Registration and Variability of Side Scan Sonar Imagery

by

John William Nicholson

B.S., United States Naval Academy (1981)

Submitted in partial fulfillment of the
requirements for the degree of

OCEAN ENGINEER

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

and the

WOODS HOLE OCEANOGRAPHIC INSTITUTION

August 1988
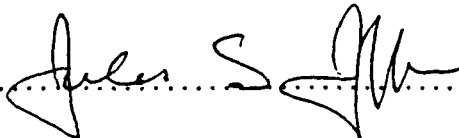
N00028-85-C-3262

© John W. Nicholson, 1988

The author hereby grants to MIT, WHOI, and the United States
Government permission to reproduce and to distribute
copies of this thesis document in whole or in part.

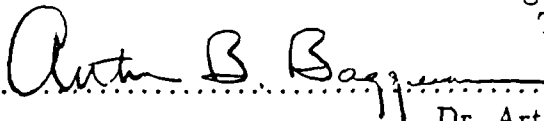Signature of Author ....
Joint Program in Oceanographic Engineering
Massachusetts Institute of Technology
Woods Hole Oceanographic Institution
August 5, 1988

Certified by ....
Dr. Jules S. Jaffe
Woods Hole Oceanographic Institution
Thesis Supervisor

Certified by ....
Dr. Arthur B. Baggeroer
Massachusetts Institute of Technology

Accepted by ....
W. Kendall Melville
Chairman, Joint Committee for Oceanographic Engineering, Massachusetts
Institute of Technology/Woods Hole Oceanographic Institution

AD-A201 895

88 11 21 093

Registration and Variability of Side Scan Sonar Imagery

by

John William Nicholson

Submitted to the Massachusetts Institute of Technology–
Woods Hole Oceanographic Institution
Joint Program in Oceanographic Engineering
on August 5, 1988, in partial fulfillment of the
requirements for the degree of
OCEAN ENGINEER

Abstract

This thesis presents the results of several experiments performed on side scan sonar equipment and imagery with the aim of characterizing the acoustic variability of side scan sonar imagery and applying this information to image rectification and registration. A static test tank experiment is presented which analyzes the waveform, power spectral density, and temporal variability of the transmitted waveform. The results of a second static experiment conducted from the Woods Hole Oceanographic Institution Pier in Woods Hole, Massachusetts permit determination of the distribution and moments of intensity fluctuations of echoes from objects imaged in side scan sonograms. This experiment also characterizes temporal and spatial coherence of intensity fluctuations. A third experiment is presented in which a side scan sonar towfish images the bottom adjacent to the pier while running along an underwater track which reduces towfish instability. Imagery from this experiment is used to develop a rectification and registration algorithm for side scan sonar images. Preliminary image processing is described and examples presented, followed by favorable results for automated image rectification and registration.

Thesis Supervisor: Dr. Jules S. Jaffe
Woods Hole Oceanographic Institution

Acknowledgements

Contents

List of Figures

6

List of Tables

Chapter 1

Introduction

1.1 Background

Man's curiosity about the seafloor and objects that may be found there has persisted for centuries, motivating the development of numerous devices for its inspection. With the increasing use of the ocean for a variety of purposes this traditional curiosity has been largely supplanted by a practical need for information about the nature of the ocean's bottom.

The most common need for such information is the safe navigation of seagoing vessels, where knowledge of seafloor morphology and the location and characteristics of isolated hazards to navigation is essential. Such hazards include stationary features such as rocky outcrops, but may also consist of time-varying features such as current driven shoaling. Man-made objects may also constitute hazards to navigation, and may also be considered time-varying due to their deposition over time [26] . Objects in this category include shipwrecks, debris dumped at sea, and undersea activities ranging from economic pursuits to the laying of mines. In the past the most common means of sensing the undersea environment has been point measurements of water depth using echo soundings or diver inspection, but the advantages of improved resolution and superior mapping rates inherent in underwater imagery is motivating the development of several imaging systems for gathering bathymetric data [7]

.

Economic pursuits such as offshore petroleum production provide further motivation for imaging underwater scenes [13] . These physical systems are subject to damage and deterio-

ration over time and require periodic inspection to ensure continuous operation. Monitoring can be performed with distributed local sensors, but such information is normally limited to physical parameters such as temperature or pressure. Imagery provides high resolution information about the condition of the system and may reveal defects which manifest themselves as changes in imagery over time.

A further application of underwater imagery is the location of objects lost at sea and resting on the seafloor. Shipwrecks, downed aircraft, and inadvertently dropped equipment frequently become the target of searches aimed at their location and salvage [18] . Searching for these objects may consist of imaging large areas of the seafloor in the area assumed to contain the object and identifying it. If the search area was imaged prior to the loss, the object's location may be revealed as an image region which exhibits change between the two images.

The aspect common to all these situations is that they require the ability to image underwater scenes and detect changes in successive images. Under normal circumstances the imaging methods of choice are photography or video. However because light has a limited range in water, efforts to develop new means of underwater imagery have focused on high resolution, high frequency acoustic systems. Perhaps the most common method of obtaining high resolution underwater images is side scan sonar, a technique commercially available since 1958 [11] . It is presently used by the petroleum industry, the military, hydrographic surveyors, and salvage operators for imaging the types of underwater scenes described above. Side-scan sonar is a remote sensing tool which generates a pictoral representation of the bottom similar to an aerial photograph. Using a narrow, high-frequency acoustic transmission the seafloor is sampled with a sample area whose size is controlled by acoustic beamwidth and pulse duration. The small size of the sample area results in a high resolution mapping of the bottom.

Comparison of side scan sonar images taken of the same bottom region at different times is occasionally performed, but for the most part this is done by visual comparison of several images. Because of the large amount of detail and information in a single side scan sonar image and the limited speed of human observers in inspecting and comparing thousands of subregions of an image it is desirable to devise an automated means of image inspection and

11

comparison. Devising such a system will require consideration of two major characteristics of side-scan sonar imagery. Because side scan images are scanned from a potentially unstable sensor this type of imagery does not represent the projection of a three dimensional scene to a two dimensional image as does a photograph or video image. As a result it generally is necessary to rectify or remap side scan image data to a reference coordinate system in order to compare multiple images of the same scene on a point by point basis [2] . Also, if side scan sonar images are to be compared on a point-by-point or feature-by-feature basis the image regions corresponding to these features must exhibit consistent intensity. If intensity fluctuations are too great it might not be possible to match corresponding features. It is therefore necessary to ascertain the nature of intensity fluctuations in side scan sonar images.

The degree to which image processing techniques have been applied to side scan sonar imagery is presently small, making the field a fertile one for research. Although standard image processing techniques may be applied to side scan sonar imagery, its peculiarities dictate that specialized processing methods such as the ones developed in this thesis also be employed.

1.2   Thesis Overview

This thesis describes the side scan sonar imaging process and techniques for automated assistance in detecting change in successive images of the same bottom as well as other related topics. Changes observed in multiple side scan sonar images of the same scene may be attributed to either changes in the bottom or to stochastic aspects of the medium and imaging system. Two experiments conducted to study this variability are described and results are presented which quantify the inherent variability of the imaging process. Measurements allow separation of system variability from medium variability .

Studies of the inherent variability of the medium, its causes, and its effect on the transmission of sound have been conducted since the World War II era work of Sheehy on the temporal variability of the amplitudes of acoustic transmissions over a fixed path [27] . Present studies including this one investigate the temporal coherence of acoustic transmissions with the desired result being the development of improved methods of underwater communications and imaging systems. The study undertaken in this thesis focuses on variability in the monostatic

12

case of the round-trip propagation of underwater sound reflected at shallow grazing angles from the bottom. This study involves shorter ranges and higher frequencies than usually encountered in the literature. To our knowledge this is the first study of its kind.

The removal of two dimensional geometric image distortions or warps which may be introduced by imaging geometry or sensor motion is known as rectification. The desired result of rectification is an image free of warps with all image points remapped to a reference image or coordinate system. Rectification of aerial and satellite images has been developed and is common [12] , but has been attempted only on a limited basis with side scan sonar imagery. Side scan sonar manufacturers including Klein Associates have developed systems which remap side scan sonar imagery to correct for towfish altitude, slant range distortion, and towfish speed variations as well as record navigation data for use in subsequent image mosaicking [19] . The Klein K-MAPS system uses sensors to determine towfish altitude and speed, and combines this data with survey ship navigation data to assemble a composite image on a geographic coordinate system. Another example of remapping side scan sonar imagery is the system developed jointly by the British Institute of Ocean Sciences (IOS) and NASA/JPL to process imagery obtained by the Geologic Long Range Inclined Asdic (GLORIA) system [24] . This system removes image distortions attributable to non-uniform survey ship course and speed by performing a computer assisted mapping of image pixels to geographic coordinates using navigation and ship's heading data recorded during the survey. In both systems the precision of data remapping is determined by navigation system accuracy.

This thesis develops a remapping algorithm using a previous image of the area as a reference to which the subsequent image is remapped. The advantage of this approach is that it eliminates the effect of navigation uncertainty on the image rectification and registration process.

Chapter 2

The Side Scan Sonar Imaging Process

2.1 Insonification

An idealized side scan sonar system is shown in fig. (2.1). The sensor, or *towfish* , is towed by the survey vessel along an underwater path. The corresponding path over the bottom will be referred to as its *bottom track* . The towfish is a slender body with tailfins for stability and an acoustic transducer on either side. The transducers are dimensioned to radiate a beam pattern which is wide in the vertical direction and narrow horizontally. The intersection of the beam with the bottom is a narrow *strip.* The narrowness of the bottom strip results in high resolution in the *axial* direction or the direction aligned with the towfish axis. The length of the pulse is short so that the length of the strip insonified at any given time is also short. As a result the *lateral* resolution or resolution in the direction perpendicular to the towfish axis is also high. The echo returns from each of these *sample areas* are received sequentially by the towfish and assembled into an image row. Between transmissions the towfish moves down the track so that on the next cycle a strip adjacent to the previous one is insonify. The bottom is therefore raster scanned to create the image. To ideally insonify the

14
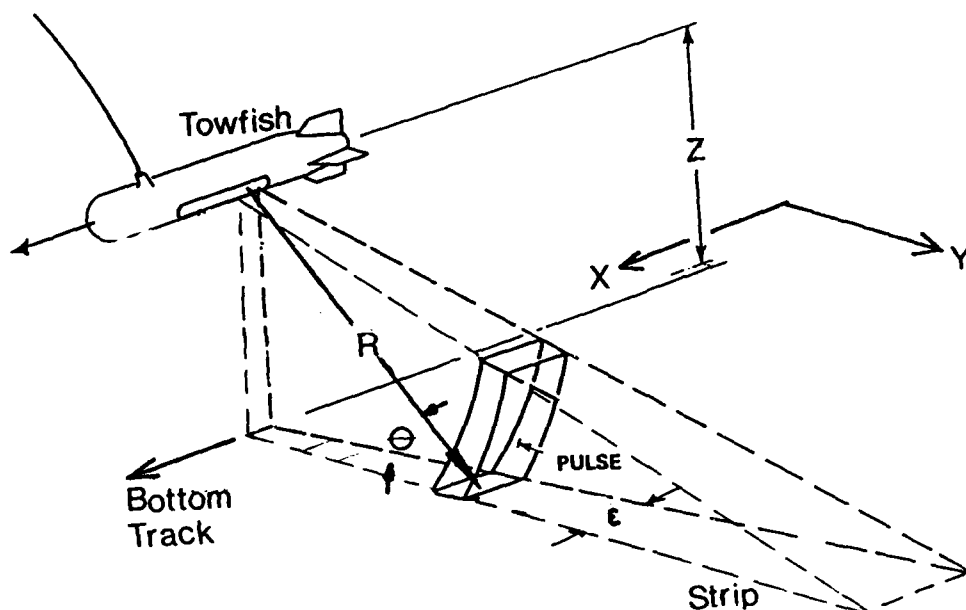
Figure 2-1: Ideal side scan sonar imaging geometry

strip the towfish would transmit a pulse insonifying three dimensional angular space defined by *grazing angle* θ and *azimuthal angle* ϕ with the following desired spatial acoustic intensity distribution:

$$p(\theta, \phi) = \begin{cases} p_o \frac{R_o}{R} & 0 \leq \theta \leq \frac{\pi}{2}, -\frac{\epsilon}{2} \leq \phi \leq \frac{\epsilon}{2} \\ 0 & \text{elsewhere} \end{cases} \quad (2.1)$$

where p is the acoustic pressure amplitude, p_o is the pressure at reference distance R_o, R is the range from the sonar transducer or *slant range* as it is commonly referred to in side scan sonar work, and ϵ is the horizontal beamwidth. The range of θ ensures complete insonification of the seafloor while preventing insonification of the air-sea interface which would contaminate the acoustic image with extraneous acoustic scattering. The variable ϕ is limited to small ϵ to achieve high axial resolution. A cartesian coordinate system is defined on the bottom where X is the axial topographic coordinate, Y is the lateral topographic coordinate, X' is the speed of the towfish along the bottom track, and Z is towfish altitude. The axial resolution

15

or axial dimension of the sample area is

$$\Delta x = \epsilon R \qquad (2.2)$$

The image representing this portion of the bottom is designated the *image matrix* . It is made up of *pixels* which are indexed by coordinates x and r where r corresponds to R. x is the row number of the image matrix and normally corresponds to X but generally not with the same fidelity that r corresponds to R. Using the Fourier transform relationship between array configuration and radiation pattern, it is impossible to generate the ideal far field beam pattern with a finite array, however a good approximation can be obtained using a rectangular array which has a large acoustic aperture in the horizontal dimension and a much smaller aperture in the vertical.The pressure field of such an array at sufficient range is described by

$$p = \frac{P_0}{R} \frac{sin\left(\frac{K_v L_v}{2}\right)}{\left(\frac{K_v L_v}{2}\right)} \frac{sin\left(\frac{K_h L_h}{2}\right)}{\left(\frac{K_h L_h}{2}\right)} \qquad (2.3)$$

where

$$K_v = ksin\alpha = \frac{2\pi}{\lambda}sin\alpha$$
$$K_h = ksin\beta = \frac{2\pi}{\lambda}sin\beta \qquad (2.4)$$

and λ is the acoustic wavelength; L_v and L_h are the vertical and horizontal dimensions of the array, respectively; and α and β are the *array angles* measured from the normal of the rectangular array in the vertical and horizontal directions, respectively. Array angles α and β are related to θ and ϕ by the relative orientation of the towfish array and the bottom.

The beam pattern is range dependent and experiences transition of the radiation field from the near (Fresnel) field to the far (Fraunhofer) field, with constant beam divergence at a constant angle valid only in the far field. The radiation field of the rectangular array is the product of vertical and horizontal components, as shown in eqn. (2.3). This separation allows treatment of transition ranges of the vertical and horizontal beam patterns separately. The transition from near field to far field regimes is not a clearly defined one. The near field of a line array is characterized by cylindrical rather than spherical spreading and by numerous regions of constructive and destructive interference. The transition to the far regime is an

asymptotic approach to spherical spreading and the regular beam pattern described by eqn. (2.3). An approximate definition of the transition range of a line array is [10]

$$r_t \simeq \frac{L^2}{\lambda} \qquad (2.5)$$

where r_t is the transition range and L is the length of the array. Values for transition range are given in table (2.1) for two different side scan sonar systems, the Klein 100 KHz and 500 KHz models.

	100 KHz	500 KHz
Array Dimensions (mm)		
Vertical	25	25
Horizontal	448	448
Wavelength (mm)	15	3
Transition Range (M)		
Vertical	0.042	0.21
Horizontal	13.4	66.9

Table 2.1: Transition ranges for Klein 100 KHz and 500 KHz towfish.

The transition of the vertical beam pattern occurs sufficiently near the transducer so that the radiation field may be considered to be in the far regime throughout the imaging volume. The horizontal beam pattern transition occurs at a significant distance considering that imaging of the seafloor typically begins in the first 10 to 20 meters. The horizontal beam pattern of the 500 KHz system is modeled in fig. (2.2). The horizontal axes are cartesian coordinates in the radiation field and the vertical axis is the pressure amplitude at that point. The near and far fields are evident at range extremes, while the transition between the two occurs in the middle ranges. At approximately 30 meters the beginning of what eventually becomes the main lobe of the far field is seen to emerge from the complicated near field. The width of the main lobe determines the axial resolution of the system, and this is plotted in fig. (2.3). Both curves represent the -3 dB or half power contours, with the wider curve showing the beamwidth of the transmitted pulse and the narrow curve showing the -3dB contour for the combined transmission and reception of an acoustic pulse assuming no array motion during the transmission cycle. This contour takes into account the fact that the same

17
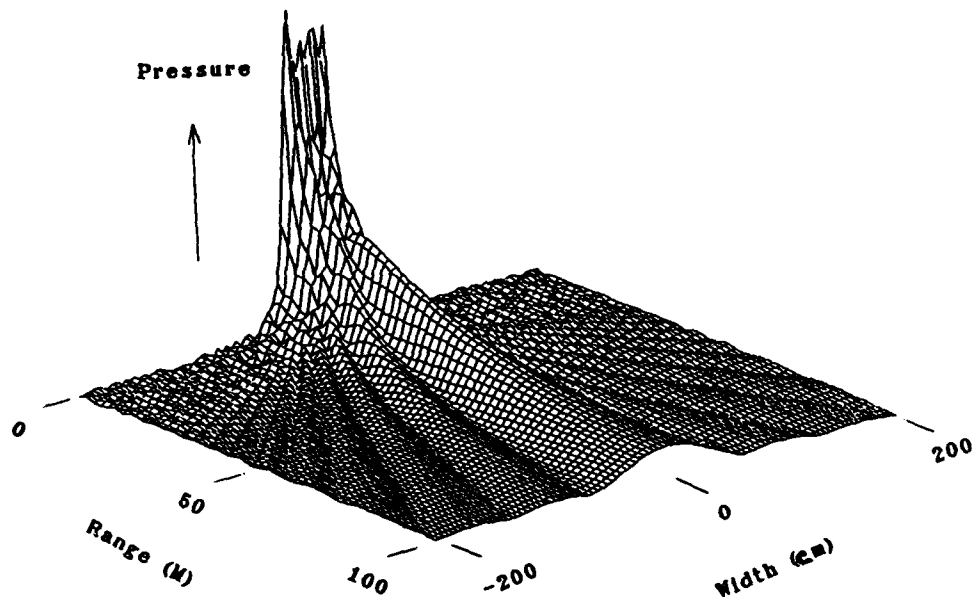
Figure 2-2: Klein 500 KHz Side scan sonar   horizontal beampattern.
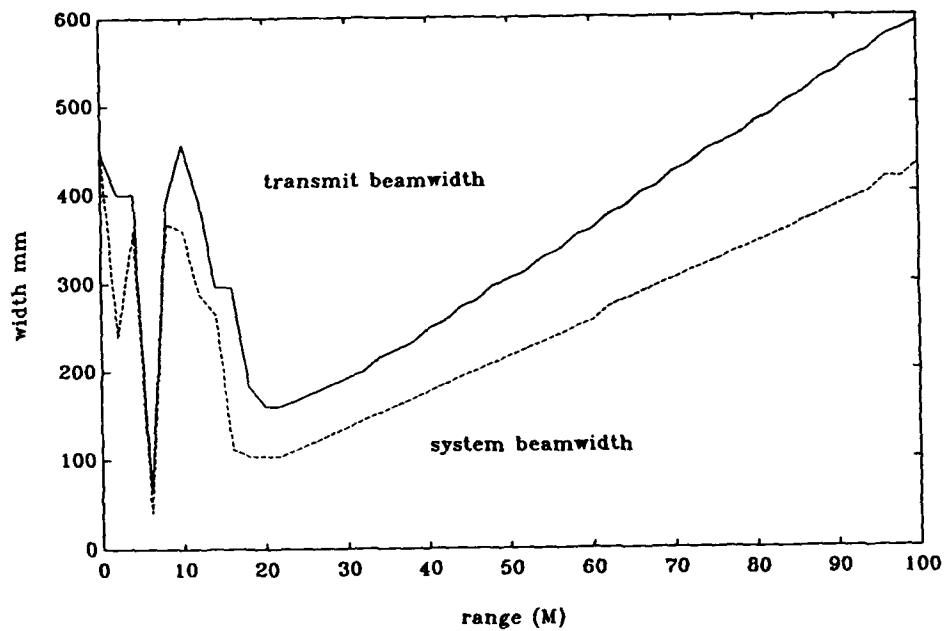
Figure 2-3: Klein 500 KHz Side scan sonar horizontal beamwidths

18

array and hence the same beam pattern is used for both transmission and reception, resulting in an overall system beam pattern that is the product of the array beam pattern. Figure (2.3) shows that although the horizontal beamwidth varies with range, once the complicated near field is overcome at approximately 30 meters the amplitude of the main lobe asymptotically approaches a linear, diverging shape. Note that this discussion does not include effects of phase.

The use of a finite rectangular array introduces sidelobes into the transmit and receive beam patterns which influence the imagery obtained. The presence of vertical sidelobes results in a departure from the ideal system described by eqn. (2.1) in two ways. The real beam is not spatially finite and therefore not limited to grazing angles below the horizontal. Because acoustic energy is radiated above horizontal the sea surface is insonified. Because of nearly perfect reflection of sound at the air-water interface, acoustic energy scattered from this surface can interfere with acoustic returns from the seafloor, despite attenuated array sensitivity in the vertical sidelobes. In the extreme case of a sidelobe oriented vertically at the surface the specular reflection of sound impinging normally on this surface results in a strong, narrow trace on the sonogram at a distance equal to the depth of the sonar towfish. When this occurs towfish depth information as well as surface clutter noise are included in the image. Additionally, eqn. (2.4) shows that beam intensity within the desired range of grazing angles is not constant. As a result portions of the seafloor which are insonified at different grazing angles will not be insonified with the same acoustic intensity. The resulting sonogram will show variations in intensity which may require compensation. Fig. (2.4) is a sonogram which shows all of these effects.

The presence of horizontal sidelobes causes image corruption in the case of strong objects or *targets* in the same manner as interaction between vertical sidelobes and the highly reflective air-sea interface does. In cases of strong targets located in low scattering backgrounds, echoes are received from the target both before and after the target is insonified by the main horizontal lobe. The sonogram subsequently records linear traces as shown in fig. (2.5). In this instance steel cylinders standing on end are located in the dark region to the right of the high intensity image region. Sidelobe returns from these targets are seen as vertical bright lines extending above and below the main lobe return. Objects displaying this signature are
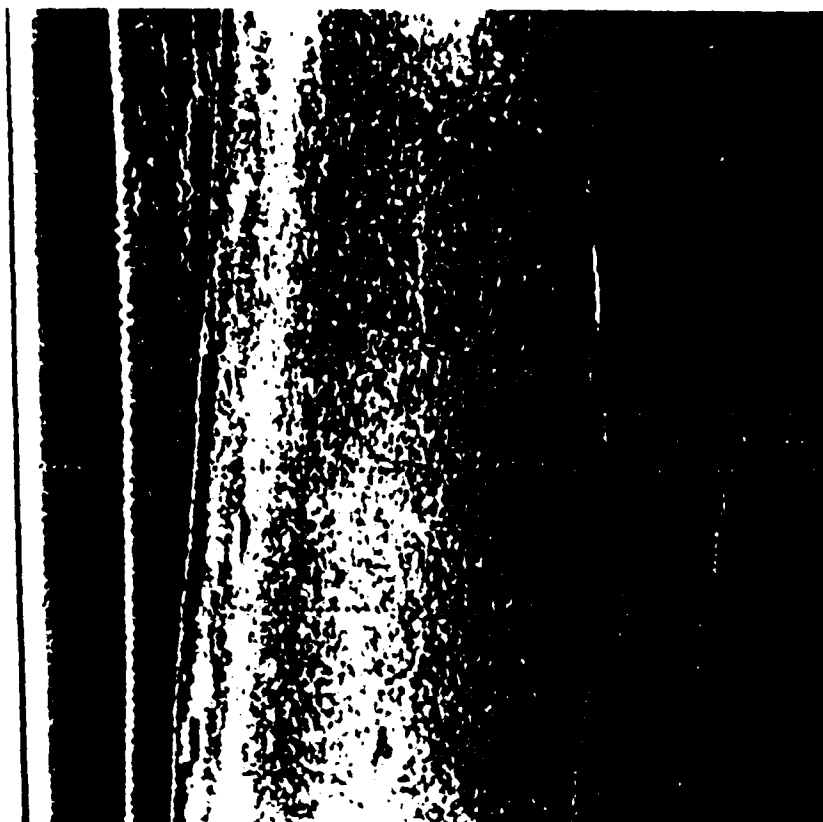
Figure 2-4: Side scan sonar image showing surface return, surface scattering, and bottom return.

Figure 2-5: Side scan sonar image showing targets with horizontal sidelobes.

usually corner reflectors or vertical cylinders. They present a specular reflecting surface to the towfish regardless of their mutual orientation. Objects located in high scatter backgrounds may not display sidelobes as they may be overwhelmed by bottom backscatter.

Lateral resolution is determined by pulse length, which is a function of system bandwidth [5]. Referring to fig. (2.1) the lateral dimension of the sample area, which is the intersection of the transmitted wave front with the bottom, is

$$\Delta y = \frac{c\tau}{2cos\theta} \tag{2.6}$$

Here c is the speed of sound in water and τ is the acoustic pulse length. Transverse resolution is a function of grazing angle and for a given value of Z improves with range.

2.2 Returned Signal

The returned signal consists of several components including acoustic energy backscattered from the bottom, energy reflected and scattered from individual objects of interest on the bottom, energy scattered from the ocean surface, and scatterers suspended in the medium. Depending on the type of survey performed the desired component is either the bottom backscattered energy or the energy from objects, the other components constituting noise terms. If the aim of the survey is to locate objects, the bottom backscattered component may also become a noise term [17]. Both components of interest have been studied extensively with the bottom backscatter continuing to be a area of research.

Bottom backscatter characteristics are not well understood due to the complex physics needed to describe an acoustic pulse interacting with a real bottom. Real bottoms are complex because of material inhomogeneities, variable roughness, and limited knowledge of the type of bottom which exists in a specified area. Therefore there is presently a lack of understanding of many of the details of high frequency acoustic scattering. The fundamentals, however, are well understood.

For most commercial side scan sonar operating frequencies the energy backscattered by the bottom is returned by material in the first few centimeters of the bottom. Studies of penetration depth for acoustic energy in common seafloor materials [31] yield the following

empirical relationship

$$\alpha = kf \tag{2.7}$$

where α is the attenuation coefficient in dB/M, k is a constant with units of dB/KHz which depends only on material, and f is frequency in KHz. The relationship has been shown to maintain its linearity over a range of 1 KHz to 1 MHz and may extend to frequencies as low as 10 Hz. Values of k have been measured on the seafloor and typically average 0.25 dB/M-KHZ, with readings as low as 0.05 dB/M-KHz for silt-clay bottoms and as high as 0.30 dB/M-KHz for sandy silt bottoms. Taking the average value of 0.25 dB/M-KHz at 100 KHz yields 3dB attenuation for a round-trip bottom penetration of 6 cm. Considering that common side scan sonar the grazing angles are normally in the range of 0 to 45 degrees it is evident that only the shallowest scatterers contribute to the backscattered energy received by typical commercial side scan sonar systems operating at frequencies of 50 to 500 KHz.

The underlying phenomenon responsible for backscatter is the reflection of acoustic energy from the various types surfaces which comprise the bottom. In the absence of any relief, the reflection of sound from a perfectly smooth bottom of determined composition is modeled by the reflection coefficient [5]

$$C_R = \frac{\rho_2 c_2 sin\theta_1 - \rho_1 c_1 sin\theta_2}{\rho_2 c_2 sin\theta_1 + \rho_1 c_1 sin\theta_2} \tag{2.8}$$

defined as the ratio of the amplitudes of the reflected and incident pressure fields and parameterized by material properties density ρ speed of sound c, and grazing angle θ. The subscripts refer to the regions forming the interface with region 1 typically water and region 2 sedimentary material. The product ρc is referred to as specific acoustic impedance. The difference in specific acoustic impedance across the interface is the primary determinant of reflectivity. The grazing angle in the bottom θ_2 is determined by Snell's Law of Refraction. Reflected energy in this model travels in the specular direction and therefore generally is not backscattered in the direction of the sound source, as in the side scan sonar case.

The first order approach to modeling non-specular backscatter is the Rayleigh model [3] , which assumes a perfectly reflecting surface parameterized by the Rayleigh Parameter $P = 2k\sigma sin\theta$ where $k = 2\pi/\lambda$ is the acoustic wave number, θ is the grazing angle, and σ is the root-mean-square surface roughness. For $P \gg 1$ the surface is considered rough and scattering occurs . This scattering is due to phase differences along different reflection

paths from this non-flat surface giving rise to constructive interference of reflected sound in non-specular directions. The amount of acoustic energy scattered is described by the bottom scattering coefficient m_b

$$I_s = SR^{-2}I_i m_b(\theta, \phi) \qquad (2.9)$$

where I_s is the far field scattered intensity, I_i is the incident acoustic intensity, S is the area of the scattering surface, and R is the distance from the scattering surface to the observation point. The scattering coefficient is therefore a function only of scattering angle. To determine the value of m_b the Method of Small Perturbations is used [3]. This assumes a scattering surface deviating only slightly from a mean (usually flat) surface, small Rayleigh parameter, and small local slopes. Also assuming scattering area dimensions large with respect to acoustic wavelength and surface correlation radius; the scattering coefficient is calculated to be

$$m_b(\theta, \phi) = 4k^2 cos^4\theta G(\chi) \qquad (2.10)$$

where $G(\chi)$ is the spatial roughness spectrum of wave vector χ. This scattering regime is called "resonant scattering" in that those spatial frequencies in $G(\chi)$ equal to the Bragg wave number $2kacos\theta$ result in scattering. The scattered sound field is generated through constructive interference of wavelets reflected from these frequencies.

The preceding development has been further analyzed and developed [20] to take bottom material properties into account. However, these derived expressions are generally applied to low frequency theory and experiments below 10 KHz [14] . In the high frequency side scan sonar case surface element dimensions become comparable with wavelength. In this scattering regime individual facets of the scattering surface having sufficient slope and orientation cause specular reflection. The backscattered field in this case is the summation of the scattered energy from all such oriented facets in the insonified region. An expression for m_b in this case is [3] :

$$m_b(\theta, \phi) = R^2(8\pi\delta^2 sin^4\theta_o)^{-1}exp\left(-\frac{cot^2\theta_o}{2\delta^2}\right) \qquad (2.11)$$

where θ_o is the grazing angle for the monostatic side scan case and δ is the mean square surface slope, which for an isotropic scattering surface is not azimuthally dependent.

Urick [30] provides a Lambertian model for the angular dependence of high frequency

backscatter which is based largely on optical backscatter and on empirical acoustic data.

$$m_b(\theta_o) = 10log\mu + 10log(sin^2\theta_o) \qquad (2.12)$$

Here μ is the backscatter strength at normal incidence for the bottom interface, this model is only valid for grazing angles of less than 45 degrees. This model is therefore applicable to side scan sonar since grazing angles greater than 45 degrees occur only at the closest ranges and are a small portion of the overall image.

The nature of high frequency backscatter strength continues to be subject to experimental investigation with measurements and analysis of material, roughness, and grazing angle dependence continuing [14] . A trend which has been observed in various experiments is that different areas with the same bottom classification show a large degree of variability in scattering strength, making knowledge of bottom type at best a tenuous predictor of scattering strength.

The preceding analysis described the static case of fixed sonar and scattering surface. When a side scan sonar image is created the sonar is moved over a wide area and numerous subregions of the bottom are sampled. One parameter of interest in this case are the statistics of this ensemble of subregions which make up the side scan image. This translates into the probabilistic distribution of pixel intensities for scattering surfaces included in the image. The probabilistic nature of this process is due to the varying numbers and spatial distribution of individual scatterers between sonar transmission cycles. One model which has been developed to describe the statistics of backscatter in both the acoustic and electromagnetic cases is the Rician model. Rician statistics describe the probabilistic nature of scattered echo intensities as well as accounting for the overall echo intensity when the return contains both energy from a scattering surface and energy from specular reflection from non-scattering targets [28] . The form of the Rician distribution is

$$p(A) = \frac{2(1+\gamma)A_e}{\langle A_e^2 \rangle} exp\left(-\frac{(1+\gamma)A_e^2 + \gamma\langle A_e^2 \rangle}{\langle A_e^2 \rangle}\right) I_o\left(\frac{2A_e[\gamma(1+\gamma)]^{1/2}}{\langle A_e^2 \rangle^{1/2}}\right) \qquad (2.13)$$

where A_e is the echo amplitude, I_o is the modified Bessel function, and γ is the ratio of coherently reflected energy echo and incoherently scattered energy in the echo. In the case of pure scattering from the bottom the Rician distribution reduces to the Rayleigh distribution. For purely coherent energy such as the case of specular reflection the Rician distribution reduces

25

to a Gaussian. Cases involving both types of echo energy form a continuum of distributions between these two cases. Applying the Rician framework to side scan sonar, specular targets reflect coherently and the bottom scatters incoherently.

The vertical projection of targets on a flat bottom gives rise to a second acoustic signature, that of the acoustic shadow. The received echo for the period immediately following the receipt of the echo from a vertical projection from the bottom is lower than adjacent bottom regions. This occurs because energy received during this period would normally be backscattered from the region of the bottom behind the object. However this energy is intercepted by the vertical projection, diminishing the amount of energy available to the associated bottom region. In the case of objects having large target strengths compared to the surrounding backscatter strength, detection occurs by the salient return from the target as shown by fig (2.5). In the opposite case of nearly equal contributions from the object and bottom backscatter the target is hard to distinguish from the backscatter, but the shadow provides detection. Fig(2.6) is such a case, where individual rocks are identifiable by their shadows while the returns from their vertical surfaces are difficult to distinguish from adjacent backscatter.

The remaining components of the returned signal are surface and volume reverberation terms. Both components are composed of the superposition of numerous acoustic echoes from large numbers of scatterers. The surface returned energy is Rayleigh distributed as it is a purely incoherent scatterer consisting of moving wave facets. This can be viewed as an example of Rician statistics for purely incoherent returns. Volume reverberation is due to scattering of acoustic energy by objects or inhomogeneities suspended in the medium. Statistics for volume reverberation are approximately Gaussian [22] .

2.3   Stochastic Effects

Energy propagation properties of water are subject to temporal variation. This variation is caused by temporal and spatial fluctuations in temperature, salinity, pressure, turbulence, and foreign material; which affect propagation loss, sound velocity, phase coherence, refraction, and reverberation. Urick [3] provides an overview of signal fluctuation in the ocean medium, and attributes open ocean fluctuations to patches of thermal inhomogeneity which act as acoustic lenses to focus and defocus transmitted acoustic signals. Variation between trans-
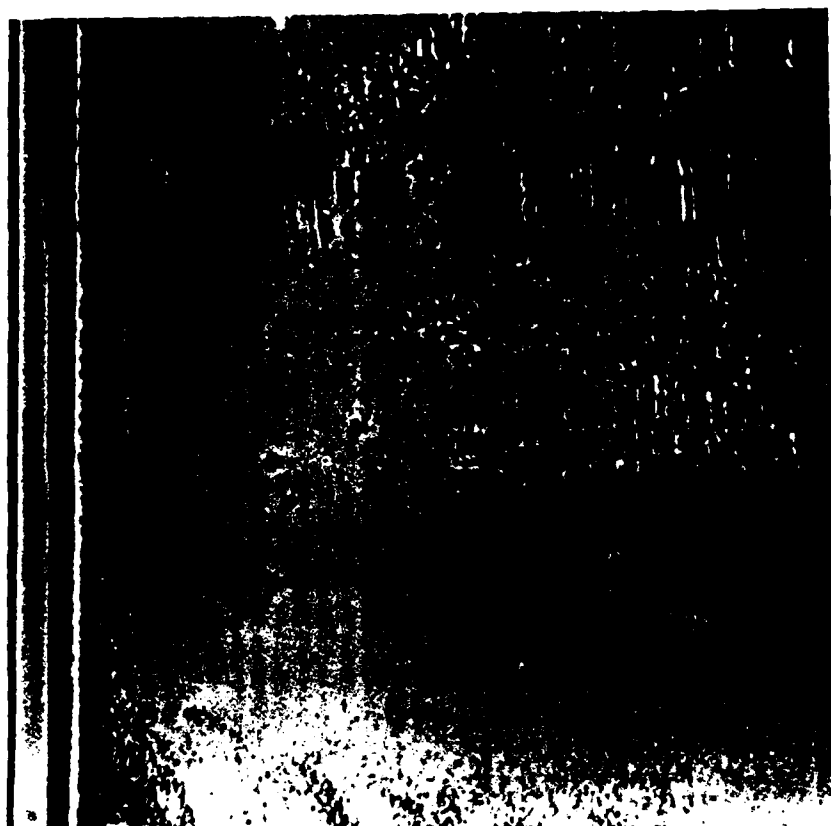
Figure 2-6: Side scan sonar targets (rocks) on high backscatter bottom

27

missions is parameterized by the Coefficient of Variation, V, defined as the signal amplitude standard deviation divided by the mean amplitude. Empirical formulas for the Coefficient of Variation specify an $r^{3/2}$ increase in V out to a transition range defined by

$$r_o = \frac{ka^2}{4} \qquad (2.14)$$

where a is the characteristic patch size. Beyond this range V increases as $r^{1/2}$. This concept was mainly employed to analyze data about frequencies below 60 KHz and ranges generally in excess of 100 M. By comparison there is little published work in high frequency acoustics which would be more applicable to the side scan case. One study of phase coherence [4] showed that for a 100KHz signal over a 48 M path length, parameters typical of side scan sonar, signal phase was coherent with a standard deviation of less than 0.31 radians.

Studies of the phase coherence of high frequency sound are discussed to a limited degree in open literature, as are studies of temporal amplitude fluctuations for frequencies and ranges outside the commercial side scan sonar range. Similarly, little information is available on temporal amplitude fluctuations in our range of interest. It is therefore necessary to investigate these fluctuations in order to quantify intensity variation in side scan sonar imagery.

2.4   Spatial Effects

2.4.1   Towfish Instability

The side scan sonar image is displayed on a video or graphic recorder upon which the returns from successive transmission cycles are plotted as sequential parallel lines. Neglecting all other spatial effects for the present discussion, this representation of the sea floor will be correct only if the insonified bottom strips are parallel and sequential. Otherwise, the image produced by the display will constitute a geometric transformation of the sonar data. There are several ways in which the towfish may deviate from the ideal case of constant velocity, constant attitude travel over the bottom and cause image to distortion.

Towfish instabilities may be divided into trajectory and attitude instabilities. Trajectory instability is the deviation from a straight, level, constant velocity path through the water. This may occur when the speed over ground X' is not constant because of currents or changing
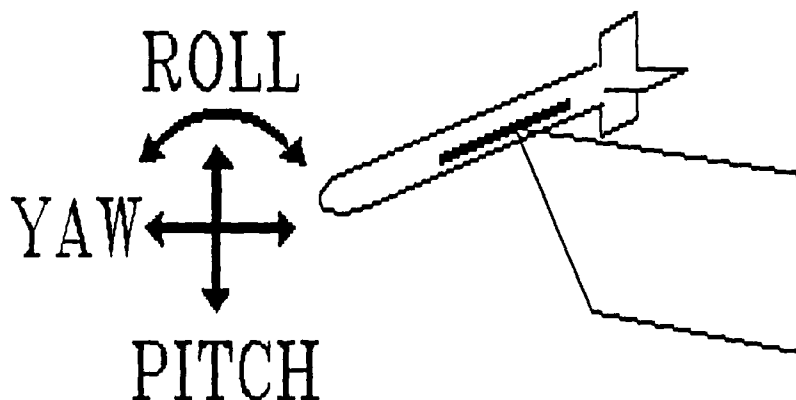
Figure 2-7: Towfish attitude instabilites

speed of the survey vessel. The image coordinates corresponding to X and R are x and r but unless the display is compensated for speed over ground variations or variations in X', x' remains constant and the image distorts along the x axis. Some systems attempt to compensate for variations in X' by varying x' in response to survey craft speed through the water, but variation in X' may be due to currents in which case survey craft speed through the water will not provide sufficient information to remove this distortion. The other trajectory instabilities are variation in Z, a variation in depth or altitude commonly referred to as heave; and variation in Y, or cross track instability which may be caused by unsteady survey vessel course or by current components perpendicular to the survey track.

Attitude instabilities are rotations of the towfish about its axial (roll), transverse (pitch), or vertical (yaw) axis (fig 2.7). These cause image distortion by displacing the acoustic beam so that it insonifies a portion of the bottom other than that perpendicular to the survey track and directly beneath the towfish. Excessive instability may also affect image line-to-line intensity in that the acoustic beam insonifies one region of the bottom and slews away before the echo is received. The result is off axis receipt of the echo with subsequent attenuation.

29

2.4.2　Slant Range Correction

A range dependent distortion occurs in all side scan images due to the fact that the image presents the range of a given bottom feature as the distance from the towfish R while the human observer is accustomed to viewing the image as a representation of the bottom in which range is the distance from the bottom track to the object Y. For $R \gg Z$ this is not normally a problem since this implies a small grazing angle and $Y \simeq R$. However for R comparable to Z the relationship

$$Y = \sqrt{R^2 - Z^2} \tag{2.15}$$

yields a larger change in Y for a given change in R. It is desirable to correct slant range distortion because this results in an image whose lateral and axial dimensions are isometrically related when compared to the bottom it represents. Correcting this geometric effect is straightforward and involves remapping and interpolation of the pixels along each image line using eqn. (2.15). The result is an image representing X and Y rather than X and R.

Chapter 3

Conduct of the Experiment

Experimental data was collected using side scan sonar systems built by Klein Associates. With the exception of the transmitted waveform variability test which was conducted at the Klein Associates test tank in Salem, New Hampshire in March 1988 all experiments were conducted along the Woods Hole Oceanographic Institution (WHOI) dock during July and August 1987. This site was chosen because of its facilities and 17 meter minimum depth of water. This allowed realistic imaging geometries to be attained while in a pierside environment that allowed a higher degree of control than the normal underway side scan survey.

3.1 Transmitted Waveform Variability Test

The aim of the test tank experiment was to determine the characteristics of the acoustic transmission in order to evaluate transmission variations as a source of variability in sonar images.

In this test a Klein model 422s-101ef 100 KHz towfish was suspended in the Klein test tank approximately 5 meters from a receiving hydrophone (fig. 3.1). To eliminate multipath interference in recorded data one the two towfish transducers was disabled so that there was only one acoustic source present. Additionally, the geometry of the towfish, hydrophone, water surface, and tank walls were adjusted such that the shortest indirect path between the transducer and hydrophone was .46 meters or .31 msec longer than the direct path assuming
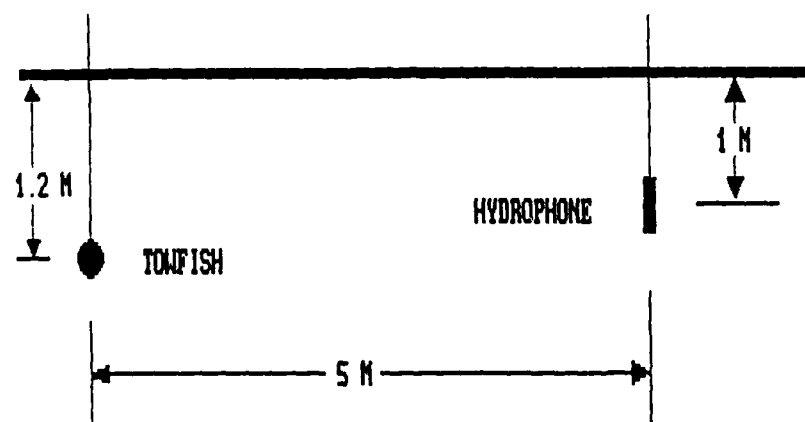
Figure 3-1: Test tank experimental configuration.

$c = 1500m/s$. Subsequent data analysis showed the pulse length to be sufficiently short to prevent multipath corruption of data.

The side scan sonar system was operated normally with towfish power and transmit key signal supplied by the model 521 recorder. The waveform transmitted by the towfish was sensed by the hydrophone and sent to the data gathering system. The data gathering system consisted of an IBM PC-AT personal computer containing a Data Translation DT-2851 frame grabber card, which was configured to accept slow-scan image data (fig 3.2). In this mode of operation the DT-2851, which is designed to acquire 512 row by 512 column images from a variety of sources, acquires the image one row at a time. The 500 KHz sampling rate and 512 column image dimension resulted in a temporal observation window of 1.024 msec. Upon filling, each image buffer was written to a hard disk file. The complete data set consisted of 7 files or 3584 transmissions records.

Because the frame grabber was designed to acquire images represented by non-negative voltage signals the hydrophone signal was first passed through a bias/scaling circuit which biased the signal to +0.3V and scaled it to fit within the 0-0.6V digitizing range of the frame grabber. The transmit keying signal generated by the sonar recorder was channeled to a Metrabyte CTM-05 counter-timer board in the computer which in turn supplied the beginning of line signal and digitizer timing signal to the frame grabber. The arrival of the transmit keying signal at the counter-timer board was delayed by a multivibrator circuit since
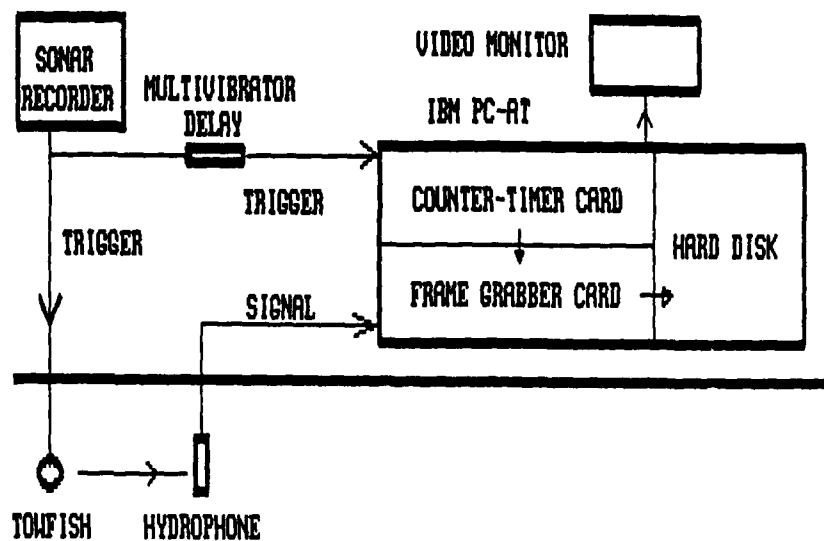
Figure 3-2: Test tank electronic configuration.

the expected time for the acoustic transmission to traverse the distance from the towfish to the hydrophone was approximately 3.3 msec, which was considerably longer than the observation time window. Monitoring of incoming data was performed with an oscilloscope and real-time video monitor display driven by the frame grabber.

3.2 Phase One Pierside Test

The objective of the phase one pierside experiment was to observe the temporal variation in echo return strength for non-varying imaging geometry. Normal side scan sonar imagery is obtained from a moving towfish with limited control over its trajectory or attitude. In this situation image variability may be attributed to variations in imaging geometry as well as to variations in system performance or properties of the medium. In this experiment the geometric effects were minimized by rigidly restraining the towfish, allowing elimination of geometric variability and observation of variability due only to system and medium effects. Unlike the normal side scan survey where the bottom strip is moved axially between transmissions, in this experiment the insonified strip of bottom was the same for all transmissions.

33
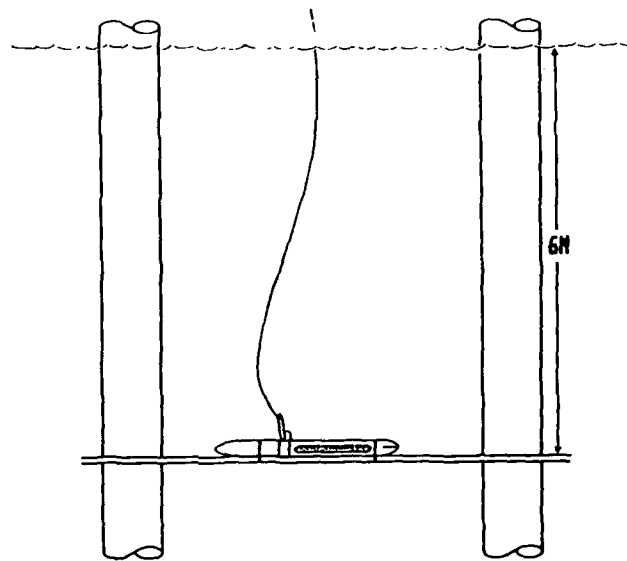
Figure 3-3: Phase one pierside experimental configuration.
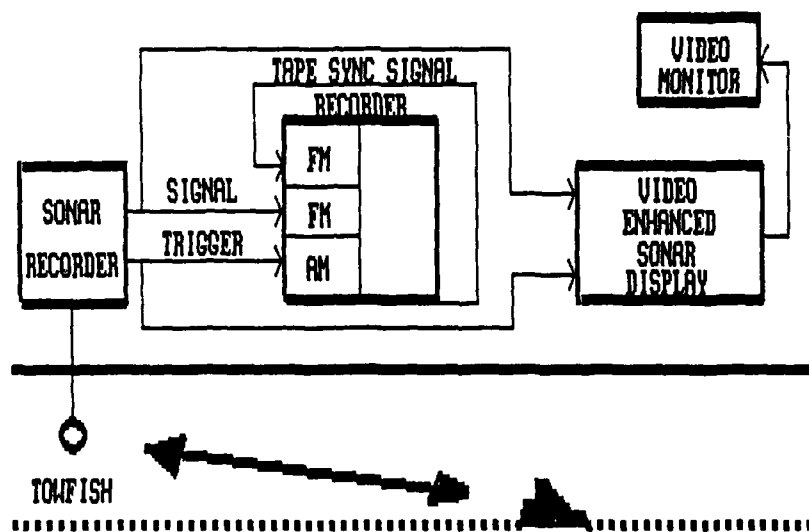
Figure 3-4: Phase one and two pierside electronic configuration

This removed image variability due to scene variations and permitted observation of temporal variability due to acoustic effects.

To eliminate towfish motion the same sonar components used in the test tank experiments were deployed from the WHOI dock with the towfish mounted on a steel box beam which spanned two concrete dock pilings 10.5 M above the bottom in 17 M of water (fig.(3.3)). The towfish was oriented parallel to the pier in a level attitude with the array axis oriented approximately 20 degrees below the horizontal. The towfish repeatedly insonified a portion of the bottom extending 100 M outward from the pier. As in the tank experiment one towfish acoustic channel was disabled to prevent data corruption due to interfering acoustic sources.

Data was recorded on a Hewlett-Packard HP-3898 instrumentation tape recorder (fig. 3.4). An FM channel and 15 ips tape speed were used to record sonar data in order to obtain a recording frequency response of 0 to 5 KHz. Other direct channels of the tape recorder were used to record the sonar keying signal and tape recorder synchronizing signal. Time varying gain (TVG), an analog signal processing feature of the Klein sonar recorder which attempts to invert the amplitude variation with range of received signal by applying a gain varying with range, was defeated in all experiments of this thesis. This was done because TVG represented an effect on the data which was difficult to quantify and potentially variable between transmit cycles. Data monitoring was done with an oscilloscope and with a video monitor through a Colmek Video Enhanced Sonar Display which converted Klein sonar recorder output to video format. Several data sets were collected during this experiment, the set analyzed in this thesis was taken on 13 July 1988 between 1302 and 1308 hours with slack currents, calm winds, and flat surface conditions. Tape recorded data was stored until the personal computer based data collection system was available to digitize it. Digitizing was done at 7.68 Khz with an anti-aliasing filter with 3.8 Khz half power bandwidth and 80 dB/ decade roll-off. The square pulse keying signal recorded on tape was corrupted by the limited recording bandwidth and was reconstituted by sending the recorded pulse to a function generator which supplied a single square pulse to the frame grabber. As in the tank experiment individual transmissions were recorded as lines in the DT-2851 image buffers but in this case 1024 samples per transmission were recorded on two adjacent image lines with 256 transmissions recorded per image file. Assuming $c = 1500 M/s$, 1 degree beamwidth,

and 0.1 msec pulse length the samples represented independent subregions of the bottom spaced at 9.8 cm intervals with approximate transverse (along beam) dimension of 7.5 cm and range varying axial (across beam) dimension of approximately 50 cm near the towfish and approximately 170 cm at 100 M range.

3.3   Phase Two Pierside Test

The objective of the phase two pierside experiment was to generate realistic side scan sonar images while exercising sufficient control over towfish attitude to prevent excessive corruption of imagery due to towfish motion. The degree of trajectory and attitude control available with the track allowed multiple imaging runs along a reproducible track, eliminating most spatial effects. These images were then used in the development of the image registration algorithm.

The experimental setup was identical to the phase one experiment with the exception of the towfish. In order to image the bottom adjacent to the pier a track was constructed which allowed the towfish to translate through the water along a straight trajectory at constant depth. Fig. (3.5) shows the track design, which carried the towfish along a 28 M run at depths below the support cable of 3 M to 15 M in 3 M increments. The support cable was approximately 2M above the water's surface. The design allowed adjustment of towfish pitch and roll angles as well as depth. The carriage was pulled down the track by a rope which was pulled by an electric winch at 20 cm/sec. Sonar pulse repetition rate was 4 Hz which resulted in a axial spatial sampling interval of 5 cm. The bottom was imaged with both 100 KHz and 500 Khz towfish, with some images taken with the test targets listed in table (3.1) present. The bottom imaged during this experiment was composed of a silty sand with scattered rocks up to 1 M in diameter.

OVERHEAD:

CABLE

|←————————————28 M————————————→| PIER

END:

CABLE

PIER   CARRIAGE

3 M

TOWFISH

SIDE:

CABLE   BLOCKS

TOWLINE

DEPTH/PITCH
ADJUSTMENT

0-15 M

ROLL
ADJUSTMENT
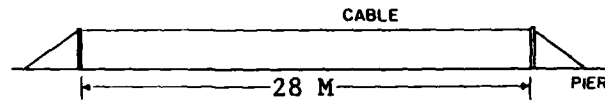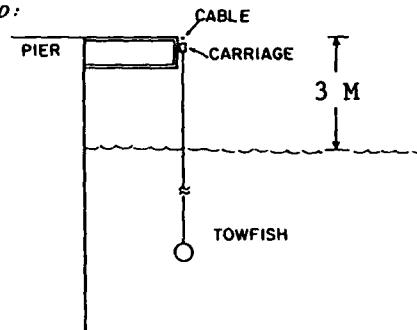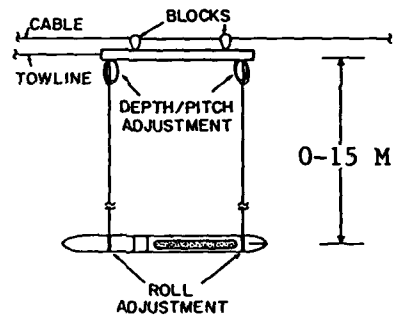
Figure 3-5: Pierside phase two experimental configuration.

Flooded steel drums:

	5 Gal.	10 Gal.	40 Gal.	55 Gal
Height (cm.):	34.3	68.6	74.9	90.2
Diameter (cm.):	28.6	36.2	47.0	57.8

Aluminum Corner Reflector

Plate Thickness (cm):	0.5
Face Height (cm):	30.4
Face Width (cm):	15.2

Table 3.1: Pierside phase two test targets

Chapter 4

Acoustic Transmission Analysis

The data taken in the Klein test tank was analyzed to quantify variability at the first step of the imaging process, the transmitted waveform. The variability exhibited here sets the lower bound on overall system variability, with effects later in the process contributing increased variability.

4.1 Waveform

The advertized characteristics of the Klein 422s-101ef 100 KHz towfish used in this experiment are a pulse length of 0.1 msec and a peak sound pressure level of 228 dB re 1 micropascal. Figure (4.1) is one of the transmitted waveforms recorded in this data set, and shows a plot of pressure versus time in microseconds. The observed waveform is a carrier of approximately 122 KHz modulated by an envelope which first experiences a linear amplitude rise then decays exponentially. As can be seen the pulse duration is approximately 100 microseconds. An analytic model for this waveform is

$$p[n] = \begin{cases} 9n\,sin(1.51n) & 0 \leq n \leq 14, \\ 126exp(-\frac{n}{15})sin(1.51n) & 14 \leq n \leq 128 \end{cases} \qquad (4.1)$$

Here n is sample number, which at the 500KHz sampling rate corresponds to 2 microseconds. This model waveform is shown in fig (4.2). A noticeable difference between the two waveforms is that the real waveform appears to have been clipped for negative pressures. The cause of the clipping is probably transducer cavitation. Clay and Medwin [5] indicate

39
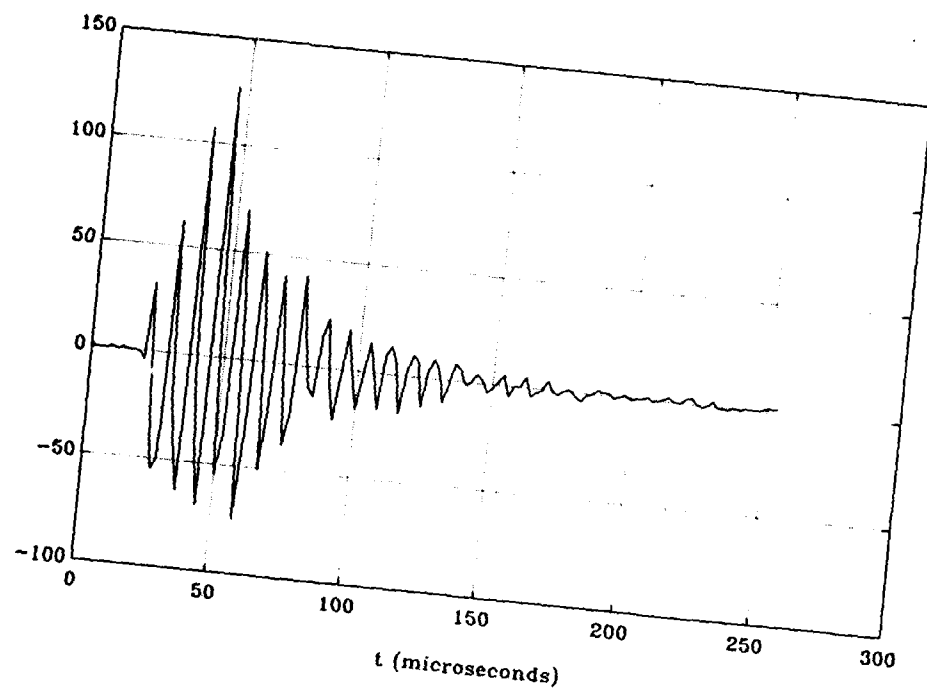
Figure 4-1: Klein 100 KHz waveform measured in Klein test tank
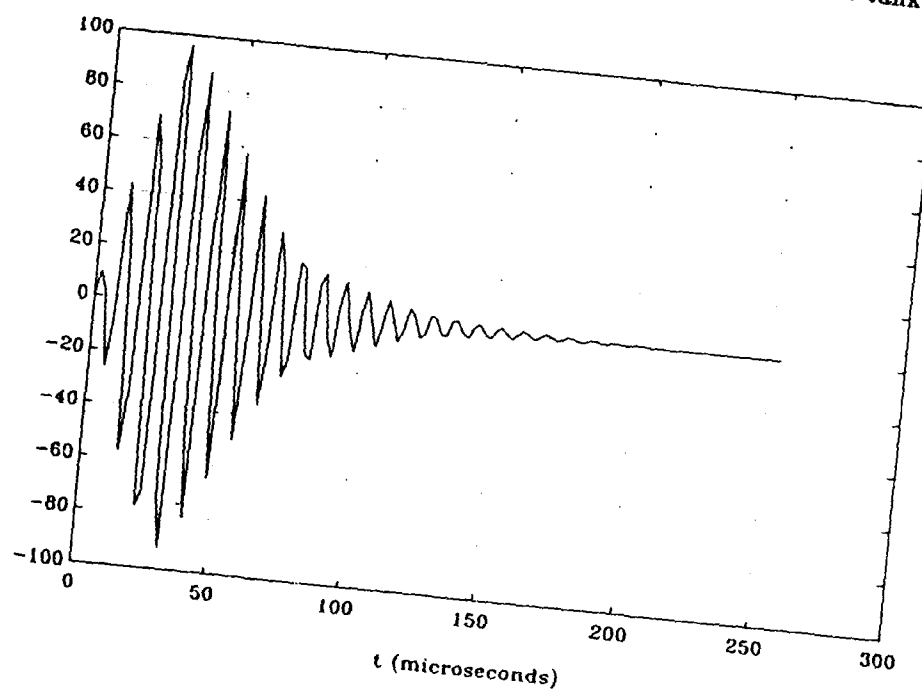
Figure 4-2: Analytic model of Klein 100KHz waveform

40

that the threshold for cavitation is a source level of approximately 220 dB re 1 micropascal. Assuming a water temperature of 70 degrees the pressure at which cavitation should occur, the saturation pressure of water at this temperature, is 0.4 psi absolute [16] . The transducer was submerged to 1.2 M giving it an ambient pressure of 16.4 psi absolute. Cavitation should therefore occur at vacuum of 16 psi relative to the ambient pressure at the transducer. Assuming that the clipping level in the observed waveform is 16 psi below the ambient pressure shown, the peak pressure occurs at 29.3 psi above ambient, corresponding to a source level of 226 dB re 1 micropascal or 2 dB from the advertized source level. It is therefore likely that cavitation is the cause of the observed transmitted waveform behavior. Calculations show that to eliminate all cavitation the transducer must be at an ambient pressure corresponding to a depth of 950 feet, so for all but very deep survey work cavitation can be expected to occur.

4.2   Energy Fluctuation

Of particular interest is the fluctuation of total energy in each transmission, knowledge of which would allow determination of expected image intensity variation due to transmission power variability. Energy of each waveform in the set was calculated as

$$E = \sum_{i=0}^{128} p[n]^2 \tag{4.2}$$

with the energy distribution shown in fig (4.3) The statistics of this distribution are shown in table (4.1) The distribution also appears to follow a low frequency sinusoidal moving average,

mean:	77300
standard deviation:	4900
standard deviation/mean	0.063

Table 4.1: 100 KHz towfish transmit energy fluctuations.

however the sinusoidal appearance is coincidental in that other data sets taken the same day show that although the envelope of the distribution does undulate it generally follows a random trajectory. Because the method used to calculate total energy was a sum of squares

41
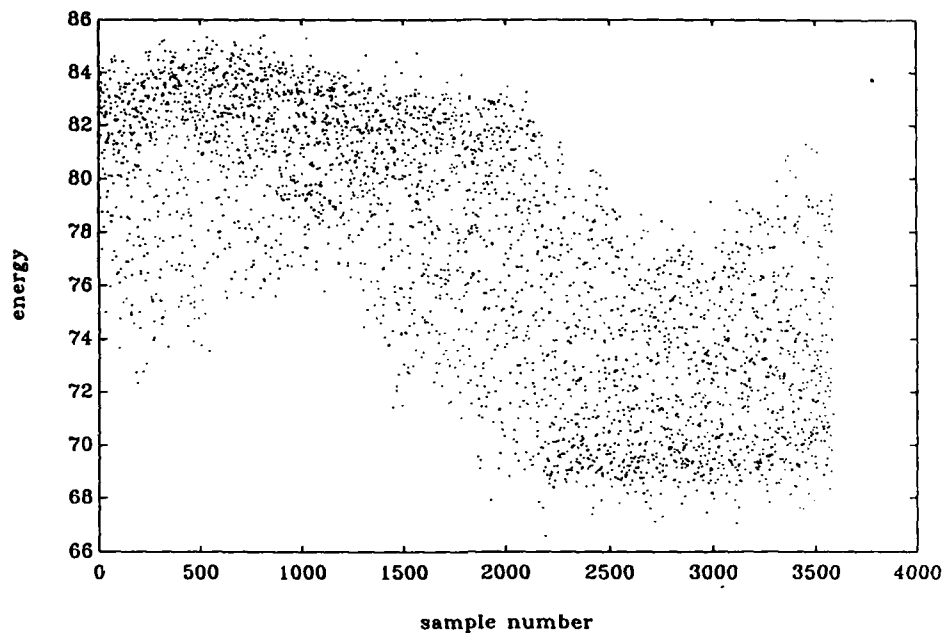
Figure 4-3: Total sonar waveform energy vs. sample number

of the data points the possibility existed that the energy distribution was faulty due to sparse sampling. The sum of squares is dominated by the few large peaks in the waveform, but since the sampling rate and carrier frequency combine to produce approximately 4 samples per carrier cycle it was thought that the energy distribution could be explained by the locations of sample points on these large peaks. Fortuitous waveforms with samples near the peaks would therefore be calculated to possess more total energy than would the same waveform with sample points shifted by $\pi/4$ in carrier phase. Testing this hypothesis consisted of repeatedly sampling the analytic waveform model with successive sampling phase shifts of 0.01 radian, resulting in 151 unique possible energy measurements since the fundamental frequency was 1.51 radians/ sample. The result of this experiment is shown in fig (4.4). Although correlation can be observed between sampling phase and total energy the statistics shown in table (2) show that the energy estimation method was not the primary cause of the observed 6% energy fluctuation.
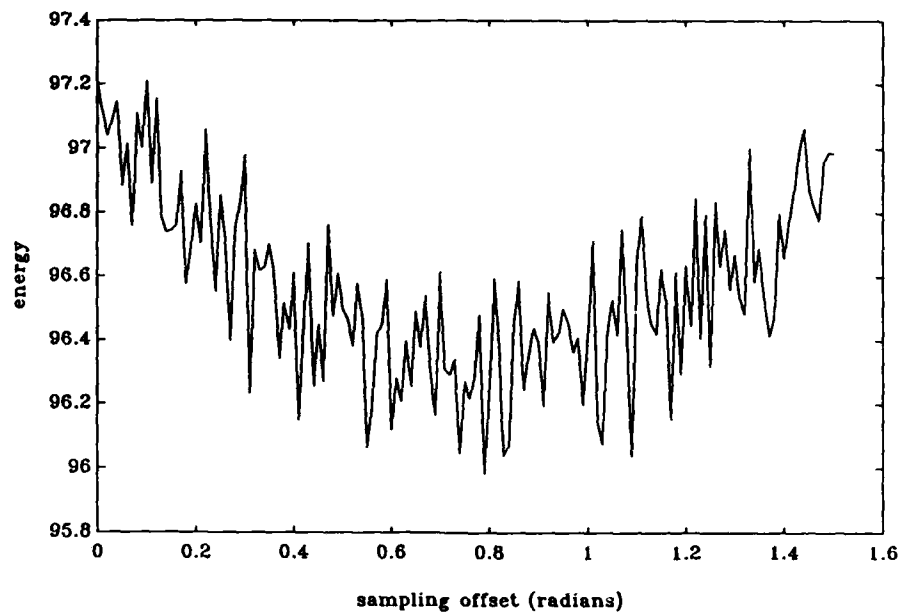
Figure 4-4: Analytic waveform total energy vs. sampling offset

mean:	96600
standard deviation:	265
standard deviation/mean	0.0027

Table 4.2: Statistics for sampling offset experiment.

4.3 Frequency Content

Having characterized the temporal variability of the sonar power level, the sonar transmission was decomposed into component frequencies for analysis of variability. The length of each record was originally 128 points, which corresponds to a infinite length waveform windowed by a 128 point windowing function with subsequent loss of frequency resolution and provides for only 128 frequency samples after Fourier transformation [23]. Each waveform was zero padded to eight times its original length, which normally interpolates between frequency samples and normally provides no new information. However, in this case the waveform decays to zero, which implies that by padding the truncated waveform the original waveform is restored and the windowing function is widened. The mean power spectral density is shown in fig (4.5). The predominant frequency is the carrier at 122 KHz with a half power bandwidth of 9.2 Khz. A significant amount of power is seen at the extreme ends of the spectrum, with local maxima at 226 KHz, 241 KHz, and at the extreme frequency 250 KHz. This region is probably the result of the generation of harmonics of the fundamental frequency due to the non-linear response of water to high amplitude pressure fluctuations [21]. Lesser energy is seen at frequencies below 122 KHz. The local maximum at 65 KHz is probably a cavitation generated subharmonic, a phenomenon described by Desantis et al. in which a frequency half the fundamental frequency is produced by cavitation [8]. Broadband redistribution of energy across the spectrum is also an effect of cavitation and is observed here. Figure(4.6) is the spectrum of the analytic wave model with clipping performed on negative pressure amplitudes to simulate the observed waveform. In this case energy is seen to be redistributed in a manner similar to the observed case. Without simulation of cavitation by clipping no observable energy occurs in these regions.

The variance of the individual spectral components is shown in fig (4.7). It shows a form similar to the mean power spectral density in fig (4.5), with peaks in the same areas on both graphs. However the region corresponding to the carrier frequency is proportionately smaller than the regions of significant energy outside the carrier. This indicates that the majority of observed variability in total transmitted energy is found at frequencies of no use to the sonar system, since these frequencies are filtered out in the recording process. The ratio of power standard deviation to mean power in this bandwidth is 0.0201, which is approximately one

frequency (KHz)
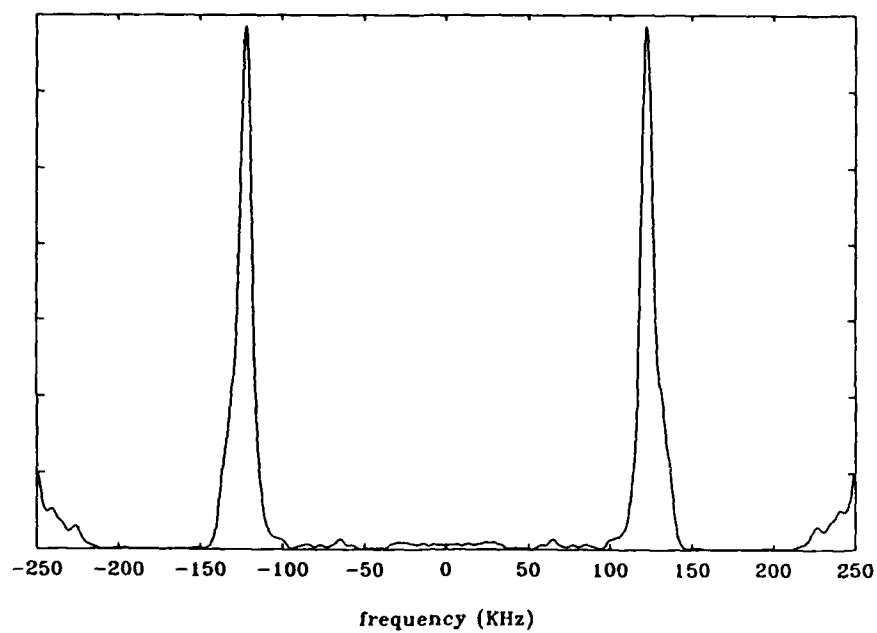
Figure 4-5: Mean transmit waveform power spectral density

frequency (KHz)
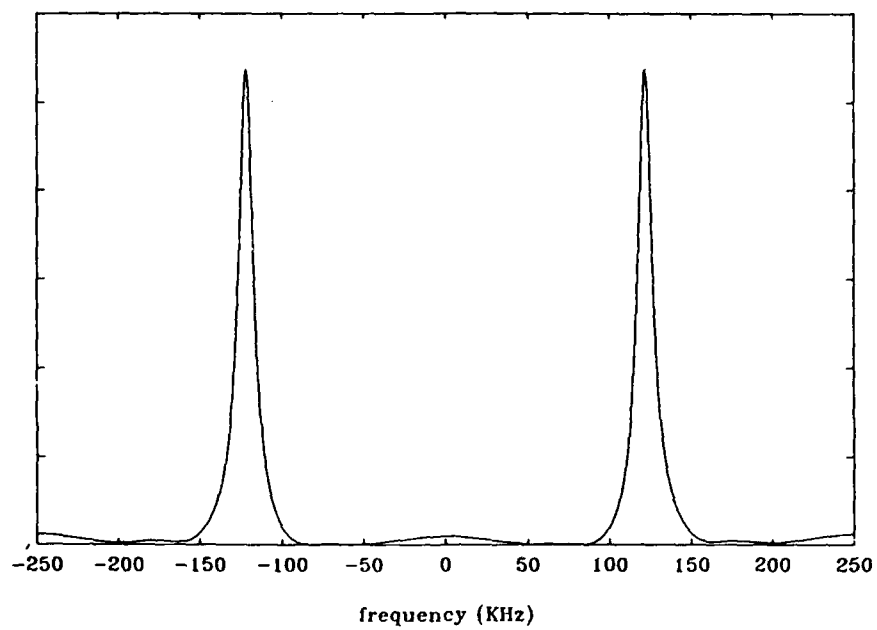
Figure 4-6: Power spectral density of fig. (4.2)

45
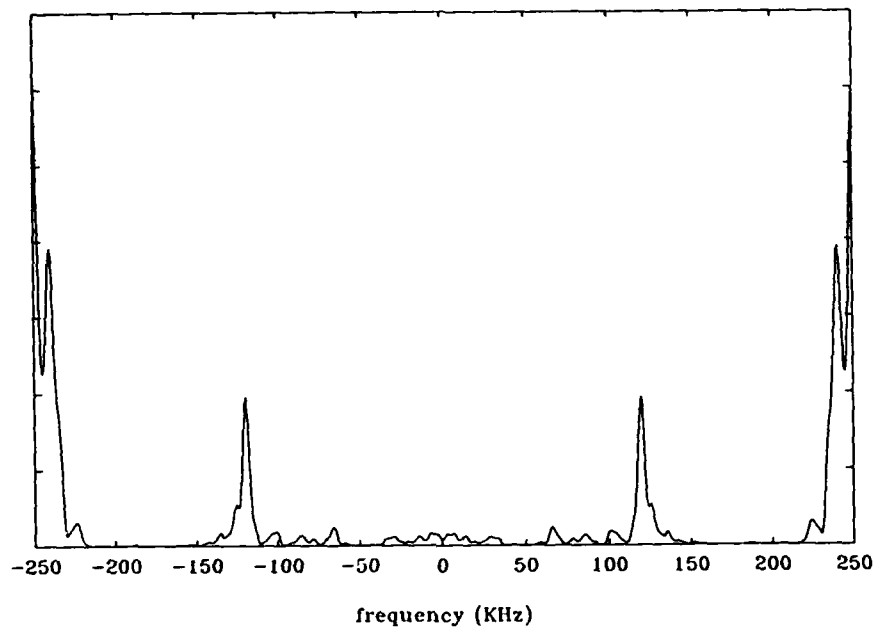
Figure 4-7: Variance of transmit waveform power spectral density

third the amount of variation found previously in the total energy distribution and which indicates that the variability within the system bandwidth is relatively small.

Having evaluated the intrinsic variability of the towfish transmitted waveform we can now account for this effect in the next chapter as we evaluate the variability contained in the images produced by this system. This will allow separation of transmission fluctuations, which are peculiar to the particular sonar system used to image an underwater scene and are therefore a controllable influence, from fluctuations caused solely by the acoustic environment and which will be present during any imaging process.

Chapter 5

Received Signal Analysis

In the previous chapter we quantified the intrinsic variability of the sonar and its transmissions. In this chapter we analyze the fluctuations which are present in the overall imaging process. Here we quantify the combined variability of the transmission, the reflection or scattering from the bottom, and the round trip through the medium. The results of the test tank experiment allow the cause of the overall system variability to be attributed either to the acoustic environment or to the sonar equipment itself.

A segment of the data from this experiment is shown in fig. (5.1). The data set consists of a single side scan sonar image originally composed of 3072 rows, each row containing 1024 samples and representing one image of the insonified bottom strip. The first three rows and rows 127, 2145, and 2683 were dropouts and were removed, yielding a 3066 row by 1024 column image matrix with the intensity level of each element $i(x, r)$ representing the acoustic pressure sensed by the transducer at that instant digitized to 8 bits. In this representation the image coordinates x, r correspond to transmission number and range bin, respectively. The nth column $i(x, n)$ of the matrix therefore represents a time series of the image pixel intensity or acoustic pressure amplitude from one independent, non-overlapping region of the bottom, while the nth row $i(n, r)$ corresponds to the same quantity for the nth transmission. In fig. (5.1) the columns, corresponding to individual bottom sample areas, are fairly consistent in intensity yet show visible fluctuation between individual rows, or points in time.

Figure 5-1: Segment of phase one pierside data
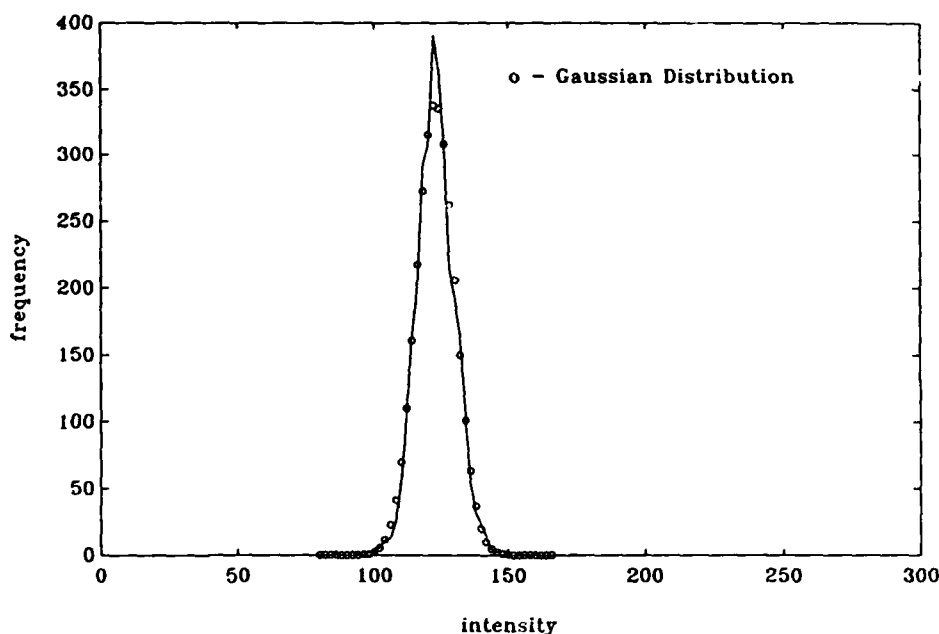
Figure 5-2: Histogram of pixel intensity for range bin 200, phase one pierside experiment

5.1 Probability Density Function

The probability density function of pixel intensity for a given column is of interest because it describes the echo stability of a single object or portion of the bottom. A representative estimated probability density function for this data set is shown in fig (5.2), a histogram of the 3066 image intensity values contained in column 200 corresponding to a range at which a strong return is received from the bottom. The distribution of intensity values appears to be approximately Gaussian when compared to the superimposed points outlining the Gaussian distribution having mean and variance equal to the sample mean and sample variance of the measured data. Compared to the Gaussian distribution the histogram contains more points at the center. For other columns in which it can be discerned the distribution is skewed slightly towards intensity values below the mean and is sufficiently unlike the Gaussian distribution that it fails the chi-square goodness of fit test [1] . This general shape is consistent regardless of column position, indicating that it is consistent for columns corresponding to volume reverberation, strong bottom backscatter, or weak bottom backscatter combined with
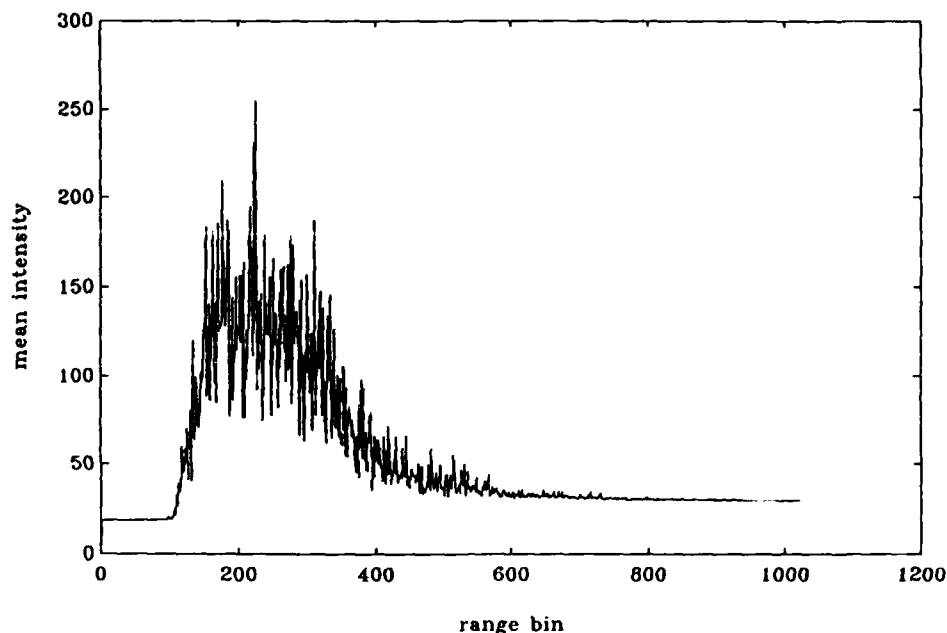
Figure 5-3: Pixel intensity mean value vs. range bin, phase one pierside experiment

volume reverberation. The skewing of the Gaussian distribution has been observed in other studies of acoustic backscatter [28] in which the Rician distribution is applied, providing for transformation of the Gaussian distribution to the Rayleigh distribution.

5.2 Distribution Moments

The mean value of each range bin n $\bar{i}(x, n)$ is shown in fig. (5.3), Empirically this can be divided into three regions. The region nearest the towfish, approximately the first 100 range bins, corresponds to ranges of 10 M or less and is generated by volume reverberation. This is because the towfish was mounted 10.5 M above the bottom. The region between range bins 100 and 600 contains the portion of the bottom with the most intense returns. The spikiness of this region is due to the differences in baskscatter strength between the various bottom subregions represented by $i(x, n)$. After column 600 the signal is greatly attenuated and increasingly noisy.

The overall shape of this curve is due to three primary influences, all functions of range.

Geometric spreading of acoustic pressure amplitude from a point source is described by

$$p(R) = p_o \left(\frac{R_o}{R} \right) \qquad (5.1)$$

Image intensity is directly proportional to pressure, which is inversely proportional to range. The reverberation limited active sonar equation for the monostatic case is [30]

$$SL - 2TL + TS - RL = DT \qquad (5.2)$$

Where SL is the sonar Source Level, TL is Transmission Loss, TS is target strength, RL is reverberation level, DT is Detection Threshold or echo signal-to-noise ratio, and all measured in dB with respect to a reference pressure and distance. Since the acoustic energy experiences a two-way trip to and from the target, twice the one-way transmission loss is deducted. In absolute terms this implies that for the side scan sonar case, acoustic pressure and image pixel intensity decay as $1/R^2$. This geometric spreading causes mean pixel intensity to decay monotonically with range. The second way in which range influences pixel intensity is the angular dependence of the bottom backscatter coefficient $m_b(\theta)$ discussed in chapter 2. Using trigonometry,

$$\theta = sin^{-1} \left(\frac{Z}{R} \right) \qquad (5.3)$$

Assuming a Lambertian surface and substituting into eqn. (2.12) yields

$$m_b = 10 \, log\mu + 10 \, log \left(\frac{Z}{R} \right)^2 \qquad (5.4)$$

showing that this effect also results in monotonically decreasing pixel intensity with increasing range. The third effect is the vertical acoustic beam pattern, which was described by eqn. (2.3). Attenuation due to the beam pattern is a function of vertical array angle α. Since the towfish array is generally oriented with its normal near horizontal, beam pattern attenuation increases with decreasing range. This effect is counter to the first two effects and results in the general trend of fig. (5.3), high attenuation at range extremes and minimum attenuation at an intermediate range.

The standard deviation of the data sequence in each range bin, $\sigma_{i(x,n)}$ versus bin number, is plotted in fig. (5.4). Its general form is nearly identical to that of fig. (5.3), mean intensity versus bin number. This indicates that the fluctuations in $i(x,n)$ could be a constant or

51
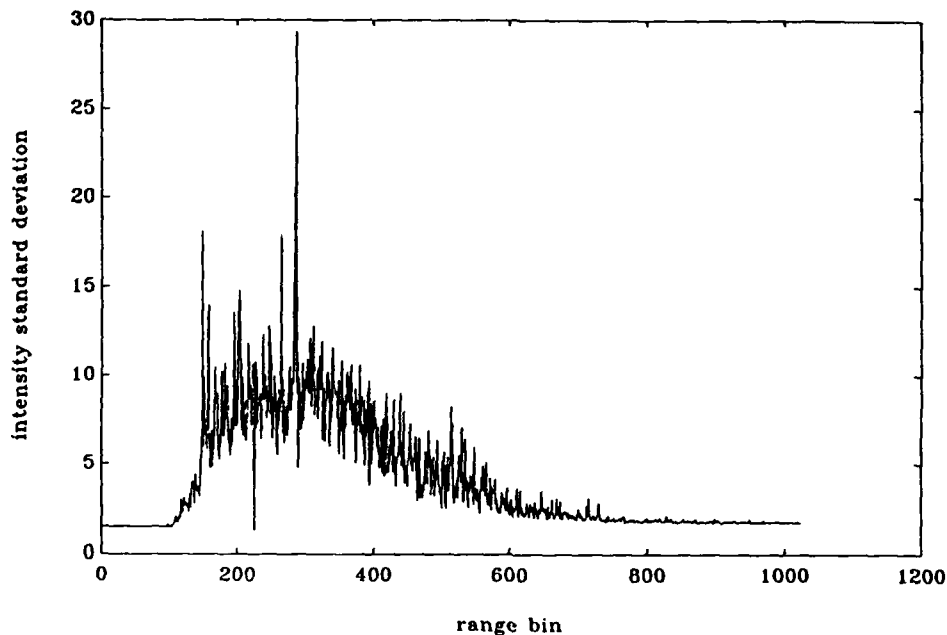
Figure 5-4: Pixel intensity standard deviation vs. range bin phase one pierside experiment

nearly constant fraction of the mean. This is seen to be approximately true in fig. (5.5) , a plot of the coefficient of variation V [31] versus bin number. The coefficient of variation is defined as the ratio of standard deviation to mean for the amplitudes of a series of acoustic transmissions. The mean value of V is approximately 0.08, with a peak centered around bin 400. The multiplicative nature of image intensity fluctuations allows for image intensity equalization in which the various regions of the image, generally regions associated with different ranges from the towfish, are scaled to yield a constant mean intensity throughout the image. Because fluctuation is approximately a fixed fraction of mean intensity, the equalized image shows nearly isotropic intensity variation statistics.

5.3   Range Bin Joint Statistics

To further evaluate the nature of image intensity fluctuations the joint statistics are evaluated for the 1024 time series $i(x, n)$. Of interest is the relationship between fluctuations in a given range bin and surrounding bins at the same point in time. Knowledge of this relationship
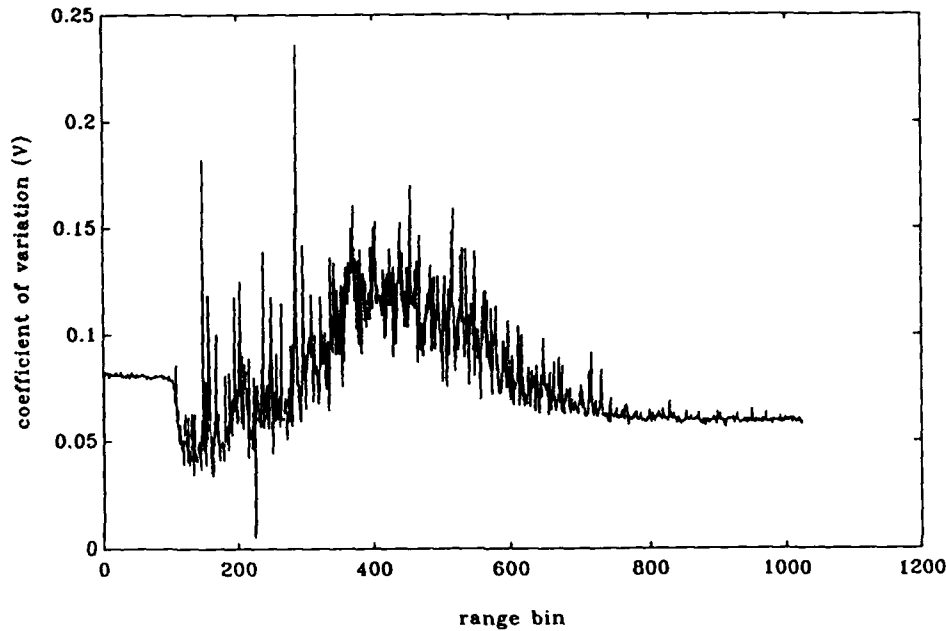
Figure 5-5: Pixel intensity Coefficient of Variation (V) vs. range bin, phase one pierside experiment

allows specification of the correlation length for the fluctuations across a region of pixels for individual acoustic transmissions. One technique for evaluating the correlation length consists of computing the correlation coefficient [25] between each $i(x, n)$ and all others. The correlation coefficient

$$C_{ab} = \frac{\sum_{j=1}^{3066}[i(j,a) - \overline{i(j,a)}][i(j,b) - \overline{i(j,b)}]}{\left(\sum_{j=1}^{3066}[i(j.a) - \overline{i(j,a)}] \sum_{j=1}^{3066}[i(j,b) - \overline{i(j,b)}]\right)^{1/2}} \qquad (5.5)$$

has a geometric interpretation as the inner product of two vectors. Treating the sequences $i(x, a)$ and $i(x, b)$ as vectors in 3066 dimension space, C_{ab} parameterizes their colinearity as the cosine of the angle between them. We use this as a measure of similarity between the various $i(x, n)$.

The C_{ab} make up a symmetric correlation matrix C whose j_{th} row or column holds the correlation coefficients for the jth column $i(x, j)$ versus all columns $i(x, n)$ of the data set. Fig. (5.6) is a typical row of C, showing $C_{300,b}$. Note that $C_{300,300} = 1$ while all other values fall between $+/- 0.4$. This is interpreted as meaning the intensity fluctuations of data set

53
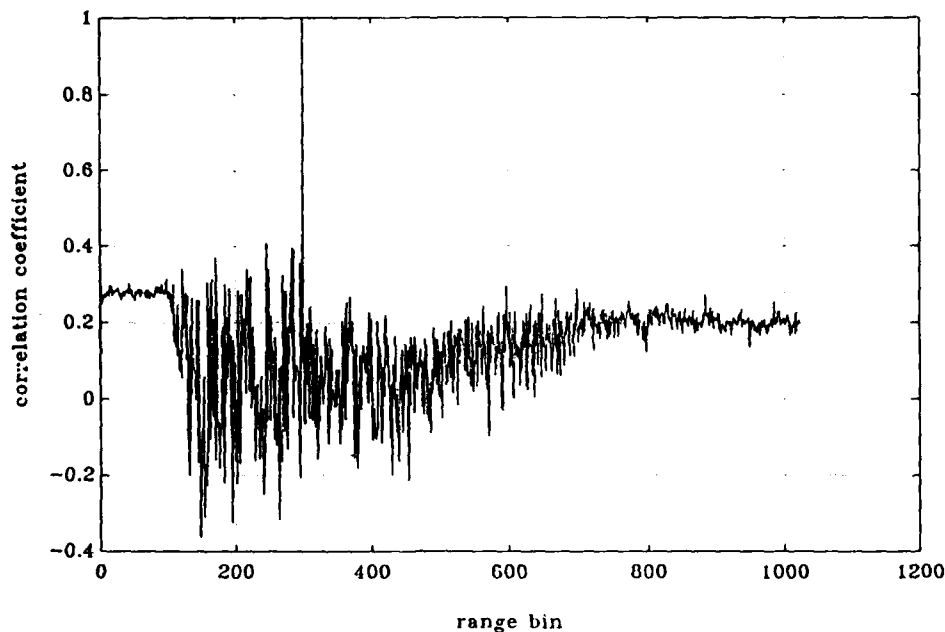
Figure 5-6: Correlation coefficients, range bin 300, phase one pierside experiment

range bin 300 are weakly correlated with the fluctuations in other bins. This weak correlation is observed across all range bins. Note also that there is no gradual drop in $C_{300,b}$ in the bins adjacent to bin 300. A gradual drop might suggest that the cause of intensity fluctuations were local inhomogeneities in the vicinity of the bottom segment imaged in column 300 which simultaneously affected the adjacent bottom segments.

5.4 Row Equalization

The acoustic transmission power fluctuations observed in the test tank experiment are a possible cause of the weak but wide range correlation of intensity fluctuations between the various range bins in the data set. Since scattered echo intensity is directly proportional to the intensity of the insonifying transmission and all range bins during one transmission cycle are insonified by the same acoustic transmission, pixel intensity variations due to transmission power fluctuations can be expected to be correlated. Total energy in each transmission was not measured during this experiment, but an estimate based on the total energy contained

54
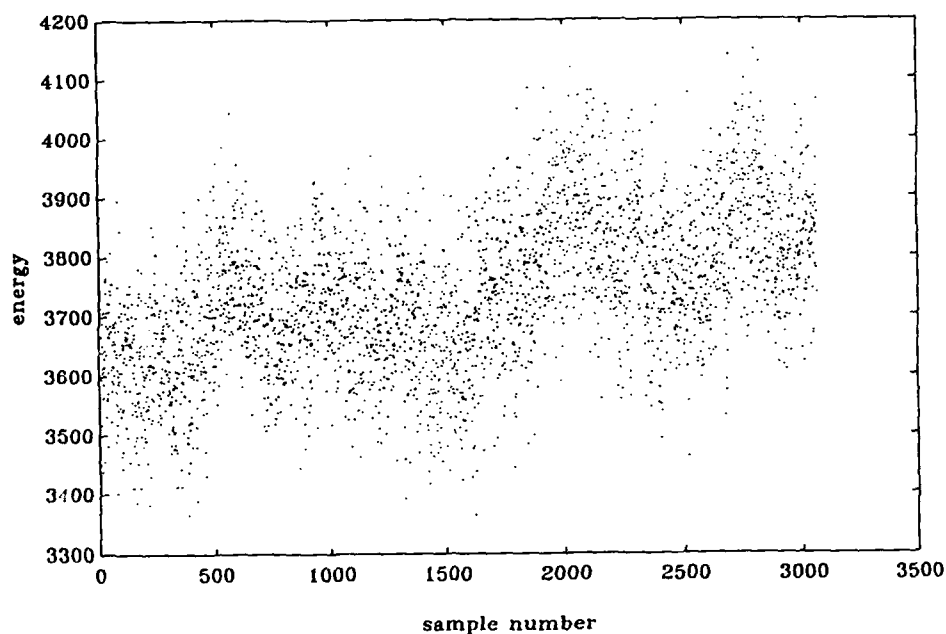
Figure 5-7: Total row energy vs. row number, phase one pierside experiment

in each row $i(n, r)$ provides satisfactory results. The total energy of each row is calculated as

$$E_n = \sum_{n=1}^{3066} i^2(n, y) \qquad (5.6)$$

Total row energy values are shown in fig. (5.7) in the same manner as they were presented for the test tank experiment. As in the test tank plot the distribution shows a low frequency undulation in mean value, with individual points scattered around the local mean. The statistics for this energy distribution are shown in table (5.1). The ratio of standard deviation to mean

mean:	3,740
standard deviation:	127
standard deviation/mean	0.0340

Table 5.1: Total row energy statistics, phase one pierside experiment

is comparable to the 0.0201 calculated in chapter 4 for transmission power fluctuations. The increased fluctuation is not unexpected considering the round-trip pathlength to the region of maximum intensity in this experiment was as much as 200 M as compared to 5 meters in

55
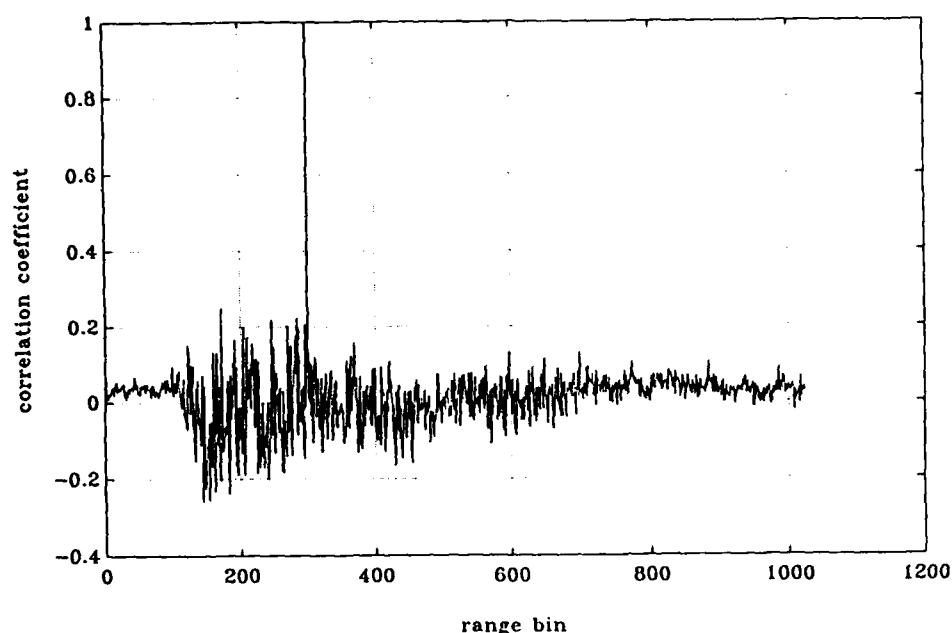
Figure 5-8: Correlation coefficients, range bin 300, phase one pierside experiment, compensated data

the test tank experiment. The additional interaction of the acoustic transmission with the medium is the probable cause.

Information about the transmission fluctuations can now be applied to the data set in order to equalize it and remove the effects of these fluctuations. The data matrix is scaled on a row by row basis, creating a new matrix with elements $i'(x, r)$ such that

$$\sum_{r=1}^{1024} i'(n, r) = \sum_{x=1}^{3066} \sum_{r=1}^{1024} i(x, r) \tag{5.7}$$

for any row n. After removing linewise image intensity fluctuations the previous analyses were performed again. Fig (5.8) shows the effect of this compensation on $C_{300,b}$. Compared to the uncorrected case the degree of correlation between column 300 and other columns is significantly reduced, indicating that transmission fluctuations are a probable cause of the weak correlation of pixel intensity fluctuations for a given transmission or side scan sonar image line. This lack of correlation leads to the conclusion that the intensity fluctuations observed at various ranges in a side scan sonar image are essentially independent, in the

56
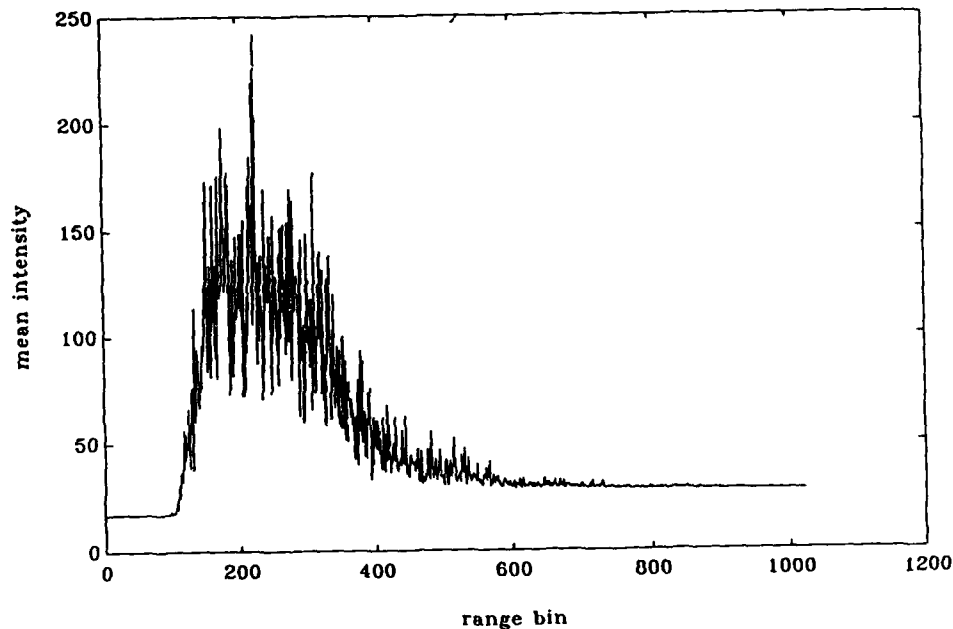
Figure 5-9: Pixel intensity mean value vs. range bin, phase one pierside experiment, compensated data

absence of change in the imaged topography.

Range bin mean, standard deviation, and coefficient of variation analyses were repeated on the row equalized matrix and are shown in figs. (5.9 -5.11). The figures show that the equalization process did not significantly change these parameters, indicating that the temporal fluctuations in pixel intensity for a side scan sonar image are generally independent of transmission energy fluctuations and cannot be attributed to them.

5.5 Power Spectral Density of Temporal Fluctuations

One final analysis of the phase 1 pierside data was the power spectral density (PSD) measurement of the fluctuations of the column intensity sequences $i(x, n)$. The fluctuation part of the sequence was generated by subtracting the mean for that range bin from each element of the sequence. Fig (5.12) is the PSD for column 400, which was smoothed by an 8 bin moving average and is a representative PSD of this data set. The PSD shows a narrow frequency spike centered around DC with a bandwidth of approximately 2 cycles within the
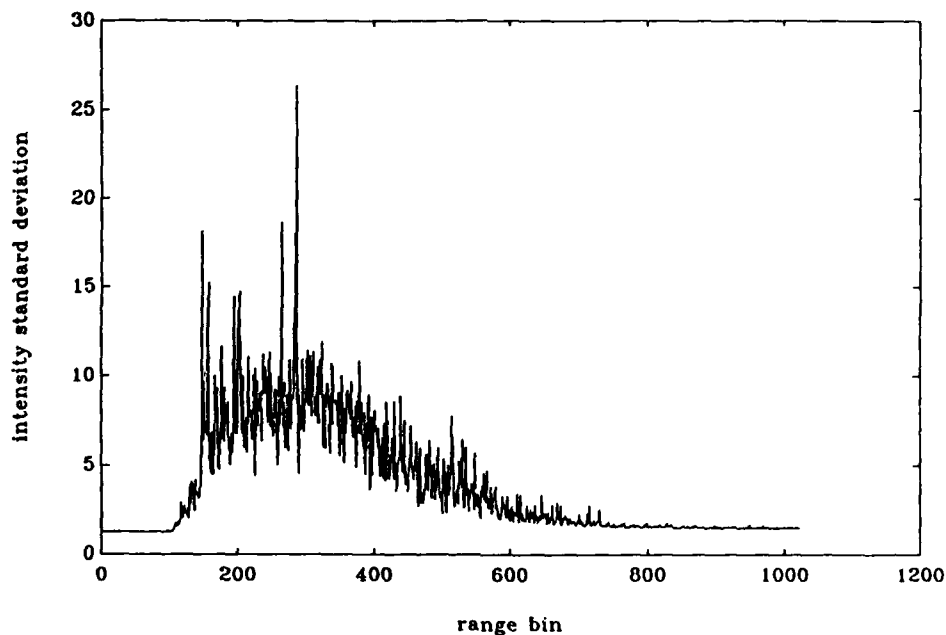
Figure 5-10: Pixel intensity standard deviation vs. range bin, phase one pierside experiment, compensated data

length of the data set with an essentially flat spectrum elsewhere. Referring to fig (5.6) this is consistent with the observed frequency of the low frequency undulations, indicating that transmission power fluctuations are again the probable cause. Since the temporal length of the data set was 613 seconds and consisted of the equivalent of 6 side scan sonar images, it may be assumed that within a single side scan sonar image the spectrum of the temporal intensity fluctuations nearly white and therefore that intensity fluctuations are temporally uncorrelated as well as spatially uncorrelated as discussed previously.
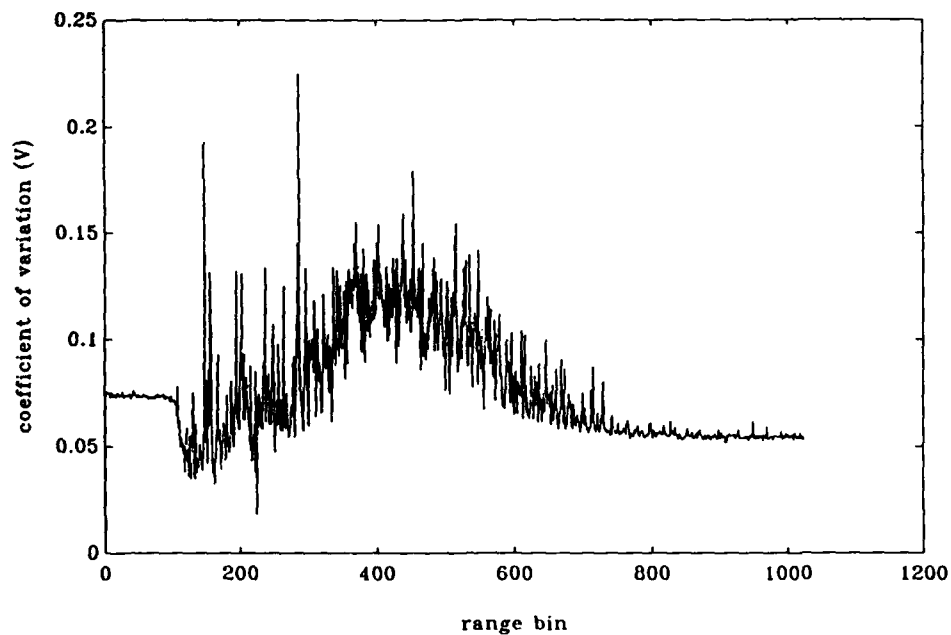
Figure 5-11: Pixel intensity Coefficient of Variation (V) vs. range bin, phase one pierside experiment, compensated data
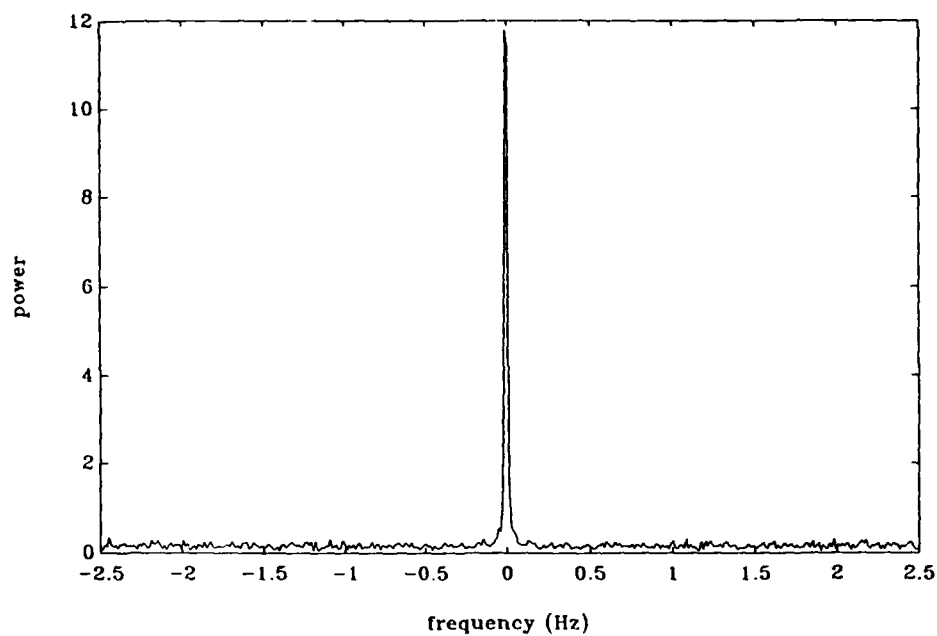
Figure 5-12: Power spectral density, range bin 400, phase one pierside experiment

59

Chapter 6

Image Rectification and Registration

In the previous two chapters we investigated the variability of the side scan sonar imaging process and determined that for the overall process the standard deviation of the intensity of an individual pixel is approximately 8% of its mean value. Additionally, we showed that for pixels representing non-overlapping segments of the bottom and consistent acoustic transmissions the fluctuations are statistically independent. Pixel intensity therefore appears stable enough to attempt matching subregions in two images of the same bottom area.

In this section we consider development of an algorithm to register two side scan sonar images. We proceed using the side scan figs. (6.1) and (6.2), two images taken during the phase two pierside experiment at a depth of 9 meters with level attitude. The same bottom is shown in both images, predominantly a silty sand with numerous natural and man-made features. The image begins with the leftmost white stripe which marks the beginning of the transmission and range equal to zero. Range increases to the right, the next vertical white stripe is the surface return, annotated "S" which shows that the towfish was approximately 6 M below the surface. Fig (6.1) also shows a school of fish at this range, marked by the brilliant return from their swim bladders. The next stripe is the first bottom return, which is seen decreasing in range as the image progresses since the image was scanned row by row from top to bottom. The bottom in this region does indeed rise due to the fact that the
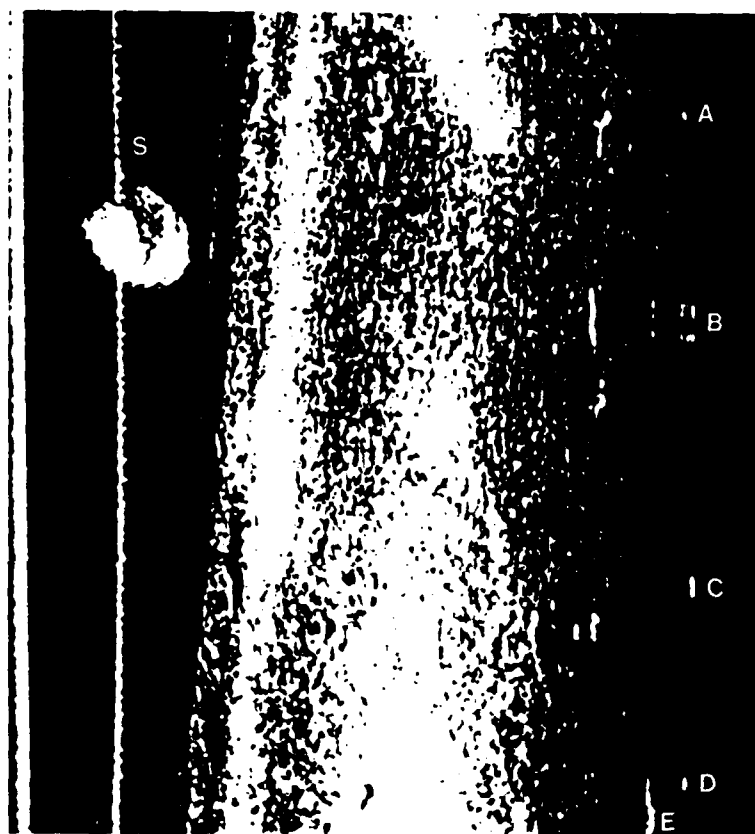
Figure 6-1: Side scan sonogram, 9 M depth, phase two pierside
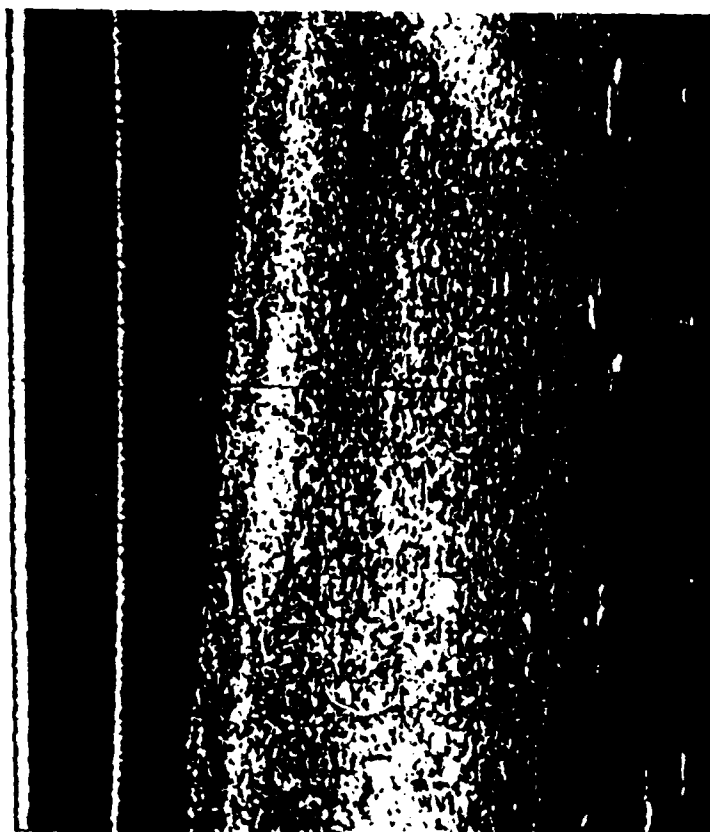
Figure 6-2: Side scan sonogram, 9 M depth, phase two pierside

track was oriented on the pier so that the towfish headed towards the shore as it was moved down the track. The discrete objects near the first bottom return are large rocks and pieces of debris approximately 1 M size. The four equally spaced objects near the right side of fig. (6.1) are the four drums described in chapter 2, with "A", "B", "C", and "D" corresponding to the 5, 10, 40, and 55 gallon drums respectively. The large linear object marked "E" is the corner reflector and its sidelobes, while the large linear object to the left of the second drum is a pre-existing bottom feature of unknown origin. The dimensions of the area in the image are 50 M in the horizontal direction and 30 M in the vertical

6.1   Preliminary Image Processing

Detection of temporal change in a bottom scene requires that at least two images be rectified and registered so that all points in the images are aligned for comparison. Preparing the images for comparison involves transformation of image coordinates and image intensities. Unless all images being compared were taken by a towfish with identical trajectory and attitude parameters, differences in these parameters produce aspect-dependent effects in the imagery obtained. Because these aspect dependent effects are due to imaging geometry and not the bottom scene itself they represent artifacts in the comparison process. An attempt should therefore be made to remove as much noise as possible in order to improve the comparison process and to separate image differences caused by differing imaging geometry from those caused by real changes in the objects in the scene.

6.1.1   Slant Range Correction

Slant range correction is illustrated in fig. (6.3). Towfish altitude Z is determined from the image by N, the column in which the first bottom return occurs. The first bottom return is distinct in 500 KHz imagery because of vertical sidelobes which insonify the bottom directly below the towfish. In this case, which provides a general example for side scan imagery, the bottom is not flat and level. N is variable along the track, which requires that image remapping of r to y coordinates use a different value of N for each image line. The number of lines involved require that this process be automated. Neglecting the effects of variable Z

TOWFISH

$$Y = \sqrt{R^2 - Z^2}$$
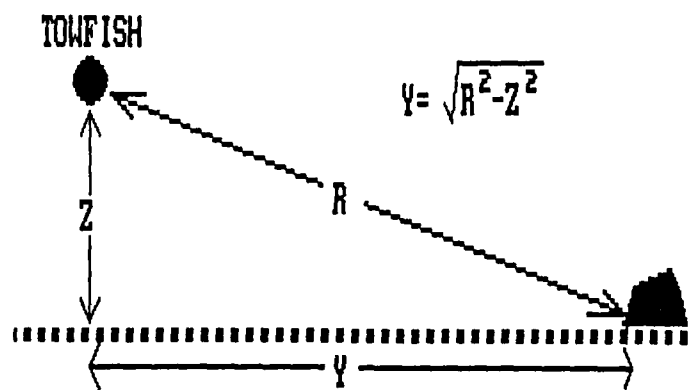
Figure 6-3: Slant range calculation

would result in remapping errors and subsequent image distortion. It should be noted that slant range correction is sensitive only to depth gradients in the direction of towfish travel. A bottom with gradient perpendicular to the direction of travel results in a constant towfish altitude.

Detection of the first bottom return is performed as a hypothesis test based on the distributions of pixel intensities in the water column adjacent to the first bottom return and in the first bottom return. The region of both images designated for testing was the entire region to the left of the surface return and fish school. These two features represented regions of gross noise in the water column and were subsequently removed from the decision process. Limited human intervention of this kind is valuable in preventing decision errors. An illustration of the hypothesis test is shown in fig. (6.4). The hypothesis test consists of choosing a threshold intensity such that pixels with intensities above this threshold are designated as "possible bottom" pixels while pixels below this threshold are not considered. The pixel of the image row under examination that is ultimately designated the first bottom return is the one for which the following condition is first satisfied. For the contiguous group of four pixels formed by the pixel under consideration and the pixels in the next three columns, at least three of the pixel intensities must be greater than threshold. This decision process prevents bright but inconsistent pixels in the water column from causing a premature decision that the bottom has been encountered, while not preventing the designation of first bottom return

64
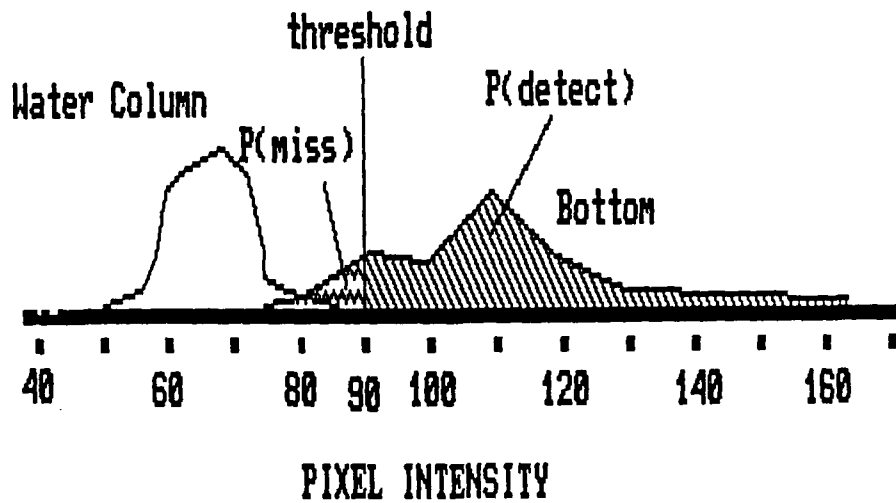
Figure 6-4: First bottom return hypothesis test

in the event of a fortuitously low or dropped out pixel.

The probabilistic description of this process is straightforward. In this specific example the threshold was set at an intensity value of 90. For individual pixels the Probability of Detection $P_p(D)$, defined as

$$P_p(D) \equiv Prob(\text{Bottom region pixel exceeds threshold}) \tag{6.1}$$

was measured to be approximately 0.91. The Probability of False Alarm $P_p(FA)$ defined as

$$P_p(FA) \equiv Prob(\text{Water column pixel exceeds threshold}) \tag{6.2}$$

was measured as less than 0.01. Basing the decision on at least 3 detections in 4 pixels is described by the binomial distribution [9]

$$P(D) = \begin{pmatrix} 4 \\ 0 \end{pmatrix} P_p(D)^4 (1 - P_p(D))^0 + \begin{pmatrix} 4 \\ 1 \end{pmatrix} P_p(D)^3 (1 - P_p(D))^1 \tag{6.3}$$

Where $P(D)$ is the Probability of Detection of the overall decision process or the probability of the overall decision process detecting the bottom when the region under consideration is the bottom. $P(D)$ was found to be 0.99. The Probability of a Miss, the probability that

65

the bottom is not detected when the region under consideration is in the bottom region was calculated to be 0.01.

Using the assumption that towfish altitude is a continuous function with a limited number of discontinuities, it is desirable to smooth discontinuities which may occur in the bottom profile sensed by this algorithm while preserving rapid changes in N. Discontinuities which would be desirable to smooth are the result of erroneous bottom detection as described above. The remapping of each row is based on the sensing of the first bottom return in that row. If the bottom profile as sensed by the algorithm is not a smooth function, linear features parallel to and in the vicinity of the first bottom return will appear jittered as different segment are mapped to various distances from the vertical edge of the image. It is desirable however to preserve rapid changes in Z. Such rapid changes could occur because of abrupt changes in bottom depth or because of rapid towfish depth changes. To accomplish this smoothing while preserving rapid changes in Z a non-linear filter known as a median filter is employed. The particular median filter used in this instance consists of a 9 point window. When centered on $N(n)$, the value of N calculated for the nth image row, the filtered value of N is [15]

$$N_{filt}(n) = Median[N(n+4), N(n+3), \ldots N \ldots, N(n-3), N(n-4)] \qquad (6.4)$$

The median filter, because it selects the median value in its window as its output value, is useful in removing spikes in the input data stream. A linear filter such as a moving average filter would attenuate the spike but still reflect its effect in the filter output. The median filter is also useful in edges, which would generally become smeared by a linear filter. It is therefore well suited for this application.

The result of the bottom sensing algorithm is shown in fig. (6.5). A white line has been plotted on fig. (6.1) where the first bottom return was sensed. Note that the line is smooth and is only affected by interference in the vicinity of the fish school.

6.1.2 Intensity Equalization

Having corrected slant range distortion, the next aspect dependent feature to be processed out of the images is intensity variation. As discussed in chapter 5 the three aspect dependent intensity effects are geometric spreading, vertical beam pattern, and angular dependence of

Figure 6-5: Fig (6.1) with bottom profile as detected by hypothesis test

bottom backscatter coefficient. The combined effect of these can be seen on fig. (6.1). The vertical beam pattern can be seen as at least 4 parallel bands of high intensity in the near to mid ranges of the image. The other two effects are evidenced by the heavy attenuation at large ranges.

Since all three effects are range dependent it is reasonable to approach this problem in a columnwise fashion. The desired result of the intensity equalization process is to determine the local bottom backscatter intensity level and to scale all subregions of the image to have the same level. In this process it is desirable to base scaling on only the background intensity and to not include individual objects in the calculation. Individual objects are normally more intense that the background while their shadows are less intense. Improper scaling to improve uniformity of image intensity may cause the brighter intensity of individual bottom features to be scaled to the desired background mean intensity, resulting in a loss of image definition.

A preliminary attempt at intensity equalization was to scale all pixels by column such that the mean intensity of each column of the resulting image had the same mean value. This proved unsatisfactory since of the three aspect dependent intensity variations, only geometric spreading is truly range dependent. The other two are grazing angle dependent, so for varying Z the point Y at which a given effect occurs on the bottom is determined by

$$Y = Z cot\theta \qquad (6.5)$$

As a result the angular effects shifted across image columns with varying Z, creating diagonal bands of constant intensity similar to those seen on fig. (6.1). The vertical extent of these bands was less than that of the window over which the mean was computed, meaning that all columns had the same mean value yet still contained intensity fluctuations. This compensation approach was therefore moderately successful only for large ranges where intensity contours are vertical.

A modification of the above approach was to shorten the region over which the mean intensity was computed from an entire column to a fraction of a column. For each pixel the mean intensity was computed for a window which extended over 1 column and 50 rows and was centered on the pixel. All image pixels were scaled such that

$$i'(x,y) = i(x,y) \left(\frac{\text{mean image intensity}}{\text{mean window intensity}} \right) \qquad (6.6)$$

By shortening the vertical extent it was hoped that the vertical dimension of the intensity fluctuations could be made larger than the window and they could be effectively removed. This was partially achieved, however the reduction in size of the window also made it more likely that the presence of small but intense regions in the window adversely affected the computation of mean window intensity. Such regions or objects caused the mean window intensity to be greater than the actual background intensity in the windowed region, resulting in greater than required attenuation. The increased intensity attenuation caused the intense object included in the window to be less distinct, as described above.

The most successful approach was to use a 51 point median filter of width 1 pixel and extending 25 rows above and below each pixel location. The limited vertical extent of the window made its sample region smaller than the vertical extent of aspect dependent intensity variations in the image, while its median filter characteristics allowed computation of the local mean intensity which was less likely to be influenced by small, high intensity regions. The scaled intensity of each pixel was then determined as

$$i'(x,y) = i(x,y) \left(\frac{\text{mean image intensity}}{\text{median window intensity}} \right) \qquad (6.7)$$

This results in the intensity equalized and slant range corrected image shown in fig. (6.6), which is fig. (6.1) after processing. This is the bottom as it would appear if viewed from above with the water removed. Having corrected these two sources of distortion we proceed with rectification and registration.

6.2 Rectification and Registration

6.2.1 Correlation

Rectification of two side scan sonar images a and b involves locating in image b by coordinates (x_b, y_b) the point in image a imaged at coordinates (x_a, y_a). As discussed earlier this has been done using navigation information recorded during the survey to tag image data which is later mapped into geographic coordinates. In this example positional information is derived from
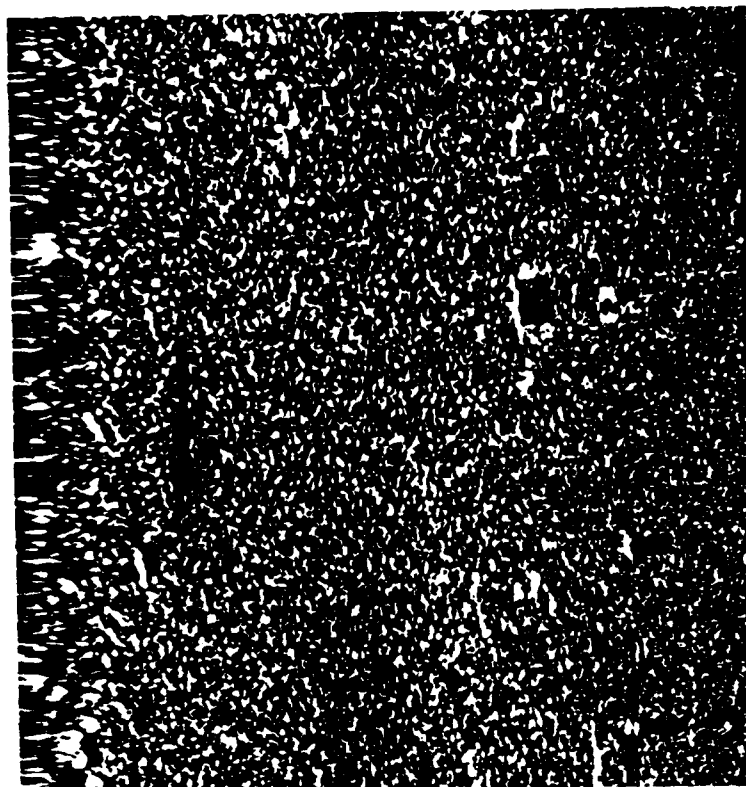
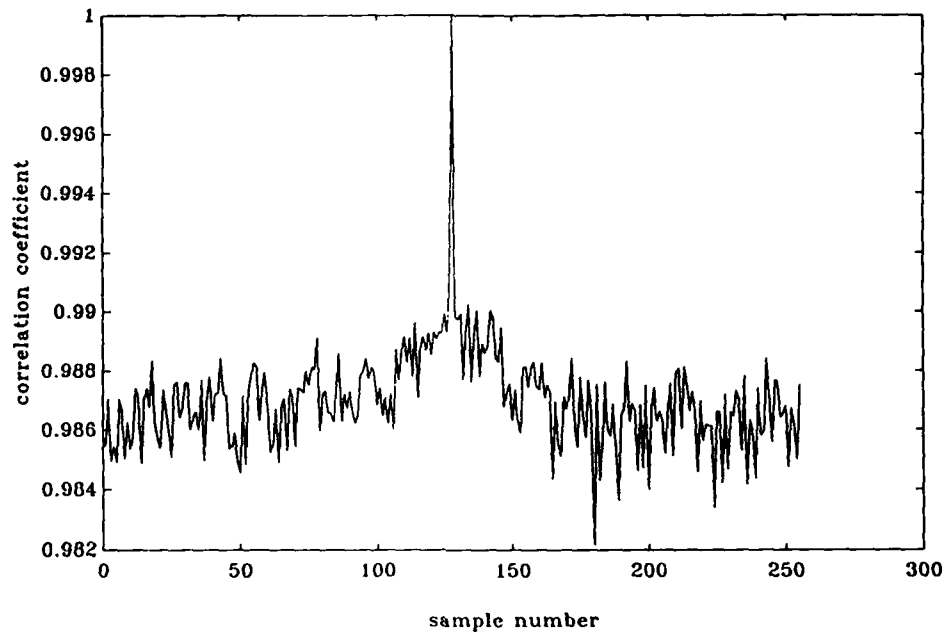Figure 6-6: Fig (6.1) after processing to remove aspect dependence

Figure 6-7: Results of correlation matching test on phase one pierside data

similarity of image subregions between the two images. The basis for comparison is the correlation coefficient, which for two subregions of equal dimension is calculated as

$$C_{ab} = \frac{\sigma_{x,y}^2}{\sigma_x \sigma_y} \tag{6.8}$$

where $\sigma_{x,y}^2$ is the pixel intensity covariance between the two image subregions and σ_x and σy are the pixel intensity standard deviations in the subregions of **a** and **b** respectively. Using Cab as a measure of similarity as in chapter 5, the best match for a windowed region in **a** is the region in **b** which Cab reaches its global maximum.

A 256 line segment of phase one pierside data was used to test the viability of this method. This data features repeated images of the same bottom segment , so the values obtained for C_{cb} in this test provide the maximum performance that can be expected of this method. In this test the correlation window consists of row 128 of the data set or a correlation window of 1 row by 1024 columns. This window was correlated against each line in the data set. The resulting plot of C_{ab} is shown in Fig. (6.7). Since in this test the line used as the template is a copy of line 128 of the data set, perfect correlation occurs at this point and $C_{ab}=1$. Note that in general $C_{ab} > 0.985$ for this ideal case.
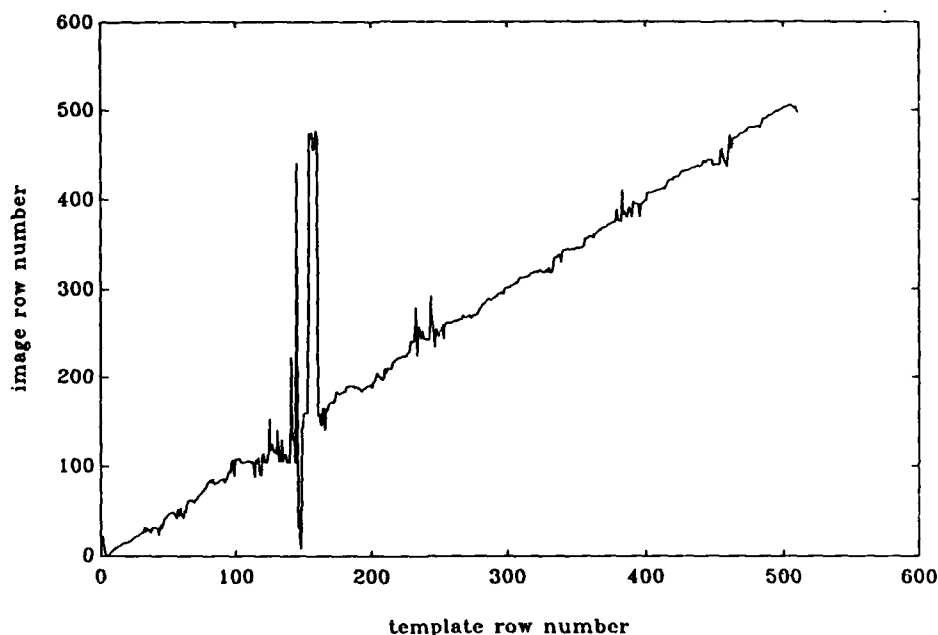
71

Figure 6-8: Correlation matching function for 1 row windows

Employing the correlation matching procedure to figs. (6.1) and (6.2) is straightforward. The uncorrected image case is presented first to illustrate the value of preliminary processing in correlation matching. Because both images were taken by the tracked towfish it can be assumed that the towfish bottom track was identical for both images with the possible exception of speed differences along the track. It is also reasonable to assume that attitude instabilities were not excessive. Because of these two assumptions it is assumed that the strips of bottom insonified by each transmission form a set of parallel lines perpendicular to the pier whose order is preserved by the uni-directional motion of the towfish. These assumptions result in the choice of individual image rows or groups of rows for the correlation window. Searching for the maximum C_{ab} is required only in the along track direction since across track motion of the towfish is eliminated by the track.

As a first approach Fig. (6.1) was designated image **a** and was used as the template. Fig. (6.2) was designated image **b**. Each of the 512 lines contained in **a** were correlated against all lines in **b** and the best match for each determined. Figure (6.8) is a plot of fig **a** line number and corresponding best matched line number in **b**. There is an approximate one-to-one correspondence between row numbers in the two regions, which would be represented by
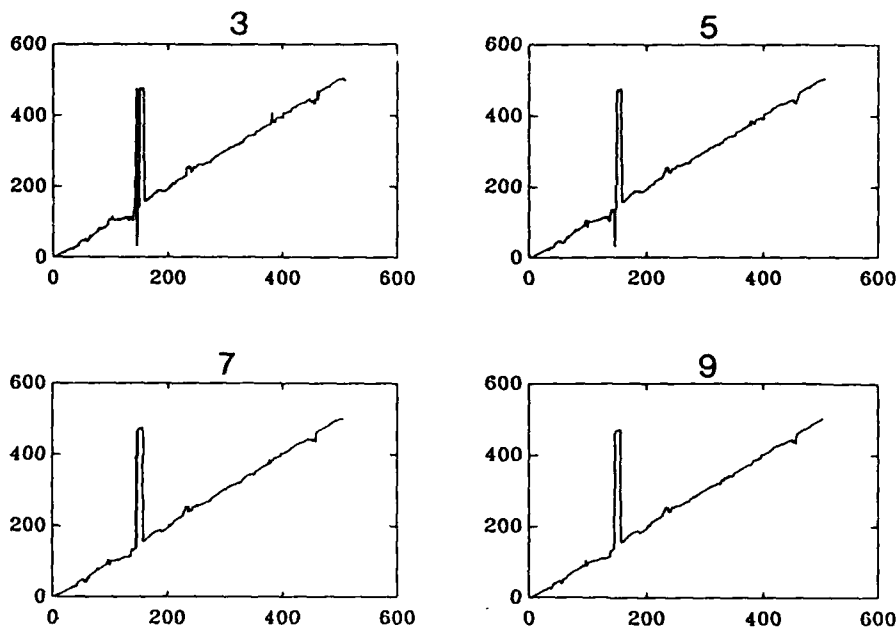
Figure 6-9: Correlation matching function for 3,5,7, and 9 row windows

a diagonal line across the graph. However there are also numerous departures from the ideal case which indicate matching errors on the order of tens or hundreds of lines.

Matching errors were found to be related to the size of the correlation window. Various window sizes, from 1 row by 512 columns to 9 rows by 512 columns, were tried with the resulting matching functions shown in fig. (6.9). Increasing the size of the correlation window has both positive and negative effects on the correlation mapping process. The positive effect is that it increases the size of the data base upon which the decision is made. The increased size not only attenuates the effects of bad data which may be included in the window by diluting it with a larger data base, it also increases the number of degrees of freedom of the decision process. Increasing the size of the template increases the number of possible combinations of pixel intensities representable in the window, which enhances the uniqueness of the region in the window and subsequently decreases the ambiguity in deciding which of several similar image regions best match the template. The deleterious effect of increasing window size is that in the presence of two dimensional geometric distortion or
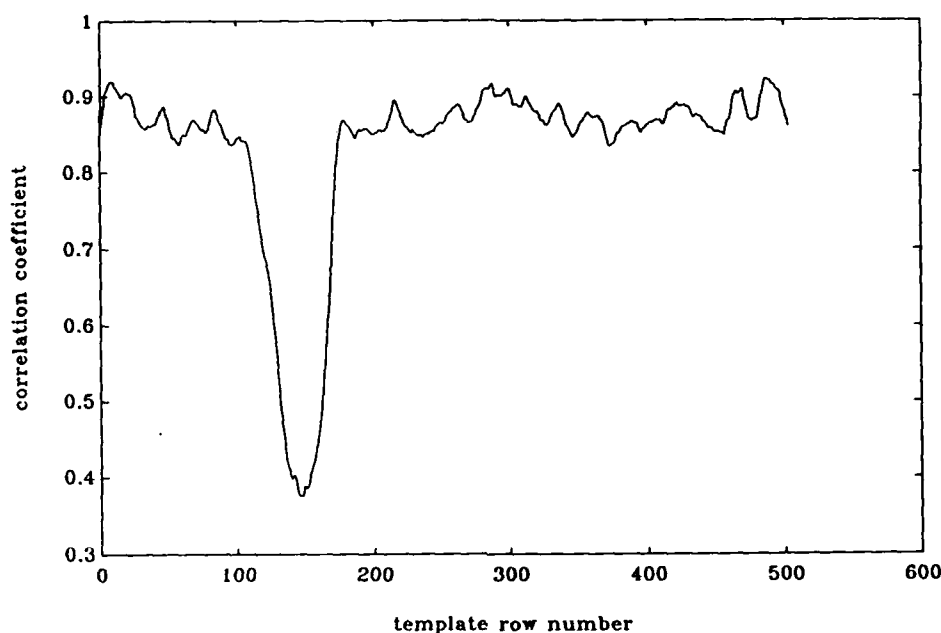
Figure 6-10: Correlation coefficient, 9 row window

warping increasing window size results in an increased probability of pixel misregistration within the window. Warping is present in all side scan imagery due to trajectory and attitude instability, resulting in a non-linear map of pixels in one image to those in another. It may be possible to center correlation windows around matching pixels in the two images, however for pixels on the periphery of the windows the probability of non-coincidence increases with increasing warp. In this study it was found that improved performance was obtained by increasing window size, but that marginal benefits were achieved by increasing the window size beyond 9 rows.

A plot of C_{ab} versus line number is shown in fig. (6.10). Note that C_{ab} is generally > 0.8 with the exception of the region corresponding to the fish school in **a**. Since no region in **b** is similar, the best matches made for this region were poor and C_{ab} was low. Both regions of this graph suggest reasons that the preliminary image processing steps described in section 6.1 are useful. The fish school is a source of noise in a portion of the image containing no information on bottom features. Removing the water column portion of the image removes a portion of the image which can provide no information about the bottom and can only introduce noise.
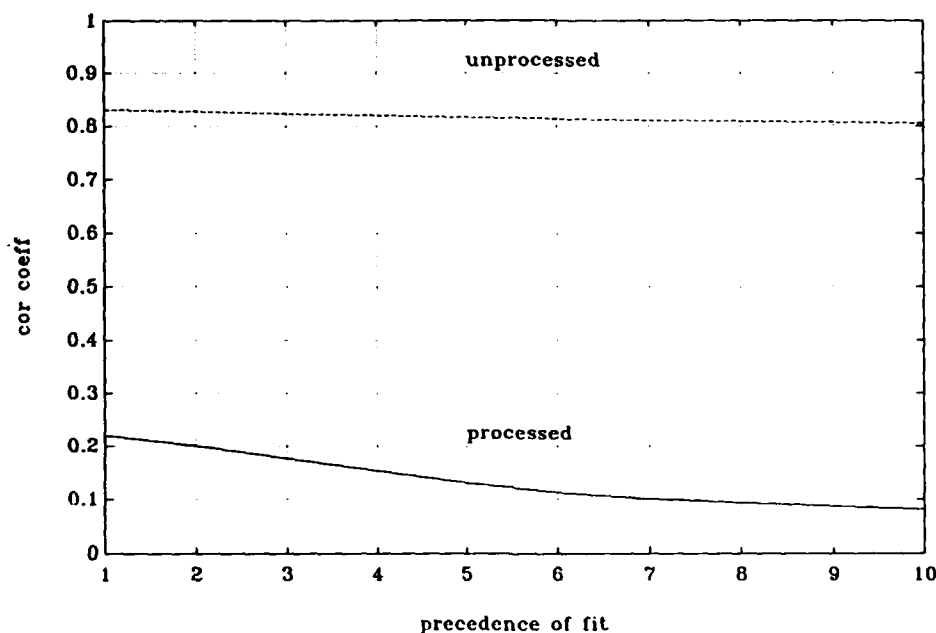
74

Figure 6-11: Effect of preliminary processing on image decorrelation

Since only noise can result in the inclusion of the water column, it is prudent to remove it. The second reason for preliminary processing is the type of information that predominates the decision process. By examining figs. (6.1) and (6.2) it can be seen that all rows in both images are very similar. All rows start with a dark region which corresponds to the water column, proceed through an intense region with intensity fluctuations, and end with heavily attenuated dark regions. This aspect dependent pattern dominates the intensity function of every line, and therefore dominates the correlation matching process. Because the correlation process is dominated by this information, it is more likely to make a decision based on towfish altitude Z than on bottom features.

In both a and b towfish altitude is a monotonic function of X. Had it not been and a particular value of Z corresponded to a non-unique value of X it is probable that the obtained matching function would not have been as near to the ideal case. Additionally, dominance of the decision process by aspect dependent information suggests that varying imaging geometry could make correlation matching of two images difficult or impossible. Figure (6.11) illustrates one effect of removing aspect dependent effects from figs. (6.1) and

75
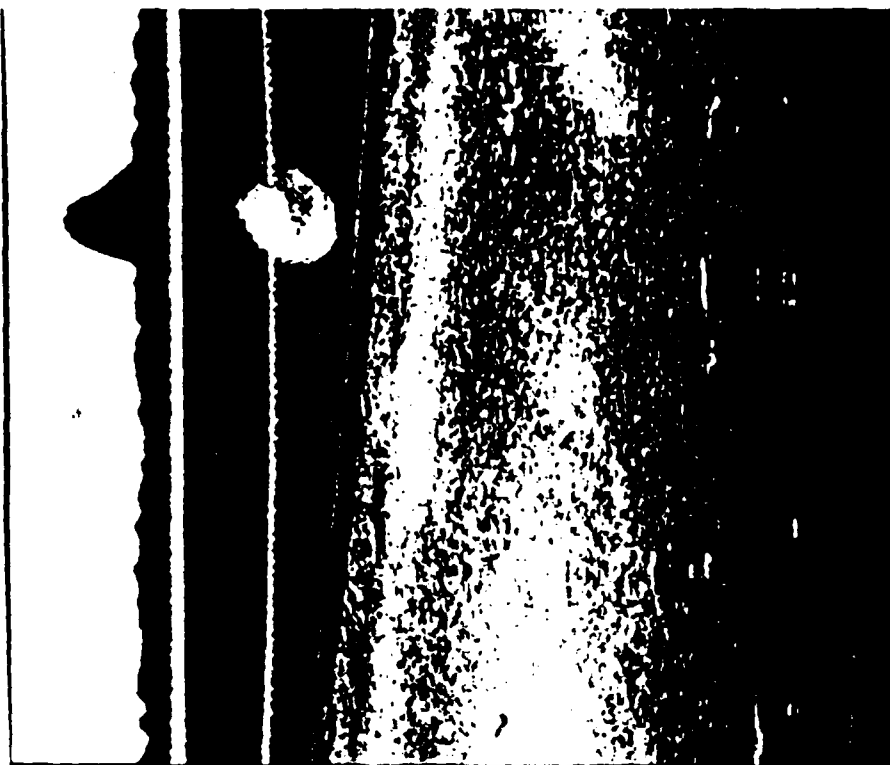
Figure 6-12: Fig. (6.1) and correlation coefficients by row

(6.2). It is a plot of mean C_{ab} versus precedence of fit, for the ten best matches between a typical 9 row from fig. (6.1) and all such windows from fig. (6.2) in both the corrected and uncorrected cases. It shows that for the matching of a given row from fig (6.1) the expected value of C_{ab} for the row in fig. (6.2) which provides the best match is approximately 0.82. The expected value of C_{ab} for the tenth best match of a row from fig. (6.2) with the same row is 0.80. The lack of decorrelation with decreasing precedence of fit suggests that there is little to distinguish the correct choice from several nearly correct choices. When compared to this curve the curve for the corrected case shows more rapid decorrelation and a more distinct best choice.

To further illustrate the effect of correction for aspect dependence fig. (6.12) shows fig.
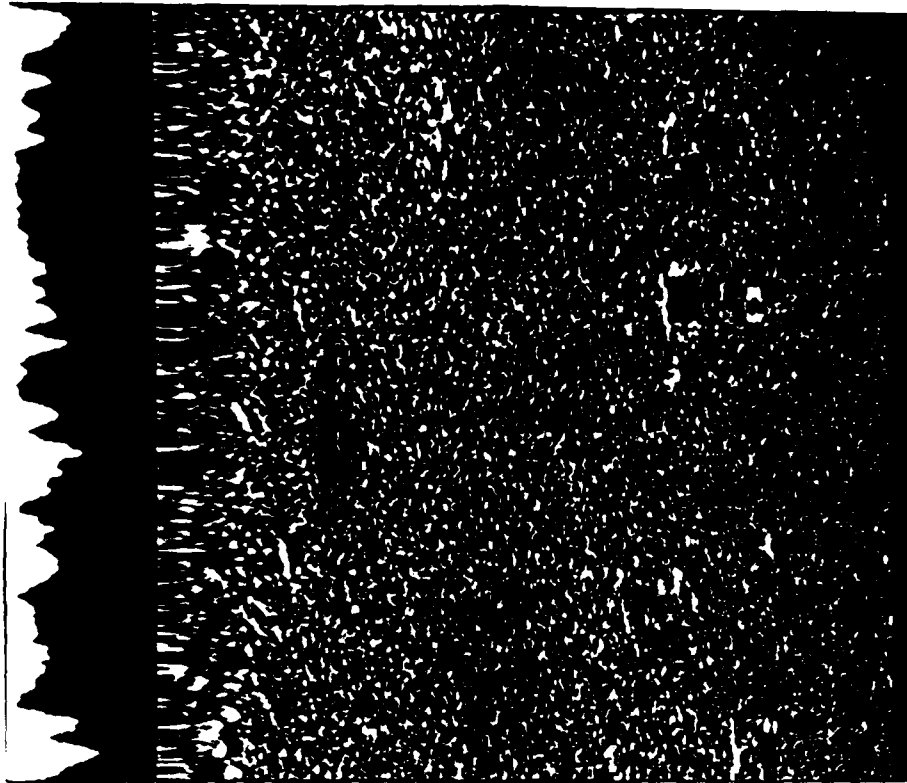
Figure 6-13: Corrected fig. (6.1) and correlation coefficients by row

(6.1) with the a graph of C_{ab} in the left margin. The width of the white line in the graph corresponds to the value of C_{ab} obtained when comparing the 9 row wide correlation window centered at the adjacent image row with all possible 9 line windows in fig (6.2). Correlation is good for all regions except the fish school. Good correlation in regions of little bottom detail further suggests that bottom detail is not the criterion upon which the decision of best fit is made.Figure (6.13) is the same type of plot for the corrected case. Note that although the values of C_{ab} are not as high as in the uncorrected case, regions of high correlation correspond with regions of greater detail.

Figure 6-14: Correlation coefficients for corrected figs. (6.1) and (6.2)

6.2.2 Rectification

Correlation matching of slant range corrected and median filter intensity equalized versions of figs. (6.1) and (6.2) was performed using a 9 row correlation window. The processed fig. (6.1) is designated the template image **a** while the processed version of fig (6.2) is designated the sample image **b**. We 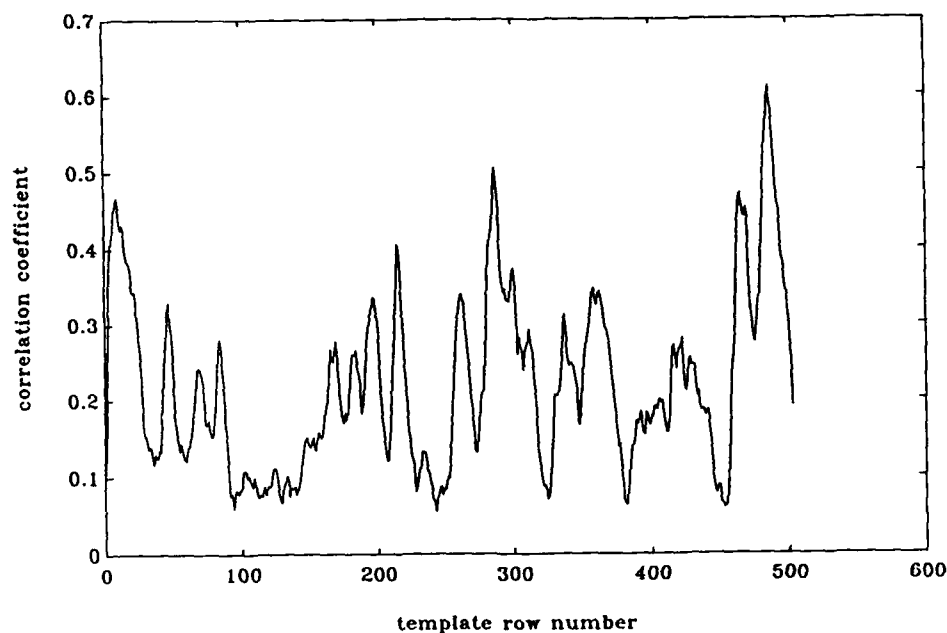now proceed to map the features of **b** to the locations of the corresponding features in **a**. The correlation curve and matching function are shown in figs. (6.14) and (6.15) respectively. The matching function is very near ideal, with a uniform mapping of rows of **b** to the rows of **a** except for occasional excursions. Note that relatively low values of C_{ab} do not necessarily cause matching errors, but that when matching errors occur they generally do so in regions of low C_{ab}. The connection between low C_{ab} and potential for mismatching is exploited in two rectification schemes.

During rectification **b** undergoes a warp along the x axis in order to remap its rows to the same image locations as they are found in **a**. Both methods employed to perform this remapping take advantage of C_{ab} as a measure of quality of the match. The first method is
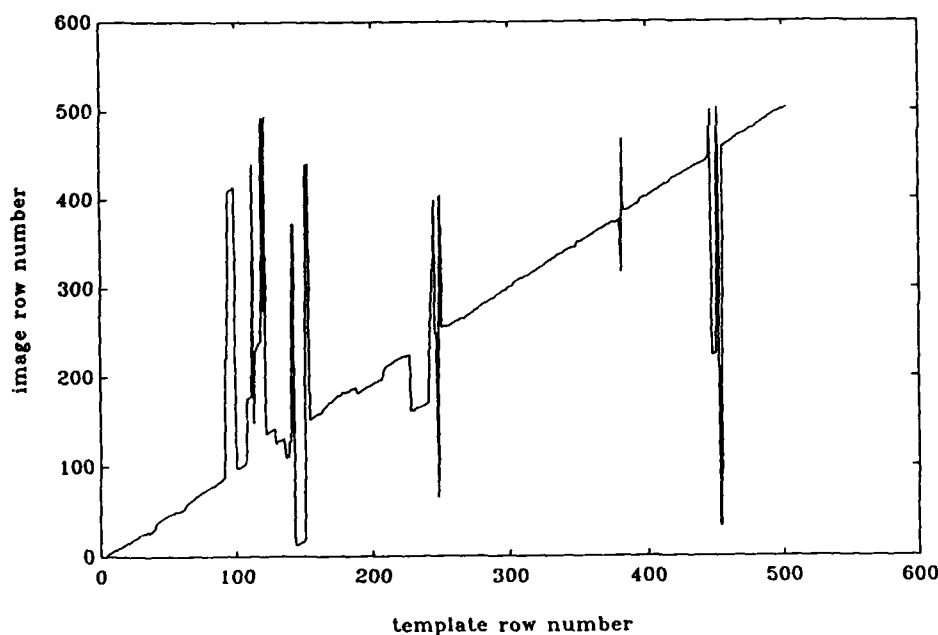
Figure 6-15: Matching function for corrected figs. (6.1) and (6.2)

called restrained mapping and proceeds as follows. For the first iteration the rows x_a and x_b in a and b which result in a global maximum for C_{ab} are located. Map x_b to row x_a in the remapped image b' and remove row x_a from further consideration since that template position has been filled with its associated row from the sample image. On the next iteration with x_a removed from consideration again find the global maximum of C_{ab} and the associated x_a and x_b. On this and subsequent iterations the current x_a, or remapping location, is compared with the nearest filled row in the remapped image both above and below the current x_a (fig. 6.16). If the ordering by row number of these three rows is the same as the associated row numbers x_b of the sample image rows from which these rows in the remapped image were taken, the current x_b is mapped into x_a in the remapped image. If the ordering is found not to be the same between these three rows in the remapped image and their relative locations in the sample image the current row is not mapped. In any case the current x_a is removed from further consideration.

This scheme is designed to make use of the assumption that the towfish did not reverse direction during either image and did not suffer sufficient attitude instability to cause rows

79

SAMPLE IMAGE   REMAPPED IMAGE

PREVIOUS MATCH

CURRENT ROW   VALID REGION FOR
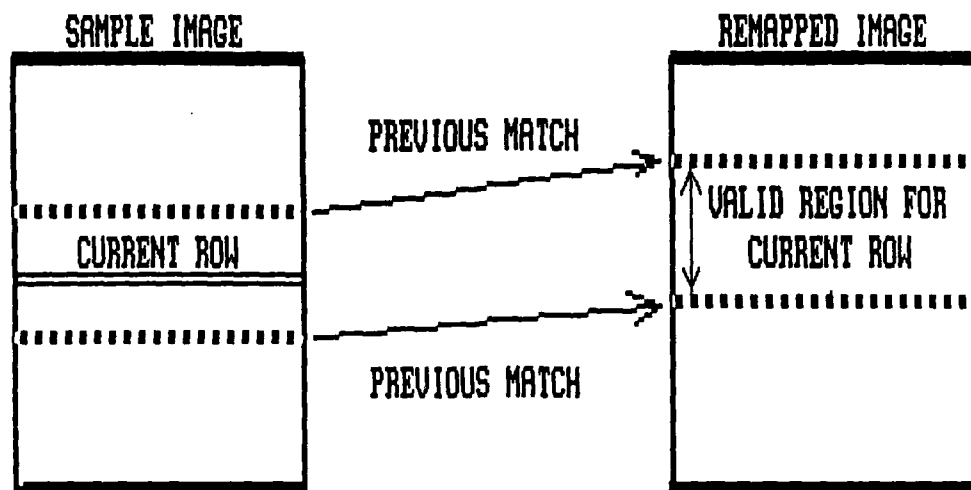   CURRENT ROW

PREVIOUS MATCH

Figure 6-16: Restrained remapping criteria

either the template and sample image to be taken out of sequence. Out of sequence lines are
then assumed to be mismatched and therefore should not be included in the remapped image.
Once the process is complete blank rows in the remapped image are filled by interpolating
across the gaps with yet unmapped lines that are in sequence with the bordering mapped
lines. If no lines remain unmapped between the two lines bordering the gap the gap is filled
by linear interpolation between the bordering lines. This method is hampered by gaps and
discontinuities in the remapped image which are caused by fortuitous mapping errors early
in the remapping process which establish a faulty ordering of remapped image rows and later
prevent mapping of correctly matched lines into the region because of the apparent incorrect
order.

The second scheme employed is called thresholded remapping. In the first step all rows
x_b of the sample image b for which the match to lines in a has an associated C_{ab} greater than
a prespecified threshold are mapped into the remapped image. In the second step all unfilled
rows of the remapped image are filled by interpolation of the yet unmapped lines of the sample
image, preserving the original order found in the sample image. This method disregards
explicit concerns about line ordering and subsequently provides for smoother remapping of
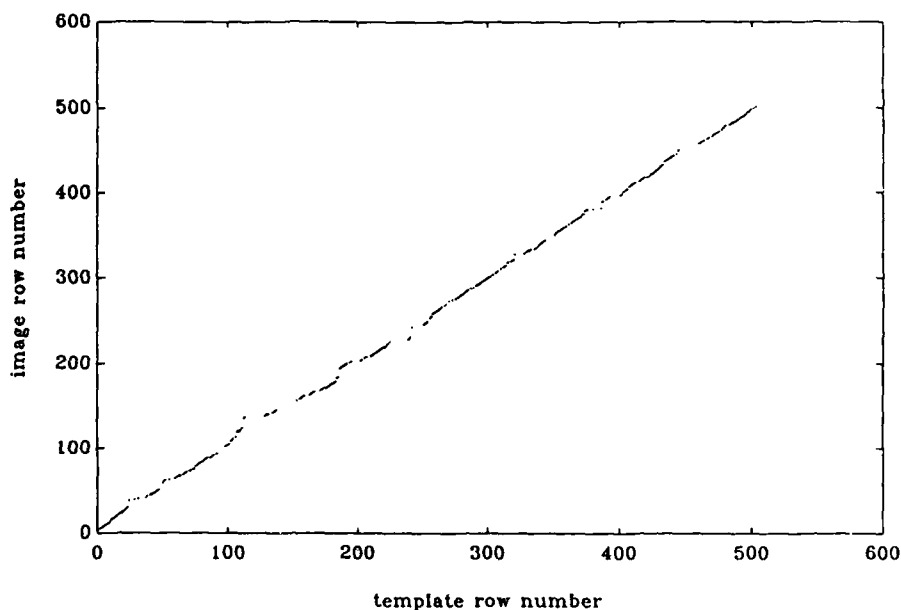the sample image.

80

Figure 6-17: Initial line mapping, restrained remapping.

Both these approaches use an initial mapping of a portion of the rows of b followed by interpolation of the remaining rows. Figs. (6.17), (6.18), and (6.19) are graphs illustrating the initial mapping of lines in the restrained, 0.4 threshold , and 0.25 threshold cases, respectively.

The restrained method maps the most lines initially, but the mapping function appears rougher than that of the 0.25 threshold case. Note that since all three preferentially use matches with high values of C_{ab} the excursions noted in previous matching functions are not present.

Four remapped images are shown for comparison. Figure (6.20) is the compensated and "blindly" remapped fig (6.2) in which each row was remapped to its best fit in the template without regard for either row order or C_{ab}. It is included for comparison purposes only. Figure (6.21) is the remapped image obtained by restrained remapping. Stratiations are evident in several region of the image where gaps existed after the initial remapping stage and were filled by interpolation. Figure (6.22) is the threshold remapped image whereby the initial mapping included only rows for which the associated value of C_{ab} was greater than 0.04. Figure (6.23) uses the same remapping scheme but uses a threshold value of 0.25. In
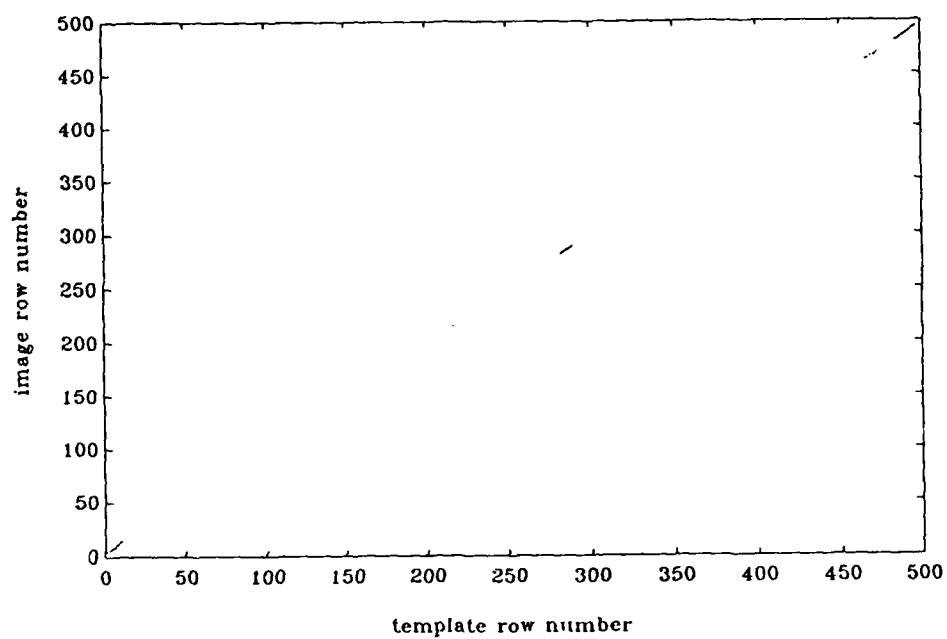
81

Figure 6-18: Initial line mapping, 0.4 threshold remapping.
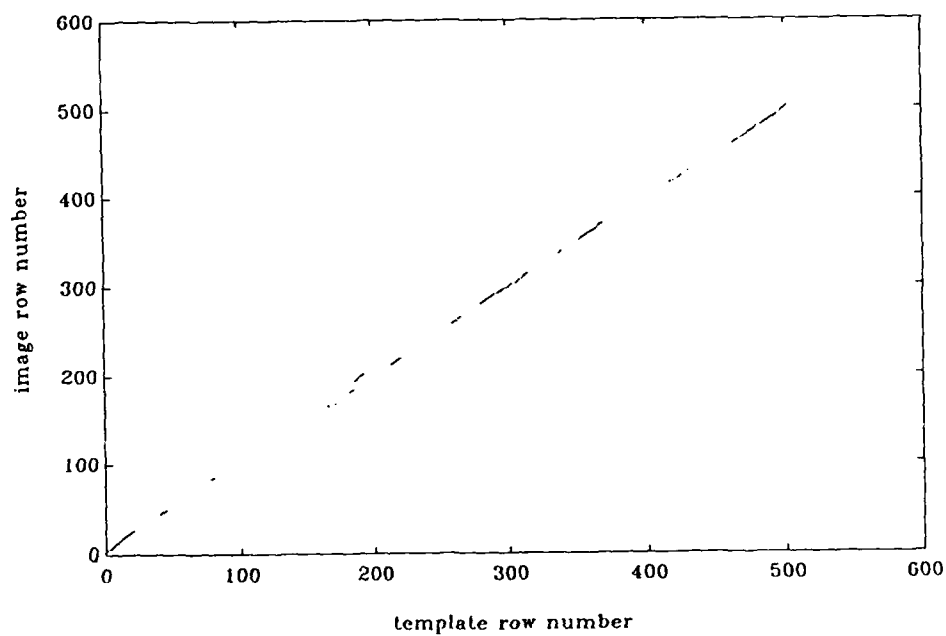
Figure 6-19: Initial line mapping, 0.25 threshold remapping.
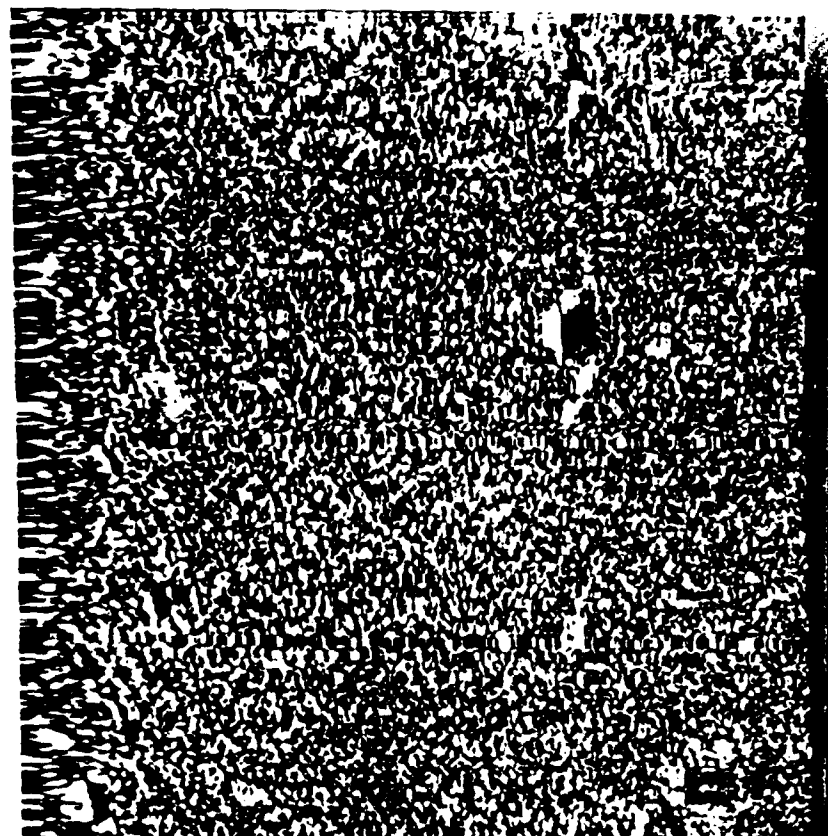
Figure 6-20: "Blind" remapping of compensated fig. (6.2)
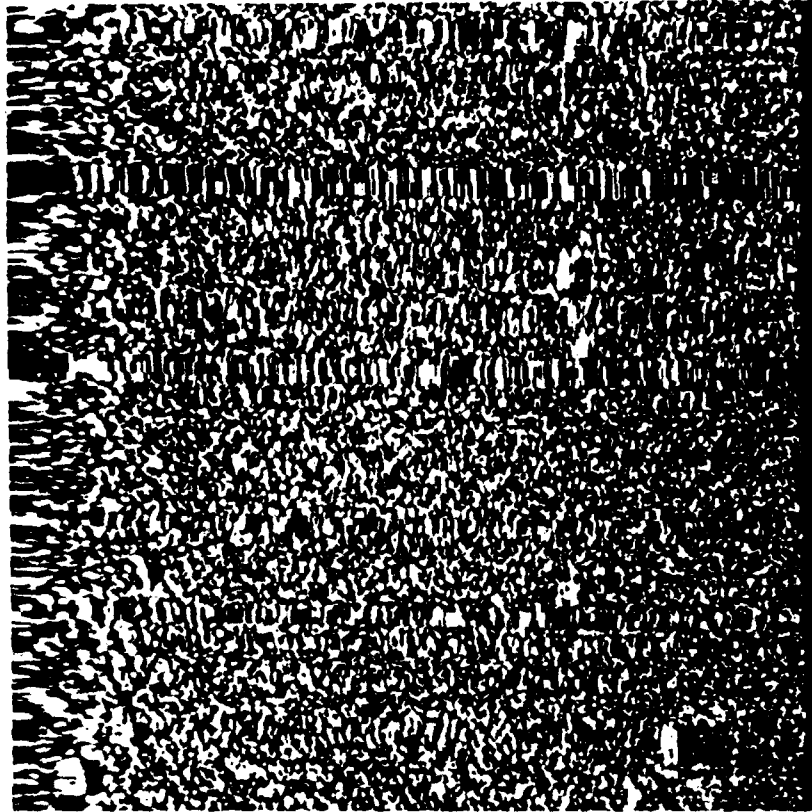
Figure 6-21: Restrained remapping of compensated fig. (6.2)
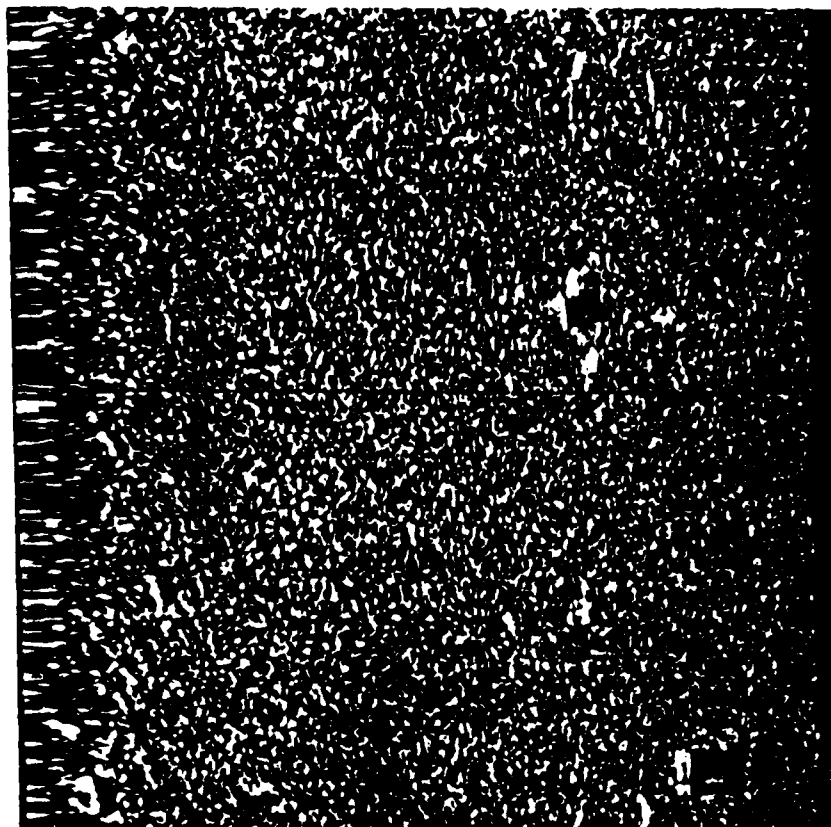
Figure 6-22: 0.4 Threshold remap of compensated fig. (6.2)
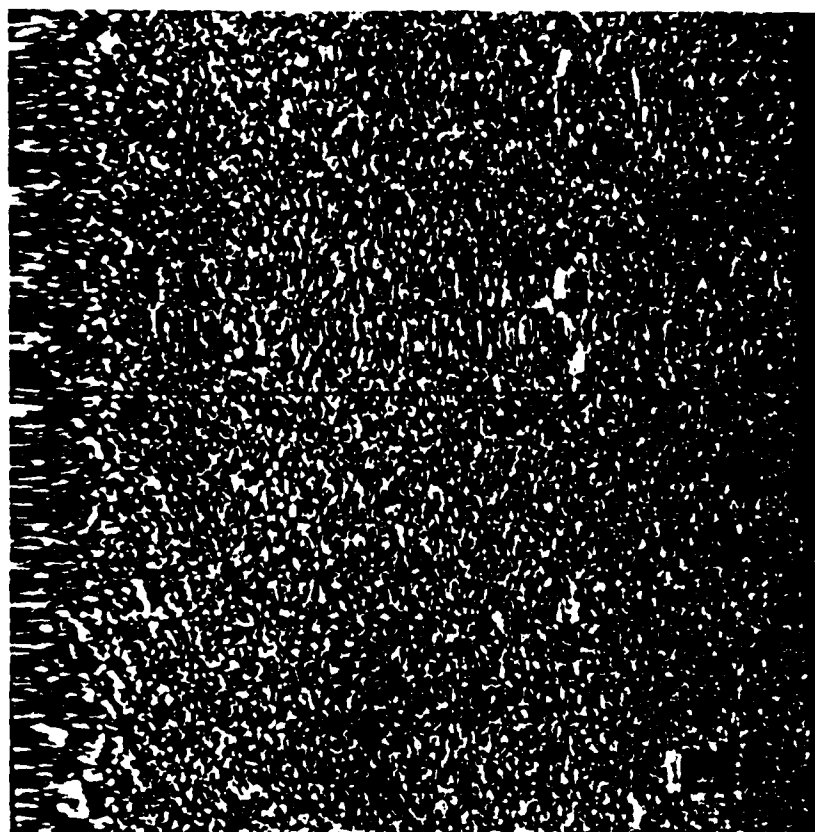
Figure 6-23: 0.25 Threshold remap of compensated fig. (6.2)

both threshold remapped images the rows for which C_{ab} is above the threshold are marked by light line segments along the right edge of the image. Both threshold remapped images have an improved appearance as compared to the restrained remapped image.

6.2.3   Registration and Comparison

We have rectified b by four different methods, and proceed to measure the degree to which the features in b now coincide with the features in a. Registration in this case has already been performed in that the extent of the bottom imaged here is a single image frame which has been rectified to conform to a single template image a. Registration therefore consists only of aligning the two image boundaries.
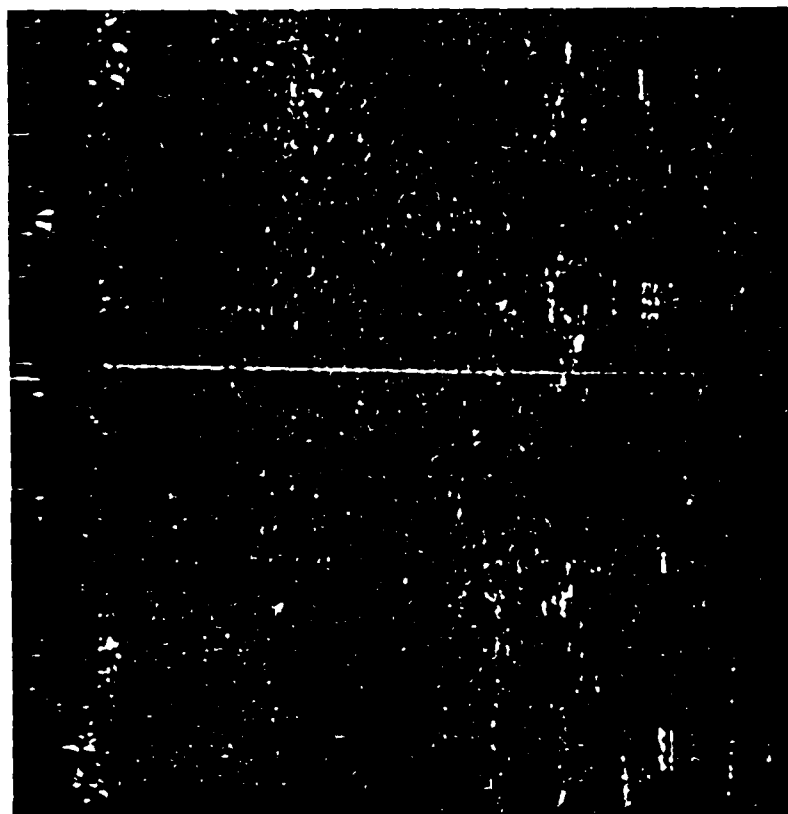
Figure 6-24: Difference image, blind remapping

As one means of comparison the remapped images b' are individually registered with a and a difference image d is generated as

$$d(x, y) = |a(x, y) - b'(x, y)| \qquad (6.9)$$

The difference images d are shown as figs. (6.21) through (6.24) for the blind remap, restrained remap, 0.40 threshold remap, and 0.25 threshold remap cases respectively. None of the images appears to be significantly better than the others, so a numerical comparison was performed in order to quantify the rms pixel differences between the two images was performed The rms pixel difference was computed as
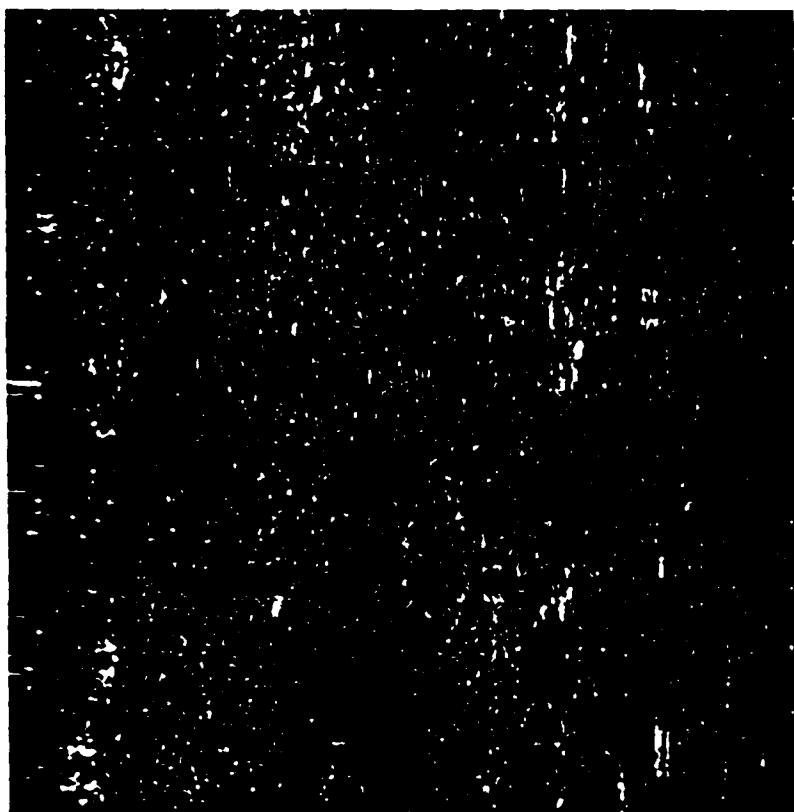
Figure 6-25: Difference image, restrained remapping
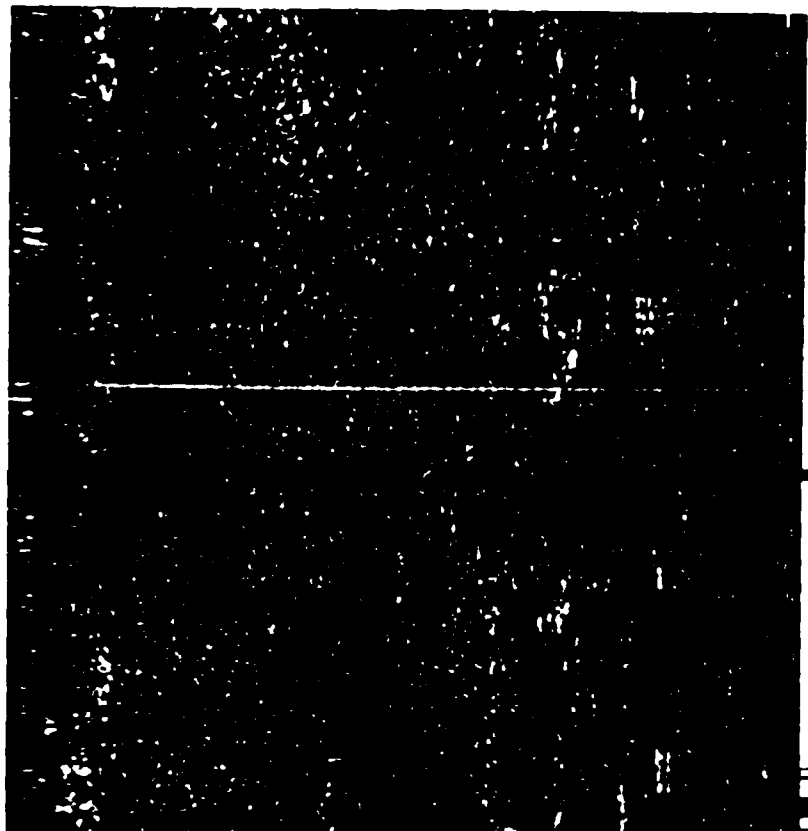
Figure 6-26: Difference image, 0.4 threshold remapping
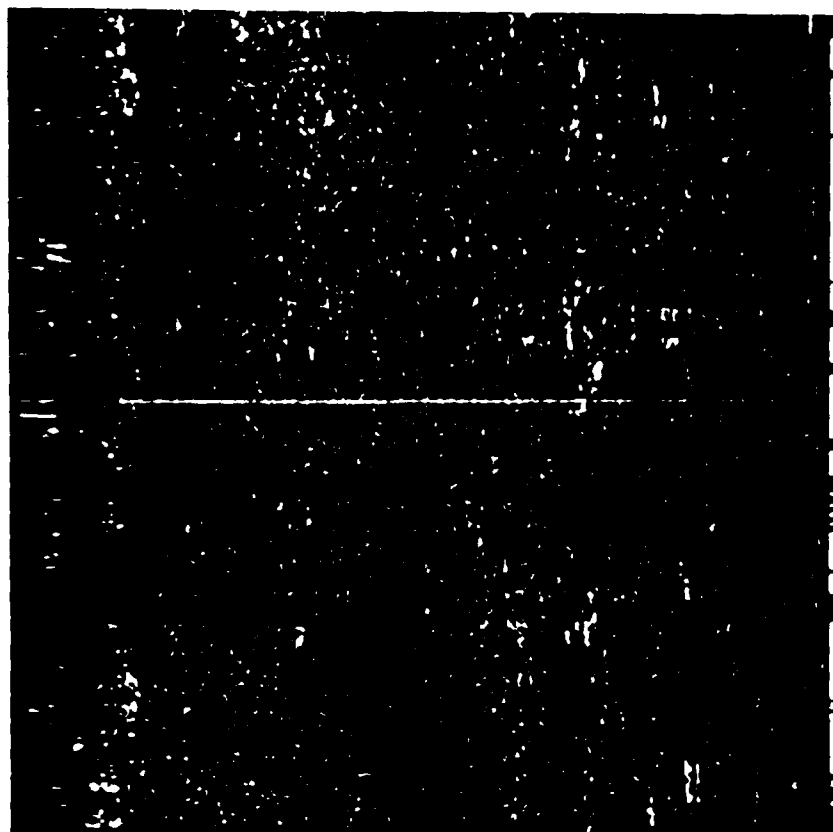
Figure 6-27: Difference image, 0.25 threshold remapping

$$\Delta i = \left(\frac{1}{512^2} \sum_{x=1}^{512} \sum_{y=1}^{512} [i_a(x,y) - i_b(x,y)]^2 \right)^{1/2} \tag{6.10}$$

The values of Δi were computed for these four remappings as well as for a manual registering of **a** and **b** whereby the image viewing window alternately displayed both images and relative positions of the two images were varied until a minimum of apparent jitter occurred on the display. Measured Δi are shown in table 6.1. The blind method should be expected

Rectification Method	Δi
Manual	19.9
Blind	19.1
Restrained	20.8
Threshold (0.4)	20.7
Threshold (0.25)	20.7

Table 6.1: Rms pixel differences for various approaches to remapping

to provide the best result as it used the best matched line in all cases without regard for line order. Both methods used here provided results comparable but slightly less accurate than manual registration. This is reasonable in this case since the images used were close to rectified at the outset, as evidenced by fig. (6.15). In a case with more severe warping the manual registration of two unrectified images would be expected to produce degraded results.

For further comparison **a** and **b** were intentionally misregistered by 10 pixels in both the x and y directions. The resulting Δi was 21.9 . This misregistration corresponds to a length on the bottom of 1.25 M, which suggests an approximate upper error bound on the methods presented here.

Assuming the information derived in chapter 4 can be applied here, a theoretical lower bound to Δi can be computed. Δi can be redefined as

$$\Delta i = E\left([i_a - i_b]^2 \right)^{1/2} \tag{6.11}$$

Here i_a and i_b correspond to pixel intensities in **a** and **b** respectively and $E()$ denotes the expected value. These intensities may be decomposed into non fluctuating and fluctuating components.

$$i_{a,b} = \overline{i}_{a,b} + \tilde{i}_{a,b} \qquad (6.12)$$

Proceeding with the computation

$$
\begin{aligned}
\Delta i &= \mathrm{E}\left(\left[(\overline{i_a} + \tilde{i}_a) - (\overline{i_b} - \tilde{i}_b)\right]^2\right)^{1/2} \\
&= \mathrm{E}\left(\left[\overline{i_a} - \overline{i_b} + \tilde{i}_a - \tilde{i}_b\right]^2\right)^{1/2} \\
&= \mathrm{E}\left(\left[\tilde{i}_a^{\,2} - 2\tilde{i}_a\tilde{i}_b + \tilde{i}_b^{\,2}\right]\right)^{1/2} \\
&= \sqrt{\mathrm{E}(\tilde{i}_a^{\,2})} - 2\sqrt{\mathrm{E}(\tilde{i}_a\tilde{i}_b)} + \sqrt{\mathrm{E}(\tilde{i}_b^{\,2})} \qquad (6.13)
\end{aligned}
$$

During equalization mean values of both images were set to 100, so the mean value terms dropped out. In the last chapter the fluctuation were found to be spatially independent, so the cross term becomes zero. Also determined in the last chapter was that pixel intensity fluctuations were approximately 8% of the mean pixel value. In these processed images the mean value was set at 100, yielding

$$\Delta i = \sqrt{(0.08 \cdot 100)^2} + \sqrt{(0.08 \cdot 100)^2} = \sqrt{128} = 11.3 \qquad (6.14)$$

which would occur under ideal conditions. The actual results therefore show more variability than predicted by the pierside phase 1 results.

Chapter 7

Conclusion

7.1 Results

In this thesis we have been able to quantify many of the aspects of side scan sonar normally not encountered in present literature.

Test tank experimental results reveal that the waveform of the particular sonar used consists of a 122 KHz carrier modulated by a decaying exponential with a bandwidth of 9 KHz. Within this bandwidth power fluctuations for the transmitted acoustic pulse are approximately 2%, with increased fluctuation outside this band due cavitation and non-linear effects of the medium. Total power calculations performed for the overall system through analysis of imagery indicates that system power fluctuations amount to 3.4%. This measure of fluctuation compounds transmit power variability with variability induced by the medium.

Analysis of imagery from repeated insonification of the same bottom features shows that intensity fluctuations for echoes from individual features are significant. Fluctuation is range dependent and averages 8% of mean image pixel intensity. Minimum fluctuation of approximately 4% occurs at 15 M range while a maximum of 14% occurs at 40 M. Compensating the experimental data for transmit power fluctuations shows that system power fluctuations are not the cause of observed image intensity fluctuations. The multiplicative rather than additive nature of the fluctuations and experiment configuration indicate that surface scatter interference is also not the primary cause. These fluctuations were further shown to be spatially and temporally uncorrelated. The lack of correlation between pixels allows application

of the Central Limit Theorem of probability theory for analysis of image features made up of several pixels. The Central Limit Theorem states that the mean value of the sum of n identically distributed independent random variables a_i having mean $\overline{a_i}$ and standard deviation σ_{a_i} is $n\overline{a_i}$. The standard deviation of the sum is $\sqrt{n}\sigma_{a_i}$. In the case of a image feature made up of several pixels which are assumed to have nearly identical statistics, the overall ratio of standard deviation to mean for this region of pixels equals $\frac{\sqrt{n}}{n} = \frac{1}{\sqrt{n}}$. The fluctuations for an image region defining a bottom feature therefore can be expected to be significantly less than that of an individual pixel.

The expectation that image regions defining individual features display relatively modest fluctuation allows development of a correlation matching approach to identifying similar subregions in separate side scan sonar images of the same bottom scene. We have demonstrated that corresponding subregions of two images of the same bottom region can be matched consistently, as evidenced by matching function obtained from the phase two pierside experiment. In order to make meaningful matches, preliminary processing including slant range correction and intensity equalization must be performed to remove imaging geometry induced aspects which otherwise dominate the matching process and generally interfere with matching of bottom features. Matching of subregions allows image rectification or remapping of image data to the coordinate system of another image. Such remapping allows feature-by-feature comparison of two images.

The results of remapping the features of one image to another show that automated alignment of features or registration is comparable to manual registration. A fair amount of difference exists between the remapped image and the image to which it was remapped for all methods attempted, including manual registration.

7.2   Future Work

All work described in this thesis either assumed no towfish instabilities or was done with data taken from experiments in which towfish stability was controlled. The deleterious effect of towfish instabilities is appreciated by undersea surveyors and efforts are normally made to minimize these. However, in circumstances where instabilities are unavoidable their effects should be understood. The rectification and registration presented here relied on the assump-

tion of a straight, stable towfish path that was repeatable. When such assumptions cannot be made the algorithm must become more general. Under conditions of varying bottom track location and orientation the validity of the assumption that rows of one image could be mapped to rows in another without further processing is lost. For the case of bottom tracks that are parallel as in this study yet may be displaced laterally the correlation search routine must include searching along the row length as well as between rows in order to find the best match. Bottom tracks which are angularly displaced require that images be rotated prior to correlation or that the search routine rotate the correlation window as well as translate it through the image while attempting to locate the best match. In the general case the best correlation window is probably square or round, and is translated and rotated throughout the image while attempting to locate the best match. Remapping through the use of image control points and polynomial warping is currently used in satellite remote sensing to rectify images taken from various aspects [12] . An adaptation of this procedure to side scan sonar imagery is possible and should be attempted.

With an increased number of degrees of freedom additional removal of aspect dependent image effects may be required. The preliminary processing discussed in this thesis removes a portion of the aspect dependent effects, however the appearance of individual *bottom* features due to imaging geometry has not been dealt with. The imaging process maps three dimensional objects into a two dimensional projection which is a function of aspect. The correlation routine in the general case should be sufficiently robust to identify the changed aspect of the imaged object. Additionally, the location and size of the acoustic shadow associated with an object on the bottom is determined by the imaging geometry. A generalized search routine should be able to recognize this effect as well and compensate for it. To remove the aspect dependence of target appearance and shadow it might be worthwhile to investigate transforming the side scan sonar image to an image in which features are represented symbolically.

The rms pixel differences calculated for the various rectification and registration schemes presented here show approximately 70% more difference between images than could be explained by the variability measured during the phase one pierside experiment. There are two probable causes for this excess variability which were not investigated during this experiment.

Although the phase two pierside equipment was designed to limit it,it is possible that small amounts of attitude instabilities such as roll, pitch, and yaw existed during the experiment. These further complicate the ability to match image subregion by causing intersecting and repeated bottom strips. The degree to which these corruptions affect the rectification and registration process and algorithms for their removal should be investigated .

Another possible cause which should be investigated is the effect of speckle and its relationship to imaging geometry. Speckle is a feature common to all coherent imaging systems such as side scan sonar and arises from constructive and destructive interference of individual scatterers within a single image element or pixel [6] . Experimental evidence shows that because of speckle small changes in imaging geometry can result in changes in the pixel intensity pattern which greatly exceed change in intensity that would be expected in incoherent imagery such as with white light [29]. The degree to which this occurs in side scan sonar imagery is unknown, but it is possible that small deviations in towfish location may result in dramatic image differences and therefore lower correlation.

Bibliography

[1] J. S. Bendat and A. G. Piersol. *Random Data: Analysis and Measurement Procedures.* Wiley–Interscience, New York, 1971.

[2] F. M. Berkson and C. S. Clay. Transformation of side scan records to a linear display. *Int. Hyd. Rev.*, 50:55–59, 1973.

[3] L. M. Brekhovskikh and Y. P. Lysanov. *Fundamentals of Ocean Acoustics.* Springer–Verlag, Heidelberg, West Germany, 1982.

[4] J. T. Christoff et al. Measurement of the temporal phase stability of the medium. *J. Acoust. Soc. Am.*, 71(6), June 1982.

[5] C. S. Clay and H. Medwin. *Acoustical Oceanography.* John Wiley and Sons, New York, 1977.

[6] J. C. Dainty, editor. *Topics in Applied Physics Vol. 9, Laser Speckle and Related Phenomena.* Springer–Verlag, Berlin, West Germany, 1984.

[7] P. N. Denbigh. A review of rapid depth measuring techniques and the development of a bathymetric side scan sonar. In W.G.A. Russell-Cargill, editor, *Recent Developments in Side Scan Sonar Techniques*, Central Acoustics Laboratory, University of Cape Town, Cape Town, South Africa, 1982.

[8] P. Desantis et al. The use of subharmonics. *J. Acoust. Soc. Am.*, (46):514, 1967.

[9] A. W. Drake. *Fundamentals of Applied Probability Theory.* McGraw–Hill, New York, 1967.

[10] Ira Dyer. Fundamentals and applications of underwater sound-course notes. 1987. Notes from course number 13.851, given at the Massachusetts Institute of Technology.

[11] B. W. Flemming. A historical introduction to underwater acoustics with special reference to echo sounding, sub-bottom profiling, and side scan sonar. In W.G.A. Russell-Cargill, editor, *Recent Developments in Side Scan Sonar Techniques*, Central Acoustics Laboratory, University of Cape Town, Cape Town, South Africa, 1982.

[12] E. M. Hall. *Computer Image Processing and Recognition*. Academic Press, New York, 1979.

[13] R. F. Henderson. Recent developments in side scan sonar and applications in the offshore oil field. In *Offshore Technology Conference, 7th Annual Proceedings*, pages 733–742, May 1975.

[14] D. R. Jackson et al. High frequency bottom backscatter measurements in shallow water. *J. Acoust. Soc. Am.*, 80(4):1188–1199, October 1986.

[15] B. I. Justusson. Median filtering: statistical properties. In T. S. Huang, editor, *Topics in Applied Physics Vol. 43, Two Dimensional Digital Signal Processing II, Transforms and Median Filters*, chapter 5, Springer–Verlag, Heidelberg, West Germany, 1981.

[16] J. H. Keenan et al. *Steam Tables*. John Wiley and Sons, New York, 1969.

[17] M. R. Kenny and J. D. Mix. *Side Scan Sonar Target Detection in the Presence of Bottom Backscatter*. Master's thesis, Naval Postgraduate School, Monterey, California, September 1983.

[18] R. M. Kessler. USS Recovery Completes F-16b Salvops. *Faceplate*, April 1981.

[19] M. Klein. A modular sonar system for seabed mapping. In W.G.A. Russell-Cargill, editor, *Recent Developments in Side Scan Sonar Techniques*, Central Acoustics Laboratory, University of Cape Town, Cape Town, South Africa, 1982.

[20] E. Y. Kuo. Wave scattering and transmission at irregular surfaces. *J. Acoust. Soc. Am.*, (36):2135–2142, 1964.

[21] O. H. McDaniel. Harmonic distortion of spherical sound waves in water. *J. Acoust. Soc. Am.*, (38):644, 1965.

[22] V. V. Ol'shevskii. *Characteristics of Sea Reverberation*. Consultants Bureau, New York, 1967.

[23] A. V. Oppenheim and R. W. Shafer. *Digital Signal Processing*. Prentice-Hall, Englewood Cliffs, New Jersey, 1975.

[24] Peter R. Paluzzi et al. Computer rectification and mosaicking of side-looking sonar images. In *Eleventh Annual Offshore Technology Conference Proceedings*, Houston, Texas, May 1979.

[25] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw–Hill, New York, 1984.

[26] D. H. Peterson. Side scan sonar - an alternative to wire drag for item investigation. In *Proceedings of the Canadian Hydrographic Service Centennial Conference*, Government of Canada, Ottawa, Ontario, 1983.

[27] M. J. Sheehy. Transmission of 24kc sound from a deep source. *J. Acoust. Soc. Am.*, 22(24), 1950.

[28] T. K. Stanton. Sonar estimates of seafloor microroughness. *J. Acoust. Soc. Am.*, 75(3):809–818, March 1984.

[29] G. E. Trahey et al. Speckle pattern correlation with lateral aperture translation: experimental results and implications for spatial compounding. *IEEE Trans. Ultrason. Ferroelec. and Freq. Cont.*, UFFC–33(3):257–264, May 1986.

[30] R. J. Urick. *Principles of Underwater Sound*. McGraw-Hill, New York, 1983.

[31] R. J. Urick. *Sound Propagation in the Sea*. Peninsula, Los Altos, California, 1982.