Wavelet Based Similarity Measurement Algorithm for Seafloor Morphology

By

Ilkay Darilmaz

S.B. in Mechanical Engineering

Turkish Naval Academy, 2002

Submitted to the Department of Mechanical Engineering in Partial Fulfillment of the Requirements for the Degree of

Master of Science in Naval Architecture and Marine Engineering

Master of Science in Mechanical Engineering

at the

Massachusetts Institute of Technology

June, 2006

Signature of Author:.....
Department of Mechanical Engineering
May 19, 2006

Certified by:.....
Nicholas M. Patrikalakis
Kawasaki Professor of Engineering
Thesis Supervisor

Accepted by:.....
Lallit Anand
Professor of Mechanical Engineering
Chairman, Department Committee on Graduate Students

Wavelet Based Similarity Measurement Algorithm for Seafloor Morphology

By

Ilkay Darilmaz

Submitted to the Department of Mechanical Engineering on May 19, 2006

in partial fulfillment of the requirements for the degrees of

Master of Science in Naval Architecture and Marine Engineering

Master of Science in Mechanical Engineering

Abstract

The recent expansion of systematic seafloor exploration programs such as geophysical research, seafloor mapping, search and survey, resource assessment and other scientific, commercial and military applications has created a need for rapid and robust methods of processing seafloor imagery. Given the existence of a large library of seafloor images, a fast automated image classifier algorithm is needed to determine changes in seabed morphology over time. The focus of this work is the development of a robust Similarity Measurement (SM) algorithm to address the above problem. Our work uses a side-scan sonar image library for experimentation and testing. Variations of an underwater vehicle's height above the sea floor and of its pitch and roll angles cause distortion in the data obtained, such that transformations to align the data should include rotation, translation, anisotropic scaling and skew. In order to deal with these problems, we propose to use the Wavelet transform for similarity detection. Wavelets have been widely used during the last three decades in image processing. Since the Wavelet transform allows a multi-resolution decomposition, it is easier to identify the similarities between two images by examining the energy distribution at each decomposition level. The energy distribution in the frequency domain at the output of the high pass and low pass filter banks identifies the texture discrimination.

Our approach uses a statistical framework, involving fitting the Wavelet coefficients into a generalized Gaussian density distribution. The next step involves use of the Kullback-Leibner entropy metric to measure the distance between Wavelet coefficient distributions. To select the top N most likely matching images, the database images are ranked based on the minimum Kullback-Leibner distance. The statistical approach is effective in eliminating rotation, mis-registration and skew problems by working in the Wavelet domain. It's recommended that further work focuses on choosing the best Wavelet packet to increase the robustness of the algorithm developed in this thesis.

Thesis Supervisor: Nicholas M. Patrikalakis, Ph.D.
Kawasaki Professor of Engineering

Acknowledgements

A journey is easier when you travel together. Interdependence is certainly more valuable than independence. This thesis is the result of two years of work whereby I have been accompanied and supported by many people. It is a pleasant aspect that I have now the opportunity to express my gratitude for all of them.

The first person I would like to thank is my Thesis supervisor Nicholas M. Patrikalakis, Ph.D. I have known him as a sympathetic and principle-centered person. He could not even realize how much I have learned from him. Besides of being an excellent supervisor, he was as close as a relative and a good friend to me. He always helped me any kind of problems I had. I am really glad that I have come to get know Nicholas M. Patrikalakis in my life.

I also would like to thank to Dr. Wonjoon Cho who kept an eye on the progress of my work and always was available when I needed his advises.

Very special thanks to my friends, who never leave me alone during my education in MIT. A real friend is one who walks in when the rest of the world walks out. Dimitrios Panagiotidis, Siddarth Kumar, Sunil Doddabasanagouda, Piyush Hari and Srinivasan Nallasivan. Each one of you has different place in my heart and you will always stay there.

The chain of my gratitude would be definitely incomplete if I would forget to thank the first cause of this chain, my family. I love you more than anything in this life and I know you will always be there for me whenever I need you. You hold my hand when I stumble and show me the way.

And my guzeller guzeli, don't know what to do without you...

This work was supported in part by NOAA via the MIT Sea Grant College Program.

Ilkay Darilmaz
Cambridge, Massachusetts. 2006

Table of Contents

List of Figures

Chapter 1: Introduction

1.1 Motivation

The Us Commission on Ocean Policy professes a "vision for the future" that entails "strengthening science and the generating high quality data" for policy makers managing ocean resources. This focus on data quality underscores the need for the information that can be interpreted and acted upon decision making that will affect the future of the resource. Ideally such databases are easy to understand, offer insight into the fundamental physical processes affecting the region of interest, have a resolution that is commensurate with viewing and interpreting the area, and come from several sources so as to increase the spectrum of information used in generating the data product. Perception is the process of achieving understanding, a major step forward from merely collecting data and we anticipate our work will improve the perceptive capabilities of image analysis techniques the community of ocean scientists currently relies on promote policy making. In the past decade in the United States, academic institutions have pioneered remote sensing technologies to improve the understanding of the marine environment.

1.2 Mathematical Background

Recently, Wavelets have been used in many signal processing as well as image processing applications because of their unique properties, both in theory and application. In terms of theory, Wavelets can easily provide resolution of signals using the various wave form bases. Unlike the Fourier transform, the Wavelet transform lets us identify signals using linear combinations of different forms of Wavelet bases shown as in equation (3).

In the Fourier series expansion, a function $f(x)$ is represented by [27]:

$$f(x) = a_0 + \sum_{k=1}^{\infty}(a_k \cos(kx) + b_k \sin(kx)) \tag{1}$$

where each coefficient is computed from[27]:

$$a_0 = \frac{1}{2\pi}\int_0^{2\pi} f(x)dx \qquad a_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos(kx)dx \qquad b_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin(kx)dx \tag{2}$$

In the Wavelet transform, the mother Wavelet translated by various values like "j" and "k" is, given by respectively [25];

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k) \tag{3}$$

Where, $\psi_{j,k}$ is the mother Wavelet and "j" is is for dilation and "k" is for translation in equation (3) and (4). Essentially, the Wavelet transform is the product of selected mother Wavelet type with original signal.

$$T(j,k) = \int_{-\infty}^{\infty} f(t)\psi_{j,k}(t)dt \tag{4}$$

Recently, the Wavelet applications in image processing have been growing rapidly. In this thesis, Wavelet transformation is mainly used for sonar imagery analysis. We can visualize the fundamental idea in Wavelet transform by considering diffraction of light by a simple prism. The light passing through the prism is decomposed into several color segments. Seven primary colors provide a basis for representing light sources. In the same fashion, Wavelets are bases for representing functions or signals [17]. Here, the decomposition is very useful because it allows us to analyze every segment separately, similar to the "divide and conquer" method. In many cases, one would like to approximate a signal given its basis expansion [1].

Image processing can be considered as 2–D signal processing, say an image can be represented by using a matrix giving the color intensities. Image data source, shown in Figure (1) is obtained by Klein image library [7] and decomposition has been processed by using Matlab Wavelet Toolbox. Columns and rows of this matrix can also be considered as 1-D discreet signals. The significant Wavelet coefficients are represented in white color densities. The most important point here is that the locations of these significant coefficients demonstrate a visible geometrical correlation as they form simple curves. Thus, 2-D Wavelet transform is good at catching edge points, but the smoothness along contours. After Wavelet transformation, we end up with filtered image, which has less details comparing to the original image. After filtering an image, remaining significant features gives us the location of the significant Wavelet coefficients. Same idea has been used in Matlab Image Processing Toolbox to perform morphological operations [1].

We can consider a simple model for natural images that are composed of piecewise smooth regions separated by smooth boundaries, as shown in Figure (1). In image processing, it is known that boundaries of shapes provide the most useful information. This model obviously ignores the important image information, namely texture [8]. In the 2-D piecewise smooth model, the discontinuities are generated by edges whereas edges are often gathered along smooth contours that are created by typically smooth boundaries of physical objects.

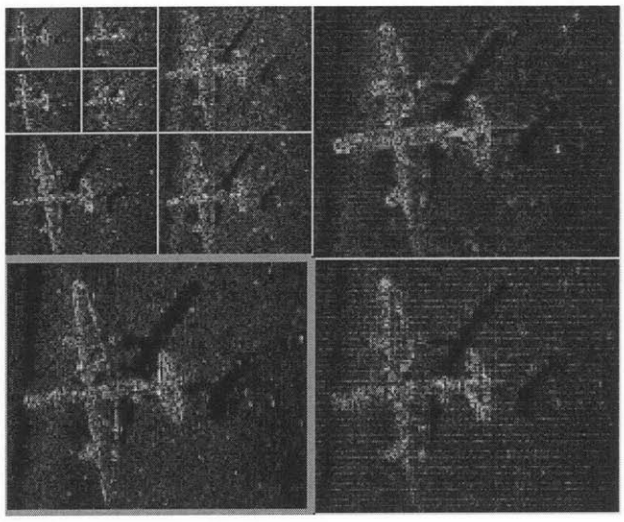

Figure 1: Example of Wavelet Decomposition of a Side Scan Sonar Image.

1.3 Problem Statement

Robust applications of Wavelet transformation in image processing area explicitly can be considered as solution for problems in sonar images. In our work, large library of seafloor side scan sonar images processed. In collection phase of the seafloor imagery, there are clearly random problems depending on the height of the sonar above the seafloor [19], pitch and roll angle. Sea elevation is described by cosines function which has phase angle (Φ) in terms of probability density function.

This random behavior shown in Figure (2) will cause some problems in side scan sonar images in many different ways, namely skew, stretch, rotation or translation, etc. These kinds of problems directed us to develop a fast automated image classifier algorithm which can determine changes in seabed morphology over time [13]. Following chapters will clearly explain how Wavelets can be used to reduce the computational complexity and how they can fix the

problems explained above. Also considering having large size of image library, it makes it a lot harder to check every side-scan seafloor image by human effort, say distinguishing buried objects with other rocks and sand ripples for instance. Here developed algorithm helps us to saving human effort and time.

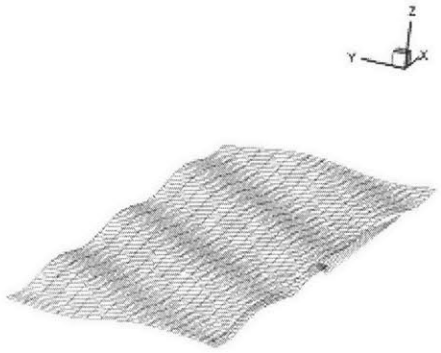

Figure 2: Representation of Random Sea-State and Waves

1.4 Side Scan Sonar Images

1.4.1 Understanding the Side Scan Sonar

Side scan sonar differs from the classic sonar in many aspects. For instance, instead of measuring the depth to the ocean bottom, side scan sonar retrieves information about sea floor composition by considering the different sound absorbing and reflecting features of different materials. Some types of material, such as metals, are very efficient at reflecting acoustic pulses.

Clay and silt, on the other hand, do not reflect sound well. Strong reflectors create strong echoes, while weak reflectors create weaker echoes.
Knowing these characteristics, you can use the strength of acoustic returns to examine the composition of the sea floor [10]. Reporting the strength of echoes is essentially what side scan sonar is designed to do. Combining bottom-composition information provided by side scan sonar with the depth information from range-finding sonar can be powerful tool for examining the characteristics of an ocean bottom [19].

1.3.2 How Side Scan Sonar Works

Side scan sonar employs much of the same hardware and processes as a conventional depth sounding sonar. Pulses are transmitted by a transmitter using a projector (or array of projectors), and hydrophones receive echoes of those pulses from the ocean floor and pass them to a receiver system [17]. Where side scan sonar differs from a depth-sounding system is in the way it processes these returns. To briefly explain, the transmitter of the single-beam echo sounder produces an acoustic pulse, or ping, which is transferred into the water by the projector, see Figure (3). The ping expands as a spherical wave from its source, striking the sea floor at the point closest to the projector source [19]. The ping is then reflected in a return spherical wave, part of which is detected by the hydrophone.

The receiver records the returned echo, which can be illustrated by a plot of amplitude as a function of time, see Figure (4). The time between the transmission of the ping and the reception of its echo is used to compute the range to the sea floor.

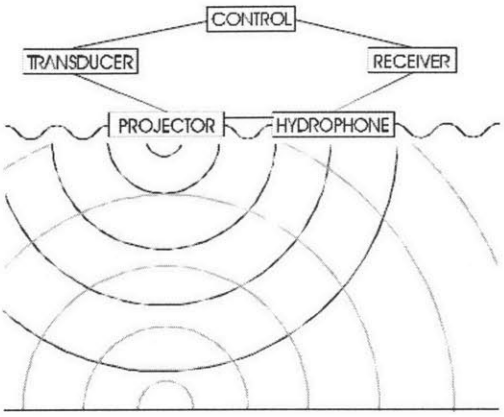

Figure 3: Single-Beam Echo Sounding System

This simplified picture does not consider what happens to the transmitted pulse after it first strikes the bottom, because a single-beam echo sounder is only interested in the time between transmission and the earliest return echo.

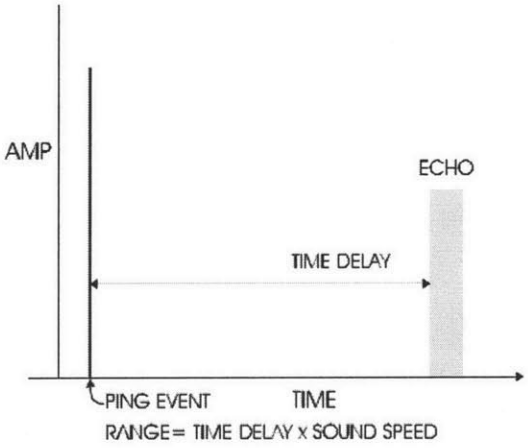

Figure 4: Plot of Amplitude as a Function of Time [19]

Let us take a look at what changes in the transmitted pulse case as it continues its spherical propagation. Figure (5) shows the single-beam echo sounder schematic at a time shortly after the ping first intersected with the bottom. Here, the first return echo is on its way back to the hydrophones. But the ping is still expanding, and in our schematic is intersecting the bottom at two points [17]. These points also create echoes. They are weaker than the first pulse, and they occur after it does. It is easy to see that as the pulse front continues to propagate, it will produce a continuous series of weakening echoes in time. If the receiver continues to collect these echoes, it will see the amplitude versus time sequence similar to the one pictured in Figure (6), with a strong first bottom echo followed by a declining slope of continuous returns. Thus far, this discussion has ignored the characteristics of the ocean bottom. Although unstated, the assumption to this point has been that the bottom is perfectly flat and perfectly uniform [19].

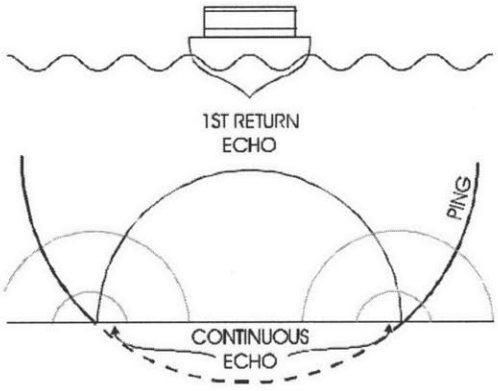

Figure 5: Single-Beam Echo Sounder Schematic [19]

Consider now what happens when some detail is added to the ocean bottom. It redraws the schematic (shown in Figure (7)) with a spherical pulse front some time after the first bottom echo. In this case, however, a box represents detail on the bottom.

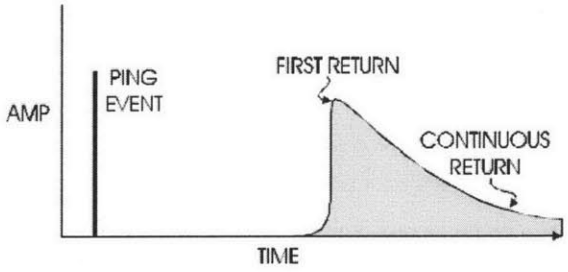

Figure 6: Amplitude versus Time Sequence [19]

This box is more efficient at reflecting acoustic energy than the surrounding uniform bottom. Because of this, the energy of its echo is slightly higher than that of the surrounding bottom, which causes a small jump in the amplitude measured by the receiver at the time of its return.

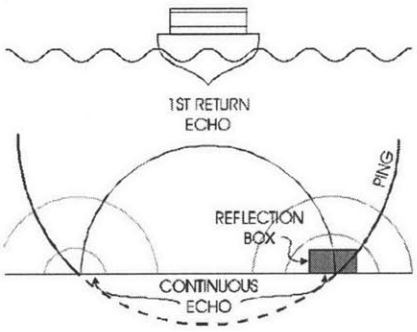

Figure 7: Schematic with a Spherical Pulse Front with a Detailed Bottom [19]

An observer looking at the amplitude versus time sequence in Figure (8) would be able to discern from this jump that some discrete feature exists on the bottom.

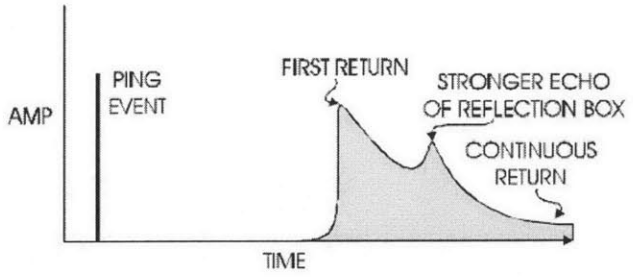

Figure 8: Amplitude vs. Time Sequence [19]

In reality, of course, ocean bottoms can be very complex. They may be composed of many different components, all with different acoustic reflection characteristics. True amplitude versus time sequences observed by receivers are similarly complex [19].

1.3.3 Basic Side Scan Sonar

Basic side scan sonar described above is not very useful as a practical instrument [19]. While it provides the times of echoes, it does not provide their direction. The side scan sonar detects a bottom feature. From the amplitude versus time plot in Figure (8), an observer can tell there is a highly reflective feature on the bottom. From the time difference between the first echo and the echo of the reflective feature, the observer can compute the range to the feature from the sonar. This is all the information the observer has.

Figure (9) shows the bottom in this example redrawn from an overhead view. The "X" marks the point on the sea floor where the ping strikes first, creating the first return echo. After striking at the "X," the spherical pulse front continues to propagate, intersecting the bottom in an expanding circle, eventually striking the bottom feature represented by the box. The observer can compute the range R between the "X" and the box, but otherwise does not know where on the ring of the expanding wave front the box is. In the real world, where sidescan sonars are used to locate objects such as rock formations or sunken vessels on the bottom, it is easy to see that this simple depth sounder amplitude information is too limited to be useful [19].

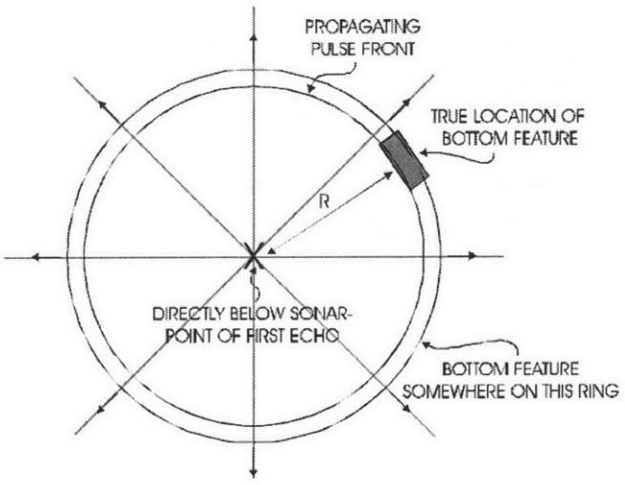

Figure 9: Overhead View of the Bottom [19]

Most side scan sonars deal with this problem by introducing some directivity to their projected pulses, and, to some degree, their receivers. This is done by using a line array of projectors to send pulses. The long axis of the line array is oriented parallel to the direction of travel of the sonar survey vessel. Generally, the arrays are towed behind the ship. Figure (10) shows a survey vessel towing a line array, and the propagation of circular pulses from that array [19].

These pulses first strike the bottom directly below the sonar system, and then their intersection with the sea floor travels away from the array in either direction. Echoes are returned first from the bottom directly below the array, followed by echoes from points farther and farther away along a line perpendicular to the array axis and the survey ship's track.

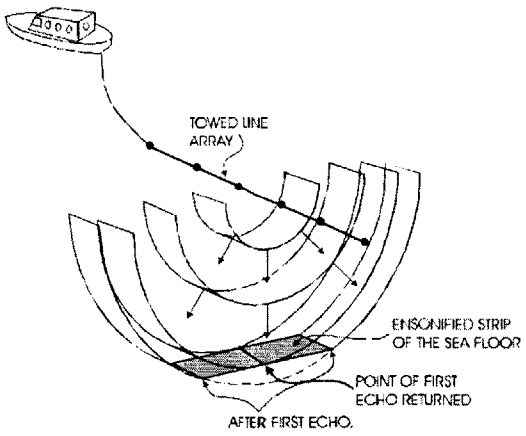

Figure 10: Survey Vessel Towing an Array [19]

Side scan sonar can receive the echoes from the bottom using two hydrophones; each set up to receive the returns from one side of the ship's survey track. In reality, instead of one projector and two hydrophone arrays, two line arrays are used as transducers. One transmits a fan beam towards the port side and then listens to its returns, and the other does the same on the starboard side.

When a pulse front first strikes the bottom below the survey vessel, an echo returns from only the one point of intersection. Later, however, the pulse front intersects the bottom at two points, one on each side of the ship.

The echoes from these two points would occur simultaneously, and a single hydrophone system would have no way of distinguishing between them. If, however, two separate hydrophones each record the echoes from only one side of the ship, there would be no such confusion.

The name side scan is used because these sonars were originally built to be sensitive to echo returns from bottom locations on either side of a survey ship, instead of directly below as was the case for a traditional single-beam depth-sounder. In practice, sidescan sonars tend to mount both the line array and hydrophones on a towfish, a device that is towed in the water somewhat below the surface behind a survey vessel. As the survey vessel travels, the towfish transmits and listens to the echoes of a series of pulses. The echoes of each pulse are used to build up amplitude versus time plot for each side of the vessel. To adjust for the decline in the strength of echoes due to attenuation, a time-varying gain is applied to the amplitude values so that sea floor features with similar reflectivity have similar amplitudes [19].

Eventually the noise inherent in the system which remains constant becomes comparable to the amplitude of the echo, and the amplified trace becomes dominated by noise. Recording of each trace is usually cut off before this occurs so that the next ping can be transmitted. Each trace is converted into a series of gray shades on a plotting device. The shade of gray is determined by the amplitude. The port and starboard traces are plotted next to and facing each other so that the ship's track is at the center of the picture, and subsequent pairs of traces are plotted in series to build up a picture of the bottom.

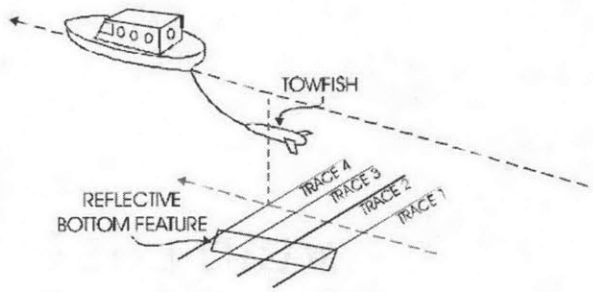

Figure 11: Side Scan Sonar Measuring a Featured Ocean Floor with Four Pings [19]

1.3.4 Limitations of Traditional Side Scan Sonar

In general the sea floor is not always flat. This concept introduces uncertainties in the interpretation of sidescan data. Recall that the only information traditional sidescan sonar provides is a series of echo amplitudes as a function of time. The directivity of the sonar indicates which direction these echoes were coming from. But an observer cannot really be sure of the exact location of the sources of these echoes. In this case, the sonar is surveying over an area that has a sharply sloped bottom.

The first place the circular pulse fronts strike the bottom is not directly below the survey vessel, but at a point on the sloped bottom well off to the side. This is the point that will cause the earliest amplitude on the amplitude versus time trace.

The point directly below the sonar will cause an echo at a slightly later time. At this same time, two points on the slope will be causing an echo. Because they occur at the same time, the echoes of these three widely separated points will be confused on the trace. It is clear that by looking at only the amplitude versus time trace in this example, an observer would not have an accurate picture of the sea floor. Because of this problem, traditional sidescan sonars are really only useful for flat or near-flat bottoms, or when used in combination with bathymetry information [19].

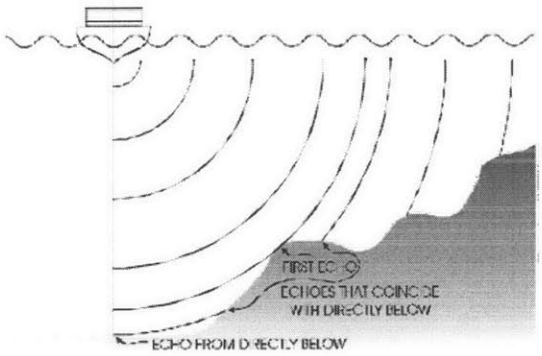

Figure 12: Half of Two-Hydrophone Sidescan Sonar [19]

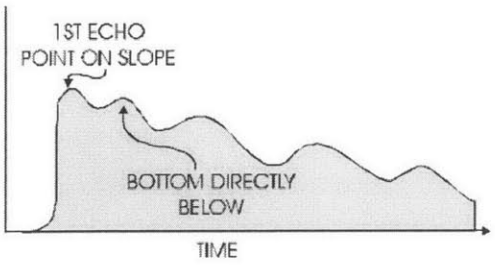

Figure 13: Amplitude vs. Time Plot for the Situation [19]

Chapter 2: Wavelet Transform

2.1 History of Wavelets

Wavelets are a powerful mathematical tool for hierarchical multi-resolution analysis of functions. They have been effectively utilized in many diverse fields, including approximation theory, signal processing, physics, astronomy, and image processing. Wavelets have also been applied to a wide variety of problems in computer graphics.

Before 1930, the main branch of mathematics leading to Wavelets began with Joseph Fourier with his theories of frequency analysis, now often referred to as Fourier Transform.

After 1807, by exploring the meaning of functions, Fourier series convergence, and orthogonal systems, mathematicians gradually were led from their previous notion of frequency analysis to the notion of scale analysis. That is, analyzing $f(x)$ by creating mathematical structures that vary in scale. Constructing a function, shift it by some amount, and change its scale. Apply that structure in approximating a signal. Now repeat the procedure. Take that basic structure, shift it, and scale it again. Apply it to the same signal to get a new approximation. And so on. It turns out that this sort of scale analysis is less sensitive

to noise because it measures the average fluctuations of the signal at different scales. The first mention of Wavelets appeared in an appendix to the thesis of A. Haar (1909). One property of the Haar Wavelet is that it has compact support, which means that it vanishes outside of a finite interval. Unfortunately, Haar Wavelets are not continuously differentiable which somewhat limits their applications [8].

In the 1930s, several groups working independently researched the representation of functions using scale -varying basis functions. By using a scale-varying basis function called the Haar basis function Paul Levy, a 1930s physicist, investigated Brownian motion, a type of random signal. He found the Haar basis function was superior to the Fourier basis functions for studying small complicated details in the Brownian motion. Another 1930s research effort by Littlewood, Paley, and Stein involved computing the energy of a function.

The computation produced different results if the energy was concentrated around a few points or distributed over a larger interval. This result disturbed the scientists because it indicated that energy might not be conserved [14]. The researchers discovered a function that can vary in scale and can conserve energy when computing the functional energy. Their work provided David Marr with an effective algorithm for numerical image processing using Wavelets in the early 1980s between 1960 and 1980 [11].

The mathematicians Guido Weiss and Ronald R.Coifman studied the simplest elements of a function space, called atoms, with the goal of finding the atoms for a common function and finding the "assembly rules" that allow the reconstruction of all the elements of the function space using these atoms. In 1980, Grossman and Morlet, a physicist and an engineer, broadly defined Wavelets in the context of quantum physics [20]. These two researchers provided a way of thinking for Wavelets based on physical intuition.

In 1985, Stephane Mallat gave Wavelets an additional jump-start through his work in digital signal processing. He discovered some relationships between quadrature mirror filters, pyramid algorithms, and orthonormal Wavelet bases. Inspired in part by these results, Y.Meyer constructed the first non-trivial Wavelets [16]. Unlike the Haar Wavelets, the Meyer Wavelets are continuously differentiable; however they do not have compact support. A couple of years later, Ingrid Daubechies used Mallat 's work to construct a set of Wavelet orthonormal basis functions that are perhaps the most elegant, and have become the corner stone of Wavelet applications today [18].

2.2 From Fourier to Wavelet Transform

In 19th century, the French mathematician J.Fourier showed that any periodic function can be expressed as an infinite sum of periodic complex exponential functions. Many years after he had discovered this remarkable property of functions, his ideas were generalized to first nonperiodic functions, and then periodic or non-periodic discrete time signals. It is after this generalization that it became a very suitable tool for computer calculations. In 1965, a new algorithm called the Fast Fourier Transform was developed and the Fourier Transform (FT) became even more popular [27].

The definition of FT is given by:

$$F(w) = \int_{-\infty}^{\infty} f(t) e^{-jwt} dt \qquad (5)$$

And the inverse transform is

$$f(t) = \int_{-\infty}^{\infty} F(w) e^{jwt} dw \tag{6}$$

The information provided by the integral, corresponds to all time instances, since the integration is from minus infinity to plus infinity over time. This is why the Fourier Transform is not suitable if the signal has time varying frequency, i.e., the signal is non-stationary. This means that the Fourier Transform tells whether a certain frequency component exists or not [27].

This information is independent of where in time this component appears. Therefore a linear time frequency representation called Short Time Fourier Transform was introduced. In the Short Time Fourier Transform, the signal is divided into small enough segments that the segments can be assumed to be stationary. For this purpose, a window function is chosen. The width of this window must be equal to the segment of the signal where its stationarity is valid.

The important feature of Short Time Fourier Transform is the width of the window that is used. The width is also called the support of the window. The narrower we make the window, the better the time resolution, and the better the assumption of stationarity, but the poorer the frequency resolution.

The problem with Short Time Fourier Transform (STFT) is the fact whose roots go back to what is known as the Heisenberg's Uncertainty Principle [23]. This principle originally applied to the momentum and location of moving particles can be applied to time-frequency information of a signal.

Simply, this principle states that one cannot know the exact time-frequency representation of a signal, i.e., one cannot know what spectral components exist at what instances of times. What one can know are the time intervals in which certain bands of frequencies exist, which is a resolution problem. Therefore the problem is a result of choosing a window function, once and for all, and uses that window in the entire analysis. The answer, of course, is application dependent.

If the frequency components are well separated from each other in the original signal, then we may sacrifice some frequency resolution and go for good time resolution, since the spectral components are already well separated from each other. However, if this is not the case, then it is a very difficult to find a good window function [14,16,19]. Although the time and frequency resolution problems are results of a physical phenomenon and exist regardless of the transform used, it is possible to analyze any signal by using an alternative approach called Wavelet transform. Wavelet Transform analyses the signal at different frequencies with different resolutions.

Every spectral component is not resolved equally as was the case in the Short Fourier Transform. Wavelet Transform is designed to give good time resolution and poor frequency resolution at high frequencies and good frequency resolution and poor time resolution at low frequencies. This approach makes sense especially when the signal at hand has high frequency components for short durations and low frequency components for long durations.

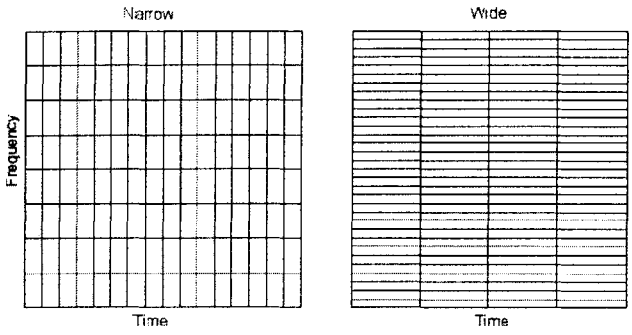

Figure 14: Narrow and Wide-Windowed Fourier Tiling of the Time-Frequency Plane

There are two main differences between the Short Fourier Transform and the Continuous Wavelet Transform:

1. The Fourier transforms of the windowed signals are not taken [21,27].

2. The width of the window is changed as the transform is computed for every single spectral component, which is probably the most significant characteristic of the Wavelet transform.

The continuous Wavelet transform of a signal, f (t), is defined as follows,

$$C(a,b) = \int_{-\infty}^{+\infty} f(t)\psi^*_{a,b}(t)dt \qquad (7)$$

where

$$\psi_{a,b}(t) = a^{-\frac{1}{2}}\psi(\frac{t-b}{a}) \tag{8}$$

is a window function called the mother Wavelet, a is a scale and b is a translation. The term Wavelet means a small wave. The smallness refers to the condition that this (window) function is of finite length (compactly supported). The wave refers to the condition that this function is oscillatory [27]. The term mother implies that the functions with different region of support that are used in the transformation process are derived from one main function, or the mother Wavelet. In other words, the mother Wavelet is a prototype for generating the other window functions.

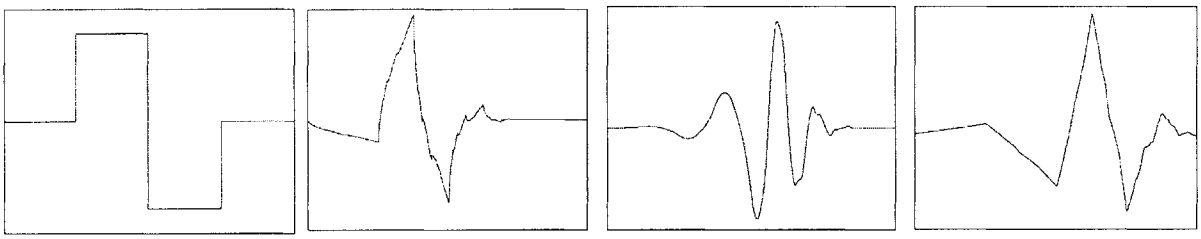

Figure 15: Example for Different Families of Wavelets

The term translation is used in the same sense as it was used in the Short Time Fourier Transform. It is related to the location of the window, as the window is shifted through the signal.

This term, obviously, corresponds to time information in the transform domain. However, we do not have a frequency parameter, as we had before for the Short Time Fourier Transform. Instead we have a scale [12,15,27].

Scaling, as a mathematical operation, either dilates or compresses a signal. Smaller scales correspond to dilated (or stretched out) signals and large scales correspond to compressed signals. However, in the definition of the Wavelet transform, the scaling term is used in the denominator, and therefore, the opposite of the above statements holds. The relation between scale and frequency is that low scales correspond to high frequencies and high scales to low frequencies.

Regardless of the dimensions of the boxes, the areas of all boxes, both in Short Fourier Transform and Wavelet Transform are the same and determined by Heisenberg's inequality [27].

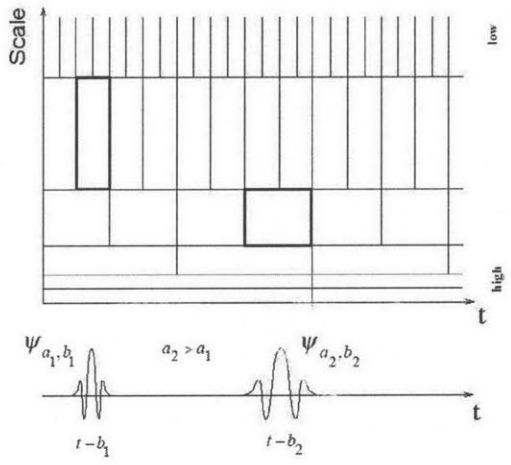

Figure 16: The Two Basic Wavelet Operations-Scale and Translation

2.3 Three Types of Wavelet Transform

Three types of analysis are: the continuous, semi discrete and discrete time analysis. The distinction among the various types of Wavelet Transform depends on the way in which the scale and shift parameters are discretized. In this section we will look closer at the three type of these possibilities.

2.3.1 Continuous Wavelet Transform

At the most redundant end, one has the Continuous Wavelet Transform (CWT), defined and explained in the previous chapter. For Continuous Wavelet Transform the parameters vary in a continuous fashion. The Wavelets must satisfy the admissibility condition; in particular it must have zero mean.

The condition is also crucial to be Continuous Wavelet Transform invertible on its range. The inverse transform is given by relation [27]

$$f(t) = \frac{1}{K_\psi} \int_{-\infty}^{+\infty} \int_{0}^{+\infty} C_{(a,b)} * \psi_{(a,b)} \frac{dadb}{a^2} \tag{9}$$

It also satisfies the admissibility condition:

$$K_\psi = \int_{-\infty}^{+\infty} \frac{|\psi(w)|^2}{w} dw < \infty \tag{10}$$

Where ψ is defined in *equation (8).*

The algorithm of CWT can be described as follows:

1. Take a Wavelet and compare it to a section at the start of the original signal.

2. Calculate a coefficient *C(a,b)*, that represents how closely correlated the Wavelet is with this section of the signal. The higher *C* is, the greater the similarity. Note that the results will depend on the shape of the Wavelet you choose.

3. Shift the Wavelet to the right and repeat steps 1 and 2 until you've covered the whole signal.

4. Scale the Wavelet and repeat steps 1 through 3.

2.3.2 Semi-discrete Wavelet Transform

In practice, it is often more convenient to consider Wavelet Transform for some discretized values. The transform will be reversible if the corresponding set of templates defines a Wavelet frame. In other words, the Wavelet must be designed such that [27]

$$A\|f\|^{2\leq}\sum_{a,b}|< f,\psi(a,b) >|^2 \leq B\|f\| \tag{11}$$

Where A and B are two positive constants called frame bounds and $< f,\psi(a,b) >$ is the signal showing the function and the mother Wavelet. Notice that we must still integrate to get Wavelet coefficients, the $f(t)$ is still a continuous function.

2.3.3 Discrete Wavelet Transform

Here, we have discrete function *f(n)* and the definition of discrete Wavelet transform (DWT) is given by [8].

$$C(a,b) = C(j,k) = \sum_{n \in Z} f(n)\psi_{j,k}(n) \tag{12}$$

where

$$\psi_{j,k}(n) = 2^{-j/2}\psi(2^{-j}n - k) \tag{13}$$

The parameters a,b in equation(8)are defined in such a way that a $= 2^{j}$,b$= k2^{j}$. Sometimes the analysis is called dyadic as well. The inverse transform is defined as: [8]

$$f(n) = \sum_{j \in Z}\sum_{k \in Z} C(j,k)\psi_{j,k}(n) \tag{14}$$

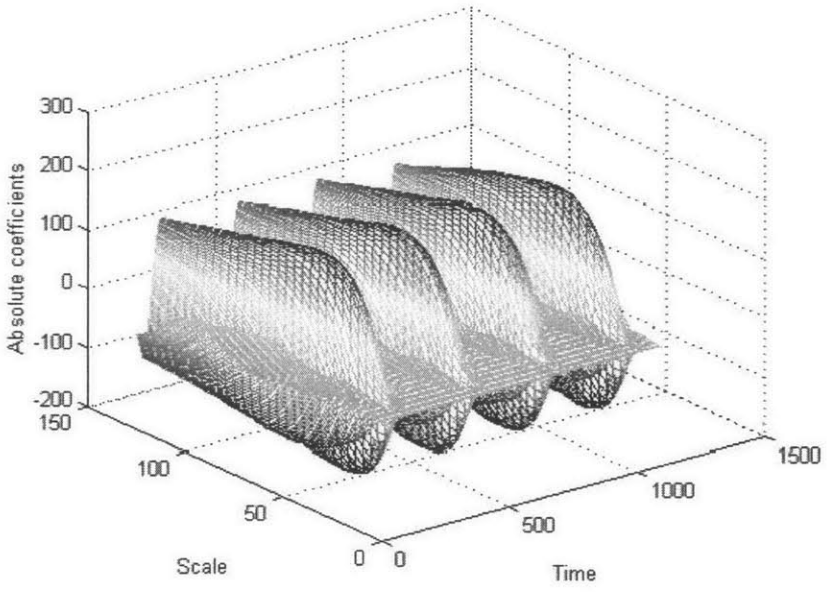

Figure 17: Example of Continuous Wavelet Analysis

2.4 Implementation

In this section, brief information about choosing the appropriate Wavelet decomposition will be provided. Knowing the application process, one needs to know which implementation should be used. First step is to understand, if energy of the signal is finite, not all values of decomposition are needed to exactly reconstruct the original signal, provided that you are using a Wavelet that satisfies the admissibility condition. A continuous-time signal is entirely characterized by the knowledge of the discrete transform. In such cases, discrete analysis is sufficient and continuous analysis is redundant. When the signal is recorded in continuous time or on a very fine time grid, both types of analysis are possible.

Continuous analysis is often easier to interpret, since its redundancy tends to reinforce the traits and makes all information more visible. This is especially true of very subtle information. The analysis gains in "readability" and in ease of interpretation what it loses in terms of space saving. Discrete analysis ensures space-saving coding and is sufficient for the synthesis.

2.5 Multiresolution Analysis

In this section, most important feature of Wavelet analysis which is multi-resolution will be discussed. This discussion will show how discrete signals are synthesized by beginning with a very low resolution signal and successively adding on details to create higher resolution versions, ending with a complete synthesis of the signal at the finest resolution. This is known as Multiresolution analysis (MRA). MRA is the heart of Wavelet analysis. In order to understand the idea fully, we must first discuss some elementary operations that can be performed on signals. Simple examples are given below to describe MRA. Figure (17) demonstrates a good example of Multiresolution feature of Wavelets applied on images [25].

Given two signals $\mathbf{f} = (f1, f2, \ldots, fN)$ and $\mathbf{g} = (g1, g2, \ldots, gN)$, we can perform the following elementary algebraic operations:

Addition and Subtraction: The sum f + g of the signals f and g is defined by adding their values:

$$\mathbf{f} + \mathbf{g} = (f1 + g1, f2 + g2, \ldots, fN + gN). \tag{15}$$

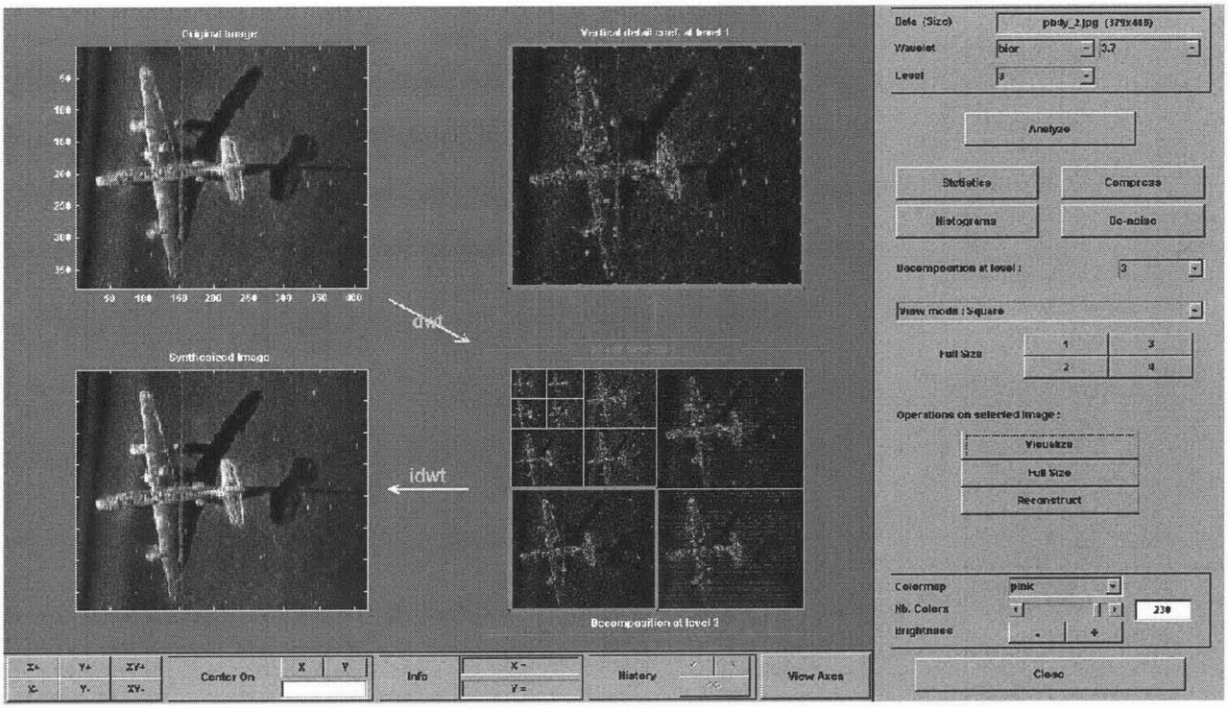

Figure 18: Wavelet Decomposition Using Matlab Wavelet Toolbox

Their difference **f** − **g** is defined by subtracting their values:

$$\mathbf{f} - \mathbf{g} = (f1 - g1, f2 - g2, \ldots, fN - gN) \tag{16}$$

Constant multiple: A signal **f** is multiplied by a constant c by multiplying each of its values by c. That is,

$$c\,\mathbf{f} = (cf1, cf2, \ldots, cfN). \tag{17}$$

For example, by repeatedly applying the addition operation, we can express a signal [9,10,12]

$$\mathbf{f} = (f1, f2, \ldots, fN) \quad \text{as follows:} \tag{18}$$

$$\mathbf{f} = (f1, 0, 0, \ldots, 0) + (0, f2, 0, 0, \ldots, 0) + \ldots + (0, 0, \ldots, 0, fN). \tag{19}$$

Then, by applying the constant multiple operation to each of the signals on the right side of this last equation, we obtain

$$\mathbf{f} = f1(1, 0, 0, \ldots, 0) + f2(0, 1, 0, 0, \ldots, 0) + \ldots + fN(0, 0, \ldots, 0, 1). \tag{20}$$

[7] This formula is a very natural one; it amounts to expressing **f** as a sum of its individual values at each discrete instant of time. If we define the elementary signals $\mathbf{V}01, \mathbf{V}02, \ldots, \mathbf{V}0N$ as

$$\begin{aligned} \mathbf{V}01 &= (1, 0, 0, \ldots, 0) \\ \mathbf{V}02 &= (0, 1, 0, 0, \ldots, 0) \\ &\vdots \\ \mathbf{V}0N &= (0, 0, \ldots, 0, 1) \end{aligned} \tag{21}$$

then the last formula for **f** can be rewritten as

$$\mathbf{f} = f_1\mathbf{V}01 + f_2\mathbf{V}02 + ... + f_N\mathbf{V}0N \tag{22}$$

This equation is called the natural expansion of a signal f in terms of the natural basis of signals **V**01,**V**02, . . . ,**V**0N [9,10,12].

This formula shows that the signal f can be expressed as the sum of two signals that we shall call the first averaged signal and the first detail signal. That is, we have

$$\mathrm{f} = \mathrm{A}1 + \mathrm{D}1 \tag{23}$$

where the signal **A**1 is called the *first averaged signal* and is defined by

$$\mathbf{A}^1 = \left(\frac{a_1}{\sqrt{2}}, \frac{a_1}{\sqrt{2}}, \frac{a_2}{\sqrt{2}}, \frac{a_2}{\sqrt{2}}, \ldots, \frac{a_{N/2}}{\sqrt{2}}, \frac{a_{N/2}}{\sqrt{2}}\right) \tag{24}$$

and the signal **D**1 is called the *first detail signal* and is defined by

$$\mathbf{D}^1 = \left(\frac{d_1}{\sqrt{2}}, \frac{-d_1}{\sqrt{2}}, \frac{d_2}{\sqrt{2}}, \frac{-d_2}{\sqrt{2}}, \ldots, \frac{d_{N/2}}{\sqrt{2}}, \frac{-d_{N/2}}{\sqrt{2}}\right) \tag{25}$$

2.5.1 Multiple Levels

The second level of a MRA of a signal **f** involves expressing **f** as

$$\mathbf{f} = \mathbf{A}^2 + \mathbf{D}^2 + \mathbf{D}^1 \tag{26}$$

Here A^2 is the second averaged signal and D^2 is the second detail signal.

$$\mathbf{A}^1 = \mathbf{A}^2 + \mathbf{D}^2 \tag{27}$$

This formula expresses the fact that computing the second averaged signal A^2 and second detail signal D^2 simply consists of performing a first level MRA of the signal A^1 Because of this, it follows that the second level averaged signal A^2 satisfics [9,10,12]

$$\mathbf{A}^2 = (\mathbf{f} \cdot \mathbf{V}_1^2)\mathbf{V}_1^2 + (\mathbf{f} \cdot \mathbf{V}_2^2)\mathbf{V}_2^2 + \cdots + (\mathbf{f} \cdot \mathbf{V}_{N/4}^2)\mathbf{V}_{N/4}^2 \tag{28}$$

and the second level detail signal D^2 satisfies

$$\mathbf{D}^2 = (\mathbf{f} \cdot \mathbf{W}_1^2)\mathbf{W}_1^2 + (\mathbf{f} \cdot \mathbf{W}_2^2)\mathbf{W}_2^2 + \cdots + (\mathbf{f} \cdot \mathbf{W}_{N/4}^2)\mathbf{W}_{N/4}^2 \tag{29}$$

2.6 Discrete Images

The 2D data that we shall be working with are discrete images. A discrete image **f** is an array of M rows and N columns of real numbers:

$$\mathbf{f} = \begin{pmatrix} f_{1,M} & f_{2,M} & \cdots & f_{N,M} \\ \vdots & \vdots & \ddots & \vdots \\ f_{1,2} & f_{2,2} & \cdots & f_{N,2} \\ f_{1,1} & f_{2,1} & \cdots & f_{N,1} \end{pmatrix} \tag{30}$$

The values of **f** are the $M\,N$ real numbers $\{ f_{j,k} \}$. It should be noted that the way in which the values of **f** are displayed in the array on the right side of the equation, is not the most commonly used one. We chose to display the values of **f** in this way because it corresponds well with the case where **f** is an array of sample values:

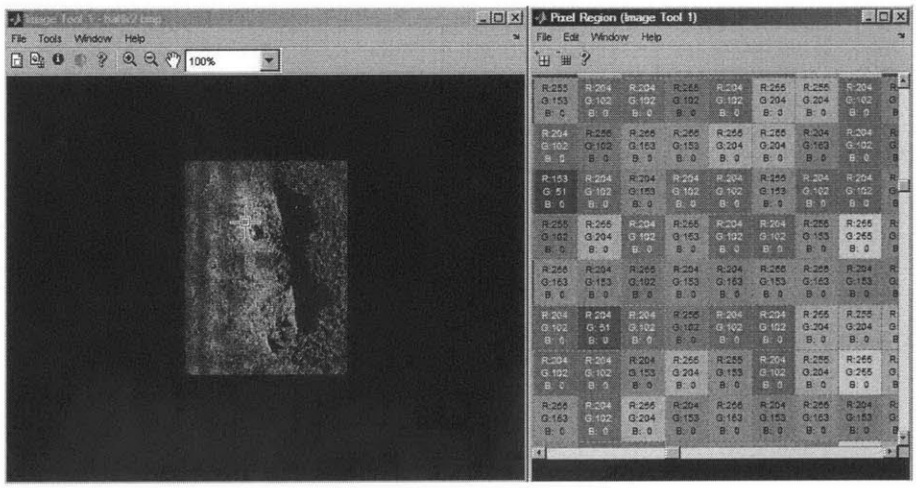

Figure 19: Discrete Image Representation in Matlab Image Processing Toolbox

$$\mathbf{f} = \begin{pmatrix} g(x_1, y_M) & g(x_2, y_M) & \cdots & g(x_N, y_M) \\ \vdots & \vdots & \ddots & \vdots \\ g(x_1, y_2) & g(x_2, y_2) & \cdots & g(x_N, y_2) \\ g(x_1, y_1) & g(x_2, y_1) & \cdots & g(x_N, y_1) \end{pmatrix} \tag{31}$$

of a function $g(x, y)$ at the sample points (x_j, y_k) in the Cartesian coordinate plane. Just as with discrete 1D signal, it is frequently the case that a discrete 2D image is obtained from samples of some function $g(x, y)$. It is often helpful to view a discrete image in one of two other ways. First, as a single column consisting of M signals having length N [9,10],

$$\mathbf{f} = \begin{pmatrix} \mathbf{f}_M \\ \vdots \\ \mathbf{f}_2 \\ \mathbf{f}_1 \end{pmatrix} \tag{32}$$

with the rows being the signals

$$\begin{aligned} \mathbf{f}_M &= (f_{1,M}, f_{2,M}, \ldots, f_{N,M}) \\ &\vdots \\ \mathbf{f}_2 &= (f_{1,2}, f_{2,2}, \ldots, f_{N,2}) \\ \mathbf{f}_1 &= (f_{1,1}, f_{2,1}, \ldots, f_{N,1}). \end{aligned} \tag{33}$$

Second, as a single row consisting of N signals of length M, written as columns,

$$\mathbf{f} = (\mathbf{f}^1, \mathbf{f}^2, \ldots, \mathbf{f}^N) \tag{34}$$

with the columns being the signals

$$\mathbf{f}^1 = \begin{pmatrix} f_{1,M} \\ \vdots \\ f_{1,2} \\ f_{1,1} \end{pmatrix}, \mathbf{f}^2 = \begin{pmatrix} f_{2,M} \\ \vdots \\ f_{2,2} \\ f_{2,1} \end{pmatrix}, \ldots, \mathbf{f}^N = \begin{pmatrix} f_{N,M} \\ \vdots \\ f_{N,2} \\ f_{N,1} \end{pmatrix} \tag{35}$$

2.7 2-D Wavelet Transforms

Similar ideas in one-dimensional case can be considered for two-dimensional Wavelet transform. In the following section, the representation of images in matrix form has been described. It is now easier for us, to think about the two-dimensional Wavelet after understanding the matrix format of images. As we know, images are composed of pixels, which include color density information. Knowing the color density distribution, which is called histogram, in terms of numerical values, images can be represented by matrices. Each column or each row of the matrix can also be visualized by a series of numbers. These discrete values then considered as signals, so it's a transition from image processing to signal processing. Fourier transforms or in our case Wavelet transform that can be used in signal processing, can also be used in image processing depending on the results that we are looking for. For instance, Cosine transform mostly to be used in filtering, Fourier transform for compression and Wavelet transform can be used in both cases successfully. As we mentioned above similar ideas applies in two dimensional and one dimensional case [5].

A 2D Wavelet transform of a discrete image can be performed whenever the image has an even number of rows and an even number of columns. A 1-level Wavelet transform of an image **f** is defined, using any of the 1D Wavelet transforms that we have discussed, by performing the following two steps [27]:

Step 1: Perform a 1-level, 1D Wavelet transform, on each row of **f**, thereby producing a new image.

Step 2: On the new image obtained from Step 1, perform the same 1D Wavelet transform on each of its columns.

It is not difficult to show that Steps 1 and 2 could be done in reverse order and the result would be the same. A 1-level Wavelet transform of an image **f** can be symbolized as follows:

$$\mathbf{f} \longmapsto \begin{pmatrix} \mathbf{h}^1 & | & \mathbf{d}^1 \\ \hline \mathbf{a}^1 & | & \mathbf{v}^1 \end{pmatrix} \tag{35}$$

where the sub images h^1, d^1, a^1, and v^1 each have M/2 rows and N/2 columns. We shall discuss the nature of each of these sub images. The sub image a1 is created by computing trends along rows of f followed by computing trends along columns; so it is an averaged, lower resolution version of the image [17].

Since a 1D trend computation is $\sqrt{2}$ times an average of successive values in a signal, and the 2D trend sub image a^1 was computed from trends along both rows and columns, it follows

that each value of a^1 is equal to 2 times an average of a small square containing adjacent values from the image .

A useful way of expressing the values of a1 is as scalar products of the image f with scaling signals, as we did in the 1D case; we shall say more about this later in the section. The h1 sub image is created by computing trends along rows of the image f followed by computing fluctuations along columns [25].

Consequently, wherever there are horizontal edges in an image, the fluctuations along columns are able to detect these edges. This tends to emphasize the horizontal edges. This discussion should make it clear why we shall refer to this sub image as the first horizontal fluctuation. The sub image v^1 is similar to h^1, except that the roles of horizontal and vertical are reversed.

Chapter 3: Image Retrieval Using Wavelet Transform

3.1 Different Algorithm Categorizations in Image Matching

3.1.1 Color Histograms

Color histograms are basically, demonstrates the color density distribution of an image. Histograms are widely used for content-based image retrieval. Their advantages are efficiency, and insensitivity to small changes in camera viewpoint.

However, a histogram is a coarse characterization of an image, and so images with very different appearances can have similar histograms.

Described technique is for comparing images called histogram refinement, which shows additional constraints on histogram based matching. Histogram refinement splits the pixels in a given bucket into several classes, based upon some local property. Within a given bucket, only pixels in the same class are compared.

Histogram refinement can be used to distinguish images whose color histograms are indistinguishable. All images are scaled to contain the same number of pixels M. Discretizing the color space of the image such that there are N distinct (discretized) colors.

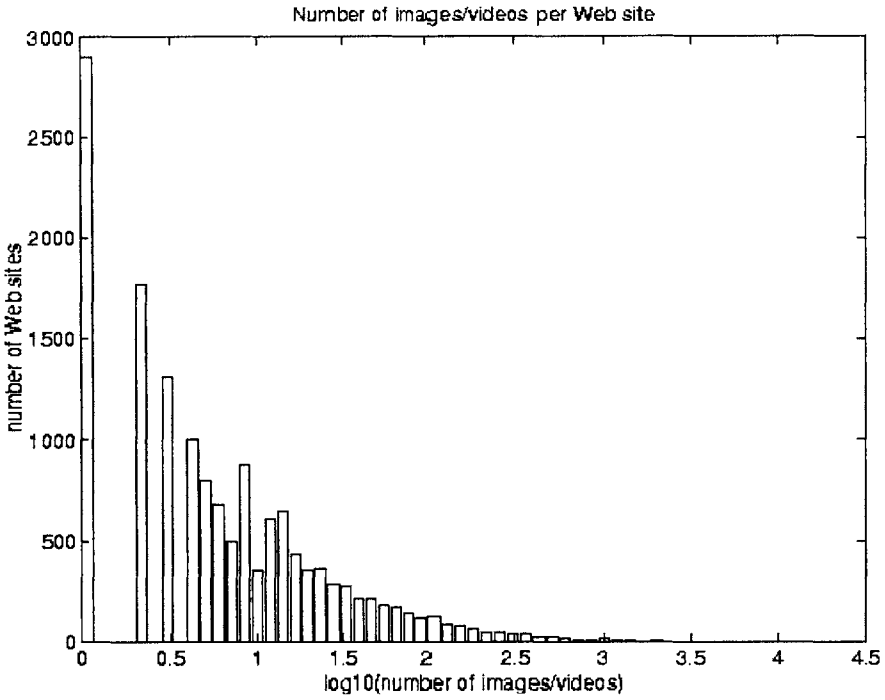

Figure 20: Color Histogram- Color Density Distribution of an Image/Video

A color histogram H is a vector h h1; h2; : : : ; h ni, in which each bucket hj contains the number of pixels of color j in the image. Typically images are represented in the RGB color space, with a few of the most significant bits per color channel. For a given image I, the color histogram HI is a compact summary of the image. A database of images can be queried to the most similar image to I, and can return the image *I0* with the most similar color histogram HI*0*. Color histograms are typically compared using the L1-distance or the L2-distance, although more complex measures have also been considered.

3.1.2 Energy methods

This method, measures energy (possibly weighted) at the output of filter banks as extracted features for texture discrimination. The basic assumption of this approach is that the energy distribution in the frequency domain identifies a texture. Wavelets are core technology in the next generation of still image coding format, JPEG-2000 the choice of Wavelet features enables the implementation of retrieval systems that can work directly in the compressed domain. Other possible transforms are Wavelet packets and Wavelet frames. Finding good similarity measures between images based on some feature set is a challenging task. The ultimate goal is to define similarity functions that match with human perception.

On the other hand, many current retrieval systems take a simple approach by using typically norm-based distances (Euclidean distance) on the extracted feature set as a similarity function. Any reasonable similarity functions defined on the features space should perform well. Sometimes, weighting factors are necessary to normalize extracted features over the entire database to comparable ranges so that they have approximately the same influence on the overall distance.

3.2 Our Approach and Developed Algorithm

Previously we address the problems, in seafloor image similarity measurement and image retrieval of disturbed images (e.g rotated, scaled, skewed) depending on the sonar carrier's elevation from sea bottom or depending on the weather and sea conditions. Considering these problems following algorithm has been implemented. We consider jointly the problems of Feature Extraction, namely multiresolution ability of Wavelet Transform and Similarity Measurement which is basically the statistical approach in image retrieval.

Considering the two related retrieval tasks Feature Extraction and Similarity Measurement as estimation and detection problems, respectively, provides us with defining similarity functions on the feature space. Well-known examples in similarity measurement are the use of histograms to capture the distribution of image features such as color etc. Generalized Gaussian density functions to represent texture images in the Wavelet domain in previously in related works. The model parameters are estimated using a method of moment matching, and the similarity function is again defined as weighted Euclidean distances on extracted model parameters.

The statistical approach fits nicely into this case, since a texture image is often regarded as a realization of an underlying stochastic process. In the end, we will briefly discuss how such approach can be applied to other features and integrated into more general image retrieval systems.

The outline will be as follows. In the next section, we set up the content based image retrieval problem in a general statistical framework. Then we will apply this framework to the Wavelet-based image retrieval application where Wavelet coefficients in each sub band are independently modeled by a generalized Gaussian distribution (GGD). We will also present experimental results in following section, which indicate a significant improvement in the retrieval of images of seafloor objects.

3.3 Algorithm Settings

In general settings, our algorithm is invariant to discussed image registration problems. We are initially searching for invariants, which will allow us to distinguish similar objects in seafloor environment (e.g. rocks, sand ripples) The problem of searching for the top N images similar to a given query image from a database of total M images (N << M) can be formulated as a multiple hypotheses problem. The query image Iq is represented by its data set x = (x1, x2, . . . , xL), which is typically obtained after a preprocessing stage. Each candidate image in the database $Ii : i = 1, 2, \ldots, M$ is assigned with a probability Hi. The goal is to select, depending on rankings, among the M possible hypotheses the N best ones that describe the data x from the query image. To select the N top matches from those M hypotheses we can use the multiple hypotheses testing argument. We first choose the best one among the M possible hypotheses {H1, H2, . . . , HM}, and then we choose the next best one among the remain (M−1) hypotheses. We assume that all prior probabilities of the hypotheses are equal and for each recursive step the optimum rule is to choose the hypothesis with the highest likelihood among the possible ones.

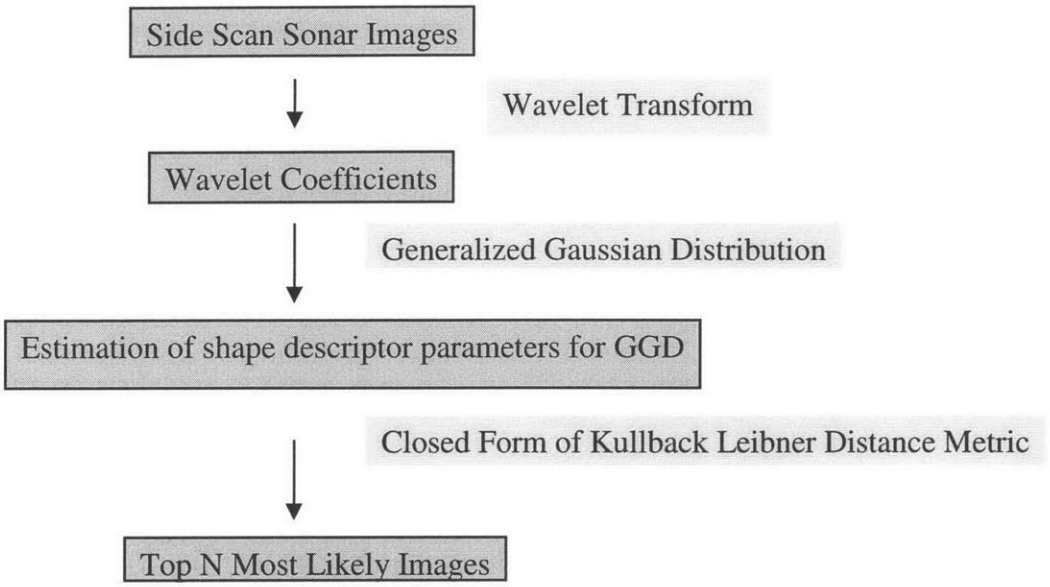

Figure 21: Side Scan Sonar Images Retrieval Algorithm Setup

This is referred to as the maximum likelihood (ML) selection rule. In this approach, the conditional probability density p(X|Hi) is modeled by a member of a family of probability density functions (PDF's), denoted by p(X; θi) where θi is a set of model parameters. With this setting, the extracted feature for the image Ii is the estimated model parameter ˆθi, which is computed in the Feature Extraction (FE) step. We denote the space of model parameters as Θ. Consider the query data x = (x1, x2. . . xL) as an independent and identically distributed sequence from the model p(X; θq) of the query image [1,8].

$$\frac{1}{L}\log p(x;\theta_i) = \frac{1}{L}\sum_{j=1}^{L}\log p(x_j;\theta_i) \tag{36}$$

$$= \int p(x;\theta_i)\log p(x;\theta_i)dx \tag{37}$$

This can be seen as equivalent to minimizing the Kullback-Leibner distance (KLD) or the relative entropy between the two PDF's

$$\left|D(p(X;\theta_q)\|p(X;\theta_i))\right| = \int p(x;\theta_q)\log\frac{p(x;\theta_i)}{p(x;\theta_i)}dx \tag{38}$$

For consistent estimator, we could employ the ML estimator, which means that for the query image, it computes:

$$\theta_q = \arg\max \log p(x;\theta) \tag{39}$$

The similarity measurement step can be computed entirely on the estimated model parameters, which are typically small in size, so that it can meet the timing constraint. That is another advantage of our algorithm. The method allows us the use of any feature data and statistical models for indexed images. We will use Kullback-Leibner distance to compare images, which will be extensively discussed on the next section. Combining the KLD's from multiple data sets, such as from different channels or feature sets, we can use the chain rule which states that the KLD between two joint PDF's p(X, Y) and q(X, Y) is

$$D(p(X,Y)\|q(X,Y)) = D(p(X)\|q(X)) + D(p(Y\ |X)\|q(Y|X)) \tag{40}$$

3.4 Wavelet Subband Coefficients and Generalized Gaussian Density

Comparing to the statistical framework, it is relatively easier, if some preprocessing is applied to the images. This kind of preprocessing can be carried out by transformation of image pixel values into a suitable space where simple models with a small number of parameters can describe the data. In our case, RGB images transform to the gray scale images to reduce the computational complexity. Wavelets have recently emerged as an effective tool to analyze texture information and since Wavelets are used in major future image compression standards and are also shown to be prominent in searching for images based on color and shape, a Wavelet-based texture retrieval system can be used effectively with a compression system and retrieval systems. The assumption is that the energy distribution in frequency domain identifies texture; traditional approaches computed energies of Wavelet subbands as texture features. Commonly, L1 and L2 norms are used as measures [8].

More specifically, given the Wavelet coefficients xi,1, xi,2, . . . , xi,L at the i-th subband, typically the following two values are used as features:

$$f_i = \frac{1}{L}\sum_{j=1}^{L}\left|x_{i,j}\right| \tag{41}$$

$$f_i = (\frac{1}{L}\sum_{j=1}^{L} x^2{}_{i,j})^{1/2} \tag{42}$$

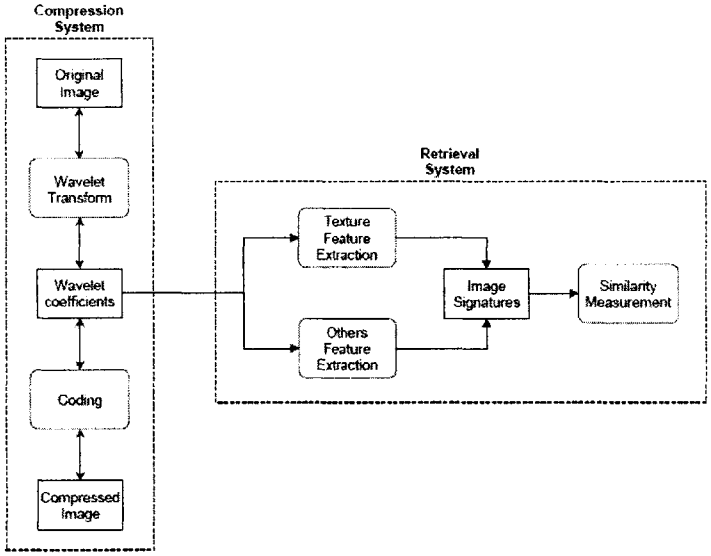

Figure 22: Block Diagram of the Wavelet-Based Image Matching Algorithm [1].

A natural extension of the energy method is to model a texture by the densities of Wavelet subband coefficients. Probability Density Function approximation for the density of coefficients at a particular subband produced by various type of Wavelet transforms may be achieved by two parameters of the generalized Gaussian density (GGD) which are:

$$p(x;\alpha,\beta)=\frac{\beta}{2\alpha\Gamma(1/\beta)}e^{-(|x|/\alpha)^{\beta}} \tag{43}$$

where;

$$\Gamma(z)=\int_0^{\infty}e^{-t}t^{z-1}dt, z>0 \tag{44}$$

Here α models the width of the PDF standard deviation, while β is proportional to the decreasing rate of the peak. α is referred to as the scale parameter however β is called the shape parameter. The GGD model contains the Gaussian PDF's as special cases, using β = 2 and β = 1, respectively. It is time to introduce the ML estimator for GGD. Lets define the likelihood function of the sample x = (x1. . . xL) [1,8].

$$L(x;\alpha,\beta)=\log\prod_{i=1}^{L}p(x_i;\alpha,\beta) \tag{45}$$

Where α and β are parameters to be estimated. In this case the following likelihood equations have a unique root in probability, which is maximum-likelihood estimator;

$$\frac{\partial L}{\partial\alpha}=-\frac{L}{\alpha}+\sum_{i=1}^{L}\frac{\beta|x_i|^{\beta}\alpha^{-\beta}}{\alpha}=0 \tag{46}$$

$$\frac{\partial L}{\partial \beta} = \frac{L}{\beta} + \frac{L\psi(1/\beta)}{\beta^2} - \sum_{i=1}^{L} (\frac{|x_i|}{\alpha})^{\beta} \log(\frac{|x_i|}{\alpha}) = 0 \tag{47}$$

Ψ is gamma function which can be described as $\Psi(z) = \Gamma'(z)/\Gamma(z)$. From above equations, one can pull the α from the first equation which is

$$\alpha = \left(\frac{\beta}{L} \sum_{i=1}^{L} |x_i|^{\beta} \right)^{1/\beta} \tag{48}$$

If we plug α into second equation we will end up with an equation in terms of only the parameter β which can be solved by using any kind of numerical root finding algorithm.

$$1 + \frac{\psi(1/\beta)}{\beta} - \frac{\sum_{i=1}^{L} |x_i|^{\beta} \log|x_i|}{\sum |x_i|} + \frac{\log(\frac{\beta}{L} \sum_{i=1}^{L} |x_i|^{\beta})}{\beta} = 0 \tag{49}$$

In our case we used Newton-Raphson root finding algorithm, which basically uses derivatives of the function and searches for the trends where the sign of the function changes. For these kind of numerical root finding algorithm , a good initial guess , takes us faster solution.

Chapter 4: STATISTICAL FRAMEWORK

4.1 Probability and Likelihood

The concept of likelihood, introduced by Sir R. A. Fisher, is closely related to the more common concept of probability. Probability is generally observing the events. For example, for an unbiased coin, the probability of observing heads is 0.5 for every toss. This is taken to mean that if a coin were tossed a large number of times then we would expect, on average, to find half of the time the coin landed heads, half of the time tails.

There are certain laws of probability that allow us to make inferences and predictions based on probabilistic information. For example, the probabilities of different outcomes for a certain event must always add up to 1: if there is a 20% chance of rain today, there must be an 80% chance of no rain. Another very common law is that if two events are independent of one another, then the probability of certain pairs of outcomes will be the product of the two outcomes by themselves: if we toss a coin twice, the probability of getting 2 heads is 0.5 times 0.5 = 0.25 [28].

4.2 Parameters and Distributions

The probability of observing events, such as the outcome of a toss of a coin is assuming some kind of model, even in this simple case. In the case of a coin, the model would state that there is some certain, fixed probability for the particular outcomes. This model would have one parameter, p the probability of the coin landing on heads. If the coin is fair, then p=0.5. The probability of observing an event is given specific parameter values for the model. In this simple case, if p =0.5, then the probability of the coin landing heads on any one toss is also 0.5. In the case of this simple example, it does not seem that one gained very much. It seems to be merely calling what was previously a simple probability the parameter of a model. However, this way of thinking provides a very useful framework for expressing more complex problems [28].

4.3 Model Fitting and Maximum Likelihood

We will start with introducing the concept of likelihood. If the probability of an event X dependent on model parameters *p* is written

$$P(X \mid p) \tag{50}$$

then we would talk about the likelihood

$$L(p \mid X) \tag{51}$$

that is, the likelihood of the parameters given the data. For most sensible models, we will find that certain data are more probable than other data. The aim of maximum likelihood estimation is to find the parameter values that makes the observed data most likely. This is because the likelihood of the parameters given the data is defined to be equal to the probability of the data given the parameters. If we were in the business of making predictions based on a set of solid assumptions, then we would be interested in probabilities - the probability of certain outcomes occurring or not occurring [28].

However, in the case of data analysis, we have already observed all the data: once they have been observed they are fixed, there is no 'probabilistic' part to them anymore (the word data comes from the Latin word meaning 'given'). We are much more interested in the likelihood of the model parameters that underlie the fixed data. To re-iterate, the simple principle of maximum likelihood parameter estimation is this: find the parameter values that make the observed data most likely. How would we go about this in a simple coin toss experiment? That is, rather than assume that p is a certain value (0.5) we might wish to find the maximum likelihood estimate (MLE) of p, given a specific dataset [28]. Beyond parameter estimation, the likelihood framework allows us to make tests of parameter values. For example, we might want to ask whether or not the estimated p differs significantly from 0.5 or not.

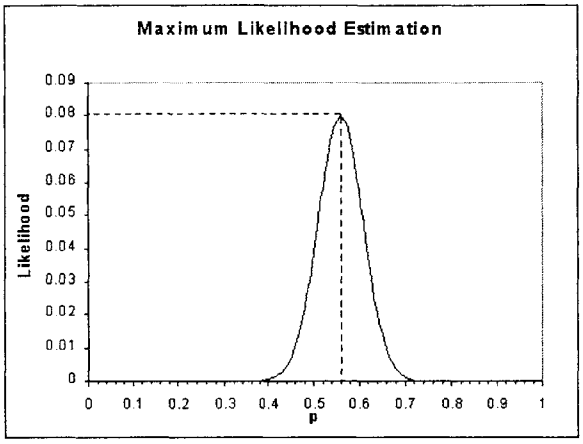

Figure 23: Maximum Likehood Estimation [28]

4.4 Generalized Gaussian Density Distribution

The normal distribution is a convenient model of quantitative phenomena in the natural and behavioral sciences. A variety of psychological test scores and physical phenomena like photon counts have been found to approximately follow a normal distribution. While the underlying causes of these phenomena are often unknown, the use of the normal distribution can be theoretically justified in situations where many small effects are added together into a score or variable that can be observed.

The normal distribution also arises in many areas of statistics: for example, the sampling distribution of the mean is approximately normal, even if the distribution of the population the sample is taken from is not normal. In addition, the normal distribution maximizes [28]

information entropy among all distributions with known mean and variance, which makes it the natural choice of underlying distribution for data summarized in terms of sample mean and variance. The normal distribution is the most widely used family of distributions in statistics and many statistical tests are based on the assumption of normality.

In probability theory, normal distributions arise as the limiting distributions of several continuous and discrete families of distributions The modeling of probability density distributions of coefficients produced by Wavelet transform subbands may be efficiently achieved by adaptively adjusting the parameters of a generalized Gaussian density (GGD) function. The estimation of GGD parameters may be carried out either by use of the moment method, entropy matching or in a maximum likelihood (ML) framework. In all these approaches, the existence and the uniqueness of the parameters are based on asymptotic behavior, that is the sample is supposed to be sufficiently large. However, in signal and image processing, we often deal with small samples for which the existence of the parameters is unknown. After recalling the two main approaches to compute GGD parameters, we give a necessary and sufficient condition for the existence of the parameters in a ML framework.

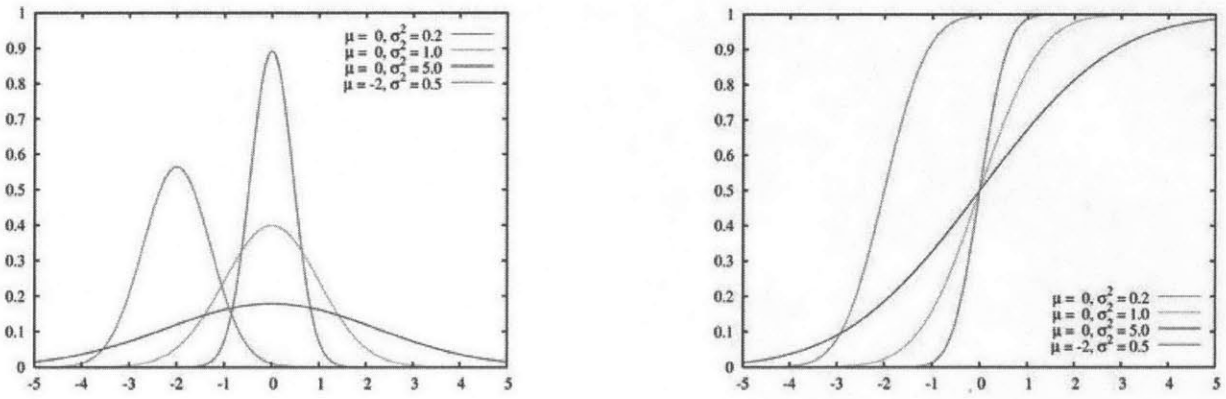

Figure 23: Generalized Gaussian Density and the Cumulative Distribution Function

There are various ways to specify a random variable. The most visual is the probability density function, which represents how likely each value of the random variable is. The cumulative distribution function is a conceptually cleaner way to specify the same information, but to the untrained eye its plot is much less informative. Equivalent ways to specify the normal distribution are: the moments, the cumulate, the characteristic function, the moment-generating function, and the cumulate-generating function. Some of these are very useful for theoretical work, but not intuitive.

4.5 Calculation of Shape Parameter in GGD

We will start this section by introducing first and second mode of distributions which are shown as

$$m_1 = \int |x| P_{\alpha,\beta}(x) dx \tag{53}$$

and the second moment is shown as

$$m_2 = \int |x^2| P_{\alpha,\beta}(x) dx \tag{54}$$

Where $P_{\alpha,\beta}$ is defined in equation (43)

through

$$\alpha = m1\frac{\Gamma(\frac{1}{\beta})}{\Gamma(\frac{2}{\beta})} \qquad \text{and} \qquad \beta = F^{-1}\left(\frac{{m_1}^2}{m_2}\right) \qquad (56), (57)$$

where

$$F(x) = m1\frac{\Gamma^2(\frac{2}{x})}{\Gamma(\frac{3}{x})\Gamma(\frac{1}{x})} \qquad \text{And once more} \qquad \Gamma(t) = \int_0^{+\infty} e^{-x}x^{t-1}dx \qquad (58),(59)$$

Considering the discrete case moment often represented as summations;

$$m_1 = \frac{1}{L}\sum_i |x_i| \qquad \text{and} \qquad m_2 = \frac{1}{L}\sum_i |x_i|^2 \qquad (60),(61)$$

As shown before

$$\alpha_{Discrete} = m_1\frac{\Gamma(\frac{1}{\beta})}{\Gamma(\frac{2}{\beta})} \qquad \text{and} \qquad \beta_{Discerete} = F^{-1}\left(\frac{{m_1}^2}{m_2}\right) \qquad (63),(64)$$

4.6 Kullback-Leibner Similarity Measurement Method

In this section, we will discuss the Kullback-Leibner similarity measurement algorithm which is mainly used for matching entropy given for different system. This method is pretty common in image matching considering statistical approach. Having the 2-D Wavelet transforms of each sea floor images, we defined our Generalized Gaussian Distributions which are namely Wavelet subband coefficients, in terms of only two parameters which are α and β.

Substituting α and β parameters inside the Kullback-Leibner distance, we end up with having a closed form of distance metric which can arranged as [1,8,28]

$$D(p(;\alpha_1,\beta_1)\|p(;\alpha_2,\beta_2)) = \log(\frac{\beta_1\alpha_2\Gamma(1/\beta_2)}{\beta_2\alpha_1\Gamma(1/\beta_1)}) + (\frac{\alpha_1}{\alpha_2})^{\beta_2}\frac{\Gamma((\beta_2+1)/\beta_1)}{\Gamma(1/\beta_1)} - \frac{1}{\beta_1} \quad (75)$$

The similarity measurement between two Wavelet subbands can be computed very effectively using the model parameters.

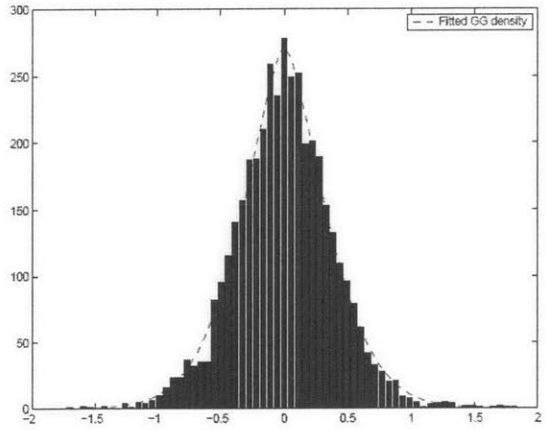

Figure 24: Wavelet Subband Coefficient Histogram Fitted with a Generalized Gaussian Density

Furthermore, applying with the reasonable assumption that Wavelet coefficients in different subbands are independent, the overall similarity distance between two images is precisely the sum of Kullback-Leibner Distances. In this case α's are defined as scale factors and the β's are defined as shape parameters, relating to curvature or inclination of the Generalized Gaussian distribution's trend.

4.7 Experiments with Wavelets GGD Model

We used 48 seafloor (man-made objects) images, obtained from Klein online side scan sonar image library [7]. During the experimentation we used Matlab Image Processing toolbox the trim the images to (512x512) gray scale images and rotated, skewed and stretched them to

simulate the variation of Auv's height above the sea floor, or the pitch-roll variations. We used Bior3.7 (from bio-orthogonal Wavelet family) Wavelet basis which has a unique shape goes very well with any kind of random signals. Having a large library with different shape of objects, we need to choose a basis that can fit robustly to any kind of randomness. In image processing and Wavelets applications of images, bior3.7 type Wavelet basis has been widely used previously.

Furthermore, to eliminate the effect of common range in the gray level of sub images from a same original image and to make the retrieval task less biased, each sub image was individually normalized to zero mean and unit variance before the processing.

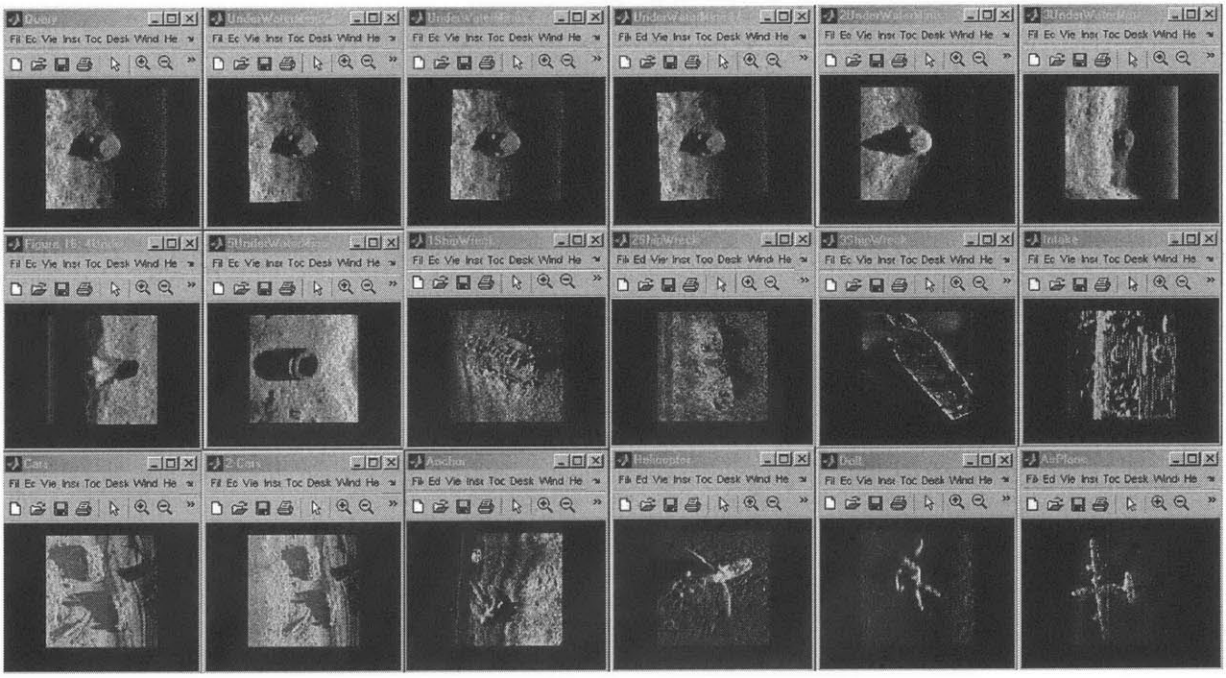

Figure 25: Experimental Results (Image Retrieval Using Wavelet Transform)

4.8 Matching Effectiveness

Above method has been implemented in a large library consist of side-scan sonar images. We also pre-processed these images using Matlab image processing toolbox. Then we evaluated the performance in terms of the average rate of retrieving relevant images as a function of the number of top retrieved images.

Qualitative evaluation of our method was carried out by visually examining the images of retrieval results. Understanding the similarity between two or more images, by using computer based algorithms relatively has much more complexity and difficulties, comparing with the judgment of bare human eye.

However working with Wavelet subband coefficients allows us to reduce that computational complexity since we only end up with compressed images and reduced number of Wavelet coefficients rather than having whole image and with all information within.

If we take a look at the experimental results which is shown in Figure (25), one can easily can see that, candidate object which is placed at the upper left corner, an underwater mine. Our aim is to recognize that object, depending on our sample image library. Having a large map of sea-floor, we previously detect this object (image filtering will be discussed at the conclusions and future work section) and crop our area of interest from the whole map. After applying developed algorithm and sorting the results in qualitative sense, consecutive seven images are likewise underwater mines. An interesting point in this case is all of the consecutive seven images have rotation problem. Another aspect of the results is, two images placed at the lower left corner are sorted meaning that they are most likely similar images. Previously, we applied skew and stretch problem on these images. They show us how our algorithm responded on these kinds of image matching problems. This subject will be covered more at the conclusion

and the future work sections. During the numerical calculations, Matlab image processing toolbox and Matlab Wavelet Toolbox have been used.

Chapter 5: Conclusions

Summary and Recommendations for Future Research

Wavelet based similarity measurement algorithm has been successively applied in image matching. Wavelet sub band coefficients have been used previously for solving the image matching problems. Same idea used in this work for solving the rotation, skew, stretch problems in side scan sonar images. Depending on the various affects of the water waves, and the wind, motion of the vessel which carries the side-scan sonar will tend to move to six fundamental directions. (Pitch, roll, yaw, heave, etc.) Motions of the vessel can be reduced by using sea-keeping instrumentation. Damping of these motions will surely help to reduce the effects of randomness.

After processing each image, (query and sample image library) using wavelet decomposition, collected wavelet subband coefficients matched successively between similar images. Another question is how to determine the area of interest and eliminating rocks and sand ripples. Distinguishing the man-made objects is a question of image filtering problem. As a future work in this case, developing another algorithm to fix this problem can be considered. During the experiments, various disadvantages have been determined. Working on the ship wrecks, we have seen that some ship wrecks are buried under the sand and it can be partially observed.

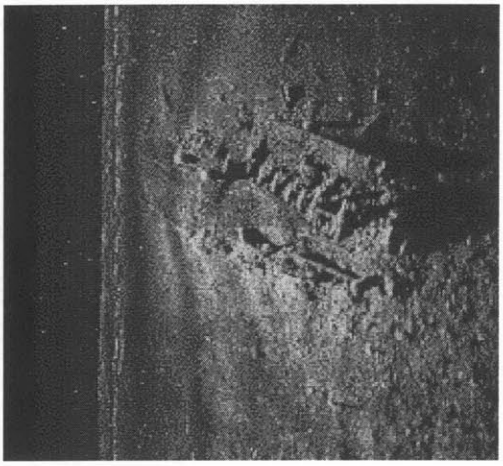

Figure 26: Partially Buried Ship Wreck in Sea Bed [7]

In this case, partial matching problem can be a good solution. Another solution can also be increasing our sample image library as possible as we can, will surely fix this problem. Different kinds of scanning patterns will also cause partial matching problems which are shown in the figure (27).

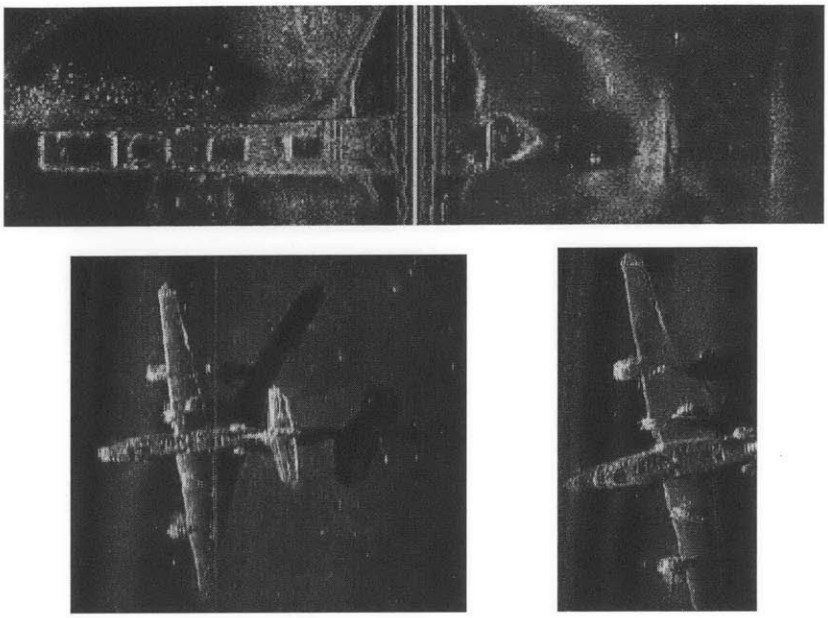

Figure 27: Partial Matching Problems [7]

\

Another disadvantage is depending upon motion of the side-scan sonar; there might be shadow problems, which can also be solved by applying image filtering techniques.

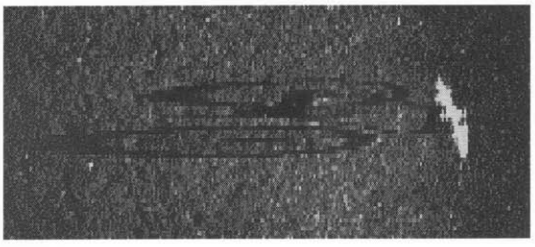

Figure 28: Shadowing Problems in Side Scan Sonar Images. [7]

References

[1] Do, M.N.; Vetterli, M., "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance," Image Processing, IEEE Transactions on , vol.11, no.2pp.146-158, Feb 2002

[2] Chang, T., Jay Kuo, C.C., 1993. Texture analysis and classification with tree-structured wavelet transform. IEEE Trans. Image Process. 2 (4), 429-440.

[3] M. Unser. Texture classification and segmentation using wavelet frames. IEEE Transactions on Image Processing, 4(11):1549--1560, 1995

[4] E.P.Simoncelli and W.T.Freeman, The steerable pyramid, a flexible architecture for multi-scale derivative computation Proc 2nd IEEE Int'l Conf on Image Processing, Washington, DC. Oct 1995.

[5] B.S.Manjuanth and W.Y.Ma, Texture features for retrieval of image data , IEE Trans, PAMI'1996

[6] S.C.Zhu,Y.N.WU and D.Mumford, Frame ,filters, random filed. Maximum entropy, towards a undefined theory for texture modelling, Int'JComp Vision1999

[7] Web site: http://www.l-3klein.com/image_gallery/image_gallery.html

[8] M,DO and M.Vetterli Wavelet-Based texture retrieval using generalized Gaussian density and Kullback Lienbler Distance,submitted to IEEE Trans'Image Proc1999

[9] Charles E. Jacobs, Adam Finkelstein, David H. Salesin, Fast Multi-resolution Image Querying, University of Washington, SIGGRAPH95. 1995

[10] Ibrahim El Rube, Maher Ahmed, Mohamed Kamel, Affine Invariant Multi-Scale Wavelet Based Shape Matching Algorithm Proceedings of the 1st Canadian Conference on Computer and Robot Vision (CRV'04) - Volume 00 Pages: 217 - 224 , 2004

[11] Marius Tico,Eero Immonen, Pauli Kuosmanen, Finger Print Recognition Using Wavelet Features, IEEE Transactions on Pattern Analysis and Machine Intelligence Volume 25 , Issue 8 (August 2003) Pages: 1009 – 1014,Tampere, Finland.

[12] Pawan Jain, Wavelet Based Multi-resolution Histogram for Fast Image Retrieval, TENCON 2003. Conference on Convergent Technologies for Asia-Pacific Region

15-17Oct.2003 Volume:2,81-585Vol.2 ISSN:ISBN: 0-7803-8162-9 India Institute of Technology, Bombay
[13] Jane You, I EEE , and Prabir Bhattacharya, IEEEA Wavelet-Based Coarse-to-Fine Image Matching Scheme in a Parallel Virtual Machine EnvironmentPublicationDate:Sep2000 Volume:9 ,Issue:9 page(s):1547-1559 ISSN: 1057-7149

[14] James Ze Wang, Gio Wiederhold, Oscar Firschein, Sha Xin Wei, "Wavelet-Based Image Indexing Techniques with Partial Sketch Retrieval Capability," adl, p. 13, Fourth International Forum on Research and Technology Advances in Digital Libraries (ADL '97), 1997.

[15] R. Barber, W. Equitz, M. Flickner, W. Niblack, D. Petkovic, and P. Yanker. Efficient query by image content for very large image databases. In COMPCON, pages 17--19. IEEE Computer Society Press, Los Alamitos, CA, 1993.

[16] M. abramovitz and I.A Stegun, Handbook of Mathematical Functions, Dover Publications, New York, 1972.

[18] B.Aiazzi L. Alparone and S. Bartoni, Estimation Based on entropy Matching for Generalized Gaussian PDF Modeling, IEEE Signal Publication Date: Jun 1999 Volume: 6, Issue: 6 On page(s): 138-140 ISSN: 1070-9908

[19] P. Cervenka and D. de Moustier, Sidescan Sonar Image Processing Techniques. IEEE Journals of Oceanic Engineering, vol. 18, pp. 108-122, 1993.

[20] K. A. Birney and T. R. Fischer. On the modeling of DCT and subband image data for compression. IEEE Transactions on Image Processing, 4(2):186--193, 1995.

[21] A. Laine and J. Fan, "Texture classification by wavelet packet signatures," IEEE Trans. Pattern Recognit. Machine Intell., vol. 15, pp. 1186-1191, 1993.

[22] J. R. Smith and S.-F. Chang, "Transform features for texture classification and discrimination in large image databases," in Proc. IEEE Int. Conf. Image Processing, 1994

[23] B. S. Manjunath and W. Y. Ma, "Texture features for browsing and retrieval of image data," IEEE Trans. Pattern Recognit. Machine Intell., vol. 18, pp. 837-842, Aug. 1996.

[24] R.L.Joshi and T.R. Fischer Comparison of Generalized Gaussian and Laplacian Modeling in DCT Image coding, IEEE trans image Processing

[25] S.Mallat A Theory for Multiresolution Signal Decomposition: the Wavelet Presentation IEEE, Trans

[26] P.Moulin and J Liu Analysis of Multiresolution Image Denoising Schemes using Generalized Gaussian and Complexity Prior, IEEE Trans Info Theory vol pp 45

[27] Wavelets and Filter Banks , Gilbert Strang, Nguyen April 1996

[28] http://statgen.iop.kcl.ac.uk/bgim/mle/sslike_1.html