

Building Qualitative Elevation Maps From Side Scan Sonar Data For Autonomous Underwater Navigation

D. Langer, M. Hebert

The Robotics Institute,
Carnegie Mellon University,
Pittsburgh, PA 15213

Abstract

Deriving a terrain model from sensor data is an important task for the autonomous navigation of a mobile robot. This paper describes an approach for autonomous underwater vehicles using a side scan sonar system. First, some general aspects of the type of data and filtering techniques to improve it are discussed. We then proceed to derive an estimated bottom contour, using a geometric reflection model and information about shadows and highlights. Several techniques of surface reconstruction and their limitations are presented. We also describe a method of feature extraction which is important for future data matching/fusion procedures.

1 Introduction

Underwater automation is becoming an increasingly important area of research in robotics owing to a growing field of potential applications. The goal of underwater automation is to build Autonomous Underwater Vehicles (AUV) that are able to navigate through an unknown environment, build maps, and act within the environment without outside intervention. Like any autonomous system, AUVs must have the capabilities of cognition, perception, and action. Promising solutions have been proposed for cognition [11], and for low-level navigation [6]. Perception, however, remains an open issue owing partly to the nature of the sensors, and partly to the nature of the underwater world.

In this paper, we address the problem of perception for an AUV, that is the building of representations of the environment from sensor data. A sonar sensor is an obvious choice for underwater machine vision applications since so far systems based on light (laser, camera) do not compare in performance. Building models from sonars for terrestrial navigation has been studied extensively [5]. For underwater navigation, many models have been proposed for representing the terrain [7] and for gathering terrain representations in a higher level data structure [3]. The problem of processing the sensor readings to construct those representations remains an issue due to the nature of the sensors [10]. In this paper, we investigate algorithms for constructing a representation of the bottom contour, the 'qualitative elevation map', that is suitable for autonomous mapping and navigation. The representation can be included in a comprehensive map representation system such as the ones described in [11] or [3]. Our emphasis, however, is on constructing reliable low-level terrain representation from sensor

data. Some previous work in this area has been done by *Cuschieri and Hebert* [4] and *Stewart* [10]. *Rigaud and Marce* developed in [9] a method of acoustic data accumulation/fusion in order to determine the absolute location of an ROV.

Currently we use data obtained by a side scan sonar system. The analog output data from the system is digitized and transformed into images of standard dimensions of 450 rows x 500 columns. Each row represents one scan of the sea bottom. The intensity of the returned echoes is recorded as intensity of each pixel at a certain instant in time as the sound wave propagates along the bottom. The intensity returned depends mainly on: Angle of incidence of sonar beam on bottom surface, Reflectance properties of bottom surface and Attenuation of ultrasound in sea water.

Taking into account the resolution of the sensor and velocity of sound in water, each pixel represents 15 cm in column direction. Right now it is assumed that the sonar fish travels along a straight line. Thus adding several scans builds up an image by rows, the translation between rows being 25 cm. Hence each pixel represents a projected area of dimensions 25 cm x 15 cm (row x col). Fig. 5 shows the digitized sonar data.

Section 2 describes methods, we have implemented for filtering the raw data and extracting features. Section 3 describes methods for computing relative elevations and approximate surface contours, using a particular surface reflection model.

2 Image Processing

Before attempting to reconstruct surfaces from sonar images, the raw images must be processed. This image processing stage has three goals: noise reduction, feature extraction, and shadow detection.

2.1 Noise reduction

The images obtained tend to be very noisy (Fig. 5). The reason is that a side scan sonar does not measure only the TOF of the first echo received for computing the shortest distance to a certain point in space, but records all echoes. Hence multiple echoes from a single scatterer are recorded at different instances in time and for each scan lead to false intensity peaks in space. Multiple echoes from several scatterers may add at the same point in space and time and thus make the situation even worse (speckle noise).

Standard image operators, like gradient operators tend to fail if applied to this type of data unless some prefiltering is done.

Our approach to noise reduction is to apply a series of filtering operations, each of which addresses a particular type of noise. This approach gives us control over the amount of smoothing, noise reduction, and data enhancement that is achieved.

The following filters are currently applied:

2-D 3x3 Median Filter: The purpose of this filter is to remove isolated spurious measurements (spikes or dropouts) without disturbing the rest of the data.

Modified Median Filter with variable window size : The previous median filter removes one pixel wide spikes and dropouts but does not remove spurious data that occupies a larger region. Using a larger median filter would remove larger regions but would also eliminate important details in the image. To get around this problem, a modified median filter is applied, removing only valleys, i.e. negative extrema, but leaving peaks intact. The idea being that peaks are likely to correspond to a point of interest, whereas valleys are mostly drop outs or arise due to speckle noise. Also, peaks due to noise generally tend to be much lower and will be removed by filling up the valleys between them. The filter is applied several times; increasing the window size each time.

Graduated Non Convexity (GNC) Filter : Once isolated spurious returns are removed, the next stage is to remove some of the speckle noise that is essentially a random distribution in the data. A standard smoothing operator would do the job but it would also smooth out discontinuities in the data that may correspond to important features. A good compromise between smoothing and edge preservation is achieved by the GNC algorithm [1]. The general concept of this algorithm is to find a function that minimizes an energy that combines the distance between the function and the input data, a measure of the smoothness of the function, and a measure of the average extend of the discontinuities. The result is such that edge discontinuities are preserved while the remaining surface is smoothed.

The GNC Filter is applied once again for each column to ensure greater consistency between rows. The effect of the applied filters are shown in Figure 1.

2.2 Shadow Detection

Shadows are an important source of information in a sonar image. Similar to optic shadows cast by objects when illuminated from one direction, acoustic shadows are formed behind large elevations or objects on the sea floor. Sound waves travel generally along a straight line over short distances in the same medium. The principles of acoustic shadow formation are therefore essentially the same as in geometric optics. Hence no distinct echoes can be observed in an acoustic shadow zone since it is shielded by the preceding contour. Therefore shadows are an important source of information in a sonar image, since they indicate the presence of an object or a change in bottom contour. Ideally the signal level within a shadow region should be zero. In reality, however, this is not true because of the previously discussed problems with noise.

The occluding edge in front of a shadow region is in most cases characterized by a sharp drop in signal level. The algorithm uses a 1-D window and operates on each row individually. It classifies a pixel as belonging to an occluding edge if a sharp signal drop is

detected and signal values within the prospective shadow region are below a certain threshold. The shadow is then defined as the region between the occluding edge and the point where the sign of the slope of the signal turns from negative to positive. A lot of false shadows are nevertheless detected by this procedure. They are eliminated as follows: By using a histogram, the probable signal values for real shadows are determined, since here values are more evenly distributed than in other regions. Because of noise in the data that has not been smoothed by the applied filters, it is also possible that an occluding edge is detected even though there is no real shadow. These false shadows are eliminated by comparing the average signal values preceding and following the occluding edge. Since the sonar fish moves quite close to the bottom, shadows tend to be long for areas of distinct elevation. Hence, if the signal level does not stay low for quite some distance with respect to the signal level preceding the occluding edge, the detected shadow is classified as being false and removed.

A clustering algorithm then assigns each shadow detected in a row to a particular shadow group in the image. The results are shown in Figure 6.

Finally it should be mentioned that shadows can also be used to locate and predict the shape of features on the sea bottom. Especially, if we have more than one viewing direction of the same area, approximate boundaries of objects or elevations can be extracted.

3 Estimation of Bottom Contour

3.1 Introduction

Recovering the bottom contour from side scan sonar data poses a highly underconstraint problem, which may not lead to a unique solution without making strong assumptions about the observed surface. Some of the difficulties encountered are discussed in the following paragraphs.

The sonar data provides essentially three parameters:

1. The orientation, x-y position and depth Z_F of the sonar fish for each scan line.
2. The range R to a surface patch.
3. The intensity (echo level) of the received echo for that particular surface patch.

It should be noted that the known range R is the slant range to a particular surface patch, but that the angle of inclination φ to the horizontal is not known. Hence the reflecting surface patch can lie on any part of an annular section with range R as radius. The horizontal and vertical extension of the annular section is determined by the horizontal and vertical field of view of the sonar at range R , i.e. the aperture of the sonar (see Fig. 2).

Combining and matching different data sets, as mentioned before, provides a means to eliminate wrong data and gain more knowledge about the area of interest. However, since we only know the slant range for each surface patch but not the direction of measurement, it is virtually impossible to find corresponding data points between two images accurately. In order to avoid this correspondence problem we would actually need a fixed x-y-z position in space for each data point. Hence we will use the return signal intensity (echo level) at a certain range to compute the angle

of inclination α of the surface normal for a particular reflecting surface patch. α is measured with respect to the line of sight to the respective surface patch (see Fig. 3). Once α is known, we can attempt to find the actual bottom contour. The reflectance model and method of surface reconstruction is described in detail in the following sections.

3.2 Surface reflection model

Sound waves behave in many ways similar to electromagnetic waves like light waves. They propagate according to Fermat's principle and obey Snell's law of refraction. In the case of reflection we can distinguish between diffuse and specular reflection. The wavelength of sound is much longer compared to the wavelength of light. Hence surfaces appear much smoother to sound than to light since the ability to resolve two neighboring peaks on a rough surface depends on angle of incidence α and wavelength λ . A surface will reflect an incident sound wave specularly when the following condition is true:

$$H \cos \alpha \leq \lambda/4 \quad (1)$$

where H is the period of the surface irregularities.

Also for specular reflection, angle of incidence is equal to angle of reflection. Further, for a rigid reflecting surface, pressure (amplitude) of incident wave is equal to pressure of reflected wave.

However, real surfaces are generally non-rigid reflectors and a part of the incident sound energy will be transmitted into the new medium. The amount of incident intensity reflected depends on a material property of the two media and is given by the reflection coefficient k_r :

$$k_r = \frac{I_r}{I_i} = \left[\frac{\rho_2 c_2 - \rho_1 c_1}{\rho_2 c_2 + \rho_1 c_1} \right]^2 \quad (2)$$

where I_r is the reflected intensity, I_i is the incident intensity, ρ_j and c_j are the density of medium j and the velocity of sound in medium j , respectively.

Source and receiver are at the same position for a side scan sonar. Therefore the sensor only receives the intensity reflected back in a direction towards it. To determine the angle of reflection from the received intensity we use a model similar to the one proposed by Torrance-Sparrow, described in [8]. The model is based on geometrical optics but it also gives a fairly good approximation of the behavior of sound waves for the environment considered here.

The present model takes into account two components of reflection, diffuse and specular. The diffuse lobe is represented by the Lambertian model. A Lambertian surface scatters incident energy uniformly in all directions. The energy radiated is therefore determined by the incident energy which is proportional to $\cos \alpha$, where α is the angle of incidence. The specular lobe represents specular reflection from a rough surface. The surface is assumed to be a collection of facets whose slopes are described by a Gaussian distribution model. The surface thus described, scatters incident energy in the form of a lobe that is symmetric about the specular direction (see [8] for details). Since source and receiver are at the same position for a side scan system, α is angle of incidence and also angle of reflection. The relationship between reflected

intensity I_r and α can then be written as a linear combination of the two:

$$I_r = C_{dl} \cos \alpha + C_{sl} G(\alpha) \frac{1}{\cos \alpha} \exp \left(-\frac{\alpha^2}{2\sigma_\alpha^2} \right) \quad (3)$$

The constants C_{dl} and C_{sl} denote the relative strength of the diffuse and specular component and are also a measure of the surface albedo k_r . σ_α is a factor that describes surface roughness. $G(\alpha)$ is a geometric attenuation factor (refer to [8] for a detailed explanation). For the sonar sensor geometry, it is given by:

$$G(\alpha) = \min(1, 2 \cos^2 \alpha) \quad (4)$$

(See Fig. 3 for a definition of α).

3.3 Inclination of the surface normal

The sonar actually measures the time of flight for a reflected sound pulse which is then converted to distance:

$$d = c_w \frac{\tau}{2} \quad (5)$$

where c_w is the velocity of sound in water. A change in c_w affects the propagation of sound in water and the accuracy of our measurement. Parameters affecting c_w , sorted in order of importance, are: Temperature, Salinity and Depth (Pressure). (Refer to [12]).

Since the sonar fish is towed relatively close to the sea bottom, variations due to changes in pressure can be neglected. Moreover the presently used side scan sonar covers only a short range, which means that temperature and salinity distributions will be fairly uniform and hence can be neglected too. However, a salinity gradient affecting measurements will exist in areas where freshwater and sea water mix.

Therefore it can be assumed that in our case the sound waves travel along a fairly straight line and attenuation of the signal is mainly due to geometrical spreading losses.

This means, that the initial intensity I_o of the pulse sent at time $t = 0$ does not stay constant over the whole range but changes proportional to a $\frac{1}{R^2}$ fall off with increasing distance R from the sensor. The attenuation is then given by the following formula:

$$I_r = \frac{1}{4\pi R^4} I_o k_A (2R \tan(\eta/2)) \quad (6)$$

I_r being the intensity of the received signal for an ideal reflector, taking into account geometric spreading losses, and η being the horizontal field of view of the sonar. k_A is a factor adjusting for the area of the scatterer as it appears to the sonar [2].

To keep the signal values of the received intensity within a convenient range, they are generally converted to decibel:

$$E_r = 10 \log I_r, \quad E_o = 10 \log I_o$$

where, for a specific range, I_r is the received signal intensity, E_r is the echo level of the received signal from the sonar and E_o the initial echo level.

It is assumed that the time varying gain of the side scan system performs an ideal correction for geometric attenuation on the received signal according to eqt. 6. Then I_r^* is the received intensity, scaled to unity and compensated for spreading losses:

$$I_r^* = 10^{(0.1E_r^*)}, \quad E_r^* = K_g \cdot (E_r - E_o) \quad (7)$$

where K_s is a factor to adjust the scale of the digitized signal. Substituting I_s^* at a certain range into Equation 3, the angle of inclination α between the sensor's line of sight and the surface normal at that point can be estimated. Since Equation 3 cannot be directly resolved for α , a look up table is calculated for positive values of alpha (Equation 3 is symmetric). The constants C_d , C_s and σ_α are surface reflectance parameters, which are estimated at the moment. The estimation is done by visually comparing our knowledge about the actual surface with the surface reconstructed by the methods described later. The reflectance parameters are then adjusted as necessary. Eventually, we intend to derive tables containing these parameters for different types of surfaces.

3.4 Reconstruction of the approximate surface contour

The following parameters are now available to help in estimating the bottom contour:

- Position and orientation of sensor in world coordinates
- Angle of inclination of the surface normal
- Range corresponding to each computed angle of inclination

Nevertheless the problem is still underconstrained and even though we have computed the angle α , we do not have a unique orientation of the surface normal for each surface patch. The surface normals possible for a particular angle α will actually form a cone around the line of sight. Since the actual reconstructed surface is a combination of individual patches and there still is a number of orientations for each patch, there will also be a large number of different surfaces possible. A global minimization procedure imposing smoothness constraints on the surface can thus be used to find the final estimate of the bottom contour. An alternative solution is to assume boundary conditions, specifically the slope of the terrain at the first recorded echo, and to propagate the angular constraint to the entire map. We now describe two such propagation algorithms: a 1-D algorithm that integrates the angular constraints along each scan line separately, and a 2-D algorithm that integrates the constraint over the whole map at once.

At the starting point, which corresponds to the closest measured range, the following assumptions are made:

- The nearest (first) echo recorded corresponds to the depth below the sensor, i.e. angle φ between line of sight and x-axis = 90° , and $Z_E = Z_0$.
- The surface normal at this point is along the line of sight of the sensor, i.e. $\alpha = 0$. This assumption means that all the orientations are computed relative to the orientation of the surface at the first recorded data point.

Also all surface normals are taken to lie in the x-z plane, with α being measured counterclockwise from the surface normal (see Fig. 4).

Now for a point at position X_i and range R_i the gradient $\tan \theta$ of the surface in x-direction is calculated from angle α .

$$\varphi = \arcsin \frac{Z_E}{R}, \quad \theta = \pi/2 - (\alpha + \varphi)$$

The surface at X_i , R_i can thus be represented by a straight line segment as shown in Figure 5. To find the next point at X_{i+1} , an intersection is found between this straight line and a circle with radius R_{i+1} , the sensor being at the center. The advancing wave front is modeled here as the section of a circle rather than as an annular section. This simplification is made since the horizontal field of view of the sensor is only 2° . Similarly all following n points for a scan line are found and the resulting 1-D surface segment is represented by a combination of line segments. Shadows are included as shown in Fig. 4. The line of sight φ is known at the occluding edge and the extraction of the shadow length was described in Section 2. The method is repeated for each scan line.

Since there are generally no large surface discontinuities in a local area between neighboring scan lines, the surface built up by the 1-D segments will be reasonably smooth. The method is obviously prone to error accumulation and its performance will also decline with the percentage of surface inclined in y-direction. Nevertheless if there is an inclination in y-direction, it will still show up as a general tendency.

The method was tested on simulated data, consisting of three intersecting planes inclined at different angles to each other. The results obtained here proved to be reasonable enough as to serve as an input to further processing. The algorithm was then used on real data and the result is shown in Fig. 7. Since the algorithm is only 1-D, an inconsistency between rows can be observed, especially the further we get away from the starting point owing to error accumulation.

In order to reduce these errors, we can introduce a further geometric constraint and extend the present method to 2-D:

Let us represent the surface in cylindrical coordinates: $\varphi = \varphi(R, y)$, where

$$x = R \cos \varphi, \quad y = y, \quad z = R \sin \varphi \quad (8)$$

At the data points (R, y) , the dot product between the surface normal \vec{n} and the range vector \vec{R} in the x-z plane is given by

$$\vec{R} \cdot \vec{n} = \|\vec{R}\| \cos \alpha \quad (9)$$

Note that the starting points for all vectors \vec{R} correspond to the respective positions of the sonar fish.

Expressing \vec{n} as a function of the partial derivatives of the surface and substituting cylindrical coordinates in the above equation, we finally get:

$$\tan^2 \alpha \left(\frac{\partial \varphi}{\partial R} \right)^2 - \left(\frac{\partial \varphi}{\partial y} \right)^2 = \frac{1}{R^2} \quad (10)$$

If we can solve the above differential equation, we can obtain the azimuth angle φ as a function of R and y , which are known from the data. In practice, we have to work with discrete values of R and y , that is: $R = \delta R, 2 \delta R, \dots, n \delta R$, and $y = \delta y, 2 \delta y, \dots, n \delta y$. The functions $\varphi(R, y)$ and $\alpha(R, y)$ become arrays φ_{ji} and α_{ji} where j and i correspond to $y = j \delta y$ and $R = i \delta R$ respectively. By using finite differences, $\frac{\partial \varphi}{\partial R}$ becomes $\frac{\varphi_{j+1,i} - \varphi_{j,i}}{\delta R}$ and $\frac{\partial \varphi}{\partial y}$ becomes $\frac{\varphi_{j+1,i} - \varphi_{j,i}}{\delta y}$. Substituting the finite differences into equation 10 and observing that $R = i \delta R$, we obtain:

$$\tan^2 \alpha \left(\frac{\varphi_{j+1,i} - \varphi_{j,i}}{\delta R} \right)^2 - \left(\frac{\varphi_{j+1,i} - \varphi_{j,i}}{\delta y} \right)^2 = \left(\frac{1}{i \delta R} \right)^2 \quad (11)$$

Equation 11 shows that φ_{j+1} depends only on φ_{ji} and φ_{j+1i} . Assuming that the values at the boundary, φ_{j0} are known, this suggests an algorithm to propagate the values of φ by, starting at column $i = 1$, iteratively computing φ_{ji+1} for all j 's as a function of the values computed at step i . Equation 11 gives a formal way to ensure both that the angular constraints are satisfied everywhere, and that the resulting surface is smooth, *i.e.* it has consistent values between scan lines.

3.5 Remarks

There still remain some additional problems in the approaches introduced so far. The main problem is that some of the parameters of the reflectance model are difficult to estimate. However, it is possible to put bounds on the acceptable values on the parameters which would then be converted into bounds on the terrain shape through the reconstruction process. Finally, the reconstruction algorithms tend to produce unsmooth surfaces due to the discretization of the signal and the approximation of derivatives by first-order differences. One improvement would be to include a smoothness constraint in the reconstruction algorithm in a way similar to [1].

4 Conclusion

Side Scan Sonar data tends to be very noisy, making it difficult to achieve good results using standard image processing techniques. Using the described filtering methods, however, we were able to obtain reasonably smooth data, which still contained all the important information. From the filtered data, we can then extract features, like high intensity points, that can be used in future to match different sets of data. By using information about shadows, high intensity points and reflective properties, we have started developing algorithms to compute boundaries of objects and an approximate surface contour for the terrain scanned by the sonar. Due to the nature of side scan sonar data it will not be possible to build an accurate 3-D elevation map, but using the methods described in this paper, we are able to build what we call a 'Qualitative Elevation Map'. This map contains information about elevation constraints of the terrain surface, boundaries of distinct objects and regions with true and relative elevations. This map does not provide a detailed high resolution description of the environment, but it contains enough information to enable an autonomous vehicle to navigate safely in unknown terrain. By matching data taken from different viewpoints, we will be able to reduce the amount of unknown areas and also eliminate false information due to noise.

Hence our future work will consist of extending the present algorithms to compute all the information needed for the qualitative elevation map and developing a robust matching algorithm. Techniques have been demonstrated on data from a conventional side-scan sonar but our goal is to port them to a fast imaging sonar. The sonar mapping module will eventually be used in the navigation and mapping of a vehicle being developed at FAU.

Acknowledgments This work is part of a joint project with Professor Cushieri at Florida Atlantic University who provided us with the sonar data. The authors would like to thank Shree Nayar for valuable comments on the reflection model.

References

- [1] A. Blake and A. Zisserman. *Visual Reconstruction*. MIT Press, 1987.
- [2] L. Camp. *Underwater Acoustics*. Wiley-Interscience, 1970.
- [3] S. G. Chappell. A Simple World Model for an Autonomous Vehicle. In *Proc. Sixth International Symposium on Unmanned, Untethered Submersible Technology*, 1989.
- [4] J. M. Cushieri and M. Hebert. Three-Dimensional Map Generation from Side-Scan Sonar Image. *Journal of Energy Resources Technology*, 112, 1990.
- [5] A. Elfes. A Sonar-Based Mapping and Navigation System. In *Proc. IEEE Conference on Robotics and Automation*, 1986.
- [6] E. M. Geyer and J. A. D'Appolito. Characteristics and Capabilities of Navigation Systems for Unmanned Untethered Submersibles. In *Proc. Sixth International Symposium on Unmanned, Untethered Submersible Technology*, 1987.
- [7] M. Hebert. Terrain Modeling for Autonomous Underwater Navigation. In *Proc. Sixth International Symposium on Unmanned, Untethered Submersible Technology*, 1989.
- [8] S.K. Nayar, K. Ikeuchi, and T. Kanade. Surface Reflection: Physical and Geometrical Perspectives. Technical Report CMU-RI-TR-89-7, The Robotics Institute, Carnegie Mellon University, 1989.
- [9] V. Rigaud and L. Marce. Absolute Location of Underwater Robotic Vehicles by Acoustic Data Fusion. In *Proc. IEEE Conference on Robotics and Automation*, 1990.
- [10] K. Stewart. Three-Dimensional Modeling of Seafloor Backscatter from Sidescan Sonar for Autonomous Classification and Navigation. In *Proc. Sixth International Symposium on Unmanned, Untethered Submersible Technology*, 1989.
- [11] C. Thorpe and J. Gowdy. Annotated Maps for Autonomous Land Vehicles. In *Proceedings of DARPA Image Understanding Workshop*, Pittsburgh PA, September 1990.
- [12] R. Urlick. *Principles of Underwater Sound*. McGraw-Hill, 1983.

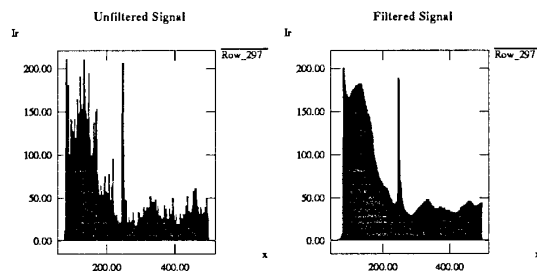


Figure 1: Echo signal of a scan before and after filtering

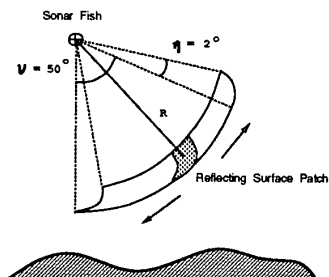


Figure 2: Sonar field of view

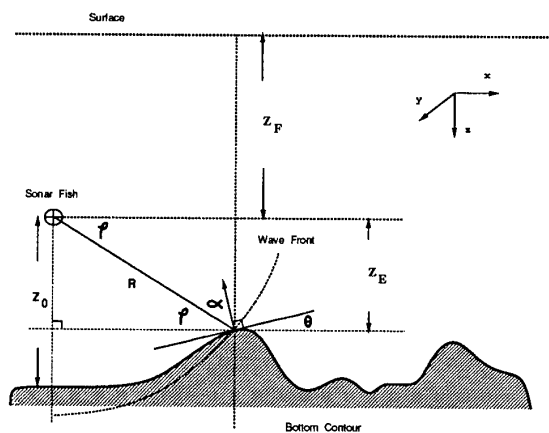


Figure 3: Sensor geometry

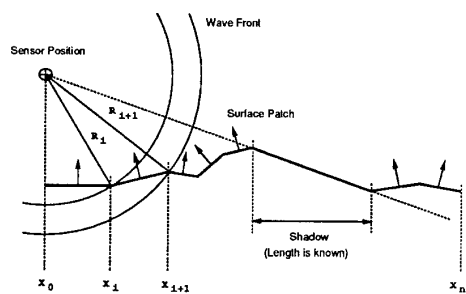


Figure 4: Surface reconstruction

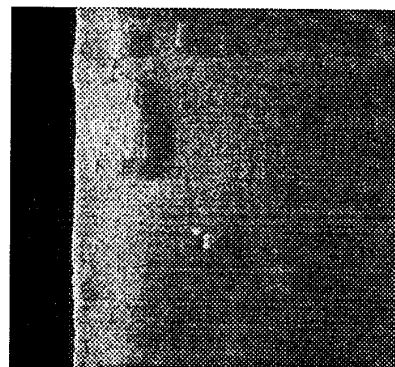


Figure 5: Unprocessed sonar image

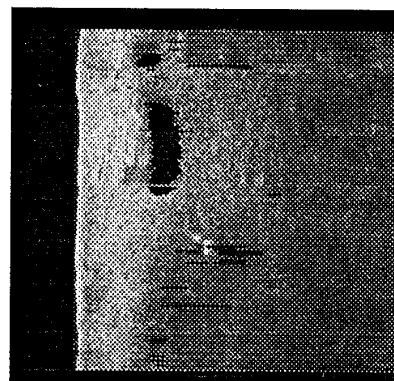


Figure 6: Shadows detected in sonar image

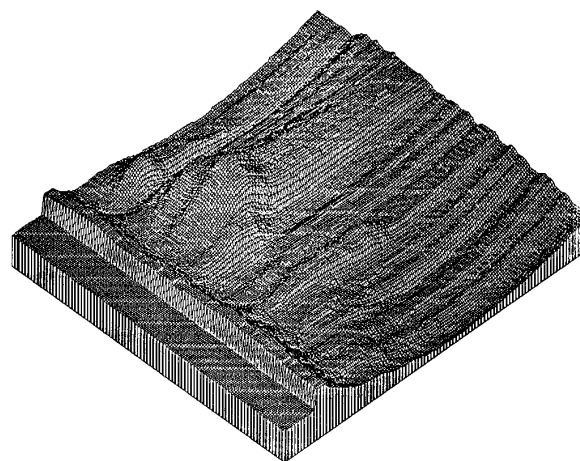


Figure 7: Isoplot of reconstructed surface