Building Elevation Maps From Underwater Sonar Data

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.

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 [13], 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 [11]. 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 [13] 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* [11]. *Rigaud and Marce* developed in [10] 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 480 rows x 512 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). Figure 6 shows the digitized sonar data.

Section 2 describes algorithms, we have implemented for filtering the raw data. Section 3 describes the scattering of sound underwater and a geometric reflection model. Finally section 4 describes methods for computing relative elevations and approximate surface contours, using the previously introduced reflection model. Limitations of the methods and results are discussed.

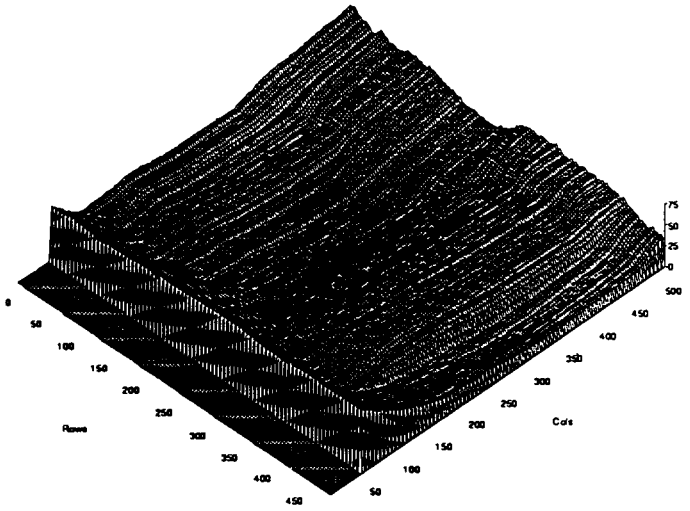

Figure 7: Isoplot of reconstructed surface (Method 1): Viewing East

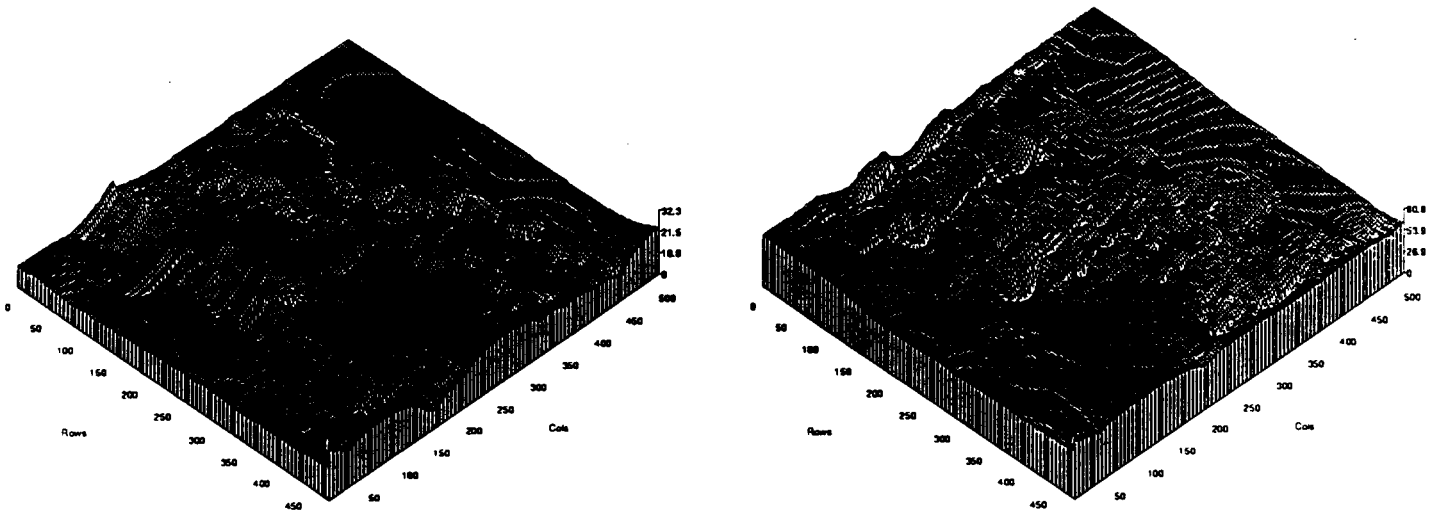

Figure 8: Isoplot of reconstructed surface (Method 2): (left) Viewing East, (right) Viewing South

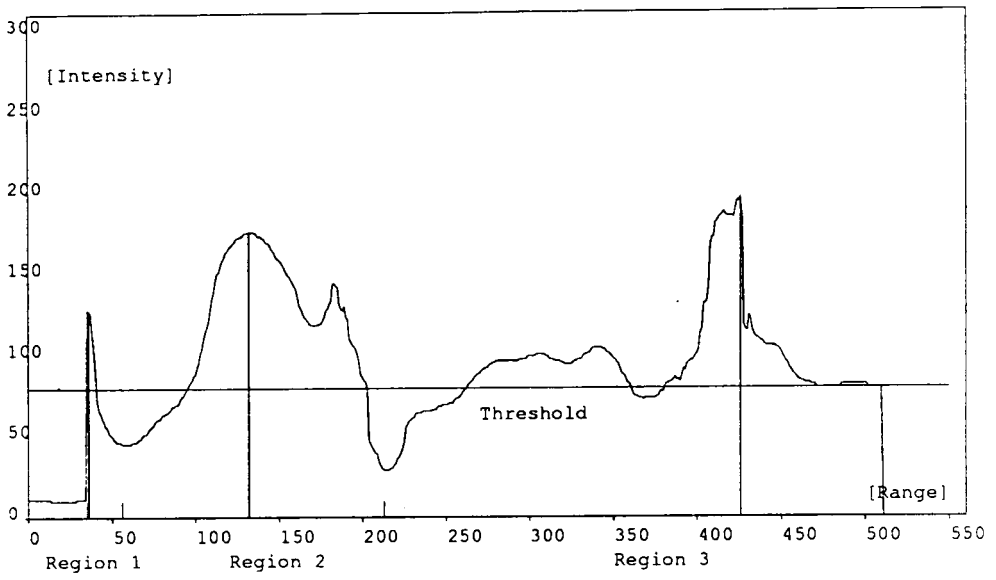

Figure 3: Thresholding and segmentation

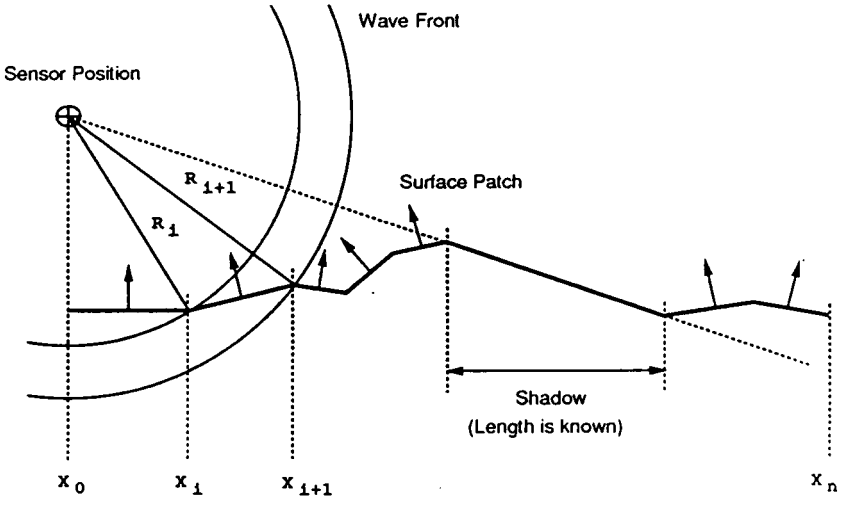

Figure 4: Surface reconstruction (Method 1)

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 (Florida Atlantic University).

Acknowledgments This work was supported in part by the National Science Foundation under Grant BCS-8816194 and 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 Umanned, 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 Umanned, Untethered Submersible Technology*, 1987.

[7] M. Hebert. Terrain Modeling for Autonomous Underwater Navigation. In *Proc. Sixth International Symposium on Umanned, Untethered Submersible Technology*, 1989.

[8] D. Langer and M. Hebert. Building Qualitative Elevation Maps from Side Scan Sonar Data for Autonomous Underwater Navigation. In *Proc. IEEE Conference on Robotics and Automation*, 1991.

[9] 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.

[10] V. Rigaud and L. Marce. Absolute Location of Underwater Robotic Vehicles by Acoustic Data Fusion. In *Proc. IEEE Conference on Robotics and Automation*, 1990.

[11] K. Stewart. Three-Dimensional Modeling of Seafloor Backscatter from Sidescan Sonar for Autonomous Classification and Navigation. In *Proc. Sixth International Symposium on Umanned, Untethered Submersible Technology*, 1989.

[12] Demetri Terzopoulos. Computing visible-surface representations. Technical Report A.I. Memo No. 800, Massachusetts Institute of Technology, AI Laboratory, 1985.

[13] C. Thorpe and J. Gowdy. Annotated Maps for Autonomous Land Vehicles. In *Proceedings of DARPA Image Understanding Workshop*, Pittsburgh PA, September 1990.

[14] K. E. Torrance and E. M. Sparrow. Theory for Off-Specular Reflection from Roughened Surfaces. *Journal of the Optical Society of America*, 57, 1967.

[15] R. Urick. *Principles of Underwater Sound*. McGraw-Hill, 1983.

which means that usually only few data points are selected, or using a low threshold, which leads to more data points, but also to much more noise. In the case of a more or less uniform signal intensity across an entire row, however, the threshold tends to be set too high and only very few data points are used. Therefore, if the computed threshold selects less than 25 % of all signal values in a row, it is set such that 25 % are selected. Hence assigning $\alpha = 0°$ at the maximum intensity of a region and $\alpha \approx 80°$ at the threshold, we can now compute I_o in the reflection model as discussed in section 3.3, using parameters E_o^* and K_g (Refer to equation 9). We choose $\alpha \approx 80°$ at the threshold, since the threshold is based on the histogramm and therefore a measure of background noise and 'noise' due to multiple echoes from interreflections. Thus the reflectance model parameters can be calculated separately for each region and we can now proceed to reconstruct the surface geometrically. Let us represent the surface in cylindrical coordinates: $\varphi = \varphi(R, y)$, where

$$x = R\cos\varphi, \quad y = y, \quad z = R\sin\varphi \tag{10}$$

The surface is then discretized in φ as shown in Fig. 5, where $\Delta\varphi = 50° / (\text{No. of range bins } R_i - 1)$. This representation simplifies enforcement of the angular constraints imposed by the sensor (Refer to assumption 1 above). It should be noted that input data points are given as a function of R and y. Hence discretization in φ means that we have to subsample in R. Starting at a zero incidence point, surface patches are now computed in the direction of increasing R. R_{i+1} can then be computed, using the sine rule and the value of α at R_i (Refer to Figure 5):

$$R_{i+1} = R_i \frac{\sin(\alpha + \pi/2)}{\sin\beta}, \qquad \beta = \pi/2 - \alpha - \Delta\varphi \tag{11}$$

Angle α is given by the reflectance model. The following surface patch is then computed using the value of α at R_{i+1}. Similarly surface patches for decreasing R are computed. If α is close to 0°, $R_{i+1} \approx R_i$ and thus $\Delta R < 1$. In this case we set $\Delta R = 1$. Similarly as φ approaches 90°, ΔR becomes very small. In this case φ is increased along a plane from the last data point until $\Delta R > 1$ and the next data point can be accessed (Refer to Figure 5). The procedure is repeated for each zero incidence point in a row and subsequently for all rows in the image.

Each surface patch satisfies the geometric constraints given by the vertical field of view of the sensor and α. The reconstruction process is also consistent with the reflection model, where the received intensity I_r is defined per unit solid angle. One problem remains in determining the relative elevation of each surface region with respect to depth Z_o and relative to each other. At present, we make the assumption that all hills and valleys are small deflections from a horizontal plane, which is positioned at depth Z_o. Thus for a particular set of ranges R_i, we can compute a corresponding range in φ (Refer to Figure 5). Each region is then computed within that φ range.

We thus obtain a sparse set of surface points, concentrated around zero incidence points. In order to get an estimate of the complete surface and ensure smooth interpolation between data points, regularization techniques are used. These techniques have become fairly popular in computer vision in recent years. A thorough introduction is given for example in [12]. The input for these algorithms are sparse data points, such as the ones obtained from the reconstruction process described above. A function is defined that computes the energy at a point in terms of the derivatives of the surface and the distance to data points. Points that minimize this energy function are then substituted as surface points. In this way the method ensures a smooth surface that is close to the original data points at the same time. Another advantage is that interpolation and smoothing are done concurrently. Using a cartesian coordinate system, the surface can be represented by $z = F(x, y)$ and an energy function can be defined in the general case as follows:

$$E = \kappa_1 \int \left(F_{xx}^2 + 2F_{xy}^2 + F_{yy}^2\right) dxdy + \kappa_2 \int \left(F_x^2 + F_y^2\right) dxdy + \lambda \int (z_D - F(x,y))^2 dxdy \tag{12}$$

where z_D are valid points from the data. Surface points z should then minimize the condition $\frac{\partial E}{\partial z} = 0$. The procedure is then applied iteratively until it converges to a minimum solution.

The images acquired by the sonar are quite large. In order to reduce computation time, a simpler version of the energy function is used at present, given by the following equation:

$$E = \int \left(F_x^2 + F_y^2\right) dxdy + \lambda \int (z_D - F(x,y))^2 dxdy \tag{13}$$

Factor λ determines how much smoothing is done on valid data points and should be selected between 0 and 1. Results are shown in figure 8, using $\lambda = 0.1$.

cannot be computed from this information unless we know the incident intensity I_i. However, I_o is a relative measure of the reflective properties of different surfaces.

The constants P_{dl}, P_{sl} and σ_α are surface reflectance parameters. At the moment we assume reasonable values for these, depending on the type of sea bottom considered. The reflectance parameters can be adjusted as necessary, depending on the performance of the algorithms.

Generally values of $\sigma_\alpha = 30°$, $P_{dl} = 0.7$ ($\Rightarrow P_{sl} = 0.3$) worked well for the surfaces considered. Eventually, we intend to derive tables containing these parameters for different types of surfaces.

4 Reconstruction of the approximate surface contour

4.1 Introduction

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 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 one type of such propagation algorithms: 1-D algorithms that integrate the angular constraints along each scan line separately. Finally a short overview is given about 2-D techniques that integrate the constraint over the entire map at once.

4.2 1-D Integration

A first approximation of the bottom contour can be provided by an integration in one dimension along the image column direction. In this case constraints between rows are not considered, which leads to a simpler surface recovery procedure. In order to obtain a more accurate result, the resulting surface should therefore be refined afterwards by a 2-D method. All algorithms described here are purely based on sensor geometry. Using this geometry, the following assumptions are made:

1. The vertical field of view of the sonar $\nu = 50°$. The sonar beam is tilted down $\approx 10°$ from the horizontal. Thus we assume that for the nearest (first) echo recorded, $\varphi = \varphi_o = 30°$. Hence the depth below the sensor $Z_o = R_o \sin\varphi_o$ (see Fig. . 2).
2. All surface normals are taken to lie in the x-z plane, with α being measured counterclockwise from the surface normal (see Fig. 4).
3. The received intensity for a particular range R is radiated from a single point on the surface. Note that this may not be the case in reality because of interreflections. However, at present this assumption is made to simplify the reconstruction process.

Note that these assumptions are made in order to put some necessary constraints on the problem. However, as a result the reconstruction process can be fairly inaccurate, depending on the actual shape of the surface. Especially the inclination of the surface in y-direction may not be computed correctly because of the second assumption. Following, we describe two different 1-D integration methods: The first one uses uniform reflectance parameters for the entire image. The second one provides a more differentiated approach by computing different surface albedos for different regions in an image. The above assumptions are made for both reconstruction methods.

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. 2). Once α is known, we can attempt to find the actual bottom contour. The reflectance model 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:

$$H \cos \alpha \leq \lambda/4 \tag{1}$$

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:

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

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 [9]. 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 [9] 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:

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

The constants C_{dl} and C_{sl} denote the strengths of the diffuse and specular component depending on incident intensity 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 [14] for a detailed explanation). For the sonar sensor geometry, it is given by:

$$G(\alpha) = \min(1,\, 2\cos^2 \alpha) \tag{4}$$

(See Fig. 2 for a definition of α).