the classification of sonar data. Carmichael *et al.* [37] use fractal estimates over 3 octaves for the same data used by Pace in [167]. The results are better as more local features are considered (per window) rather than considering the Power spectrum of the whole sonar swathe.

When raw uncalibrated sonar data is considered, [2] the problem can be considered as that of texture-image classification (detailed in chapter 7). Multiscale texture analysis techniques

[2]This is not the approach considered in this work, for reasons explained later and relating to taking sonar direction, and seabed depth into consideration.

are widely used in this area. An early study of wavelet transforms for texture analysis was described by Mallat in [144]. Later Henke-Reed and Cheng [102] applied a wavelet transform to texture images, using the energy ratio between frequency channels as features. Gabor filters have also been used to extract multiscale features as used by Bigun *et al.* [21], du Buf *et al.* [75] and Jain *et al.* [114]. Chang and Kuo [39] developed a tree-structured wavelet transform algorithm for texture classification. Tang and Stewart [201] used multiscale features obtained from wavelet packets and Fourier transforms and proved that these features can contain discriminating information for sonar texture classification. They also highlighted the necessity to develop an automatic approach that can select which scales are most relevant for texture discrimination. GLCMs (explained in chapter 7) are widely used for sonar texture classification. They are calculated over neighbourhoods of pixels and used as features for classification [107]. The window size considered for studying these neighbourhoods of pixels is similar to the concept of scale, and allows viewing features that are relevant over a certain scale [107]. As the literature is very wide, this paragraph only gives a few examples of using multiscale properties for textural discrimination. Chapter 7 goes more into the details of these methods, and discusses their advantages and disadvantages.

The backscattering coefficient received by the sonar depends on both physical and geometrical factors. The dominant physical factors are the impedance and velocity contrast between the sea water and the seabed material, while the geometrical ones concern as well as the grazing angle, the large and small scale roughness of the sea floor. The features observed in sonar images are a function of all these parameters. In Somers [190], it is stated that the vast majority of seabeds, even if composed of microscopic particles, are not acoustically smooth. In the sense that the radius of curvature of the interface is always much larger than the acoustic wavelength. This is the reason why, in this study, seabed sediments (being of smaller particles than the acoustic wavelength) do not appear acoustically smooth. In the previous chapter, the factors affecting seabed backscattering were divided into surface and volume scattering effects. For the high frequency sonars used for this work, and in the areas excluding near nadir areas,

the effect of volume scattering is much smaller than surface scattering [110]. Surface scattering for very complex seabeds as the ones considered in this work, covers a large range, from micro-bathymetric features of the surface to apparent variations of the bathymetry (such as sand waves). This variation of bathymetry is due to many reasons, these include currents and sediment deposits, as well as ship motion (in very shallow areas), and animal activity. This process of surface scattering is the main reason behind the appearance of different textures in received backscattered signals. In his analysis of fine scale seabed morphology on Swathe sonar data, Hughes Clarke [48] notes the following: Both seabed roughness and topography form part of the complete roughness spectrum, so there is no real delineation between these two factors. The type of the sediment is also of great importance, due to the factors that change with each sediment type. The most important ones are acoustic penetration in sediments, the compressional velocity ratio and loss parameter and the density ratio (explained in chapter 2). The change of these properties leads to a change of backscattering strength from sediment to another.

3.2.1 Effect of the Sonar Parameters on the Features Selected

This section will investigate the effect of several sonar parameters on the multiscale features introduced in this chapter. Several sonar studies use 'raw' sonar data to study sonar image texture. This approach, in many cases, neglects the process by which the sonar images are formed and the physical parameters of the problem. This section addresses several aspects of this problem (and their effects on the features observed): the effect of sonar resolution, look direction and range on sonar features.

Effect of Sonar Resolution and Range

The range resolution of a sidescan sonar is related to the bandwidth of the transmitted pulse (varying from 10 cm at 100 kHz to 10 m at 6 kHz). The along track resolution is either the

width of the horizontal beam on the ground or the distance travelled by the transducer during the reception interval, whichever is worse. At close ranges, the latter is worse and the bottom is often under-sampled with the possibility of aliasing or missing features [190]. At far ranges, the resolution degrades. The effect of this is that the picture elements take on a larger aspect ratio parallel to the track. Sonar designers generally design swathe sonars to have the narrowest possible horizontal range of angles, and a large enough vertical range of angles to cover the total swathe to be imaged. In this work, and as explained in section 3.1.1, transmission loss (caused by varying range) is corrected for, assuming a spherical spreading. Very close areas to the sonar and very far areas are not considered for classification. These corrections try to remedy for the variation of sonar resolution due to range, but can not provide a totally exact correction for the effect of range on the features observed. However, on a global scale, the range effects can be compensated for, and produce negligible distortions on a local scale as deduced in several sonar studies [18].

Effect of Look Direction

In several sonar studies, sonar images are used to derive textural features that can describe seabed types. In the majority of these texture based methods, the actual process that results in the generation of the sonar image is ignored. The illuminant direction is fundamental to texture analysis, and its importance has been analysed for the classification of photographic images [40, 135, 1, 209].

This orientation effect has been largely ignored in sonar studies, and its noted as an important factor that can deteriorate the performance of texture based classifiers, as in [194]. Reed [178] states that GLCM features are rotation invariant, but rotates the texture rather than analysing the view-dependent sonar orientation.

Several studies attempted to remedy for this effect. In Bell *et al.* [18], the authors develop a frequency domain model for examining relative orientation effects of the sonar and seabed features. This model is a linearisation of Lambert's law, which the authors use as an approxi-

mation of backscattering from high frequency sonar.[3] The model is tested on a range of towfish heights, pulse lengths and beam patterns. The experiments prove that the sonar image acquisition process acts as a directional filter of seabed texture. This model can be developed to alleviate some of the directional effects by using it to derive a compensating filter (which also depends on some of the survey parameters).[4]. In [77], a Lambertian model is used to reconstruct seabed height information from sidescan sonar data. Features selected from contour map images are then used for classification. This approach can also alleviate the problems caused by using sidescan texture on its own, and neglecting effects of the problem's geometry.

The model explained in section 3.1.1, does not provide corrections for look direction. Hence, features looked at from different directions appear different in a sonar image (sand ripples for example). The approach used in this work to remedy for this problem is to consider features over windows of a certain size. In spite of having an averaging effect and causing a loss of resolution, this method provides a way to extract features that appear at a certain scale. High energy of a window over a certain scale indicates that features of that scale are found in this window and they are selected for classification. So rather than a pixel classification, a window (or patch) classification can be less variable with directionality changes.

When seabed features are of a large scale, such as rocks, acoustic shadows will appear. Since a direct modelling of the effect of bathymetry on sidescan features is not used in this work, areas with acoustic shadows will not be used for classification.

Effect of Beam Pattern and Normalisation

In sidescan bathymetric sonars, and as mentioned previously in this chapter, the sonar beam is designed to have the narrowest possible range of angles and a large enough vertical (or

[3]This approximation has been used in several studies [18, 77, 140], but in general assumes a seabed of relatively smooth bathymetry. More on the demonstration of Lambert's law as a good approximation of high frequency bottom backscattering can be found in [25, 155, 109].

[4]An interesting future direction here would be to combine this model or a similar one with currently used models that correct for sonar signals, such as that in section 3.1.1 (derived from the sonar equation)

across track) range of angles to cover the total swathe to be imaged. The effect of sidelobes on consecutive pings is corrected for online for the sonar used in this work, and the along track ping repetition rate is designed to take care of this problem. As for the across track variation of the intensity values with the grazing angle, it was explained in section 3.1.1. The Directivity index losses versus grazing angle and sample number are calculated for each ping, and depend on the grazing angle between the sonar beam and the seafloor, as explained in figure 3.2.

The variation of the sonar footprint with range and depth is explained also in section 3.1.1, and described as a loss due to the variation of the insonified areas. Correcting for this effect, as well as that of directivity and transmission loss produces an averaging effect on the received signal, as can be observed in figure 3.5. This effect, if observed on a series of received signals will give a more homogeneous processed sidescan image than the raw sidescan data. The drawback of applying several of these corrections is that the final image can not be considered as a raw sonar data image that can be used for a texture study. On the other hand, applying these corrections provides a method of correcting for several survey parameters that are rarely taken into consideration in sonar texture studies, and in fact cause a deterioration of several textural techniques that do not correct for variation due to depth, range and grazing angle variation.

The variation of sonar signal due to sonar motion, which was explained in detail in section 2.4.1 is dealt with by the system used in the surveys provided for this work. The rotational parameters (roll, pitch and yaw) are measured for each ping. The raw sonar signal data and the bathymetric signal finally registered are corrected for the variation of these parameters, by the sonar system, and in real-time.

3.3 Summary and Discussion

In this chapter and the previous one, several parameters that affect the sonar signal are discussed. A model to correct for variations due to slant range, directivity and beam foot-print is discussed. This model uses standard corrections and is described in several studies [207].

As bathymetry is available for the case of sidescan bathymetric sonar datasets provided for this work, it is used to derive the parameters of the correction model for each ping. By using the bathymetry provided, the corrected backscattered amplitude is plotted versus the grazing angle for each point in the ping. For the bathymetric resolution provided by the sidescan bathymetric datasets, our results confirm Goodfellow's results: If the bathymetry provided is not of adequate resolution, it is difficult to use backscattering amplitude only (after using the bathymetry for corrections) as a feature that can discriminate between seabed types. Looking at multiscale textural features is a widely used approach in sonar image classification. This chapter summarises several studies that have used multiscalar features for sonar classification. Parameters that affect how features are viewed are also discussed. A correction for some of these parameters is included in the model used (such as the sonar-beam foot-print). However, several of these parameters are considered a general sonar problem and definitely deteriorate the performance of several classifiers. An example is look-direction, where features viewed at different orientations appear different in the sonar image. The approach that will be considered in the next chapters is that of using small patches (or windows) of the survey for classification. Features are calculated per patch rather than per pixel. In addition to summarising datasets, this approach remedies some of the problems caused by varying look direction and range. Features are selected from the windows after applying the correction model described in this chapter. Thus, 'processed' rather than 'raw' data is used for analysis, as the aim of this work is to benefit from the bathymetric data provided. The next chapter introduces wavelet transforms as a valid method to select multiscalar features, its application to sonar data will be detailed in chapters 6 and 10.

Chapter 4

A Review of the Wavelet Transform and Wavelet Analysis

Wavelets are mathematical functions that separate data into different frequency components, and then define each component with a resolution matched to its scale. They have advantages over traditional Fourier methods in analysing physical situations where the signal contains discontinuities and sharp spikes. Their applications range from detecting and measuring signal variations accurately, to approximation of a signal with fast computational algorithms [219, 55], to compression of signals and images [141], de-noising [54, 73] and filtering signals. This chapter briefly explains the wavelet theory that is used in this research. More details are found in Mallat [142], Daubechies [63] and Chui [47]. Chapters 6 and 7 both use the concepts introduced in this chapter. In chapter 6, Wavelet analysis is used to extract features for the classification of 1-D signals. Chapter 7 extends the ideas presented in this chapter to develop 2-dimensional Wavelet Transforms that can be used for texture classification.

4.1 Time-Frequency Representation

Wavelets were introduced about 50 years ago to fill a gap between two extreme ways of representing a signal : its time expansion which can be seen as an expansion over the Dirac masses:

$$f(t) = \int_R f(u)\delta(t-u)du \tag{4.1}$$

and its frequency representation or Fourier transform which is its expansion over the complex exponentials:

$$f(t) = \int_R \widehat{f}(w)e^{iwt}dw \tag{4.2}$$

The first representation gives an information of maximal time resolution : the value $f(t)$ represents a signal intensity at a time t but contains no frequency information, due to the fact that the Dirac distribution function $\delta(t)$ has an infinitely high spatial resolution but no frequency resolution. The converse is true for complex exponentials e^{iwt}, where frequency information is accurately given without any time information. The question is then to find a way to represent signals between these extremes in which the two types of information are given explicitly. A theoretic limit to this representation is the Heisenberg inequality (uncertainty principle), which states that the energy spread of a function and its Fourier transform cannot be simultaneously arbitrary small. Two approaches have been used to solve this problem:

- Gabor's approach: Gabor defined elementary time-frequency atoms as waveforms that have minimal spread in a time-frequency plane. Gabor atoms are constructed by translating in time and frequency a time window g :

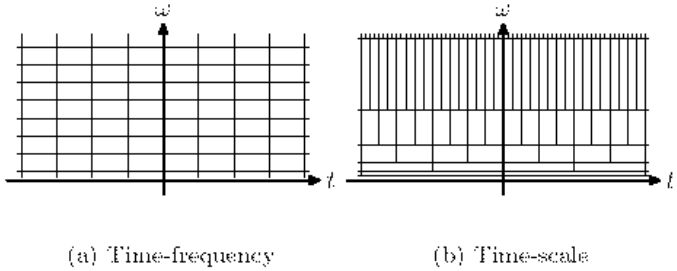
  $$g_{u,\xi} = g(t-u)e^{i\xi t} \tag{4.3}$$

  The energy of $g_{u,\xi}$ is concentrated in the neighbourhood of u over an interval of size σ_t, measured by the standard deviation of $|g|^2$. The energy spread of its Fourier transform $\hat{g}_{u,\xi}$ is therefore localized near the frequency ξ with a frequency spread σ_w. In a time-frequency plane, the energy spread of the atom $g_{u,\xi}$ is represented by a Heisenberg rectangle centred

at (u, ξ) (with a time width σ_t, frequency width σ_w). The uncertainty principle proves that this area satisfies:

$$\sigma_t \sigma_w \geq 1/2 \tag{4.4}$$

This inequality is minimum when g is a Gaussian, in which case the atoms $g_{u,\xi}$ are called Gabor functions. The windowed Fourier transform defined by Gabor correlates a signal f with each atom $g_{u,\xi}$, the representation is a time frequency spectrogram.



(a) Time-frequency (b) Time-scale

Figure 4.1: (a) represents the Gabor time-frequency approach where the Heisenberg rectangles have a fixed size. More resolution is gained in (b) when the basis function is scaled to provide better resolution at high frequencies.

- The 'scale' approach : where the localisation of g can be modified with a scaling. Suppose $g_s(t) = s^{\frac{-1}{2}} g(t/s)$ is our scaled version of g, the Heisenberg time and frequency width of g_s are respectively $s\sigma_t$ and $\frac{\sigma_w}{s}$. The area of the Heisenberg box is not modified but it is dilated by s in time and compressed by s in frequency.

The main differences between the two approaches is that in the Gabor approach, the paving of the time frequency area is that of rectangles that are all translates of the same one, whereas in the 'scale' approach, higher resolution can be obtained at different scales (figure 4.1). It is represented in a time-scale graph called a scalogram. This idea introduces the Wavelet transform.

4.2 The Wavelet Transform

4.2.1 Main Ideas

A wavelet is a function $\psi \in L^2(R)$ with a zero average ($L^2(R)$ is the space of finite energy functions):

$$\int_{-\infty}^{+\infty} \psi(t)dt = 0 \tag{4.5}$$

Which is dilated with a scale parameter s and translated by u:

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}}\psi(\frac{t-u}{s}) \tag{4.6}$$

Its Fourier transform is:

$$\hat{\psi}_{u,s}(w) = e^{-iuw}\psi^*(sw) \tag{4.7}$$

The wavelet transform is an operation that transforms a function by integrating it with modified versions of the kernel function ψ (the mother wavelet). The modifications are translations and compressions of the mother wavelet. For a function to be a mother wavelet, it must be admissible:

$$c_\psi = \int_{-\infty}^{+\infty} \frac{|\psi(w)^2|}{w}dw < \infty \tag{4.8}$$

where c_ψ=admissibility constant to the function ψ(t). Given an admissible function $\psi(t)$, the wavelet transform (WT) $Wf(u,s)$ of a function $f \in L^2(R)$ at time u and scale s can be written as :

$$Wf(u,s) = \int_{-\infty}^{+\infty} f(t)\frac{1}{\sqrt{s}}\psi^*(\frac{t-u}{s})dt \tag{4.9}$$

It could be considered as a convolution with $\bar{\psi}_s(t) = \frac{1}{\sqrt{s}}\psi^*(\frac{-t}{s})$:

$$Wf(u,s) = f * \bar{\psi}_s(u) \tag{4.10}$$

The inverse transform is:

$$f(t) = \frac{1}{C_\psi}\iint_{R^2} Wf(u,s)\frac{1}{\sqrt{s}}\psi(\frac{t-u}{s})\frac{ds}{s^2}du \tag{4.11}$$

When $Wf(u,s)$ is known only for $s < s_o$. To recover f we need a complement of information corresponding to $Wf(u,s)$ for $s > s_o$. This is obtained by introducing a scaling function ϕ (also known as the father wavelet) that is an aggregation of wavelets at scales larger than 1, its Fourier transform is defined by:

$$\left|\hat{\phi}(w)\right|^2 = \int_1^{+\infty} \left|\hat{\psi}(sw)\right|^2 \frac{ds}{s} \tag{4.12}$$

The scaling function $\left(\phi_s(t) = \frac{1}{\sqrt{s}}\phi(\frac{t}{s})\right)$ can be interpreted as the impulse response of a low pass filter. By convoluting f with the scaling function, its low frequency approximation at a scale s can be obtained:

$$Lf(u,s) = \left\langle f(t), \frac{1}{\sqrt{s}}\phi(\frac{t-u}{s}) \right\rangle = f * \bar{\phi}_s(u) \tag{4.13}$$

$\langle f, g \rangle$ represents the inner product of two functions given by :$\int_{-\infty}^{+\infty} f(t)g^*(t)dt$.

So the wavelet extracts the detail of the signal at specific scales and the scaling function gives the low frequency approximation of the signal at that scale.

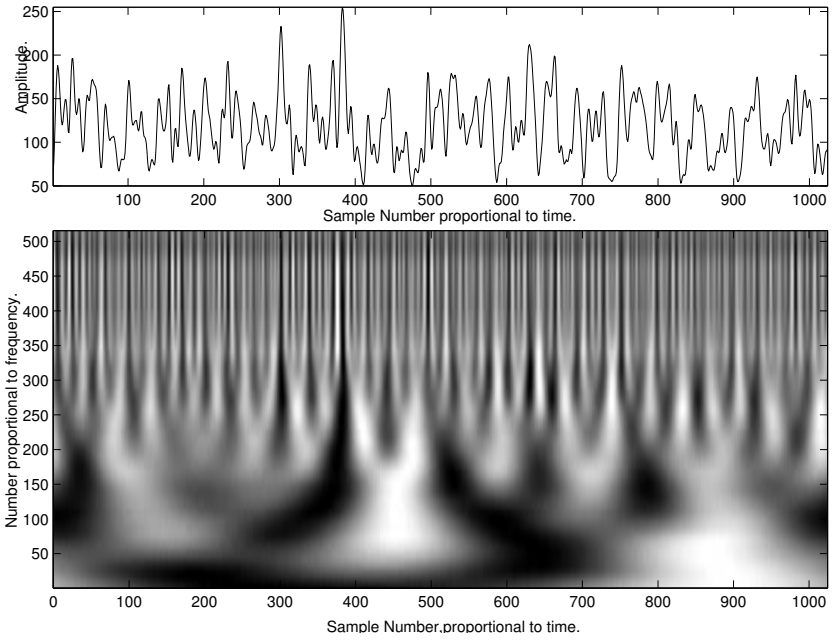


Figure 4.2: An example of the continuous wavelet transform of a signal, the x-axis is the number of the samples (proportional to time), the y-axis is the scale.

4.2.2 Construction and Calculation of the Wavelet Transform

Mallat [142] draws on the techniques of subband coding to describe a method that allows the design of a certain wavelet for a particular application.

Starting with an approximation of a function f (f is the signal under study), θ a unit resolution function that best approximates f is chosen. The approximation of the signal f is defined to be a multiresolution approximation across an embedded grid of approximations. The approximation at a certain scale is the orthogonal projection of f on a space V_j. The orthogonal projection of a function f is the function $f_j \in V_j$ that minimises $\|f - f_j\|$. A sequence $\{V_j\}_{j \in Z}$ of subspaces $L^2(R)$ is introduced.

To compute the projection of f on each subspace V_j, a basis for V_j is required. The unit resolution function θ can be used to deduce the scaling function ϕ. This function is dilated and translated to construct an orthogonal basis of each space V_j. The approximation of f is a low pass filtered version of $f(t)$ obtained by correlating f with ϕ.

From the scaling function ϕ, the low pass filter $h(t)$ is derived (the conjugate mirror filter), from which a complementary high pass filter $g(t)$ can be deduced. Figure 4.3 presents the filters in the frequency domain, the low pass $h(w)$ and high pass $g(w)$ which has a cutoff identical to h.

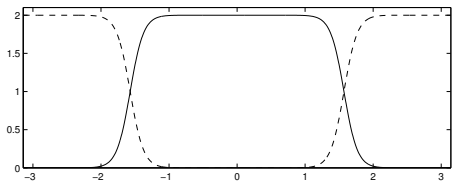


Figure 4.3: The Conjugate mirror filters g and h in the frequency domain. The solid curve gives the square of $h(w)$ on $[-\pi, \pi]$, for a cubic spline multiresolution, the dotted line is the square of $g(w)$.

The wavelet ψ is derived from the scaling function ϕ and the conjugate mirror filters h and g.

In practise, the filters h and g are used for a fast computation of the wavelet coefficients

of a signal $f(t)$. The approach decomposes successively each approximation $P_{V_j}f$ [1] into a coarser approximation $P_{V_{j+1}}f$ plus the wavelet coefficients carried by $P_{W_{j+1}}f$. Since $\{\phi_{j,n}\}_{n\in Z}$ and $\{\psi_{j,n}\}_{n\in Z}$ are orthogonal bases of V_j(the 'low frequency' approximation space) and W_j (the 'detail' space) respectively, the projection in these spaces is characterised by :

$$a_j[n] = \langle f, \phi_{j,n}\rangle \text{ and } d_j[n] = \langle f, \psi_{j,n}\rangle \tag{4.14}$$

These coefficients are calculated with a cascade of discrete filterings with h and g followed by a factor 2 subsampling, giving the signals a_j which are approximations of the original signal at each scale, and d_j the detail signals made explicit at that scale. Inversely, a fast inverse wavelet transform reconstructs progressively each a_j by inserting zeros between samples a_{j+1} and d_{j+1}, filtering and adding the output (refer to figure 4.4).

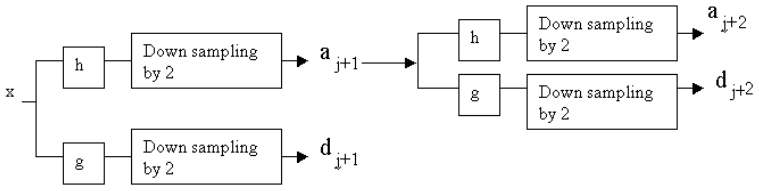


Figure 4.4: A fast wavelet transform is computed with a cascade of filtering with h and g followed by a factor 2 subsampling.

The Discrete Wavelet Transform (DWT)

Both scale and time can be discretised. in the calculation of the wavelet transform, leading to the Discrete Wavelet Transform (DWT). If $f(t)$ is a function uniformly sampled at intervals N^{-1} over [0,1], its wavelet transform can only be calculated at scales $N^{-1} < s < 1$. The discrete wavelet transform (DWT) is computed at scales $s = a^j$ with $a = 2^{\frac{1}{v}}$, which provides v intermediate scales in each octave $[2^j, 2^{j+1})$. If $\psi(t)$ is a wavelet whose support is included

[1]The Orthogonal projection of f on V_j

in $[-K/2, K/2]$ for all $2 \leq a^j \leq NK^{-1}$, a discrete wavelet scaled by a^j is defined by $\psi_j[n] = \frac{1}{\sqrt{a^j}}\psi(\frac{n}{a^j})$. To avoid border problems, $f[n]$ and $\psi_j[n]$ are treated as periodic signals of period N, the DWT can be then written as a circular convolution $\bar{\psi}_j[n] = \psi_j^*[-n]$:

$$Wf[n, a^j] = \sum_{m=0}^{N-1} f[m]\psi_j^*[m-n] = f * \bar{\psi}_j[n] \tag{4.15}$$

This circular convolution is calculated using the FFT (fast Fourier transform algorithm) which requires $O(Nlog_2N)$ operations, thus making the calculation of the WT a fast process.

Translation Invariance and The Dyadic Wavelet Transform

In pattern recognition, it is important to construct signal representations that are translation invariant: when a pattern is translated, its numerical descriptors should be translated and not modified. The Continuous wavelet transform (CWT) and the windowed Fourier transform provide translation invariant representations. For the CWT:

$$Wf_\tau(u, s) = f_\tau * \bar{\psi}_s(u) = Wf(u - \tau, s) \tag{4.16}$$

However, if we uniformly sample the translation parameter u the translation invariance might be destroyed. One way around this problem is to select a sampling interval $a^j u_o$ that is small enough so that the samples of $f * \bar{\psi}_{a^j}(t)$ are approximately translated when f is shifted.

To construct a translation invariant wavelet representation, the scale is discretised but not the translation parameter u. The scale can be sampled along a dyadic sequence $\{2^j\}_{j\in Z}$ to simplify the numerical calculations, so the dyadic WT of $f \in L^2(R)$ is defined by:

$$Wf_\tau(u, 2^j) = \int_{-\infty}^{+\infty} f(t)\frac{1}{\sqrt{2^j}}\psi(\frac{t-u}{2^j})dt = f * \bar{\psi}_{2^j}(u) \tag{4.17}$$

$$\text{with : } \bar{\psi}_{2^j}(t) = \psi_{2^j}(-t) = \frac{1}{\sqrt{2^j}}\psi(\frac{-t}{2^j}) \tag{4.18}$$

In many applications, sampling the scale dyadically does not provide a perfect reconstruction of the starting signal. In dyadic scales, some information could be missed as it would be lying

between the chosen scales. An option is to adapt the scaling to the main features of the signal, or to choose a finer scaling parameter that can pick up main signal features. This will be observed in chapter 7 for 2-D image texture analysis.

Mallat [142] describes a method for constructing a fast dyadic wavelet transform by using filter-banks and inserting zeros between signal samples. This algorithm is called *algorithme à trous* and is implemented in the Wavelab[2] package used in this work.

4.3 Wavelet Bases

$B = \{g_n\}_{n \in Z}$ is the basis for a class of functions if every function in the class can be expressed as a sum of the functions in B :

$$f = \sum_n \alpha_n g_n \text{ where } \alpha_n = \langle f, g_i \rangle \tag{4.19}$$

Suppose $B = \{g_n\}_{n \in Z}$ is an orthonormal function basis (i.e. $\langle g_i, g_j \rangle \geq 0$, $i \neq j$ and $\|g_i\| = 1$), a linear approximation of a signal f over the first M functions of the basis is given by:

$$f_M = \sum_{n=0}^{M-1} \langle f, g_n \rangle \, g_n \tag{4.20}$$

The approximation error $\varepsilon[M]$ is given by the sum of the remaining squared inner products:

$$\varepsilon[M] = \|f - f_M\|^2 = \sum_{n=M}^{\infty} |\langle f, g_n \rangle|^2 \tag{4.21}$$

Considering the Fourier domain, the Fourier basis is an efficient approximation of uniformly smooth signal with an approximation over the M lower sinusoids.

In Wavelet bases, the signal is projected over the M larger scale wavelets, equivalent to approximating the signal at a fixed resolution. The function f_M can be improved by the choice of the M functions ψ_n depending on many factors. The next section explores the elements that are usually taken into consideration in the selection of wavelet-type.

[2]Wavelab is a library of functions designed to implement a variety of computational algorithms related to wavelet analysis. It can be freely downloaded from the statistics department website in Stanford university. http:/www-stat.stanford.edu/~wavelab/

4.3.1 Selection of Wavelets

The selection of the mother wavelet ψ depends very much on the nature of the signals analysed and the goal of signal processing. Some wavelets are better than others for specific applications. A wavelet chosen for compression of signals might be a different one from that chosen for singularity detection. Factors to be considered in wavelet selection include the number of vanishing moments, the compact support of the wavelet and its regularity.

The vanishing moment property makes it possible to analyse the local regularity of the signals. A signal is regular if it can be locally approximated by a polynomial [142]. Mallat presents a condition that relates the pointwise regularity of a signal to the decay of its WT modulus. If f is regular and ψ has enough regular moments, the wavelet coefficients $|\langle f, \psi_{j,n}\rangle|$ are small at fine scales 2^j. In other words, the higher the degrees of vanishing moments a basis has, the better it models the smooth part of the signal [82].

The compact support of the wavelet is also a factor taken into consideration. To minimise the number of high amplitude coefficients, the support size of ψ must be reduced.

The regularity of the wavelet influences the error introduced by thresholding or quantising the wavelet coefficients. The low regularity of the wavelet might result in fractal-like shapes in the reconstructed signals or images.

Some examples of mother wavelets are: (shown in figure 4.5):

- The Haar wavelet, which is the wavelet with the shortest support. It is not well adapted to approximate smooth functions.
- The D4 Daubechies wavelet, Daubechies wavelets have the property of minimum-size support for any given number of vanishing moments.
- The Coiflet wavelet which has a large number of moments equal to zero because of its sharp ridge pulse. It is often used to detect transients in signals.

- Symmlet is a symmetric version of the Daubechies wavelet family [29] (used in many studies to analyse vibrational spectra).

More wavelet types can be found in Mallat [142], Daubechies [63] and Chui [47]. The selection of wavelet type for this work is done experimentally as several wavelet types are tested on sonar signals in chapter 6, section 6.2. The wavelet type that provides the best separation between classes as tested by different classifiers is chosen.

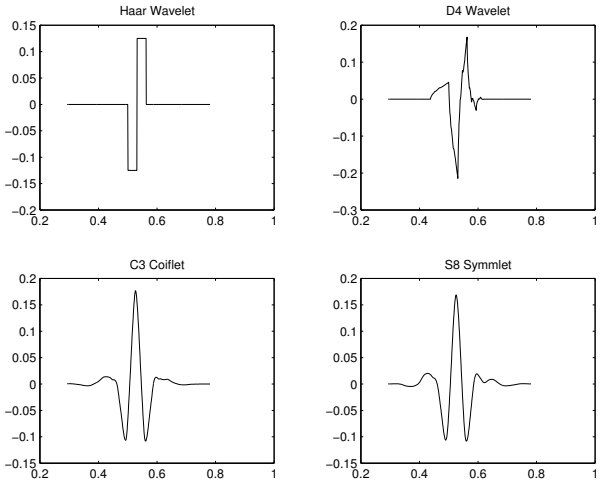


Figure 4.5: This figure presents some of the commonly used wavelets. The Haar, Daubechies, Coiflet and Symmlet wavelets.

4.3.2 Examples of Wavelet Bases

One of the bases used in many studies is the wavelet packet transform (WPT), which is a generalised version of the Wavelet Transform that permits an adaptive time frequency tiling. It retains both the low and the high frequency subbands, performing a decomposition upon both at each stage. As a result, the tiling of the time frequency plane is configurable and the partitioning of the frequency axis may take many forms to suit the need of the application. Starting with a signal x of length N samples, the first level of decomposition generates the low pass and high pass subbands (hx and gx respectively), the second level generates four

subsequences h^2x, ghx, hgx, g^2x. The process is repeated j times resulting in jN coefficients. This iterative process generates a binary wavelet packet tree structure where the nodes of the

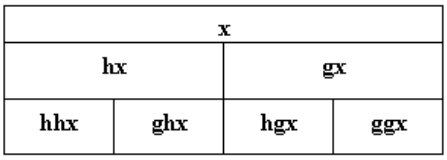


Figure 4.6: The first two levels of decomposition in a Wavelet Packet Transform. The low pass operations (h) to the left and the highpass operations (g) to the right.

tree represent subspaces with different frequency localisation characteristics.

Another basis that can be constructed is the Cosine Packet Transform (CPT), which can be regarded as the conjugate operation of the WPT. Local trigonometric functions partition the time axis, and yield an orthonormal basis on each temporal interval. The subbands in the CPT are instead *intervals* since each subspace in the decomposition is partitioned with respect to time rather than frequency. More information on the CPT can be found in [10]. A disadvantage of the CPT over the WPT is that it is more computationally complex and requires a larger number of calculations [82].

Both Wavelet packets and Cosine Packet transforms have a set of redundant subspaces in the binary tree. The next section explains a few ideas on the selection of a 'best' basis from a redundant set.

4.3.3 Wavelet Bases Selection

An efficient way of representing a signal should give large magnitudes along a few axes and negligible magnitudes along most axes when the signal is expanded into the associated basis. To determine the best basis, it is necessary to evaluate and compare the efficacy of many bases chosen from the whole set of redundant bases. The best basis chosen for representing a signal

could not be the ideal basis for discriminating between the signal and another. Hence, the separation of basis selection into two categories:

Signal Compression

The best basis selection algorithm, generally used for signal compression, operates on a single signal and its binary packet tree decomposition along an orthonormal basis. It was proposed by Coifman and Wickerhauser [219, 55] as a divide and conquer search of the binary tree (pruning algorithm). It uses the fully decomposed tree, starts at the lowest level and eliminates branches until an optimal solution is found. The cost function is based on entropy (explained in section 5.3), since the goal is the maximisation of information with respect to the given signal. The reader is referred to [218, 143] for more detail.

Classification

In classification, the main aim is to observe differences between classes rather than to capture the main features of signals. Saito [182] uses relative entropy between classes as a discriminant measure. The basis which maximises the discriminant measure on time-frequency energy distributions of the different classes is chosen as the best basis. This method is known as the Local discriminant basis (LDB) and described in [183, 182].

This thesis explores the idea of relative entropy (section 5.3) as a measure of class discrimination between wavelet signal decompositions, although a less redundant basis than the WPT is chosen. The discriminant measures are used with a 1-D wavelet decomposition of signals from different classes to indicate which subbands provide the best class distance, as will be explained in chapter 6.

4.4 Summary

This chapter has provided the background necessary to develop wavelet analysis as a means for extracting features for signal classification. The chapter starts with an explanation of time frequency tiling and the need for a basis to do so; the chosen basis should be able to adapt to different frequencies in the signal. Wavelets are introduced as a means to do so. The mathematics and concepts of wavelet theory are explained briefly. The construction of a fast Wavelet transform and the discretisation of both time and scale provide means for efficient calculation of the wavelet transform. Using wavelets as bases for signal decomposition is possible because they provide perfect reconstruction of signals. The selection of wavelets for different applications is discussed, and some examples of mother wavelets are given. The idea of bases constructed from the decomposition of a certain signal is developed, to include more redundant decompositions such as the Wavelet Packet transform. The selection of a 'best basis' for signal compression and classification are considered separately. Relative entropy is used as a measure of the separation between the wavelet decomposition of signals belonging to different classes, and the best basis includes bands that provide maximum class separation.

Chapter 5

Methods for Feature Analysis

5.1 Background

The improvements in processing and storage of data have led researchers to look for new approaches in representing data and selecting the important information from it. This work deals with sonar data which is a good example of data that is easily stored in large quantities. Methods of summarising relevant information from sonar data are investigated in this work. The aim is to reduce the dimensionality of data, keeping the most relevant information from it. Thus, avoiding the 'curse of dimensionality' since increasing the dimensionality of feature space rapidly leads to very sparse data that provides very poor representation of the important features [22, 19].

This chapter summarises steps that are needed for efficient classification of sonar data following feature extraction [1]. Feature extraction is detailed in the next few chapters. In this chapter, however, the problem that is considered is how to deal with data-features after the extraction stage, going through the following steps:

[1] What is meant by feature extraction is the generation of features from information. Many authors use it to describe the selection of relevant features from a certain number of features. This process is described as feature selection in this work rather than extraction. So the two terms describe two different processes.

- Feature Representation and Visualisation, especially of multidimensional data. Methods used include Sammon's mapping and Principal Component Analysis (PCA).
- Feature Selection: A very important step to avoid data redundancy since some features are more relevant than others for classification.
- Classification: It can either be supervised or unsupervised. Methods used for each case will be explained briefly.

5.2 Feature Representation and Visualisation

This step serves as an initial step to estimate which features are relevant for an efficient classification. Selecting a few features from a large number and plotting them in a scatter plot of 2-D or 3-D can only give an idea of class separation if the other features contribute much less to classification. In the case of a large number of features, feature selection only for the purpose of visualisation becomes a very demanding process. Mapping the features into a lower dimension observable by a user can aid in testing the features extraction method. The goal in dimensionality reduction is to preserve as much of the relevant information as possible. Typically, data mapping is unsupervised, although it is possible to perform supervised visualisation. At least as many as eight major families of mapping techniques can be distinguished [188]. However, they are data dependant and depend on why the reduction is required, (for example: visualisation, classification, merging...). This section will only cover the two types of data mapping used in this work and the reader is referred to [188] and [22] for more details on other data mapping techniques. The two types considered are Sammon's mapping and Principal Component Analysis (PCA).

5.2.1 Principal Component Analysis

Although this section describes PCA as a means to visualise multidimensional data, it is an efficient generic method to summarise data and will be also used in the feature selection stage 5.3. PCA offers an option of mapping a dataset of large dimension (M) into a smaller dimension dataset (d), keeping as much information as possible about the initial dataset. The best approximation of the initial dataset is defined to be that which minimises the sum of squares of the errors over the whole dataset. The sum of squares error is considered between the original vectors and the approximation vectors. It is shown in [22] that the minimum occurs when the $M - d$ smallest eigenvalues are chosen and their corresponding eigenvectors are discarded. The problem with choosing only d eigenvalues as the relevant ones is that some information is discarded. The selection of vectors from PCA will be discussed in section 5.3 in the process of feature selection.

In practice, the algorithm proceeds by first computing the mean of vectors x^n and then subtracting off this mean. Then the covariance matrix is calculated and its eigenvalues and eigenvectors are found. The eigenvectors corresponding to the d largest eigenvalues are retained and the input vectors x^n are projected onto the eigenvectors to give the components of the transformed vectors z^n in the d dimensional space.

Principal component analysis is discussed at length in Jollife [115], and is also known as the Karhunen-Loéve transformation. The drawbacks of PCA are mainly that it assumes the data is multivariate normal so that the variance matrix is a sufficient statistic. In addition to that, PCA could miss the nonlinear structure of the data, and is affected by outliers.

5.2.2 Sammon's Mapping

The approximate structure preservation in this method is maintained by fitting N points in the lower dimensional space d such that their interpoint distances approximate the corresponding distances in the M-space [184].

Suppose that we have N vectors in an M-space designated $x_i, i = 1, ...N$ and corresponding to these we define N vectors in a d-space (d is usually considered to be 2 or 3) designated by $y_i, i = 1, ...N$. Let the distance (any distance measure could be used, but normally the Euclidian distance is used) between the vectors x_i and x_j in the L-space be defined by $d^*_{ij} \equiv dist[x_i, x_j]$ and the distance between the corresponding vectors y_i and y_j in the d-space be defined as $d_{ij} \equiv dist[y_i, y_j]$.

An initial d-space configuration is assumed (a random one in this case although it would also be a good idea if principal component analysis was used to save computing time). The d-space distances are calculated and used to define an error E which represents how well the present configuration of N points in d-space fits the N points in the L space:

$$E = \frac{1}{\sum_{i<j}[d^*_{ij}]} \sum_{i<j}^{M} \frac{[d^*_{ij} - d_{ij}]^2}{d^*_{ij}} \tag{5.1}$$

The next step is to adjust the d-space configuration so as to decrease the error. This is done by iterating until a minimum in error is reached.

Sammon's nonlinear mapping belongs to the nonanalytical type of transformations. It creates images only for the points from the data ensemble. To add a new point to the data set, one has to recompute the whole transformation. Therefore, it is suitable for the multidimensional structure of data-preservation in low dimensional spaces rather than for classifier design [78]. Hence, Sammon's mapping is used in this study for data visualisation rather than classification.

5.3 Feature Selection

Reducing the number of irrelevant/redundant features in a feature set both reduces the running time of a learning algorithm and improves learning. This could be very significant for many applications. This work deals with classifying sonar data, so feature selection is done with the aim of choosing the best features that provide class discrimination as well as a good approximation of the original class distribution. Reduction of dimensionality and discarding irrelevant

features lead in many cases to better classification results. There are two approaches to feature selection - variable transformations and variable selection.

Variable transformation includes methods like PCA (5.2.1), where a number of features is transformed to a lower dimensionality space. The approach will be used in this work for reducing feature number. To choose the number of principal components to keep, the first eigenvalues that contain more than 90% of the variation are chosen. The values of the eigenvalues (λ_i) corresponding to each principal component can be used to measure what percentage of the total variation (p) is contained in the chosen principal component (among a total of N principal components):

$$p = \frac{\lambda_i}{\sum_{i=1}^{N} \lambda_i} \tag{5.2}$$

Starting with a feature set of size N, if an optimal subset of features is required without knowing how many are relevant, the total number of subsets to be explored is 2^N. When N is large, the exhaustive approach is out of question. Therefore, various feature selection methods have been designed to avoid exhaustive search while aiming at the optimal subset.

Variable selection techniques can be split into two categories - filter methods and wrapper methods. Filter methods determine whether features are discriminative using heuristics based on characteristics of the data. Wrapper methods make use of the classification algorithm that will ultimately be applied to the data in order to evaluate the discriminative power of features. Wrapper methods generally result in better performance than filter methods because the feature selection process is optimised for the classification algorithm to be used. However, they are generally too computationally expensive to be used if the number of features is large because each feature set considered must be evaluated with the trained classifier. For this reason, wrapper methods will not be considered in this study. Filter methods are much faster than wrapper methods and therefore are better suited to high dimensional data sets. Dash *et al.* [62] recognise four steps in a typical feature selection method:

1. The choice of an evaluation function to evaluate the subset under examination;

2. a generation procedure to generate the next candidate subset;

3. a stopping criterion to decide when to stop, and

4. a validation procedure to check whether the subset is valid.

These steps will be detailed in the next subsections.

5.3.1 Evaluation Functions

The selection criteria that should ideally be considered for classification problems is the probability of misclassification. This could in principle be calculated by using either parametric or non-parametric techniques to estimate the posterior probabilities for each class [97]. Evaluation of this criterion directly is generally too complex, and the standard approach is to consider class separability instead [22]. A set of variables in which the classes are well separated are expected to be a good set of variables to use as inputs to a classifier. Many measures can be used to test class separability. The ones used in this work are (considered between each two feature vectors f_k and f_l)

1-The Euclidian distance:

$$J_{k,l} = \sum_{i=1}^{n} (f_k(i) - f_l(i))^2 \tag{5.3}$$

2-The Fisher criterion or distance:

$$J_{k,l} = \frac{(mean(f_k) - mean(f_l))^2}{std(f_k)^2 + std(f_l)^2} \tag{5.4}$$

The Fisher distance solves the problem of high class asymmetry that is not taken into consideration in simple Euclidian measures. The Fisher criterion is derived from Fisher's Linear discriminant ([22], [86]) which is basically an optimal choice of direction for projection of data to a lower dimension. It is not a discriminant itself, but can be easily used to construct a discriminant.

3-Cross Entropy (also known as the relative entropy, Kullback-Liebler distance, or I divergence):

$$D_{k,l} \approxeq \sum_{i=1}^{n} f_k(i) \log \frac{f_k(i)}{f_l(i)} \tag{5.5}$$

To obtain a more symmetric measure of cross entropy, the entropy formula is modified [182]:

$$J_{k,l} \approxeq D(f_k, f_l) + D(f_l, f_k) \tag{5.6}$$

Cross entropy is a useful measure used in information theory to measure the similarity of information in sequences. It can be used for signal analysis, to observe how similar two signals are. Saito uses cross entropy for a selection of a best basis that can provide the best discrimination between classes ([182], [183]). In this study, cross entropy is used in feature space to detect which features provide the best discrimination between classes.

For the total number of features, the distance between each 2 types (k, l) is measured, and then the mean distance between all pairs of features is used as a class separability measure:

$$D = mean(\sum_{k,l=1}^{t} J_{k,l}) \tag{5.7}$$

If the full criterion of misclassification rate was to be used, the general performance of the system should improve as the number of features is reduced (reducing the number of the redundant ones) until an optimal subset of features is reached. If fewer features are retained, the performance will degrade. The problem with many selection criteria such as most criteria based on class separability is the incapability of modelling this phenomenon. As extra variables are added, the evaluation function J always increases. This is known as the monotonocity property. If both A and B are sets of features with the size of $A >$ size of B then

$$J(A) \geq J(B) \tag{5.8}$$

This property prevents evaluation functions from being used to compare the discriminative abilities of sets with different sizes since equation 5.8 means that deleting features cannot reduce the error rate. In this work, evaluation functions will be used to compare feature subsets with the same number of elements.

5.3.2 Generation of Feature Subsets

The reader is referred to [22] and [62] for a review of feature subset selection. This section will only describe sequential techniques used in this work.

The simplest method to select d out of M features, is to select the features which are individually the best. This is decided by evaluating the selection criterion on each one alone. This method is only truly valid in the case of independent features, and for a selection criterion that can be expressed as a function of the selection criteria per feature.

Sequential Forward Selection (FS) is a more general approach that can be used. It starts with each feature on its own and selects the one that gives the largest value for the selection criterion. One feature is added at each stage, on condition that it gives the largest increase in the value of the selection criterion. The drawback of this method is that there are cases where two features give good discrimination together, but poor discrimination individually. In such cases, the FS method might never find this combination.

Sequential Backward Elimination (BE) on the other hand starts with the whole set of features. In each stage the feature which gives the least reduction in the selection criterion is deleted from the set. This overcomes the problem of the FS mentioned above, but it is still not guaranteed to be optimal. BE requires more operations than FS, since the number of features considered is more than or equal to d whereas the FS algorithm considers features less than or equal to d.

There are many variations of these algorithms, an example includes adding h features at a certain stage using the FS algorithm, and deleting others using the BE.

5.3.3 Stopping Criterion and Validation

The number of features finally selected as input to a classifier should lead to a better performance than using the whole set of features as input. If performance degrades, it could mean that the number of features selected is too low and cannot contain all the relevant information.

In this case, the number of features could be increased. Validation is usually done by comparing the final results of the classifier for different feature subsets, and verifying that the selected subset gives the best classification rates under these conditions.

5.4 Classification

The final stage in designing an efficient classification scheme is the choice of classifier. The choice is dependent on the type of data, the availability of training data, and the type of classification needed. If a training set is available or alternatively parts of the data with an apriori known type, a classifier can be trained using this set. The classifier can be used to classify other 'testing' data, or the rest of the unknown dataset in a *supervised classification.* If known data is not available, the most common option is to observe clusters in feature space in an *unsupervised classification.*

Supervised Classification can be thought of as inferring regularities from a set of training examples, then using the learnt infornmation to classify the 'test' data set. Much research has been devoted to the study of various learning algorithms which allow the extraction of the underlying regularities of the data. If the learning is successful, the intrinsic regularities of the data can be captured in the values of some parameters of the learning machine. For a polynomial classifier, the parameters learnt are the coefficients of the polynomial. For a logistic regression classifier, the parameters learnt are the weights that linearly map the inputs to the output function (a good summary is found in [22], chapter 3). For a linear parameter model (also known as generalised linear model GLM), nonlinear functions of the input are mapped to the output function, and the weights are learnt in the training stage. The next section explains the design of MLP (Multilayer perceptrons) that also learn network weights in the training stage.

The realisation that the higher the dimension of the space is, the easier it is to find a hyperplane that linearly separates the data forms the basis for Support Vector Machine (SVM)

methods. The main idea is to map each vector into a much higher dimensional space where the data can be linearly separated. Training points which do not affect the boundary can then be discarded. Extensive literature on this method can be found in http://www.kernel-machines.org, a few references are the following: [33, 133, 189, 138, 57].

For a radial basis function classification (RBF), a set of basis functions is introduced. The activation of these hidden units is determined by the distance between the input vector and the prototype vector. The output is a linear combination of the basis functions of the networks, and the weights of this linear combination are learnt in the training stage [22, 163, 164, 165].

For a high dimensional space, there is a risk of overfitting as a large number of basis functions is normally used, each one with an associated parameter. This motivated a Bayesian treatment of the problem, where priors can be put on the output functions themselves. One advantage of function space priors is that they can impose a general smoothness constraint without being tied to a limited number of basis functions. This motivated the Gaussian Process 'GP' method, where a Gaussian Process is defined over y the input to the sigmoid function (in a classifier). However, it is not easy to perform inferences and adapt GP models as in regression problems [138], because the likelihood function of the target data given the GP representing the inputs is not a Gaussian function (due to the use of sigmoids). Williams and Barber [220] used Laplace approximations to implement classifiers based on Gaussian Process priors whereas Neal [156] used Markov Chain Monte Carlo methods. Methods of developping Gaussian Process classification to multiclass problems are described in Neal [156], Gibbs and MacKay [91], and Williams and Barber [220]. The methods for both supervised and unsupervised classification have been studied extensively ([76], [22], [185], [138]), and the previous paragraph highlighted some of the most widely used classifiers. The next few paragraphs describe the methods used in this thesis. Having an optimal classifier is necessary to the successful classification of sonar data used in this work. On-going work is investigating the use of SVM as well as other kernel methods for the classification of sonar data [2]. The classification software in this work used

[2] done after thesis submission

some of the software available in the Netlab toolbox [154][3].

5.4.1 Supervised Classification

The K-Nearest Neighbour Approach (KNN)

The main problem considered in KNN is how to measure the similarity of a sample to a class given k neighbours. The Euclidian measure (5.3) can be used between each two samples in feature space as a similarity measure. Starting with each sample, and given k nearest neighbouring points to it, the simplest way to determine the classification is by using the majority rule. An unclassified case is assigned the category most frequently appearing among the k nearest neighbours. For the KNN, the factor k acts as a smoothing factor and there is an optimal choice for k.

The supervised KNN approach starts form centres assigned using the training dataset, and uses the KNN approach to classify the rest of the dataset.

The Multilayer Perceptron (MLP)

One of the most general and most widely used neural network models is the 'multi-layer perceptron (MLP)'. The basic building block of this model is the perceptron introduced as a mathematical analogue of the biological neuron. The MLP is formed by connecting successive layers of adaptive weights. Connections run from every unit in one layer to every unit in the next layer, with no other connection permitted.

The output of the jth hidden unit a_j (including the biases- refer to [22] for more detail) can be written as a weighted linear combination of the d input values (an extra term is added to include the bias):

$$a_j = \sum_{i=0}^{d} w_{ji}^{(1)} x_i \tag{5.9}$$

[3]http:/www.ncrg.aston.ac.uk/netlab/

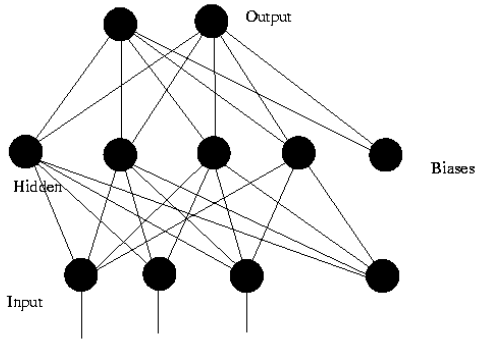


Figure 5.1: An example of a feed forward MLP neural network with 2 layers. The bias parameters in the first and second layer are shown as weights from an extra node with an activation value of 1.

where x_i is the ith input and w_{ji} is the weight connecting the input i to hidden node j. The activation of hidden node j, z_j is obtained by transforming the sum in 5.9 using an activation function g :

$$z_j = g(a_j) \tag{5.10}$$

The final outputs of the network are the result of weighting the hidden node outputs (z_j) with another layer of weights and using an activation function h on the final layer. h and g can be the same function, or two different functions. Thus the complete function represented in figure 5.1 can be written as:

$$y_k = h(\sum_{j=0}^{M} w_{kj}^{(2)} g(\sum_{i=0}^{d} w_{ji}^{(1)} x_i)) \tag{5.11}$$

The activation functions can be chosen to be a number of functions including linear and sigmoidal functions as shown in figure 5.2. It has been shown that a two layer MLP is capable of forming an arbitrarily close approximation to any non-linear mapping ([61]) provided that the activation function is continuous, smooth and monotonically increasing. The sigmoid function (figure 5.2) has these properties.

MLPs are trained on a representative dataset (training data). Known examples, consisting of input patterns and corresponding output patterns, are repeatedly offered to the network during

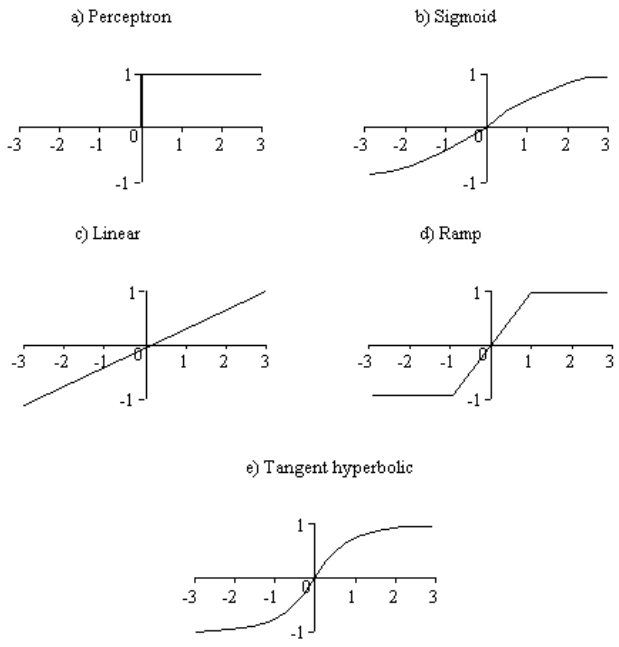


Figure 5.2: Examples of functions that can be used as activation functions in an MLP.

the training phase. Most training algorithms involve an iterative procedure for minimisation of an error function, with adjustments to the weights being made in a sequence of steps. The 'back-propagation' algorithm is normally used to train an MLP. It consists of two passes through the MLP. The first pass is a forward pass that calculates the output to a given input pattern. The error between this output and the desired output is also calculated. The second pass of the algorithm is a backward pass that propagates the error backward, and adjusts each weight so that the actual output of the MLP moves closer to the desired output.

Optimisation of MLPs The problem of learning in neural networks is usually formulated in terms of the minimisation of an error function E. This error is a function of weights and biases of the network, grouped together in a weight vector w. A search in weight space is performed, to locate the global minimum of the error without falling into stationary points and local minima. Many algorithms have been developed for the optimisation of neural networks and finding the global minimum. A few will be summarised here.

The *gradient descent* algorithm starts with an initial guess of the weight vector. It is updated automatically, such that at each step, a short distance is moved in the direction of the greatest rate of increase of the error. The gradient of the error is calculated at each step. One of the main problems of gradient descent is that at most points on the error surface, the local gradient does not point directly towards the minimum, so the algorithm takes many steps to reach a minimum.

A way of optimising this is by using *line search* techniques in weight space. A search direction is considered in the weight space, and the minimum of the error function is found along that direction.

Another algorithm that is used is the *conjugate gradient* algorithm. A sequence of search directions is chosen such that each new direction is conjugate to all previous directions. This approach solves one of the problems encountered in *gradient descent,* where the error oscillates on successive steps while making little progress towards the minimum. The proof can be found in [22].

Newton's method is another algorithm that can be used. It calculates a direction pointing to the minimum of the error at each step. This direction is known as Newton's direction and can be written as : $-H^{-1}g$, where g is the gradient of the error function and H is the Hessian matrix given by:

$$H_{ij} = \left(\frac{\partial E}{\partial w_i \partial w_j}\right)_w \tag{5.12}$$

Quasi Newton methods build up an approximation of the inverse Hessian and thus offer a significant computational optimisation.

When using an optimisation algorithm, some choice must be made of when to stop the training process. The choice can be to stop after the process has performed a fixed number of iterations, or when a predetermined amount of CPU time has been used. It can also be dependent on the error itself ; the algorithm can stop when the error falls below a certain threshold or when the relative change in error function falls below some specified value.

5.4.2 Unsupervised Classification

The Kmeans Algorithm

A simple clustering technique that can be used to assign a set of centres to reflect the distribution of datapoints is the Kmeans clustering algorithm [151]. The number of centres K must be decided in advance. Kmeans partitions the datapoints into K disjoint subsets S_j containing N_j data points, to minimise the sum-of-squares clustering function given by [22]:

$$J = \sum_{j=1}^{K} \sum_{n \in S_j} \|x^n - \mu_j\|^2 \tag{5.13}$$

where μ_j is the mean of the data points in set S_j and is given by:

$$\mu_j = \frac{1}{N_j} \sum_{n \in S_j} x^n \tag{5.14}$$

The batch version of the algorithm begins by assigning the points at random to K sets and then computing the mean vectors of the points in each set. Next, the points are assigned to a new set according to which is the nearest mean vector. The means of the sets are recomputed at each step. The procedure is repeated until there is no further change in the grouping of the datapoints.

5.5 Summary

This chapter presented a view of techniques used in this work for feature analysis. Methods of feature visualisation were first introduced, then a review of feature selection techniques and feature classification methods. An efficient classification procedure is a combination of feature selection and classification. The selection of method of classification is dependent on the application considered, and on the availability of training data. Examples of using these techniques will be detailed in chapters 6 , 7 and 10.

Part I

1-D Sonar Signal Analysis

Chapter 6

1-D Sonar Feature Extraction and Application to Datasets

6.1 1-D Feature Extraction

This chapter deals with 1-D sonar signals rather than sonar images. This approach is useful in several cases. An example is when sonar features need to be observed in real-time, and features need to be extracted per ping as the ship is moving, rather than waiting for the whole survey data. Another use of this method is when exact geo-referencing is not available for the received data.

Methods for classifying seabeds using 1-D signals can be divided into two categories, parametric and non parametric methods. Parametric methods used in acoustic seafloor characterisation directly estimate morphological parameters of the seafloor by physically modelling the acoustic backscatter from a sea bed. These models study the interaction of sound with the medium in terms of acoustic propagation and scattering. Thus, parametric methods invert these models to obtain the respective parameters of the sediments [112, 111, 159, 152, 171].

As explained in chapter 3, section 3.1.1, the results of using backscattered amplitude for the classification of the sidescan bathymetric data supplied for this work confirm the results

obtained by Goodfellow [93]. If bathymetry is not of adequate resolution, it can lead to misclassifications if it is used to correct for sonar parameters. The resulting 'processed' backscattered amplitude, used on its own for classification, shows overlap between sediment classes in the sidescan bathymetric datasets.

Previous studies invloved in using multiscale spectral features for sonar data are given in section 3.2. The next sections use the wavelet transform introduced in chapter 4 as a means of calculating multiscale spectral features. The application of wavelet analysis to underwater seabed classification is a new approach explored in this chapter. Results on both sidescan and sidescan-bathymetric sonar datasets are presented and discussed.

6.1.1 The Use of Spectral Features for Sonar Classification

The spectral method of Pace and Gao [167] for extracting features from sonar signals can be summarised in the following steps:

- The DC component is first removed from the signals by using a regression line for each trace. A cosine bell taper is used to suppress boundary samples on each swath.
- The Power spectrum modulus of each of the traces is calculated by using the FFT (Fast Fourier Transform) then the log of the Power spectrum P_{NL} is calculated and normalised.
- Windows are considered from P_{NL}. Features are described as the ratio of energy in each window over the total energy in the normalised log power spectrum. The selection of windows was determined experimentally by Pace based on maximising class distance.

Reut [179] and Pace [167] also investigated the use of the power cepstrum [1] of signals, but noted that using the spectrum provided better class discrimination.

[1]The power cepstrum can be obtained by performing the direct Fourier transform on the logarithm of the averaged power spectrum and taking the modulus squared of the resulting complex quantities.

Tamsett [200] suggests an alternative method of feature extraction from the power spectrum. The method fits an appropriate low pass filter response curve to observed spectra. Parameters of the filters are then optimised and one or more of the optimised parameters are used for defining features. Tamsett's method is less susceptible to changes in the noise levels of the signals compared to Pace's method, but is computationally slower.

Although these methods provide very good classification for the dataset considered, they have a few disadvantages. The main disadvantage is that features are considered per swath. If an area is composed of many types, the boundaries within the same swath can not be recognised. Windows for calculating the power spectrum are normally considered over the whole signal, probably because small windows are affected by noise.

Another disadvantage is the process of feature selection, since in [167], the window limits are selected experimentally. With different datasets, new window boundaries might have to be considered. Thus, the process of feature selection is not automatic.

A way of localising spectral features in time is by considering the wavelet decomposition of a certain signal. The construction of the wavelet transform is detailed in chapter 4. Wavelets allow frequency features of a signal to be split over several scales, the selection of windows over certain areas in the scalogram (refer to figure 4.2) allows the localisation of features. The next few sections will explore feature extraction from wavelet transforms of sonar signals as well as selecting wavelet features automatically.

6.1.2 Feature Analysis using the 1-D Wavelet Transform

Observing scalograms of the wavelet transforms of signals from two different seabed types (from the Hopvågen dataset), as in figure 6.1, the user can note certain differences. Visual discrimination using scalograms, however, is subjective depending on the user and it is hard to achieve consistency, especially when thousands of signals are considered. An alternative is to use an automatic classifier such as a neural network, but the problem in this case is the large

dimensionality of the wavelet transform features even if both scale and time were discretised.

For efficient generalisation and training of a classifier, as well as visualisation, only a small number of coefficients should be used to describe a signal. For a visual display, two or three is a realistic number for correct visualisation.

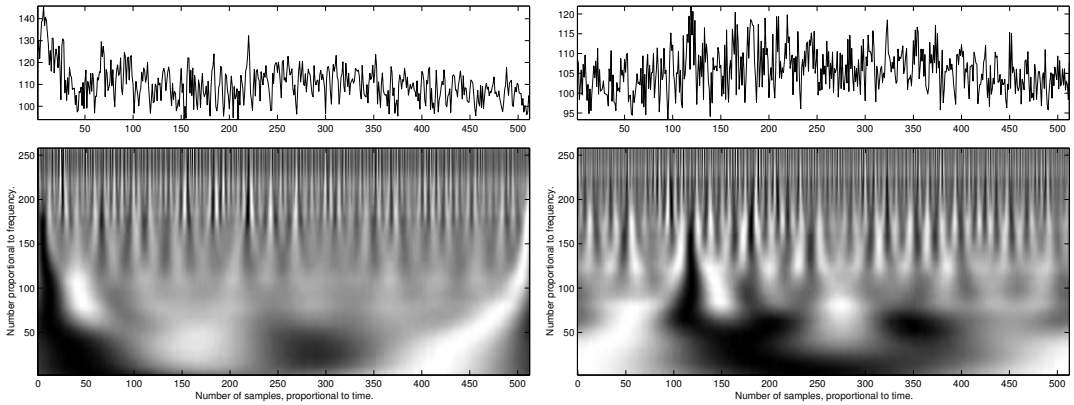


Figure 6.1: This figure presents the continuous wavelet transform of two signals from the Hopvågen site. The upper graph, is the processed amplitude of the received signal, plotted versus the number of sample. The bottom graph is the scalogram of the signal. The x-axis is the number of the samples (proportional to time), the y-axis is the scale. The change of wavelet coefficients is noticed as we go up in the scalogram, and higher frequency components are selected.

To reduce the amount of data obtained from the one-dimensional wavelet transform, many approaches have been considered. Chen *et al.* [42] use wavelets to discriminate between different passive sonar signals radiated by ships. The approach followed in their work was to select tonal features from the wavelet transform spectogram and then give all these features to a neural network for classification. To reduce the amount of data, in an application closer to ours, Kundu *et al.* [127] threshold the wavelet transform coefficients and select those wavelets with maximum amplitude as features for classification. Telfer *et al.* [202] also use wavelet scalograms for studying acoustic backscatter signals. They use Gaussian patches in the scalogram, choosing by eye those which they believe to contain the most discriminatory information.

The method used in this work to reduce the number of coefficients consists of using a window

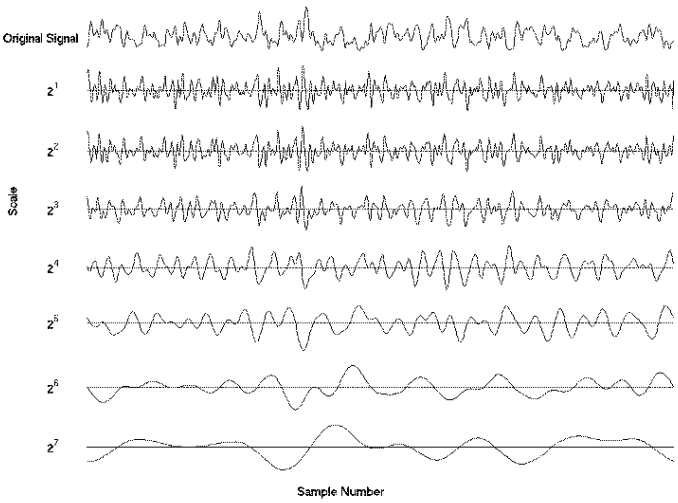


Figure 6.2: Using the dyadic wavelet transform of a processed signal from a stone sea-bed. The original signal is represented above, then the dyadic wavelet transform is shown over several scales.

to find properties of the individual wavelet coefficients Wf over a section of time (energy window $[a, b]$). The time-section represents a range of grazing angles. Wavelet transforms are squared at all scales of the scalogram and their mean is calculated over the interval $[a, b]$. This is similar to calculating the 'energy' of a signal at a certain frequency (or scale), considering a fixed time window (as in figure 6.2). The feature S_k at scale k is:

$$S_k = mean(Wf[i,k]^2 \tag{6.1}$$

For a discrete wavelet transform of a signal f and in the case of dyadic scales[2]:

$$S_k = mean(Wf[i,2^k]^2) \tag{6.2}$$

Another feature which proved to provide good discrimination between classes is the standard deviation SD_k of wavelet transform energy in a window $[a, b]$ over a scale k.

[2] i is the sample number and k is the scale nuber in the wavelet decomposition of a signal.

$$SD_k = \text{standard deviation}(Wf[i,k]^2) \tag{6.3}$$

The limits of the time interval $[a, b]$ are chosen depending on the total number of samples, and are chosen to avoid clipping of the signal, as well as near nadir and very far datapoints. However, choosing very small time intervals does not give enough information to be considered as a feature for a classifier due to the noise present in sonar data.

It is worth noting that the signal used for wavelet analysis was chosen to be the processed sidescan signal after correcting for physical factors and sonar directivity. The reason for not using the raw signal is that the wavelet transform would be affected by seabed depth and sonar directivity since the amplitude would vary with both these factors. The selection of wavelet type for this work is given experimentally in section 6.2.

Feature Visualisation

To compare the different wavelet bands, scatter plots of the selected features can be plotted in 2-D or 3-D. All the bands cannot be observed simultaneously. Sammon's mapping (explained in section 5.2) is used to observe the projection of features on a 2-D scatter plot. To observe feature clusters from Pace's dataset for example, Sammon's mapping was used to transform data to a lower number of bands. From 7-D to 2-D, the classes suffer from overlap, but Sammon's method proved to be a good method to observe the classes projected from 4-D to 2-D as in figure 6.3. This figure shows the projection of 4 highest frequency bands over 2 axes.

Feature Selection and Classification

The Backward elimination algorithm explained in section 5.3 was used to select wavelet bands that give the maximal distance between classes. The cross entropy and Fisher measures were used to measure class distances. As for the classification stage, the choice was between neural networks or simpler clustering algorithms such as the KNN algorithm. The choice is dependent

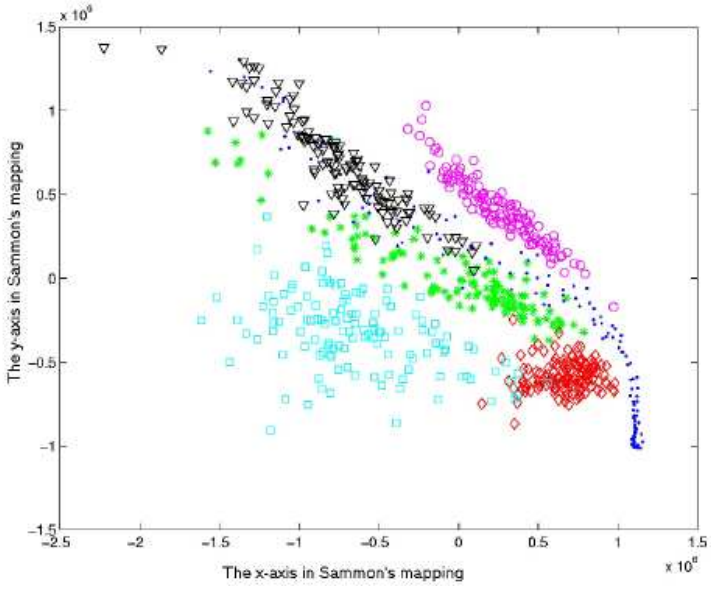


Figure 6.3: Sammon's mapping applied on 4 bands (the first four), to transform the classes to a 2-D space. (Clay=*, gravel=., mud=diamonds, rock=o, sand=squares, stone=triangles.)

on the availability of training data and the size of the testing data that is being classified. If the testing dataset is large (a whole survey for example), neural networks require a large training set. An alternative is to select cluster centres from training points then use a clustering algorithm for classification (KNN). In the case of Pace's dataset, the number of training points was sufficient for training a neural network classifier, whereas in the Hopvågen survey, the small number of grabs compared to the size of the dataset led to the use of clustering algorithms.

6.2 Wavelet Type Selection

To select an appropriate wavelet type for this work, several wavelets were tested on Pace's dataset (since it is groundtruthed), and the wavelet that provided the best overall classification was selected. The results of using scatter plots to observe class differences by using wavelets is given in figure 6.4. The wavelet transform was calculated over 6 scales for each ping resulting in the features S_k and SD_k for each scale. Thus, each ping is described by a feature vector of dimension 12. Only the first 6 bands were considered since very low frequency scales in the wavelet decomposition tend to approximate the general shape of the signal and do not provide much information on the important details that are found on high scales. By observing figure 6.4, we can deduce that the dyadic wavelet seems to provide a better class separation than other wavelet types used. The features selected for display in these scatter plots are the features selected by the backward elimination algorithm as the features providing the best discrimination.

The advantage of the dyadic wavelet (with continuous time) over other discrete time wavelets is that it provides translation invariance. As previously explained in section 4.2.2, the scale is discretised dyadically, but not the translation parameter, providing a continuous representation across scales. This probably leads to 'keeping' more discriminating information, across several scales as proved in figure 6.4. Verification of wavelet type selection using neural network (selecting the best wavelet according to classifier results) is given in section 6.3

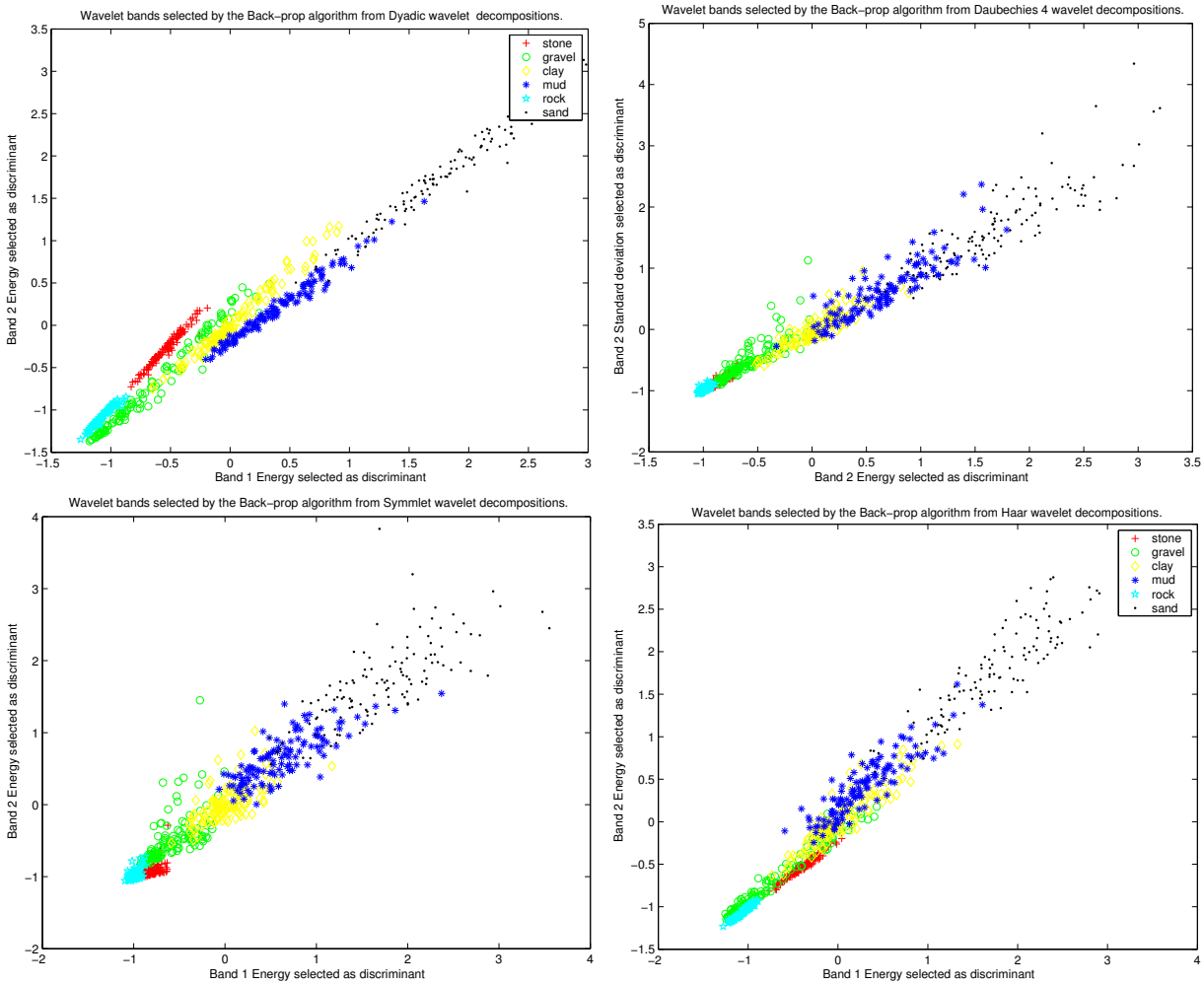


Figure 6.4: Using several wavelets to select features from the Pace dataset (from above: Dyadic wavelet transform using spline wavelets, Daubechies 4, Symmlet 4 wavelet and Haar 4.) The best two discriminating bands are selected by the Backward elimination algorithm. The symbols are given in the first figure.

6.3 Results on Pace's Datasets

Pace's dataset is introduced in section 2.3.1. The aim of this section is to derive an automatic method using wavelets for the classification of this dataset.

6.3.1 Classification by Using Neural Networks (MLP)

The same method of feature extraction from the wavelet transform as in the 6.2 section was used in this case. The dyadic wavelet was used to extract features, since it provided better class separation as observed in the scatter plots of figure 6.4.

A training dataset of 40 pings of each type of sediment in the dataset was considered. The testing dataset was of 50 pings. The Backward elimination algorithm (BE) was used on the training dataset to extract the n most discriminant bands using Fisher's distance [3]. In each case, the n selected bands were used as inputs to an MLP neural network (see section 5.4). Several configurations for MLPs were considered but the best results were obtained for MLPs with the following configurations:

- The MLP is made of 2 layers and 7 hidden nodes.
- The number of training cycles is 2500.
- The number of outputs is 6 (corresponding to the number of classes).
- The algorithm used for optimising the MLP is the scaled conjugate gradient algorithm.

Table 6.1 displays the results of conducting several experiments on the MLP mentioned above. The main reason for using several numbers of inputs is that although the BE algorithm can decide which are the best features to use for classification, it cannot deduce their optimal number. An efficient way to test for the optimal number is to use the classifier itself and observe the success rates obtained by using different feature subsets. The average success rate of several experiments is displayed in the third column of table 6.1.

Figure 6.5 shows the success rate of an MLP network versus the number of inputs. The two curves shown are for two MLPs of 7 and 8 hidden nodes. The optimal number of features can be selected where the graphs show a maximum. The maximum is observed when the number

[3] Using Euclidian distance gave the same results (same bands selected) as the Fisher distance for almost all cases considered in this section.

of inputs is around 6 (although the graph fluctuates a lot). This is probably due to the reason that small feature subsets (2 for example) do not contain enough information and the whole set of features (12 for example) has a lot of redundant information. Using 6 inputs to the MLP, the average success rate over 10 experiments is 97.7%, although the maximum success rate is 98.5%. The results of several experiments were observed and the overlap was mainly between the two classes of clay and gravel. The other 4 classes (namely mud, stone, rock and sand) were in general well separated.

These results are comparable to the results obtained by both Pace and Tamsett [167, 200] when dealing with this dataset. The main advantage is that the selection of features in this study is an automatic one, and user intervention is minimum. Another advantage is that more time-frequency resolution can be obtained by using wavelets compared to Fourier transforms. Thus, the windows selected can be adapted to be areas of valid grazing angles.

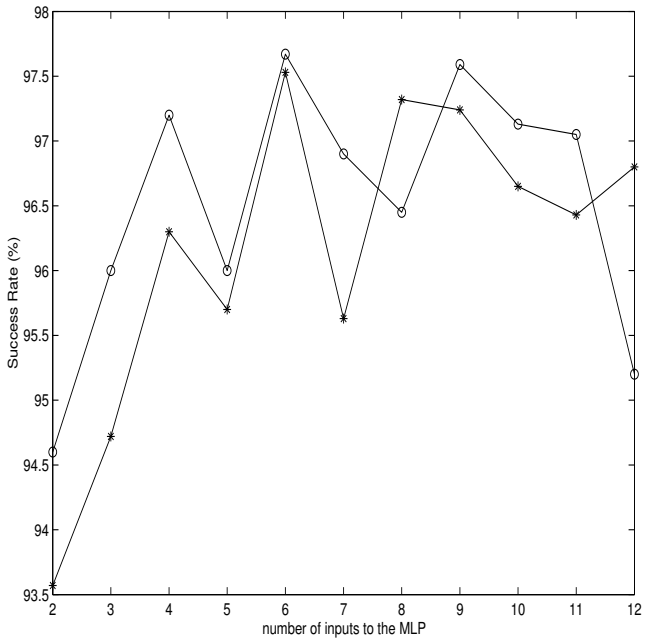


Figure 6.5: The success rate versus the number of inputs using 2 cases of MLPs (testing Pace's data). Circles are the results of using the MLP with 7 hidden nodes, stars are the results of using an MLP with 8 hidden nodes.

Number of Inputs (n)	Features Selected by BE	Results (%)
2	S_1 and S_2	94.6
3	S_1, S_2 and SD_2	96
4	S_1, S_2, SD_1 and SD_2	97.2
5	S_1, S_2, SD_1, SD_2 and SD_3	96
6	S_1, S_2, S_3, SD_1, SD_2 and SD_3	97.7
7	S_1, S_2, S_3, SD_1, SD_2, SD_3 and SD_4	96.9
8	S_1, S_2, S_3, S_4, SD_1, SD_2, SD_3 and SD_4	96.5
9	S_1, S_2, S_3, S_4, S_5, SD_1, SD_2, SD_3 and SD_4	97.6
10	S_1, S_2, S_3, S_4, S_5, SD_1, SD_2, SD_3, SD_4 and SD_5	97.1
11	S_1, S_2, S_3, S_4, S_5, S_6, SD_1, SD_2, SD_3, SD_4 and SD_5	97.1
12	all S_k and SD_k	95.2

Table 6.1: The results an MLP with 7 hidden nodes and 2 layers, with a different number of input in each case according to the best bands selected by the Backward elimination (BE) algorithm. The last column is the average success rate over 10 experiments. S_k is the mean wavelet energy at a scale k, and SD_k is the standard deviation of the energy at a scale k.

6.3.2 Neural Networks(MLP) For Wavelet Type Selection

The optimal MLP found in the previous paragraph (with the optimal number of hidden nodes) is used here to test the success rates obtained from using several types of wavelets to extract multiscale features. Table 6.2 gives an average of the success rates obtained by using the same procedure described above for extracting multiscale features then using an MLP for classification. The only parameters changed are the wavelet types (given in the first column of table 6.2. The other parameters are chosen as the optimal parameters from the previous section, example 6 inputs with 7 hidden nodes). 7 experiments were run using each wavelet type and the results averaged. Although the results are close, they give more evidence that the dyadic wavelet trans-

form using spline wavelets gave better results than other discrete wavelet transform methods (which was observerd in the scatter plots of figure 6.4).

A disadvantage of using MLPs is that several parameters are selected experimentally and finetuned during the classification stage, such as the number of hidden nodes, and the training cycles. Other methods of classification such as kernel methods and support vector machines (SVM) could be used to overcome some of these problems.

Wavelet Type	Average Success rates using MLP (in percent)
Dyadic (Spline wavelet)	97.79
Daubechies 4	96.39
Symmlet 4	96.77
Coiflet 4	96.68
Haar 4	97.2

Table 6.2: The results of using an MLP with 7 hidden nodes and 2 layers, using several wavelets to extract signal features. 7 experiments were done for each wavelet type, the averaged success rates are given in the last column.

6.4 Results on the Hopvågen Dataset

This section uses 1-D wavelet analysis for the study of sediment types in the Hopvagen dataset (presented in section 2.3.2). The advantage of using this set is that data can be visualised versus the exact GIS position of each point unlike the Pace dataset. Grab samples are provided over areas in the survey and can be used for the training and validation stages of classifiers.

Feature Observation The wavelet transform was performed over the whole survey, ping by ping and the features S_k and SD_k extracted. The signals used for wavelet analysis are the

signals obtained after correcting for losses, thus describing the backscattering strength of the seabed.

Figures 6.6 and 6.7 show the feature S_1 and SD_1 respectively interpolated over both sides of the sonar displayed versus eastings and northings of each swath. The highest frequency band statistics were selected for display since the BE algorithm picked these bands in most cases considered.

6.4.1 Relationship between Wavelet Features and Sediment Types

A number of different physical parameters such as grain size and the ratio of volume to surface scatter affect the backscattering intensity and distribution, and hence the wavelet features. Since the wavelet transform picks out information at different spatial scales, correlation is expected between the wavelet features at different scales with grain size. To study this relationship, several areas around given grab samples were selected for training. Although more grab samples were provided for this work, they were located in areas quite far from areas surveyed, or in areas of no bathymetry information, so they were discarded. The grab samples are given in table 6.3 with their corresponding grain size. The type of these sample was selected using the Udden Wentworth table given in appendix A. The grab samples are numbered in figure 6.12 , which also gives the result of classification on this dataset (explained in the next paragraph). The selection of areas around grab samples is done automatically. As several grab samples do not fall exactly in the areas surveyed, the closest surveyed area (with a maximum limit of $5m$) is selected for study. A maximum of 20 pings near the grab sample is usually selected

Scatter plots were plotted of each averaged (between pings) wavelet scale versus the grain size. Examples are given in figure 6.8. Wavelet energies of low scales (representing high frequency features) seem to provide better correlation with grain size than low frequency scales. In general, wavelet means also provide more correlation that wavelet standard deviation.

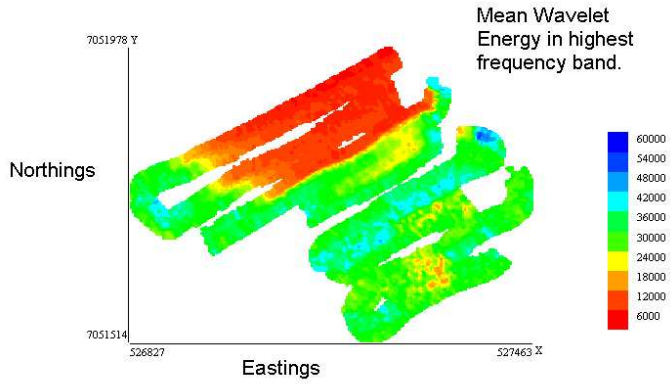


Figure 6.6: The figure displays the wavelet energy in the highest frequency band (S_1) over the survey. The left and right pings are studied separately and the results are displayed versus the easting and northing position of each point.

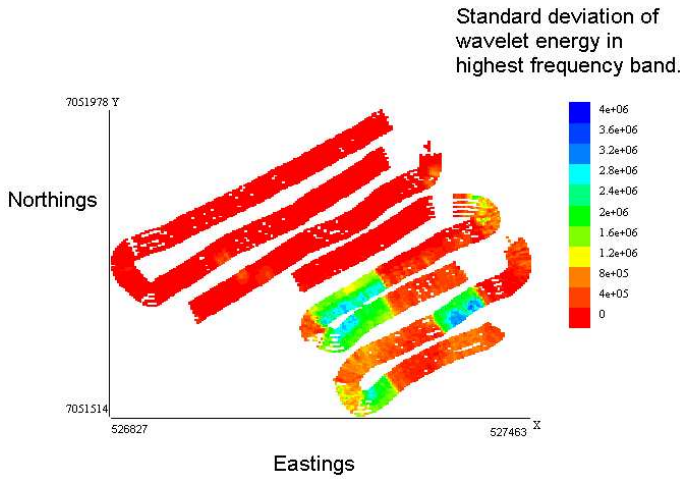


Figure 6.7: The figure presents the standard deviation of the wavelet transform coefficients in the highest frequency band (SD_k) displayed in the survey. The left and right pings are studied separately then the results are displayed versus eastings and northings of each point.

Point Number	Location Eastings	Location Northings	Type	Grain Size in μm
1	526878	7051782	1-Med Sand	289
2	526963	7051700	2-Fine Sand	196
3	526990	7051690	2-Fine Sand	200
4	527008	7051850	3-V. Fine Sand	108
5	527099	7051753	4-Mud/Silt	63
6	527125	7051854	4-Mud/Silt	58
7	527138	7051827	4-Mud/Silt	56
8	527144	7051665	1-Med Sand	572
9	527225	7051569	1-Med Sand	243
10	527181	7051886	4-Mud/Silt	65

Table 6.3: Location and grain size of samples selected from training areas for the Hopvågen dataset, also their classification according to the Wentworth table into four classes.

To get more insight on the correlation of wavelet features with grain size, the matlab function *corrcoef* was used. This function calculates the linear correlation coefficients between two variables. This function also gives a probability p, which tests the hypothesis of no correlation between the variables. Correlations where p is of a small value (less than 0.05) are considered for analysis. Table 6.4 summarises the results. The first four features (namely the first four wavelet means) provide a relatively good correlation with grain size. Correlation decreases when standard deviation features are considered, or when low frequency wavelet bands are considered for analysis. The features that provided high correlation with grain size were in most cases identical to the features selected by the BE algorithm.

To get more insight on the relationship between wavelet features and sediment type, pings were selected around grab areas of each type (numbered in figure 6.12) and used for analysis. First the Backward elimination algorithm was used to find the two most discriminating features,

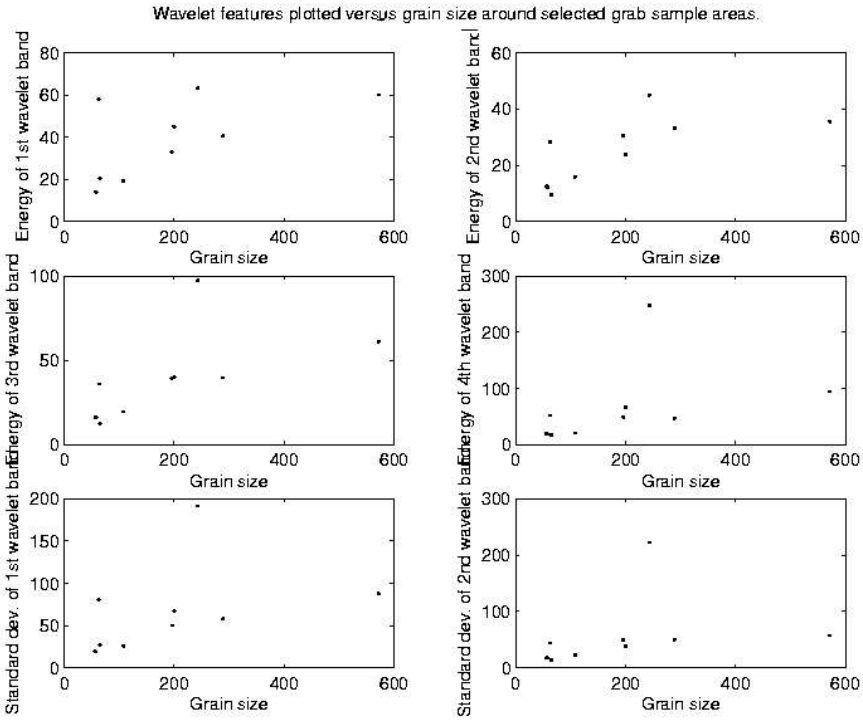


Figure 6.8: Plots of wavelet features versus grab samples. The first four are wavelet energies, and the last two are standard deviation features (as given in the axis labels).

according to Fisher's criteria. These features were plotted for each point and colorcoded in scatter plots (figures 6.9 and 6.10). The features selected were the second and third wavelet band energies. Each ping in the selected areas is represented as a point in these scatter plots. Figure 6.9 is a plot of nine points selected from table 6.3 excluding point 9. In figure 6.10, all the grabs in this table are used. Some remarks on the class distributions in these figures are as follows:

- Point 5, which belongs to class 4 (silt/mud-colorcoded as red) seems to totally converge with class 3 of fine sand (green). Looking at table 6.3, the grain size is similar to that of the silt/mud points, and the error in this case is most probably due to ping selection, or due to having more than one type of sediment per ping. However, the other silt/mud points seem to cluster very well (the red points).

Feature	Cross Correlation with Grain size	p-value
S_1	0.62	0.05
S_2	0.68	0.03
S_3	0.61	0.05
S_4	0.42	0.22
S_5	0.23	0.52
S_6	0.19	0.59
SD_1	0.43	0.21
SD_2	0.31	0.37
SD_3	0.28	0.43
SD_4	0.23	0.5
SD_5	0.17	0.62
SD_6	0.17	0.63

Table 6.4: The first column is the feature, the second the correlation coefficient and the third the probability of having no correlation.

- The very fine sand class (in yellow) has several points overlapping with the silt/mud class, although in general, it can be viewed as a class on its own.
- The fine sand class (in green) clusters very well for both grabs selected.
- Point 1, which is of medium sand, has some overlap with the green class of fine sand. Point 9, of medium sand also has a lot of variability between the pings selected. This is why it was not used in figure 6.9, to observe the other inter-class variations. In figure 6.10, some samples belonging to Point 9 overlap with the other medium sand points (in blue), whereas some samples seem to have higher energies suggesting that the area is a mix of medium sand, and a coarser type of sand or gravel.

From the analysis of correlation and the observation of scatter plots, a clear trend is apparent, supporting the view that scale features are useful for sediment discrimination.

Classification of known samples Good visual separation on a scatter plot suggests that a neural network or other classifier can be trained to discriminate between classes. Pings used in the previous section are used for testing classifier performance. For this section a KNN classifier was used (with varying values of k). To test the performance of the classifier, the k-fold cross validation method can be used as a figure of merit (there are several others, [22] gives a few examples). In this method, the data set is divided into l subsets. In each iteration, one of the l subsets is used as the test set and the other $l-1$ subsets are put together to form a training set. Then the average error across all l iterations is computed. The advantage of this method is that it matters less how the data gets divided. Every data point gets to be in a test set exactly once, and gets to be in a training set $l-1$ times. The variance of the resulting estimate is reduced as l is increased.

Each input to the KNN is a vector containing 5 features selected by Backward elimination (which was run on the training data). The outputs are the class labels. Several values of k were tested for the KNN algorithm. The results of average classification (using k-fold cross validation) are given in figure 6.11 for several values of k. Low values of k provide higher rates of classification, with l in this case chosen to be 16, and each subset of the data containing 12 samples. The best results are for $k = 2$, and with 96.34% of correct classification.

Using this value of k, the KNN algorithm is used again, using averaged pings around training areas for training. The test set is a set of pings of mixed types (180 samples), containing 41 samples of Medium sand, 41 of fine sand, 23 of very fine sand and 75 of mud/silt. This experiment is done to deduce which types are overlapping and thus leading to misclassification. In this case, less knowledge of types is assumed and only 5 pings (averaged) around each sample are used for training. This is mainly because in most surveys, large areas around grabs cannot be assumed to be of the same type (as done in the previous experiment). The confusion matrix

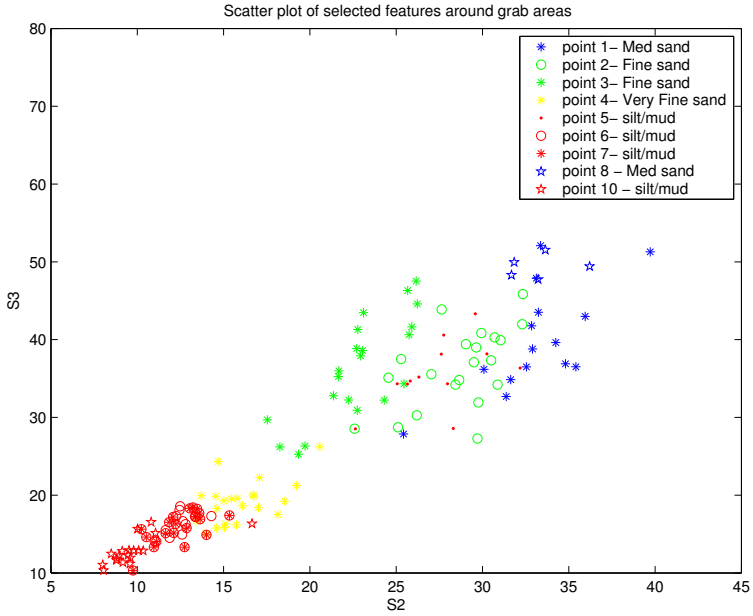


Figure 6.9: Scatter plot of the selected features from the Hopvågen dataset, around the given grab samples in figure 6.12, only nine grabs are selected for better visualisation (from table 6.3).

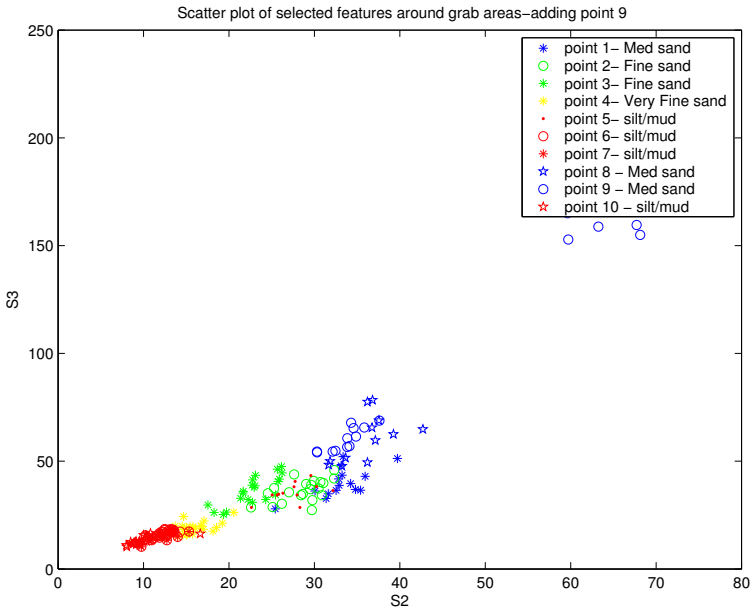


Figure 6.10: Scatter plot of the selected features from the Hopvågen dataset, around the given grab samples in figure 6.12, all ten grabs in 6.3 included.

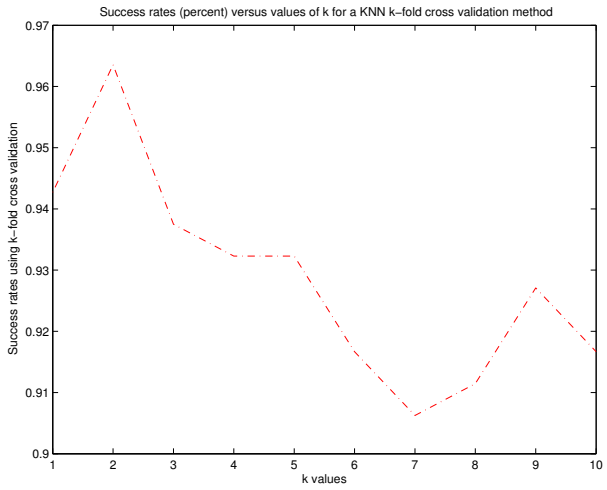


Figure 6.11: Average results of the KNN using k-fold as a cross validation method for several values of k.

is given in table 6.5, and the average classification success rate is 90.56%. From this table we can notice that the classes that overlap are mostly the medium sand and the fine sand (the blue and green classes in the scatter plots). Some samples of mud/silt are misclassified also as types of sand. In the scatter plot of figure 6.9, these could be points from areas around Point 5 which seem to converge with the fine sand class (for the features selected). Using more features (as used in this case), would probably lead to less ambiguity between classes.

6.4.2 Classification of the Whole Survey and Comparison with Other Methods

The same parameters selected in the previous paragraph were used to classify the whole survey, which consists mainly of areas of unknown types. The first step was to process the signals in the survey and correct for depth, range and insonified area variation. Then, the wavelet transform of each ping was calculated using the dyadic wavelet transform, since it provided better class separation in the Pace dataset. 12 features describe each ping, and they include the wavelet means and standard deviations of six wavelet bands. The training set is the same one as before,

True classes\Predicted classes	Medium sand	Fine sand	Very fine sand	mud/silt
Medium sand	36	4	0	0
Fine sand	4	35	0	3
Very fine sand	0	1	23	3
mud/silt	1	1	0	69

Table 6.5: Confusion matrix on the selected samples from the four classes in the Hopvågen dataset using KNN, the total classification success rate is 90.56%.

using 10 grabs for training, as detailed in table 6.3. KNN was used with $k = 2$, and 4 outputs, and the Backward elimination algorithm to choose the most discriminating features (5 inputs). The results are displayed in figure 6.12.

Geo-referencing and Feature representation Using 1-D signal classification, each signal is given 1 class label. It is assumed that each ping is composed of only 1 type of sediment (which definitely causes classification problems), and only a valid range of grazing angles is used for classification. To represent these features for a survey, each point of the ship trajectory should have one point representing the class label 'seen' from that point. To make it easier to compare results with other methods, figure 6.12 represents each ping by 3 points, $15m$ apart from each other, just to stress that each ping is considered of one type. The georeferencing data given for each ping is used to plot the starting and finishing point of each ping.

There is definitely an effect introduced by ship motion, which is mainly observed in curves in the ship trajectory, such as the area near the grab-point 1 in figure 6.12. While the ship motion describes a curve, the inside pings overlap, whereas the outside pings have large spaces between them. Some sonar systems disregard these areas as invalid, and some use interpolation to represent these areas. In this work, these areas are kept for analysis, and interpolation is used to represent final classified areas.

Comments on the classification results Some comments on the classification results:

- Areas around points 6,7 and 10 are correctly classified as mud/silt. So are areas around points 2 and 3, classified as fine sand. Correct classification is also in areas around points 8 and 9 (medium sand).
- Areas close to point 4 are classified as mud/silt rather than very fine sand. This could be due to the area there being mixed between these two types, or could be that the pings selected for training are somewhat far from the grab itself (so there is a training error there).
- Point 5 is obviously misclassified, but that is expected, considering that in scatter plots before (figure 6.9), pings from this area totally converged with the fine sand class (in green).
- Areas around point 1 are classified as fine sand. This could be probably due to a training error, or selection of training pings that actually belong to an overlapping class.
- By observing some grab samples in areas that were not used for training, we can deduce that in general, the method does provide correlation with sediment type. However, the main drawback of the method is that features are considered per ping and inter-sediment boundaries are not detected. This could be a reason for several misclassifications.

Comparison with Other Techniques Another sonar survey conducted in the Hopvågen area was a survey that used the ECHOplus sonar which is one of the most commonly used commercial packages for surveying (introduced in appendix B). This sonar uses a single beam transducer to both transmit and receive an acoustic energy pulse directly beneath the sonar. The results of using the ECHOplus sonar include a classification of the seabed according to Hardness and Roughness features (also explained in appendix B). Figure 6.13 displays the

ECHOplus results from [15]. The classes in each image are clustered according to correspondence of the values with grab samples in the area. In these two images, the classification differs in some sections between the hardness and roughness and we were not presented with a classification of this survey that uses both parameters. As explained in the appendix, this method considers only areas under the sonar for classification. Thus, interpolation was used by Bates et al. in [15] to represent the areas that are not surveyed in figure 6.13. One very important difference between this work and that of [15] is that of selecting class boundaries between sediment types. The authors in [15] cluster their data in an unsupervised way for each of roughness and hardness, then use clusters to classify their sediment types. In this work, the Wentworth scale is used to classify sediments, which provides a more geo-physically accurate method. It is worth noting that the color codes are different in the ECHOplus results of figure 6.13 and in the results of figure 6.12. Yellow in the ECHOplus results is used to describe very coarse sand. In the data provided for this study, training samples of very coarse sand were not available so this class is not present. The red class in the ECHOplus classification is split into two classes in this study: the red class (with mainly mud and silt) and the yellow class (for samples of very fine sand). The only way to validate both methods is to compare the classified areas with the grab samples obtained. Both methods show a good correlation with grabs, but both fail in a few points.

Real-time Application of the Method

In this work, wavelets were used to analyse sonar datasets that have been already stored and verified with grab samples from the areas considered. This is a first step to test the method and compare to ground-truthed data. Due to the speed of computation of the wavelet transform, online application of the method is possible. The real-time application would include preprocessing the pings according to the bathymetry, performing the wavelet transform and storing the coefficients. After this step, only the coefficients are stored instead of the whole sonar signal, thus, saving a large amount of space. The classification stage, however, can be

done offline, or by using clustering techniques online, where each new point is added to a known cluster. This approach can be investigated if an online application is of interest.

6.5 Summary and Discussion

This chapter provides a novel approach to classifying backscattered sonar signals. The pre-processed signals are used and analysed by using wavelets. This method provides a better observation of both spectral and temporal features than standard spectral techniques using the Fourier transform. The selection of features from the wavelet transform is automatic using backward elimination algorithms. The method can be used in real-time, although in this chapter, the interest was in the validation of the method and the application on already surveyed datasets.

Two datasets were considered, one is a sidescan dataset (Pace) and the other a sidescan bathymetric dataset (Hopvågen). The results of classification on the sidescan dataset were about 98% correct classification. As for the Hopvågen dataset, wavelet features provided good class separation as observed in scatter plots and in results of neural network classification (with more than 90% correct classification). The features selected by the method provided good correlation with the grain samples in the training areas.

One of the difficulties in assessing work of this type is the paucity of ground truth information. However the method has shown a high success rate over a well understood data set, and generally good correlation between classified areas and grab samples. An advantage of using signal data from a single ping is that the method may be extended to an on-line system. However a limitation is that the classification degrades rapidly if the data set includes a transition region between different types of sediment, and the data is too noisy for such regions to be distinguished. Other limitations are those of artifacts induced by fish motion, sonar look direction and the validity of the sonar corrections applied to pre-process the signals. A two dimensional approach, which segments and classifies a corrected sidescan image attempts to

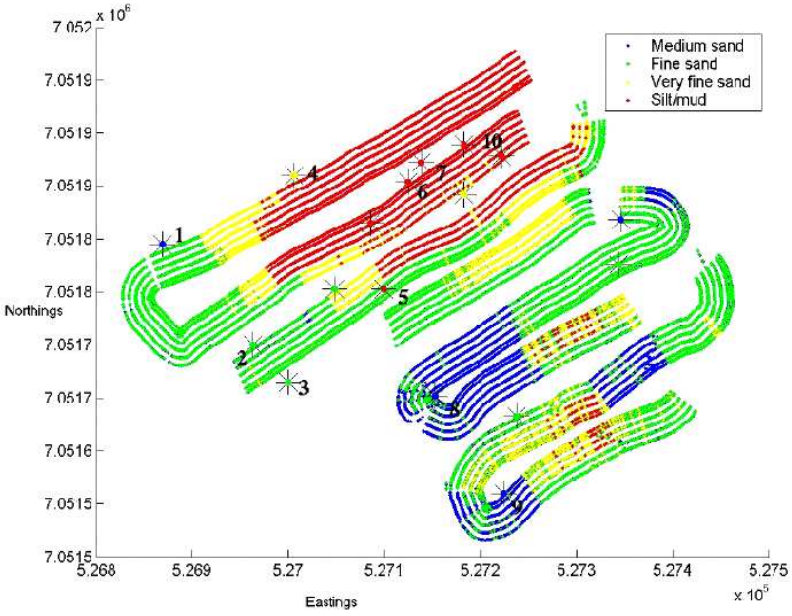


Figure 6.12: The figure presents the result of classification of the dataset after assigning the centres of the classes to known values (according to grain size from the Udden Wentworth table [216]), and running a KNN classifier on the whole set. The numbered grabs are those selected for training.

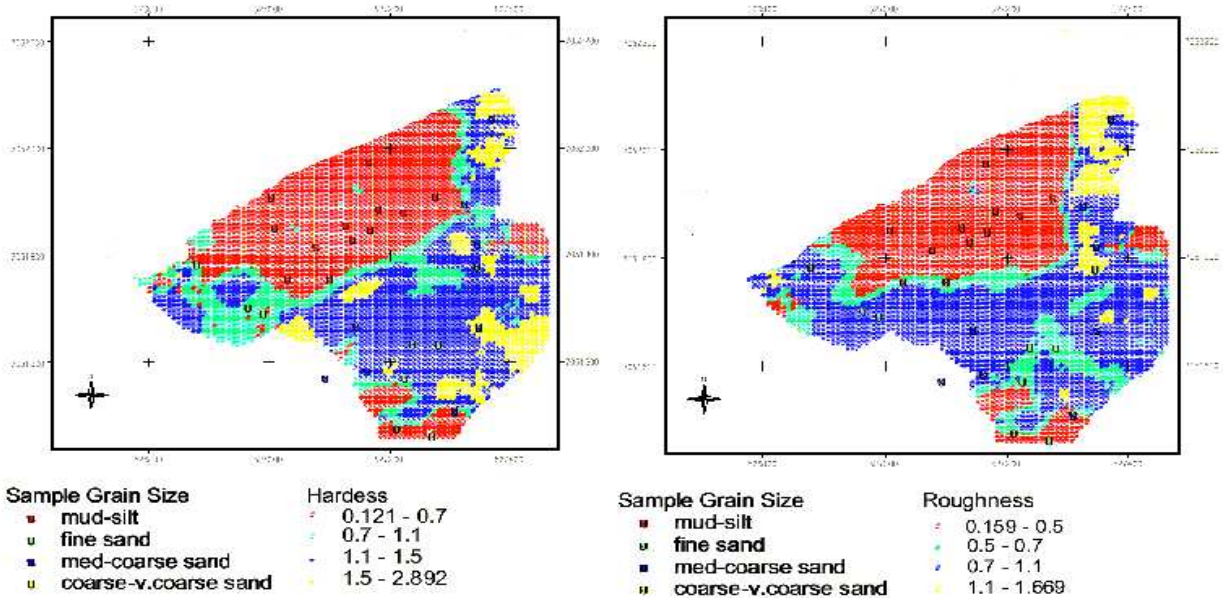


Figure 6.13: The hardness and roughness features from the area considered. The grab samples are displayed and colorcoded. The colorcodes are different from 6.12 and given in section 6.4.2.

solve the problem of having several sediment types per ping. This approach is explained in the next part of this thesis. Information from series of pings is used and a method that combines features from both sidescan and bathymetry is described.

Part II

2-D Sonar Analysis

Chapter 7

Methods for Texture Feature Analysis

7.1 Background

Several sonar classification studies have shown that the use of gray level features (local amplitude) is inadequate to classify and segment sonar images. This is due to the sensitivity of pixel amplitude to the bathymetry of the seabed. In Cochrane [52], adjacent areas of rocky seafloor could not be differentiated from areas of thin sand on the basis of acoustic backscatter (i.e. grey level) alone (due to using a sonar frequency of 100 kHz which is transmitted through thin layers of sediments and reflected from the underlying rock interface). This is why additional features are needed, such as textural features that are a function of multiple pixel values and their relative spatial positioning [12]. 2-D textural features offer an extension to the 1-D frequency properties that were used to classify seabeds in chapter 6. Since 2-D methods observe series of pings, they provide a method to segment boundaries between different sediment types, which was not possible with the 1-D approach.

Due to the vast amount of data present in sonar surveys, sonar data analysis presents an interesting data-mining problem. The approach followed in this work is a feature based one, rather than a statistical modelling based approach. In addition to the need to summarise data [194], this is also due to the variability of sonar data that can be found in a survey, and the

relative difficulty of finding a model that could match all the texture types present in a survey, when compared to models used for photographic image texture modelling. Sonar clutter adds to the problem causing many types to appear similar.

Textural features will be combined with other features that depend on amplitude and bathymetry variation (if bathymetry is available) in chapter 10. This chapter, however, presents an introduction to textural feature selection, and an overview of several feature based methods used for texture classification. These methods will be tested on sonar datasets in chapter 8. The chapter also sheds the light on texture methods using spatial continuity between filter subbands, and several other methods used for texture classification (such as Markov random field methods MRF) in section 7.6.

7.2 Introduction

Texture segmentation has been studied intensively on images from various sources, including biomedical, radar, sonar and photographic images. Texture classification approaches can be broken down into three categories: statistical, structural and model based. Statistical approaches attempt to characterise textures in a probabilistic sense. They can be spatially or spectrally defined. Examples are GLCM features (gray level co-occurrence matrices), filter banks (Gabor), Laws filters, Discrete Wavelet filters, Complex Wavelet filters, among many others. Statistical methods have shown success on data similar to that used in this work; hence their adoption for this thesis.

Structural methods are distinctive from statistical approaches since they consider textures to have two fundamental components: a basic primitive that comprises the texture and the spatial organisation of these primitives. The primitive is usually a fixed template that can be moved to match a certain image pattern [51]. Since sonar images seem to have no predictable, consistent repeating pattern, structural methods are not considered for this work.

Model based methods provide another generic approach to texture analysis. By fitting some

analytical function to texture, the texture characteristics can be captured. The parameters of the model provide a specification for the texture under study. Typically the analytical function is based on a two-dimensional stochastic process or random field. Markov Random Field (MRF) models [46, 41, 118, 145] have been used successfully for texture analysis, as well as fractal based approaches [37]. The main aim of the study is to obtain a robust, low complexity texture analysis method that can deal with large amounts of data, and not be influenced by huge depth variations (i.e. the textural features should be similar for the same type of seabed whatever the depth is). Thus we do not aim to study different types of textural features for the sake of comparison only, since there has been a large number of comparative studies.

In one of the studies, Randen [175] compared heuristically designed filter banks (including Laws filters, Dyadic Gabor filters, Wavelet Transforms Packets and Frames, Discrete Cosine Transforms, Quadrature Mirror filters) with optimised filter banks (including Optimised Gabors) and non filtering approaches (GLCM) on a number of images. Two important results were deduced from that study:

- No single approach performed best or was very close to the best for all the different images. Thus no single approach may be selected as the "winner" of the study. Hence the selection of texture segmentation method must be dependent on the images being studied. Textural techniques for high resolution photography, or minimal-noise radar images may not be the optimal solution for other noisy applications such as sonar images.

- Complexity of the algorithm, and feature dimensionality are also factors to be taken into consideration since they affect the classifier. This is evident in the case of GLCMs that have a very high computational complexity.

Taking these results into consideration, the next sections presents an overview of several texture segmentation techniques, with a summary of the advantages and disadvantages of using each method, and an emphasis on their application to sonar imaging.

7.3 GLCM (Gray Level Co-occurrence Matrices)

The co-occurrence probabilities provide a second-order method for generating texture features (as introduced by Haralick [98]). The matrix contains the conditional joint probabilities of all pairwise combinations of gray levels given two parameters: interpixel distance (δ) and interpixel orientation (θ) (measured counter clockwise from the horizontal). The interpixel distance δ compares to 'scale' in filtering applications. Large interpixel distances between considered pixels correspond to large scale textures and vice versa. The GLCM matrix components C_{ij} are defined by:

$$C_{ij}(\delta, \theta) = \frac{P_{ij}}{\sum_{i,j=1}^{G} P_{ij}}$$

where P_{ij} represents the number of occurrences of gray levels g_i and g_j, and G is the total number of gray levels. The sum in the denominator represents the total number of gray level pairs within a particular window given a particular (δ, θ). A different GLCM is required for each (δ, θ). A complete set of co-occurrence statistics would cover all values of δ and θ over a meaningful range. Typically only four orientations are used : 0, 45, 90 and 135 degrees. The orientation 180 degrees is redundant to 0 degrees, 225 to 45, etc.. The average of four orientations may be used if invariance to spatial rotations is desired. If G is large, the number of pixel pairs contributing to each element C_{ij}, in C would be low, and the statistical significance poor. On the other hand, if the number of gray levels is low, much of the texture features may be lost in image quantisation. Different statistical information can be obtained from the GLCM, Haralick [98] suggests 14 features describing the two dimensional pdf, these features provide information on:

- shift invariant statistics (important for the classification not to be a function of tone), Features such as Homogeneity, and Entropy (table 7.1) provide this information.
- the degree of smoothness of the texture. The closer the entries of C are closer to the diagonal, the smoother the texture. This is indicated by features such as contrast and

Feature Name	Definition
Angular second moment	$\sum_{i=1}^{G}\sum_{j=1}^{G} C_{ij}^2$
Entropy	$\sum_{i=1}^{G}\sum_{j=1}^{G} C_{ij}\log C_{ij}$
Contrast	$\sum_{i=1}^{G}\sum_{j=1}^{G} C_{ij}(i-j)^2$
Homogeneity	$\sum_{i=1}^{G}\sum_{j=1}^{G} \frac{1}{1+(i-j)^2} C_{ij}$
Correlation	$\sum_{i=1}^{G}\sum_{j=1}^{G} \frac{ijC_{ij}-\mu_x\mu_y}{\sigma_x\sigma_y}$

Table 7.1: The most commonly used GLCM features.

inverse difference moment (homogeneity) (table 7.1)

- correlation between gray level pairs (Correlation feature).

Table 7.1 summarises some of the co-occurrence features, where μ_x, μ_y, σ_x and σ_y in the 'Correlation' measure are the means and standard deviations corresponding to the following distributions:

$$C_i(x) = \sum_{j=0}^{G-1} C_{ij}$$

$$C_i(y) = \sum_{i=0}^{G-1} C_{ij}$$

Randen [175] uses angular second moment, contrast, correlation and entropy as his main features. While in the sonar domain, Blondel argues that the most relevant features are entropy and homogeneity [24]. This is deduced after a detailed study to assess the usefulness of more than 25 textural indices [23]. Cochrane *et al.* [52] also use these two features for their sidescan sonar classification. Kenneth Stewart [194] uses more than two features, arguing that each application has its own set of "suitable features", and it is very rare that any one measure would be able to describe subtle textural differences.

The major drawback in using GLCM features is that it is a very time consuming method. To calculate the features over several values of θ (usually 0°, 45°, 90° and 135°) and then to use a few values of δ, a long time is required (especially in the case of large images). In comparison

to other approaches, Randen [175] deduces that GLCM did not perform better than many filtering methods, and the computational complexity of this approach should be kept in mind.

To reduce the computational demands, a number of techniques can be used such as grey level quantisation, avoiding pixel by pixel segmentation and limiting the number of features. However, this might affect the features and make the classification worse. Clausi [50, 49] suggests using a linked list approach for the calculation of GLCM features. This approach avoids storing elements with zero probability, and uses a linked list instead of a large matrix containing all the cases (i.e. all pairs of pixels that are found at distance δ and angle θ). Each list node is a structure containing the two co-occurring gray levels (i, j), their probability of co-occurrence and a link to the next node on the list. The features are then calculated using the list (note that it is asymmetrical- also to save space) which is faster than using the whole matrix.

Due to the complexity problem, only a few inter-pixel distances are used for the GLCM calculation, which results in a major disadvantage, as the features selected would not be able to pick up textural properties across different scales.

GLCMs have been used successfully in many studies to analyse a few images rather than large datasets (as in the sonar surveys used here). As a starting point, we used photographic images (of sizes 100×100 pixels) instead of the sonar data, and GLCMs performed well on them. Using GLCM features on several selected (small) sonar areas, then clustering the GLCM features into classes proved to be a successful though very time consuming approach. In the case of large sonar surveys, much more processing power was needed to deal with the GLCM complexity, even though the linked list approach was used. The interpixel distances were chosen from $1 - 3$ and the orientations from ($0°$, $45°$, $90°$ and $135°$).

In conclusion, despite the success of the GLCM approach in image texture analysis, the drawback of this method is the time complexity especially when dealing with large sonar files. Filtering approaches were preferred for this study due to their efficiency in picking up multiscale textural features in much less time.

7.4 Gabor Filters

7.4.1 The Importance of Multiscale measures in Texture Classification

Multiple resolution techniques, such as filter banks, have the ability to decompose the image into relevant textural features that can be classified accordingly. Multi-channel filtering mimics the performance of the human visual system (HVS), which is extremely complex. A summarised explanation for the interconnection is as follows [51]: Receptor cells are found on the back inner surface of the eyes. From there, electrical impulses are passed through various nervous pathways through the optic nerve and finally to the visual cortex. Within the cortex, it is known that various cells perform different types of processing on the incoming signals. In cats, the tests done by Hubel and Wiesel [106] proved that individual simple cells are tuned to specific orientations. So as a bar is rotated through a visual field, individual simple cells respond only when a bar is within a certain range of frequencies. The range is not more than 30° for a given cell. Campbell and Kulikowski [35] carried this concept further and demonstrated that humans have both orientational and spatial frequency sensitivity. Experiments have shown that the frequency bandwidth of simple cells in the visual cortex is about one octave. This HVS multi-channel filtering model agrees with the design of Gabor filters and wavelet analysis. Results from the research done by Rao and Lohse [176] provide additional insight to this idea but from a human perception slant. Subjects were asked to classify different textured pictures based solely on their own perception. Results indicate that people use the following high level features for texture discrimination:

- Repetition (dependent on frequency).
- Directionality (representative of orientation).
- Image Complexity (relating to the consistency of the texture)

Hence to design a machine learning approach that would recognise textures, the starting point would be to imitate some of the functions of the HVS. Combining features that indicate degrees of directionality, repetition and complexity is an aim of many multiresolution filters used for texture classification (Gabors, Wavelets, QMF filter banks). The next section explains Gabor filters and covers some of the methods used to optimise their performance.

7.4.2 2-D Gabor Filters for Texture Segmentation

These filters have been shown to posses optimal localization properties in both spatial and frequency domain and thus are well suited for texture segmentation problems. Gabor filters have been used in many applications, such as texture segmentation, target detection, fractal dimension management, document analysis, edge detection, retina identification, image coding and image representation. A Gabor filter can be viewed as a sinusoidal plane of particular frequency (U, V) and orientation, modulated by a Gaussian envelope $g(x, y)$. It can be written as [214, 213]:

$$h(x, y) = g(x, y)e^{-j2\pi(Ux+Vy)} \tag{7.1}$$

where $g(x, y)$, the 2-D Gaussian can be defined as:

$$g(x, y) = \frac{1}{2\pi\sigma_g^2} e^{-\frac{(x^2+y^2)}{2\sigma_g^2}} \tag{7.2}$$

In many studies, $g(x, y)$ is assumed to be circularly symmetric for simplicity. The spatial-frequency response $H(u, v)$ of the Gabor filter is:

$$H(u, v) = G(u - U, v - V) \tag{7.3}$$

where:

$$G(u, v) = e^{-2\pi^2\sigma_g(u^2+v^2)} \tag{7.4}$$

The Gabor filter is essentially a bandpass filter centred about a frequency (U, V), with bandwidth determined by σ_g. The spatial extent, or scale, of $h(x, y)$ is also determined by σ_g.

σ_g determines the scale of the envelope of $h(x, y)$ but does not scale the centre frequency. The factors σ_g, U and V can be used to scale the Gabor filters over n scales, covering a number of orientations in each scale. The figures 7.1 and 7.2 give an example of scaling Gabor filters and using different orientations. To make calculations faster, the image is usually first correlated with the even part of the Gabor filter, then the odd part. The magnitude and phase of the correlated image can then be calculated.

After using a bank of Gabor filters, the resulting images are generally smoothed with a Gaussian filter matching the size of the Gabor filter at that scale. It has been well established that the Gaussian post-filter reduces the error in texture segmentation [28].

Figure 7.1: Gabor filters scaled over three scales, and are used to capture fine to coarse textural features.

Figure 7.2: Gabor filters in the same scale but with orientations 0 degrees, 45 degrees and 90 degrees respectively.

Although a bank of Gabor filters can be used to cover the total spatial frequency domain, this method suffers from data redundancy, especially if the whole range of frequencies and orientations is used. This problem has been addressed by many researchers, and the next section summarises some of the solutions offered.

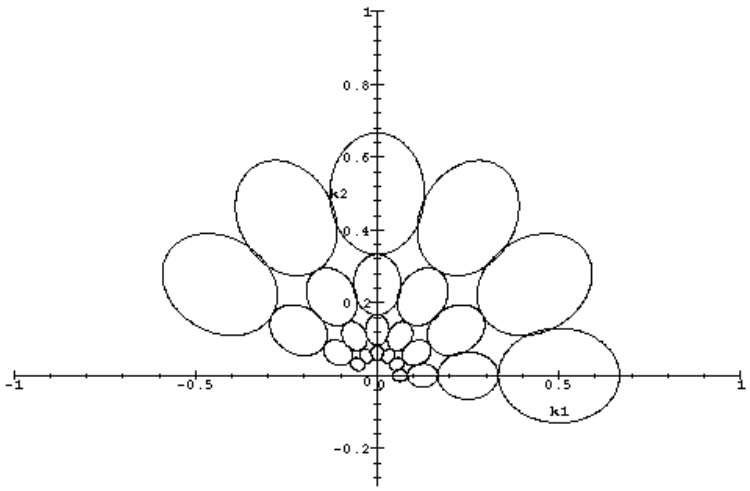


Figure 7.3: Half-value plot of the Gabor filters in the frequency plane tuned to different frequencies and orientations (30 degree resolution-each ellipse represents one of the filters used resulting in one filtered image).Gabor filters in this case divide the frequency plane over 4 scales and 6 orientations per scale.

7.4.3 Optimisation of Gabor Filters

Bovik [28] applies a simple peak finding algorithm to the image power spectrum (before filtering), this is used to guide the filter centre frequency to match that of the image. Both multiresolution amplitude and phase features are used for texture classification. By locating discontinuities in texture phase, information not available in the amplitude envelopes is made explicit. Variations in texture phase can give more information: small variations correspond to regions of smooth texture phase, and large variations correspond to sudden transitions from one dominant phase to another. Bovik's approach, however, is optimised for image representation and does not guarantee efficient feature separation. Weldon and Higgins [213] optimise Gabor filters with respect to feature separation. Segmentation error between known texture images is used to select the Gabor centre frequency. Their initial approach was to design a single optimal Gabor filter that can discriminate between two textures [215]. A measure of the total output power of the Gabor-filtered image is used to estimate its statistics assumed to be Rician

(of initially two textures), and then these statistics are used to vary the central frequency of the filter to offer maximum class separability (maximum expected average output energy for the two textures). The approach is developed in [214] to cover multiple Gabor filters and segment more than two textures. A candidate set of filters is first considered, then a multivariate Gaussian statistical model is used for the vector output of all the filter channels. The error is considered to be that of class separation between the Gaussians of each two textures, and the filters that give the maximum distances are selected from the candidate set. Bresch [30] also optimises filter banks for texture recognition. For the design of filter banks, the variances of the frequency components are maximised. The hyper- volume spanned by the normalised feature vectors representing different textures must be maximised as well. Thus, class separation is also used in this case to guide the design of the filter bank.

Many studies have been done on feature extraction from Gabor filtered images. A comparative study done in [126] compares several types of extracted features, including energy features, complex moment features, the filtered image itself (i.e. the results of the Gabor filtering with no averaging or analysis) and thresholded Gabor features. The deduction is that in most cases, post-Gabor processing increases discrimination effectiveness. Since energy features performed well in most studies, they will be used in this work.

Instead of using a Gabor filter, Kovesi [124, 125] suggests the use of Log Gabor filters. These filters use the log of a Gaussian instead of a Gaussian. Log Gabors are preferred over other wavelets since they allow arbitrarily large bandwidth filters to be constructed while still maintaining a zero DC component in the even symmetric part of the filter (a zero DC component cannot be obtained in Gabor filters over 1 octave). Log Gabors offer the advantage of including more harmonics with low frequencies. In addition to that, real textures often have a linearly decreasing log power spectrum [221], so Log Gabors are more adapted to this problem.

In the comparative study done by Randen and Husoy [175], it was deduced that the results obtained by optimising Gabor filters, following Weldon's methods (as mentioned above) were

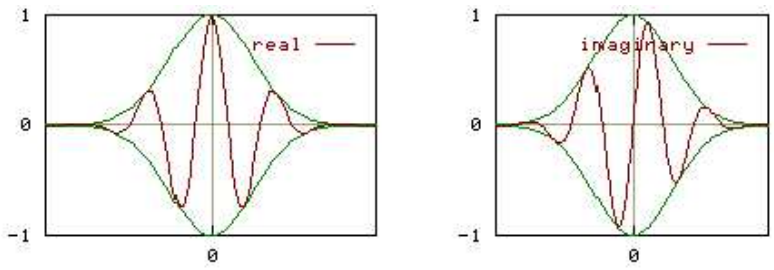


Figure 7.4: The real and imaginary parts of a 1-D Gabor filter

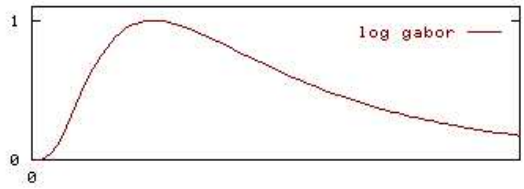


Figure 7.5: The Log Gaussian, used in Log Gabor filters.

fairly good, although some problems were observed when dealing with similar textures. It was also deduced that the commonly used dyadic decomposition (or octave band decomposition) is generally not superior to other decompositions, as the texture features for the images can be most prevalent in intermediate frequency bands. Thus, it could be a good idea to use a subband decomposition that is not dyadic (as in the case of the Wavelet Packet transform).

7.4.4 Application of Gabor Filters

This section describes the use of Gabor filters for texture image classification. The main aim is to use them for the underwater sonar images considered in this study. The classification system designed follows the following steps:

1. A log Gabor filter bank is generated for a determined number of scales and orientations at each scale. The starting centre frequency of the filter and its wavelength should be specified.

2. The image is convolved with the filter bank.

3. The filtered images are convolved with a smoothing Gaussian of the same size as the Gabor function at each scale.

4. An averaging window (of pre-defined size) is used to find the average around each pixel.

5. A Kmeans clustering algorithm is used to cluster the resulting pixels into two classes.

Since texture boundaries are rarely available in sonar images, the method was applied (as a first step) to photographic images of different textures. Figure 7.6 shows the results for the following selected parameters:

-filter orientations: 3 angles for each scale.

-the number of scales is 2.

-the moving window size is selected to be only 5 pixels.

In this figure, the small features of these images are classified whereas large features are ignored. Varying the filter parameters and the size of the averaging window lead to different classification results as observed in figure 7.7. In this figure, a larger window and more scales are considered (4 scales, window=12 pixels), picking up larger features. As noticed in these two examples (figures 7.6 and 7.7), the classification is very image dependent, so a set of parameters that work well for one of the images may not be the ideal set for another image. [1]

7.4.5 To Use or Not to Use

As a general conclusion to this section, Gabor filters have both advantages and disadvantages when used for sonar texture analysis. The following is a summary, focusing on the use of Gabor filters for the datasets used in this work:

[1]Note that this experiment is just a simple experiment to observe effect of the size of the filters chosen on the image features picked, more on this and experiments with sonar data are given in chapter 8.

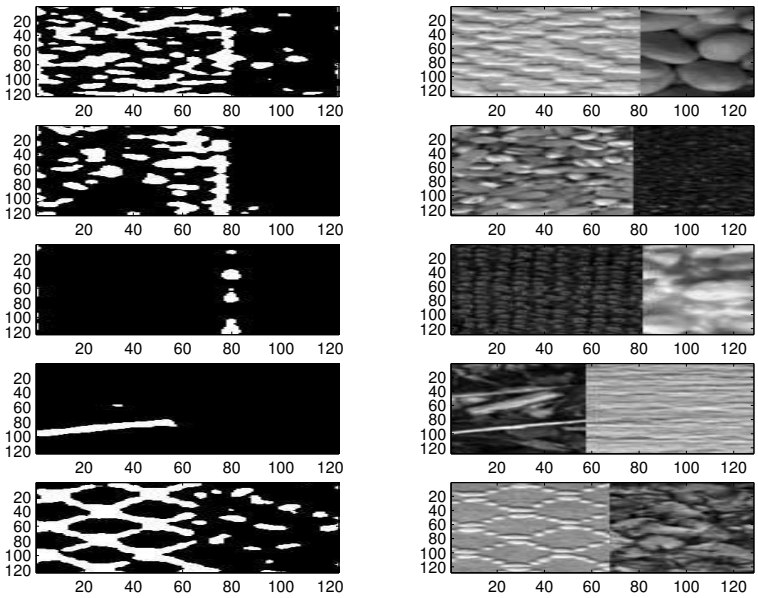


Figure 7.6: Results of using log Gabor filters to classify the photos, the filter orientations were 3 for each scale, the number of scales is 2, the moving window size is selected to be only 5 pixels, this is why the small features of these images are classified whereas large features are ignored.

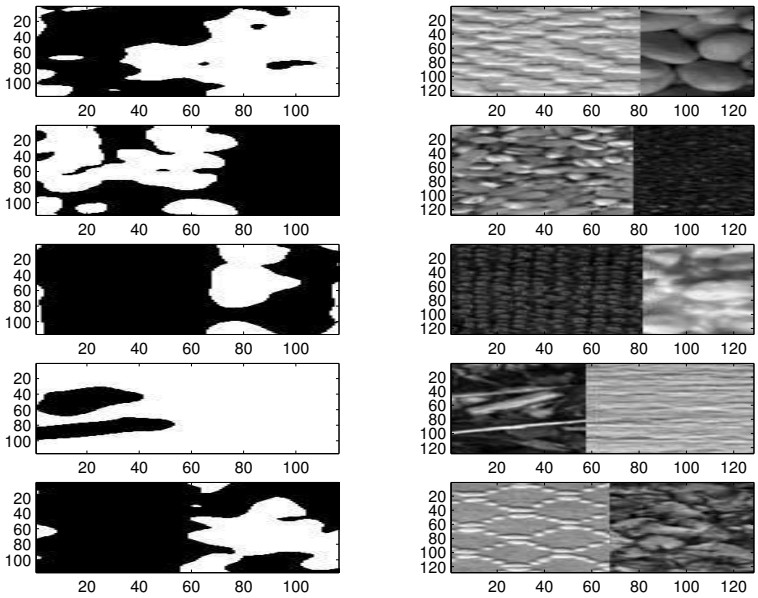


Figure 7.7: Results of using log Gabor filters to classify the photos, the filter orientations were 3 for each scale, the number of scales is 4, the moving window size is selected to be 12 pixels. Large features are picked up in these images, but some images are better classified than others.

- The sonar datasets used in this work are quite diverse, which makes it difficult to select optimal orientations that can provide sediment-class separation. The data does not generally provide dominant periodic features (as in the photographic images of figures 7.6 and 7.7), and the sonar files are quite large. Thus, feature selection is done after the filtering stage and the whole set of Gabor filters is used in filtering. Feature selection uses given grab samples, or validated areas (if available) as training data.

- In Randen [175], the Gabor filter was outperformed in texture classification experiments by other filtering techniques (wavelets in a few images). The author argues that a filter with optimal joint resolution in both frequency and spatial domains (Gabor) does not have to be the optimal filter for classification.

- Being not necessarily dyadic, Gabor filters can provide better resolution at high frequency scales, which might be needed for class separation in the sonar datasets.

7.5 2-D Wavelet Analysis for Texture Classification

Multiresolution techniques tend to transform an image into a representation containing both spatial and frequency information. Many methods were developed to make the best of multiresolution properties, including Gabor filters, Gaussian pyramids and subband filtering. Wavelet analysis provides a more formal, solid and unified approach to multiresolution approximation [63, 142] as presented in chapter 4. In [186], statistical features of textures are studied across multiple scales, and it is concluded that a texture can be characterised by these features.

1-D Wavelets are presented in detail in chapter 4. This section presents an extension of 1-D wavelets to the 2-D case, focusing on the use of wavelets for texture analysis.

7.5.1 2-D Wavelet Analysis

An extension of 1-D wavelet expansion is used for the construction of a two-dimensional multiresolution approximation of an image. What follows is a mathematical explanation of the method.

Instead of using only a wavelet ψ and a scaling function ϕ, to obtain one high pass detail signal and one low pass approximation at each scale (as in section 4.2.2), products of these functions can be used to obtain a 2D multiresolution.

The scaling function ϕ is associated to a 1-D approximation $\{V_j\}_{j\in Z}$. Let $\{V_j^2\}_{j\in Z}$ be the separable 2D multiresolution defined by $V_j^2 = V_j \otimes V_j$, and W_j^2 be the detail space corresponding to the orthogonal complement of the lower resolution approximation space V_j^2 in V_{j-1}^2:

$$V_{j-1}^2 = V_j^2 \oplus W_j^2 \tag{7.5}$$

To construct a wavelet basis of the two dimensional detail space W_j^2, Mallat [142] defines three wavelets:

$$\psi^1(x) = \phi(x_1)\psi(x_2) \tag{7.6}$$

$$\psi^2(x) = \psi(x_1)\phi(x_2) \tag{7.7}$$

$$\psi^3(x) = \psi(x_1)\psi(x_2) \tag{7.8}$$

He proves that these wavelets form an orthonormal basis of W_j^2. The three 2-D wavelets extract image details at different scales and orientations. The original image a_j is thus represented by a set of subimages at several scales. The resultant images at each scale are the approximation image a_{j+1} and the three detail images d_{j+1}^1, d_{j+1}^2 and d_{j+1}^3. The separable wavelet expressions in equations 7.6, 7.7 and 7.8 imply that in the 2-D frequency domain:

$$\hat{\psi}^1(w_1, w_2) = \hat{\phi}(w_1)\hat{\psi}(w_2) \tag{7.9}$$

$$\hat{\psi}^2(w_1, w_2) = \hat{\psi}(w_1)\hat{\phi}(w_2) \tag{7.10}$$

$$\hat{\psi}^3(w_1, w_2) = \hat{\psi}(w_1)\hat{\psi}(w_2) \tag{7.11}$$

$\left|\hat{\psi}^1(w_1, w_2)\right|$ acts as a filter that selects low horizontal frequencies w_1 and large vertical frequencies w_2 (the detail in the filtered image d^1_{j+1} corresponds to horizontal edges). $\left|\hat{\psi}^2(w_1, w_2)\right|$ is large at high horizontal and low vertical frequencies (so the detail in the filtered image d^2_{j+1} corresponds to vertical edges). $\left|\hat{\psi}^3(w_1, w_2)\right|$ is large at high horizontal and vertical frequencies (hence the detail in the filtered image d^3_{j+1} corresponds to high frequencies in both directions-such as corners).

To construct a fast two-dimensional wavelet transform, Mallat [142] uses the conjugate mirror filters $h[m]$ and $g[m]$ corresponding to the 1-D wavelet ψ. Images are decomposed by a cascade of four-filter banks. Figure 7.8 is an example of using separable filters for the decomposition of an image a_j. The down arrow means that the image is subsampled by keeping only the pixels where both coordinates are even. Signal processing is performed on the low resolution image, and the detail images at each scale. The processed signal is recovered by another cascade of filter banks (figure 7.9). The up arrow means that the size of the image is doubled by inserting black (zero) rows and columns of pixels between every row and column of the source image. Figure 7.10 gives an example on separable wavelet transforms and decomposing an image at several scales.

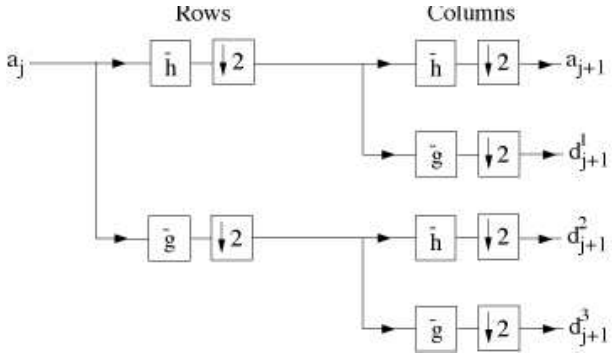


Figure 7.8: Decomposition of an image a with 6 groups of 1 dimensional convolutions and subsamplings along the image rows and columns.

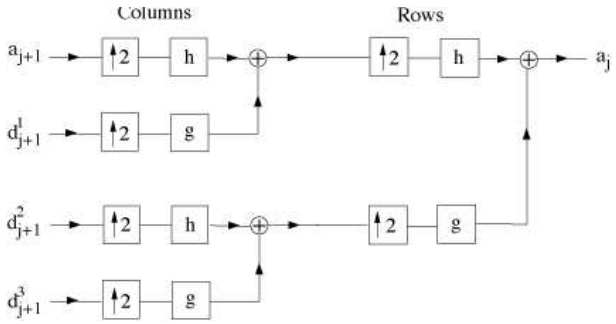


Figure 7.9: Reconstruction of an image a by inserting zeros between the rows and columns of the lower resolution detail and approximation, and filtering the output.

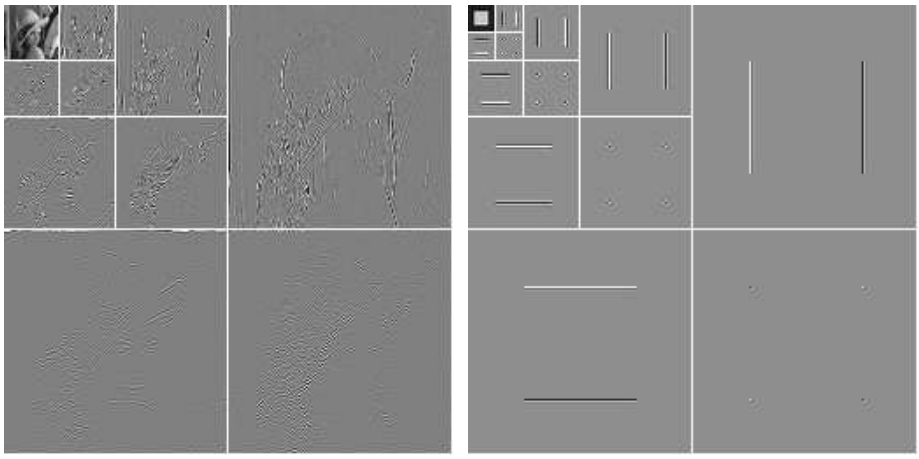


Figure 7.10: Seperable wavelet transforms of two images, decomposed respectively on 3 and 4 octaves. In each scale, the horizontal features are selected in a subimage, the corners in another, and the vertical features of the original image in a third subimage (This is easily observable in the square image on the right). The final upper image represents what is left after filtering over all the scales, and thus presents a low frequency component. Black, gray and white pixels correspond respectively to positive, zero and negative wavelet coefficients. Since images are downsampled at each scale, the resulting images are two times smaller than the original at that scale. The figures are an example from Mallat's book [142].

7.5.2 2-D Wavelets for Texture Classification

The wavelet transform energy at a certain scale j is high if the wavelet orientation and scale match particular features of the image at that scale. Thus, wavelet coefficients at various scales can be used to discriminate between textures by observing their statistics. Figure 7.11 shows an image with two different textures, the right part is that of a piece of wood having more horizontal lines than the left part, which is of woodchips. By observing the multiscale wavelet analysis of this image (using a Symmlet wavelet), discriminating features can be obtained between the two textures. Horizontal filters can pick up the horizontal lines across appropriate scales. Statistics of the filtered images, such as local energy or standard deviation can be used to classify the two textures and segment the boundary between them.

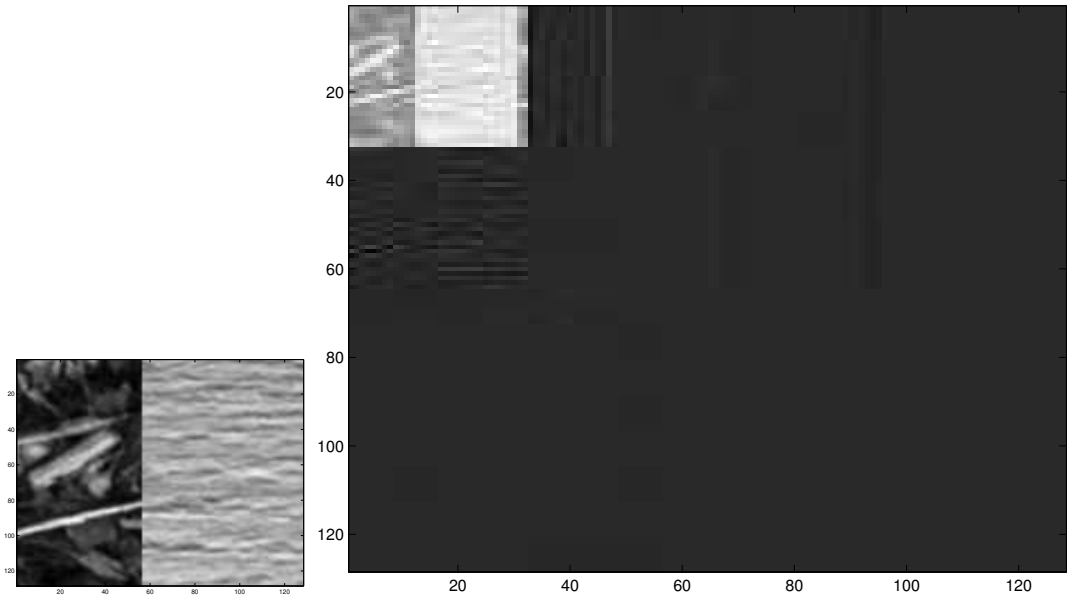


Figure 7.11: The left image is that of two different textures, of wood chips (right) and wood (left). The right one is a wavelet decomposition using a separable decomposition starting with a Symmlet wavelet.

However, one of the problems with the DWT (Discrete Wavelet Transform), is that due to the dyadic nature of the scales adopted, some textural features are omitted since they lie between two scales. The work by Chang and Kuo [39] indicates that texture features are

most prevalent in intermediate frequency bands and that octave decomposition is not optimal. Several extensions to the DWT to improve selectivity have been introduced with a finer discretisation of the scale parameter. One of the extensions uses a tree structured decomposition (the wavelet packet), which performs a DWT on every detail image obtained at every scale [55, 217, 129]. The 1-D version of the Wavelet packet was discussed in chapter 4. As for its extension to 2D, it follows the same explanation of the normal DWT, and uses separable wavelets, so that 1-D filtering is actually done (twice for each scale) using a mother wavelet and its corresponding scaling function (figure 7.12). The wavelet transform basis is a subset of the wavelet packet representation, which provides better frequency localisation especially at high frequencies.

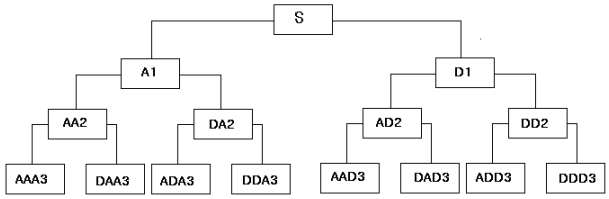


Figure 7.12: The Wavelet Packet Transform, where each image at each scale is filtered by a highpass (to obtain detail) and a low pass filter.

One of the main problems in applying wavelet packets for texture analysis is that of redundancy. Among all the resulting wavelet bands (or images in this case), a basis has to be selected which gives the greatest discrimination between the classes being studied. Chang and Kuo [39] developed a tree-structured wavelet transform algorithm for texture classification, which is similar to the wavelet-packet best basis algorithm by Coifman and Wickerhauser [55]. This algorithm uses entropy as a measure to select the subbands that best represent the signal (or original image). Thus a wavelet basis becomes an *adaptive* wavelet basis [198, 202, 219, 55]. The selection of a 'best basis' was discussed in chapter 4.

Tang and Stewart [201] compare Wavelet packets with Fourier features for sonar imagery and note that the latter performs better than the former. In their case, they did not select a

best basis from the wavelet packet transform, but derived features from all resulting images by using principal component analysis. Their expansion was dyadic in the wavelet packet case. In the Fourier transform, they localised areas in the transform where they measured the energy for a selected frequency band (they select a radius and angle in the 2-D frequency plane). So as a result, their Fourier frequency localisation was better than the wavelet packets and lead to a better classification. This does not mean that wavelets are necessarily a more complicated method than Fourier and lead to worse results! Wavelets can be used with a better frequency localisation, and definitely provide much better spatial resolution.

Another way of overcoming the discretisation of scale is to consider a non-dyadic representation, where the wavelets can be scaled according to the features that are being captured in the image. This is very similar to the Gabor approach mentioned in the previous section. It is expected to result in a better classification, especially in the supervised classification case, where specific textures are studied, the filters matched to them, and then used to find these textures in other images. Of course, the disadvantage of this approach is in time consumption.

A disadvantage in using wavelets, whether DWT or wavelet packets, is that only three orientations are selected at every scale (horizontal, vertical and diagonal). For some applications, more orientations are needed to provide class discrimination. For example, Munjanath and Ma [153] compare Gabors with DWT and deduce that Gabors work better for their image retrieval problem, mainly due to considering more orientations. Recently, a new complex wavelet has been proposed by Kingsbury ([119]) which gives a fast way of computing Gabor like wavelets. The similarity is in the fact that Complex wavelets can use multiple orientations at each scale. They have separate subbands for positive and negative orientations, and can distinguish between lines at 45° and -45° for example. The Complex Wavelet Transform attains these properties by replacing the tree structure of the conventional wavelet with a dual tree. At each scale, one tree produces the real part of the Complex Wavelet coefficient, while the other produces the imaginary parts. By using even and odd filters alternately in the trees it is possible to achieve overall complex impulse responses with symmetric real parts and antisymmetric

imaginary parts. The filters are designed to give a number of desired properties including strong discrimination between positive and negative frequencies.(Discrimination between positive and negative frequencies is not possible while using real wavelets). This important property means that separable Complex Wavelets can separate between information in the first and second quadrants in the two-dimensional frequency response. The main advantage as compared to Gabor filters is their speed of computation.

In addition to the orientation advantage over DWT, Complex Wavelets are approximately shift invariant (meaning features in subbands are more robust to translations in the image), whereas the DWT is sensitive to the precise position of image features with respect to sampling grids used by the decimators at the filter outputs. This feature gives Complex Wavelets an advantage when applied to many areas of image processing such as motion estimation and de-noising [120, 139].

For texture classification, de Rivaz [67, 66] has compared the performance of the DWT (with severe downsampling), the NDWT (similar to the DWT but the bands are averaged over all positions of the origin to solve the translation invariance problem) and the Dual Tree Complex Wavelet transform (DT-CWT). He deduces that the DT-CWT performs better than the other two for all the photographic images considered. The reason is the improvement in directionality detection and solutions to aliasing problems. The next section addresses one of the main questions that arise from having both phase and amplitude information after filtering an image (whether using Complex Wavelets or other types of wavelets).

7.5.3 How Important is Phase Information in Texture Problems?

In the study of image reconstruction from partial Fourier transform information, Oppenheim shows several examples in which an image is reconstructed only from its original phase information and arbitrary magnitudes, the results [161, 160] evidently confirm that phase contains most of the content of the image. However, in [199], the authors show several cases where the

global amplitude spectrum is also essential for specifying the particular content of natural images. Thus, their conclusion is that although the average amplitude spectra of different natural images may be similar in their overall form, a realistic description of the amplitude spectra must also incorporate the particular way in which the energy is distributed across different orientations, that, added to the phase information would give sufficient description of the image. What follows is a brief description of phase features and a discussion of their use in texture analysis.

Looking at a 1-D Fourier expansion of a signal, the phase is constant across the main features of the signal, such as edges. Using a 2-D image analysis, if the local phase information was captured, most of the information on the position of corners and edges will be captured. This provides a very efficient edge detector used by Kovesi [123] who proves that the results are far better than commonly used edge detectors (like the Canny edge detector, for example). Kovesi [123, 124, 125] describes a phase congruency measure for signals and images, that can capture the presence or not of a feature. He uses two filters in quadrature (similar to quadrature mirror filters used in wavelets), and calculates the phase congruency of all the image locating the main features. More on this method is found in [124, 125][2]. The rich phase information has lead many researchers to investigate textural features obtained from phase analysis [81, 75]. So far, however, all results in textural analysis research seem to show that the textural content of the phase information is low [201].

The work done on natural images usually focuses on images with large areas of constant amplitude, where a phase shift between these areas indicates an image feature (a transition between sky and land, etc...) The general shapes of Fourier magnitudes for these natural images are quite similar with most energy concentrated in the low frequency area. Thus, most structural information is carried by phase features rather than amplitude features. For reconstruction, phase plays the major role.

For textural images, phase information is a special property of each of the individual texture

[2]website http://www.cs.uwa.edu.au/~pk/

samples, whereas in the previous case, it indicated features in a mostly constant amplitude image. Tang *et al* [201] argue that statistically the phase signals for all texture samples are essentially random regardless of texture classes and do not offer a significant discrimination power for texture signals. Absolute phase values, in their opinion, are similar to the noise signal that we normally try to remove. However, phase differences between different frequencies reflect the relative positions of harmonic functions and would offer information on texture. This is difficult to obtain, since, no matter how small the energy of the frequency component is, its phase value can be anything in a period of 2π. This equal value property renders the phase information useless for natural texture classification. For synthetic textures, however, phase information can be extracted (example in [75] where phase information is extracted from Gabor filters).

In this study a few experiments were done on series of sonar pings, and Gabor filters were used to observe phase textural features, such as the average phase and its standard deviation. As observed as in [201], phase derived features did not offer much discrimination between the underwater seabed classes. The reason behind this is probably the randomness of backscattered sonar images (due to random scatterers in the sea bed).

7.5.4 Textural Features and Wavelet-Type Selection

To obtain features which reflect scale dependent properties, a feature is extracted from each scale separately. An appropriate quantity is the energy [186, 39, 153]. To summarise features in this work, a window of size ($[a, b]$; $N = a * b$) is selected around each pixel and the energy of the wavelet transform is calculated as follows:

$$S_{i,j}(x, y) = \frac{1}{N} \int_1^a \int_1^b W^2(x, y) dx dy \tag{7.12}$$

This is done for all sub-images and the resulting (energy) images are used for classification. Several studies investigate alternative measures, but no general conclusion in favour of a par-

ticular measure can be drawn out of them. Laine and Fan [129] compared energy and entropy and found the latter to be less suitable. Munjanath and Ma [136] used absolute values and found that adding a second variance measure improves performance. So another feature we will use is the standard deviation of Wavelet Transform coefficients per window:

$$SD_{i,j}(x,y) = \text{standard deviation } (W(x-a:x+a,\ y-b:y+b)) \tag{7.13}$$

Selecting both energy and standard deviation features per scale and per pixel (or window of pixels) leads to a large number of features. The feature selection techniques described in chapter 5 will be used to select features from the resulting images.

In [85], many wavelet filters are compared (although the authors limit their work to separable and real wavelet filters). These filters include Haar filters, Daubechies of different sizes and Battle-Lemarie filters. The general conclusion is that the filter (or wavelet) type did not have a major effect on the texture classification results (as long a wavelet that can pick up the fine features of the image is used). In this study also, the type of filter did not have a major effect on the results. Other factors were much more important, such as the number of subbands, the starting size of the filter and the size of the averaging window.

7.6 More Texture Classification Techniques

As texture modelling and classification is a very wide subject, this section sheds the light on a few directions that this area of research has taken. The discussion will be divided into two sections. The first section describes filtering approaches, including methods using statistical models that benefit from the multiscale dependency of filtered images, as well as approaches that deal with varying viewpoint and illumination issues and the use of class separation to guide the selection of filters. The second section explains approaches that use methods based on Markov Random field modelling of local pixel neighbourhoods in images, as well as other methods that use local information directly without any filtering (such as [20]).

7.6.1 Filter Based Methods

In addition to wavelet and Gabor filters explained in this chapter, several other filters have proved to be efficient for texture classification problems. These include Quadrature mirror filters, which are often used in image coding, and can be applied to texture segmentation problems [175]. This is, like the class of wavelet filters, a very broad class of filters, incorporating both infinite impulse response (IIR) and finite impulse response filters (FIR). Based on the energies in the subspaces of the filtered images, Chen and Kundu [128] classify sets of textured images. The discrete cosine transform DCT which is popular in image coding due to its good performance and fast implementation, can also be used efficiently for texture feature extraction [175]. However, the main problem with several filtering approaches (including wavelets and Gabor filters), as noted by De Bonet and Viola [26, 64] is that several of them fail to model textures with long scale structure that is neither local in space or frequency. Other approaches that use simple local measures at different scales implicitly assume that textures are self similar at every scale. These assumptions do not allow classification methods to benefit fully from the multi-resolution features provided by filtering approaches and several alternative methods based on statistical modelling have been suggested.

In [26], the authors argue that there is a strong cross-scale dependence between filter coefficients between the wavelet coefficients of an image. The assumption of Independence has also been challenged in the work of Zhu *et al.* [225] and Buccigrossi and Simoncelli [31]. De Bonet and Viola [26] model the joint distribution in an image as a coarse to fine probabilistic process, where decisions made at high frequencies are conditioned on those made at low frequencies. The conditional distributions are estimated from example images in a flexible, non parametric fashion that is able to capture complex and multi-modal distributions. Gaussian mixture models have been used extensively to model wavelet coefficients across multiple scales. An example is in [212], where the authors use a finite generalized Gaussian mixture to model the overall distribution of the coefficients. An unsupervised learning procedure is developed to

quantify the histogram through a tripled adaptive algorithm including detection of the number of local kernels, approximation of the shape of local kernels, and estimation of model parameter values. Bouman *et al.* [27] use a Gaussian autoregressive model for the observed image. The MAP (maximum a-posteriori) estimate of the region image at the coarsest resolution is approximated first, using iterated conditional modes. The result is then used as an initial condition for segmentation at the next finer level of resolution, and the process is continued until individual pixels are classified at the finest resolution. Noda *et al.* [157] use Markov Random Fields over a wavelet pyramidal lattice to capture significant intrascale and interscale statistical dependencies in the observed image.

Hidden Markov Tree Models are parametric models with discrete hidden states and Gaussian observable variables that can be used for separable 2D wavelet transforms [181]. Crouse *et al.* [59] proposed a Hidden Markov Tree (HMT) as a statistical model capable of modelling the multiresolution properties of wavelet coefficients. This model also captures the joint statistics of the wavelet transform (including *persistence* across the branches of the wavelet tree-which describes how large coefficient values cascade across the tree). Choi *et al.* [46] adapted this model for texture segmentation, since it is quite simple but captures singularities. Lu *et al.* [134] introduced an HMM-type global dependency into the model, termed the HMM-HMT, achieving an improved segmentation performance. The HMM-HMT model was modified by Won *et al.* [222] to be applied to the Complex Wavelet Transform, thus benefiting from the shift invariant properties of Complex wavelets, and demonstrating competitive performance. Similar to a HMM on a tree, Sajda *et al.* [166] develop a probabilistic network over image spaces. The model developed employs a multiscale pyramid decomposition to factor images across scale and a network of tree structured hidden variables to capture long range spatial dependencies. Bayesian networks have also been used to model texture images. Stainvas *et al.* [192] introduce a Bayesian network classifier which has a similar form to the HMT models, as it has a mixed discrete hidden state and continuous observation variable model. Wavelet coefficient independence is assumed within a scale. Gabor filters are used to capture orienta-

tion properties, and to find the parameters of the model using the Expectation Maximisation algorithm (EM). The image is then classified using a Bayesian classifier.

Other methods that use a probabilistic model for filter outputs include work by Kunishi and Yuille [122] and Heeger and Bergen [101]. In [122], the authors use several filters capturing scale, edges and colour . Then, they look at the empirical joint probability distribution of these filters for each class of images. The classifier used is a Bayesian classifier which performs well for the set of images provided, but has a few drawbacks. The main one is that the amount of data needed for training grows exponentially with the number of filters used. In [101], a bank of orientation and spatial frequency selective filter is used, and the authors describe a method of synthesizing textures by matching distributions (or histograms) of filter outputs.

A problem that can cause a variation of observed textures from the same class is that of varying viewpoint and illumination. Leung and Malik [130] address this problem and suggest the use of 3D textons. The idea of 3D textons is to form a small vocabulary of micro-structures made of the most informative filter response vectors with respect to a certain texture class. Once this vocabulary is built, a surface of any material is represented as a spatial arrangement of symbols from this vocabulary. Varma and Zisserman [210, 209] also tackle the problem of varying illumination and viewpoint, and base their model on statistical distribution of clustered filter responses. By comparing several filters, the authors conclude that rotationally invariant filters outperform other types of filters (such as those used by Leung and Malik). The classification method is similar to that in [130]. However, clustering is done in a much lower dimensional space and weak isotropy is used to compensate for viewpoint changes, which leads to better classification rates.

In section 7.4.3, methods of optimizing Gabor filter design with respect to class separation between several textures were highlighted [214, 30]. This guides the selection of filters from a filter set, or the design of filters that can best capture discriminating features. In section 7.5.2, wavelet packets were summarised as a method that can overcome some of the drawbacks of using a wavelet basis. Adaptive approaches were also explained, which provides a way to

overcome the redundancy of wavelet packet bases and select the basis that allows for the best classification.

7.6.2 Using Local Pixel Neighbourhoods

The supremacy of large scale filter banks for texture synthesis was brought into question by Efros and Leung [80], who demonstrated that better results can be obtained by using local pixel neighbourhoods directly. A Markov random field model (MRF) was assumed, and the conditional distribution of a pixel given all its neighbours was estimated by querying the sample image and finding all similar neighbourhoods. Zelensy and Van Gool [224] also used MRF for image synthesis. They used subsets of all available cliques present in a neighbourhood, to sample from a parametric MRF. Although interest in MRF models for tackling image processing problems can be traced back to the work of Abend [2], only recently have appropriate mathematical tools for exploitation of the full power of MRF in image processing been developed. Gibbs distribution has been used in several studies to characterise the MRF, examples are in [58, 90, 53]. Paget and Longstaff [170, 169, 168] also use MRF, but present a model of a non-parametric multiscale MRF that can capture characteristics of a wide variety of textures. They propose a new texture classification scheme based on a significance test, without requiring a set of predefined classes. Varma and Zisserman [208] also use the joint distribution of intensity values over extremely compact neighbourhoods for texture classification and prove that the performance surpasses several of the state of the art filter bank based classifiers such as those presented in [209, 60, 131].

In [20], the authors use image patches representing different textures directly (without feature selection) as inputs to a Multilayer Perceptron, that learns from the training data and can classify images without being affected by minor pattern variations (such as rotation, scaling or blurring). Their argument is that the MLP works in a way like our visual system, so they avoid the calculation of a feature set. This study gives more evidence that local image properties

can be used directly for texture analysis.

7.7 Conclusions

This work uses texture features as an additional feature that could help in sonar survey classification. There is a difference between the aim of this work, which is to classify a survey (given a few grab samples in most cases), and several texture based sonar image segmentation studies. The images considered for classification are those obtained after processing and correcting for several physical factors. Thus, they are not exactly 'raw' sidescan sonar images. The training stage could suffer from a lot of errors, as training areas around grabs are automatically selected to be of the same type. The aim of this chapter was to give an overview of methods used for texture classification, and to view them from a sonar analysis point of view, deducing which methods can be applied for this work. Some of the main points are the following:

- An efficient texture classification method is needed, which can extract features from series of sonar pings without being affected by depth variations. It must not be very dependent on amplitude variations of seabed areas, and should recognise similar textures of seabeds at different depths.

- Since the amount of data analysed in this study is rather large, statistical methods are a more appropriate selection than model based techniques. Although model based techniques provide a lot of advantages in many image texture studies, models usually require a training stage with known data. In this study, a method which is less dependent on the training stage is required, since many sonar surveys have no validated data that can be used for training.

- Due to the great advantage of multiresolution methods, Wavelet analysis and Gabor filter analysis are good candidates for this study.

- Computational complexity must also be taken into consideration. Filtering approaches (Gabor filters and wavelets) perform well here since the computational demand is much lower than other statistical methods such as GLCMs.

Chapter 8

Sonar Texture Analysis - Applications to Datasets

8.1 Background

After presenting a review of textural analysis methods in chapter 7, this chapter applies these methods on two sonar datasets. The main aim of this chapter is to experiment with several filter based classification techniques and then observe the effect of varying the filter parameters. The results from these experiments will aid in selecting filters for the classification of the sidescan bathymetric datasets in the next chapter. Although the datasets used in this chapter are sidescan sonar datasets taken at a different frequency from the sidescan bathymetric sonar datasets, they present areas that are groundtruthed (Pace's dataset), or areas that can be distinguished by visually observing the texture differences (the coral reef dataset). Thus, these datasets can be used to understand the importance of each parameter uesd to distinguish between seabed classes. The two datasets considered in this chapter are:

- A sidescan dataset (Pace's dataset, see section 2.3.1) of 48 kHz. The series of pings were displayed consecutively and used as a sonar image [167].

- A sidescan sonar dataset of artificial reefs in the seabed (section 2.3.1).

The sidescan bathymetric sonar datasets which will be analysed in chapter 10 are much larger datasets where the only evidence of seabed type is a series of grab samples in nearby locations. This chapter is a preliminary stage to test the texture techniques before moving to more complex and much larger datasets.

This chapter does not aim to compare several texture techniques and deduce the best one, but observes filter-based texture methods that can be useful for sonar classification. Note that the main aim of this work is not to classify sidescan datasets by using texture techniques, but to use textural features as an additional feature, selected per window, and applied on bathymetry-corrected (processed) signals. This work focuses on finding a method that can deal with sonar data per survey (rather than per sonar image), and select features that are discriminant, given several varying parameters, such as sonar motion, depth and resolution.

Earlier classification results were presented using 1-D methods of analysing single sonar pings in chapter 6. The advantage of using 2-D textural features is that the method can draw boundaries between different types of seabeds, whereas the 1-D methods do not localise boundaries correctly.

8.2 2-D Sonar Classification Scheme

The 2-D method that will be followed to analyse both sidescan datasets in this chapter and the sidescan bathymetric datasets in chapter 10 can be summarised in the flow diagram in figure 8.1. The steps (in a very general form) are the following:

- Specifying parameters for classification (User Defined Input). These parameters vary with the method used, but generally include the size of the averaging window, the size of the filters (wavelets or Gabor filters) and the number of scales and orientations used.

- The preprocessing stage for the correction of data.

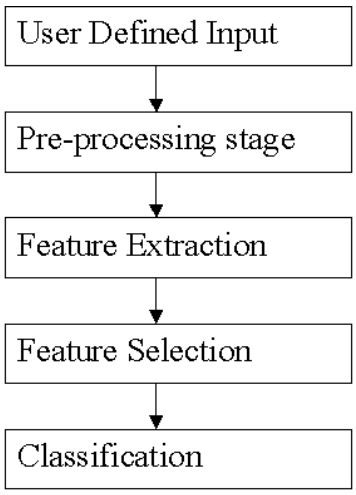


Figure 8.1: Flow diagram of the classification scheme

- Feature extraction, where features vary with the type of dataset used.
- Feature selection from the extracted feature vectors. Several methods are used here depending on the availability of ground-truthed data (the methods are detailed in section 5.3).
- The classification stage, which can be supervised (if ground-truthed data is available) or unsupervised. Then the final classified datasets are displayed.

8.3 Pace's Dataset

8.3.1 Introduction

The dataset used in this section is the same sidescan dataset used previously for the study of 1-D signal analysis. More details on the dataset can be found in section 2.3.1. The advantage of using this dataset is that the true type of each section is known, so an assessment of misclassification can be made. The misclassification error guides the selection of window size, filter size, the number of filter subbands, and in the case of Gabor filters, the number of useful

orientations. The interclass distance can also be measured to give an idea on class-separation.

Another advantage is that since the data types are known in advance, training data can be extracted from the dataset and used to train a neural network such as an MLP, or a nearest neighbour classifier.

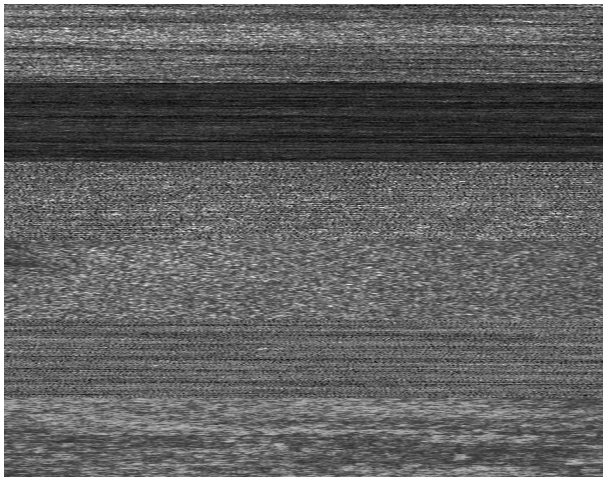


Figure 8.2: Series of pings from Pace's dataset. The classes are (from above) clay, gravel, sand, stone, mud and rock. The series of pings are displayed consecutively. The vertical direction is *across* track and the horizontal is *along* track.

Although this dataset was used by many authors (Pace [167] and Tamsett [200]), it is difficult to compare their classification methods, since in each case, different feature extractors were used as well as different classifiers. In addition to that, the classification techniques used in this section are 2-D techniques whereas the techniques used by both Pace and Tamsett depend on the 1-D power spectrum of each ping. The approach here follows Carmichael [37] in the selection of 2-D windows for measuring properties of the areas. The advantage over 1-D methods is that sediment boundaries can be observed and used to segment seabed areas. A 2-D data set is synthesised from the set of sonar data in Pace's dataset. 120 pings of each of clay, gravel, sand, stone, mud and rock are used to form a sonar image to be analysed (figure 8.2). Each type is assigned a certain class number for classification. Preprocessing this dataset was done online using a time varied gain filter but the details of the online corrections were not

provided.

This section compares the performance of both the DWT and Gabor methods on Pace's dataset. The area is divided into windows of user-defined size. A compromise must be made between large windows that provide a lower resolution and good classification rates, and small windows, where a better resolution is obtained on the expense of not picking up some of the features, leading to a worse classification than large windows in some cases. Feature selection uses both PCA and backward elimination algorithms. Due to the presence of ground-truthed data, the classification is a supervised one using the KNN algorithm (section 5.4). The dataset is divided into training and testing data (for both DWT and Gabor filters). The training data is used to assign centres for the KNN clusters. The whole image is filtered across $nscale$ scales, and on each scale, windows of sizes $(winh, winv)$ are used, the last line of windows of each series is used as a training set for classification.

8.3.2 DWT Analysis

In chapter 7 section 7.5.1, the two dimensional wavelet transform was introduced. By using separable wavelets, the 2 dimensional wavelet decomposition of an image can be obtained. The scale used in this section is a dyadic one. A Symmlet wavelet is used to construct the wavelet transform. The use of DWT for the classification of Pace's dataset will be explained following the flow chart in figure 8.1 in the following steps:

User defined parameters The user defined parameters that were varied in this case were the number of bands in the wavelet transform and the size of the averaging window. The number of bands was varied from 2 to 5 and the sizes of windows were chosen from the following : (30,30), (40,40), (50,50), (10,100), (10,300) and (20,300). These numbers are vertical and horizontal pixel numbers respectively.

Preprocessing Done online as mentioned in the previous section.

Feature Extraction The features used from the Wavelet transform are the energy and the standard deviation of wavelets within a user-defined box (as defined in section 7.5.1). Thus, feature-images are calculated from a starting image and decomposed along $norient$ orientations ($norient = 3$, in the case of DWT), and $nscale$ scales, the number of resulting feature-images is $numb = 2(norient \times nscale)$. These feature images are transformed into feature vectors to be used for classification.

As the size of the wavelets was increased, boundary artifacts were created in the image since the wavelet was not adapted to deal with boundary problems, resulting in high amplitude boundary pixels. To overcome this problem, a minimum filter was employed around the boundaries to replace very high amplitude pixels with the minimum of a neighbouring box. The parameters for this algorithm were the size of the box and the threshold at which high amplitude pixels should be replaced.

Feature Selection, Classification Results and Figures First, the window means were used alone as a feature vectors. The backward elimination (BE) algorithm (chapter 5) was used to reduce features using Fisher's discriminant measure and the dataset was classified using KNN. 6 features were selected from the total number of features in each case to maximise the distance between the training class centres. The results are displayed in table 8.1 and figure 8.3.

The wavelet standard deviation was included as a feature in addition to wavelet means. Using BE as in the previous case, the results improved as observed in table 8.2 and figure 8.4.

The best results were obtained for windows of size (20,300) and a scale decomposition of 3 scales using both standard deviations and means. These parameters gave a classification rate of 98.9%. The (20,300) windows in table 8.2 correspond to approximately $44m$ in the along track and $25m$ in the across track, thus covering an area of 1096 m^2. For higher resolution, good results were obtained for windows of size (40,40) pixels and 3 scale decompositions (corresponding to an area of $88.2m$ along track and $3.34m$ across it).

The results for using PCA analysis are displayed in table 8.3 and figure 8.5 for the means only and table 8.4 and figure 8.6 for means and standard deviations. In both tables, 3 principal components were used. However, the results are inferior to the BE algorithm in most cases considered.

Figures 8.7 and 8.8 consider the two best cases mentioned above and display the results of classification for both PCA and BE algorithms.

The results for a correct choice of window size are comparable to the results obtained using the 1-D method in chapter 6. Accurate GPS data was not available to locate the pings, so an error could have been induced from the fact that data is not geographically accurate. However, in [37], the same dataset was used without any reference to geo-referencing problems. It appears therefore, that for Pace's dataset, using 2D DWT does not necessarily improve classification accuracy, but has the advantage of detecting boundaries between sediment types.

nscales- window-size(BE)	(30,30)	(40,40)	(50,50)	(10,100)	(10,300)	(20,300)
2	69.9	77.3	74.2	72	80	88.9
3	86	91.3	86.6	80.7	92.42	96.7
4	82	91	94	74	79.8	88
5	81.2	90.3	82.5	75	81.3	91

Table 8.1: The results of using the backward algorithm for feature selection (DWT). 6 features were retained and used for a KNN classification. The numbers are the rates of correct classification compared to the original image. The features used for classification are the wavelet means.

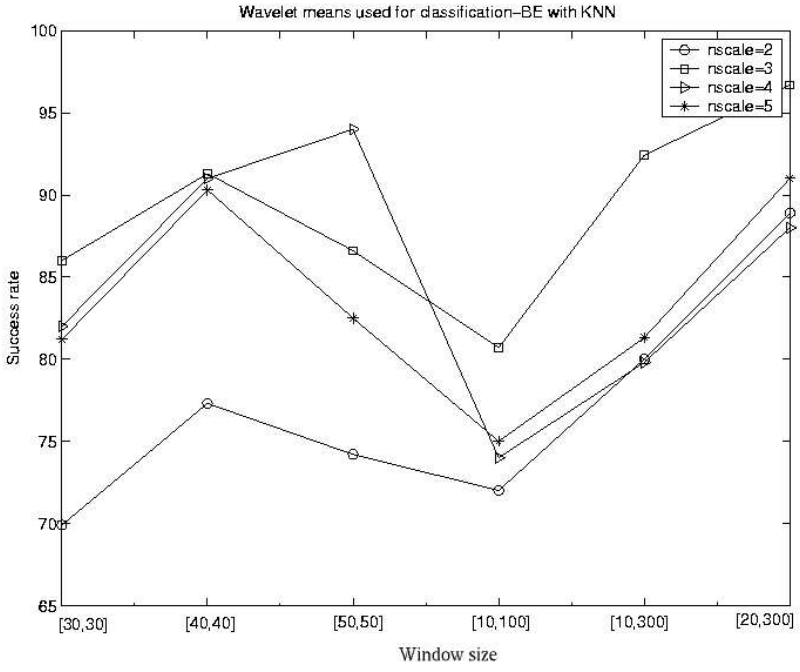


Figure 8.3: The results of using the backward algorithm for feature selection (DWT). 6 features were retained and used for a KNN classification. This figure presents the results from table 8.1

nscales- window-size(BE)	(30,30)	(40,40)	(50,50)	(10,100)	(10,300)	(20,300)
2	86.9	91.7	90.8	77.9	83.83	91.1
3	84.5	91.3	89.2	76.36	88.4	98.9
4	72.5	80.3	83.3	70.3	81.3	88.9
5	58.5	71.7	74.2	77.2	83.8	90

Table 8.2: The results of using the backward elimination algorithm for feature selection (DWT). 6 features were retained and used for a KNN classification. The numbers are the rates of correct classification compared to the original image. The features used for classification are the wavelet means and the standard deviations.

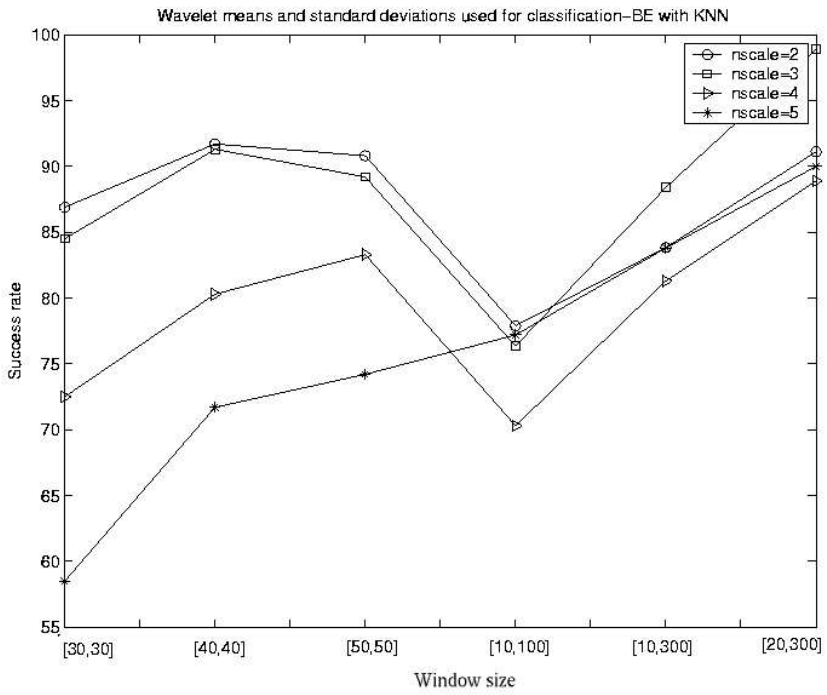


Figure 8.4: The results of using the backward elimination (BE) algorithm for feature selection (DWT). 6 features were retained and used for a KNN classification. This figure presents the results from table 8.2 and includes wavelet standard deviations as features used for classification.

nscales- window-size(PCA)	(30,30)	(40,40)	(50,50)	(10,100)	(10,300)	(20,300)
2	79.7	89	81.7	73.7	80.8	86.7
3	79.3	92.3	80.8	74.1	76.4	78.9
4	75.2	81.7	87.5	74.2	76.3	76.7
5	68.5	73.7	72.5	75	87.4	87.8

Table 8.3: The results of using the PCA algorithm for feature selection (DWT). 3 features were retained and used for a KNN classification. The numbers are the rates of correct classification compared to the original image. The features used for classification are the wavelet means.

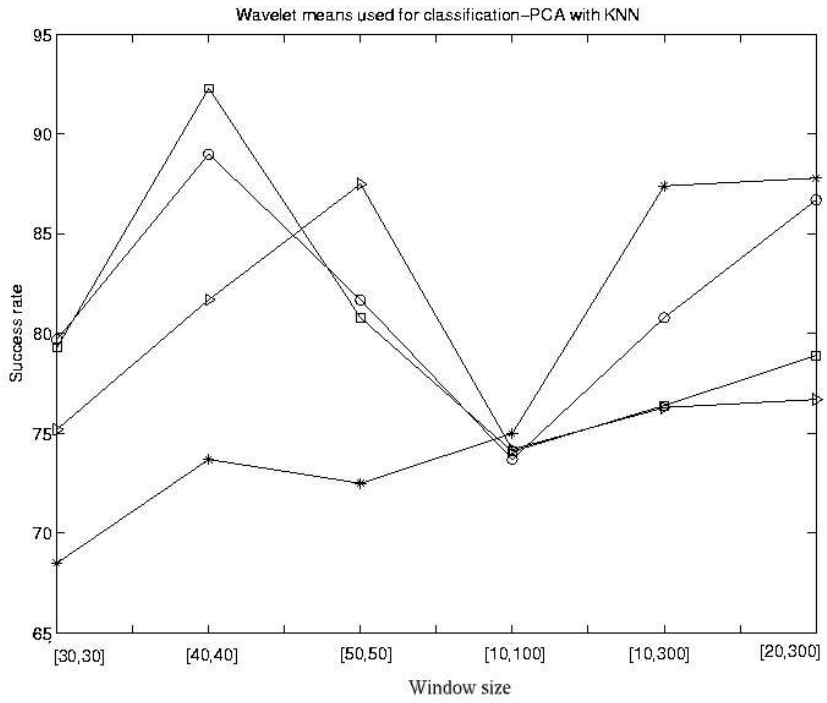


Figure 8.5: The results of using the PCA algorithm for feature selection (DWT). This figure presents the results from table 8.3 and uses wavelet means only for classification.

nscales- window-size(PCA)	(30,30)	(40,40)	(50,50)	(10,100)	(10,300)	(20,300)
2	79.4	87.7	79.2	69.24	72.2	80
3	73.2	82.7	81.7	73	78.3	81
4	71.4	77	89.2	71.7	77.3	76.7
5	66	78.7	74.2	78.3	86.9	86.7

Table 8.4: The results of using the PCA algorithm for feature selection (DWT). 3 features were retained and used for a KNN classification. The numbers are the rates of correct classification compared to the original image. The features used for classification are the wavelet means and standard deviations.

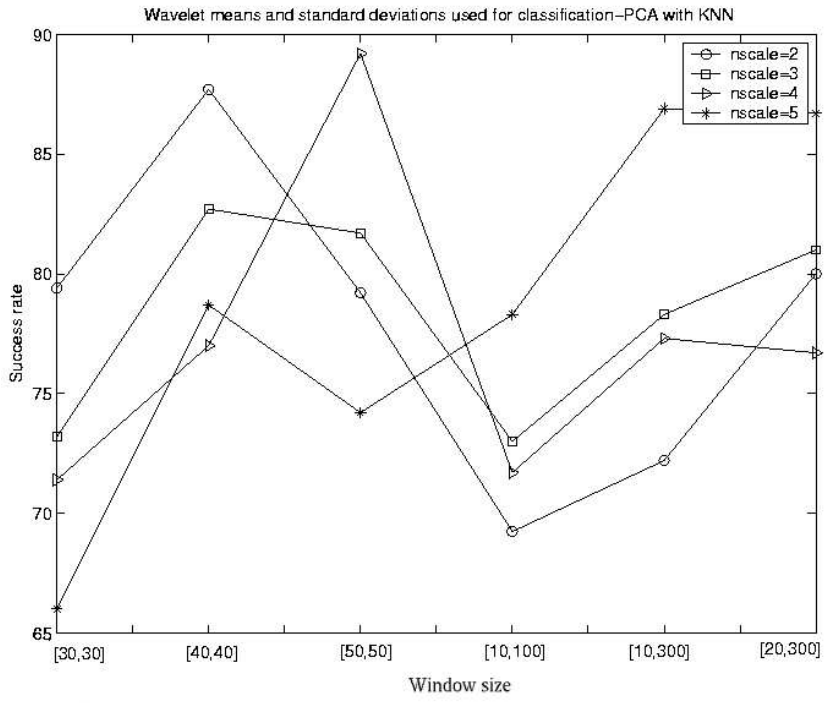


Figure 8.6: The results of using the PCA algorithm for feature selection (DWT). This figure presents the results from table 8.4 and uses wavelet means and standard deviations only for classification.

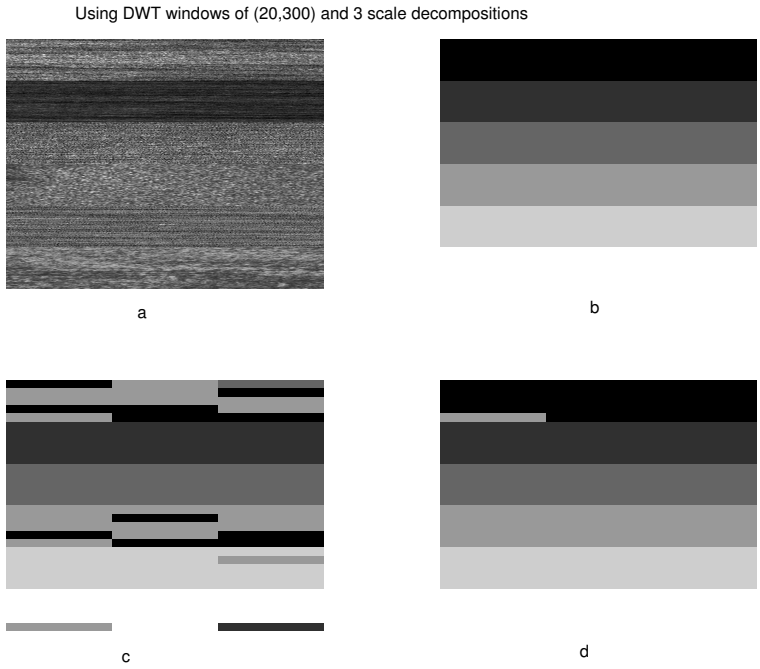


Figure 8.7: The results of classification using a window of (20,300) and a decomposition of 3 bands using both means and standard deviations as features. Figure (a) is the dataset to be classified. Since the classes are known they are assigned a value in (b). The results of classification using PCA are in (c) and the backward elimination algorithm in (d) with a success rate of 98.9%.

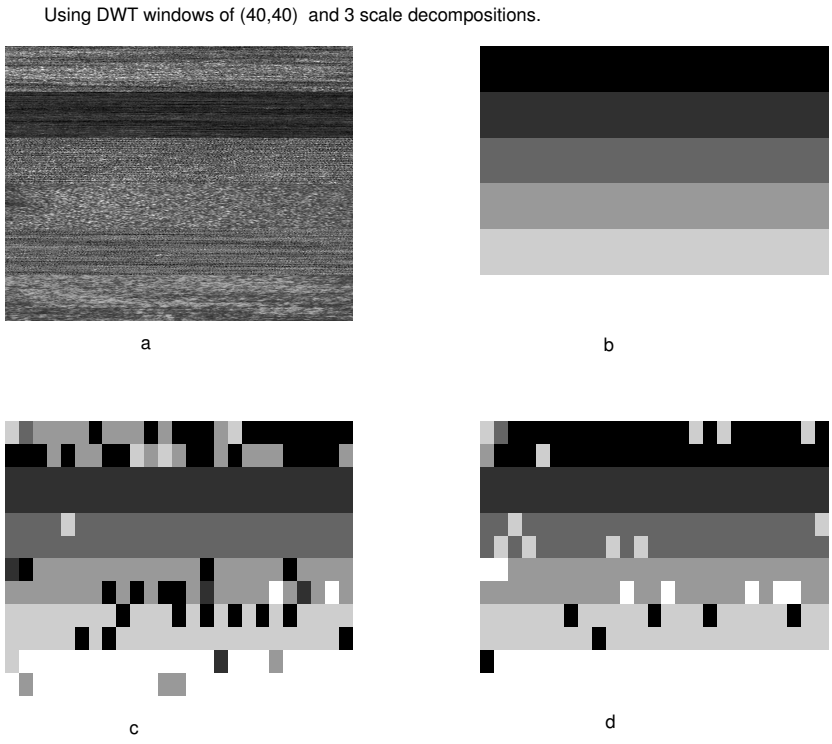


Figure 8.8: The results of classification using a window of (40,40) and a decomposition of 3 bands (also using both means and standard deviations). Figure (a) is the dataset to be classified. Since the classes are known they are assigned a value in (b). The results of classification using PCA are in (c) and the backward elimination algorithm in (d).

8.3.3 Gabor Analysis

In chapter 7, section 7.4, Gabor filters were introduced as an efficient method for texture classification with many advantages:

- The scale can be dyadic or not and so the user can choose to scale consecutive bands as appropriate for the application.
- The number of orientations per scale can be varied. Different orientations (and observation of the differences between them) can lead to a better classification especially in patterns with a dominant orientation.

- The size of the filter can be varied.

The advantages of Log Gabor filters over Gabors were explained in section 7.4. To construct Log Gabor filters, Kovesi's approach was followed in this study[1]. Kovesi uses an algorithm that splits the 2-D frequency spectrum into *norient* angles, then multiplies them by a log Gaussian function (section 7.4) that is scaled between consecutive scales and is selected optimally to divide the frequency space without redundancy. Figure 8.9 is an example of constructing one of the 2-D Log Gabor filters.

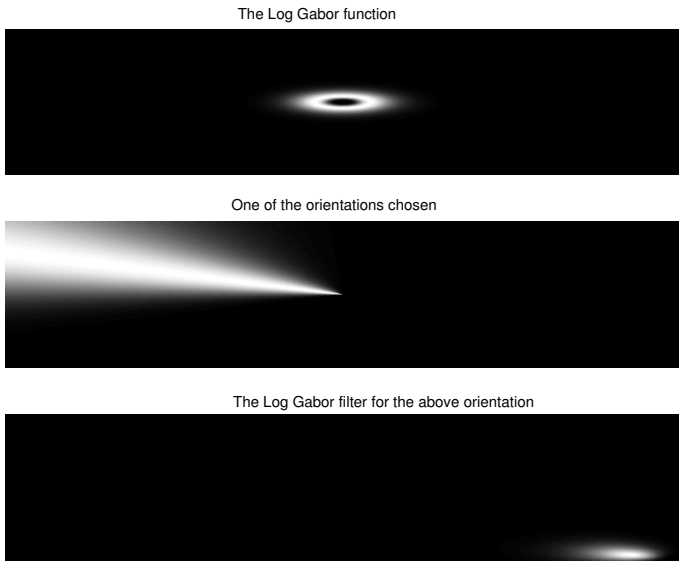


Figure 8.9: Construction of a Log Gabor filter from a log Gabor function and an exponential that divides the frequency space into several orientations. Only one of the filters is shown in the figure.

The main aim of using Gabor filters in this section is to observe the effect of orientation and decomposition on the classified results. The parameters varied for Gabor filters were the number of scales, and the number of orientations. The number of scales were chosen from [2,3,4,5] and the number of orientations per scale from [3,4,6]. The window size was fixed at (20,300) since this window size gave the best results in the DWT study. The preprocessing and feature selection stages are the sane as in the DWT, detailed in the previous paragraph. The

[1] www.cs.uwa.edu.au/~pk

results are given in table 8.5 and figure 8.10 for the BE algorithm and in table 8.6 and figure 8.11 for the PCA. In both tables, the mean of wavelet scale energy is used as a feature. The results of introducing standard deviation as a second type of feature are given in table 8.7 and figure 8.12 for the BE algorithm and in table 8.8 and figure 8.13 for PCA.

Figures 8.14 and 8.15 consider the best 2 cases from the above tables using Gabor filters with a fixed window size of (20,300).

nscales-norient (BE)	3	4	6
2	77.8	88.9	96.6
3	81.1	83.3	94.4
4	81.1	83.3	94.4
5	81.1	83.3	94.4

Table 8.5: The results of using the backward elimination algorithm for feature selection after filtering the image with Gabor filters. 6 features were retained and used for a KNN classification. The numbers are the rates of correct classification compared to the original image. This table stresses the effect of orientation on the Pace dataset the results improve greatly when more orientations are added. The features used for classification are the wavelet transform means.

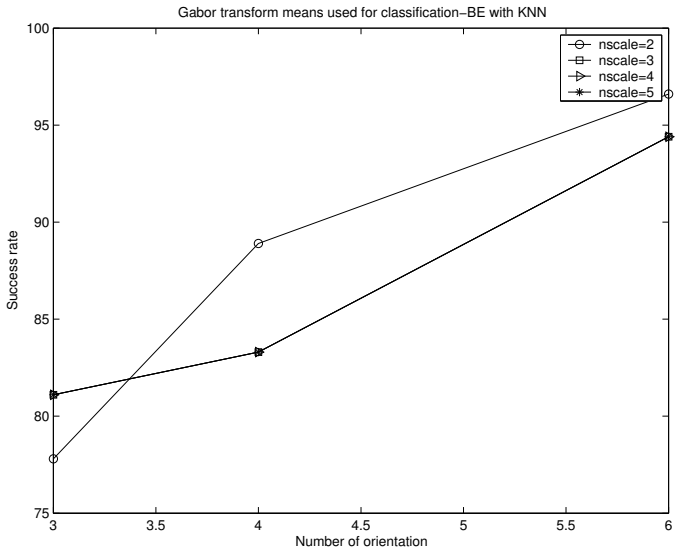


Figure 8.10: The results of using the backward elimination algorithm for feature selection after filtering the image with Gabor filters. This figure presents the results from table 8.5.

nscales-norient (PCA)	3	4	6
2	76.7	82.2	88.9
3	71.1	84.4	88.9
4	78.9	86.7	88.9
5	83.3	87.8	91.1

Table 8.6: The results of using the PCA algorithm for feature selection after filtering the image with Gabor filters. The numbers are the rates of correct classification compared to the original image. This table also stresses the effect of orientation on the Pace dataset the results improve greatly when more orientations are added. The features used are also the wavelet transform means.

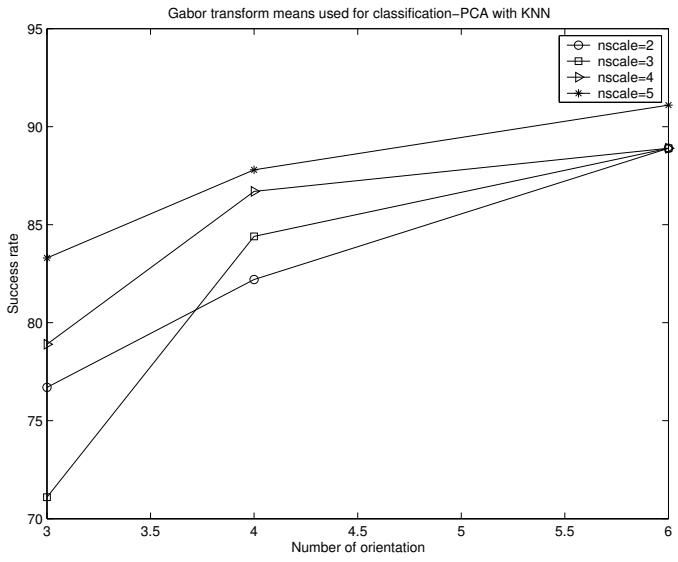


Figure 8.11: The results of using the PCA algorithm for feature selection after filtering the image with Gabor filters. This figure presents the results from table 8.6.

nscales-norient (BE)	3	4	6
2	80	91.1	91.1
3	77.8	64.4	90
4	77.8	64.4	90
5	77.8	64.4	90

Table 8.7: The results of using the backward elimination algorithm for feature selection after filtering the image with Gabor filters. 6 features were retained and used for a KNN classification. The numbers are the rates of correct classification compared to the original image. The features used are both means and standard deviations.

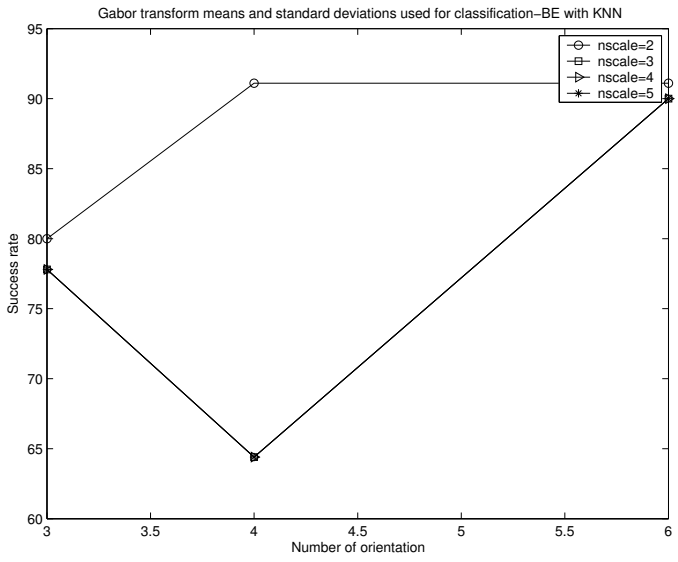


Figure 8.12: The results of using the backward elimination algorithm for feature selection after filtering the image with Gabor filters. This figure presents the results from table 8.7.

nscales-norient (BE)	3	4	6
2	85	78.9	78.9
3	73.3	83.3	87.8
4	78.9	84.5	84.5
5	83.3	86.7	90

Table 8.8: The results of using PCA for classification after filtering with Gabor filters. The numbers are the rates of correct classification compared to the original image. The features used are both means and standard deviations.

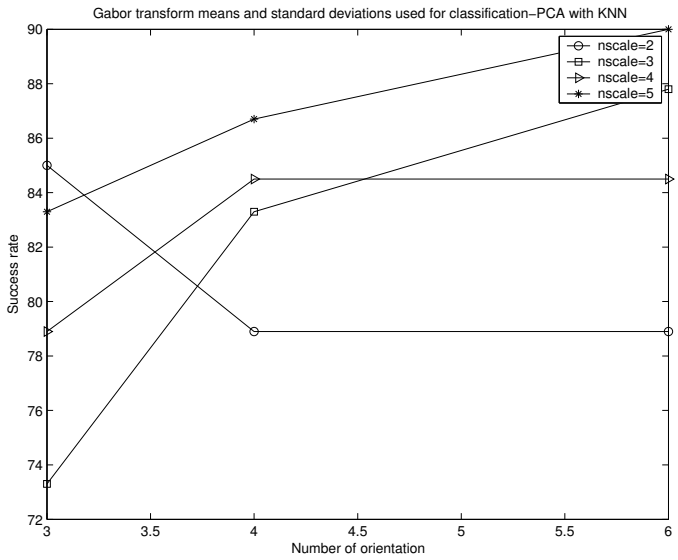


Figure 8.13: The results of using PCA for feature selection after filtering the image with Gabor filters. This figure presents the results from table 8.8.

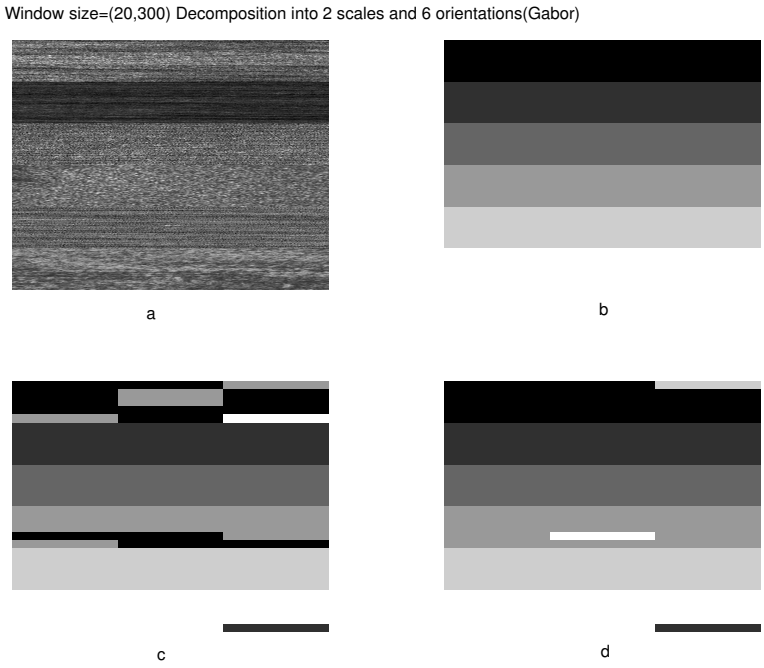


Figure 8.14: The results of classification using a window of (20,300) and a Gabor decomposition of 2 scales and 6 orientations per scale. Figure (a) is the dataset to be classified. Since the classes are known they are assigned a value in (b). The results of classification using PCA are in (c) and the backward elimination algorithm in (d).

8.3.4 Comparison of the Two Methods

What was observable in both methods is that in many cases, adding more scales caused redundancy rather than better classification. In the DWT case for example, the results for 4 scales are better than 5 for both the PCA and the BE cases.

For the Gabor filters, orientation proved to be very important as the classification rates increased as more filters at different orientations were added. This can be easily observed in tables 8.5, 8.7, 8.6 and 8.8. An increase of orientation filters from 3 to 6 per scale caused the classification rate to improve by more than 20% in some cases.

For the DWT, adding the standard deviation of wavelet transforms as a new feature improved the classification rates in the backward elimination (BE) algorithm method. For PCA,

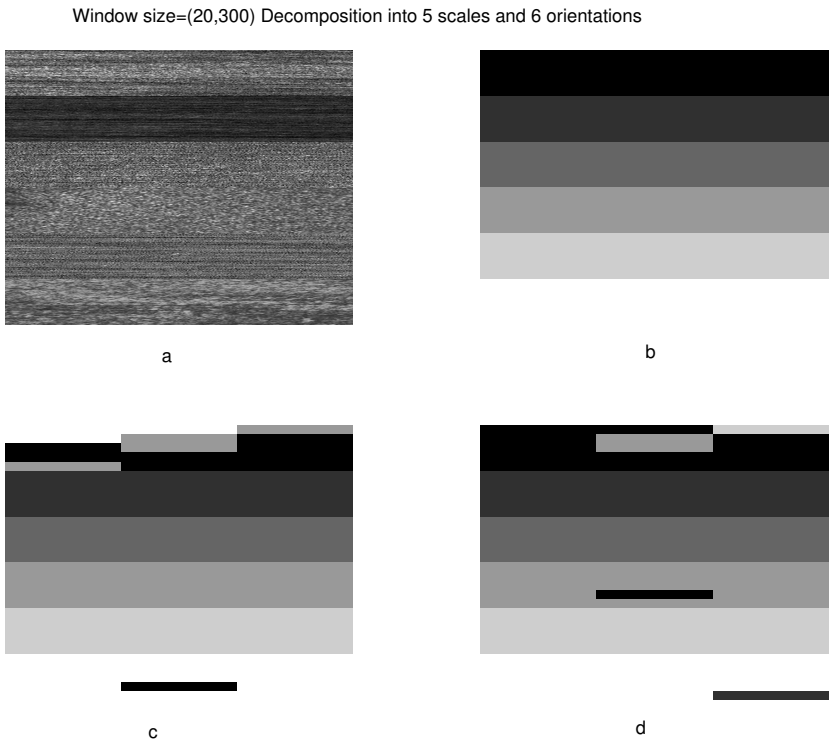


Figure 8.15: The results of classification using a window of (20,300) and a Gabor decomposition of 5 bands and 6 orientations per band. Figure (a) is the dataset to be classified. Since the classes are known they are assigned a value in (b). The results of classification using PCA are in (c) and the backward elimination algorithm in (d) with a classification rate of 94.4%.

the results were almost the same, or even worse in some cases. Probably adding these extra features lead to a data redundancy in some cases as can be seen when tables 8.3 and 8.4 are compared.

In most misclassification cases, the class that was mostly misclassified is the first one, namely clay. Looking at figure 8.2, it is observable that this class is not of uniform texture as a window is moved across it. Thus, it is highly likely that at certain areas this class could be misclassified. Larger windows solve this problem, and include more data from each class, but lower the classification resolution. With larger windows, the classification between types was better ($> 98\%$) and approached the results of the 1-D classification. The reason why the 1-D technique performed better in general cases was probably because it considered the whole

signal and averaged non consistent parts of it (such as the clay pings). In the 2-D technique the selection of window size played a major role, and lead in some cases to misclassification.

The main benefit of using the 2-D method is therefore the detection of boundaries between sediments (depending on the size of windows selected), whereas in 1-D the classification assumes uniform sediment type across the ping.

8.4 Artificial Coral Reefs

8.4.1 Introduction

This section uses the artificial reef dataset introduced in section 2.3.1. The main aim is to discriminate between the two types of reefs in figure 2.5 and the sediment background. Three different areas are defined as seen in figure 2.5: an area of coarse texture (in the upper reef, labelled type1), a less-coarse textured area (in the lower two reefs, labelled type3) and an area of surrounding sediment with smoother texture (labelled type 2). Figure 8.16 shows in addition to the original coral reef image an expert's classification of the coral reef area into an area of coarse texture (black), a less coarse area (white), and the surrounding (grey). What follows is a comparison of varying parameters using supervised and unsupervised classification to separate these three types:

- For the supervised approach, training features are chosen from selected windows in the image containing samples from the three types. The backward elimination algorithm is then used on the training set to determine the set of relevant features. The KNN approach is used for classification starting with the relevant features as centres.

- For the unsupervised approach, PCA is used to derive the most important features in the decomposed images. The final principal component features are used for unsupervised classification using the Kmeans algorithm.

The last section uses spatial information to include the influence of surrounding areas on classifying an area of artificial coral reef in these images.

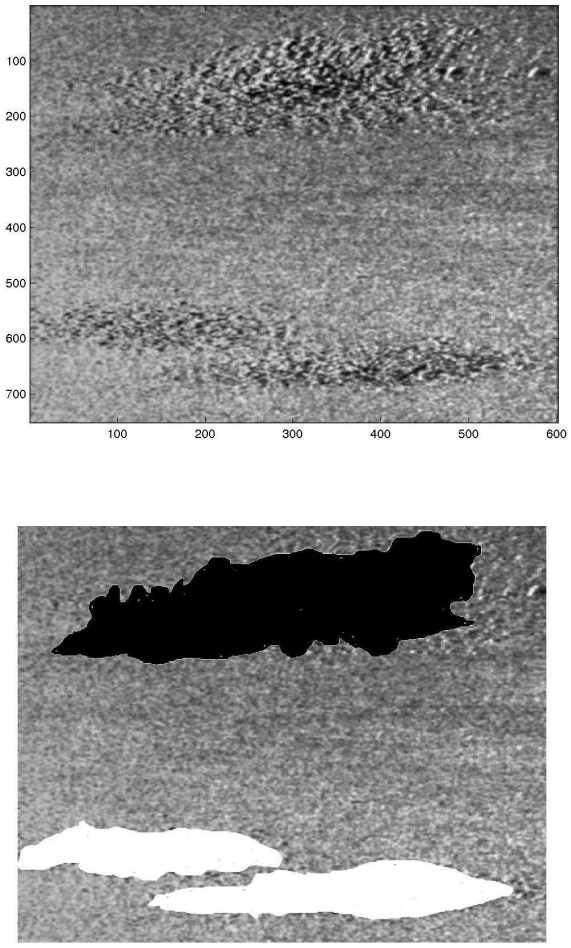


Figure 8.16: The artificial coral reef image (above), and an expert classification into 3 classes: coarse texture coral reef (black), coral reef of medium coarse texture (white), and surrounding area (gray).

8.4.2 DWT Results

The diagram in figure 8.1 will be used to explain the classification method in both the supervised and unsupervised case.

Supervised Classification

- Feature Extraction: The features used from the Wavelet transform are the energy and the standard deviation of wavelets within a user-defined window. The starting image was decomposed into *norient* angles and *nscale* scales. Resulting in $numb = 2(norient \times nscale)$ feature images. The feature images are transformed to feature vectors for classification.

- Feature Selection and Classification: For the supervised part of the study, training areas for three types are visually selected from the sonar image. Features are extracted from the training areas to form training vectors. The cluster centres are selected from these vectors leading to a centres matrix of dimension (number of types)$\times numb$ which is used as input to a feature backward elimination algorithm (BE). This algorithm uses Fisher's criterion to decide which features give the best separation between classes (from the centres matrix). The cluster centres are used to classify the dataset by using the KNN algorithm (see chapter 5).

Unsupervised Classification

The feature extraction stage is the same as that for supervised classification. The feature selection stage is different assuming no prior knowledge about the classes present in the sonar image. Thus, PCA is used to select the principal components of the *numb* features and Kmeans is used to cluster the data. Each group of pixels (box) is assigned a certain class after classification.

Variation of Parameters and Effect on Classification

The parameters to be varied in both the supervised and the unsupervised case are:

- The size of the starting wavelet.

- The number of scales to be used in the wavelet decomposition.

- The size of the window used to calculate features.

Effect of the size of the starting Wavelet Values of the size of the starting wavelet were selected from 4, 6, 8 and 10. The window selected for calculating the features was fixed at 10×10. The results are displayed in figure 8.17 for both the supervised and the unsupervised cases. The ideal classification is when the upper reef (type 1) is classified as 'black' in the classified image, the lower two reefs (type 3) as white, and the background (type 2) as gray. For the supervised classification, the upper reef is mostly separated from the surrounding despite the sonar image being quite noisy. However, there are a few misclassifications between areas in the two reef types (types 1 and 3), mainly due to texture similarity in these areas.

The results of the unsupervised methods are also in figure 8.17. The classification in this case seems more influenced by the initial selection of the wavelet size than in the supervised classification. More areas in the background are misclassified as the second type of coral reef (with not very coarse texture).

Effect of the number of scales The number of wavelet scales to be used for classification is varied from 2 to 5 ($nscale$). The result can be observed in figure 8.18 for both the supervised and non supervised cases.

Effect of the window size Varying the window size leads to a variation in resolution of the results. Figure 8.19 shows the results. A good window size is one that does not totally destroy the resolution of the initial image (large windows such as the 20x20 case) nor leave lots of noise in the classified image (very small windows that increase the rate of misclassification).

Comments

Due to lack of ground-truth for this dataset, the only proof of correct classification is user drawn boundaries given in figure 8.16. The ideal classification would be a gray background, with the upper coral reef in black and the lower ones in white. The main problem was separating the reefs from each other as they have very similar texture even for a trained expert. So it is no

surprise to see areas of similar texture in both reefs classified as the same type (coloured either black or white in the resulting images). The next section uses spatial information to improve the classification. Other alternatives could be using region-growing or a Bayesian classifier that uses neighbourhood information. However, from a textural point of view, the method provides a reasonable classification assuming the only factors affecting a region's classification are its textural properties.

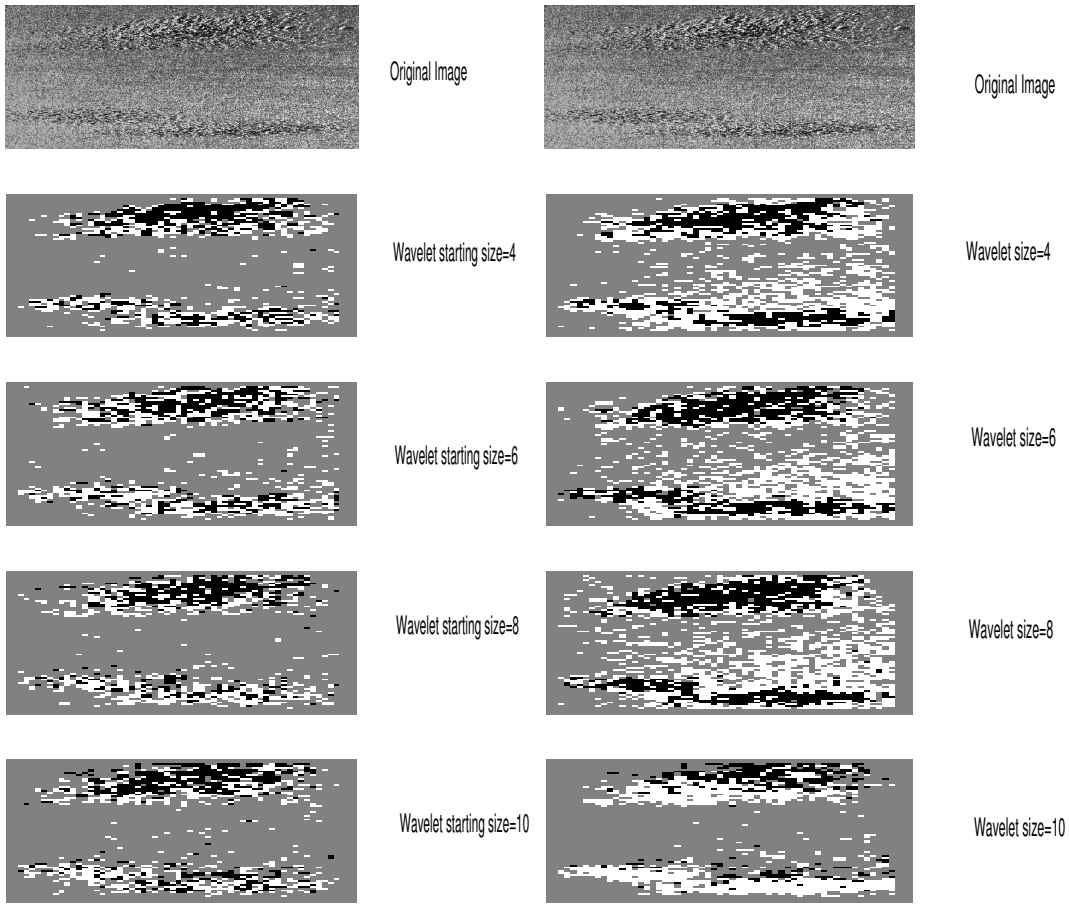


Figure 8.17: The upper figure is that of the original coral reef. The images on the left are the result of using a suprevised method (DWT), and the right images are the results of the unsupervised method (DWT). The parameter varied was the starting size of the Wavelet.

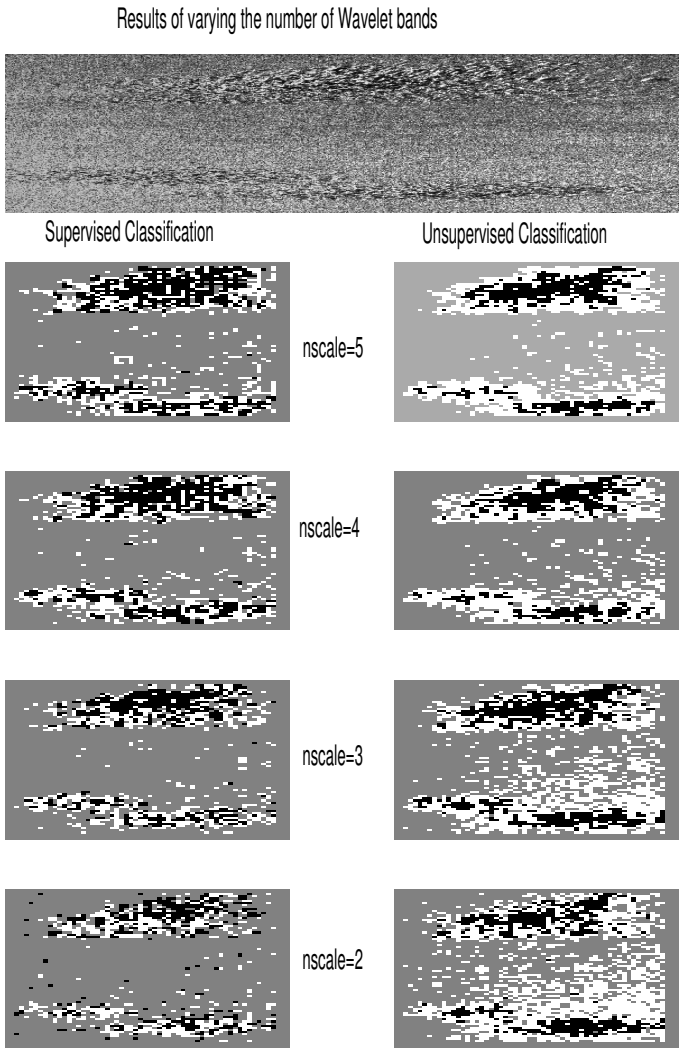


Figure 8.18: The upper figure is that of the original coral reef. The images on the left side are the results of using a supervised method (DWT) where the training classes are selected as boxes from the survey by the user. The images on the right side are the unsupervised results (DWT). In the first experiment, 5 wavelet bands were used, then the number was lowered in each case.

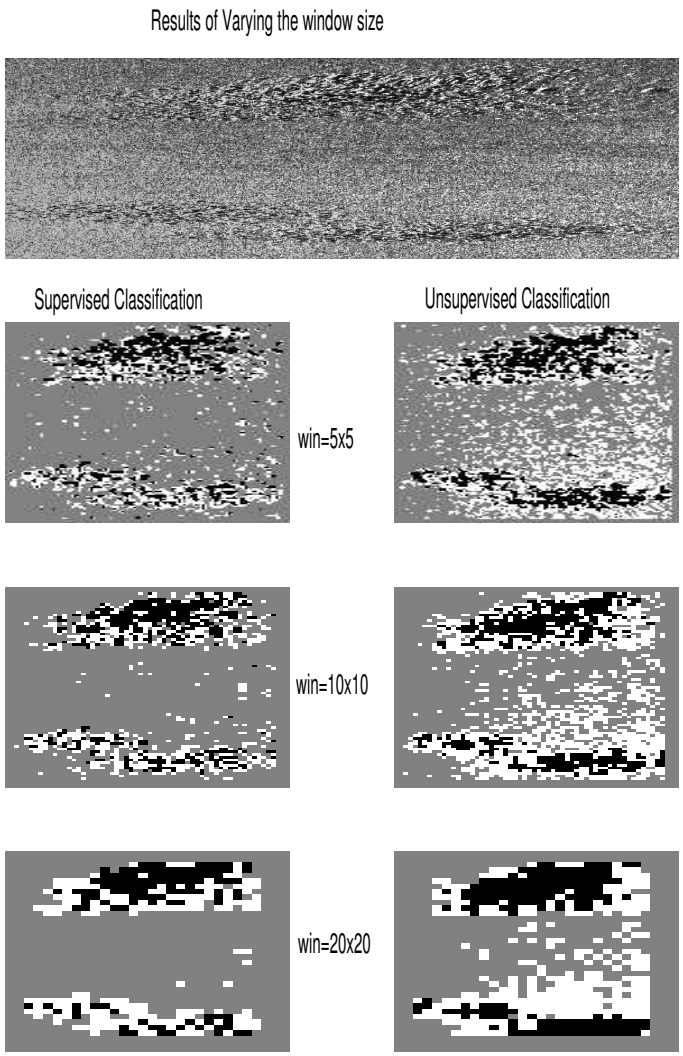


Figure 8.19: The upper figure is that of the original coral reef. The images on the left side are the results of using a supervised method(DWT) where the training classes are selected as boxes from the survey by the user. The images on the right side are the unsupervised results. In the first experiment, window sizes of 5x5 were used, the window size was doubled to 10x10 then 20x20.

8.4.3 Gabor Results

The method used for classification in the Log Gabor domain is the same as that used in the DWT case. The only difference is the parameters that were varied:

- The number of orientations per scale.
- The scaling between consecutive scales, whether dyadic or not.
- The starting size of the filter.

One parameter was varied at a time keeping the others constant.

In the first case, the number of orientations per scale *norient* is varied from 6 to 4 to 3 and the classification results are shown in figure 8.20. The scaling between classes was dyadic, and only 2 scales were used. The effect of adding more orientations did not affect the results a lot, most probably due to the features in the original image being quite random and more dependent on scale selection rather than orientation.

In the second case, the scaling factor was chosen from the following set [1.3, 1.6, 2(dyadic)] keeping the other parameters constant. The results are shown in figure 8.21.

Finally, the starting size of the filter was varied. Results are displayed in figure 8.22 for filter sizes chosen from [4 6 8].

8.4.4 Comparison of the two methods

From observing the results on the artificial coral reef dataset, the following conclusions can be deduced:

- The choice of the number of scales is important. Misclassifications from too small a number can be observed in the unsupervised case of figure 8.18. However, adding scales is not always helpful. Figure 8.18, for example, shows that having 5 scales or 4 produces almost the same results. The number of scales must be chosen to include significant

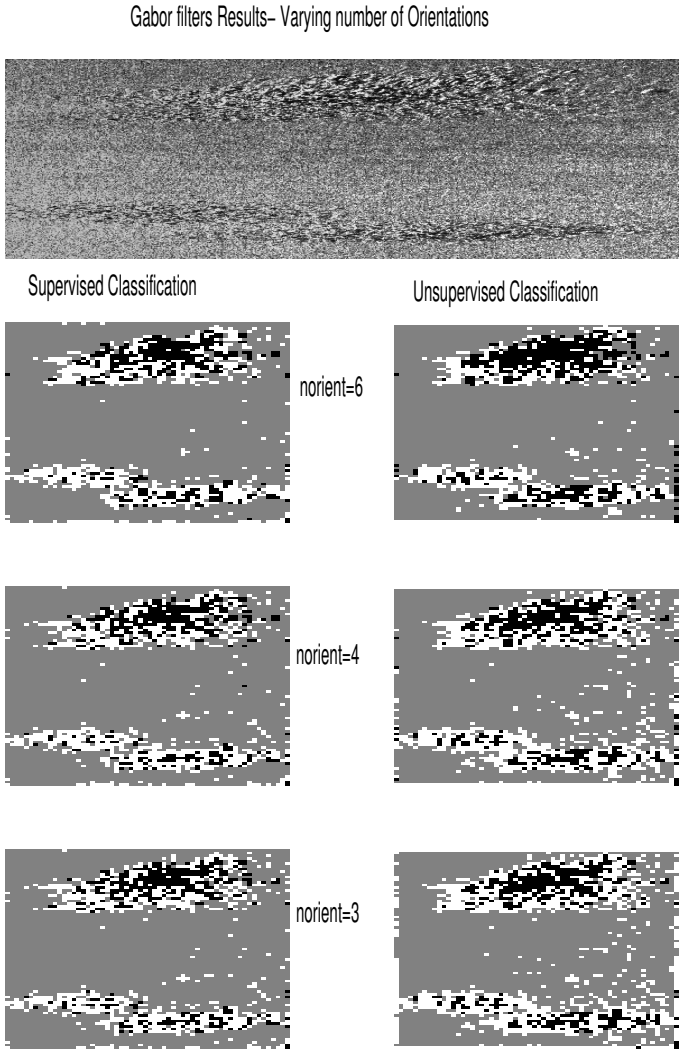


Figure 8.20: Supervised (left) and unsupervised (right) results by using Gabor filters on the reefs dataset. The number of orientations per scale varies from 6 to 4 to 3.

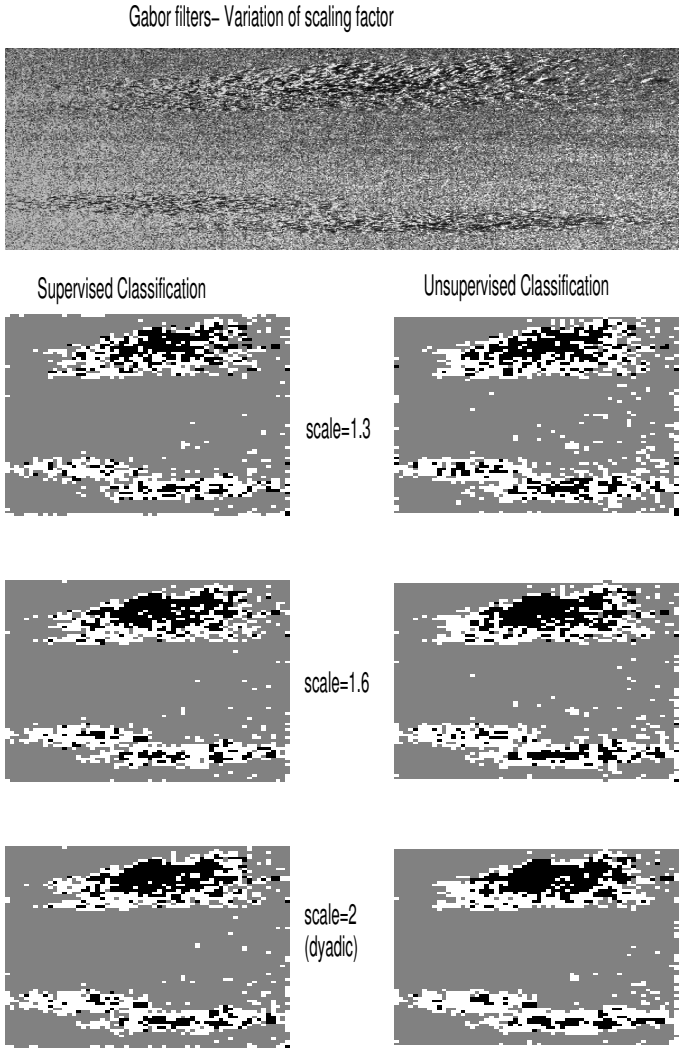


Figure 8.21: Supervised (left) and unsupervised (right) results by using Gabor filters on the reefs dataset. The scale multiplication factor was chosen from {1.3, 1.6, 2}

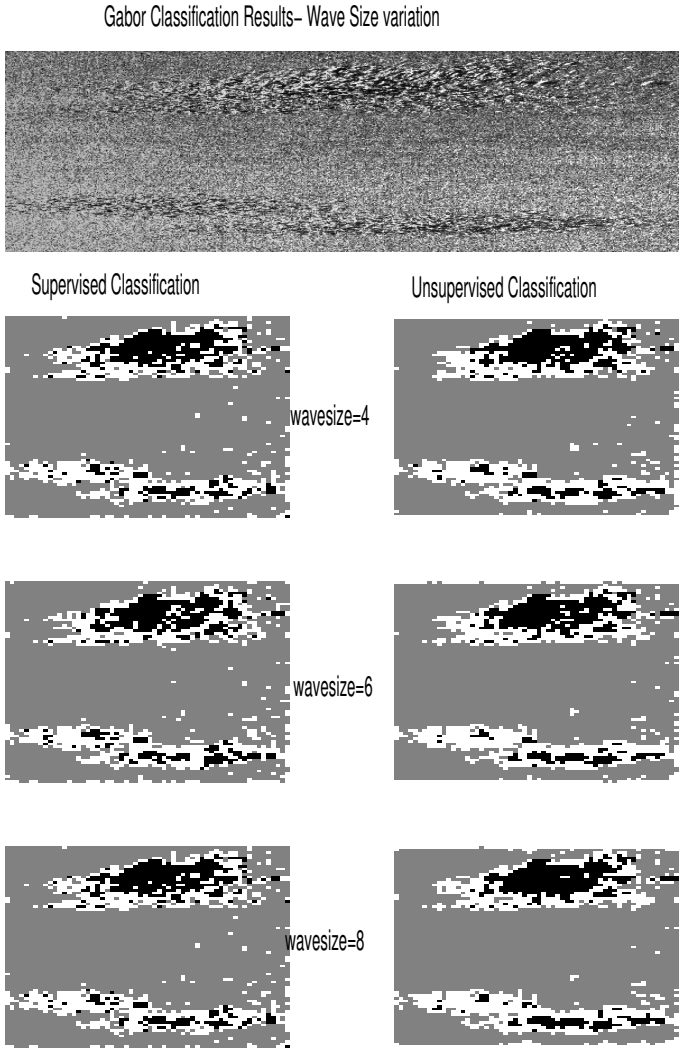


Figure 8.22: Supervised (left) and unsupervised (right) results by using Gabor filters on the reefs dataset. The starting filter was chosen to be from the following {4, 6, 8}

features to discriminate between the classes but not so many as to produce redundancy in the feature set.

- Window size is an important factor, as seen in figure 8.19. A careful selection must be made to avoid noise and to provide a good resolution.

- The number of filter orientations in the Gabor case did not have a major influence on the results. For the coral reef image, adding more Gabor filter orientations per scale did not improve the discrimination between the classes.

- Non dyadic scales can be a better choice for some datasets, since they can pick up more high frequency features. For the coral reef image, however, scaling factor variation did have a major influence on the results.

8.4.5 Using Neighbourhood Information to Improve Classification Results

The previous section presented experiments with several filter parameters with both DWT and Gabor filters, to observe the effect of varying several parameters on the classification of the artificial coral reef dataset. However, one problem was evident in all the classified images from figure 8.18 to 8.22. Classification using texture properties per window does not take into consideration the information present in the neighbourhood of a window. There are several methods that tackle the problem of texture segmentation (as referenced in chapter 7). The method that is used in this work can be summarised in the following steps:

- First, the image is classified as given in the previous section (using KNN) and a chosen type of filter (Gabor or DWT) and filter parameters.

- From the initial classification, a 'prior' labelling of each window can be obtained. A neighbourhood of size N is selected around each window. The number of windows belonging

to each class are counted per neighbourhood. The ratio of the count over neighbourhood size is used as a probability of a window belonging to a certain class given its surrounding (pw). An example is given in figure 8.23.

- The probability of a window to belong to a certain class (C_1) is given to be proportional to $exp(-d^2)$, where d is the distance between the feature vector for that window and the centre of the class. Thus, the closer a window is to the class center the higher its probability is.

- The probability of belonging to a certain class is weighted by the probability due to neighbourhood information pw. Thus, a probability of belonging to a class (C_1) is given more certainty if it is in an area of a high number of windows with class C_1, and low certainty otherwise.

- The new weighted probabilities are used for classification. [2]

Results Using Neighbourhood Information

The experiments done in the previous sections are repeated using neighbourhood information. Several values of N, the neighbourhood size were tested. The value that gave the closest results (by visual inspection) to the expert classified image, was a 4 by 4 neighbourhood. This value was selected for the next set of experiments. Results in this section are presented for DWT varying the following parameters, and adding neighbourhood information as previously explained:

- Wavelet starting size, in figure 8.24.

- Window size, in figure 8.26.

[2] Altough an exact Bayesian formulation of this method is possible, the aim of this section is to illustrate the benefits of using neighbourhood information to improve the classification results. Texture segmentation using Bayesian techniques is itself a very wide subject and references are given in chapter 7.

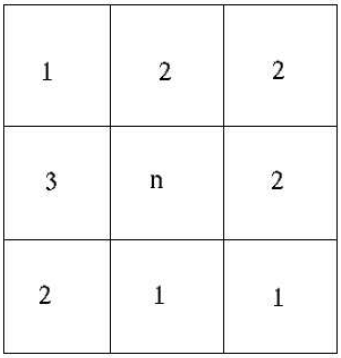


Figure 8.23: The probability of a window belonging to a certain class given its neighbourhood. The classes in this figure are $1, 2$ and 3. The probablility pw is given as a count of class label over total number of windows in the selected neighbourhood. The neighbourhood here is given to be 3 by 3, but the same method applies to larger neighbourhoods.

- Number of wavelet scales, in figure 8.25.

For Gabor filters, neighbourhood information is used (same size as before) and the parameters varied are:

- The number of orientations per scale, in figure 8.27.
- The scaling factor between consecutive scales, in figure 8.28.
- The starting Gabor wavelet size, in figure 8.29.

Note that the results are only given for the supervised case (where a selected area of the coral reefs is used for training), an extension to the unsupervised case follows exactly the same methods, but uses class clusteriung information to determine class centres (rather than training

areas). By observing the results, the method provides a straight forward way of incorporating neighbourhood information in texture classification. The results for the wavelet filters seem to be better than those for Gabor filters, and adding orientations does not seem to offer a great benefit in this case.

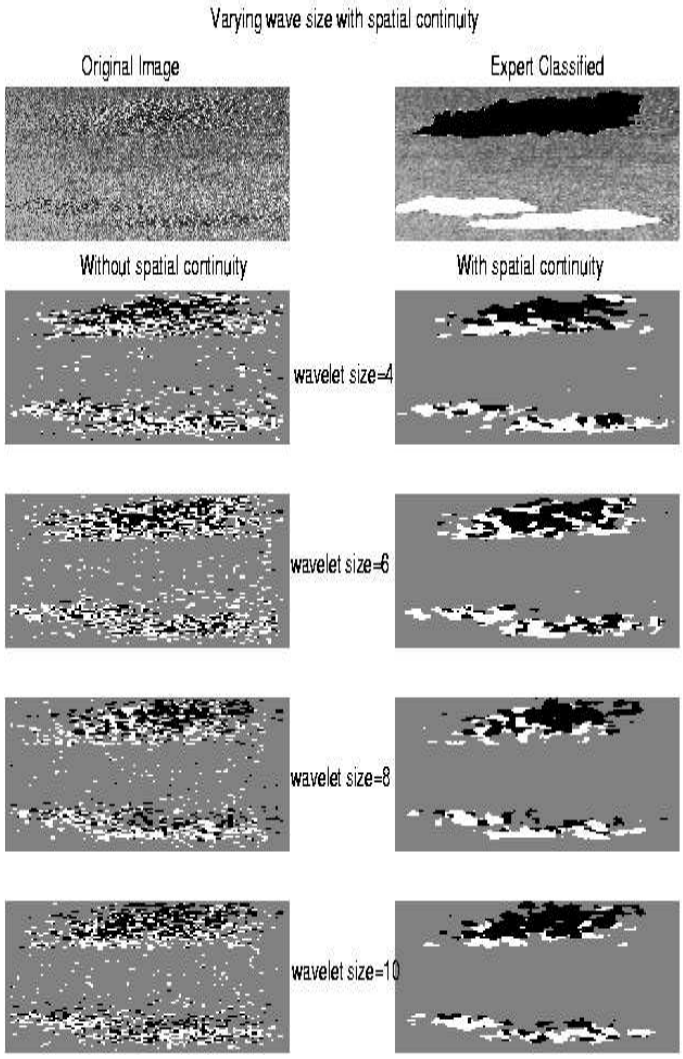


Figure 8.24: The upper two figures are of the original coral reef and the expert classified image. The images on the left side are the results of using a supervised method (DWT) where the training classes are selected as boxes from the survey by the user. The images on the right side are the supervised results improved by using neighbourhood information. The parameter varied is the wavelet starting size.

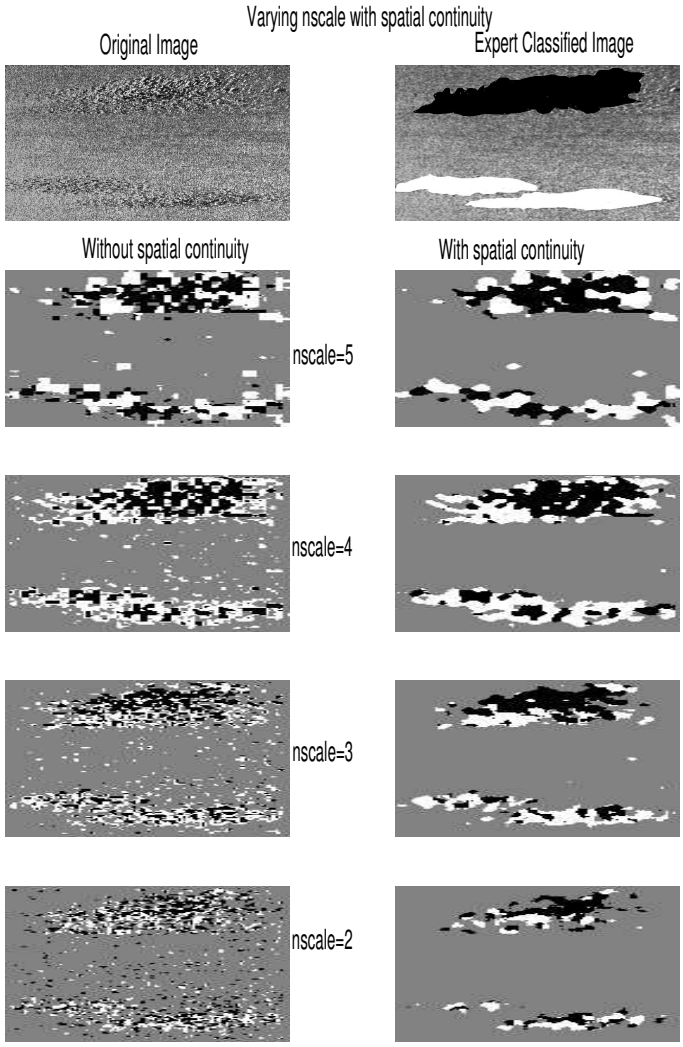


Figure 8.25: Similar to figure 8.24, but the parameter varied is the number of wavelet bands (DWT).

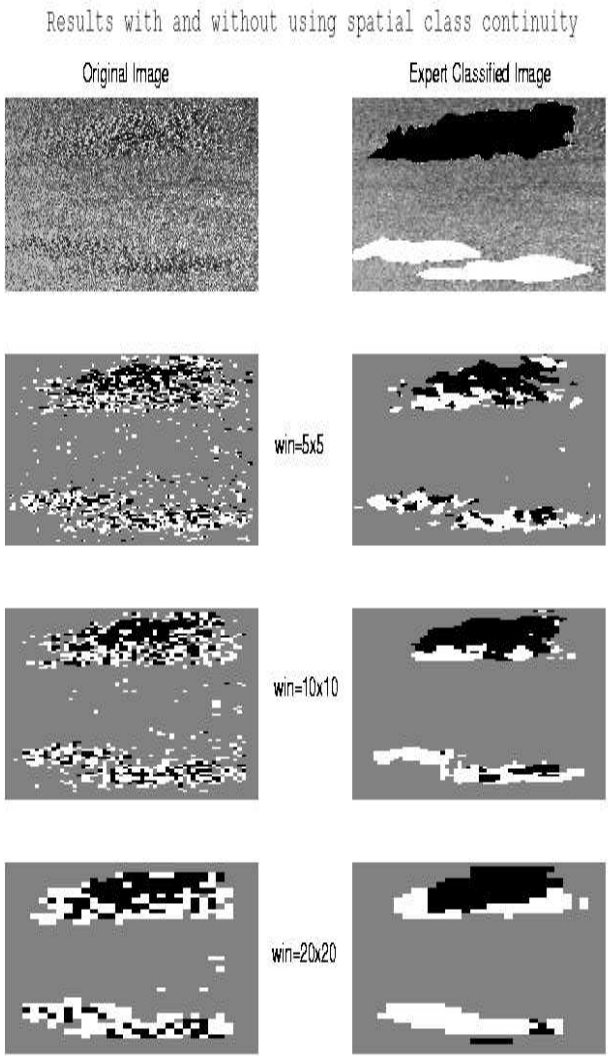


Figure 8.26: Similar to figure 8.24, but the parameter varied is the size of the window used for feature extraction (DWT).

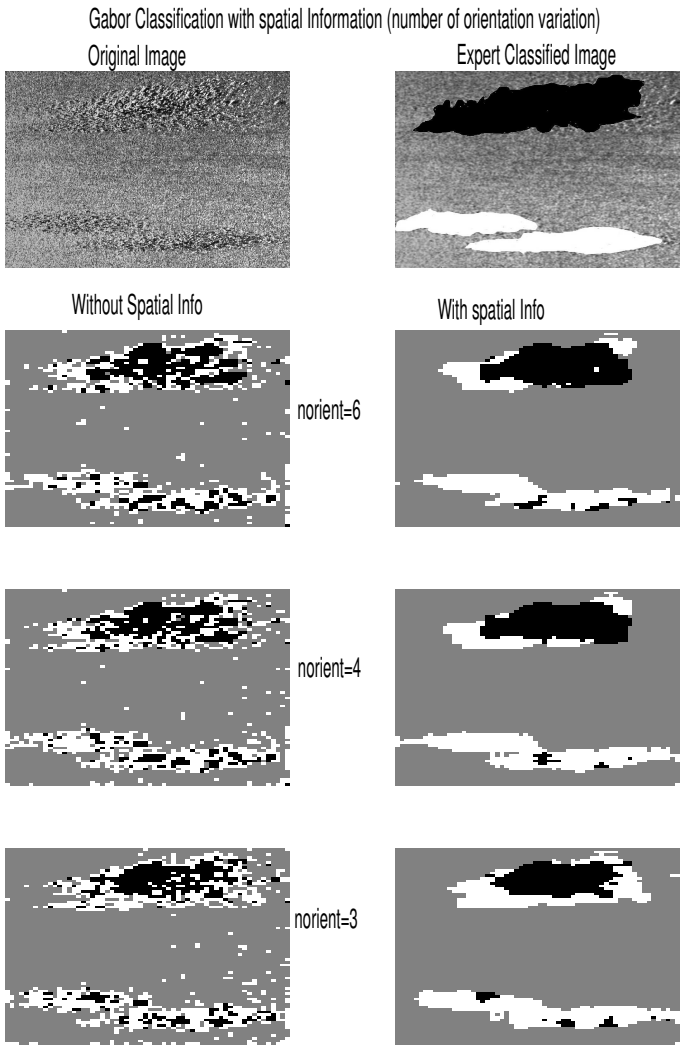


Figure 8.27: The upper two figures are of the original coral reef and the expert classified image. The images on the left side are the results of using a supervised method (Gabor filters) where the training classes are selected as boxes from the survey by the user. The images on the right side are the supervised results improved by using neighbourhood information. The parameter varied is the number of orientations per scale.

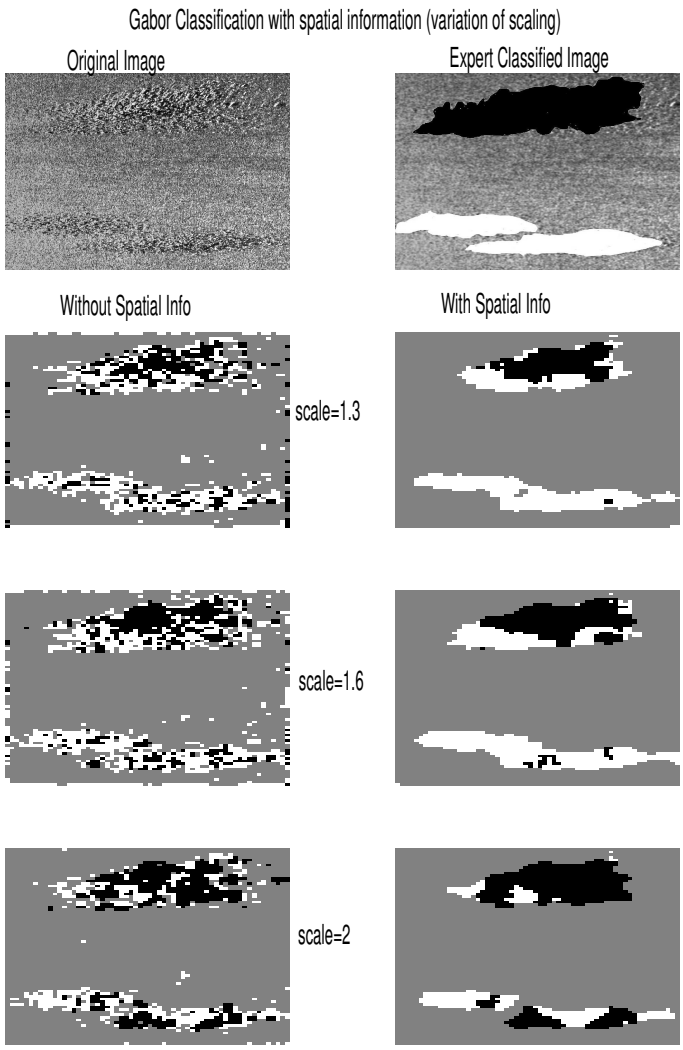


Figure 8.28: Similar to figure 8.27, but the parameter varied is the scale multiplication factor.

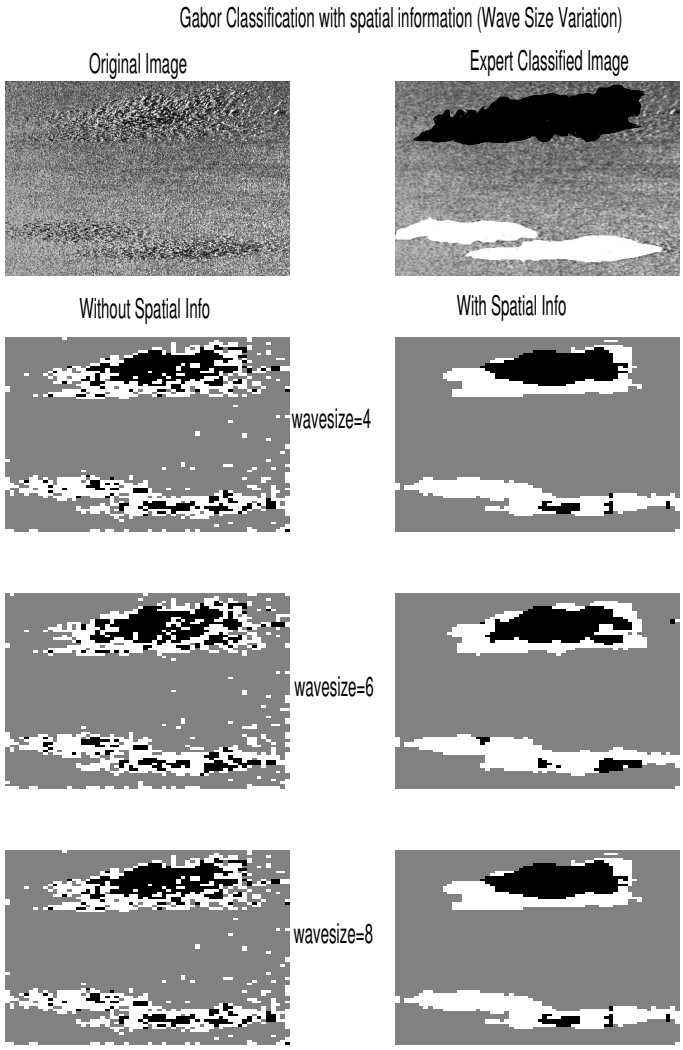


Figure 8.29: Similar to figure 8.27, but the parameter varied is the starting filter size.

8.5 Summary

Two datasets were classified using both the Discrete wavelet transform and the Gabor filters (Log Gabors). For the artificial coral reef dataset, the method was divided into a supervised part, where windows were selected in the image and used to train a KNN neural network, and an unsupervised part using Kmeans clustering. The performance of both methods was observed in each case. For both the DWT and Gabor filters, parameters were varied and the final classified images were compared. Spatial information was incorporated and the results proved to have more visual similarity to the expert classified image.

The sidescan sonar dataset had classes assigned types in advance, so the results could be compared to the a-priori given classification. For this dataset, feature selection was done using both PCA (Principal Component Analysis) and Backward elimination methods (BE). These methods were compared with respect to the final classification stages. The classification was supervised in both cases, and a number of pings was considered for training and setting the centres of clusters. The filter parameters were varied for each of the DWT and the Gabor transform to obtain the best classification of the given dataset. Good rates of classification were obtained (more than 98%).

The conclusions after varying several parameters were:

- The inclusion of orientation dependent features was useful for the sidescan dataset but did not offer great improvements for the artificial coral reef dataset which has more random features. Hence, the use of orientation dependent Gabor filters is useful mainly for datasets with clear non-isotropic features with dominant orientations (such as sand waves).

- The addition of more scales while filtering proved unnecessary after a certain level and classification rates went down. A level of decomposition should be selected where data redundancy is avoided.

- The use of standard deviation of wavelet transform energies in addition to their mean improved the results in many cases.

- In general, the BE algorithm gave better results than PCA in selecting features for classification.

- Window size is an important factor. Windows should be chosen to give good resolution as well as provide good class discrimination. The selection of windows offers the advantage of observing boundaries between different types (which was not possible in the 1-D classification).

- To improve segmentation, using neighbourhood information helped draw class boundaries between the coral reefs and their surrounding areas. Errors due to only using textural information regardless of the window position were reduced.

Compared with the 1-D classification method, the 2-D filtering methods offer the advantage of using information from consecutive pings for classification. Boundaries are observed better and the classification results are similar if the right filtering parameters are selected.

Chapter 9

Bathymetric Features for Sonar Classification

'Only a fool tests the depth of water with both feet'

African proverb

The advantage of using a sidescan bathymetric sonar over a simple sidescan sonar is that a map of the seafloor bathymetry can be provided. Bathymetric information can help in recognising many seabed areas that show similar features. The aim of the 2-D part of this study is to combine various features for a better classification of seabed areas. These features include backscattering amplitude features (explained in chapter 2), textural features (in chapter 7) and bathymetric features which will be detailed in the next few sections.

The first section of this chapter introduces several of the approaches used to derive bathymetry-related features from sonars. The main contribution of this chapter, however, is the introduction of new features derived from the bathymetry of high frequency sonars. These features will be used in chapter 10 for the classification of sidescan bathymetric sonar data.

9.1 Why use bathymetry for Sonar Analysis?

The generation of bathymetric data may serve different purposes when integrated into an environmental survey. The bathymetric profile provided in both multibeam and sidescan bathymetric sonars can be used to map pixels of the sonar image correctly. Distortions in the acoustic image can be corrected by using the bathymetric profile and plotting an accurate plot of the relief [38, 36]. To do an accurate study of backscatter versus grazing angle, the bathymetry has to be taken into account, to correct for the effect on the amplitude profile and to plot each point in the amplitude profile where it is actually located in the seabed. The assumption in most cases is that of a flat seabed, but in Cervenka *et al.* [38] the geometry of the bathymetric profile is used to reposition the pixels.

Bathymetry is an easily viewed backdrop over which other data can be draped and studied, particularly when the data is presented in a GIS format, or in any griddable format. Features in the seabed such as channels and outcrops can be observed better when draped over a bathymetric map of the same area. SEA's RTS2000 offers the option of draping the sidescan image over the bathymetric profile, a facility which may aid a user in observing underwater features. Figure 9.1 is an image from Loch Laxford in Scotland (see section 2.3.2) where the sidescan profile is draped over the bathymetry. The image on the right shows the slope of the bathymetry, measured and plotted for each point to give an idea of its bathymetry variation. Although visualisation methods combining sidescan sonar data with bathymetry data have had some success in portraying complex areas [84], these results are qualitative and do not allow the user to explore the data in a truly geophysical sense [150]. More measures are required to obtain features from the bathymetry that are useful for classification [88].

Bathymetric classification is usually used for large scale features in the seabed. Thus, it is more relevant for low frequency sonars ($10 - 15\ kHz$) that map large features at significant depths (such as Ocean ridges). These large scale features can offer an indication of ocean currents as well as other geophysical information relating to the geological formation of a certain

area and its composition. An example of this is the crust next to mid-ocean ridges which is quite rough in texture and generated through time by a combination of tectonic and volcanic processes, forming abyssal hills characterized by a series of approximately parallel ridges and troughs [103].

To obtain features for classification, large amounts of data are normally needed to gather statistics of seafloor roughness, which is defined in terms of the spatial power spectra of the topography. Sidescan sonar imagery is capable of resolving smaller features but lacks the information base that bathymetry provides. Hence, to deal with this problem, a new approach was suggested by Mitchell *et al.* [150]. This approach combines features from both bathymetry and sidescan. The authors select a training set from a gridded set, and perform a comparison using a set of data attributes related to both sidescan and bathymetry. Sidescan features include the echo strength after correcting for physical factors and bathymetric features include the slope and topologic curvature of the region. Using these bathymetric features and using a paraboloid to measure coefficients proved to be a more efficient way to detect bathymetric orientation compared to contour maps usually used in bathymetric analysis [150].

In [103], variograms are introduced for use in bathymetric classification. The variogram is defined as a spatial structure function. Considering elevation to be a regionalised variable $z(x)$, defined through out a two-dimensional region D, the variogram is calculated in bins according to:

$$\gamma(h) = \frac{1}{2n} \sum_{i=1}^{n} [z(x_i) - z(x_i + h)]^2 \tag{9.1}$$

where $z(x_i)$ and $z(x_i + h)$ are samples taken at locations x_i and $x_i + h \in D$ respectively, and n is the number of pairs separated by the vector h. Thus, the variogram gives a good indication of the spatial correlation in elevation. Features derived from the variogram indicate surface roughness types and can be used for classification. For example, in [103], the authors observe variograms of training data, then use them to extract parameters to identify different types. Sediment ponds for example yield a smooth variogram with overall low values, whereas

abyssal hill variograms have large values. Thus, these two classes can be distinguished from each other. In automated classification, a window is moved over the study area and parameters are extracted from variograms, then used to compose feature vectors and finally assign a class of seafloor type.

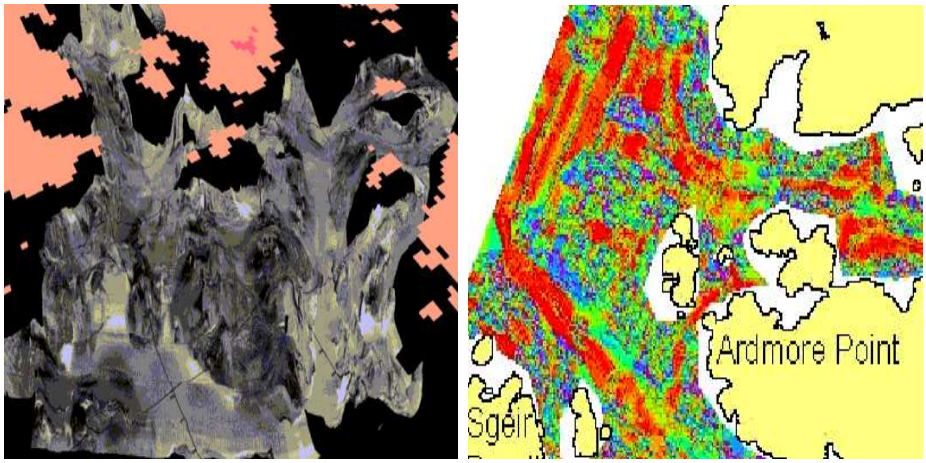


Figure 9.1: The figure on the left is the sidescan of a part of the Laxford dataset draped over the bathymetric profile. The figure on the right is the slope of the bathymetry plotted for each point. The slope can provide the user with an measure of where there is a large or small variation of bathymetry. The figures were provided by R. Bates (University of St Andrews).

Multidirectional classification is also used in [103] to make use of the geologic information and to find the main directions of a certain area. In addition to that, an automated approach is introduced to determine the most distinctive direction automatically and use it to distinguish an area from its surroundings.

Fox and Hayes [88] develop a 2-dimensional spectral model of bathymetry. Parameters of the model can be directly related to seafloor properties corresponding to isotropic (non directionality dependent) roughness, anisotropic roughness (directionality dependent), orientation of components and spectral roll off. The model is fitted to 2 dimensional spectra of gridded bathymetry. A test area is selected and cluster analysis is used to classify the classes. The method is used to classify large seabed features and an applied to areas such as the Juan de

Fuca Ridge [87].

If bathymetry is not available, an alternative approach using sidescan imagery can be used to detect bathymetric features of a certain area. In [77], the authors use the geometry of the sidescan image formation process and the reflectivity model to recover 3-D surface information. Lambert's reflectivity model is assumed since it provides a good fit to backscattering from roughness and volume together [112]. The drawback of this model is that it is a very simplified approximation of real sediment reflectivity. The 3-D surface information is then used to draw contour maps that join points of equal elevation. Contour maps are useful for identifying directional features such as ripples. The authors classify their areas into rock and sand ripples by using features from contour maps.

9.2 Application to the Sidescan Bathymetric Dataset

The previous section describes the methods generally used for extracting information from bathymetry of a certain seabed area. The major applications are to analyse features at usually large depths using a sonar of low frequency. This study uses a sidescan bathymetric sonar at relatively shallow depths and rather high frequencies (compared to the applications in section 9.1). The bathymetric features usually extracted from low frequency bathymetric images cannot be applied directly to the datsets considered since the bathymetric images are of lower quality. This fact makes it a very difficult to derive discriminative features from the bathymetric images, compared to very efficient sidescan sonar textural features.

To obtain bathymetric images, the sidescan bathymetric sonar is used as explained in chapter 2. The corrections done on the received signal are detailed and the final result is an interpolated bathymetry profile for each ping leading to the bathymetry image. The bathymetric images are of lower resolution than the sidescan images, but they are useful to observe general trends in seabed bathymetry. After observing both sidescan and bathymetric images of certain areas provided, the following can be deduced:

1. In many cases, and after observing both sidescan and bathymetric images obtained from different types of seabed areas, it is evident that the change of bathymetry in a certain area affects pixel-strengths in sidescan images. This fact is rarely taken into consideration in sonar classification studies, mainly due to the unavailability of bathymetric data. Variation of depth in a certain area can definitely cause a change in pixel strength, and if not corrected for or taken into account it can lead to misclassification. This is the main reason why the processed sidescan data is used in this study rather than the raw sidescan data.

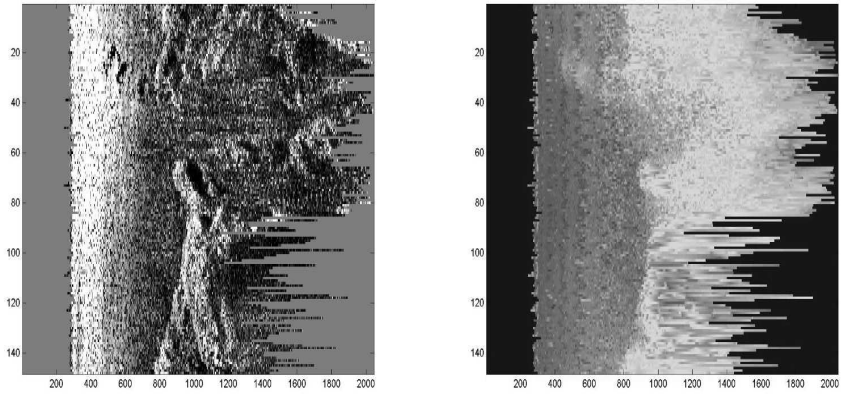


Figure 9.2: The image on the left is the sidescan amplitude image of a seabed area, the right one is its bathymetric profile. The change of pixel strength in this case is due to the variation in bathymetry. Some areas in the bathymetric image (in this case the bottom right) do not have enough resolution and thus they are not used for classification. In the amplitude image, clipping also has to be taken into account and clipped points are not used for classification.

2. Bathymetric slope by itself can provide information on the variation in depth of the area. It can be used to study the effect of depth variation on other features (such as textural features). But as a discriminative measure on its own, it is highly unlikely to be useful especially in small areas. The Hopvågen dataset (introduced in chapter 6), for example, was composed of vast areas of smooth sediments. Slope was not an efficient measure in

this case, but it could be in cases where certain sediment (or rock) types have a dominant slope.

9.2.1 Bathymetric Features

In this study, and to summarise huge amounts of data, bathymetric features are calculated per window. A discriminative bathymetric feature observed for many datasets considered is the following:

$$D(winx, winy) = \text{standard deviation}(d(x + areax, y + areay)) \tag{9.2}$$

where a window is considered in the bathymetric image (d) around point (x, y). This window covers an area of $(areax, areay)$. The standard deviation of the bathymetry in the window is measured and saved as a feature for the box in the location $(winx, winy)$. This feature expresses the variability of a certain seabed type bathymetry in a chosen window. Rocks for example express more variability than smooth sediments such as mud unless there were external factors causing the sediments to move, or deposit. So studying the variability can be an indication of seabed type. Calculating the bathymetric feature per selected window helps 'smooth' the bathymetric data, which in general, has a lot of outliers. The resolution is ofcourse dependent on the window size. Figure 9.3 shows the bathymetry mean and bathymetry standard deviation measured per window for the same area selected in figure 9.2.

9.3 Summary

This chapter presents a brief summary on using bathymetry as a backdrop for other sonar data, then a review of the use of bathymetric classification for low frequency sonar data sets. Bathymetric features are rarely used as features on their own for classification of high frequency sonar data. This chapter introduces a method of using high-frequency bathymetric features for classification of sidescan-bathymetric dataset used in this thesis. With a sufficient resolution, the variation of bathymetry in selected windows can be an indicator of seabed type. The use

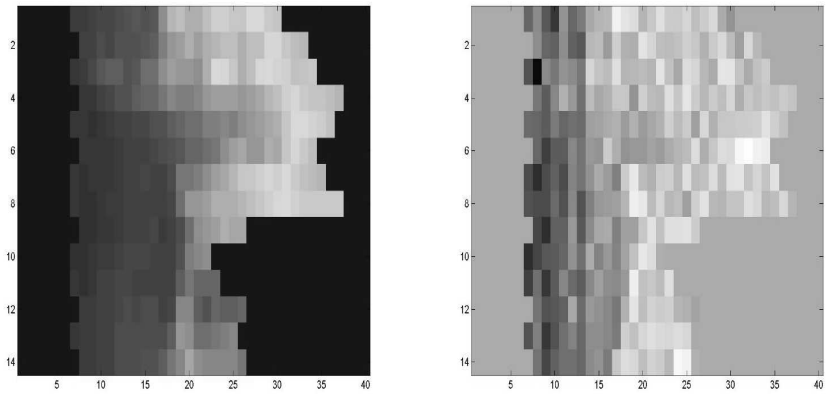


Figure 9.3: The mean of bathymetry per feature box (left), and the standard deviation D (right). The mean itself hardly gives any information besides the general sloping of the seabed. The standard deviation is used as a discriminating feature.

of these features is not limited to sidescan-bathymetric sonars, and can be used for any type of sonar that provides bathymetry with a reasonable resolution such as multibeam sonars.

Chapter 10

Classification of Sidescan Bathymetric Data

10.1 Background

For a good analysis of sidescan bathymetric data, a sonar expert generally observes several aspects of the data provided by the sonars. To understand the data efficiently, textural differences between different areas are observed as well as variations in pixel strength and bathymetric graphs. Interpretation is done across several dimensions and relies on the experience of the interpreter to ensure an accurate result [150]. Visualisation methods have been used to combine sidescan sonar data with bathymetry (as in RTS2000- the SEA sidescan bathymetric software package). These methods allow the user to observe many aspects of the data at the same time. However, they remain qualitative and are highly dependent on the user's expertise.

An objective method of sonar analysis is required which can automatically use the different features offered by advanced sonars. The importance of this method becomes more evident in use for underwater vehicles, and in multifrequency sonars. In these two cases, it is very difficult for a user to observe many aspects of the data at the same time, and a processing method has to be used.

Chapters 7 and 8 explored the use of textural features for the classification of seabeds. In chapter 9, bathymetric features were derived to analyse sonar data. This chapter presents a new method of combining all the above features for a better discrimination between sediment classes. Properties of sidescan data and bathymetric images have been used separately in several studies for sediment classification. As far as we are aware, only one study attempted to combine features from both [150]. However in [150], the sidescan features relied only on amplitude and not texture, and the dataset samples which were considered had clear bathymetric differences between each other (due to the use of low frequency sonar in deep sea areas). This chapter presents a method for high frequency sonar classification, using an automatic feature combination and classification method. In the first part of this chapter, the classification scheme is presented (methods of classification) with an emphasis on feature selection and extraction. In the second part, the results are presented for two sidescan bathymetric datasets. The first dataset is the Hopvågen dataset presented in chapter 2. The second dataset is a part of a survey done in Laxford for Scottish National Heritage [14] (also introduced in chapter 2). The final results for each set are a comparison of the classification scheme results with grab samples from the area, and other classification systems.

10.2 Sidescan Bathymetric Classification Scheme

In this section, we will refer to the general classification scheme presented in chapter 8, figure 8.1. The classification scheme can be divided into the following steps:

- The user input stage: Several parameters have to be assigned by the users before running the process on the whole survey. These parameters include the texture filter properties, the grab sample positions in the survey and the number of sediment types (they will be detailed in the next section).

- The preprocessing stage: to obtain a processed sidescan image and a processed bathymetry

image (this stage is detailed in section 3.1.1).

- The feature extraction stage: This stage includes the extraction of textural features, amplitude features and distance variation features. The area is divided into user defined windows and features are extracted for each window.

- The feature selection stage: If grab samples of known types are available from the survey, a grab localisation step is required. This step searches the whole dataset for given grab samples and chooses training areas (of user defined size) around the grabs. The training areas are used to select the features that provide the maximum discrimination between the sediment classes (using methods described in section 5.3). If grab samples are not available, an algorithm is used (such as PCA) to reduce the dimensions of the dataset. The features then would be the reduced feature vectors (of user defined size).

- The classification stage: This stage also depends on the availability of grab samples. If grab samples are available, then the classifiers use the training areas as cluster centres and classify the survey into sediment classes (using KNN). If grab samples are not available, an unsupervised classification is used. The classifiers assign a class to each window in the dataset, then display the windows in their exact GPS position.

10.3 Data Input and Pre-Processing

Several parameters have to be assigned before running the method on the whole survey, these parameters are:

1. The window sizes $(areax, areay)$. For the whole survey, the features are extracted per window and each window is described with a feature vector of size $nfeat$.

2. If feature selection algorithms are used, the number of selected features ($nselectedfeat$), which is selected from the total number of features $nfeat$ should be specified. In un-

supervised classification, $nselectedfeat$ is the number of useful principal components to keep.

3. The type of wavelet to be used in texture analysis filtering, and the number of wavelet scales.

4. The location and type of grab samples if they are available, and the size of the training area selected around each grab sample.

5. The total number of sediment classes that the data should be divided into ($ntype$).

As for the processing stage, the raw sonar data is processed as detailed in chapter 3. This processing takes care of clipped points, near nadir and far-distance data. The final output of the preprocessing stage is a series of corrected amplitude pings (representing backscattering strength) and a series of bathymetric profiles corresponding to them. By arranging these pings in order of received time, two images of the area can be obtained [1]:

- The backscattering strength image (BS) which is obtained after correcting the intensity image for losses.
- The bathymetry image (d).

10.4 Feature Extraction

The whole survey is then divided into user defined windows of size ($areax, areay$) and features are extracted per window. The main reason for choosing this approach over doing a pixel by pixel classification is the need to summarise data. Dealing with large datasets (each line of surveyed data has a minimum size of 100 Mb), it becomes a very computationally intensive

[1]Note that arranging pings by received time and not taking into consideration the exact gridding of the pings in the survey, could lead to classification errors, especially when sonar motion effects are significant. This issue will be discussed for the classification of the datasets considered in this chapter (section 10.6).

task to do a pixel classification, and the requirements are much more than the current workstation can provide. In addition to that, a good selection of window size averages the data, reducing noise and preventing erroneous data from being classified into a new class, although there is a compromise in the resolution of the classified image. The extracted features are of three types: bathymetric features, backscattering strength features and textural features.

10.4.1 Bathymetric Features

Chapter 9 introduces features that can be derived from bathymetric images of high frequency sonars for the classification of several seabed areas. The feature D, describing the standard-deviation of bathymetry per window, was derived in section 9.2 and is used in this chapter as a bathymetric feature of seabed areas.

10.4.2 Backscattering Strength Features

The most commonly used method of sonar analysis using multibeam or sidescan bathymetric sonars is to use the amplitude of backscattered signals to discriminate between different seabed types. The backscattered signal is corrected for physical factors as detailed in chapter 2, section 3.1.1, and one of the outputs of the preprocessing stage is the backscattering strength signal (processed sidescan image). Different types of sediments have different absorption and reflection strengths for the same grazing angle, and so they have different backscattering strengths. These values can be used to separate the sediments into different classes ([3], [148], [147], [93], [36]).

The following two features are calculated per window of size $(areax, areay)$ in the backscattering strength (BS) image:

$$M(winx, winy) = mean(BS(x + areax, y + areay)) \tag{10.1}$$

$$M_s(winx, winy) = \text{standard deviation}(BS(x + areax, y + areay)) \tag{10.2}$$

These features are displayed in figure 10.1 for the same area shown in figure 9.2.

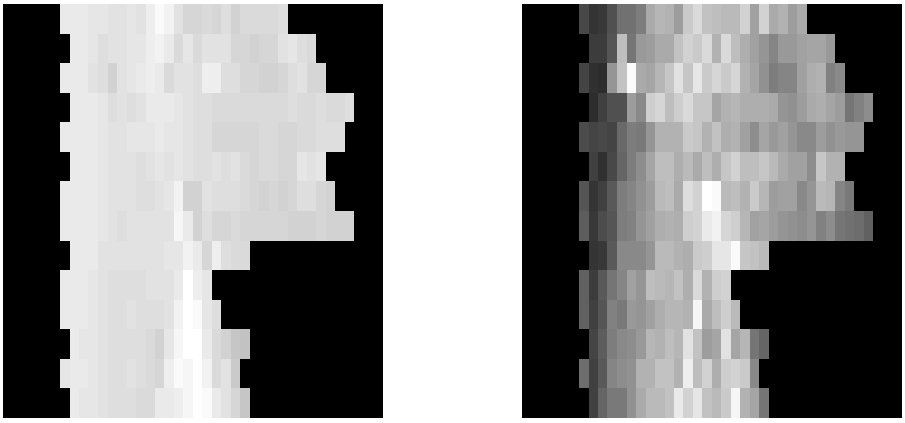


Figure 10.1: For a window of 50 pixels horizontally, and 10 pixels vertically, the figures (a) and (b) are the mean feature image (M) and the standard deviation feature image (M_s) of the selected area respectively.

10.4.3 Textural Features

Chapters 7 and 8 investigate the use of textural features for the classification of sidescan sonar datasets. Textural information can be used for the classification of sidescan bathymetric data as well, since the received amplitude images can be treated as sidescan amplitude images. However, the images used in this work are the backscattering strength images (BS), rather than the raw sonar data images. In the backscattering strength images, the effect of clipping is reduced, so is that of pixel-amplitude variation due to bathymetry.

In the cases considered in chapter 7 and in several other studies [175], the selection of a method to derive textural features proved to be dependent on the type of texture to be classified. If the texture evidently has dominant orientations that can be used as a discriminant feature, filters of multiple orientations are chosen (such as Gabors, or Complex Wavelets). If the features span a number of scales, then a multiscale approach is required. The speed of the process and computational complexity are also factors to be taken into consideration.

The DWT was chosen for this section, since a multiscale approach was needed to focus on the high frequency scales. The images are in general very random and do not display large

orientation differences (the Hopvågen dataset, for example).

Textural features are calculated per window in the survey. The Backscattering strength image is filtered across $nscales$, and on each scale the resulting three filtered images are used to obtain feature images. The features selected for classification are the means of wavelet energies within these windows $S(winx, winy)$[2] (as described in chapter 7, equation 7.12). The feature analysis of an image results in $3(nscale)$ texture-mean feature images $S_{(i,j)}(winx, winy)$ where i is the number of the scale and j indicates which orientation is chosen.

Boundary conditions resulting from sudden jumps in the image, and leading to large wavelet transform coefficients in the filtered images are corrected. Acoustic shadow areas are not considered for classification, as they form a class of their own and lead to misclassification. Acoustic shadow areas are areas that are not insonified by the sonar due to the geometry of the seafloor.

10.5 Feature Selection and Classification

10.5.1 Supervised Classification

In many of the datasets used in this study, several grab samples were available from around the survey. The grab samples are obtained by sending divers and are verified with a study of the sediment grain size in the region. It is assumed that areas around the grab samples are of the same type (this assumption was verified by observing sidescan images). The training stage consists of searching for the nearest area in the survey that is closest to the grab sample. According to the grab samples, the area is divided into a number of types ($ntype$) using a standard sediment classification table such as the Wentworth scale [216]. If the grab sample itself does not exist in the selected area of the survey, a range is allowed which in most cases is between 10 to $15m$. The areas around the grab samples are used as training areas of user

[2] $winx$ and $winy$ are the coordinates of the centre of the window that is used to calculate the features.

defined size ($nwinh$, $nwinv$) and the features derived at these areas are used as cluster-centres for classifiers. The cluster centres are stored in the centres matrix of size $ntype$ (number of types) and $nfeat$ (number of features).

The selection of $nselectedfeatures$ features from $nfeat$ is done by observing the centres matrix and choosing the features that provide the best discrimination between the classes. The choice is done by applying the Backward Elimination (BE) algorithm on the centres matrix (explained in detail in section 5.3). Several class-distance measures were compared in this study, including Euclidian, Fisher and Entropy measures. The best results, however, were obtained for Fisher's distance (when compared to grab sample types in the datasets considered).

For the classification stage, the KNN algorithm is used (refer to 5.4.1). The data is clustered according to its proximity to the centres. The classified class-labelled data is transformed from vectors to 2-D, plotting each point in its geographical position. The grab samples are then class labelled and also plotted on the final survey graph.

10.5.2 Unsupervised Classification

If grab samples are not available from the survey, an unsupervised classification method is used. Principal Component Analysis (PCA) is applied to the whole dataset. The number of principal components that is retained for the classification stage is chosen according to the amount of variance present in the remaining principal components (section 5.2.1). The result is a dataset with a reduced number of features. The Kmeans algorithm is used to cluster the data (section 5.4.2) into $ntype$ classes. The classified data vectors are then transformed to 2D and plotted against their exact geographical position.

10.5.3 Data Gridding and Visualisation

To view the final classified data, each window is assigned a certain number corresponding to a seabed class. The windows are then displayed in an image corresponding to their exact position

in Eastings and Northings. The data can be interpolated to present a general overview of the classification in the area (as in figure 10.14). It is worth noting that gridding is done in the final stage and not as a pre-classification step. This definitely introduces an error, especially when the sonar trajectory is a curving one. If that is the case, pings inside the curve will converge and the area is surveyed more than once. The pings on the outer side of the curve will be quite sparse. The next sections will discuss the problem in detail, give an example where the trajectory is a curving one, and propose a few methods that can be used to deal with this problem.

10.6 Results of Sidescan Bathymetric Classification on The Hopvågen Dataset

The Hopvågen dataset was introduced in section 2.3.2 and used for testing 1-D algorithms in chapter 6.

The classification scheme described in the last section was applied to this dataset. This section presents a brief analysis of the feature images, an observation of scatter plots around known areas and an analysis of errors and the relevance of features. Supervised classification is used in this section, assuming that areas around the grabs are of the same type. The grab areas are used for training and the dataset is then classified.

Due to system requirements, parts of the whole Hopvågen dataset were considered instead of the whole set. These parts included areas containing the grab samples, with about 100 pings to the left and right, in addition to a few areas in other locations.

10.6.1 Feature Image Analysis

The two types of images considered for feature analysis are the processed sidescan image and the bathymetry image. The method explained in section 10.4 was used over a series of pings in

these two types of images to derive feature images. To display the feature images, a gridding program was used to interpolate between the points and display the feature images versus their exact location (Eastings, Northings). The features investigated in this section are as follows:

- The mean of the amplitude of the processed sidescan image in windows of selected size (M).
- The standard deviation of the processed sidescan image in windows of selected size (M_s).
- The bathymetry standard deviation in selected size windows (D).
- The textural features consisting of the mean of the Wavelet transform energies (using DWT) over 3($nscale$) images ($S_{i,j}$), where i represents scale and j orientation.

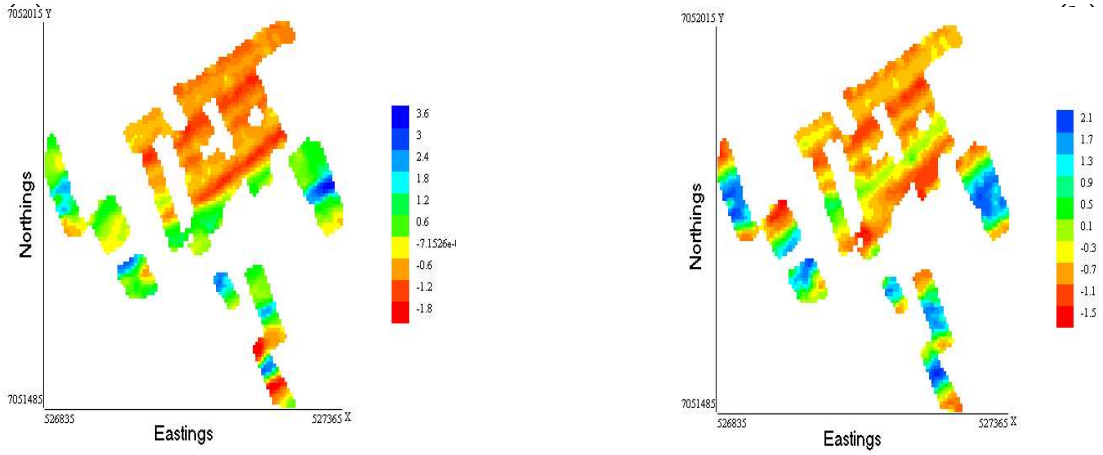
Figure 10.2 shows the mean (M) and standard deviation (M_s) of the amplitude of the Hopvågen area (windows are 50×10).

Figure 10.3 shows the bathymetry standard deviation (D) in the selected size windows. Thus, it indicates whether these areas exhibit bathymetric changes or not.

Figure 10.4 shows 6 textural feature images (S). Each image represents the image resulting from using a chosen wavelet scale and orientation. This figure considers only 2 scales leading to 6 images (same window size as above).

The Effect of Window size on Features

As window size is increased, the resolution of features is affected. Effects of irregularities are less visible in the final classified set if features are averaged in larger windows. Varying the window size also affects the boundaries between sediments in the final classified images. This is expected since larger windows have less resolution, and could contain two or more types.



(a) **(b)**

Figure 10.2: Using a window of (50 horiz,10 vert) pixels, (a) is the mean amplitude feature image (M) and (b) is the standard deviation feature image (M_s) of the whole Hopvagen dataset.

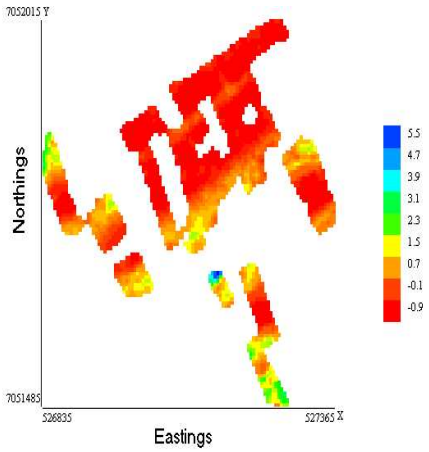


Figure 10.3: The distance variation per window (D) (the size chosen was 50×10).

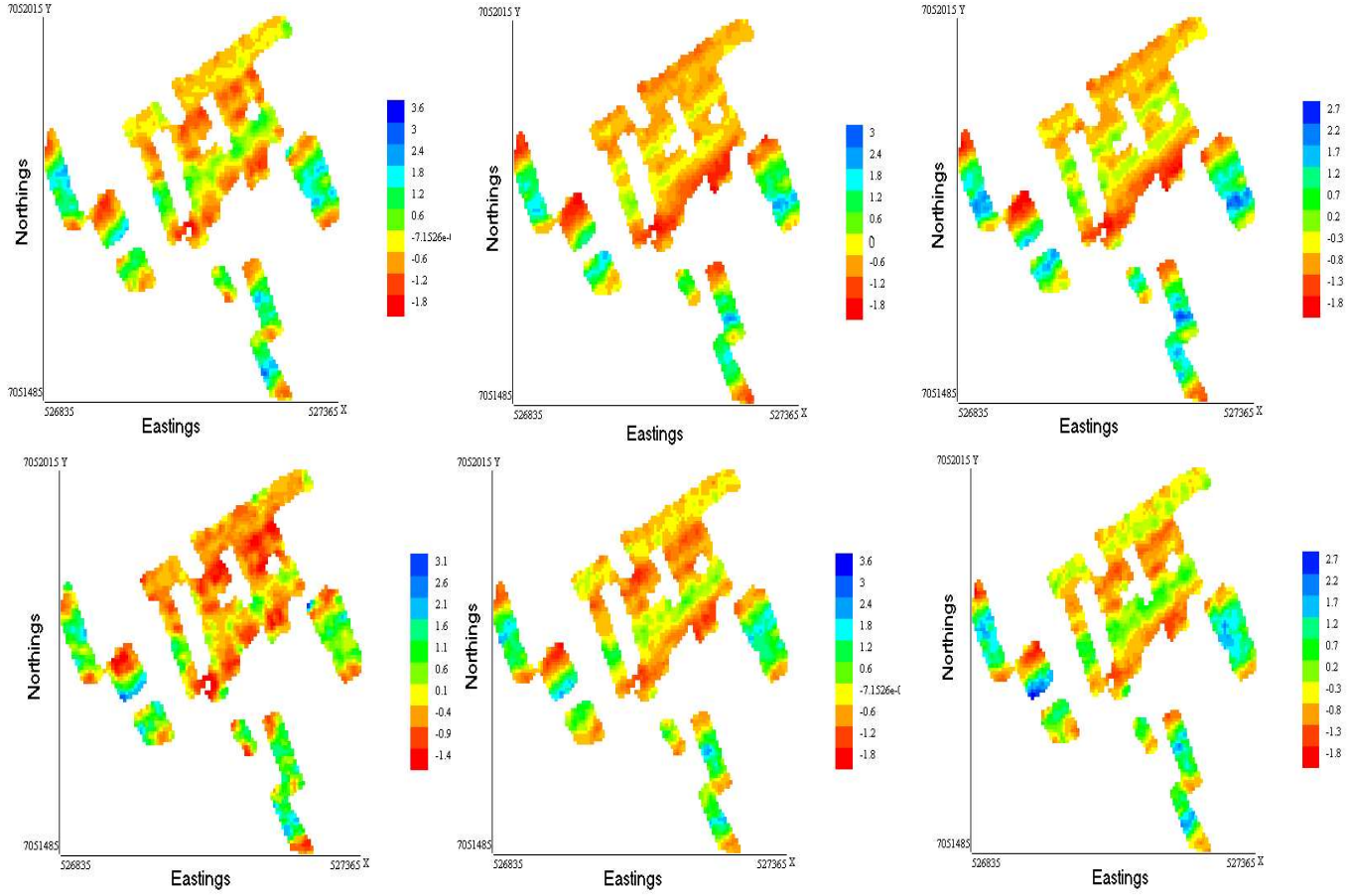


Figure 10.4: The Textural images ($S_{i,j}$) of the dataset considered over 2 scales leading to 6 images (since the DWT is considered). ($S_{1,1}$, $S_{1,2}$, $S_{1,3}$, $S_{2,1}$, $S_{2,2}$, $S_{2,3}$)

10.6.2 Classification Results

Scatter Plots and Analysis of Sediment Classes

The same 10 grabs used in chapter 6, table 6.3 are used in this section as 'known' areas. Around each grab, a number of windows is selected (15 in this case), to represent samples belonging to the same type. Selecting windows, rather than whole pings (as in chapter 6) can introduce more variability between samples. On the other hand, it improves the resolution of the final classified areas. To observe inter and intra class variations in a scatter plot, the BE algorithm using Fisher's distance was used to select the 2 features that provide the largest class separation. These were selected to be the bathymetry variation feature D and the texture feature $S_{1,3}$ ($nscale = 1$, $norient = 3$). The features for areas around the 10 grab samples are plotted in figure 10.5. This figure can be compared to figure 6.9 in chapter 6, and uses the same colour coding for each point. However, each point here represents features per window rather than per ping. From this figure, the following can be deduced:

- From observing clusters of points, it can be deduced that the medium sand class (in blue) clusters quite well (observing points 1, 8 and 9), so does the fine sand class (points 2 and 3-in green).

- As for the silt/mud class, points 6,7 and 10 cluster quite well. Point 5 however (red dots), seems to converge with the fine sand class (as in the ping analysis case), which could mean that the area there belongs to this class, and the grab sample was taken over a small non-representative area.

- The very fine sand class (yellow) seems to totally converge with the silt/mud class. This could be due to a wrong selection of windows in that area. Averaging over a ping in figure 6.9 allowed the separation of pings belonging to this class from other class-clusters.

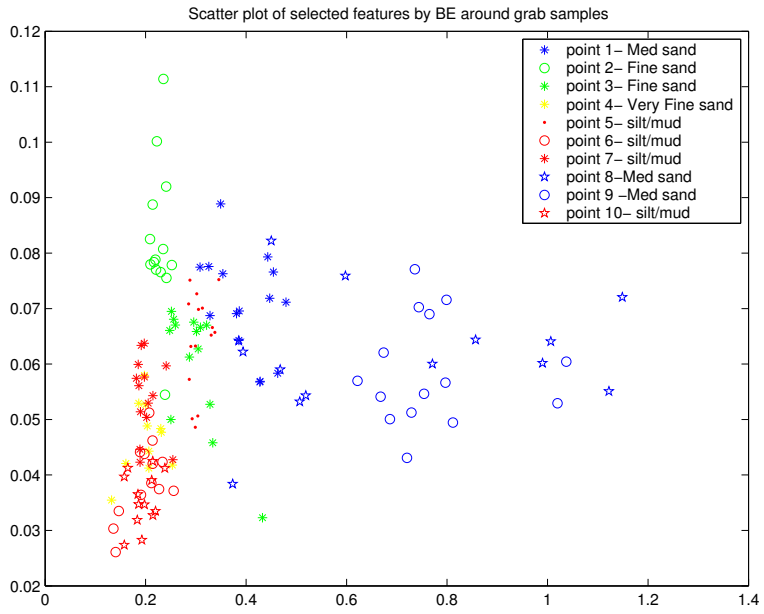


Figure 10.5: Features D versus $S_{1,3}$ for areas around the 10 grab samples selected from the Hopvågen dataset.

Classification Figures of Merit and Relevance of Features

To get an idea of how well the features selected can help in separating sediment classes, a KNN classifier is used in this section. The set of windows around the 10 grab samples is collected in a training set with a label assigned to each window. The k-fold cross validation method is used as a method that can give a figure of merit on the performance of the classifier. This method was explained in section 6.4.1. The training set is randomly arranged, then divided into 11 subsets of 18 samples each. One subset is used as a test set and the others for training. This is repeated for all the subsets and the results are averaged. The experiment is repeated with a KNN classifier, varying the following parameters:

- The number of inputs (*ninput*) to the classifier. For each value of *ninput*, BE selects the most discriminating features. The results for varying *ninput* from 2 to 8 (all features but one) is given in table 10.1.

- The k value for the KNN classifier.

Number of features	Features Selected by BE
2	D and $S_{1,3}$
3	D , $S_{1,3}$ and $S_{2,3}$
4	D, $S_{1,2}$,$S_{1,3}$ and $S_{2,3}$
5	D, M, $S_{1,2}$,$S_{1,3}$ and $S_{2,3}$
6	D, M, $S_{1,2}$,$S_{1,3}$, $S_{2,3}$ and $S_{2,1}$
7	D, M,M_s, $S_{1,2}$,$S_{1,3}$, $S_{2,3}$ and $S_{2,1}$
8	D, M,M_s, $S_{1,2}$,$S_{1,3}$, $S_{2,3}$, $S_{2,1}$ and $S_{2,2}$

Table 10.1: BE selected features from the Hopvågen training dataset.

Having a look at table 10.1, some interesting facts can be observed. Feature D, which is that of bathymetric variation is selected as a discriminant feature from the beginning. Textural features are selected in the next few steps. The amplitude standard deviation M_s is only selected, when the algorithm chooses 7 or more parameters out of 9. This could give a guide on the importance of each feature for the classification process. Figure 10.6 shows the correct k-fold cross validation rates for k varying from 1 to 10, and the number of inputs varying from 1 (using only D) to 9 (all features). What is interesting is that using only 2 parameters (D and $S_{1,3}$), the results are quite close to using the whole set. The best results are for using the whole set of inputs and $k = 4$ for the KNN algorithm.

Using the k-fold method to observe classifier performance, it is interesting to observe the effect of using each type of features for classification on its own. This is given in figure 10.7. If amplitude information is considered on its own for classification, the results are below 65%. This confirms the results of chapter 3, where using amplitude on its own caused overlap between several sediment types. Adding bathymetric information improves the results by 20% in all cases considered, which shows the advantage of using such a feature. By adding textural information,

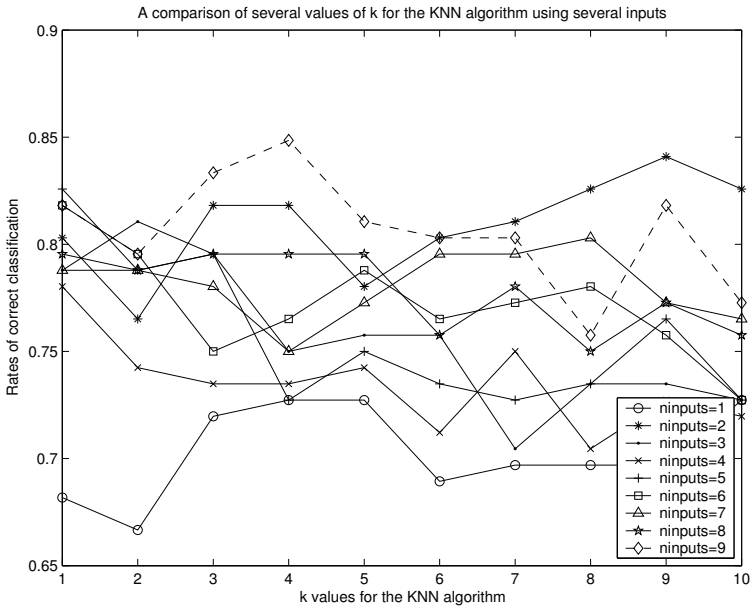


Figure 10.6: Success rates using the k-fold cross validation method on the KNN classifier. Each line represents the success rates for a number of inputs (*ninputs*), plotted versus the value for k in the KNN algorithm.

the results improve and attain the highest rates (up to 85%).

To observe misclassification rates between classes, a training set and a test set are selected from the above training set (of areas around the grabs). The size of the training set is 16, balanced between the four sediment types, to assign class centers. The test set is made of 97 randomly selected samples. The best case scenario of figure 10.6 is used for classification, so all nine features are used with $k = 4$. The confusion matrix is given in figure 10.8. The rates obtained are lower than those obtained when whole pings were used for classification (in chapter 6), but again the gain is in the resolution of the classified areas. Note that the test dataset is quite unbalanced due to the nature of the whole dataset used, with 32 samples of medium sand, 21 of fine sand, 8 of very fine sand and 36 of mud/silt. Most confusions are in the very fine sand class as half the samples are misclassified. The classes of mud/silt and medium sand are in general well classified.

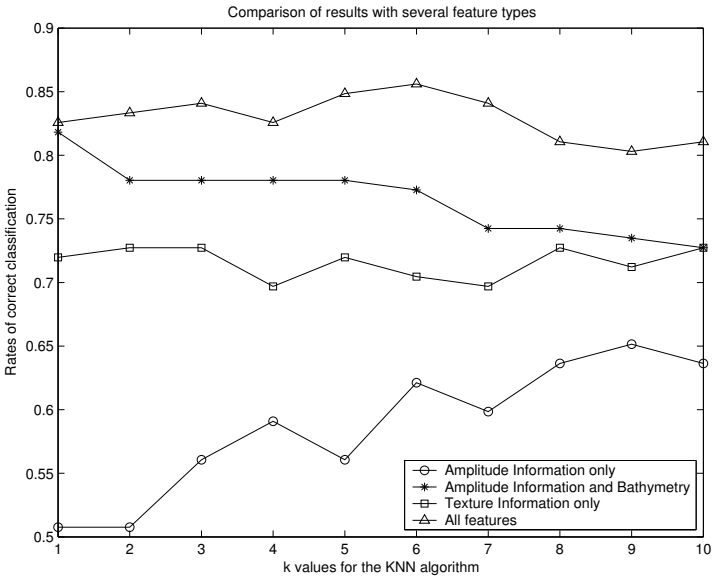


Figure 10.7: Success rates using the k-fold cross validation method on the KNN classifier. A comparison of success rates using amplitude features only, amplitude and bathymetry, texture features and all the features.

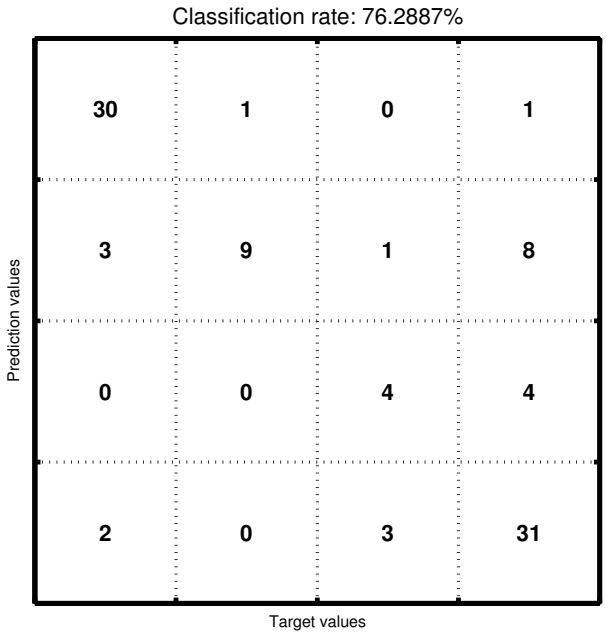


Figure 10.8: Confusion matrix on 97 samples from the Hopvågen dataset, using KNN with $k = 4$ and $ninputs = 9$.

Classification Results on the whole dataset

The training set consisting of samples around the 10 grabs is used to train a KNN classifier and classify the whole dataset. The optimal results from before are used as parameters of the neural network. The number of inputs is chosen to be 9 with $k = 4$. Figure 10.9 shows the results of classifying the whole dataset. The grab samples that are used for training are numbered from 1 to 10. It is interesting to observe the areas around the grab samples that were not used for training (i.e the grabs that are not numbered in figure 10.9).

The medium sand class is in general well-classified and the classification matches the type of the close-by grab samples. The very fine sand class is in general misclassified. This is most probably a training error. As observed in the scatter plot of figure 10.5 and the confusion matrix in figure 10.8, the samples belonging to this class and that were selected for training almost totally converge with the silt/mud class. The mud/silt grabs in general match the nearby classified areas.

A problem with the final classified area, and since spatial information is not used (i.e windows are classified according to their features regardless of their position in the survey), is that of variability within small areas. Using spatial information as done in section 8.4.5 could help smooth the final classification results, and solve this problem.

Errors of the Method

There are several factors that can lead to misclassification when dealing with a large survey and training on a selected set of grab samples. Here is a list of a few:

- As explained in chapter 2, sonar motion could lead to errors, especially when it introduces patterns on the sidescan images. In this work, sonar motion is corrected for before the pre-processing stage. Although this is good enough for the Hopvågen dataset, it might not be enough to correct for some sonar datasets with severe sonar rotation effects. A solution then would be the application of image filtering techniques on the sidescan images

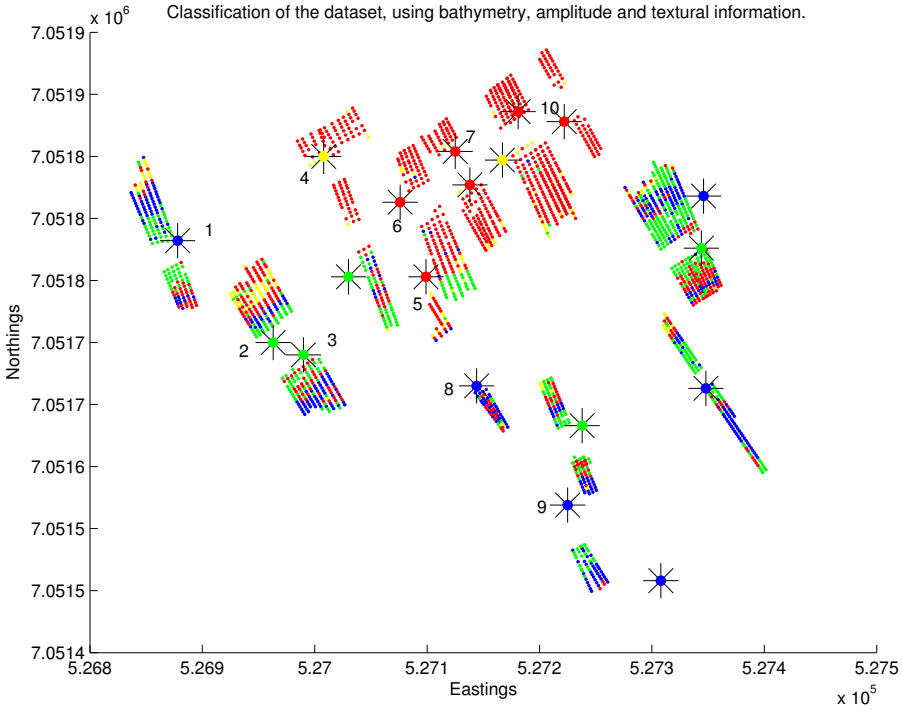


Figure 10.9: Result of classification using the 2-D features. The Backward elimination algorithm was used using Fisher distance between classes. The grab samples are indicated with circles around them. The grab samples used for training are numbered from 1 to 10.

before analysis.

- Sonar direction with respect to seabed features can lead to misclassifications especially with large scale features that appear different depending on sonar direction.

- Incorrect Georeferencing and gridding, especially in areas of sparse pings can be a very important cause of misclassification. In the images used for analysis (the processed amplitude and the bathymetric image), the pings are plotted in the order they are received. The final results are gridded to their exact GPS position. This is a problem for areas where the trajectory of the sonar is a curving one, or where pings overlap due to sonar rotation. The effects of sonar trajectory will be observed in the next section, where the sonar trajectory is chosen to be a curving one.

- Size matters when it comes to window size selection, which can lead to averaging features in each window although they might not be from the same sediment class. Small windows were selected for this dataset, but a further observation of the effect of window size on feature selection and classification performance is required.

Despite these errors, and as proved by the scatter plots and the classifier results presented in this section, this work offers a method that can benefit from bathymetric, textural and amplitude information to correctly classify seabed areas.

Comparison with Other Techniques

The advantage of the 2-D approach over the 1-D is the location of boundaries between different sediment types. However, the resolution in the 2-D analysis is dependent on the chosen window size. The final classified image (figure 10.9) can be compared with figure 6.12 in chapter 6 (using the same colour-coding). Note that for the 1-D case, the classification was done per ping. Each ship position 'sees' one type of sediment, and figure 6.12 interpolates this over an area of $50m$ starting from the ship position. In figure 10.9, the classification is done per window and geo-referenced to its exact position in the survey.

In section 6.4.2, a comparison was made between the 1-D classification results, and the results of using the ECHO-Plus system for classification. Important factors to consider when comparing the 2-D results (in figure 10.9) and the Echo-Plus results (figure 6.13) are the following:

- The Echo-Plus method uses areas under the sonar for classification (single ping method as given in appendix B), whereas we are using the whole ping information for the 2-D classification scheme.

- Class labels are given by Bates et al. [15] according to clusters after doing an unsupervised classification. In this chapter, we used the Udden-Wentworth table [216] to assign class labels for areas around the grabs.

- In the Echo-plus results, an extra class is used (figure 6.13 in yellow) which is the coarse sand class. Grab samples for this class were not available for this work. As mentioned in section 6.4.2, the 'red' class used in the Echo-plus survey is divided into two subclasses in this work, since the Wentworth scale divides sediments belonging to this class into two subclasses: mud/silt (in red) and the very fine sand class (in yellow).

Due to the paucity of ground truth data, one of the ways to assess classification performance in the whole dataset is to observe areas near grab samples that were not used for training. The 2-D classification scheme (figure 10.9) shows good performance for areas near the grabs as explained earlier. The 1-D classification scheme (figure 6.12) does not have a sufficient resolution to note differences within a ping, as features are averaged per ping. The Echo-Plus method shows some confusion between classes of fine sand and medium sand. By comparing the roughness and hardness images in figure 6.13, we can note several mis-classifications near grab samples.

10.7 Results of Sidescan Bathymetric Classification on the Laxford Dataset

10.7.1 Introduction

The Laxford dataset was introduced in 2.3.2. This study considers only a small part of the whole survey. The part considered is a part of the outer Loch known to be varying between circalittoral rock or boulders and circalittoral sand and silty gravel. It is displayed as a rectangle in figure 2.7 and magnified in figure 2.8 (chapter 2).

The main aim of this section is to use an unsupervised classification method to cluster the two classes and provide boundary location between them.

10.7.2 Feature Image Analysis

The window selected to calculate features was of size 50×10 pixels corresponding to a resolution of $3.65m$ in the along track direction and $2.5m$ in the across track direction. Figure 10.10 is the sidescan image of the considered area in Loch Laxford. It is worth noting that the sidescan pings are plotted one after the other as received and not gridded to the exact GPS position as in figure 10.10. As figure 10.13 shows, the trajectory of the sonar in this area of the survey is a curving one. The pings towards the inside of the trajectory overlap and the outer ones are sparse. Thus, there are some errors in figure 10.10. In this figure, regions A and B contain pings that cover the same seabed area, although they are plotted in the sidescan image as separate areas. It is interesting to observe the performance of the classifier for such a case. The time-received information (both amplitude and bathymetry) is the one used for analysis. So features are calculated for corrected backscattered and bathymetric images, without taking gridding into consideration. For display purposes, values for windows from overlapping pings are averaged, and values are filled in for sparse pings. This is shown in image 10.11 for the mean (M) and standard deviation (M_s) of backscattered amplitude, and in figure 10.12 for the

distance variation (D) feature and one of the texture feature images (scale 1). The features in these images are gridded to their exact position.

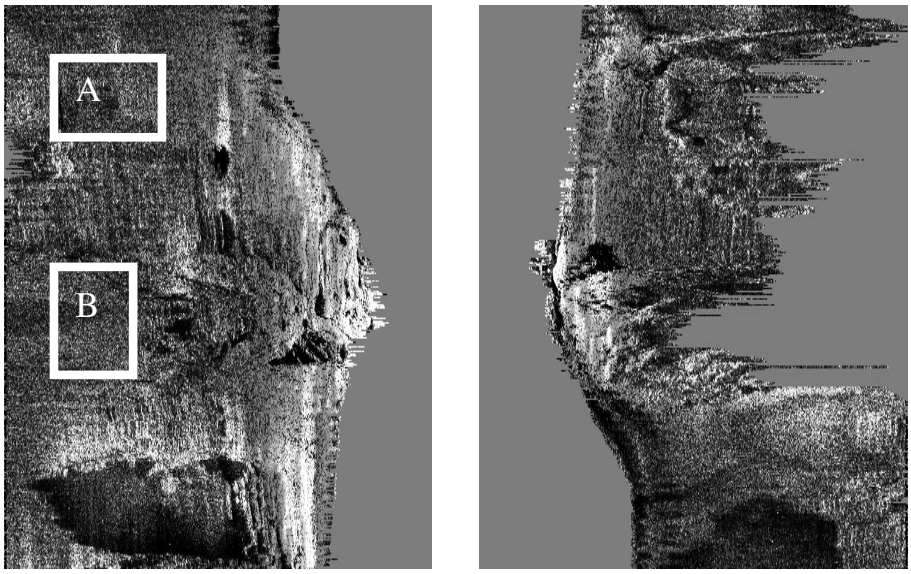


Figure 10.10: The sidescan images of the considered areas in Laxford. The pings are plotted one after the other as received, and the exact position of the points is not taken into consideration in these plots. Areas A and B represent areas that contain information from overlapping pings. An ideal gridding would plot these areas as one versus the exact GPS position.

10.7.3 Classification Results

PCA was used to extract features from the un-gridded feature images and the resulting PCA images were used as final feature images. The Kmeans algorithm was used to cluster the data into 2 classes representing circalittoral rock or boulders and circalittoral sand and silty gravel. Each window of the data is finally classified and displayed versus its exact position in the survey (figure 10.13, with no interpolation). The area shown in the box in figure 10.13 shows windows classified as both rocks and gravel (red and blue). The reason is that, for analysis of the dataset, the amplitude and bathymetric images were used without exact gridding. Thus, areas (such as A and B in image 10.10) contain pings from the same area (the rectangle in figure 10.13), as

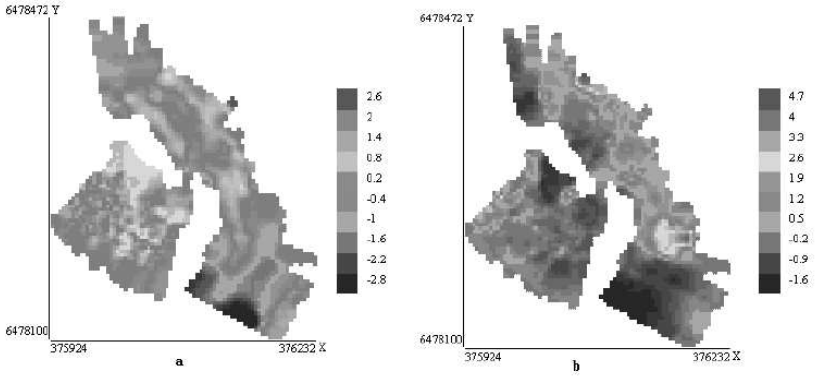


Figure 10.11: The amplitude feature images of the selected area. (a) is the mean backscattered processed amplitude and (b) is the standard deviation.

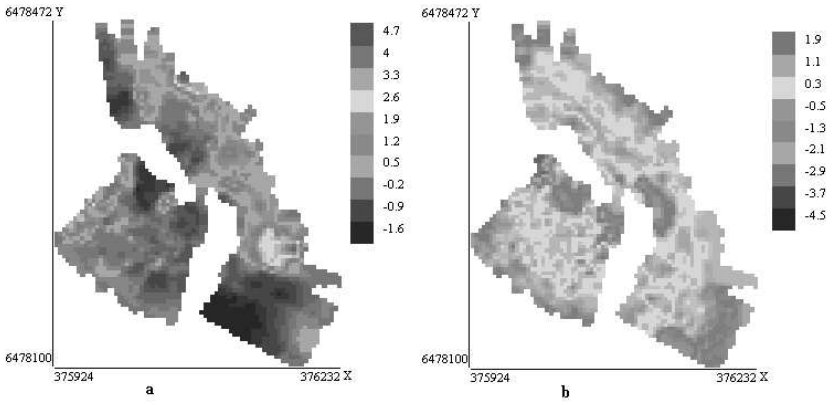


Figure 10.12: The depth variation of the area in figure (a). One of the textural feature images (scale 1) in figure (b).

well as other pings. This is definitely a problem when curving sonar trajectories are considered for analysis. Several commercial system avoid classifying these areas.

In this study, we decided to average points per GPS position after classification, as given in figure 10.14. Areas between sparse pings were interpolated. Areas of red and blue in figure 10.14 indicate areas belonging to each of the classes. The colours in between can be regarded as a way of expressing the uncertainty level.

Figure 10.15 shows the area as classified in [14] into the two classes.[3]

A method of data validation that can be used for this dataset is to compare the final classified data (figure 10.14) with the sidescan images (in figure 10.10) to draw boundaries between the sediment types according to differences in texture. However, other factors must be considered such as the corrected amplitude feature image (figure 10.11) and the bathymetric variation image (figure 10.12, which proved to be of great importance for the Hopvågen dataset). Although this approach is subjective, it is probably the only method that can be used where the underwater area is unknown. By comparing the final classified area to the sidescan images, the rock class which is coarser than the surrounding areas in the sidescan images is mostly classified correctly (colour-coded as the red class in figure 10.14).

In addition to the errors due to gridding, the error-causing factors that were mentioned in the previous section are applicable in this case. These errors relate to sonar motion, sonar direction and window size selection.

10.8 Conclusions

This chapter presented a novel classification system that makes use of several features provided by sidescan bathymetric sonar data. Bathymetric information, used generally for amplitude correction, was used to derive features that proved to help in sediment discrimination. The

[3]The problem here is that we were not provided with the exact class label per survey point from [14], so our comparison was to the sidescan images as well as the image provided (figure 10.15).

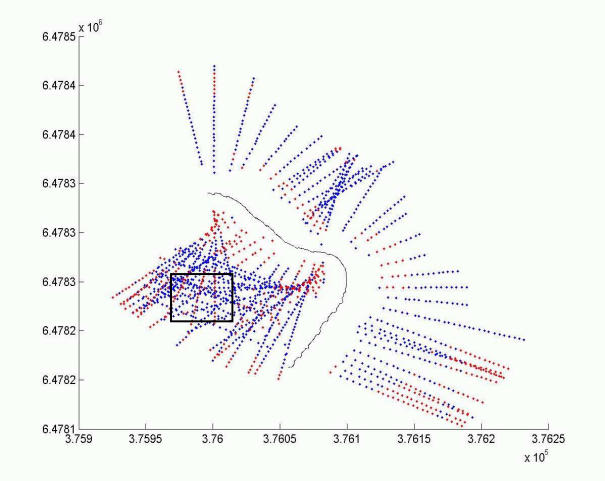


Figure 10.13: The selected area classified into two types with no interpolation between the points. The classes are rocks (red) and sand and gravel (blue).

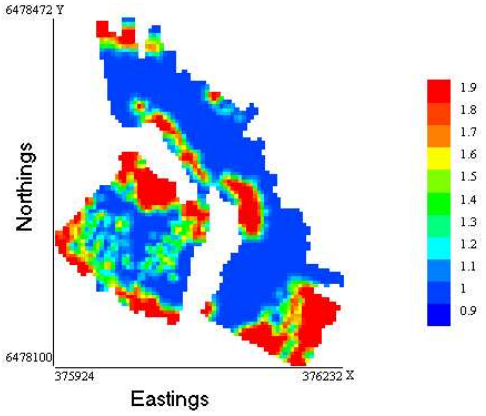


Figure 10.14: The interpolated classified area in Laxford. The classes are rocks (red) and sand and gravel (blue). The colors in between indicate the windows classified between these two types.

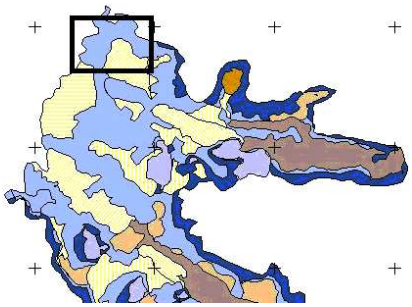


Figure 10.15: The area as classified in [14] into two classes of circalittoral rock or boulders and circalittoral sand and silty gravel.

corrected amplitude data is used to extract amplitude and textural features calculated per window. The suggested method can select the features that offer the most class separation by using training areas near 'known' grab samples.

When tested on the Hopvågen dataset, the method proved to have good classification rates. The k-fold cross validation method was used to assess classification rates by varying the training and test set from areas selected near the grab samples (of known types). The importance of using bathymetric features for classification is highlighted in figure 10.7, where classification was improved greatly when bathymetric features were included with amplitude information. Adding textural features also caused an improvement of classification rates. The method provides better boundary location between sediments compared to using the 1-D techniques presented in chapter 6. However, the method has a few problems. The images used for analysis use pings plotted as they are received during the survey.Whereas the classified areas are plotted versus their exact GPS position. This leads to problems as highlighted in section 10.7 for the Laxford dataset. A way around this problem is to use proper gridding programs from the beginning, and use the gridded images for classification. This, however, will depend on the gridding program iteslf, and its ability to interpolate sparse data and average overlapping data. The problem does not end here. Correct gridding might affect some features appearing in the sonar images (such as shadow areas). A 3-D modelling of seafloor features might be used to

correct for this problem. This is beyond the scope of this thesis, but it would be necessary to deal with some of these problems, if the method were to be applied to curving sonar trajectories in sidescan-bathymetric surveys.

Better classification is expected if the methods in this chapter were applied to datasets with better bathymetric resolution, such as multibeam datasets or sidescan bathymetric sonar datasets with improved resolution.

Chapter 11

Discussion and Conclusions

11.1 Summary and Analysis of Results

The work that has been carried out in this thesis can be divided into several ideas: An analysis of sonar parameters and the correction of sidescan bathymetric data using the sonar and survey parameters, an investigation of features that can be used for sonar classification and the development of a classification method that can use the extracted features to classify sidescan bathymetric data (sidescan data was used for validation and testing of the features selected). Sonar data can be considered as a series of 1-D pings, or an image of consecutive amplitude and bathymetric information.

Analysis and Correction of Sonar Data

The thesis analyses several of the factors that affect backscattered sonar data. These factors relate to sonar motion, beam shape and losses due to range and bathymetry. A model that corrects for these parameters was used to correct for sidescan bathymetric data. Several studies use a precise modelling approach to deduce the variations of backscattered amplitude due to sediment type [3, 148, 147, 36]. However, the bathymetric signal used for such models should be quite precise, and should represent an exact variation of the seabed. The depth signal used

in this work was not of adequate resolution, but could give an idea of the general variation of the seabed bathymetry. Chapter 3 investigated the use of corrected amplitude features for classification, and confirmed Goodfellow's [93] results: backscattering amplitude can not provide efficient class separation for sidescan bathymetric datasets if used alone as a feature for classification.

Feature Extraction

Bathymetric variations at both microbathymetric and larger scales cause the appearance of features at different scales in received backscattered sonar signals. Methods of extracting these features were explored in this work. Wavelet transforms at different scales can give an indication of the presence of discriminant features due to a certain sediment type. Due to the high dimensionality of the wavelet transform, a feature extraction technique was required. The one used in this work considered windows over selected areas of the signal and measured the wavelet mean energy and standard deviation per window over several scales. A careful selection of the windows was done to avoid areas where the data is clipped, as well as near nadir and very far points. Feature selection methods were used to find which scales of the wavelet transform provided the best class discrimination, making the process of feature selection an automatic one. The chosen bands were observed in scatter plots and used as neural network inputs for both the training and testing stages.

The method was first applied to the classification of sidescan data from Pace's dataset (48 kHz processed sidescan). The six sediment classes (sand, mud, clay, gravel, stone and rock) were well separated when selected wavelet features were plotted in scatter plots. The dataset was divided into training and testing data and several experiments were conducted with MLPs for data classification using wavelet features. The maximum success rate was 98.5% when the optimal number of inputs was selected. The success rate was used as a guide for optimal input feature selection.

Pace's dataset was used as a guide for selecting features from the sidescan bathymetric

dataset (Hopvågen). The first step was to observe wavelet features of pings selected from areas near grab samples. The scatter plots showed good separation between the sediment classes in general. However, some overlap was noticed between classes of very fine sand and silt/mud classes. This could have been due to a training error or the fact that in the process of selecting features per ping, information is averaged and class boundaries are lost. These issues were analysed in chapter 6, where figures of merit were discussed to evaluate classification performance, and correlation between sediment types and selected features were studied. Overall, a clear trend was apparent, and a correlation is found between multiscale features and grain size (indicating sediment type). Using a cross validation method, such as the k-fold method to divide the samples (near the 'known' areas), into training and test sets proved that classification rates were quite high in most cases considered (more than 90%) .

To classify the whole dataset, areas near the grabs were used as training data. Since the training data set was of a small size compared to the whole survey, clustering techniques had to be used instead of MLPs (or other neural networks requiring larger training sets). The KNN algorithm was used to cluster the data using the training points as cluster centres. The results of classification correlated with most of the grab samples provided. The final classified datasets using this method were compared with the results of using a single echo sonar on the same area. The results were similar in many areas but showed a few differences. This might have been due to the fact that features were described per swathe in this study, whereas they are generally described for the area directly under the sonar in echo sounder studies. The way the sediment classes were divided (we used the Wentworth scale) could have also been another reason for classification differences.

One advantage of using a 1-D signal analysis, is that the method can be extended to an on-line system due to its speed, and processing of each ping on each own. However, the classification degrades rapidly if several sediment types are contained in the same ping. 1-D data is usually too noisy to identify sediment boundaries and an approach using consecutive pings has to be considered. Thus the next part of the thesis considered 2-D sonar images and

bathymetric maps for feature analysis.

Sonar Textural Features

Looking at 2-D sonar images, it has been shown that the use of gray level amplitude features (local amplitude) fails in many cases if used alone for classification [52], mainly because of the high dependence on bathymetric variation. Textural analysis extends 1-D spectral features to 2-D and offers features that could aid in sediment discrimination as well as boundary location, so it was investigated further in this work.

Gabor filters and wavelets were compared for this work to extract the multiscale features present in backscattered seabed signals from sidescan and sidescan-bathymetric images. In both cases, filter parameters were adapted to provide a better classification. Due to the large amount of data usually present in sonar surveys and the general paucity of grab samples, feature selection techniques were used in this work after the filtering stage rather than during it. Because of noisy data and the lack of ground-truth information in general, adapting the filter parameters to selected areas during the filtering stage can be problematic, as is the application of wavelet statistical models on training area images. The textural features extracted to describe the sediment areas are the mean and standard deviation of the filtered images (in both the Wavelets and Gabor cases) over selected windows and scales. Before moving to sidescan bathymetric data, texture classification methods were tested on an artificial image made of consecutive sidescan sonar pings from Pace's set (the six types) and an image of artificial coral reefs. The classification was done per window of user defined size. For Pace's set, several parameters were varied to observe their effect on the classification. The following was deduced:

- An optimal window size should be chosen to compromise between small windows that do not contain enough information, and large windows that average a lot of useful information.

- An optimal number of decomposition scales should be selected as a large number would

also lead to redundancy.

- Boundary problems arise when convolving with Wavelets and have to be corrected.

When applying the DWT (discrete wavelet transform) to Pace's dataset, the use of both wavelet standard deviation and means improved the results compared to only using the latter. The best classification rates were obtained when the Backward elimination (BE) algorithm was used to select features (using Fisher's criterion for class separation among several methods tested). The results were 98.9% correct classification for the optimal window size and scale chosen.

When Gabor filters were used to classify the same dataset, the parameters varied were the number of scales and the number of orientations per scale. An optimal case was also chosen, where the orientations and scales select the most discriminating information. The best results were also obtained for the Backward elimination algorithm using the Fisher criterion ($\sim 97\%$). Adding more orientations per scale improved the results in all cases considered (in some cases up to 20%). In general, the best results from using the DWT and Gabor filters were close.

For the artificial coral reef dataset, the aim was to separate two different types of reefs from each other and from the surrounding sediment. The method was divided into a supervised part, where windows were selected in the image and used to train a KNN neural network, and an unsupervised part using Kmeans clustering. The performance of both methods was observed in each case. For both the DWT and Gabor filters, parameters were varied and the final classified images were compared. The main deductions were that the choice of an optimal number of scales is important in all cases considered, so is the selection of an optimal window size. On the other hand, and unlike the Pace dataset, the increase of the number of orientations per scale (Gabor filters) did not play a major role in improving the results, probably because of the random orientations present in the dataset. Dyadic and non dyadic scales were compared for Gabor feature extraction and their effects on the results were not significant. Neighbourhood information was used to finally obtain a classification of the coral reef images that includes

information describing the position of each window with respect to surrounding windows. This method gave more visual correlation between the final-classified coral reef datasets and the expert-classified images.

The use of texture classification methods on sidescan datasets was used as a guide for the selection of a method to classify sidescan bathymetric data.

Sediment Classification by using 2-D sidescan bathymetric sonar features

The use of bathymetry was extended in this thesis to provide features that can aid in classification of high frequency data (it has been used before to classify low frequency data (10-15 kHz), and for the detection of large features in the seabed such as ocean ridges and abyssal hills). Several bathymetric features were explored. Bathymetric slope in selected windows, and average bathymetry did not provide much information about the nature of the seabed under study. However, the variation of bathymetry in chosen windows proved to be useful to discriminate rough from smooth seafloors (rocks and mud for example). Thus, the bathymetry standard deviation per window was introduced as an extra feature.

Bathymetric features were combined with amplitude dependent features (spatial and spectral) for a better classification of sidescan bathymetric data. Large surveys were divided into user defined windows, to improve data retrieving and feature selection. The features selected for each window were the following:

- The average intensity amplitude and standard deviation per window in the 'processed' images (backscattering strength images).
- The bathymetric standard deviation per window in the bathymetry image.
- The mean and standard deviation of 2D wavelet transform coefficients obtained from filtering the 'processed' amplitude images. Wavelets were selected instead of Gabor filters since the results of using the two methods were very close for sidescan image analysis, and wavelets proved to be faster. Note that the images used for textural analysis are

the 'processed' amplitude images rather than the 'raw' images generally used in sidescan-texture studies.

These features were combined in one feature vector per window in the survey. In datasets where grab samples were available, areas around them were used as training areas (supervised classification), otherwise clustering techniques were used to classify the data (unsupervised classification). The Backward elimination (BE) algorithm using Fisher's criterion was used for the supervised case and the Kmeans algorithm for the unsupervised case.

The classification was tested on the Hopvågen dataset using the provided grab samples as training data. The best results were obtained for BE using Fisher's criterion. Good correlation was observed between the results obtained and the given grab samples. Inter and intra class variations were analysed, and samples from the best scales selected by BE were used in scatter plots. The classes of fine sand, medium sand and mud/silt cluster well in general. The samples used to test the very fine sand class completely converge with the silt/mud class. This could be due to a wrong selection of windows near the grab sample, where areas are assumed to be of the same type (but in reality they can vary). The k-fold method is used as in the 1D case to test classifier performance.

A test is done on using amplitude features alone, then adding bathymetric features and doing a k-fold KNN classification on a training area. The results prove that bathymetry improves classification by 20% in some of the cases. Adding textural features improves classification rates and in general, the best classification rates are obtained when all three types of features are used. However, due to dividing the dataset into small windows, less averaging is done than in the 1-D case, and the rates of correct classification are lower due to data variability. The resolution of the final classified data is of course much higher than the 1-D case, and the method offers a better way of boundary location. The images used for analysis are of time received sonar information, whereas the final classified data is plotted versus exact GPS positions of the classified windows. This can cause a lot of variability in the results, especially if the sonar

trajectory is a curving one. To observe this effect, a part of the Laxford dataset with a curving trajectory is chosen for this work. Unsupervised classification is used for this dataset, and a problem while analysing time-received data is that several pings that are considered different, actually cover the same region in the survey. These pings, however are mixed with others, and the sidescan image does not exactly portray the regions surveyed. Since in this work, features are averaged per window, the effect might be lowered. In final classification areas, interpolation is done between windows, which in a way, averages results and fills in empty areas. If final results are colorcoded (as in figure 10.14), the color-coding can be used to observe areas of certainty in classification (areas in red or blue), and areas of ambiguity (colours in between).

11.2 Discussion and Errors

Bathymetric information, when provided in sonar surveys, is generally only used for the correction of sonar signals. This work provided a new method of using the corrected signals for feature extraction, as well as investigating features from the bathymetry itself. The importance of using bathymetric features was explained in chapter 10, where these features improved the overall classification of a dataset selected from 'near-grab' areas. However, this method depends on the resolution of the bathymetry and would probably deteriorate if the bathymetry was of a lower resolution. The method is applicable to both sidescan-bathymetric and multibeam sonars. If bathymetric information is not available, the method would rely only on amplitude and textural information.

Some features selected in this work might be irrelevant for classification, an example is areas of similar texture and different backscattering strength, where extracted textural features would not be important. In this case, the feature selection algorithms would select the most relevant features and ignore the redundant or irrelevant ones. This automatic way of selecting features avoids selecting features per sediment class, or fine-tuning feature selection to the properties of the sonar used in a certain survey.

Although the method showed good performance and correlation with the sediment types considered for this study, several sources of errors can be observed. Some of these errors are present in most sonar surveys, and developing a general solution for them remains an open problem in the sonar community. Some of these errors are:

- Look direction of the sonar with respect to seabed features. This issue was highlighted in section 3.2, referring to several studies that have addressed this problem. In this work, no specific modelling of seabed features versus sonar direction was attempted. However, features were averaged over a window (or a ping in chapter 6) which makes extracted features less variable with directionality changes, although it decreases the resolution of the survey classification.

- Sonar motion can be a source of error as discussed for the Laxford dataset, in the previous chapter. Data is analysed in its time-received order, assuming a rather uniform trajectory. The final results are gridded versus their exact GPS position. If the sonar's trajectory is a curving one, errors can be introduced as pings can be overlapping or sparse. This was explained in detail and a few methods that correct for this effect were suggested.

- Averaging features per window, or ping can be a source of problems, especially if the window (or ping) includes several sediment types. In addition to that, another source of error could be using a fixed window size for each survey in this work. A method of adjusting window sizes to obtain a maximum class separation should be developed. Some features appear at certain scales, and considering smaller or larger scales for analysis can be misleading. There are several methods to select the scale of features for texture classification, examples are [116, 117] as well as other references given in chapter 7. However, an analysis of optimal scale per feature would be necessary. Window size for one of the features selected might not be the optimal scale for another type of feature (an example is a difference between bathymetric feature scale and texture scale), a further investigation of this issue is required.

- The model used in chapter 3 to correct for sonar variations is an approximation model of seabed backscattering and is quite a standard model used in the sonar community. Further study of this model is required, since some assumptions (such as insonified area variation) might be invalid in some cases.

11.3 Future Work

Seabed classification is of great importance to many industries and research areas. Its application in dredging, environmental surveys and underwater archaeology are only a few examples. This work attempted to shed the light on learning from sonar data for a better seabed classification, using features present in both sidescan amplitude and depth graphs. This section enlightens some of the improvements and future applications that could be done to this work.

A significant difficulty encountered in this work was data validation. Validation with more datasets would help to improve the algorithms as well as verifying the results. However, the problem is that not every underwater survey is provided with grab samples. Surveys that are well ground-truthed would be of great importance in testing the algorithms of this thesis further (both the 1-D and 2-D approaches). The use of an experimental tank for measuring backscattering from different materials under controlled conditions would also help observe the relationship between sediment type and backscattering features.

The correlation between seabed parameters and features selected in this work is also an issue to be observed further. A parametric approach could be used (as a first trial in a controlled environment) to find the relationship between wavelet features and seabed parameters including the following : roughness, porosity, seabed depth and absorption of sound. What would be really interesting is to observe how these parameters affect the specular (frequency domain) features of the reflected signals (a lot has been done on amplitude profile measurements [171], but frequency features remain an area of further investigation).

The lack of processing speed was a factor hindering the 2-D method (for example to the

whole Hopvågen dataset). Faster and more efficient processors could speed the process and lead to the application on more datasets (the files were quite large, each survey ship line was about $150Mb$).

Due to the speed of computation of the wavelet transform used for feature extraction in this work, online application of the method is possible. The online application (for the 1-D case) would include preprocessing the pings according to the bathymetry, performing the wavelet transform and storing the coefficients. After this step, only the coefficients are stored instead of the whole sonar signal, thus, saving a large amount of space. The classification stage, however, can be done offline, or by using clustering techniques online, where each new point is added to a known cluster. For the 2-D case, a set of consecutive pings is used to derive features. Thus, an online 2-D application, would mean the processing of a set of pings (50 for example), as new pings are being registered. Features would be grouped per cluster, and the survey colour-coded according to the clustered sediment types. An online application of the method is possible since the user defined parameters (such as window size and number of clusters) can be set a priori, and both the processing and classification processes are totally automatic.

The visual display of surveyed and classified areas is also an area to be improved if there is an interest in marketing the classification system as a product. In this work we used extra programs (Matlab, etc..) to plot each classified window in its exact GPS position, but to be marketed as a seabed classification software package, visualisation algorithms should be included.

This work tested several feature selection and classification techniques. For feature selection, methods such as the Backward elimination algorithm and the PCA were compared. However, the area of feature selection is a very developed research area on its own and there are several techniques that can be used to optimise the selection of features [22, 113], including 'Floating Search Methods' [173, 191] among many others. Their application for sonar feature selection is an interesting future application requiring further investigation. Feature classification in this work included the use of MLPs and clustering techniques such as Kmeans and KNN.

Optimisation can be done to the Kmeans algorithm to improve the classification process, an example is the Kmeans-Fisher classifier used by Clausi [50]. These algorithms can be applied to improve the classification scheme.

In chapter 8, neighbourhood information was used to add certainty to the classification of a window if its 'neighbours' belonged to a certain class. Distances between feature vectors and class centres were used to obtain a measure of 'closeness' to the class centres, using the training data. This method was investigated for the coral reef dataset. An application of this method to the sidescan-bathymetric classification problem is possible. Neighbourhood information could be used to add to a window's certainty of belonging to a certain class. This could decrease the variability of the final classsified results (figure 10.9). Ultimately, a Bayesian formulation of the problem is possible, where both Eucledian distance to class center and neighbourhood information can be used to assign a posterior probability of classifying a certain window to its corresponding class.

The algorithms explored in this thesis have many possible applications. Here are a few ideas:

- The detection of water flow direction in chosen areas. Oriented filters such as Gabor filters or complex wavelets can be used for that purpose. A set of consecutive sonar images could be considered to observe the changes in water flow direction during a given period of time. The effect of water flow is usually observable on seabed features such as sand waves. The orientations (and maybe 'best scale') of the filters can be used as parameters to be fitted according to the direction of the sand waves in the considered image. (An example is figure 11.1, of a 900kHz sidescan sonar image of sand ripples).

- The observation of biological species on the seafloor and their effect on the backscattered signals. The effect of environmental factors (such as pollution) on biological species could also be studied.

- As several sonar images were observed, objects in the water column backscattered sound before reaching the seabed. This backscattered information can be used to observe properties of the water column. A textural study of the water-column reflected signals can be used as a guide for studying biological distributions in certain areas. This can be applied to fish schools and algae growth.

- The coral reefs observed in this work were artificial ones. However, an interesting application lies in the observation of 'real' coral reefs. Their deterioration through time is a subject to be researched. This could be also an application of registration techniques (mainly used in bio-medical image studies) to locate the same coral reef across several surveys and observe how it has changed.

- Registration could also be used to observe features in the seabed covered in several survey lines. Wavelet image features could aid in recognising objects such as shipwrecks appearing in many surveys.

- The observation of sediment transport in certain regions could be an area of application. If sediment particles are suspended in the water column, a difference in the received signal (probably in the high frequency domains) could theoretically be observed. This is due to the backscattering and interference effects caused by these sediment particles. The observed signals could be used to detect whether sediments are present in the water column or not.

- Several sonars use two types of frequencies, which could offer a 'double' application of the methods suggested in this work. Low frequencies could be used to obtain a prior classification of areas. Then, the higher frequency would offer more resolution and would select more features at microbathymetric scales. This could lead to a wider range of features selected, and thus better classification rates.

Figure 11.1: Sand Ripple sidescan image obtained by a very high resolution sonar at 900 kHz, by Marine sonic (http:/www.marinesonic.com/archives.html). The sonar we used (234 kHz) was of a much lower resolution, and sand waves were not that evident.

Appendix A

The Wentworth Sedimentary Grain Scale

Millimeters	µm	Phi (ϕ)	Wentworth size class	
		-20		
4096		-12	Boulder (-8 to -12ϕ)	
1024		-10		
256		-8	Pebble (-6 to -8ϕ)	
64		-6		
16		-4	Pebble (-2 to -6ϕ)	
4		-2		Gravel
3.36		-1.75		
2.83		-1.50	Granule	
2.38		-1.25		
2.00		-1.00		
1.68		-0.75		
1.41		-0.50	Very coarse sand	
1.19		-0.25		
1.00		-0.00		
0.84		0.25		
0.71		0.50	Coarse sand	
0.59		0.75		
1/2 0.50	500	1.00		
0.42	420	1.25		
0.35	350	1.50	Medium sand	
0.30	300	1.75		
1/4 0.25	250	2.00		Sand
0.210	210	2.25		
0.177	177	2.50	Fine sand	
0.149	149	2.75		
1/8 0.125	125	3.00		
0.105	105	3.25		
0.088	88	3.50	Very fine sand	
0.074	74	3.75		
1/16 0.0625	63	4.00		
0.0530	53	4.25		
0.0440	44	4.50	Coarse silt	
0.0370	37	4.75		
1/32 0.0310	31	5	Medium silt	
1/64 0.0156	15.6	6	Fine silt	
1/128 0.0078	7.8	7	Very fine silt	
1/256 0.0039	3.9	8		Mud
0.0020	2.0	9		
0.00098	0.98	10		
0.00049	0.49	11	Clay	
0.00024	0.24	12		
0.00012	0.12	13		
0.00006	0.06	14		

Figure A.1: The Udden Wentworth grain size scale for sediments.

Appendix B

The Echo Sounder (Single Beam Sonar)

The echo sounder [1] has been used for a number of decades to measure bathymetry and record reflecting objects such as fish in the water column. The received acoustic amplitude can also be used for seafloor classification [45]. An echo sounder consists of a sonar transducer used to both transmit and receive an acoustic energy pulse directly beneath the sonar. The intensity of the reflected signal depends on the impedance ratio between water and the reflector and the angle that the reflector makes with the acoustic pulse. A hard seafloor would reflect more energy than a soft one, or one that is at an angle to the transmitted acoustic pulse. Although the echo sounder has been shown to produce high resolution depth data, it only produces information for targets directly beneath the sonar. Thus, it is necessary to extrapolate between survey line tracks to produce coverage maps of the seafloor [15].

The strength of acoustic energy returned from the seafloor with a single beam sonar has been used to classify the bottom type. A number of methods have been proposed for this and include those by Jackson and Briggs [110] who used the backscattered energy from the echo to infer bottom roughness. Orlowski [162] used a method that integrated parts of the multiple echo signature from the seafloor to provide information on its characteristics. Burns et al. [34]

[1] This type of sonar, although not used for classification in this work, is used for result comparison. This section is only a brief summary of its functions.

and Chivers et al.[45] developed a classification system based on the first echo and the second echo or first multiple echo from the seafloor. The ECHOplus sonar system is a commercial package of this classification system.

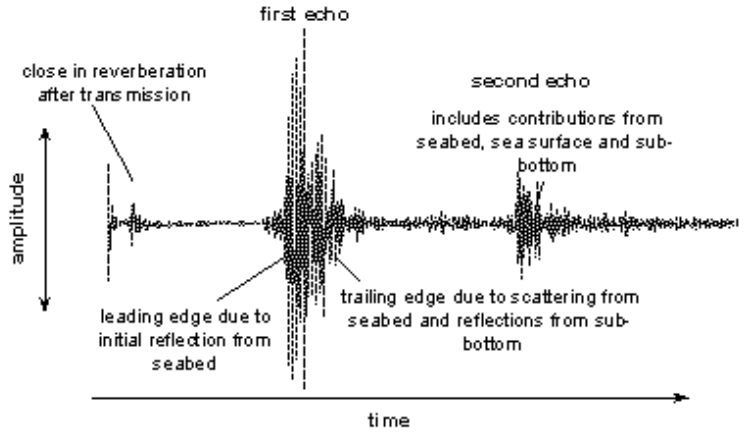


Figure B.1: Typical Echo Sounder output [15].

Figure B.1 shows the typical signal output of the Echo sounder. On the left is a signal of usually small amplitude resulting from close in structure and air bubbles beneath the transducer. In the middle is the first echo from the seabed, composed of: the initial reflection from the seabed beneath the transducer, backscatter from an area surrounding the point below the transducer and reflections from the sub-bottom. On the right can be seen the second echo from the seabed which has undergone an additional trip to and from the sea surface.

Information from the first echo is used to characterise seabed roughness by measuring the energy backscattered. To avoid contamination of the backscattered energy with energy that has been reflected from below the transducer, only the tail of the first echo is used in the analysis. The second echo is not just a delayed version of the first. Information from the second echo is used to characterise seabed hardness. There are at least two theories to explain the underlying physical mechanisms. Chivers *et al.* [45] assume that the ray paths for the second echo undergo two reflections at the seabed and a single scattering at the sea surface. The amount of reflection

is related to the difference in impedance between the sea water and the sea bottom (i.e. the hardness of the sea bottom). Heald and Pace [99] explain the configuration as a bistatic one. Because of the sea surface, the transmitter and receivers are virtually displaced by twice the water depth. The impact of this is that the receiving transducer is in the near field scattering zone of the seabed and the scattered energy is therefore driven by the reflection and hence hardness properties of the seabed. Both theories support the fact that the harder the seabed, the more energy appears in the second echo [2].

[2]The reader can refer to the following website for more information on the ECHOplus system : http:/www.sea.co.uk/marine_echoplus.html

Appendix C

SEA's Sonar Specifications

This section sumamrises some of the basic properties of the SEA sidescan bathymetric sonars used in this work. The sonars are of two frequencies 117 and 234 kHz.

The SEA Submetrix 2000 Series Sonars uses a pair of transducers, one facing port and the other to starboard. Each transmitter transmits a beam of azimuth width 1°, that can cover a large range. The amplitude is measured at one point of the transducer (thus acting as a sidescan sonar), whereas the phase is measured across many points and compared to obtain the depth data. As the ship moves, data is recorded in data files containing the following [196, 197]:

- the ping number

- the sample number (a ping is made of 1024 samples)

- which transducer is being used (port or starboard)

- the ship's position (in UTM coordinates, Eastings and Northings), its heading, roll, pitch and heave.

- the tide.

- the sample's received amplitude, its coordinates (Eastings and Northings) and its measured depth.

- a flag for each sample indicating its acceptability according to system filters.

Before starting a survey, the system has to be calibrated. Roll, pitch, heading, position and

Sonar Frequency	117 kHz	234 kHz
Maximum swath width	600 m	300 m
Maximum water depth	200 m	100 m
Range vs depth	15 times	15 times
Across track resolution	0.15 m	0.075 m
Accuracy versus depth	1% or 0.3 m	1% or 0.3 m
Transmit pulse length	68 μs to 1ms	34 μs to 500 μs
Ping repetition rate	max of 25 pings/sec	max of 25 pings/sec

Table C.1: The specifications of the SEA sidescan bathymetric sonar used in this work. Parameters in this table are given as a guide, and may vary according to survey conditions.

depth offsets are corrected. Factors such as the speed of sound profile, the position parameters, the sensor parameters and the tide are all recorded as survey parameters. Some filters are supplied with the sonar package for data selection. They include the following:

- a low amplitude filter to reject low amplitude depth data.

- a box filter that needs input parameters including the minimum and maximum depth and the maximum horizontal range.

- a range filter (the user enters the maximum slant range).

- a phase confidence filter that rejects non consistent phases.

- a phase-smoothing filter, that smoothes the high frequency 'wiggles' in the phase signal.

- an across track median filter.

- an along track filter that learns the shape of the gain across successive pings and removes irrelevant samples.

These filters are implemented in the software package (RTS2000) that is used for online visualisation of the sonar surveys.

Bibliography

[1] M. J. Chantler A. Penirschke and M. Petrou. Illuminant Rotation Invariant Classification of 3D Surface Textures using Lissajous's Ellipses. In *Texture2002, The 2nd International Workshop on Texture Analysis and Synthesis, Copenhagen*, pages 103–107, June 2002.

[2] K. Abend, T.L Harley, and L.N. Kanal. Classification of binary random patterns. *IEEE Tram. Inform. Theory*, 11:538–544, 1965.

[3] D. Alexandrou and D. Pantzartzis. A methodology for acoustic seafloor classification. *IEEE Journal of Oceanic Engineering*, 18(2):81–86, April 1993.

[4] L. Atallah and P.J. Smith. Using wavelet analysis to classify and segment sonar signals scattered from underwater sea-beds. *Special Issue on Underwater Acoustics of ACTA ACUSTICA united with ACUSTICA*, June 2002.

[5] L. Atallah and P.J. Smith. Using wavelet analysis to classify and segment sonar signals scattered from underwater sea-beds. In *Proceedings of ECUA 2002- The 6th European conference on Underwater Acoustics-Gdansk, Poland*, pages 59–64, June 2002.

[6] L. Atallah and P.J. Smith. Using wavelet analysis to classify and segment sonar signals scattered from underwater sea-beds. *International Journal of Remote Sensing*, 24(21):4113–4128, November 2003.

[7] L. Atallah and P.J. Smith. Automatic seabed classification by the analysis of sidescan sonar and bathymetric imagery. *IEE Proceedings - Radar, Sonar And Navigation -accepted for publication*, 2004.

[8] L. Atallah, P.J. Smith, and C.R. Bates. Wavelet Analysis for the classification of Hopvagen bay - Norway using a Bathymetric-Sidescan sonar. *Marine Geophysical Researches*, 23(5):431–442, 2002.

[9] J.M Augustin, R. Le Suave, X. Lurton, M. Voisset, S. Dugelay, and C. Satra. Contribution of the multibeam acoustic imagery to the exploration of the sea bottom. *Marine Geophysical Researches*, 18:459–485, 1996.

[10] Pascal Auscher, Guido Weiss, and Mladen Victor Wickerhauser. *Local Sine and Cosine Bases of Coifman and Meyer and the Construction of Smooth Wavelets.* Academic Press, Boston, 1991.

[11] V.M. Babovic and R. Gopakumar. Seabed recognition using neural networks. *D2K Technical report 0399-1, Danish Hydraulic Institute*, March 1999.

[12] H. Barad, A.B Martinez, B.S. Bourgeois, and E.J. Kaminsky. Acoustical boundary location through textural analysis of multibeam bathymetric sonar data. *Marine Technology Society Journal*, 27(1):24–30, 1993.

[13] C.R. Bates. A review-archaeological geophysical remote surveying for the seafloor. *Sedimentary Systems Research Group-University Of St Andrews.*

[14] C.R. Bates, G.C Moore, D.B. Harries, W. Austin, and J. Mair. Broad scale mapping of sublittoral habitats in Loch Laxford, Scotland. *Scottish Natural Heritage Commissioned Report, No. F01AA401A*, pages 1–65, 2002.

[15] C.R. Bates and E.J. Whitehead. Echo plus Measurements in Hopvagen bay, Norway. *The Oceanography Society, Biennial Scientific Meeting*, April 2001.

[16] R. A. Beck. Automatic seabed classification using sidescan sonar. *Proceedings of I.O.A*, 13(3):141–148, 1991.

[17] J. Bell. *A model for the simulation of sidescan sonar.* PHD Thesis, Heriot-Watt University, 1995.

[18] J.M. Bell, M.J. Chantler, and T. Wittig. Sidescan sonar:a directional filter of seabed texture. *IEE Proceedings-Radar, Sonar and Navigation*, 146(1):65–72, February 1999.

[19] R. Bellman. *Adaptive Control Processes: A guided tour.* New Jersey: Princeton University Press, 1961.

[20] U. Bhattacharya, B.B. Chaudhuri, and S.K. Parui. An MLP based texture segmentation method without selecting a feature set. *Image and vision computing*, 15:937–948, April 1997.

[21] J. Bigun and J.M.H. du Buf. N-folded symmetries by complex moments in gabor space and their application to unsupervised texture segmentation. *IEEE PAMI*, 16(1):80–87, 1994.

[22] Christopher M Bishop. *Neural Networks for pattern recognition.* Oxford University Press, Great Clarendon Street, Oxford, OX2 6DP, 1995.

[23] P. Blondel, J.C. Sempere, and V. Robigou. Textural Analysis and Structure Tracking for Geological Mapping: Applications to Sonar Images from Endeavor Segment, Juan De Fuca Ridge. In *Proceedings of Oceans '93 IEEE-OES*, pages 209–213, 1993.

[24] Ph. Blondel. Automatic mine detection by textural analysis of COTS sidescan sonar imagery. *Int. J. Remote Sensing*, 21(16):3115–3128, September 1999.

[25] H. Boehme and N.P. Chotiros. Acoustic backscattering at low grazing angles from the ocean bottom. *J. Acoust. Soc. Am*, 84(3), 1988.

[26] J. S. De Bonet and P. Viola. Texture recognition using a non parametric multiscale statistical model. In *Proc IEEE Computer Vision and pattern recognition conference*, 1998.

[27] C. Bouman and B. Liu. Multiple resolution segmentation of textured images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(2):99–113, 1991.

[28] A. C. Bovik, M. Clark, and W. S. Geisler. Multichannel texture analysis using localized spatial filters. *IEEE Trans. Pattern Anal. Machine Intell*, 12(1):55–73, January 1990.

[29] Prof. Michael Brady. *Transform methods for signal and image processing.* University of Oxford-Course notes, 2001.

[30] M. Bresch. Optimising filter banks for supervised texture recognition. *Pattern Recognition*, 35:783–790, 2002.

[31] Robert W. Buccigrossi and Eero P. Simoncelli. Image compression via joint statistical characterization in the wavelet domain. In *Proceedings ICASSP-97 (IEEE International Conference on Acoustics, Speech and Signal Processing)*, number 414, Munich, Germany, 1997.

[32] W.S. Burdic. *Underwater Acoustic System Analysis.* Prentice-Hall Signal Processing Series, 1990.

[33] Christopher J. C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2):121–167, 1998.

[34] D. Burns, R. C. Chivers, and H. Sisk. Analysis of Echosounder Signals for Sea Bed classification. In *Ultrasonics International - Institute of Acoustics*, pages 423–429, 1989.

[35] F. W. Campbell and J. J. Kulikowski. Orientation selectivity of the human visual system. *J. Physiol*, 187:437–445, January 1966.

[36] G. Canepa, E. Pouliquen, N.G. Pace, A. Figoli, and P. Franchi. Validation of a seafloor segmentation algorithm for multibeam data. In *Proceedings of ECUA 2002*, volume 1, pages 89–94, June 2002.

[37] D.R. Carmichael, L.M. Linnett, S.J. Clarke, and B.R. Calder. Seabed classification through multifractal analysis of sidescan sonar imagery. *IEEE Proc.-Radar.Sonar Navig*, 143(3):140–148, June 1996.

[38] P. Cervenka, C de Moustier, and P. F. Lonsdale. Geometric Corrections on Sidescan Sonar Images based on Bathymetry. Application with SeaMARCII and Sea Beam Data. *Marine Geophysical Researches*, 16:365–382, August 1994.

[39] T. Chang and C.C.J. Kuo. Texture analysis and classification with tree structured wavelet transform. *IEEE Trans. Image. proc*, 2(4):429–441, 1993.

[40] M.J. Chantler and G. McGunnigle. The response of texture features to illuminant rotation. In *ICPR 2000, 15th International Conference on Pattern Recognition, Barcelon*, pages 955–958, September 2000.

[41] R. Chellappa and S. Chatterjee. Classification of textures using gaussian markov random fields. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 33(4):959–963, August 1985.

[42] C.H. Chen, J.D. Lee, and M.C. Lin. Classification of underwater signals using wavelet transforms and neural networks. *Mathematical and computer modeling*, 27(2):47–60, January 1998.

[43] W.D. Chesterman, P.R. Clynick, and A.H. Stride. An acoustic aid to sea bed survey. *Acustica*, 8:285–290, 1958.

[44] W.D. Chesterman, J.M.P. St.Quinton, Y. Chan, and H.R. Mathews. Acoustic surveys of the seafloor near hong kong. *Int. Hydro. Rev*, 44(1):35–54, 1967.

[45] R.C. Chivers, N. Emerson, and D. Burns. New acoustic processing for underway surveying. *Hydrographic Journal*, 42:8–17, 1990.

[46] H. Choi and R. Baraniuk. Multiscale image segmentation using wavelet domain hidden markov models. *IEEE Trans. Image Proc*, 10(9):1309–1321, September 2001.

[47] C.K. Chui. *Wavelets: A Mathematical Tool For Signal Analysis*. SIAM, 1997.

[48] J.E. Hughes Clarke. The effect of fine scale seabed morphology and texture on the fidelity of swath bathymetric sounding data. In *Proceedings Canadian Hydrographic Conference 1998, Victoria.*, pages 168–181, 1998.

[49] D. Clausi and M Ed. Jernigan. A fast method to determine co-occurrence texture features. *IEEE trans. Geosc and Remote sensing*, 36(1):298–300, January 1998.

[50] D. Clausi and yongping Zhao. Rapid extraction of image texture by co-occurrence using a hybrid data structure. *Computers and Geosciences*, 28:763–774, September 2002.

[51] D. A. Clausi. *Texture Segmentation of SAR Sea Ice Imagery. Ph.D. Thesis, Department of Systems Design Engineering*. University of Waterloo, Waterloo, Ontario, Canada, 1996.

[52] G.R. Cochrane and K.D. Lafferty. Use of Acoustic classification of sidescan sonar data for mapping benthic habitat in the Northern Channel Islands, California . *Continental Shelf Research*, 22:683–690, 2002.

[53] F. S. Cohen and D. B. Cooper. Simple parallel hierarchical and relaxation algorithms for segmenting noncausal markovian fields. *IEEE Transactions on Pattern and Machine Intelligence*, 9:195–219, 1987.

[54] R.R. Coifman and D.L. Donoho. Translation invariant de-noising. *Yale University and Stanford University report*, 1995.

[55] R.R. Coifman and M.V. Wickerhauser. Entropy based methods for best basis selection. *IEEE Trans. Info. Theory*, 38:713–718, 1992.

[56] W.T. Collins and J.M.Preston. Multibeam seabed classification, quester tangent corp. *International Ocean Systems*, 6(4):12–15, July 2002.

[57] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995.

[58] G.R. Cross and A.K. Jain. Markov random field texture models. *IEEE Transactions on Pattern and Machine Intelligence*, 5:25–39, 1983.

[59] M. Crouse, R. Nowak, and R. Baraniuk. Wavelet-Based statistical signal processing using Hidden Markov Models. *IEEE Transactions on Signal Processing*, 1997.

[60] O.G. Cula and K.J. Dana. Compact representation of bidirectional texture functions. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR'01)*, pages 1041–1047, 2001.

[61] G. Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals and Systems*, 2(4):303–314, 1989.

[62] M. Dash and H. Liu. Feature selection for classification. *Intelligent Data Analysis*, 1:131–156, March 1997.

[63] I. Daubechies. Ten lectures on wavelets. In *CBMS-NSF Regional Conference Series in Applied Mathematics 61*, Philadelphia, 1992. SIAM.

[64] J.S. de Bonet and P. Viola. A non parametric multiscale statistical model for natural images. *Advances in Neural Information Processing*, 10, 1997.

[65] C. de Moustier. Oblique incidence classification methods. *OMG Coastal Multibeam training course notes, Ocean Mapping Group, Department of Godesy and Geomatics Engineering, University of New Brunswick, NB*, 2000.

[66] P. de Rivaz. *Complex wavelet baesd image synthesis and analysis.* Dphil thesis- University of Cambridge, 2000.

[67] P. de Rivaz and N. Kingsbury. Complex wavelet features for fast texture image retrieval. *Eng Department- Cambridge University*, 2001.

[68] C. deMoustier, P.F. Lonsdale, and A.N. Shor. Simultanuous operation of the seabeam multibeam echo sounder and the seamarcii bathymetric sidescan sonar system. *IEEE J. Ocean. Eng.*, 15(2):84–94, 1990.

[69] P.N. Denbigh. A bathymetric sidescan sonar. *Proc.Ultrasonics Int*, pages 321–326, 1979.

[70] P.N. Denbigh. Phase only sidescan sonar for underwater mapping. *Acoustic Letters*, 1:84–87, 1984.

[71] P.N. Denbigh. Swathe bathymetry: Principles of operation and analysis of errors. *IEEE Journal of Oceanic Engineering*, 14(4):289–298, October 1989.

[72] P.N. Denbigh. Signal processing strategies for a bathymetric sidescan sonar. *IEEE Journal of Oceanic Engineering*, 19(3):382–390, July 1994.

[73] D. Donoho. Non linear wavelet methods for recovery of signals, images, and densities from noisy and incomplete data. *Different Perspectives on Wavelets*, 1993.

[74] D.T. Donovan and A.H. Stride. An acoustic survey of the seafloor south of dorset and its geological interpretation. *Philos.Trans.R.Soc*, pages 299–330, 1961.

[75] J. M. H. du Buf and P. Heitkaemper. Texture features based on gabor phase. *Signal Processing*, 23:227–244, 1991.

[76] R. O. Duda and P. E. Hart. *Pattern Classification and Scene Analysis.* John Wiley and sons, Toronto, 1st edition, 1973.

[77] E. Dura, J. Bell, and D. Lane. Seafloor characterisation using contour maps recovered from side-scan sonar. In *CAD/CAC 2001 conference*, 2001.

[78] W. Dzwinel. How to make sammon's mapping useful for multidimensional data structure analysis. *Pattern Recognition*, 27(7):949–959, 1994.

[79] T.M. Edgecock. Towards gain invariant seabed classification. *Proceedings of I.O.A*, 16(6):217–224, 1994.

[80] Alexei A. Efros and Thomas K. Leung. Texture synthesis by non-parametric sampling. In *ICCV (2)*, pages 1033–1038, 1999.

[81] J. O. Eklundh. On the use of fourier phase features for texture discrimination. *Comput. Graphics. Image. Process*, 9:199–201, 1979.

[82] K. Englehart. *Signal Representation for Classification of the Transient Myoelectric Signal.* Ph.D. Thesis University Of New Brunswick, 1998.

[83] P. Etter. *Underwater Acoustic Modeling.* E and FN Spon, 1995.

[84] J.A. Farre and W.B.F. Ryan. A 3-D view of erosional scars on the U.S. mid-Atlantic continental margin. . *AAPG Bull*, 69:923–932, 1985.

[85] N. Fatemi-Ghomi, N. Palmer, and M.Petrou. Performance evaluation of texture segmentation algorithms based on wavelets. *Report-University of Surrey*, 1996.

[86] R.A. Fisher. The use of multiple measurements in taxonic problems. *Annals of Eugenics*, 7:179–188, 1936.

[87] C.G. Fox. Objective classification of oceanic ridge-crest terrains using two dimensional spectral models of bathymetry: Application to the Juan de Fuca Ridge. *Marine Geophysical Researches*, 18(6):707–728, 1996.

[88] C.G. Fox and D.E. Hayes. Quantitative methods for analysing the roughness of the sea floor. *Rev. Geophys*, 23:1–48, 1985.

[89] D. Gabor. Theory of communication. *J. Instr. Electr. Eng*, 93(3), 1946.

[90] S. Geman and D. Geman. Stochastic relaxation, gibbs distributions and the bayesian restoration of images. *IEEE Transactions on Pattern and Machine Intelligence*, 6:721–742, 1984.

[91] M. N. Gibbs and D. J. C. MacKay. Variational gaussian process classifiers. *IEEE-NN*, 11(6):1458, November 2000.

[92] M.F. Glenn. Introducing an operational multibeam-array sonar. *Int. Hyd. Review*, 47(1):35–39, 1970.

[93] I.T. Goodfellow. *Analysis Of Coregistered bathymetric and sidescan data.* PHD Thesis, University Of Bath, 1996.

[94] J.A. Grant. Filtering, digitisation, logging and presentation of depth. *Proc. 3rd Hydro. Symp. of Defence*, 1985.

[95] J.A. Grant and R. Schreiber. Modern swathe sounding and subbottom profiling technology for research applications:the atlas hyderosweep and parasound systems. *Marine Geophysical Researches*, 12:9–19, 1990.

[96] R.G. Haines. Sea room for the super tanker. *Navy*, 75(6):189–190, 1970.

[97] D. J. Hand. *Discrimination and Classification.* New York: John Wiley, 1981.

[98] R. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Trans. Systems, Man Cybernetics*, 3:610–621, November 1973.

[99] G.J. Heald and N.G. Pace. Implications of a bi-static treatment for the second echo from a normal incidence sonar. In *Proceedings 3rd European Conference on Underwater Acoustics, Bath*, pages 649–654, 1996.

[100] M.J.P. Heaton and W.G. Haslett. Interpretation of lloyd mirror in sidescan sonar. *Proc. Soc. Underwater. Tech*, 1(1):24–38, 1971.

[101] David J. Heeger and James R. Bergen. Pyramid-based texture analysis/synthesis. In *SIGGRAPH*, pages 229–238, 1995.

[102] M.B. Henke-Reed and S. N. C. Cheng. Cloth texture classification using the wavelet transform. *J. Imaging and Sci. Technology*, 37:610, 1993.

[103] U.C. Herzfeld and C.A. Higginson. Automated geostatical seafloor classification-principles, parameters, feature vectors and discrimination criteria. *Computers and Geosciences*, 22(1):35–52, 1996.

[104] T.J. Hickley. Narrow beam transducing sounding system. *Int. Hyd. Review*, 43(1):37–42, 1966.

[105] E.A. Howson and J.R. Dunn. Directional echo sounding. *J. Inst. of Navy*, 14(3):348–359, 1961.

[106] D. H. Hubel and T. N. Wiesel. Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *J. Neurophysiol*, 28:229–289, 1965.

[107] V.A.I. Hunvenne, Ph. Blondel, and J.-P. Henriet. Textural analyses of sidescan imagery from two mound provinces in the porcupine seabight. *Marine Geology*, 3190:1–19, 2002.

[108] A.N. Ivakin. Models of scattering for remote acoustic sensing of the seafloor. *IOA Conference on Acoustical Oceanography*, 23(2):268–275, April 2001.

[109] D.R Jackson. Models from scattering from the seabed. In *Proc. IOA*, volume 17, 1995.

[110] D.R. Jackson and K.B. Briggs. High frequency bottom backscattering versus sediment volume scattering. *J. Ac. Soc. Am*, 92(2):962–977, August 1992.

[111] D.R. Jackson, K.B. Briggs, K.L. Williams, and M.D. Richardson. Test of models for high-frequency seafloor backscatter. *IEEE Journal of Oceanic Engineering*, 21(4):458–469, October 1996.

[112] D.R. Jackson, D.P. Winebrenner, and A. Ishimaru. Application of the composite model to high-frequency bottom backscattering. *Journal of the Acoustical Society of America*, 79(5):1410–1421, May 1986.

[113] A. K. Jain and D. Zongker. Feature-selection: Evaluation, application, and small sample performance. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19(2):152–157, February 1997.

[114] A.K. Jain and F. Farrokhnia. Unsupervised texture segmentation using gabor filters. *Pattern recognition*, 24:1167–1186, 1991.

[115] I. T. Jollife. *Principal Component Analysis*. Springer-Verlag, New York, 1986.

[116] T. Kadir, D. Boukerroui, and M. Brady. A mathematical analysis of the scale saliency algorithm, 2003.

[117] T. Kadir and M. Brady. Scale saliency : A novel approach to salient feature and scale selection. In *ICCV*, pages 25–28, 2003.

[118] C. Kervrann and F. Heitz. A markov random field model-based approach to unsupervised texture segmentation using local and global spacial statistics. *IEEE Trans. Image Processing*, 4(6):856–862, June 1995.

[119] N. Kingsbury. The dual tree complex wavelet transform: a new efficient tool for image restoration and enhancement. In *Proc. European Signal Processing Conf*, September 1998.

[120] N. Kingsbury and J. Magarey. Wavelet transforms in image processing. *Signal Processing and Prediction I, EURASIP, ICT Press, Prague 1997*, pages 23–34, 1997.

[121] V.O. Knudsen, R.S. Alford, and J.W. Emling. Underwater ambient noise. *J. Mar. Res.*, 7:410, 1948.

[122] S. Konishi and A.L. Yuille. Statistical cues for domain specific image segmentation with performance analysis. In *CVPR'00*, pages 1125–1132, 2000.

[123] P. Kovesi. Image features from phase congruency. *Technical Report- University of Western Australia*, pages 1–30, 1995.

[124] P. Kovesi. *Invariant Measures of Image Features from Phase Information-Phd Thesis.* University of Western Australia, 1996.

[125] P. Kovesi. Symmetry an asymmetry from local phase. *Tenth Australian Joint Conference on Artificial Intelligence*, 1997.

[126] P. Kruizinga, N. Petkov, and S.E. Grigorescu. Comparison of texture filters based on gabor filters. *Proceedings of the 10th International Conference on Image Analysis and Processing*, pages 142–147, September 1999.

[127] A. Kundu, G.C. Chen, and C.E. Persons. Transient sonar signal classification using hidden markov models and neural nets. *IEEE Journal of Oceanic Engineering*, 19(1):87–99, January 1994.

[128] A. Kundu and J.L Chen. Texture classification using qmf bank-based subband decomposition. *GMIP*, 54:369–384, 1992.

[129] A. Laine and J. Fan. Texture classification by wavelet packet signatures. *IEEE Trans. Pattern analysis and machine learning*, 15(11):1549–1560, 1995.

[130] Thomas K. Leung and Jitendra Malik. Recognizing surfaces using three-dimensional textons. In *ICCV (2)*, pages 1010–1017, 1999.

[131] Thomas K. Leung and Jitendra Malik. Representing and recognizing the visual appearance of materials using three-dimensional textons. *International Journal of Computer Vision*, 43:29 – 44, 2001.

[132] L.N. Lieberman. Origin of sound absorption in water and sea water. *J. Acoust. Soc. Am*, 20, 1948.

[133] Y. Lin. Support vector machines and the bayes rule in classification, 1999.

[134] J. Lu and L. Carin. HMM Based multiresolution image segmentation. In *Proc. ICASSP 2002*, volume 4, pages 3357–3360, May 2002.

[135] A.Penirschke M. J. Chantler, G.McGunnigle and M. Petrou. "Estimating Lighting Direction and Classifying Textures. In *BMVC 2002, Cardiff*, September 2002.

[136] W.Y. Ma and B. S. Munjanath. Texture features and learning similarity. In *Proc IEEE Computer Vision and pattern recognition conference*, pages 256–259, June 1996.

[137] E. MacCurdy. *The Notebooks Of Leonardo Da Vinci*. Garden City Publishing, 1942.

[138] D. MacKay. *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press, 2003.

[139] J. Magarey and N. Kingsbury. Motion estimation using complex wavelets. *Technical Report TR-226, Cambridge University Engineering Department, 1995*, 1995.

[140] N. C. Makris. Estimating surface orientation from sonar images. In *High Frequency Acoustics in Shallow Water, edited by N.G. Pace, Kluwer, Dordrecht*, volume 94, 1997.

[141] S. Mallat. Applied mathematics meets signal processing. *Documenta Mathematica. Extra Volume:International Congress of Mathematics, Berlin.*, pages 1–20, 1998.

[142] S. Mallat. *A wavelet tour of signal processing.* Academic Press, 24-28 Oval Road, London NW1 7DX,UK, 1999.

[143] S. Mallat and Z. Zhang. Matching pursuit with time-frequency dictionaries. *IEEE Transactions on Signal Processing*, pages 837–842, December 1993.

[144] S.G. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE PAMI*, 11:674–693, 1989.

[145] B. S. Manjunath and R. Chellappa. Unsupervised texture segmentation using markov random field models. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13(5):478–482, May 1991.

[146] H. Medwin and C. S. Clay. *Fundamentals Of Acoustical Oceanography.* Academic Press, 1998.

[147] Z.H. Michalopoulou and D. Alexandrou. Bayesian modeling of acoustic signals for seafloor identification. *Journal of The Acoustical Society of America*, 99(1):223–233, January 1996.

[148] Z.H. Michalopoulou, D. Alexandrou, and C. De Moustier. Application of neural and statistical classifiers to the problem of seafloor characterisation. *IEEE Journal of Oceanic Engineering*, 20(3):190–197, July 1995.

[149] M. Mignotte, C. Collet, P. Perez, and P. Bouthemy. Hybrid genetic optimisation and statistical model-based approach for the classification of shadow shapes in sonar imagery. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(2):129–141, feb 2000.

[150] N. Mitchell and J. E. Hughes Clarke. Classification of seafloor geology using multibeam sonar data from the Scotian shelf. *Marine Geology*, 121:143–160, September 1994.

[151] J. Moody and C.J. Darken. Fast learning in networks of locally-tuned processing units. *Neural Computation*, 1(2):281–294, 1989.

[152] P.D. Mourad and D.R. Jackson. High frequency sonar equation models for bottom backscatter and surface loss. In *Proceedings of Oceans 89*, volume 4, pages 1168–1175, September 1989.

[153] B. S. Munjanath and W. Y. Ma. Texture features for browsing and retrieval of image data. *IEEE Trans. Patt. Anal. Mach. Int. Special Issue on Digital Libraries*, 18(8):837–842, 1996.

[154] I. Nabney. *Netlab: Algorithms for Pattern Recognition.* Springer, 2001.

[155] N.C.Mitchell. Quantitative backscatter measurements with a long-range side-scan sonar. *IEEE J. Oceanic Engineering*, 14:368–374, 1989.

[156] R.M. Neal. Monte carlo implementation of gaussian process models for bayesian regression and classification. *Technical Report No. 9702, Dept. of Statistics, University of Toronto*, pages 1–24, 1997.

[157] H. Noda, M. Shirazi, and E. Kawaguchi. MRF- based texture segmentation using wavelet decomposed images. *Pattern Recognition*, 35:771–782, February 2002.

[158] J.A. Nystuen and M.J. McPhaden. The beginnings of operational marine weather observations using underwater ambient sound. *IOA Conference on Acoustical Oceanography*, 23(2), April 2001.

[159] J.A. Ogilvy. *Wave scattering from random rough surfaces.* Adam Hilger, 1991.

[160] A. V. Oppenheim, M. H. Hayes, and J.S. Lim. Iterative procedures for signal reconstruction from fourier transform pairs. *Opt. eng*, 21:122–127, 1982.

[161] A. V. Oppenheim, J.S. Lim, and S.R. Curtis. Signal synthesis and reconstruction from partial fourier-domain information. *J. Opt. Soc. America*, 73(11):1413–1420, 1983.

[162] A.O. Orlowski. Application of multiple echoes energy measurement for evaluation of seabottom type. *Oceanologica*, 19:61–78, 1984.

[163] M.J.L. Orr. Introduction to radial basis function networks. *Centre for Cognitive Science-University Of Edinburgh*, pages 1–67, April 1996.

[164] M.J.L. Orr. Matlab functions for radial basis function networks. *Institute Of Adaptive and Neural Computation, Division of Informatics, Edinburgh University*, June 1999.

[165] M.J.L. Orr. Recent advances in radial basis function networks. *Institute Of Adaptive and Neural Computation, Division of Informatics, Edinburgh University*, pages 1–23, June 1999.

[166] C. Spence P. Sajda and L. Parra. A multi-scale probabilistic network model for detection, synthesis and compression in mammographic image analysis. *Medical Image Analysis*, 7(2):184–204, 2003.

[167] N.G. Pace and H. Gao. Swathe seabed classification. *IEEE Journal of Oceanic Engineering*, 13(2):83–90, April 1988.

[168] R. Paget. Nonparametric markov random field models for natural texture images, 1999.

[169] R. Paget and D. Longstaff. Texture synthesis via a noncausal nonparametric multiscale markov random field, 1997.

[170] Rupert Paget and I. D. Longstaff. Texture synthesis and unsupervised recognition with nonparametric multiscale markov random field models.

[171] E. Pouliquen, O. Bergem, and N.G. Pace. Time-evolution modeling of the seafloor scatter 1. concept. *Journal of The Acoustical Society of America*, 105(6):3163–3141, June 1999.

[172] E. Pouliquen, M. Trevorrow, Ph. Blondel, G. Canepa, F. Cernich, and R. Hollet. Multi Sensor Analysis of the seabed in shallow water areas: overview of the MAPLE'2001 experiment. In *Proceedings of ECUA 2002*, volume 1, pages 21–29, June 2002.

[173] P. Pudil, J. Novovicov, and J. Kittler. Floating search methods in feature selection. *Pattern Recognition Letters*, 15:1119–1125, November 1994.

[174] G.D. Quartly, J.W. Gregory, T.H. Guymer, K.G. Birch, G.W. Jones, and S.J. Keogh. How reliable are acoustic rain sensors? *IOA Conference on Acoustical Oceanography*, 23(2):142–148, 2001.

[175] T Randen and J. H. Husoy. Filtering for texture classification: A comparative study. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 21(4):291–310, April 1999.

[176] A. R. Rao and G. L. Lohse. Identifying high level features of texture perception. *CVGIP: Graphical Models and Image Processing*, 55(3):218–233, May 1993.

[177] Lord Rayleigh. *Theory Of Sound Vols 1 and 2.* Dover Publications, 1945.

[178] T.B. Reed and D. Hussong. Digital image processing techniques for enhancement and classification of seamarc ii sidescan sonar imagery. *Journal of Geophysical Research*, 94:7469–7490, 1989.

[179] Z. Reut, N.G. Pace, and M.J.P. Heaton. Computer classification of seabed by sonar. *Nature*, 314:426–428, 1985.

[180] G.S. Ritchie. British hydrography since cook. *The Naval Review*, 58(1):126–134, 1970.

[181] J. K. Romberg, H. Choi, and R. G. Baraniuk. Bayesian wavelet domain image modelling using Hidden Markov trees . In *1999 International Conference on Image Processing (ICIP '99) , Kobe, Japan*, pages 1–5, October 1999.

[182] N. Saito. *Local Feature Extraction and Its Applications Using a Library of Bases.* D. Phil. Thesis - Yale University, 1994.

[183] N. Saito and R.R. Coifman. On local feature extraction for signal classification. *Applied Analysis-Special issue of Zeitschrift fur Angewandte Mathematik und Mechanik*, pages 453–456, 1996.

[184] J.W. Sammon. A nonlinear mapping for data structure analysis. *IEEE Trans. on Computers*, C-18(5):401–409, May 1969.

[185] R. Schalkoff. *Pattern Recognition: Statistical, Structural and Neural Approaches.* John Wiley and sons, Toronto, 1st edition, 1992.

[186] P. Scheunders, S. Livens, G. Van de Wouver, P. Vautrot, and D. Van Dyck. Wavelet-based texture analysis. *International J. of Comp. Sc and Inf. Management*, dec 1997.

[187] C.E. Shannon. The mathematical theory of communications. *Bell System Tech. J.*, 27, July 1948.

[188] W. Siedlecki, K. Siedlecka, and K. Sklanski. An overview of mapping techniques for exploratory pattern analysis. *Pattern Recognition*, 21(2):411–429, 1988.

[189] N. Smith, M. Gales, and M. Niranjan. Datadependent kernels in svm classification of speech patterns, 2001.

[190] M.L. Somers and A.R. Stubbs. Sidescan sonar. *IEEE Proceedings*, 131 F(3):243–256, June 1984.

[191] P. Somol, P. Pudil, J. Novovicov, and P. PaclGk. Adaptive floating search methods in feature selection. *Pattern Recognition Letters*, 20:1157–1163, 1999.

[192] I. Stainvas and D. Lowe. A generative probabilistic oriented wavelet model for texture segmentation. *Neural Computing Research group, Aston University - Report*, pages 1–16, October 2001.

[193] S. Stanic, R.R. Goodman, K. Briggs, N. Chotiros, and E. Kennedy. Shallow-water bottom reverberation measurements. *IEEE Journal of Oceanic Engineering*, 23(3):203–209, July 1998.

[194] W. Kenneth Stewart, M. Jiang, and M. Marra. A neural network approach to classification of sidescan sonar imagery from a midocean ridge area. *IEEE Journal of Oceanic Engineering*, 19(2):214–224, April 1994.

[195] A.H. Stride. A linear pattern on the seafloor and its interpretation. *J. Mar. Biol. Assoc*, 38:313–318, 1959.

[196] SEA (Submetrix). *2000 SERIES online user's guide*. SEA(Group). Ltd, 2000.

[197] SEA (Submetrix). *Grid2000 manual*. SEA(Group). Ltd, 2000.

[198] H. H. Szu. Mathematics of adaptive wavelet transforms: Relating continuous with discrete transforms. *Optical Engineering*, 33(7):2111–2124, 1994.

[199] Y. Tadmor and D.J. Tolhurst. Both the phase and the amplitude spectrum may determine the appearance of natural images. *Vision Research*, 33(1):141–145, 1993.

[200] D. Tamsett. Sea-bed characterisation and classification from the power spectra of sidescan sonar data. *Marine Geophysical Researches*, 15:43–64, April 1993.

[201] X. Tang and W. Kenneth Stewart. Optical and sonar image classification: Wavelet packet transform vs fourrier transform. *Computer Vision and Image Understanding*, 79:25–46, February 2000.

[202] B.A. Telfer, H.H. Szu, G.J. Dobeck, J.P. Garcia, H. Ko, A. Dubey, and N. Witherspoon. Adaptive wavelet classification of acoustic backscatter and imagery. *Optical Engineering*, 33(7):2192–2203, July 1994.

[203] E.I. Thorsos. The validity of the kirchoff approximation for rough surface scattering using a gaussian roughness spectrum. *J. Acoust. Soc. Am.*, 83:78–92, 1988.

[204] D.G. Tucker. Directional echo-sounding. *Int. Hyd. Rev*, 37(2):43–53, 1960.

[205] R.J. Urick. Low frequency attenuation in deep ocean. *J. Acoust. Soc. Am*, 35:1413, 1948.

[206] R.J. Urick. Some directional properties of deep water ambient noise. *Naval Research Labratory report 3796*, 1951.

[207] R.J. Urick. *Principles Of Underwater Sound For Engineers.* McGraw Hill, 1967.

[208] M. Varma and A. Zisserman. Texture classification: Are filter banks necessary? *ICCV 2003*, pages 1–8, 1983.

[209] M. Varma and A. Zisserman. Classifying images of materials: Achieving viewpoint and illumination independence. In *ECCV (3)*, pages 255–271, 2002.

[210] Manik Varma and Andrew Zisserman. Statistical approaches to material classification.

[211] A.G. Voronovich. *Wave scattering from rough surfaces.* Springer-Verlang, 1999.

[212] Y. Wang, H. Li, J. Xuan, S. C. B. Lo, and S. K. Mun. Modeling of wavelet coefficients in medical image compression. In *1997 International Conference on Image Processing (ICIP '97) 3-Volume Set-Volume 1*, 1997.

[213] Thomas P. Weldon and William E. Higgins. Multiscale rician approach to gabor filter design for texture segmentation. In *ICIP (2)*, pages 620–624, 1994.

[214] T.P. Weldon and W.E. Higgins. Design of multiple gabor filtres for texture segmentation. *Proc. ICASSP*, May 1996.

[215] T.P. Weldon, W.E. Higgins, and D.F. Dunn. Efficient gabor filter design for texture segmentation. *Pattern recognition*, 1996.

[216] C.R. Wentworth. A scale of grade and class terms for classic sediments. *Journal of Geology*, 30:377–392, 1922.

[217] M.V. Wickerhauser. *Acoustic signal Compression with Wavelet Packets in Wavelets: a tutorial in theory and applications. C.K. Chui, Ed.* Academic Press, San Diego, 1992.

[218] M.V. Wickerhauser. *Adapted Wavelet Analysis from Theory to Software.* A. K. Peters, Wellesley, MA, 1994.

[219] M.V. Wickerhauser. Best adapted wavelet packet bases. *Proceedings of the Symposia in applied mathematics*, 1996.

[220] C.K.I Williams and D. Barber. Bayesian classification with gaussian processes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(12):1342 – 1351, 1998.

[221] P. S. Williams. *The automatic hierarchical decomposition of images into sub-images for use in image recognition and classification. Phd thesis.* University of Western Australia, 1999.

[222] J. H. Won, K. Pyun, and R. M. Gray. Hidden Markov Multiresolution Texture Segmentation using Complex Wavelets. In *Proceedings of the International Conference on Telecommunication, ICT-2003, Tahiti*, February 2003.

[223] P.M. Woodward. *Probability and information theory with application to radar*. Elmsford, N.Y. :Pergamon Press Inc., 1955.

[224] Alexey Zalesny and Luc Van Gool. A compact model for viewpoint dependent texture synthesis. *Lecture Notes in Computer Science*, 2018:124–143, 2001.

[225] S. Zhu, Y. Wu, and D. Mumford. Filters random fields and maximum entropy (frame): To a unified theory for texture modeling.