Technical Report 1143

Multisensor Modeling Underwater with Uncertain Information

W. Kenneth Stewart, Jr.

MIT Artificial Intelligence Laboratory

This blank page was inserted to preserve pagination.

MULTISENSOR MODELING UNDERWATER WITH UNCERTAIN INFORMATION

by

W. Kenneth Stewart, Jr.

Submitted to the Department of Ocean Engineering on July 5, 1988
in partial fulfillment of the requirements for the Degree of
Doctor of Philosophy in Oceanographic Engineering

ABSTRACT

This thesis develops an approach to the construction of multidimensional stochastic models for intelligent systems exploring an underwater environment. The important characteristics shared by such applications are: real-time constraints; unstructured, three-dimensional terrain; high-bandwidth sensors providing redundant, overlapping coverage; lack of prior knowledge about the environment; and inherent inaccuracy or ambiguity in sensing and interpretation. The models are cast as a three-dimensional spatial decomposition of stochastic, multisensor feature vectors that describe an underwater environment. Such models serve as intermediate descriptions that decouple low-level, high-bandwidth sensing from the higher-level, more asynchronous processes that extract information.

A numerical approach to incorporating new sensor information—*stochastic backprojection*—is derived from an incremental adaptation of the summation method for image reconstruction. Error and ambiguity are accounted for by blurring a spatial projection of remote-sensor data before combining it stochastically with the model. By exploiting the redundancy in high-bandwidth sensing, model certainty and resolution are enhanced as more data accumulate. In the case of three-dimensional profiling, the model converges to a "fuzzy" surface distribution from which a deterministic surface map is extracted.

Computer simulations demonstrate the properties of stochastic backprojection and stochastic models. Other simulations show that the stochastic model can be used directly for terrain-relative navigation. The method is applied to real sonar data sets from multibeam bathymetric surveying (Sea Beam), towed sidescan bathymetry (Sea MARC II), towed sidescan acoustic imagery (Sea MARC I & II), and high-resolution scanning sonar aboard a remotely operated vehicle. A multisensor application combines Sea Beam bathymetry and Sea MARC I intensity models. Targeted real-time applications include shipboard mapping and survey, a piloting aid for remotely operated vehicles and manned submersibles, and world modeling for autonomous vehicles.

Thesis Co-Supervisor: Arthur B. Baggeroer, Ph.D.
Title: Professor of Ocean Engineering and
Professor of Electrical Engineering and Computer Science

Thesis Co-Supervisor: W. Eric L. Grimson, Ph.D.
Title: Matsushita Associate Professor of Electrical Engineering and Computer Science

ACKNOWLEDGEMENTS

Funding for this research was provided by the Sea Grant Program of the Massachusetts Institute of Technology, the Office of Naval Research, and the Monitor Marine Sanctuary Program of the National Oceanic and Atmospheric Administration.

Support for the Artificial Intelligence Laboratory's work in the areas of sensor interpretation and autonomous vehicles is provided in part by an Office of Naval Research University Research Initiative grant under contract N00014-86-K-0685.

Contents

ABSTRACT 2

Chapter 1 INTRODUCTION 7

1.1 WORKING UNDERWATER 9
1.2 AN ANALOGY 12
1.3 A MODEL-BASED PERSPECTIVE 13
1.4 EMERGING TECHNOLOGIES 14
1.5 A DIFFERENT APPROACH 17
1.6 DOCUMENT ORGANIZATION 19

Chapter 2 MODELING ISSUES 20

2.1 INFORMATION PROCESSING 20
2.2 TYPES OF MODELS 22
2.3 REPRESENTATIONS 23
2.3.1 Types of Representation 23
2.3.2 Levels of Representation 25
2.3.3 Spatial Representation 26
2.4 MODELING UNCERTAINTY 29
2.5 HISTORICAL PERSPECTIVE 32
2.5.1 Reconstruction from Projections 32
2.5.2 Incremental Reconstruction 36
2.5.3 Terrestrial Robotics 37
2.5.4 Underwater Robotics 40

Chapter 3 MODELING THE UNDERWATER ENVIRONMENT 43

3.1 SPECIFIC UNDERWATER CONSTRAINTS 43
3.2 REPRESENTATIONAL DESIGN 44
3.3 ARCHITECTURAL OVERVIEW 48
3.3.1 Vector Modeling 48
3.3.2 Open-Loop Modeling 50
3.3.3 Mapping Feature Vectors Through Model Space 52
3.3.4 A Sonar Modeling Example 53
3.4 MODELING A SONAR EVENT 56
3.5 STOCHASTIC BACKPROJECTION 60
3.6 GLOBAL MODELING 65
3.6.1 Binary Model 65
3.6.2 Continuous Model 69

3.7 SUMMARY OF IMPORTANT POINTS 72

Chapter 4 COMPUTER SIMULATIONS 74

4.1 SENSOR MODELS 74
4.2 OPEN-LOOP MODELING 83
4.3 CLOSING THE LOOP 93
4.4 SUMMARY OF IMPORTANT POINTS 98

Chapter 5 APPLICATION DATA SETS 100

5.1 MULTIBEAM BATHYMETRIC: SEA BEAM 100
5.2 TOWED SIDESCAN: SEA MARC I 105
5.3 TOWED SIDESCAN: SEA MARC II 109
5.3.1 Intensity Model 110
5.3.2 Bathymetric Model 114
5.4 PROFILE SCANNING SONAR: MESOTECH 114
5.5 SUMMARY OF IMPORTANT POINTS 121

Chapter 6 MULTISENSOR MODELING 123

6.1 ISSUES 123
6.2 CURRENT TECHNIQUES 126
6.2.1 Optical Intensity Modeling 126
6.2.2 Acoustic Intensity Modeling 127
6.3 SEAFLOOR BACKSCATTERING 128
6.4 THE SIDESCAN PROBLEM 131
6.5 LAMBERTIAN SCATTERING 135
6.6 SEA MARC/SEA BEAM MODEL 135
6.7 SUMMARY OF IMPORTANT POINTS 145

Chapter 7 CONCLUSIONS 147

7.1 SUMMARY OF RESULTS 147
7.2 LIMITATIONS OF CURRENT IMPLEMENTATION 148
7.3 ISSUES TO EXPLORE 151
7.3.1 Iterative Techniques 151
7.3.2 Bootstrap Modeling and Positioning 152
7.3.3 Strategies 152
7.3.4 Virtual Models 153
7.3.5 Dynamic Modeling 153
7.3.6 Multifeature Modeling 153
7.3.7 Other Representations 154
7.3.8 Multidimensional Feature Extraction 154
7.3.9 Bookkeeping 155
7.4 MODELING APPLICATIONS 156

ACKNOWLEDGEMENTS 158

REFERENCES 159

Figures

Figure 1.1: A generic multisensor exploratory probe. 8
Figure 2.1: Parallel ray projection geometry. 33
Figure 3.1: Information flow for open-loop modeling. 50
Figure 3.2: Detection envelopes for a sonar event. 58
Figure 3.3: Geometry for three-dimensional modeling. 66
Figure 3.4: Geometry for two-dimensional modeling. 70
Figure 4.1: Data-flow diagram. 75
Figure 4.2: Probability distribution for a sonar event. 76
Figure 4.3: Omnidirectional reconstruction: scanning geometry. 77
Figure 4.4: Omnidirectional reconstruction: one event. 78
Figure 4.5: Omnidirectional reconstruction: two events. 78
Figure 4.6: Omnidirectional reconstruction: convergent. 80
Figure 4.7: Down-look reconstruction: scanning geometry. 81
Figure 4.8: Down-look reconstruction: ramp response. 82
Figure 4.9: Ramp response: surface estimates. 82
Figure 4.10: Down-look reconstruction: impulse response. 83
Figure 4.11: Impulse response: surface estimates. 83
Figure 4.12: Mount St. Helens terrain database. 85
Figure 4.13: Terrain modeling simulation with low error. 86
Figure 4.14: Terrain modeling simulation with high error. 86
Figure 4.15: Probability plane: sparse data. 87
Figure 4.16: Probability plane: 6° sonar. 88
Figure 4.17: Probability plane: scanning laser. 88
Figure 4.18: Probability plane: convergent. 89
Figure 4.19: Probability plane: horizontal. 90
Figure 4.20: Probability profile (generic). 90
Figure 4.21: Convergent probability histogram (generic). 91
Figure 4.22: Probability plane: interpolated and bit-clipped. 92
Figure 4.23: Perspective view of seascape relief map. 93
Figure 4.24: Low-error, 3°-beamwidth seascape model. 93
Figure 4.25: Moderate-error, 12°-beamwidth seascape model. 94
Figure 4.26: Scan correlations for stochastic positioning. 96
Figure 4.27: Depth map of artificial seascape. 96
Figure 4.28: Navigation error for stochastic positioning. 97
Figure 5.1: Sea Beam profiling geometry. 101
Figure 5.2: Sea Beam bathymetric tracks. 102
Figure 5.3: Sea Beam high-error tracks. 103
Figure 5.4: Sea Beam bathymetric surface. 104
Figure 5.5: Sea MARC I scanning geometry. 105
Figure 5.6: Sea MARC I: raw linear map. 107
Figure 5.7: Sea MARC I: averaged linear map. 107
Figure 5.8: Sea MARC I: grid-averaged map. 108

Figure 5.9: Sea MARC I: backprojected and corrected map. 108
Figure 5.10: Sea MARC II: sidescan intensity (large-scale). 111
Figure 5.11: Sea MARC II: bathymetry (large-scale). 111
Figure 5.12: Sea MARC II: sidescan intensity (medium-scale). 112
Figure 5.13: Sea MARC II: bathymetry (medium-scale). 112
Figure 5.14: Sea MARC II: sidescan intensity (small-scale). 113
Figure 5.15: Deep Drone profiling geometry. 115
Figure 5.16: Depth map of *USS Monitor*: transverse tracks. 116
Figure 5.17: Depth map of *USS Monitor*: longitudinal tracks. 117
Figure 5.18: Depth map of *USS Monitor*: composite tracks. 117
Figure 5.19: Depth map of *USS Monitor*: extrapolated. 118
Figure 5.20: Depth map of *USS Monitor*: filtered and scaled. 118
Figure 5.21: *USS Monitor*: Photo mosaic of wreck site. 119
Figure 5.22: *USS Monitor*: Transverse section through turret. 119
Figure 5.23: Perspective view of *USS Monitor* from the east. 120
Figure 6.1: Sidescan shadow transitions. 132
Figure 6.2: Sea MARC II: sidescan map showing intensity transitions. 132
Figure 6.3: Sea MARC II: relief map for intensity transitions. 133
Figure 6.4: Sidescan geometry artifacts. 133
Figure 6.5: Sea MARC I intensity map of Clipperton area. 136
Figure 6.6: Sea MARC I intensity map of Clipperton area (from *Kastens et al.*). . .137
Figure 6.7: Sea Beam contour map of Clipperton area. 138
Figure 6.8: Sea Beam contour map of Clipperton area (from *Gallo et al.*). 139
Figure 6.9: Sea Beam perspective view of Clipperton area. 140
Figure 6.10: Sea MARC/Sea Beam synthetic sidescan maps. 142
Figure 6.11: Normalized Sea MARC/Sea Beam intensity map. 143
Figure 6.12: Sea MARC/Sea Beam multisensor model. 144

Chapter 1

INTRODUCTION

In an age in which we have mapped the far side of the moon, still less than a tenth of one percent of the ocean floor has ever been seen by human eyes. Yet, an increasing use of the oceans has required a rapid expansion of our abilities to image the seafloor at a range of scales and resolutions. Recent developments in advanced remote systems promise to extend our human perception to the deeper ocean regions, but the ability of these systems to conduct successful and efficient research, exploration, survey, work, or inspection demands an acute capability to "sense" and model the undersea environment in real time.

Yet, as our understanding of subsea processes is refined and our questions become more subtle, the limitations of individual sensors become more apparent. Considering the full scope of a detailed site survey, for example, a gamut of sensors over different scales of range, resolution, and raw data types must be accommodated. Such a mission is represented by Figure 1.1, which shows an underwater vehicle equipped with a suite of remote sensors. These might include different sonars (obstacle avoidance, down-look, sidescan), cameras (video, digital still), a scanning laser, and sensors to measure gravity, magnetic fields, temperature, salinity, and so on. Though a tethered remotely operated vehicle (ROV) is represented, the intended scenario also applies to a free-swimming autonomous underwater vehicle (AUV) or towed instrument sled.

In all cases, this generic exploratory probe is capable of collecting an enormous amount of multisensor data as it moves through the undersea terrain. The technology to generate this information flow is here today; the challenge lies in developing new methods to integrate the data and to construct high-level models of the environment that can be used by man and machine alike. Though there are basic differences between sonar, video, and laser scanning, there is still much common ground in data acquisition, signal processing, digital representations, archiving, and presentation. What we need to

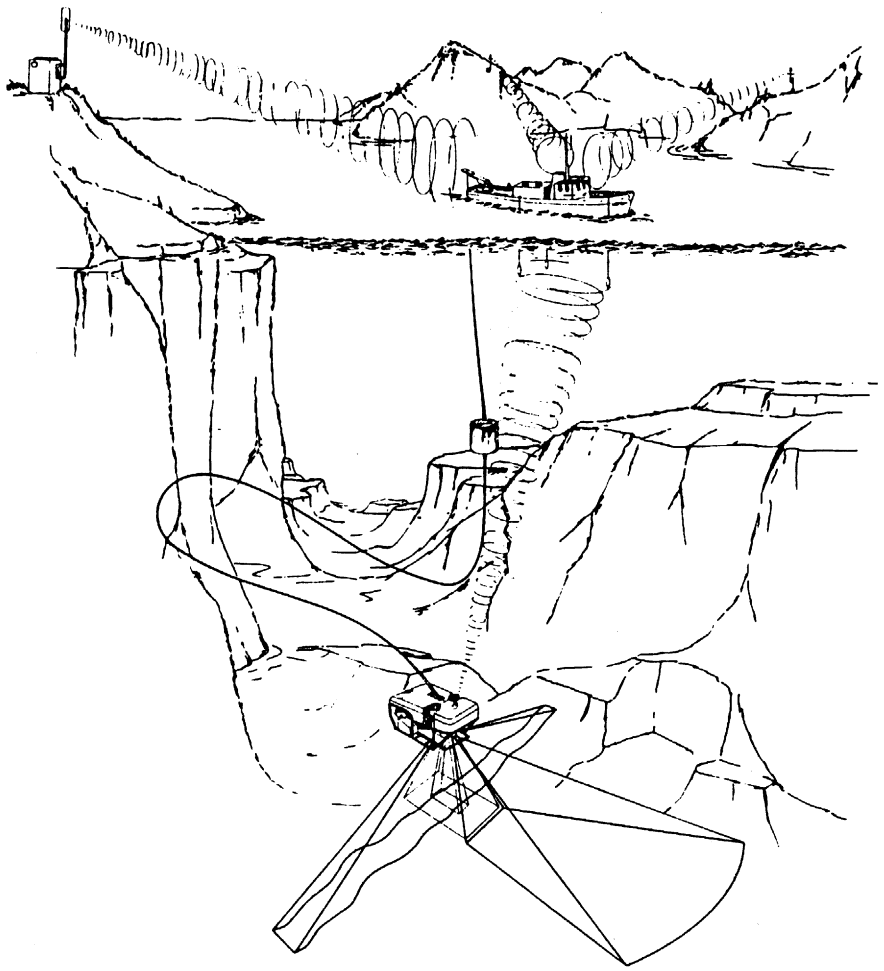

Figure 1.1: A generic multisensor exploratory probe.

take advantage of this commonalty for the synthesis of multisensor data is a consistent framework for information management.

Such is the problem I address in this thesis—constructing multidimensional models of the undersea environment with real-time multisensor data. Though I am mainly motivated by the needs of intelligent, autonomous systems exploring an unknown terrain, the approach is relevant to man-in-the-loop systems (ROV's and submersibles), and towed or shipboard mapping. The important characteristics shared by such applications are: real-time constraints; unstructured, three-dimensional terrain; high-bandwidth sensors providing redundant, overlapping coverage; lack of prior knowledge about the environment; and inherent inaccuracy or ambiguity in sensing and interpretation.

The approach taken in this thesis is to form a model as a three-dimensional spatial decomposition of cubical volume elements, or voxels. Associated with each voxel is a stochastic, multisensor feature vector that represents the properties within the small region. The model is an intermediate, numerical description that decouples low-level, high-bandwidth sensing from the higher-level, more asynchronous processes that extract deterministic information—for operator displays, obstacle avoidance, or path planning, to give a few examples. As new sensor information is acquired, it is merged using a technique I call *stochastic backprojection*; this is derived from an incremental adaptation of the summation method for image reconstruction. Error and ambiguity are accounted for by blurring a spatial projection of remote-sensor data before combining it stochastically with the model.

By exploiting the redundancy in high-bandwidth sensing, the model's certainty and resolution are incrementally enhanced as more data accumulate. This is in contrast with traditional approaches that rely on extensive postprocessing to eke out information from sparse data sets. Also, by taking advantage of complementary information from different sensors, more complete and more accurate models can be built, with less effort than for an exhaustive analysis of single-sensor data. For the real data sets considered in later chapters, the computational efficiency is such that cost-effective applications are feasible, and the quality and resolution of the models are appropriate to each.

The approach I take to modeling research relies mainly on a qualitative, visual assessment of results. On the one hand, this is important for man-in-the-loop applications that are subject to the same criteria of relevance and utility. On the other hand, vision has the highest bandwidth of all our senses, and offers a practical way to digest the large volume of information that a model contains. Such an approach has allowed me to quickly define the "envelope" of stochastic modeling—to look at the big picture and spot important determinants of performance.

In the rest of this chapter I provide a more detailed background on the problem, discuss current methods, and expand on the basis for my approach. The first section describes the underwater work environment for the three classes of systems that I mentioned: shipboard and towed systems, ROV's and manned submersibles, and autonomous underwater vehicles. Next, I draw an analogy between an intelligent underwater system and a human being exploring unknown surroundings to stress the advantages of a model-based approach. I then discuss emerging technologies that make this approach practical, and elaborate further on my philosophy of modeling.

1.1 WORKING UNDERWATER

To further our national economic interests, shipboard and towed sensor packages will continue to play an important role in mapping and assessment of seafloor resources. Maritime defense

requirements also call for a more comprehensive approach to tactical underwater terrain assessment, dictating more sophisticated information management and multisensor approaches. Marine scientists (geologists, archaeologists, and so on) will need more complete, more accurate, and more quantitative information than is now available.

Yet, despite advances in underwater sensors and computational technology, data processing and display techniques have changed little in the last two decades. Such two-dimensional seafloor mapping tools as sidescan sonars, for example, typically use analog paper-chart recorders to generate the final mapping product. Wide-area mosaicking relies on manual "cut-and-paste" methods and photographic reproduction for data manipulation. Within the last few years, video displays, digital recording, and image-processing techniques have come into use, but the basic approach is strongly linked to traditional paper-based methods.

Three-dimensional survey methods are even less advanced. Though digital data recording is most often used, postprocessing with manual intervention at every step remains the norm. Data products emerge after weeks or months and much expense. For large-scale bathymetric surveying, systems and processing tools are usually custom developed by end users or supplied as less capable add-ons from hardware manufacturers, tailored to a specific sensor. For small-scale, higher-resolution mapping, mainly used by the offshore industry, the only practical alternatives are "do it yourself" or rely on the expensive, customized offerings of a few service organizations.

For manned systems that operate within the relatively opaque underwater medium, the need for better environmental models is most strongly felt. In particular, researchers are hampered by the lack of sensory information available to man-in-the-loop systems. The pilot of a submersible or ROV suffers a tunnel-vision effect from the restricted sensing envelope of a viewport, camera, or sonar. The ensuing disorientation has severe economic penalties in terms of work efficiency, and can lead to damage or loss of a vehicle. For such systems, real-time processing of imaging sensor data is nonexistent and there appear to be no new approaches on the horizon.

The operator of an ROV, for example, usually relies on a view offered by one or more video cameras, sometimes augmented by a scanning sonar display. Under good conditions, low-light-level cameras can have a range of ten meters, less for a color image. Commonly, though, visibility can be restricted to less than a meter, especially when working near the bottom or in strong currents. Under all conditions, the operator's perception of distance is degraded by optical distortion and monocular vision. These factors, along with a camera's narrow field of view and the apparent "sameness" of underwater scenes, can quickly disorient a person at the controls.

Sonar systems extend the range of perception, give a direct measure of distance, and open another dimension under low-visibility conditions. Sonar, however, lacks the spatial resolution of a camera and is less easily interpreted by a human pilot. In the absence of strong acoustic reflectors with distinctive geometric properties, a vehicle's position can be hard to judge from the sonar display alone. The problem is compounded by motion artifacts introduced by a dynamic platform.

A drawback to both sonar and visual techniques is the transience of information presented to the operator. Though recorded for later review, from the pilot's perspective the data are continuously discarded. It is the human's burden to assimilate the information and to form his own internal model of the surroundings. In a terrestrial environment rich in sensory information, visual, tactile, aural, and other cues arrive in a form readily integrable by a human processing system evolved to match the task. But with already degraded sensor data collapsed to a two-dimensional form for video or sonar display, the information-assimilation problem is formidable and worsened by the need for attention to a complex system and to the immediate task at hand. The best ROV pilots seem to have a heightened proprioceptive sense, which makes this job easier; but the information remains unavailable to the ROV system itself, for example, to close position loops.

More direct ways of determining position underwater suffer from other limitations. Measurements of attitude and vertical position are available from an accurate, cost-effective sensor suite, however, horizontal positioning is more problematic. Acoustic transponder networks offer repeatable performance over extended periods, but are time-consuming to position and survey; this makes them uneconomical for small jobs at widely separated sites. Multipath and shadowing further restrict their use in shallow areas or in a cluttered environment such as the inner volume of an offshore platform. Inertial packages and doppler velocimeters are becoming more affordable for routine underwater use but need external updates to offset drift.

The few autonomous underwater vehicles in operation today use little more than programmed controllers to follow a pre-established track. The dominant issues in their development have been hardware related—mainly power and communications. Usually the *autonomous* designation really means that they are untethered, and rely on low-bandwidth acoustic modems for intermittent communications with a human supervisor at the surface.

To work without human supervision, a free-swimming robot must model its environment and locate itself while exploring the surroundings, especially if traversing widely for extended periods. As AUV's evolve, they will need more sophisticated multidimensional models comprising multisensor data, which let them respond to an unpredictable environment. The machine's computational model must support such low-level behaviors as trajectory control or obstacle avoidance, and offer an approach to

more complex problems—path planning or other context-dependent strategies need some framework in which to evaluate alternatives.

What is needed is a comprehensive approach to modeling and positioning underwater—new techniques that furnish enhanced sensory cues for more efficient human piloting, and that generate information in a form suited to automatic control systems as well. A cumulative sonar model, for example, could be used to generate a screen image of the underwater terrain with a representation of the vehicle superimposed. Digital position estimates, derived from a model of local features, could be used directly for closed-loop position control; this could circumvent the need for external navigation equipment in many applications.

In terms of end use, a distinction between teleoperated and autonomous systems is largely irrelevant. There is really a continuum of function that will serve man-in-the-loop systems and facilitate a transition to more independent underwater robots. Unlike the laboratory environment, where an investigator can walk down the hall to rescue an errant machine or to observe its behavior, the ocean is a more inaccessible and hostile place, and it will force us to adopt different research strategies. The gradual relinquishment of human control will be preceded by a more interactive phase of sharing and trading control between man and machine [*Sheridan*, 1982]. Model-based imaging and positioning could ease the load of a human operator, but are prerequisite to a robot in an unknown terrain. A useful approach in both domains will help speed the evolution.

1.2 AN ANALOGY

To help put the problem in perspective, think about this analogy. Suppose we wanted to build an underwater system that could sense its environment and construct a representation that could be displayed to an operator. The vehicle might be equipped with a scanning sonar or laser rangefinder and a suite of sensors to measure pitch, roll, heading, and depth. Because of the difficulty in measuring horizontal position underwater, we would want the machine to estimate its location relative to the surroundings, and use this estimate as it adds new range measurements to the model.

Now recall the sensation of entering an unfamiliar room in near darkness—your vision is diminished and you rely mostly on touch. At first, your knowledge is limited to a few observations about the boundaries of the space and of the objects it contains. Gradually, as you move about, your awareness of spatial relationships is enhanced, and confidence in your internal model grows. Eventually, you move more quickly and freely between known positions, avoiding obstacles, with only occasional checks to correct your perceived location.

A similar internal model is responsible for our sense of visual acuity over a wide field of view, though the full resolution of a normal human eye is limited to a narrow, forward-looking cone. Optical illusions further illustrate the power of such models in shaping our human perceptions (for example, see *Cornsweet*, 1970; *Marr*, 1982). As *Winston* [1984] puts it, image understanding may be a form of "controlled hallucination," so that our perceptions are influenced by what we expect to see. My point is that this internal processing and representation results from our internal "wiring" as well as from our experience, and usually enhances an ability to deal with a complex world.

1.3 A MODEL-BASED PERSPECTIVE

This human analogy is not meant as an argument for some anthropomorphic blueprint to build the machine counterpart, but to point out important characteristics common to the two scenarios. The main idea is that each approach is centered on an internal model of the environment. For the automated version, this is an intermediate representation describing the distribution of surfaces that reflect energy from the rangefinder. Information is lost as the raw sensor data are condensed and coerced into a new form. But if the structure is more appropriate to interpretive processes that "extract" information from the model, then system productivity may be enhanced.

A model can also fill in gaps left by degraded sensors or represent regions beyond their immediate field of view. For example, a graphic display created from a sonar model could be used for piloting in low-visibility situations. Even under ideal conditions, a representation of objects outside the camera's narrow viewing envelope would reduce an operator's sensation of tunnel vision and lessen the danger of entangling a vehicle's umbilical cable. Fully concurrent modeling and positioning, like that of the semiautonomous navigator sketched in our analogy, could start to take on part of the human pilot's load.

Autonomous vehicles that develop beyond the primitive capabilities of those today will need more comprehensive systems of representation. The models will have a high dimensionality that encompasses many different sensors—redundant and complementary types. But to be successful, the modeling also must account for the noisy, blurry, inaccurate, incomplete, and sometimes conflicting reports from many nonideal sensors. As new information is added, some notion of probability, possibility, or plausibility must be maintained and updated. Deterministic conclusions, if needed, are the venue of interpretive or evaluative processes.

These processes, which reference the model and act on their "interpretations," can be seen as a mechanism for closing the information loop. In the positioning example, there is a two-way flow of information between model generation and "perception" of sensor orientation. Registration of range

returns depends on a knowledge of sensor position and attitude. Conversely, estimates of sensor location are extracted relative to the model. In such a technique there is a threshold of error beyond which the algorithms will diverge. An implicit assumption is that environmental features are distinct enough to allow unambiguous position referencing.

An AUV especially must be able operate in a region where no prior map exists and must accommodate unstructured underwater features. At powerup or in recovering from a failure, for example, the system must use some strategy to bootstrap into an awareness of the surroundings. Still, an expert knowledge of each sensor and its medium is required, and specific knowledge about the environment ought to be integrable with the model whenever useful.

As in human learning, the model should grow incrementally, converge toward some useful representation of the sensory data being conserved, and allow us to draw conclusions at any time from all information on hand. Real-time performance is an important issue since a practical system cannot adopt a stop-and-go strategy with several intervening seconds of intensive computation. For that reason, modeling and interpretation processes must be computationally efficient or have a suitable decomposition for parallel or application-specific processors. For example, by decoupling high-bandwidth, real-time sensor requirements through an intermediate model, information extractors and synthesizers may function in a more asynchronous manner suited to their task or hardware base.

The main point is that a model provides a powerful, unifying framework in which such processes can operate. The human brain is the site of much of our internal model, though research is beginning to unravel the complex interactions between later, more cognitive representation/processing and early, lower-level components, which are more closely related to our sensors [*Grimson*, 1980; *Marr*, 1982]. Most approaches to world modeling or machine vision take a high-level approach, in which the surroundings are represented as an assemblage of features—edges, corners, surfaces—or objects. In part, this may have been because of the economy of such representations in an era when sensor data were sparse and computational resources precious. As I make clear in the next two chapters, the models I consider are formed at a much earlier, or lower level.

1.4 EMERGING TECHNOLOGIES

Any technological constraints on the realization of such an approach are rapidly diminishing. New high-bandwidth, high-resolution sensors generate an enormous amount of data sometimes destined for postprocessing, but often relegated to the archives. Cost-effective sensor suites for attitude, velocity, and acceleration permit new approaches to the problems of misregistration and motion artifacts introduced by dynamic sensor platforms. And the computational resources that will let us take

advantage of this flood of information are here as 32-bit CPU's, cheap memory, and high-performance graphics.

Because of the ocean's relative opacity to electromagnetic energy, sonar has enjoyed a prosperous history since its introduction in the early part of this century [*Horton*, 1959; *Urick*, 1975]. Originally driven mainly by military applications, the field has spawned a family of systems suited to a wide range of uses [*Sutton*, 1979], from seafloor mapping and imaging [*Tyce*, 1987; *Davis et al.*, 1987] to search, classification, and navigation for submersibles and unmanned vehicles [*Cyr*, 1987]. Along with lower cost, the trends toward high information rates, narrow beam width, light weight, low power, modular design, and digital interfaces [*Baggeroer*, 1978; *Cyr*, 1987] are expanding the role of acoustic sensors in the underwater domain.

Still, acoustic methods for underwater use are less advanced than those of medical imaging [*Lee and Wade*, 1986; *Ferrari*, 1987] because of fewer driving interests, lack of fiscal support, and more severe environmental constraints [*Sutton*, 1979]. Ultrasonic techniques used in medicine include [*Havlice and Taenzer*, 1979] reflective (pulse-echo) imaging, direct transmission imaging, tomography (time-of-flight, attenuation, reflection, diffraction), holography, interferometry, and Bragg-diffraction imaging. Acoustic methods are also highly developed for applications in nondestructive testing [*Kino*, 1979]. Nevertheless, research in underwater acoustics is continuing and improvements in techniques, in technologies, and in matching system performance with human needs should lead to more effective underwater imaging systems [*Sutton*, 1979].

Once confined mainly to large, expensive military systems, sonar arrays and sophisticated processing have moved into the commercial world. Preformed-multibeam and phase-comparison sonars are supplanting fixed-beam and mechanical-scan sonars in an attempt to increase the information content and data rate of acoustic sensing. Though the theory has preceded their implementation by many years, such approaches to two- and three-dimensional sonar imaging as acoustic lenses [*Belcher*, 1987a, b], spatially encoded waveforms [*Jaffe and Cassereau*, 1988], and holographic techniques [*Collins*, 1987] portend even higher bandwidths for underwater acoustic sensing.

For long- and medium-range sensing underwater, sonar provides the only reasonable option. Though systems operating in the 1-5 MHz frequency range will also see service as high-frame-rate imaging sonars, developments in underwater scanning-laser technology offer an alternative with similar range capabilities but with higher angular resolution [*Dixon et al.*, 1983; *Klepsvik et al.*, 1987; *Henderson*, 1988]. Advances originally aimed at medical users (primarily the diode-pumped, frequency-doubled Nd:YAG laser) have greatly alleviated the power and size constraints faced by earlier researchers [*Holmes*, 1986]. New digital cameras with unprecedented sensitivity and dynamic range [*Harris et al.*, 1987] are further expanding the domain of optical imaging underwater.

Parallel developments in position and attitude measurement are enhancing our capabilities for tracking, navigation, and control of underwater platforms, though horizontal positioning will continue to be an issue [*Geyer et al.*, 1987]. Attitude sensing is a mature technology, and 3-axis measurements of angular position and velocity are available from compact, inexpensive packages. Rate sensors (mechanical gyro or gas-rate transducer, for example) updated with an absolute reference (magnetic for yaw, gravimetric for pitch and roll) can be accurate to a fraction of a degree in a strapdown package.

As the cost of laser-ring and fiber-optic rate sensors falls, the low drift rate and sensitivity of these devices will make a strong contribution to the performance of affordable inertial navigation systems [*Tusting and Caimi*, 1987; *Johnson and Eppig*, 1987]. Velocity-aided inertial navigation (using doppler or correlation velocity logs, for example), combined with a model of the platform dynamics, will let a vehicle navigate by dead reckoning for periods of time ranging from minutes to hours, even days, depending on accuracy requirements. As with attitude sensing, though, such systems drift and need a periodic update from an absolute frame of translational reference.

Along the vertical axis, pressure sensors or acoustic altimeters (up- or down-looking) offer satisfactory solutions for most applications, but lateral positioning has always posed a challenge. Acoustic positioning systems are the workhorse of the industry and include long-, short-, and super-short-baseline types. These typically span a range of accuracy and coverage from about 5-10 m at 5 km, to less than a meter for higher-frequency systems over a few hundred meters, though newer approaches offer still greater accuracy [*von der Heydt*, 1985]. Even higher-frequency systems, with baselines of a hundred meters or less, are now demonstrating an accuracy to within a few centimeters [*Hahn et al.*, 1985].

Geophysical navigation, usually by feature correlation, evolved mainly under military auspices but has started to arouse interest in the civilian sector [*Geyer et al.*, 1987]. Bathymetric navigation, where a sequence of sonar readings is correlated with a pre-stored map, is also used in seafloor mapping applications to supplement satellite fixes [*Tyce*, 1987; *Nishimura and Forsyth*, 1987]. Magnetic terrain navigation, using similar principles but a different geophysical feature, is enjoying recent interest [*Tyrén*, 1987; *Polvani*, 1987], particularly since it is a passive sensing modality and suited to covert applications. Gravimetric navigation is another possibility, but field results have been less encouraging than those using magnetic techniques [*Geyer et al.*, 1987]. The accuracy of all such techniques depends on the spatial bandwidth and distinctiveness of geophysical features.

To keep up with these high-bandwidth remote sensors, and to take advantage of position and attitude information for improved composite imaging, calls for computational resources that have been beyond the reach of many applications before now. However, the steady gains in price/performance of digital technology, especially with the advent of cheap 32-bit computing, have set the stage for more

sophisticated approaches to underwater modeling. Even real-time processing is within reach for many applications.

On the one hand, the continuing evolution of faster, more inexpensive memory devices lets us take advantage of the extended address space of new-generation microcomputers. Though unremarkable in the mainframe world, the size and complexity of our models, and of our development tools (UNIX, for example), only recently have become practical on an interactive, personal system. On the other hand, new graphics hardware and techniques offer the only reasonable approach to digesting the huge amounts of data these powerful machines can generate [*CG&A Staff*, 1987; *Stewart*, 1987b].

1.5 A DIFFERENT APPROACH

Research for this thesis proceeded from the premise that the technological obstacles to more sophisticated underwater modeling are no longer significant, or are rapidly diminishing. Also, our understanding of the physical processes that govern the sensors and their medium is such that we should be able apply this knowledge to the raw sensor data to enhance the information available to the system. What we lack, though, are the computational tools or, more precisely, the sets of tools that will let us take advantage of the diverse, high-bandwidth data at our disposal.

The results of this thesis research show that existing sensor and computational technology is such that sophisticated, high-resolution, multisensor modeling is within reach and can be accomplished in an incremental, real-time manner (see also *Stewart* 1987a, b, 1988). The basis of this new approach is a probabilistic, spatial decomposition strongly suited to amorphous, underwater terrain. Such a representation is an aggregate of sensor data obtained locally, but may incorporate prior information from other kinds of models (for example, a CAD model of an offshore structure or an underwater-terrain database). Besides a quantitative facility, an advantage of the technique is that the probabilistic framework explicitly represents the quality of information in the model, and the uncertainty imposed by the sensors, a dynamic platform, and the environment itself.

For the manned scenario described earlier, the model has been used to generate an auxiliary piloting display, from a global perspective, showing the vehicle and its relation to objects beyond the operator's field of view. The operator can benefit from a greater efficiency in task positioning, reduced transit times between work sites, and a lowered operational risk. Depth and range information can be provided using color contour maps or a shaded perspective view. Such an auxiliary display gives the pilot a more easily assimilable representation of his surroundings, and as the model's certainty increases, can be used directly for low-visibility piloting.

Other modeling results with real systems suggest benefits in large-scale underwater mapping applications—quality of the final product is improved and real-time processing reduces delay and expense in the postprocessing tedium. The technique is also demonstrated with multisensor processing of imaging (sidescan) and profiling (multibeam bathymetric) sonars, to take advantage of the complementary characteristics of each sensing modality. The stochastic modeling approach has been developed with such applications in mind, and is largely independent of scale, resolution, and sensor types.

Toward such an end, this thesis develops a philosophy of acquiring, processing, and representing information in a multisensor environment for consumption by high-level processes that interpret the information and act on it. The main tenets that shape this approach are:

1. The broad concept of a model provides a powerful framework for organizing our information about the environment and in assessing our understanding at any time. The most appropriate form is an explicit stochastic representation that accounts for the inaccuracy and uncertainty in our sensors and techniques.

2. In contrast to most approaches underwater, which often assume sparse information, there is a great deal of redundancy in this high-bandwidth data. By applying a knowledge of the sensors and media, we can exploit this redundancy to enhance the resolution and certainty of our models.

3. A realistic approach to more sophisticated undertakings must deal with information from many different sensors, redundant and complementary types. Multidimensional models and representations will be an essential part of more intelligent underwater systems.

4. There is no all-encompassing representation or processing paradigm to serve all purposes, so models and modeling processes will be largely domain-specific. At the same time, we should strive for a generality and consistency that lets us move conveniently among different representations and modeling domains as needed.

5. Model building is a simple kind of "learning" in which information is combined and accumulated to enhance the fidelity of our representation. As such, the approach should incorporate tools that are incremental and capable of real-time performance with modest computational resources.

6. In the end, postprocessing methods are capable of producing more faithful descriptions—more time and processing power can be brought to bear, and an inversion of all data can be

performed in aggregate. So for situations in which real-time feedback is essential (or useful), a tradeoff in fidelity against performance is inevitable and acceptable.

1.6 DOCUMENT ORGANIZATION

In the next chapter, what is meant by a model, in the context of this dissertation, is made more precise, and the more general modeling issues are presented. I also discuss the need for good representations and the distinction between the representation and the model itself. With this as a basis, different types of representations are discussed along with their use in modeling uncertainty. The chapter concludes with an overview of previous approaches to representing the surroundings, taken mainly from the world of terrestrial mobile robots.

Using the context developed in the second chapter, Chapter 3 introduces the more specific modeling constraints characteristic of an underwater environment, and applies them to the selection of an appropriate modeling representation. Computational and architectural issues are examined and used to define the general analytical framework. Finally, the framework is applied to examples of active sonar sensing and model building in the acoustic domain.

The two following chapters describe the results of computer simulations and field applications with a single remote-sensing modality. Chapter 4 deals mainly with computer simulations of open-loop modeling, in which all information used to build the representation is derived from sensor data. Other simulations show the feasibility of positioning with a stochastic model, and some implications of the approach are discussed. In Chapter 5, four real-world data sets, from profiling and imaging sonars, are used to confirm the results of model-building simulations.

Chapter 6 describes a higher-dimensional approach to integrating data from different sensing modalities. General issues are discussed and an overview of current techniques is given. An example using sonar bathymetry and sidescan imagery is used to illustrate the approach. The final chapter summarizes thesis results, discusses the limitations of the current implementation, and raises other issues to be addressed by future research. The chapter concludes with a discussion of more general underwater applications and forthcoming multisensor systems.

Chapter 2

MODELING ISSUES

In this chapter, I explain more precisely what I mean by models and representations, and discuss the different types, levels, and coordinate systems of each. After this discussion, I give a brief overview of alternatives for representing uncertainty in robotics applications. The remainder of the chapter is devoted to a survey of the different approaches to modeling a three-dimensional environment. This includes a discussion of image reconstruction from projections and more general incremental reconstruction techniques. Finally, I describe different methods that have been applied to mobile robots; most come from the domain of terrestrial robots, though a few approaches to autonomous underwater vehicles are considered. The purpose of this discussion is to establish a context for the ideas to be developed in Chapter 3, and to provide references to alternative methods.

2.1 INFORMATION PROCESSING

In the first chapter, I described several scenarios involving instrument platforms acquiring sensory data underwater. The manner in which these data should be stored and manipulated depends on the reasons for collecting the data and on their end use. For a geologist, a survey goal might be the construction of a bathymetric map to further his understanding of seafloor processes. For the pilot of an ROV, the information could be used to help navigate in a local area. An AUV could build an internal representation of the surroundings for path planning through an undersea terrain. The common thread is that of compiling knowledge about a previously unknown, or little-known, environment into a useful description.

In simple terms [*Marr*, 1982], we want to know what is where—to build some description of environmental properties with their spatial (and sometimes temporal) distributions. In the simplest case, this could be just the shape of the seafloor itself. More informative descriptions might include

surface scattering properties with respect to different energy sources (optical and acoustical, for example), temperature, chemical makeup, and so on. This can be extended to subsurface structure or to the composition of the seawater itself. In a general sense, such descriptions can be thought of as *models*, which embody knowledge or information about the world.

In the context of this thesis, what I mean by *model building*, or *modeling*, is the incremental aggregation of information into a stochastic description of the environment. By *information*, I mean any source of knowledge that reduces uncertainty in the model. In spirit, this is akin to *Shannon's* [1949] classic definition, though a rigorous usage is not implied. I also make a distinction between the *data* (digital bytes, for example) and what they tell us about the world. Ultimately, the information we are able to extract from the data depends on our understanding of the sensor, its medium, and how the two interact. In other words, our physical model of the sensory process is a source of prior knowledge about the world, which can be applied to the raw sensor data to enhance the information available to us. This is a forward model, in the usual sense, which guides our inversion of the data in the context of a specific model-building process.

Usually there are different ways of managing and storing the information that lend themselves to different goals or end-uses. For example, an optical image of a particular scene can be maintained as a photographic negative, a positive print, an analog signal on a video disk, or a collection of digital intensity values in computer memory. In the digital case, there are several more alternatives to consider, including two-dimensional arrays, quadtrees, run-length encoding, and different statistical coding techniques [*Pratt*, 1978; *Stoffel*, 1981].

We see intuitively that the information content for each is roughly the same; what differs is the particular scheme by which it is represented. The choice of any one for a given application depends on how the data are manipulated and on what information needs to be made explicit. A photographic print, for example, is a convenient way to represent images for human viewing; it is inexpensive, portable, and immediately understandable. Digital images offer much greater flexibility for machine manipulation and processing, but different digital representations serve different needs. Statistically-encoded image data has a compact format for storage or transmission, but a two-dimensional array is suited to an image operator such as a Fourier transform.

In overall philosophy, the approach developed over the course of my research has acquired a flavor similar to that of David Marr, who wrote as introduction to his book on vision [*Marr*, 1982]:

> Vision is therefore, first and foremost, an information-processing task, but we cannot think of it just as a process. For if we are capable of knowing what is where in the world, our brains must somehow be capable of *representing* this information—in all its profusion of color and form, beauty, motion, and detail. The study of vision

> must therefore include not only the study of how to extract from images the various aspects of the world that are useful to us, but also an inquiry into the nature of the internal representations by which we capture this information and make it available as a basis for decisions about our thoughts and actions. This duality—the representation and the processing of information—lies at the heart of most information-processing tasks and will profoundly shape our investigation of the particular problems posed by vision.

Though Marr's perspective is shaped by machine-vision research, his point about the duality of processing and representation is equally relevant to multisensor modeling, if not more so. The type of representation we use determines what information is made explicit in the model; the purposes for which a model can be used and the efficiency with which those purposes can be accomplished follow directly from that choice of representation.

Though the focus of this thesis is on building models, this is not an end in itself. Aside from providing an efficient framework for the aggregation of information, the representation must serve both human beings and machine processes that use this knowledge to understand the environment, to make decisions, and to act. For these important reasons I place much emphasis on representations in this and the next chapter. After elaborating more on models in the next section, I describe the types of representations and representational primitives as a basis for comparing the work of other researchers. With this as background the first part of Chapter 3 discusses the particular constraints of modeling underwater and their implications for the volumetric representation I adopt.

2.2 TYPES OF MODELS

A *model* is an *instance of a particular representation* that encapsulates some body of knowledge or information about an entity or process of interest. This definition is broad enough to subsume the general usage as it applies to an engineer's or artist's model that guides the realization of a full-scale project. It is a description of or substitute for the real thing. In another sense, it describes a process or abstraction, for example, an economic model of the world marketplace. An analytical model, sometimes represented by an equation or by a computational algorithm, captures an understanding of the physical world as, for example, an acoustic model of sound propagation in the ocean.

A particular physical model of importance to this thesis is referred to as a *sensor model*. This incorporates such parameters as the sampling envelope (beam pattern and look direction of a sonar, for example), resolution, accuracy, and so on. Loosely speaking, it also refers to the noise and distortion introduced by the system or the medium. The model should also include some characterization of the uncertainty in any real-world sensor.

In the fields of Artificial Intelligence (AI) and Robotics, *model* is typically used to denote a static description derived from prior information about the world. In this sense, the model is often employed for object recognition by template matching [*Grimson and Lozano-Perez*, 1983; *Shneier et al.*, 1986; *Magee and Nathan*, 1987] or for navigation by correlating sensor data with a pre-stored map of the environment [*Miller*, 1984; *Drumheller*, 1985; *Marce and Julliere*, 1986].

Since this thesis is about models and modeling, I will be explicit when my usage deviates from the definition given in Section 2.1. When not stated otherwise, though, *model* will refer to an aggregation of sensory information that describes the environmental parameters of interest. In contrast with the often-used AI/Robotics term, the models that this thesis treats are *dynamic descriptions* that reflect the *totality of information* at any moment. Such a model could incorporate prior knowledge at an initial state, but would be updated continuously as new sensory information arrived.

It is also appropriate to distinguish image processing from modeling. For the most part, image processing refers to a body of techniques that transform, encode, or transmit information already in the image data [*Pratt*, 1978; *Stoffel*, 1981]. This is not to say that an understanding of the physical basis for image formation is unimportant to the development of image-processing techniques or implementations. Rather, the priorities of image processing have to do with the images themselves, not with their use in building a description of the world.

2.3 REPRESENTATIONS

As indicated in Section 2.1, a *representation* is a *set of conventions* about how to describe a class of things; a *description*, or a model, uses the conventions of a representation to characterize some particular thing [*Marr*, 1982; *Winston*, 1984]. Though there may be many ways to represent a feature or object of interest, from a practical standpoint the choice of representation can have a strong influence on the types of processing the model can support and on the efficiency with which it can be implemented. A particular approach to the representation of knowledge must then be guided by the context in which data are acquired, manipulated, and presented.

2.3.1 Types of Representation

In general, representational schema may be broadly classified as *propositional* and *analogic* [*Ballard and Brown*, 1982]. The low-level models I begin to formulate in the next chapter use an analogic representation since they are suited to the description of physical or geometric properties of the environment. They can be used to describe spatial or temporal relationships among different

properties, and characterize each property over a range of continuous or discrete values. A bathymetric map, for example, describes the spatial distribution of depth over some area, and at each point any one of many possible depth values can be specified. Ballard and Brown characterize such analogic representations by:

- **Coherence**: Each element of a represented situation appears once, with all its relations to other elements accessible.

- **Continuity**: Analogous with continuity of time and motion in the physical world; permits continuous change.

- **Analogy**: The structure of the representation mirrors the relational structure of the represented situation; the representation is a description of the situation.

- **Simulation**: The models are interrogated and manipulated by arbitrarily complex computational procedures that often have the flavor of physical or geometric simulation.

Propositional models comprise assertions about the world that are either true or false. Such representations are most often used by high-level, semantic world models. Since I elaborate on this distinction between high- and low-level modeling in the next section, and refer to such propositional models in the later discussion of other approaches, I include here the characteristics attributed by Ballard and Brown:

- **Dispersion**: An element of a represented situation can appear in several propositions. However, the propositions can be represented in a coherent manner using semantic nets.

- **Discreteness**: Propositions are not usually employed to represent continuous change. However, they may be made to approximate continuous values arbitrarily closely. Small changes in the representation can thus be made to correspond to small changes in the represented situation.

- **Abstraction**: Propositions are true or false. They do not have a geometric resemblance to the situation; their structure is not analogous to that of the situation.

- **Inference**: Propositional models are manipulated by more or less uniform computations that implement *rules of inference* allowing new propositions to be developed from old ones.

As *Ballard and Brown* [1982] point out, each type of model derives its "meaning" differently. However, in computer implementations especially, the two representations only differ essentially in the last two points, and it is often possible to transform one representation to the other without loss of information.

2.3.2 Levels of Representation

Machine perception can be considered as a mapping of sensor input to a description of the environment. In other words, given some collection of sensory data, the problem is to attach some meaning (or to extract information) by relating it to existing models of the world. In general, this is not a direct process, but a sequence of transformations over a range of representations. The process usually proceeds in a hierarchical manner from low-level, physical descriptions to higher-level, more objective, or cognitive interpretations of the surroundings. Low-level representations and processes tend to be purely analogic; high-level representations and processes tend to be both analogic and propositional [*Ballard and Brown*, 1982]. As in the human visual system, though, the flow of information is not necessarily unidirectional. Lower-level interpretations may be guided by prior knowledge embedded in the upper levels of a hierarchy [*Marr*, 1982].

Marr's [1982] hierarchical paradigm for vision includes four coarse levels beginning with the *image*, represented as a collection of intensity values. Moving up the hierarchy, the *primal sketch* carries information about two-dimensional features in the image, including edge segments, blobs, discontinuities, and boundaries. The *2½-D sketch* describes such surface attributes as local orientation, depth discontinuities, and distance from the viewer. The *three-dimensional model* describes shapes and their spatial organization in an object-centered coordinate frame.

Ballard and Brown [1982] take a similar view starting with low-level *generalized images*—iconic, analogic representations of the input data. At this level, *intrinsic images*, which reveal physical properties of the scene (surface orientation, range, or surface reflectance), are contrasted with the raw sensor images. *Segmented images*, at the next higher level, are formed from the generalized images by gathering their elements into sets likely to be associated with objects. *Geometric representations* capture two-dimensional and three-dimensional shape. At the highest level, *relational models* are complex assemblages of representations, often using prior knowledge and models acquired before the perceptual experience.

For the purposes of this dissertation, I define a slightly different hierarchy, which overlaps with the two I have just discussed. There are two motivations for doing so. First, a characterization by images is inappropriate for many sensing modalities. An image implies an array of data samples taken in a "snapshot" mode when sensor platform dynamics can be ignored during the interval in which the image is formed. The sampling rate of a single-beam sonar (sidescan or sector-scan, for example) is limited by the speed of sound in water and, depending on the range, can take several seconds to finish one ping cycle. Tactile sensing offers another example. Second, a more general characterization

provides a context for discussing the work of other researchers who often use their own unique version. At the lower levels, this characterization is similar to that described in *Henderson et al.* [1987]:

- **Data level**: Corresponds to the raw sensor input. No inversion is performed.
- **Physical level**: Physical models are used to invert the raw data for the extraction of such low-level intrinsic properties as scattering strength, surface reflectance, or texture.
- **Feature level**: Physical- or data-level parameters are grouped locally to extract such primitive features as edges, surfaces, regions, or blobs.
- **Object level**: Lower-level parameters are used to segment the model into distinct entities.
- **Semantic level**: Feature- or object-level primitives are classified and interpreted according to a prior model and may be assigned "meaning" or inherit the propositional characteristics of their class.

In most real applications, the distinctions among such categories tend to blur. At the data level, for example, sensor subsystems often perform partial inversions based on a crude model of the medium—sonars may apply a time-varying gain (TVG) to compensate for scattering and absorption, or return one range value based on an assumed speed of sound and scattering threshold. At the higher levels, it is often difficult to classify a particular researcher's approach because of the bidirectional flow of information.

For convenience, I further divide the use of representations into low- and high-level methods. Low-level techniques mainly use the first three kinds of representations and are most often purely analogic. High-level approaches concentrate effort at the top three levels of the hierarchy, often incorporate propositional information, and tend to make use of prior models. The overlap at the feature level is intentional since some instance of this representation can be found in most implementations.

2.3.3 Spatial Representation

Spatial reasoning is recognized as an important part of many cognitive processes and germane to nearly every line of Robotics research (for a collection of papers on spatial reasoning and multisensor techniques, see *Kak and Chen*, 1987). Implicit in most representations, at all levels, is a mechanism for defining the spatial relationships among the different bits of information. At the most basic level of interest, we want to discover the shape—the geometry of a physical surface—of the seafloor and its features. For other kinds of models, we usually need to describe the spatial distribution of certain

parameters—bottom backscatter, temperature and salinity in the water column, or the flora and fauna of a benthic ecosystem, for example.

Shape information has a special character because, unlike color or visual texture information, the representation of most kinds of shape information requires some kind of coordinate system for describing spatial relations [*Marr*, 1982]. It is an intrinsic property of three-dimensional objects; in a sense, it is the primal intrinsic property for a sensory system from which many others (surface normals, object boundaries) can be derived [*Ballard and Brown*, 1982].

In designing a representational system for machine modeling, we need to consider: (1) the representation's primitives, the primary units of shape information used in the representation; and (2) its coordinate system, which defines the spatial relationships among the primitives. In particular, a robotic vehicle operating in the undersea environment must be designed to confront a three-dimensional world. For this reason, I confine the following discussion to three-dimensional representations (for further detail and expanded references to two-dimensional representations, see *Ballard and Brown*, 1982; *Winston*, 1984).

Representational Primitives

The fields of computer graphics, computer-aided design, and pattern recognition/image processing have extensively investigated the issues associated with representational schema, and the term *computational geometry* has come to identify the branch of computer science research dealing with the problems of representing, manipulating, and generating internal models of geometric objects [*Bajcsy*, 1980; *Posdamer*, 1981; *Srihari*, 1981; *Ballard and Brown*, 1982; *Winston*, 1984; *Besl and Jain*, 1985]. Broadly, three-dimensional geometric primitives may be categorized by four principal classes [*Posdamer*, 1981]:

- **Faceted representations**: Faceted representations approximate the bounding surface of an object. This is typically represented by a set of planar regions, each corresponding to part of the surface. Each region or face may be maintained as an ordered list of vertices, the connections between successive vertices being finite line segments or edges.

- **Functional representations**: Consider a function that generates points in 3-space as it is evaluated over a bounded range. There may be one, two, or three independent variables that generate a space curve, surface, or free-form solid. The surface may be used in a manner similar to the faceted representation, producing a surface model of patches joined at space-curve edges.

- **Cellular representations**: A cellular array is a regular spatial structure in which each cell is uniquely labelled by an integer triple (the indicial vector). The neighborhood associated with an indicial vector is its *volume element*, or *voxel*. Explicit geometry and implicit topology are specified by an enumeration of those cells occupied by the represented object. Such an enumeration may be specified by listing the cells, or by a three-dimensional array. Other methods of indexing include dope vectors, marginal indexing, and octrees [*Srihari*, 1981].

- **Procedural representations**: Procedural methods use solid primitives, parameterize the primitives to generate instances, and define operators for combining instances to form a model. The description of an object comprises a set of instances and the appropriate operators for combining them.

Coordinate Systems

In engineering terms, we usually think of coordinate systems being categorized as Cartesian, polar, spherical, and so on. However, a choice among these is largely determined by mathematical convenience, and transformations between any two can be carried out by purely mechanical operations. In the design of a representational system for machine perception, though, more fundamental issues must be faced when evaluating the tradeoffs between *view-centered* and *view-independent* coordinates on the one hand, and between *relative* and *absolute* coordinates on the other. Some researchers choose different combinations, and others dispense with a geometric coordinate system altogether.

View-centered coordinates offer a natural framework at the sensor level, and are appropriate to such low-level operations as image processing and feature extraction. *Marr's* [1982] primal sketch and 2½-D sketch use such a coordinate system since the data manipulation is intimately related to the process of image formation. The main problem with this approach at higher levels of processing—object recognition or terrain navigation, for example—is that the description of an object or a scene is sensitive to the viewer's position and attitude. Matching or correlation requires extensive search or iteration over all unconstrained translations and rotations.

View-independent coordinates are used to overcome this problem by establishing an *object-centered* or *global* frame of reference. In describing the shape of a highly symmetric object, a cigar, for example, the choice of an object-centered coordinate system is obvious and corresponds to the well defined principal axes. However, objects with many or poorly defined axes—like a sphere, a crumpled newspaper, or unstructured underwater terrain—lead to ambiguities [*Marr*, 1982]. Another issue in the use of such a *canonical* coordinate frame—a frame uniquely determined by the shape itself—is that the shape must be described before the frame can be set up [*Marr*, 1982].

In building an aggregate model of discrete objects, images, or static scenes, another choice must be made between a *common* (global, absolute) frame of reference and a *distributed* (local, relative) coordinate system. In the latter, each distinct entity has its own natural coordinate axes; these are usually linked by transformation matrices specifying the relative translations and rotations among the model's constituents. For high-level representations, there are advantages to this approach: each object's description is stable, unique, and completely self-contained [*Marr*, 1982]. However, at the lowest levels of modeling, which may comprise many primitives or bits of information, the extra overhead incurred by explicit representation of spatial relationships may extract a large penalty in computational performance.

Another alternative is to use a purely topological approach devoid of any geometric description. A *graph* or *network*, in which each node represents an object or primitive, can define the connectivity or adjacency in a relative sense. For certain high-level applications, this approach can distill the essence of spatial relationships and avoid much of the transformation overhead in a distributed coordinate system. For low-level representations, though, model size and processing efficiency must be considered. More significantly, the information content of such models is limited—there is no mechanism for describing the explicit geometry needed for many kinds of spatial reasoning.

2.4 MODELING UNCERTAINTY

In the first place, sensor-based methods are, by nature, probabilistic—prior information about the sensors, their media, and the conditions for research observation is always limited: that is, some properties can be described only by statistical methods. In the second place, acoustic signal and noise models are described by random processes or random fields [*Ol'shevskii*, 1978]; optical propagation and scattering in water are analogous [*Duntley*, 1963]. The inaccuracy of position sensors, quantization noise of a discrete processor, and algorithmic approximations add more uncertainty to the model.

With developments in the fields of AI and Robotics has come a realization of the complexities of the problems being addressed. This attitude has engendered a new look at the brittle, deterministic techniques of the early years and a trend toward methods that explicitly represent uncertainty. This springs partly from the need to account for less than ideal sensors and actuators, and from a recognition of the limitations in our coarse techniques.

There is also a growing appreciation for the central role of "fuzzy" representation and processing in a human brain evolved to contend with a highly dynamic and capricious environment. An inherent characteristic of the information available to human beings is that it is *imperfect* in the sense of being incomplete, uncertain, inconsistent, or otherwise not wholly suited to the judgmental task at hand

[*Stephanou and Sage*, 1987]. Computational approaches to machine perception must surmount the same obstacles.

A framework for Probability Theory has been in existence for over three hundred years [*Nutter*, 1987], though a rigorous formulation dates only to the beginning of this century, mainly based on the works of R. A. Fisher, A. Kolmogoroff, and R. von Mises [*Feller*, 1950; *Doob*, 1953]. Recently, though, many new theories have emerged that purport to overcome some limitations of classical probabilistic methods in the context of machine intelligence [*Stephanou and Sage*, 1987]. Among the objections raised against Probability as a tool for building intelligent systems are [*Cheeseman*, 1985; *Nutter*, 1987; *Stephanou and Sage*, 1987]:

- Probability is a frequency ratio, and each event has exactly one correct probability. For AI purposes, such a probability is neither attainable nor, in some cases, even interesting.

- The frequency theory restricts probability to those domains where repeated experiments are possible.

- The philosophical concept of *long run frequency* raises the questions: How long? How do you know? Why should *large numbers* (How large?) have special properties?

- The subjectivist view is based on the belief of an ideal rational subject, but: What makes someone an ideal rational subject? And how, other than by measured frequencies, can we establish the degree to which a subject ought to believe that an event will occur?

- Bayesian analysis requires vast amounts of data often unavailable or too expensive to obtain. The normal way around this is to guess.

- Prior probabilities assume more information than is given, and equate lack of evidence (uncertainty) with equal probability (from factual statistics).

Cheeseman [1985], a strong supporter of Probability, contends that all these objections can be overcome by a proper interpretation of the theory, and that no alternative framework is needed. *Stephanou and Sage* [1987] conclude that all returns are not yet in and that a definitive taxonomy of the different methods, including benefits and costs, is needed to assess the issues. *Nutter* [1987] takes a balanced view and finds a place for combined modes of reasoning about uncertainty. *Henderson et al.* [1987] report a general agreement among participants in a multisensor workshop on the use of probability models at the physics and geometry level, and that other methods may be more appropriate at the symbolic level. This thesis offers no attempt to resolve the controversy but, for the sake of later discussion, a brief overview of a few widely-used techniques follows.

Confirmation Theory [*Carnap*, 1950; *Salmon*, 1973] arose from long-standing inquiries into the nature of scientific induction—reasoning toward general principals from particular facts or instances. However, no purely logical validation of inductive reasoning has ever been demonstrated. At most, experimental results tend to confirm a theory and, in some cases, accumulated confirming evidence may elevate a general hypothesis to the status of, at least, provisional acceptability [*Salmon*, 1973]. For example, the law of conservation of energy is now taken to be a sound scientific generalization because of much confirming evidence and no compelling disconfirmations. Yet, such was the case with Newton's gravitational theory before being superseded by Einsteinian relativity. *Salmon* [1966] makes the point that induction is *ampliative*—that the whole (accepted general principle) is greater than sum of its parts (accumulated evidence). Confirmation Theory seeks to overcome this logical limitation by an incremental substantiation of an hypothesis with the accumulation of supporting evidence.

In the MYCIN system for evidential reasoning in medical diagnostics, *Shortliffe* [1976] heuristically develops the use of Certainty Factors (CF) as a direct outgrowth of Confirmation Theory. A measure of belief (MB), ranging from 0 to 1, is used to incrementally accumulate confirming evidence for a particular diagnosis; a measure of disbelief (MD) independently combines disconfirming evidence. A modified Bayesian combining formula aggregates MB and MD, and evidence is combined without regard to the order of acquisition. Then for each candidate diagnosis, a certainty factor is defined as: CF = MB - MD. This is interpreted as a confidence in the diagnosis, which ranges from -1 (complete disbelief) to +1 (complete belief). A CF of zero indicates complete uncertainty about the diagnosis.

One component of the model prescribes a method for the parallel combination of certainty factors that bear on the same hypothesis as:

$$z = \begin{cases} x + y - xy, & x,y \geq 0 \\ (x + y)/(1 - \min[|x|,|y|]), & xy < 0 \\ x + y + xy, & x,y < 0 \end{cases}$$

where x and y represent the independent certainty factors, and z is the effective certainty factor. *Horvitz et al.* [1986] point out that this combining rule is a specialization of probability because assumptions of conditional independence are imposed by the methodology.

Zadeh [1965] develops a rigorous Possibility Theory based on fuzzy sets. In essence, Zadeh extends the definition of sets to include continuous degrees of membership, and defines new set operations for manipulating them. This "fuzzification" of mathematical structures then leads to the concepts of fuzzy logic and inference. Unlike classical logic, fuzzy logic does not separate logic and probability [*Stephanou and Sage*, 1987]. Fuzzy reasoning has been adopted by researchers in a wide

range of fields [*Gaines*, 1976]. A chip for real-time reasoning with fuzzy logic has also been designed and fabricated [*Togai and Watanabe*, 1986].

In response to a need for representing imprecision in Bayesian probability values, *Dempster* [1967] introduced a concept of lower- and upper-probability bounds to contend with the subjective imprecision of uncertainty measures. *Shafer* [1976] extends this concept to a *Theory of Evidence*, and formulates a combining rule for cumulative evidence. In the Dempster-Shafer (D-S) approach, two separate intervals of uncertainty—*belief* and *plausibility*—can be assigned to each proposition. The D-S Theory of Evidence models the narrowing of the hypothesis set with the accumulation of evidence. *Zadeh* [1986] describes a simplified view of the approach and proposes an extension that links the D-S theory with a theory of fuzzy relational databases. *Gordon and Shortliffe* [1985] discuss advantages of the D-S theory over Certainty Factors, and show that MYCIN could be recast in a D-S framework.

2.5 HISTORICAL PERSPECTIVE

With this background, I devote the rest of this chapter to an overview of past and current works that relate to this dissertation. I start with a brief description of the most common techniques for image reconstruction from projections. As I mentioned in the introduction, these methods are influential to the approach I develop in the next chapter. Next, more general approaches to modeling by what I call *incremental reconstruction* are presented. Though the techniques described in this section are distinct from those preceding and from my own approach, the two sections together establish a context for incremental reconstruction from projections.

A complete survey of modeling and navigation in terrestrial robotics would be prohibitively long, but I discuss a few approaches and give references to alternative theories and implementations. In doing so, I offer a basis for comparing my work with similar efforts in the field. Some of the main components I point to are: the representational scheme and frame of reference, the dimensionality (spatial or multisensor), the representation of uncertainty (or lack of it), and the high- or low-level characterization of the approach. To conclude, I discuss a few treatments of the problem in autonomous underwater vehicles, and briefly outline where my work fits in.

2.5.1 Reconstruction from Projections

In 1917, the Austrian mathematician J. Radon proved (the Radon transform) that a two- or three-dimensional object can be reconstructed uniquely from the infinite set of all its projections [*Budinger and Gullberg*, 1974]. Since that time, the technique has been independently rediscovered several times

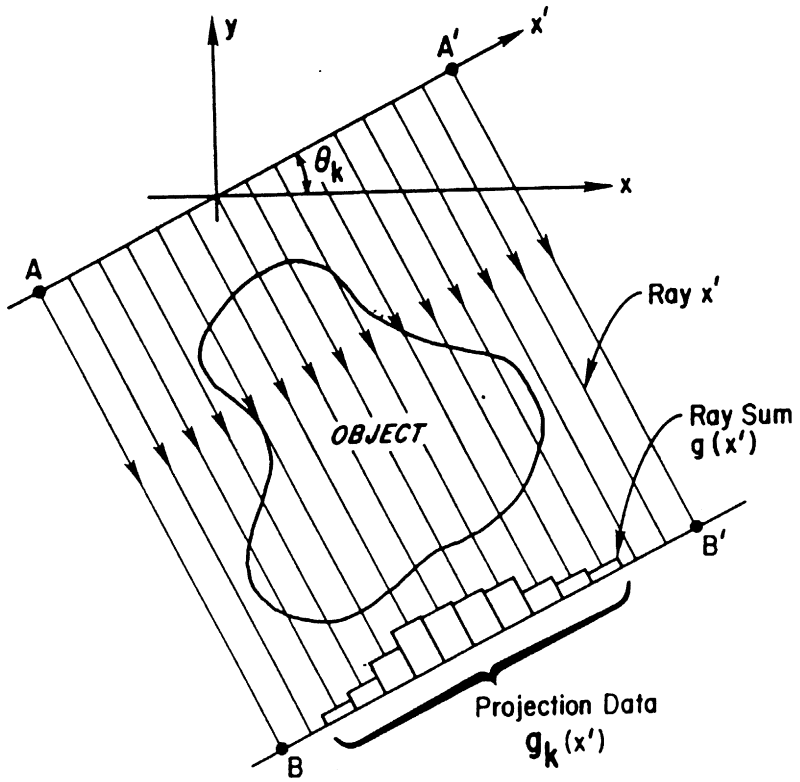

Figure 2.1: Parallel ray projection geometry.

[*Gordon et al.*, 1975], and applied to such diverse fields as medical X-ray imaging, nuclear medicine, electron microscopy, radio astronomy, and nondestructive testing (for readable tutorials, see *Gordon et al.*, 1975; *Scudder*, 1978; more complete discussions and extensive references can be found in *Budinger and Gullberg*, 1974; *Cho*, 1974; *Gordon and Herman*, 1974; *Brooks and di Chiro*, 1976; *Mueller et al.*, 1979). *Munk and Wunsch* [1979] showed that tomographic reconstruction could be used to monitor the speed of sound (and, by inference, density) over large regions of the ocean.

Digital reconstruction techniques attempt to approximate an object sampled by discrete projections. Most approaches model a three-dimensional object by stacking two-dimensional slices. However, *Mersereau and Oppenheim* [1974] show that the techniques can be extended to the reconstruction of multidimensional objects by successively applying the Projection Slice Theorem. Figure 2.1 shows the basic geometry of a one-dimensional projection that uses parallel-ray sampling (X-ray, for example) of a two-dimensional slice.

The *projection* $g(\mathbf{x}')$ of an ideal image $f(\mathbf{x})$ in the direction Θ is given by: $\int f(x',y')dy'$, where $\mathbf{x}' = \mathbf{R}_\Theta \mathbf{x}$ (bold typeface is used for vector notation: $\mathbf{x} = \{x,y\}$; and $\mathbf{R}_\Theta$ is a rotational transformation matrix). With a source at position *A* and a detector at *B*, the first projection datum is acquired. The line *AB* is called a *ray* and the measurement at *B* a *ray sum*. For an X-ray system, the ray sum is related to the integral of absorption (corresponding to density) along the ray. Data for the entire projection are obtained by moving the source and detector along *AA'* and *BB'*, sampling at fixed intervals. The lines are rotated by a small angle, $d\Theta$, and the process is repeated *N* times over an angular range of 180°. Reconstruction is an inverse problem stated as: given the projection data $g_k(x')$, $k = 0, \ldots, N-1$, construct the original image $f(\mathbf{x})$.

One possible solution is a simple matrix inversion or, if the inverse does not exist, a pseudoinverse can be taken [*Scudder*, 1978]. However, Scudder estimates that for 300 projections of 300 rays, displayed at a 300 X 300 resolution, at 1 μs/operation the inversion would take about 31 years to compute (or only 10 years or so on today's machines). Many more practical approaches to the problem have been developed—*Budinger and Gullberg* [1974] present 13 distinct categories of techniques. In general, though, these can be broadly classified as [*Brooks and di Chiro*, 1976]: (1) Backprojection (Summation); (2) Analytical Reconstruction; and (3) Iterative Analytical Reconstruction Techniques (ART).

Summation is the simplest method. Using a gridded image array, each ray sum, $g_k(\mathbf{x}')$, is distributed over all cells along the corresponding ray. For *M* cells intersected by the ray, each cell is incremented by $g_k(\mathbf{x}')/M$, a step called *backprojection*. When all rays are backprojected and the gridded array is normalized, the reconstructed image is an approximate version of the original. The result of reconstructing a point object with a discrete number of projections is a star-shaped object whose center lies at the location of the original point. This is the point-spread function (or impulse response) of a discrete backprojection.

Even with an infinite number of projections, the summation method does not produce an exact version of the original. The result is equivalent to convolving the original with the function $1/(2\pi r)$. To see this, consider a point object. The limiting case of superimposing equally-spaced straight lines through a common point is equivalent to rotating the line about that point. The weight of each point on the line is distributed over the length of the locus it sweeps out, in other words, the circumference, $2\pi r$. *Gordon and Herman* [1974] point out that a truly three-dimensional reconstruction by summation (over all spatial angles) would have an impulse response proportional to $1/r^2$. Blurring is lessened considerably because of the increased data.

The most common analytical techniques make use of the frequency domain. The Projection Slice Theorem [*Mersereau and Oppenheim*, 1974] shows that the Fourier transform of a projection is

equivalent to a central section of the transform of the original image rotated through the same angle as the projection in the spatial domain. In practice, the projections are transformed and assembled in a frequency grid, and an attempt is made to interpolate between the discrete slices. After interpolation, an inverse Fourier transform produces a reconstructed image with greater fidelity than a backprojected approximation. The Convolution-Backprojection technique partially ameliorates artifacts introduced by the Summation method. Though based on the frequency-space derivation of a convolution kernel, the method is applied in the spatial domain. The kernel (in two-dimensional space, an approximation to the Fourier inverse of $|r|$) is convolved with each projection, then the modified ray sums are backprojected.

ART [*Herman et al.*, 1973] and other iterative techniques are applied to the solution of an undetermined system of linear equations in the spatial domain. The algorithm consists of iteratively correcting the discrepancy between the measured ray sum and a calculated ray sum from the image generated by the previous iteration. Variations using additive and multiplicative corrections exist. Important modifications that take advantage of working in the spatial domain use the constraints that no image cells can have negative values, and all cells along a ray whose sum is zero must also be zero. Other versions of the technique use a Bayesian approach to incorporate prior knowledge of the object being reconstructed [*Hanson*, 1987].

Das and Boerner [1978] develop a novel approach to shape estimation of convex bodies using radar returns from multiple, nonorthogonal look angles (in a plane). Based on an application of the Radon transform, the authors use target signatures (normalized backscatter ramp response) to extract successive area projections. Numerical studies with published target signatures of a sphere (and assumed perfect registration) showed promising results. *Rockmore et al.* [1979] point out limitations of the approach and describe a method developed from a three-dimensional version of the Projection Slice Theorem (attributed to *Mersereau and Oppenheim* [1974]). The authors derive a convolution-backprojection algorithm and contend that the Das-Boerner method is a special case.

Denton et al. [1978] describe an approach to the target-association problem—identifying multiple targets viewed by multiple sensors—as an image reconstruction problem. A convolution-backprojection method (a precursor to that of *Rockmore et al.* [1979]) is developed for a three-dimensional case of active radar and a two-dimensional case of passive infrared. At a given target range, the authors use a finite plane, bounded by the sensor's beam pattern, as a simplifying approximation to the curved, shaded range surface. Monte Carlo simulations show the star-shaped impulse response caused by sparse projections, but otherwise demonstrate results that the authors claim are equal to or better than search-based algorithms. They indicate that, although fast performance was not a research issue, the technique appears to be implementable in real time.

Rockmore [1981] extends these ideas to the ocean domain and proposes the use of multiple, passive sonar arrays for surveillance, target localization, and mapping the acoustic space-time noise field. He suggests that the (steady-state) signal received from any direction could be modeled as a ray sum, and that rays from multiple look directions with one array would constitute a projection with fan-beam geometry. From multiple arrays, multiple projections would provide the data for reconstructing the acoustic emission field. In justifying his proposal, Rockmore writes:

> The more conventional method for performing this type of surveillance is to perform the frequency processing and thresholding on a per-array basis, and then combine the threshold exceedances geometrically. This procedure of thresholding *before* all signal processing is performed is in violation of sound engineering practices. Thus the performance of this technique will be generally inferior to that of tomographically combining prior to decision making.

Rockmore goes on to discuss such problems needing resolution as: beam-pattern effects—the rays are no longer lines; multiple look angles to "deconvolve" the beam pattern; sparse projections; and noise.

Norton and Linzer [1979a, b] discuss reflectivity tomography for ultrasonic medical imaging with circular and spherical arrays of omnidirectional transducers, and present a comprehensive theoretical analysis that leads to a convolution-backprojection approach to reconstruction. The authors' major assumptions are: (1) weak scattering (Born approximation); (2) uniform attenuation from absorption (which can be compensated); (3) uniform sound velocity; and (4) the object can be modeled as a collection of isotropic scatterers. A first-order expansion (shown to be valid near the center of the transducer array) is equivalent to the local intersection of linear projections in the circular case, or planar projections for the spherical array. Aside from the normal monostatic (backscattering) geometry, the authors offer a theoretical treatment of the bistatic case (over a limited range of scattering angles) in which the projection integrals are taken over elliptical range arcs.

2.5.2 Incremental Reconstruction

A limitation of the techniques just described is that they need a fixed, regular scanning geometry for analytical and computational tractability (an exception is the summation method). *Horn* [1978] describes a method for arbitrary scanning geometries, but requires that they be fixed so a convolution kernel can be derived for each geometry. However, a mobile robot cannot always adopt a regular sensing strategy, but must take an opportunistic approach to whatever information becomes available. This alternative I call *incremental reconstruction* and I describe a few three-dimensional techniques. Other, more complete robotic applications using this approach are discussed in the next section. Further references and discussion of modeling with range data can be found in *Jarvis*, [1983] and *Besl and Jain*, [1985].

Martin and Aggarwal [1983] develop an algorithm for volumetric modeling from successive two-dimensional boundary constraints. *Occluding contours* (the boundary on an object silhouette in the image plane) from thresholded camera images are used to refine a *volume-segment* representation comprising linked segments, lines, and planes. *Connolly* [1984] uses simulated range-image boundaries to incrementally construct an octree representation of three-dimensional objects. The gridded range image is converted to an intermediate quadtree representation, which is projected into the octree model. Later work [*Connolly*, 1985] develops a strategy for finding the most efficient sequence of views for model building. *Veenstra and Ahuja* [1985] project the silhouettes directly into an octree, but require nine views (corresponding to six faces of a cube and oblique views at three corners) for efficiency. All these approaches assume perfectly registered data and no uncertainty.

A long-term project at the National Bureau of Standards has produced a sophisticated modeling approach to managing a manipulator's workspace for manufacturing robotics [*Shneier et al.*, 1984, 1986; *Hong and Shneier*, 1985]. Object silhouettes from a moving camera are projected into an octree structure as *generalized cones*. The intersections of cones from multiple views successively constrain the object boundaries and implicitly represent uncertainty in size and position of each object. This technique is only one component of a high-level world modeling system from which information also flows "downward" to help resolve ambiguities.

Faugeras [1984] uses a high-level approach to modeling three-dimensional objects with range data. Planar and quadric patches are fitted to segmented clusters of range points and accumulated in a *region adjacency graph*. This contains information about points in a region, borders, and neighboring regions. Hypothesis prediction and verification, implemented in a tree search, is used for matching and localization of objects. A probabilistic formulation [*Faugeras*, 1987] uses planar primitives and motion from passive stereo vision to combine information from several viewpoints. *Amblard et al.* [1986] propose a technique for three-dimensional surface estimation from multiple stereo-camera views. The surface is modeled as planar triangles related by Markov Random Fields.

2.5.3 Terrestrial Robotics

At Stanford Research Institute in the late sixties, the mobile robot SHAKEY [*Rosen and Nilsson*, 1968; *Nilsson*, 1969; *Coles et al.*, 1969; *Rosen*, 1970] became an early testbed for research in autonomous intelligent vehicles. The high-level world model uses a gridded spatial structure, which later evolved into the quadtree representation (Rosen notes, however, that a fully divided array is more suited to such applications as path planning). The gridded model is augmented by an object-oriented *property list*, simplistically, the LISP equivalent of multidimensional feature space. Machine vision and

dead reckoning gave SHAKEY a rudimentary ability to navigate, explore, and "learn" about its environment.

Begun in 1977 at the French Laboratoire d'Automatique et d'Analyse des Systemes, research with HILARE resulted in one of the most complete and consistent mobile robot implementations to date [*Giralt et al.*, 1979, 1983; *Briot et al.*, 1981; *Chatila and Laumond*, 1985]. Using multiple sensors including video, laser, sonar, and infrared, the robot maintains an uncertain, dynamic world model. Descriptions of polyhedral objects, each with its own reference frame, are maintained in a geometry database. These are projected on a plane, and the polygons linked to form a graph of *places*. The authors introduce the idea of *fading* to propagate accumulated positioning error backwards through the topological representation. A semantic model maintains property information about objects and places so that distributed decision makers, in the form of expert processes, can access the model database for navigation, planning, and task supervision.

From the early seventies at Stanford University's Artificial Intelligence Laboratory to the present at the Carnegie-Mellon University (CMU) Robotics Institute, *Moravec* [1980, 1983] has evolved many elegant modeling approaches for mobile robots. Using stereo and, more recently, sonar to correct dead-reckoning errors, a unique, low-level approach incorporates the powerful notion of mapping probabilistic sensor distributions into model space [*Moravec and Elfes*, 1985; *Elfes*, 1986a, b]. Certainty Factors and a modified MYCIN combining formula [*Shortliffe*, 1976] are used to incrementally merge thresholded sonar returns into a stochastic, two-dimensional gridded representation. Correlation techniques are applied to compare the grid with a prior model for localization. What have come to be called *Certainty Grids* [*Moravec*, 1987a, b] have been successfully implemented by the CMU group for map making, path planning, and navigation [*Thorpe*, 1984; *Serey and Mathies*, 1986].

In a parallel effort at the CMU Robotics Institute, *Crowley* [1985a, b] used a rotating ultrasonic sensor to build a two-dimensional *composite local model* of the world. Line segments are extracted from sonar readings, adjusted with a recursive line-fitting technique, and used to update the dynamic model. A side effect of matching line segments is an error vector representing a position correction. *Drumheller* [1985] uses a similar model of line segments but applies a search-based technique [*Grimson and Lozano-Perez*, 1983] for matching segments. Drumheller reports good results despite noise and error from specular reflection, multipath, and the wide beam pattern.

Miller [1984, 1985] develops a technique for navigation and path planning in a two-dimensional space. Using a wide-angle sonar aboard a mobile robot, Miller applies search techniques and heuristic pruning to match range returns with linear features (lines, edges) of a prior model. Miller reports good performance with noisy real-world data, and enough speed to accommodate sonar data rates.

More recently, *Miller and Slack* [1987] develop a theory of *message passing* among nodes in a linked hierarchical grid for navigating in a dynamic two-dimensional space. *Bixler and Miller* [1987] propose a technique that combines vision and sonar to extract linear features for updating a wire-frame grid map of the world.

Flynn [1985] describes an application of mapping and path planning with data from sonar and infrared sensors. Using physical models of the sensors, Flynn incrementally combines real data to overcome the limitations of each sensing modality. Near edges (doorways), the better angular resolution of the infrared sensors (poor range resolution, though) alleviates the blurring introduced by the wide-angle sonar. From the refined map, an intermediate *curvature primal sketch* is extracted. In turn, this is converted to a two-dimensional polygonal representation for path planning.

Brooks [1985a, b, 1986] develops a philosophy that rejects the brittle approaches of earlier work and focuses on the realities a robot must face in a complex, dynamic environment, and on the tools to make such "artificial beings" practical. According to Brooks: errors and inaccuracies in sensors and actuators must be considered; real robots must be adaptable and tolerant of sensor failure; multiple sensors and different levels of resolution are needed; and three-dimensional representations are essential in a three-dimensional world. Brooks also argues strongly against the use of absolute coordinate systems because of cumulative errors. He contends that a relative framework is more useful, and that the design space for perception systems must reflect this.

Rao et al. [1986] address the problem of autonomous robot navigation in unexplored terrain. They develop a theory of concurrent navigation and learning of the environment, and discuss performance considerations. Their two-dimensional representational scheme uses a *modified adjacency list*, a graph structure linking labeled polygons, with no mechanism for representing uncertainty. *Marce and Julliere* [1986] discuss an alternative method for navigating in two dimensions that depends on prior knowledge in a gridded world map. Simulations of a laser rangefinder demonstrate position and heading determination using a point-matching technique.

Kuipers and Byun [1987] describe a qualitative approach to learning a topological map of *distinctive places* in a two-dimensional environment. The authors simulate a mobile robot by *hypothesizing* a sensor system and sufficiently distinct environmental features to overcome noise and error, but offer no real-world results or quantitative simulation of uncertainty. *Levitt et al.* [1987] develop a theory of *Long Term Memory* for a mobile robot, which represents visual landmarks in a distributed topological framework. The authors describe a simulation laboratory and results that suggest the utility of the techniques.

Dean [1987] develops an approach that combines elements of a *subsumption architecture* [*Brooks*, 1985b] and certainty grids [*Moravec*, 1987a, b], and points to a forthcoming implementation. *Strat and Smith* [1987] describe the *Core Knowledge System*, an architecture combining elements of relational and spatial databases with uncertain reasoning, as a basis for a forthcoming application with an autonomous robot.

Saridis and Valavanis [1987] propose the use of *entropy* as a common measure of uncertainty for organization, coordination, and execution in autonomous systems. *Farreny and Prade* [1987] describe the use of fuzzy techniques to contend with robotic uncertainty in action, perception, communication, and reasoning. *Smith et al.* [1987] use state-estimation methods to propagate error in a distributed coordinate system. *Durrant-White* [1988] develops a rigorous *stochastic topology* and describes the consistent transformation of uncertainty between relative coordinates at different nodes.

In reviewing these works chronologically, two trends seem clear: the use of multiple sensors, and the adoption of probabilistic methods. I think this reflects the heightened awareness among robotics researchers that there are no simple, all-encompassing techniques to circumvent the realities of a complex, uncertain environment. On other accounts, the kinds of representations, reference frames, and levels of approach reflect the research tracks of different investigators and the varied constraints and opportunities of different applications. On the one hand, this is a natural state of affairs for an immature field in which the methods have not yet converged toward preferred theories or implementations. On the other hand, it is unlikely there will ever be one *optimal* representational or processing paradigm. What this thesis strives for is a generality, consistency, and flexibility that is useful over a range of applications and sensory domains.

2.5.4 Underwater Robotics

Though the history of autonomous underwater vehicles extends back at least two decades [*Busby*, 1981; *Bane and Ferguson*, 1987], these first AUV's were no more than preprogrammed, free-swimming idiots. With new technology and rising military interest, though, more sophisticated AUV's have begun to appear. However, many approaches are taken directly from terrestrial methods, which may not be the most suited to the undersea domain. Most descriptions in the literature focus on power, communications, and other hardware issues, but I discuss a few of the more recent approaches that include information on modeling and representations. Broader overviews and more detailed presentations on missions and technology are given in [*Krabach*, 1983; *Thomas*, 1983; *Wang*, 1983; *Eppig*, 1985; *Stenovic*, 1986; *Bane and Ferguson*, 1987].

One of the first "intelligent" AUV's was the EAVE vehicle of the University of New Hampshire Marine Systems Laboratory [*Blidberg et al.*, 1978; *Blidberg*, 1984, 1986; *Chappell*, 1987]. The system has evolved considerably in the last decade and a new effort with the National Bureau of Standards will further enhance its capabilities [*Albus and Blidberg*, 1987]. The NBS plans to implement a high-level world model, which will include quadtree maps from prior surveys. *Orser and Roche* [1987] describe a technique for extracting topographic features (ridges, passes, ravines) from such prior maps. In a later implementation, information from multiple sensors will be used to add objects to the global model or remove them if perceived to exist no longer. Recognized objects also may have associated confidence factors, and degrees of believability and dimensional certainty.

The ANGUS vehicle is an early testbed that helped shape a sophisticated AUV program at Herriot-Watt University [*Russell et al.*, 1983]. *Russell and Lane* [1985, 1986] describe a knowledge-based system with a blackboard control structure for sector-scan sonar interpretation. The system uses the planar-bottom assumption to generate a two-dimensional sonar image taken from a stationary platform. Rule-based knowledge sources direct the application of image-processing operators to the data before interpretation. Higher-level routines evaluate such features as shadows and artifacts caused by sidelobes, multipath, and reverberation. Segmented features are tracked between images before being accepted as bona fide objects, correlated with a prior model, then labeled with size, position, orientation, and other symbolic information. The authors describe simulation results with real sonar data, and emphasize the utility of multiple representations (one-, two-, and three-dimensional) and the interaction of model- and data-driven processes.

Cuschieri and Hebert [1988] describe a simple shape-from-shading technique [*Ballard and Brown*, 1982] using sidescan sonar. Each slant-range-corrected sonar line is segmented into shadows. Shadows preceded by a high-intensity signal are assumed to indicate positive relief; shadows followed by high-intensity are taken to indicate trenches or gullies. The shadow length is used as a measure of change in relief to generate contours. These are smoothed and merged with neighboring lines using a least-squares technique. The result is an estimated bathymetric map of the imaged area. This map is passed to an image-processing routine that extracts linear hills, valleys, and ridges in a manner similar to that of *Orser and Roche* [1987]. Other researchers addressing sidescan applications for AUV's are *Glynn* [1985], and *Nichols and Jensen* [1985].

Chandé and Noon [1986] describe the proposed sensory subsystem for obstacle avoidance and navigation of the Martin Marietta AUV. The vehicle will incorporate a scanning laser and sonar array, with navigation supplied by an acoustic transponder net underwater and Loran-C or Omega on the surface. *Zonal spots* from multiple range sensors will be tracked with a Kalman filter, clustered, and compared with a stored pattern to infer distinct objects. The extracted objects will then be referred to a high-level modeler that maintains a description of the environment.

Trimble [1987] focuses on the multiprocessor architecture of the Lockheed AUV, but offers a brief description of its sensory capabilities. A *Scene Awareness* function will interface with a *Mission Manager* to provide high-level sensor management and process scheduling. This will include multisensor correlation and knowledge-based algorithms that deal with incomplete information. Acoustic and optical information will be interpreted in real time using signal processing and symbolic processing, deterministic and probabilistic. The system is intended to judge the quality and correctness of information, and may support some form of geophysical navigation. Other high-level goals include stationary-object location and identification.

Other ongoing AUV efforts include: the EAVE-WEST vehicle of the Naval Ocean Systems Center (San Diego) [*Heckman*, 1980; *Durham et al.*, 1987]; the EPAULARD vehicle of the French IFREMER [*Borot et al.*, 1983; *Jarry and Michel*, 1985; *Michel*, 1987]; the ARCS and DOLPHIN vehicles of International Submarine Engineering [*Ferguson and Jackson*, 1983; *Jackson*, 1983; *Thomas*, 1985; *Butler and Maryka*, 1987]; the SIMRAD Freeswimming Prototype [*Klepaker et al.*, 1986, 1988]; and several Japanese efforts [*Collins*, 1987].

Judging by the recent literature, it seems clear that AUV research is beginning to address issues more germane to intelligent systems, in parallel with the ongoing efforts to resolve more basic hardware difficulties. My perspective, though, is that much of the work too closely mirrors the approaches of terrestrial robotics. Granted, there is much to be learned from the accumulation of terrestrial robotics research, particularly at the higher, more cognitive levels. However, the more hostile undersea environment presents a greater challenge to intelligent systems in several ways: it is a fully three-dimensional world in all respects; environmental features are more unstructured, or amorphous; and there is more inherent uncertainty in sensing, interpretation, and localization. I elaborate on these and other constraints in the next chapter, and use them to guide the formulation of a low-level approach to modeling underwater.

Chapter 3

MODELING THE UNDERWATER ENVIRONMENT

In this chapter, I start to formulate a philosophy of multisensor modeling underwater and describe a specific framework for implementation. My primary concern is to establish a computational architecture broad enough and flexible enough to encompass a diverse range of applications. At the same time, the structure and processing must efficiently manage realistic data rates and imperfect information. Later chapters show the utility of the approach with results from computer simulations and applications with real-world data.

Using the context established in the previous chapter, I first summarize the modeling constraints specific to an underwater environment. These are used to define a representational scheme and low-level modeling processes. Computational and architectural issues are examined and used to help define the general analytical approach. The framework is applied to active sonar sensing and model building in the acoustic domain, using both binary and continuous models. I conclude with a summary of important points raised in the chapter.

3.1 SPECIFIC UNDERWATER CONSTRAINTS

As surveyed in Chapter 2, the rich body of literature on terrestrial robotics has been strongly influential in shaping a modeling paradigm for intelligent, autonomous machines. However, in extending these ideas to the subsea domain and to the practical requirements of remote underwater systems, the following considerations are noted:

- The land rover's environment is usually modeled as a two-dimensional plane allowing one rotation. For a free-swimming vehicle, the world is unarguably a three-dimensional space—six degrees of freedom must be accommodated.

- Typically, there is a prior world description used by the robot to find its relative position. For an exploratory probe, though, such models will be largely unavailable. A more general approach must treat mapping and positioning as concurrent processes.

- Most work has been confined to highly-structured environments (laboratories) with regular geometric features (smooth walls, straight edges). A practical technique for underwater use must be consistent with the more irregular, amorphous features of the environment as well as the more tractable shapes of man-made objects.

- There is more uncertainty in the relatively opaque and hostile underwater environment. Both acoustic and optical sensors are hampered by greater attenuation, distortion, and noise introduced by an inhomogeneous, dynamic medium. Currents, turbulence, and other physical disturbances preclude the assumption of a static position and attitude, even when propulsion is inactive.

- Real-time constraints are often absent. Frame rates of stereo-image processing systems, for example, are often measured in seconds, even minutes. A practical underwater system must absorb sensory data at practical rates so that higher-level processes can respond in a timely manner.

3.2 REPRESENTATIONAL DESIGN

As indicated in the last chapter, a fundamental component of modeling is the representational scheme. The overall approach taken in this thesis is decidedly low-level, as defined in Section 2.3.2. There are several reasons for this: (1) an emphasis on sensors and physical processes; (2) development of a foundation for higher-level processing; and (3) opportunities offered by new computer technology.

First, some interpretation of sensor data must take place at the physical level in any realistic implementation. Considering an active device such as a sonar or a laser, for example, the information content of raw sensor data depends on the low-level characteristics of the sensor, on the physical nature of energy propagation through the medium, and on the interaction of that energy with environmental features. Passive sensors are essentially no different. At the geometric level of interest, we are concerned primarily with shapes, boundaries, discontinuities, and so on. Though higher-level knowledge might be useful in guiding this interpretation, it is often unavailable.

Second, a firm low-level foundation can facilitate the construction of high-level models. The human visual system incorporates early processing of raw optical data before passing the enhanced information to cognitive regions of the brain. For a machine "intelligence" also, such preprocessing can filter raw sensor data and distill its essence so high-level processes may avoid an information overload. Still, a purely low-level approach is an inadequate solution to the problems a fully autonomous system must address—planning and reasoning about the world, for example. My intent has been to investigate the properties of such a low-level structure as the basis for future elaboration at higher levels.

Third, there are many unexplored avenues for research at this level. Most AI/Robotics work has managed information at the feature level or higher, partly because of the constraints imposed by previous generations of computer hardware. With the advent of cheap 32-bit computing—machines with extended address space and fast, inexpensive memory devices to match—numerical approaches to modeling become an attractive possibility. In discussing his choice of a low-level representation, *Moravec* [1987b] writes:

> Despite its effectiveness, in each instance we adopted the grid representation of space reluctantly. This may reflect habits from a recent time when analytic approaches were more feasible and seemed more eloquent because computer memories were too small to easily handle numerical arrays of a few thousand to a million cells. I think the reluctance is no longer appropriate. The straightforwardness, generality and uniformity of the grid representation has proven itself in finite element approaches to problems in physics, in raster based approaches to computer graphics, and has the same promise in robotic spatial representations.

Representational Primitive

To make an intelligent choice among the different primitives for three-dimensional modeling requires some elaboration of the physical and computational analog. In a broad sense, this includes the features of an underwater environment, the stochastic nature of modeling, and the characteristics of a digital-processing approach. Such constraints must be considered in the context of the goals to be achieved.

Environment: For navigating an underwater vehicle in three-dimensional space, the simplest model must describe the presence or absence of solid matter in enough detail to discriminate among different features of the terrain. In contrast to the structured environment often presented to a land rover, underwater seascapes are usually characterized by irregular, amorphous features. This constraint tends to preclude the use of procedural schema since a prohibitively large repertoire of primitives would be needed to describe the environment directly. The problem might be circumvented by using many small-scale instances to approximate natural terrain; but with increasing detail, the

model approaches a cellular partitioning. Similar arguments apply to faceted representations, which need many polygons to model complex surfaces.

Uncertainty: Faceted descriptions and boundary representations, in general, offer no direct method for representing uncertainty. When accommodated, error and inaccuracy are often specified as a covariance matrix associated with coordinate transformations of the vertices or control points. However, a characterization of sensor uncertainty by error distributions suggests a functional model could be applied to the geometry as well. It is unlikely, though, that one function describing an expectedly complex model could be found. A set of vector functions might be formulated to describe the different features and associated uncertainty; each could be valid over a bounded range within the model, similar to a faceted representation. However, the level of complexity grows with the level of detail required.

Process: Real-time performance in a practical implementation puts a premium on computational efficiency, which is strongly dependent on the underlying digital representation. The dynamics of the process—the model evolves as new information is incorporated—raises questions about the efficiency of a functional representation, even at a low level of detail. To merge new information with a model, each descriptive function in the affected volume of space would have to be reparameterized. Preservation of continuity dictates reconsideration of neighboring functions and, perhaps, a new partitioning of functional domains. Potentially, this is a highly expensive computational process.

Though the most effective models are likely to be application specific, an extendibility to other domains or dimensions is useful. For example, the 3-space decomposition used by a bathymetric sonar modeler should also be accessible to a two-dimensional intensity mapper processing sidescan records. A higher-level interpreter could use the slope and amplitude information to estimate localized scattering properties of the surface material. A decomposition by frequency space might be more suitable for multispectral fingerprinting with different sonars or other sensors.

Some functionality over a dynamic scale of range and resolution would enhance the assimilation of sensory data with different granularities. Interpretive and display processes could adopt a coarse-to-fine algorithm or a strategy that trades fidelity for real-time execution. Degraded performance from higher uncertainty, coarser resolution, or reduced computational assets should occur in a generally monotonic way with no catastrophic thresholds to be crossed.

These considerations argue in favor of a cellular model (or, more generally, a vector decomposition of n-space). Coherence is satisfied, it is continuous in space and time, and mirrors the physical structure of the environment as well as the internal structure of a discrete, computer-based system. It is amenable to simulation in the computational sense because it offers an efficient

representation for mapping, positioning, and imaging procedures. And the implicit spatial relationships among a model's constituent voxels afford a level of simplicity and efficiency difficult to achieve with other representations.

Coordinate System

The choice of reference frame is a thornier problem; the tradeoffs between relative and absolute coordinates are not easily resolved for a single, general approach. *Brooks's* [1986] argument against absolute coordinates—that cumulative positioning error corrupts the spatial relationships in such a frame of reference—is a persuasive contention. However, the same error will be present in any model derived under the same conditions, unless it can be removed by external (absolute) position fixes. The real issues are: how efficiently does the representation lend itself to coping with that error, and how can the model be adjusted when a fix becomes available?

Our own internal model probably uses a combination of the two reference frames. Certainly we have some sense of absolute position, and it is most evident when we perceive visually. Yet, with our eyes closed, the richness of description fades quickly and we are left with the coarse, underlying representation that probably guides our more basic cognitive processes [*Marr*, 1982]. At this level, the surroundings are remembered more generally as distinct objects with sparse features and more generic shape, and with "fuzzier" measures of spatial relationships attached—the "tall" cabinet "across" the room is "next to" the door.

For a high-level machine model in which the environment has been characterized by such objects, this topological approach is manageable, perhaps preferred. However, for the low-level approach developed in this thesis, a global coordinate system is used, except for a few instances in which a local (or view-centered) frame is adopted for analytical convenience at intermediate steps. My reasons for doing so are: (1) efficiency and consistency; (2) response to constraints imposed by unstructured terrain; (3) consideration of requirements for real-world applications.

I have already emphasized the need for computational efficiency and the advantages of a low-level structure that implicitly represents spatial organization; this is consistent with the cellular structure selected for a modeling architecture. That choice, in turn, is driven partly by a need to deal with the amorphous shapes of natural underwater terrain. Such an environment is ill suited to purely object based methods. Distinctive measures of shape are more difficult to formulate and apply, and the boundaries between natural features are not clearly defined.

My main concern, though, is to address the requirements of real-world AUV's. For a robot whose only purpose is to wander the hallways without getting in trouble, it may suffice to know only that a feature is "large" or that it is "next to" another (such concepts are more important to semantic

reasoning). But many practical applications—exploration, mapping, survey, inspection—depend on an absolute frame of reference. Other missions require an AUV to reach a specific objective, perhaps after an extended traversal, and return to its starting place. Such systems are intended to serve a human user who demands more continuous, quantitative information about size, shape, and location in the real world.

3.3 ARCHITECTURAL OVERVIEW

So far, I have established that the models with which this thesis is concerned: (1) are analogic; (2) are low-level; (3) use a volumetric primitive (two-dimensional variations using an area element are also considered); and (4) use an absolute frame of reference. I have justified these choices mainly in terms of the modeling applications—intelligent systems, the underwater environment, real-time constraints, and so on. In later chapters I offer further substantiation in the context of particular applications; but for now, I leave these issues and discuss the two remaining representational components important to this work: feature vectors and spatial indexing.

Unlike the first four components just mentioned, which provide a constant and consistent framework for all that follows, the last two depend on the modeling application, and even vary with the different processing components of the same application. For that reason, I maintain a generality in this section, the last devoted mostly to representational issues. I start with a description of a model as a set of feature vectors that represent the environmental parameters of interest, then briefly discuss spatial relationships and spatial indexing. Next, the flow of information from measurements to models is considered, and the processing steps are formulated as a sequence of transformations in which vectors are mapped from one representation to the next. In this section, I also develop a general notation used for the rest of the thesis. With this as background, the rest of the chapter begins to focus on processing—the other half of the representation/processing duality—and I offer more concrete examples of modeling specific sensors.

3.3.1 Vector Modeling

A global model of an underwater region can be formed by dividing it first into regular cubical *volume elements*, or *voxels*. If the division is fine enough, then one value can be given to any property that we want to consider within each voxel. For example, to represent the distribution of acoustic scattering strength over an underwater terrain, I use the *feature vector:*

$$\rho = \{t,x,y,z,\rho,\sigma_\rho\}$$

where t is time; x, y, and z locate the voxel in 3-space; ρ is an estimate of how much sound energy will be reflected back to a sonar receiver; and σ_ρ expresses our confidence in that estimate.

The curly braces imply that you can also consider each vector as a *set* of features associated with each small volume element; then, a *model*, or *model vector*, comprises a set of such feature vectors, for example:

$$\mathbf{M} = \{t,\mathbf{x},\rho,\lambda,\boldsymbol{\sigma}\}$$

where: $\mathbf{x} = \{x,y,z\}$

$\boldsymbol{\sigma} = \{\sigma_\rho,\sigma_\lambda\}$

and λ is an optical reflectance parameter, perhaps derived from a scanning laser. A representation of higher dimension might include camera images, temperature, salinity, magnetic field, and so on. As above, I use bold typeface as a vector shorthand, and the bold, upper-case **M** denotes a *global model*. By global model, I mean the set of all feature vectors over the region being modeled; a *local model* is just some subset. It is consistent to consider a feature vector that describes a single voxel as a local model; however, I use the term to designate larger subsets for specific purposes.

If we bound the region to be modeled and consider a fixed voxel size, then the model constitutes a finite set. Also, I consider only digital representation and processing in this thesis (discrete notation is used only occasionally for clarity), and that each feature has a finite resolution and range of possible values. With these conditions, the size of the model is deterministic—the computer storage needed has an upper bound determined by the range, resolution, and number of different features. These conditions are not required absolutely in any of the development that follows, but the property of determinism is useful because computational resource requirements can be forecast. I elaborate on this point in later sections.

In the model vector above, I make all information explicit by including the spatial coordinates of each voxel. For some types of processing, this is a convenient representation. For example, a *point process*, which requires no information about neighboring voxels, could operate on a list of feature vectors. Simple thresholding is such a process: the value of a feature at some point is compared with a threshold value that does not depend on any other point. Convolution is a *regional process*: point values are multiplied by coefficients that depend on spatial (temporal) relationships within a region.

For such operations it is usually more efficient to arrange the values in a spatially organized structure such as an array. Though convolution can operate on a list, repetitious traversals of the list to identify neighboring voxels would be computationally burdensome. However, a three-dimensional array is not the only choice to represent a three-dimensional model. As mentioned in Section 2.3.3, an

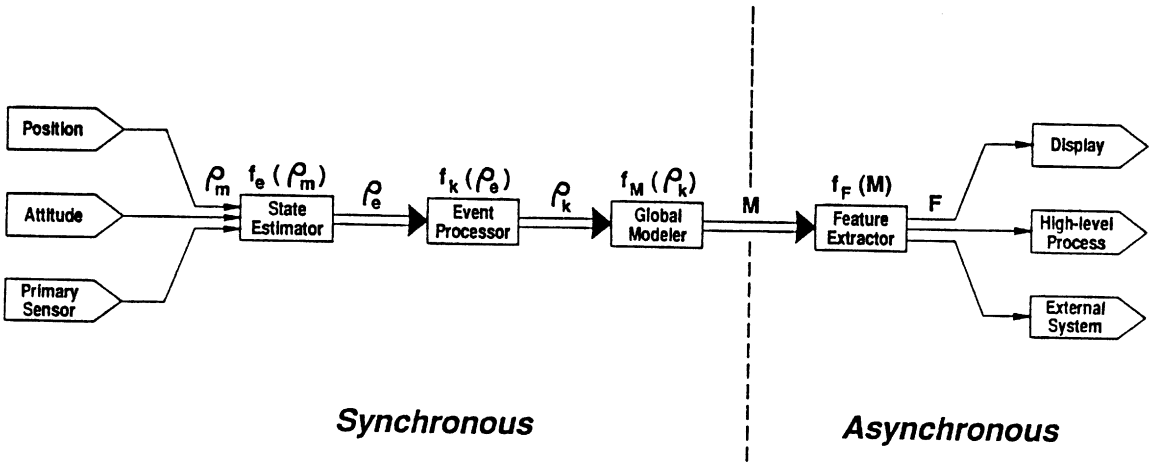

Figure 3.1: Information flow for open-loop modeling.

array only provides a technique for spatially indexing the values associated with each voxel. The spatial *relationships* are implicit in the cellular decomposition; the spatial *index* may be implicit in an array (by virtue of spatial relationships among memory locations), or explicitly represented in a list of vectors.

For the computer implementations to be described in later chapters I maintain the global model as a three-dimensional array of vectors. The choice is made for convenience and simplicity in programming development, but other indexing schemes are also applicable; the tradeoff is usually efficiency against storage (I briefly mention the use of octrees in Chapter 7). For local models and other intermediate representations, I use 2-D arrays, 3-D arrays, 2- and 3-D arrays of vectors, and vector lists of different kinds. In each instance, the choice is one of convenience or efficiency. Regardless of the underlying computer implementation, I use the explicit notation described above as a reminder of the representational flexibility.

3.3.2 Open-Loop Modeling

In the last chapter I mentioned the closed-loop nature of concurrent modeling and positioning. For the rest of this chapter, though, I consider the more straightforward problem of building a model when position measurements are available. The information flow in a general, open-loop modeling process is depicted in Figure 3.1. The flow starts with a stream of measurements, which are filtered by a state estimator before being passed to an event processor.

I define a sensor *event* as a spatially or temporally contiguous set of measurements (or state estimates). Considering a pulsed sonar, for example, an event starts with an outgoing pulse and comprises the interval ending with the start of the next pulse. During that interval, the signal passes through a region of space and may be scattered or reflected back to the receiver by one or more targets. The boundaries of that region are mainly determined by the sonar's beam pattern and effective range, and only within that space can we detect any features of interest. A pulsed-laser event is defined analogously, and the *detection envelope*, which bounds the region in which information may be derived from that event, encompasses an almost linear segment of space. The detection envelope of a temperature probe approaches a point, and an event comprises a single digitized sample.

I elaborate on events and detection envelopes in Section 3.4 and describe how events are backprojected into the model in Section 3.5; but, for now, it is sufficient to picture a stream of measurements and events from which a model is built. For an open-loop system, in which the sensors are not under high-level control, there is usually a high-bandwidth, *synchronous* stream of data that must be processed and incorporated into the model (or ignored). However, the processes that draw on the model as a source of information—a display processor or path planner, for example—are driven by needs unrelated to the relentless flow of sensor data.

Functionally, the model can be cast as an intermediate representation that decouples high-bandwidth, real-time sensors from the more *asynchronous* processes that consume information. We want an "expert" sonar process, driven by the hardware-determined data flow, to be able to build its model independent of a display expert, say, that serves a human operator's changing information needs. Rather than ignore the data if they are not considered immediately useful, they should be merged with the model whenever available; the information may turn out to be useful in retrospect.

Our human sensory processes work similarly; often, we recall something seen or heard that seemed unimportant when it was perceived in "background" mode. In fact, we are not consciously aware of most sensory data; our attention is freed for more important foreground processing. However, I do not want to push this analogy too far; our own sensing and "modeling" is highly parallel, and cannot to be matched by synchronous, serial processing on a computer. My point is that, if machine sensory/model processing can also be formulated deterministically, in the same way as modeling storage requirements, then a dedicated sensory subsystem can free-run using the low-level model to buffer higher-level processes.

This functional partitioning can also facilitate the separation of an application's hardware base into independent processing modules. For example, measurements can be filtered by a digital signal processor, and state estimates streamed to a separate event processor. The event processor may consider independent sensor events and generate a local model of each event as a vector list to be

manipulated by a global modeler. The global model might reside in shared memory where it is accessible to other functions, or a dedicated feature extractor could service requests from higher-level processes over a low-bandwidth communications channel. Such partitioning has provided the basis for an application over a broadband network, in which sonar event vectors from an ROV operating at the waterfront were received by a real-time display processor in the laboratory, half a mile away. I also expect this kind of network configuration to become the norm in a shipboard setting, where many scientists and engineers will share the enormous amount of data coming from a suite of high-resolution sensors.

3.3.3 Mapping Feature Vectors Through Model Space

Before going on to specific modeling processes in the next sections, I discuss the data flow of Figure 3.1 in more detail here. My reasons for doing so are to: give a perspective on the larger modeling framework—the "big picture"—before elaborating on the pieces; and establish the terminology and notation to avoid later distractions.

First, I adopt a vector representation of dimensionality high enough to accommodate all sensor measurements and derived features over space and time. Using the notation outlined earlier, a feature vector has the general form:

$$\mathbf{v} = \{u_0, u_1, \ldots, u_m\}$$

Second, I consider a class of functions that map vectors through model space in a hierarchical manner. Such a vector function is denoted as:

$$\mathbf{f} = \{f_0, f_1, \ldots, f_n\}$$

and a vector mapping can be expressed as:

$$\mathbf{w} = \mathbf{f}(\mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_n)$$

or:

$$w_0 = f_0(v_0) = f_0(u_0, u_1, \ldots, u_m)$$
$$w_1 = f_1(v_1)$$
$$\vdots$$
$$w_n = f_n(v_n)$$

and so on. As before, the set notation is used as a reminder that we are dealing with a deterministic number of transformations, and all vectors and mapping processes can be enumerated. Otherwise, I

have only maintained a generality that preserves an option to use the many mathematical, heuristic, and engineering tools in a manner most suited to the application.

3.3.4 A Sonar Modeling Example

At this point, I leave the more general discussion and notation with a reminder that the representation and processing to be formulated is largely suited to different sensors and modeling applications. To clarify the issues, though, I give an example of building an acoustic model from sonar data. In the rest of this section I describe the feature vectors and mapping functions of a sonar model in a general manner, then give a detailed discussion of the modeling processes (mapping functions) in the next two sections.

I consider a pulsed sonar that returns a continuous stream of discrete, intensity measurements, and assume that position and attitude measurements are also taken at an appropriate data rate. I define the *measurement vector* as:

$$\boldsymbol{\rho}_m = \{t, \mathbf{x}, \dot{\mathbf{x}}, \boldsymbol{\alpha}, \dot{\boldsymbol{\alpha}}, \rho, \boldsymbol{\sigma}\}$$

where:
$$\mathbf{x} = \{x,y,z\}$$
$$\dot{\mathbf{x}} = \{\dot{x},\dot{y},\dot{z}\}$$
$$\boldsymbol{\alpha} = \{\alpha,\beta,\gamma\}$$
$$\dot{\boldsymbol{\alpha}} = \{\dot{\alpha},\dot{\beta},\dot{\gamma}\}$$
$$\boldsymbol{\sigma} = \{\sigma_t,\sigma_x,\sigma_{\dot{x}},\sigma_\alpha,\sigma_{\dot{\alpha}},\sigma_\rho\}$$

To avoid notational complexity, I do not use the subscript *m* for all the individual measurements; the meaning should be clear as indicated and the subscripts are assumed. As before, t is time; $\mathbf{x}$ is the sensor's position in 3-space, and $\dot{\mathbf{x}}$ is the first positional derivative, or translational velocity; $\boldsymbol{\alpha}$ is a unit vector defining the sensor look direction, and $\dot{\boldsymbol{\alpha}}$ is the look-angle velocity vector; ρ is the acoustic intensity measured by the sonar, the *primary sensor;* and a bold $\boldsymbol{\rho}$ denotes the entire measurement vector.

Here, all information is explicitly represented and passed to the state estimator in the measurement vector; the intensity is tagged with time, position, and attitude. I also express a measure of confidence, σ, associated with each parameter. Normally, this would not be included with the raw measurements, but more intelligent sensors could furnish such an error estimate as part of their function. σ_t models the uncertainty in any absolute measurement of time. In a general case where

information is contributed by multiple sensors using different time bases, perhaps from widely distributed locations, it may be significant.

A *state estimator* then receives this stream of measurement vectors, applies a *state mapping function*, $\mathbf{f}_e$, and generates a stream of *state estimates*. In this transformation, the outputs from nonideal sensors corrupted by noise are combined with consideration to a physical model of the sensors and a dynamic model of the system. For example, I apply a Kalman filter [*Gelb*, 1974], a recursive, stochastic estimator of the form:

$$\rho_e(t_+) = \mathbf{f}_e(\rho_m(t), \rho_e(t_-))$$

The plus and minus signs indicate that the new estimate, made just after a measurement is taken at time t, is based on the previous estimate and the new measurement.

Measurements also may be smoothed (a weighted combination of earlier and later measurements). Such smoothing introduces a delay but, if real-time constraints are not restrictive, smoothing generally produces better results. Various approaches to smoothing and estimation have been taken, deterministic and nondeterministic [*Gelb*, 1974]. The received sonar signal is usually filtered in hardware, but more filtering or processing might also be performed by the state estimator (for example, thresholding in software to estimate the ranges of strong sonar targets). For this example, the state estimate is:

$$\rho_e = \{t, \mathbf{x}, \alpha, \rho, \sigma\}$$

Here, the first derivatives are ignored since they are only modeled to improve estimates of position and heading. If Doppler effects were of interest, $\dot{\mathbf{x}}$ could be included.

My development in this section so far has followed traditional methods. However, the *event processor*, which I describe fully in the next section, reflects a new approach to mapping a state vector (a function of time) into an *event vector* (a function of time and space). For now, though, I merely denote the *event mapping function* as:

$$\rho_k = \mathbf{f}_k(\rho_e)$$

A simple form belies the potential for a complex transformation. Implicit in the event processor is an expert's knowledge of the physical basis for that class of events. The event vector is given as:

$$\rho_k = \{\mathbf{x}, \rho, \sigma_\rho\}$$

All uncertainty—spatial, rotational, detection—has been merged into one parameter, σ_ρ, that expresses a level of confidence in the estimate of ρ at each point in space and time. The subscript k denotes a model vector of the k^{th} event (time is implicitly associated with an event).

The event vector is a *model vector* of the same kind discussed in Section 3.3.1, since it has been mapped to a form that can be merged with a local or global model. In a sense, the event comprises a local model of the event space defined by the sonar's detection envelope. In Section 3.5 I describe stochastic backprojection, the process by which events are aggregated, but indicate the modeling transformation here as:

$$\mathbf{M}_k = \mathbf{f}_M(\rho_k, \mathbf{M}_{k-1})$$

where: $\mathbf{M}_k = \{\mathbf{x}, \rho_M, \sigma_M\}$

This is a recursive process of the general kind described earlier. I would like to strengthen this analogy by likening ρ_k to a "measurement vector" that is combined with the previous "state estimate," $\mathbf{M}_{k-1}$, to generate a new estimate with information derived from the event. In all essential aspects, the global model represents our current best estimate of the modeled parameters (state variables).

From the stochastic model, the *feature extractor* can derive a "deterministic" estimate of such environmental features as shape, surface normals, acoustic scattering properties and so on. I give an example of a *feature mapping function* in Section 3.6.1, and discuss other approaches in later chapters; for now, I consider it as a "kit" containing tools of the general form:

$$\mathbf{F} = \mathbf{f}_F(\mathbf{M})$$

This can be a computationally expensive step, asynchronously executed only as the application requires, or a more economical operator applied concurrently as the model is built.

Again, the deceptively simple representation may obscure profound, domain-specific knowledge imbedded in the function. However, the sensor-oriented, physical perspective on model building presented so far should be distinguished from the *evaluative* or *interpretive* feature-extraction processes that use the model. The latter may be categorized as operators acting on intrinsic properties of the model (image processing, pattern analysis) and those that bring to bear extrinsic knowledge (context, prior data), adding information in the process.

I use the term *feature vector* in a broad sense to subsume all the previous vectors, but more specifically to denote a derived type. For example, an estimate of the seafloor surface from ρ_M, an estimate of the surface scattering albedo, or some higher-level classification or display representation

based on the two is considered to be a feature (of the environment) that may be inferred from the model. In this example, I assume the model is used to estimate surface shape, and indicate the feature vector as:

$$\mathbf{F} = \{x\}$$

or: $$\mathbf{Z} = f(x,y)$$

where **Z** is a two-dimensional terrain map that describes the distribution of depth over the modeling region.

In summary, modeling vectors and mapping functions offer a consistent analytical framework for the modeling representations and processes. The feature vectors given in this sonar example are:

$$\rho_m = \{t,x,\dot{x},\alpha,\dot{\alpha},\rho,\sigma\}$$
$$\rho_e = \{t,x,\alpha,\rho,\sigma\}$$
$$\rho_k = \{x,\rho,\sigma_\rho\}$$
$$\mathbf{M} = \{x,\rho_M,\sigma_M\}$$
$$\mathbf{F} = \{x\}$$

From each level, the transformations move generally toward a vector space of lesser dimension, more germane to the application feature(s) of interest. In the next two sections, I use this same sonar example to elaborate on the key processing elements represented by the *modeling functions*, $\mathbf{f}_k$ and $\mathbf{f}_M$, the event processor and the global modeler.

3.4 MODELING A SONAR EVENT

The central idea to be addressed here is the distribution of information in a one-dimensional, time-varying signal over a three-dimensional volume. For an active sensor with one transmitter and receiver (sonar, laser, or radar, for example), energy is projected into a region of space and may be absorbed or reflected by one or more targets. Some portion of the signal may be returned to the receiver after being attenuated or distorted by the medium and corrupted by noise. Passive sensors depend on a source of energy external to the system, perhaps originating from the target itself (passive sonar, infrared). For both, however, the signal received at any moment is generally a summation of the contributions from multiple sources or reflectors.

Because the received waveform is represented as a one-dimensional function of time, though the sound energy has passed through a three-dimensional volume of space, there is an inherent ambiguity

in localizing any target causing a return. Another kind of uncertainty arises since we may not detect all targets, for example, if they reflect only weakly or direct energy away from the receiver. There is also inaccuracy in our estimate of where the sonar is located and of the direction in which it points. Acoustic noise, an inhomogeneous medium, and other statistical characteristics of the process further diminish the information in the received signal.

Suppose we wanted to build a model of the spatial distribution of acoustic scattering in some region. One thing such a model could tell us is the location of acoustic surfaces—the bottom, a sunken ship, and so on. To match the cellular decomposition of space, the sonar's signal must be mapped to a description that can be merged with the model. This will be a probabilistic spatial distribution reflecting the uncertainties just mentioned. Because of the high data rates in many sonars, the efficiency of such a process must be considered. By applying a series of constraints, the modeling computation for each event can be reduced to a practical level for cost-effective field systems.

To examine the physical situation, I start from a version of the Sonar Equation [*Horton*, 1959; *Urick*, 1975], a system design tool that expresses the relationships in an intuitive form. An energy balance lets us write:

$$RL = PL + DI_t - TL_t + S - TL_r + DI_r$$

where: RL = receiver level
PL = power level
DI = directivity index
TL = transmission loss
S = scattering strength

and subscripts "t" and "r" denote transmitted and received (reflected) paths. The terms are logarithmic, and noise and other uncertainty are ignored for the moment. The equation simply says that the level of the transmitted signal (or energy) detected at the receiver will be proportional to the power transmitted by an omnidirectional source, if you consider the directionality of the sonar transducers, the attenuation of the medium, and the scattering properties of any targets in the region.

Given a physical model of the sonar (power level and directional characteristics of the transmitter), an acoustic propagation model (sound velocity, absorption, and spreading losses) provides an estimate of how much energy will arrive at any point in space and time. If the scattering function of a target at that location were fully known, we could also predict the time history of the detected signal considering the receiving transducer's location and directional characteristics.

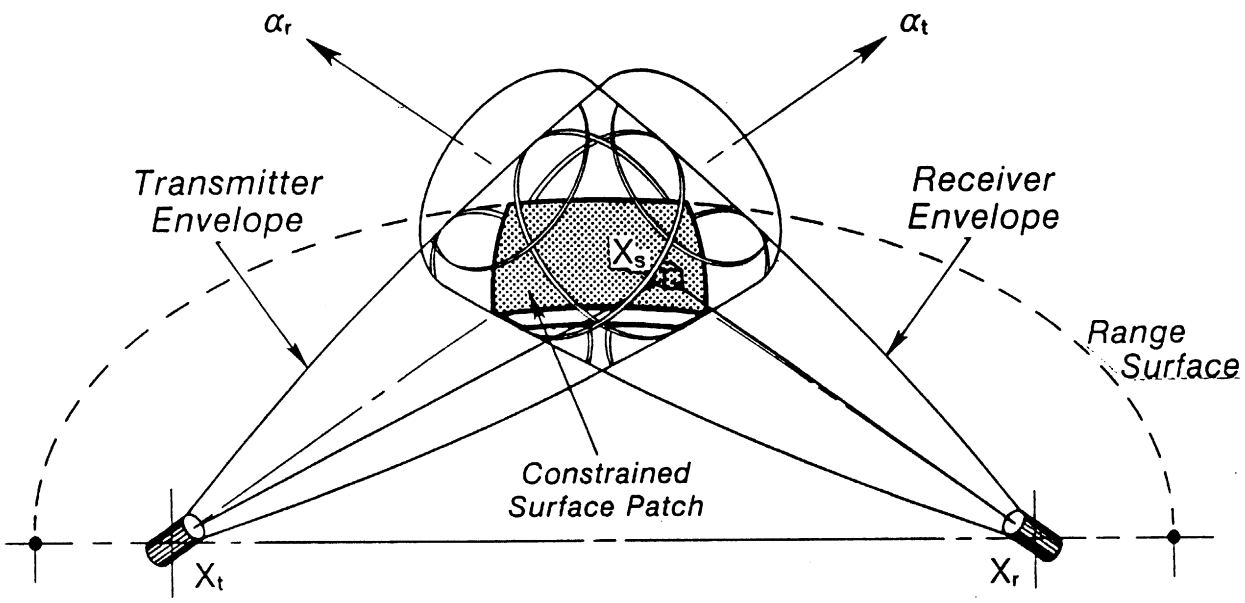

Figure 3.2: Detection envelopes for a sonar event.

Starting the event, a transmitter at x_t (refer to Figure 3.2) sends out a pulse of acoustic energy in the direction α_t ($\alpha = \{\alpha,\beta,\gamma\}$, the *direction cosines*). After a time, the attenuated pulse reaches a target at x_s, which absorbs some energy and scatters the rest. Still later, this scattered sound reaches x_r, where there is a receiver pointing in the direction α_r. This may be a different sensor, or the same transmitting and receiving transducer, which has moved since the pulse was sent.

I assume that the medium is uniform, and that the only losses come from absorption by the water and from spherical spreading as the sound moves away from the transmitter. This is a good assumption over short ranges, but becomes less valid with increasing distance, as refraction and multipath effects become important. In this homogeneous regime, though, we can talk about time and distance interchangeably, since they are related by a constant speed of sound in water. Transmission losses become a simple function of time (or range), and the transformation from a time signal to a spatial distribution is more direct. Now we can parameterize the Sonar Equation according to our model of the physical system, and rearrange terms to give:

$$S(\tau,x_s) = RL(\tau,x_r) - [PL(\tau,x_t) + DI_t(\tau,\alpha_t) + DI_r(\tau,\alpha_r) - TL(\tau)]$$

where τ is elapsed time from the start of an event. The *DI* terms now represent the sonar beam patterns, rotated to their look directions. In this treatment, I have ignored any directional dependence of the scattering term—in reality, a complex function of transmitter and receiver geometry. In later applications, I show this to be a reasonable assumption for many purposes; in other applications—sidescan sonar, in particular—the assumption leads to ambiguous results. I discuss scattering and reflection more fully in Chapter 6 but, for now, I model it as a simple target strength, with uniform scattering in all directions.

The important thing is that the terms in square brackets capture all our knowledge of the sonar system itself and of the medium in which it operates. I refer to this as the *detectivity*, since it is a measure of how strongly a target must reflect for it to be detected in the received signal. As long as the position and direction of the sensors are known, the detectivity of any voxel can be calculated directly; then the received signal can be used to infer something about the scattering distribution at any point and time.

For a given sonar event, the time origin is set to coincide with the transmission of an outgoing pulse of short duration. After the signal is transmitted, the source position and attitude become irrelevant and can be disregarded. At some later time, say t_r, the attenuated pulse is detected at the receiver. Given the position of the transmitter at the time of the outgoing pulse and of the receiver at time t_r, but without regard to the directionality of the transducers, a single target causing the return can lie anywhere on an elliptical surface of revolution. In a real situation, however, the signal may represent the contributions from multiple targets. Under such circumstances, our knowledge of the situation is limited to the summation of scattering strengths distributed over the surface. However, there is other information that can be brought to bear—our model of the transducer beam patterns.

Considering only the transmitter, an envelope may be constructed beyond which a signal is attenuated to such a point that it cannot be distinguished from noise. In other words, suppose the position and attitude of the transmitter is fixed and the receiver is moved throughout space so that it is always pointing directly at the source. In some locations the transmitted pulse will be detected. In others it will not. The locus of all points at which the pulse is just detected will form a closed surface beyond which we cannot derive any information from an event.

The boundaries are a function of power level, transmitter beam pattern, transmission loss, noise level, and receiver characteristics, excluding the receiving transducer's directivity. By fixing the receiving transducer and moving the transmitter through space, a receiver envelope is defined analogously. Outside these surfaces, the possibility of detecting any passive reflector is nil. The intersection of the source and receiver envelopes bounds a region from which a scattered return must have come; this is the *detection envelope* for the event. Combining this information with the detection

time, any targets contributing to the received signal are constrained to lie on a surface patch of the ellipsoid.

3.5 STOCHASTIC BACKPROJECTION

Considering the constrained surface patch, the signal received at t_r is, in a sense, the three-dimensional equivalent of the ray sum defined in Section 2.5.1 (see also *Norton and Linzer* [1979a, b]). In the absence of noise, it is the integral of scattering strength over the surface, modulated by power level, acoustic attenuation, and beam patterns:

$$g(t_r,\mathbf{x}) = \int_S p(\mathbf{x}_t)\upsilon(t_r)b(t_r,\mathbf{x},\boldsymbol{\alpha})s(t_r,\mathbf{x})dS$$

where: S = surface patch
p = transmitted power
υ = attenuation
b = composite beam pattern
s = scattering strength

This can be simplified slightly by noting that, for a short pulse, the power and location of the transmitter are independent of time after the pulse is sent. Also, the attenuation is constant over the surface since I assume a homogeneous medium. Finally, I assume a static environment in which the point targets do not change position, and drop the time notation since we are only considering one range. Then:

$$g_r(\mathbf{x}) = K_r\int_S b_r(\mathbf{x},\boldsymbol{\alpha})s(\mathbf{x})dS$$

where the power and attenuation terms have been combined in the constant K, and the subscript r denotes a particular range surface.

If we consider the entire signal for that event, it corresponds to a sequence of surfaces, and the time signal with a sequence of surface integrals (or summations, in the discrete case). This is a three-dimensional projection over ellipsoidal surfaces, shaded by the beam patterns, or:

$$g(\mathbf{x}'') = K\int_{S''} b(\mathbf{x}'',\boldsymbol{\alpha}'')s(\mathbf{x}'')dS''$$

where the double-prime notation indicates that the projections are taken over arbitrary translations and rotations.

With an ensemble of N such projections, not necessarily of equal spacing, we can formulate a reconstruction problem as in Section 2.5.1: given the projection data $g_k(\mathbf{x}'')$, k = 0, . . ., N-1, construct the original scattering distribution s(**x**). Important issues here are: (1) ellipsoidal surface projections; (2) attenuation and shadowing; (3) beam-pattern effects; (4) reconstruction method; (5) the role of uncertainty.

First, the curved range surfaces are inconsistent with most previous approaches to the reconstruction problem. Other researchers [*Das and Boerner*, 1978; *Denton et al.*, 1978; *Rockmore et al.*, 1979; *Rockmore*, 1981] circumvent the problem by approximating each surface as a plane. *Norton and Linzer* [1979b] begin with a spherical geometry but make simplifying assumptions leading to equivalent results. This leads to an analytical tractability but introduces some error. In the numerical approach taken here, such an approximation is unnecessary. The curved range surfaces are modeled exactly, except for some quantization error in the discrete cellular model.

Second, I assume implicitly that the distribution of targets is sparse enough that scattering at close ranges does not prevent the signal from reaching targets further away (the *weak scattering* assumption of *Norton and Linzer* [1979a, b]). This assumption is not always valid. For example, the signal received from a down-looking sonar would not be meaningful beyond the bottom return. This is accommodated easily (if the bottom can be detected) by ignoring the signal after the corresponding time. *Das and Boerner* [1978], and *Rockmore et al.* [1979] also discuss this *shadow boundary*. In a numerical approach, losses can be compensated at any range simply by inverting the assumed attenuation function, the basis for TVG (analytically, it is more difficult, but *Budinger and Gullberg* [1974] present several approaches).

Third, the beam-pattern also introduces a weighting of the surface integral. As such, a simple inversion, like that applied to the attenuation, does not work here. At the boundary of a detection envelope, the transducer's directional sensitivity approaches zero and an inverse compensation approaches infinity. Obviously, this is the wrong way to look at it. *Rockmore* [1979] also discusses the problem but offers no definitive solution. *Norton and Linzer* [1979b] assume omnidirectional sensitivity so beam-pattern effects are not considered. I take a heuristic approach described later in this section.

As discussed earlier, traditional reconstruction techniques rely on a fixed geometry and regular scanning pattern. An exception to this is the backprojection (summation) method, though it is usually formulated in that context. This is the approach I take to incremental modeling and denote it as:

$$\mathbf{M}_N = \sum_{k=0}^{N-1} f(\mathbf{x}'')g_k(\mathbf{x}'')/(N \cdot K)$$

where: $\mathbf{M} = \{x, \rho, \sigma_\rho\}$

and: $f(x'')$ is a beam pattern compensation function

In other words, the global model, $\mathbf{M}_N$, is formed by the backprojection and summation of N event vectors. With no uncertainty, the general algorithm would go as follows for each event:

1. Isolate the volume defined by the intersection of the detection envelopes and form a list of voxels ordered by discrete range surfaces.

2. For each range surface, compute the value of each voxel as the product of the inverse attenuation and the received signal for that range, and apply a beam-pattern compensation factor.

3. Backproject the list into the global model and sum the new values.

4. Normalize each accumulated value by the total number of operations on that voxel.

Other than formulating a more general numerical approach to the geometry, the main difference here is that I have adapted the backprojection algorithm to satisfy the incremental modeling criterion established earlier. A consequence of using this technique is the introduction of reconstruction artifacts, as discussed in Section 2.5.1. However, as shown in the next chapter, the method produces results that compare favorably with standard techniques.

Going now to the question of uncertainty, I first equate the beam-pattern ambiguity with uncertainty introduced by attitudinal (or pointing) error. Strictly speaking, this is incorrect. In effect, though, both sources of uncertainty limit our ability to localize a point target that contributes to a received signal. For a single event, all we can say is that the target lies somewhere on a range surface.

Consider a very narrow-beam sensor—a laser, for example. With no attitudinal error, a point target detected at some range can be localized unambiguously (or, at least, to within very tight bounds). If we allow some error in sensor attitude, then the highest probability of the target's position lies on the estimated axis of the beam, and the probability falls off away from the axis. At a given range, with no ranging uncertainty, the localization probability function forms a range surface. That function is equivalent to the convolution of the pointing-error probability density function (pdf) with the sensor beam pattern, which approaches a delta function for a laser.

For a wide-beam sensor, the target causing a return also is most likely to lie near the axis of the beam. My argument goes as follows: consider a sensor fixed in space, and a target of given scattering strength moving along a range surface. At some angular distance away from the axis, the received signal will fall below a threshold and the target will not be detected. That threshold may be set

arbitrarily in the receiver, or it may be imposed as a floor level determined by noise. For another target of lesser strength, the angular cutoff point will lie closer to the axis. If we assume a spatial distribution of targets with a distribution of scattering strengths, then for any range surface, more targets will be detected near the center. An assumption of fixed targets, and a distribution of events at different look angles, produces the same conclusion by duality.

I contend this is a true proposition though it is derived purely by logical, physical arguments. However, to establish a more concrete relationship between the received signal and any scatterers in the region requires two missing pieces of prior statistical knowledge: (1) the spatial probability distribution of point targets and (2) the probability distribution of scattering strengths among those targets. With such information in hand, an application of Bayesian techniques might be used to generate a localization pdf for each event; but for unexplored terrain, such information will not be available. Even if long-term statistics had been compiled, the probabilities would be highly variable and dependent on the direction in which the sensor was looking—at the bottom, at the surface, through the water column.

To sidestep this problem, I use the normalized beam pattern itself for the localization function—the boundary conditions are correct and it has the right general shape. With this assumption, an *angular localization function* (ALF) is formed by convolving the beam pattern with the pointing-error pdf. Other assumed error pdf's are incorporated in the same manner. This includes any uncertainty in position, range resolution, or timing. Error caused by ray bending or other distortion in an inhomogeneous medium might be accommodated also, though the computational complexity could rise significantly if these were modeled accurately.

Alternatively stated, the error pdf's are convolved with the detectivity. A physical interpretation is that any uncertainty in position or attitude, for example, will tend to smear each surface distribution through space and reduce our confidence in an estimate at any point. The outcome is that the surface now corresponds to a thin volume of space. The pdf over that space, which I call the *composite localization function* (CLF), is related to the likelihood that a point target at any voxel within the space contributed to the received signal at t_r.

A final modification to the algorithm convolves the CLF with the backprojected signal, and accumulates the CLF separately for normalization. In such a manner, the beam pattern ambiguity and other sources of error are represented; and as error approaches zero and the beam pattern narrows, the CLF approaches a delta function and the result becomes an incremental backprojection and summation. The disadvantage of such an approach is that the more explicit error characterizations are combined in a single measure of (un)certainty. However, this can be interpreted qualitatively as a measure of the information content for each voxel. The advantage is computational efficiency. For high-resolution

three-dimensional models, such a tradeoff—real-time performance against fidelity—may be acceptable. Later chapters show the outcome of these assumptions. To summarize:

$$h_a(x'') = b(x'',\alpha'')*h_\alpha(\alpha'')$$

$$h_c(x'') = h_a(x'')*h_x(x'')$$

$$H_N(x) = \sum_{k=0}^{N-1} h_k(x'')$$

$$S_N(x) = \sum_{k=0}^{N-1} h_{c_k}(x'')*g_k(x'')/(K \cdot H_N(x))$$

$$M_N(x) = \{S_N, H_N\}$$

where:
- b = composite beam pattern
- h_α = attitudinal error probability density function
- h_x = positional error probability density function
- h_a = attitudinal localization function
- h_c = composite localization function
- H_N = global certainty parameter
- S_N = global modeling parameter (acoustic signal)
- M_N = global model

We can see there are two competing forces at work here. The cumulative constraints imposed by an unmodified backprojection method tend to sharpen the model, and compensate for the degradation caused by the range-surface projections. At the same time, the model is filtered, or blurred, by purposeful convolutions with the error pdf's. One might argue that the convolutions should be omitted and a better model would result. Certainly a crisper model would be produced, but the quality would be illusory; the convolutions explicitly represent our best estimate of uncertainty. For an autonomous system, it is just as important to model the unknown as the known, and avoid unjustified assertions about the world.

The advantage of filtering at this low level is that uncertainty can be represented more accurately according to the cause, amount, and direction. Beam-pattern ambiguity and pointing error are distributed over a curved surface, ranging error along the axis—each is in proportion to its own degree of uncertainty, and in the right spatial frame as it is incorporated in the model. The alternative of filtering the finished model according to some average error—convolution with a symmetric Gaussian kernel, for example—underestimates uncertainty in some directions and overestimates in others. The argument holds also for the introduction of backprojection constraints in proper global coordinates, as the model is built. To paraphrase *Rockmore* [1979], we should process *then* threshold, or extract deterministic conclusions from the model.

The example used in this development is that of an active sonar, the bidynamic case where transmitter and receiver are separate and in motion (after discussing the constrained surface patch at t_r, I dropped the time notation as a convenience). Yet, because of its generality—a numerical formulation in the spatial (and temporal) domain—the method can be extended to passive cases (for a related discussion, see *Rockmore*, 1979) and other sensing modalities. The physical model of an underwater scanning laser, for example, could be formulated and substituted for the sonar model to produce similar results. A tactile model, built with a robotic arm, is also feasible. My intent has been to develop a general approach to incremental modeling largely independent of range, resolution, and sensor type.

3.6 GLOBAL MODELING

In this section, I give two hypothetical examples of sonar modeling that provide more elaboration of the algorithms and notation applied to specific cases. In the first type, which I call a *binary model*, a profiling sonar returns a thresholded range to target. The model uses a normalized scattering distribution representing the probability that a voxel is occupied, without regard to absolute target strength. In the end, a binary decision is made on whether the voxel is empty or full. The second example is that of a two-dimensional *continuous model*, the intensity map produced by a sidescan sonar. In both examples, and in those of later chapters, I assume that the same transducer acts as transmitter and receiver. This is usual in most applications, and is less burdensome computationally for modeling simulation.

3.6.1 Binary Model

As an example, I consider a free-swimming vehicle with a down-looking, high-frequency (say, 500 KHz) profiling sonar that models a static terrain (refer to Figure 3.3). The sonar uses a plane circular transducer as transmitter and receiver, returns a thresholded range to target, and has a maximum range $R_{max} \gg 50$ m. For a reasonably stable platform and short return times (say, range $<$ 50 m), sensor motion will be small during an event, and time variations can be neglected. I assume a homogeneous medium, stationary white Gaussian noise in all measurements, and that a Kalman filter generates the state estimates. From the Sonar Equation:

$$S(\mathbf{x}) = RL(\mathbf{x}) - [0 + 2DI(\alpha) - 2TL(r)]$$

Since the actual received signal is not available, the model is cast in a "normalized" form. Following earlier convention:

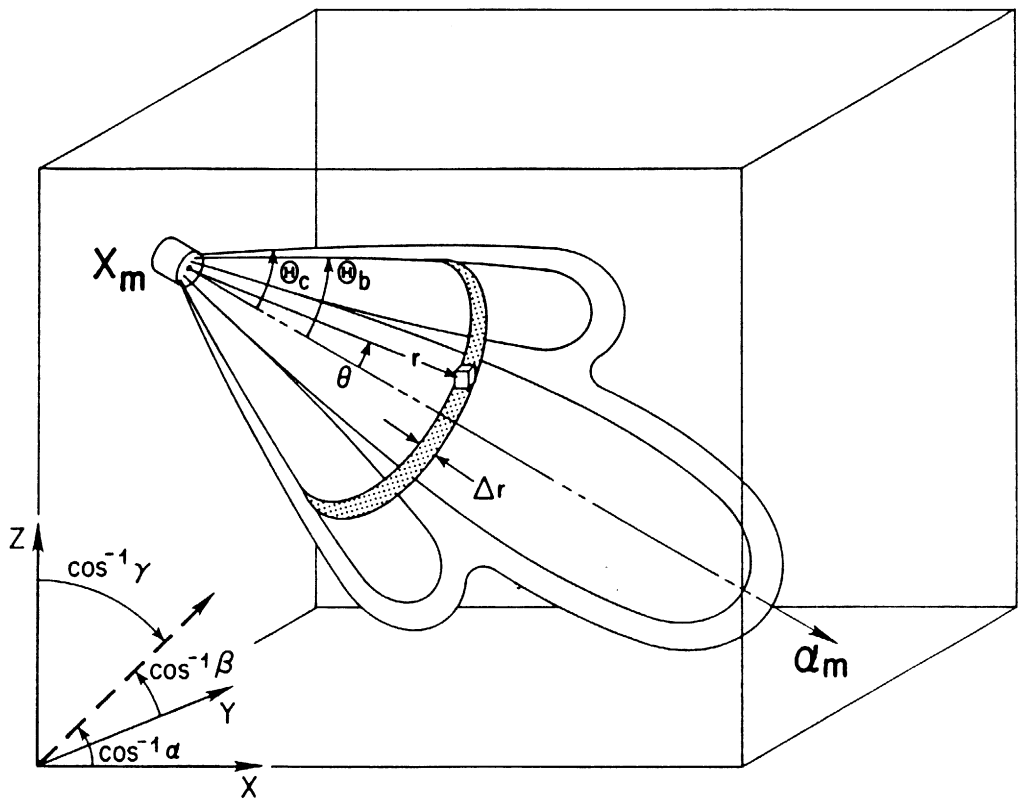

Figure 3.3: Geometry for three-dimensional modeling.

$$\rho_m = \{t, x_m, \dot{x}_m, \alpha_m, \dot{\alpha}_m, R\}$$

$$\rho_e = \{t, x_e, \alpha_e, R, \sigma_K\}$$

$$\rho_k = \{x, \rho, \sigma_\rho\}, \qquad r \leq R + f(\sigma_r), \quad \theta \leq \Theta_c$$

$$S = \{x, S, \sigma_S\}$$

$$Z = \{x\}$$

where: R = range return

$r = |x - x_e|$

$\theta = \cos^{-1}(\alpha_e \cdot (x - x_e)/r)$

Θ_b = detection-envelope half-angle

Θ_c = composite-localization-envelope half-angle

As before, ρ_m and ρ_e are the measurement and state vectors. The first derivatives of position and attitude are only measured to refine the state estimates for x_e and α_e. As indicated in the figure, α

represents the direction cosines. I use the subscripts *m* and *e* to distinguish the position and look direction of the sensor itself. The unsubscripted terms refer to parameters of each voxel, which vary as the calculations are performed throughout the localization envelope. Accordingly, **x** gives the absolute coordinates of a voxel; *r* is the range to a voxel from the estimated sensor location; $\boldsymbol{\alpha}$ is a unit vector from the estimated sensor location to a voxel; and Θ is the angle between the sensor look direction and the unit vector to a voxel.

For a plane circular transducer, the beam pattern can be modeled as a first-order Bessel function parameterized by transducer diameter and acoustic wavelength [*Urick*, 1975]. The detection-envelope half-angle, Θ_b, is taken at the second null so the first sidelobe is modeled. The composite-localization-envelope half-angle, Θ_c, is defined as the angular extent of the composite localization function, after Θ_b is blurred by the error pdf's. σ_K comprises the main diagonal of the covariance matrix for a discrete steady-state Kalman filter. **Z** is a surface depth map, a feature extracted from the global scattering model, **S**. The mapping functions are:

$$\mathbf{f_e} = \{h_K(\mathbf{x}_m,\dot{\mathbf{x}}_m), h_K(\boldsymbol{\alpha}_m,\dot{\boldsymbol{\alpha}}_m), g_R(r), \sigma_K\}$$

$$g_R = \begin{cases} 0, & r < R \\ 1, & r = R \\ \text{undefined}, & r > R \end{cases}$$

h_K = Kalman filter

$$\mathbf{f_k} = \{\mathbf{f_x}(r,\boldsymbol{\alpha}), g_c(r,\Theta), h_c(r,\Theta)\}$$

$$\mathbf{f_x} = \mathbf{x_e} + r\boldsymbol{\alpha}$$

$$g_c = h_c(r,\Theta) * g_\rho(r)$$

$$h_c = G(\sigma_r) * G(\sigma_\Theta) * b(\Theta)$$

$$g_\rho = \begin{cases} 1 - e^{-\nu r}, & r < R \\ e^{-\nu r}, & r = R \\ 0.5, & r > R \end{cases}$$

$G(\sigma)$ = normalized Gaussian distribution

σ_R = range variance

σ_r = composite range variance := $\sigma_R + |\sigma_x|$

σ_Θ = composite angular variance := $\sigma_\alpha + |\sigma_x|/R_{avg}$

R_{avg} = average sonar range := $R_{max}/\sqrt{2}$

ν = normalized signal parameter := $e^{-\nu R_{max}} = 0.5$

$$f_M = \{x, f_S(\rho_k, S_{k-1}), f_H(\sigma_\rho, \sigma_{S_{k-1}})\}$$

$$f_H = \sigma_\rho + \sigma_{S_{k-1}}$$

$$f_S = (\rho_k \sigma_\rho + S_{k-1}) / \sigma_{S_k}$$

$$f_F = \{h_f(z, S)\}$$

$$h_f = \sum_{z=0}^{\mathcal{Z}-1} zS(z) \Big/ \sum_{z=0}^{\mathcal{Z}-1} S(z), \qquad S > S_T$$

S_T = scattering threshold $\cong 0.7$

$\mathcal{Z}$ = number of voxels along depth axis

In practice, we want to avoid the computational expense that would be incurred if all convolutions were performed for each event. There is another assumption that reduces the computational cost to a practical level—for stationary noise the error pdf's are constant. And because the detectivity is also fixed for a specified sonar system and medium, some convolutions can be precalculated numerically. By taking advantage of axial symmetry, the real-time computational load can be lightened even more.

For an axially-symmetric sonar beam, h_a is precalculated at the beginning of a computer program, and accessed from a lookup table indexed by the angular offset, θ, of a voxel from the beam axis. The translational convolutions are resolved into axial (σ_r) and radial (σ_θ) components at some average range. If the radial convolution is incorporated in the lookup table, we are left with a one-dimensional convolution along the sonar axis. This is an inexpensive operation in a three-dimensional world, and can be carried out on the time signal for each event.

Note that the modeling process is almost entirely deterministic with these assumptions. Except for the one convolution, the modeling has been reduced to an ensemble of point processes. Neglecting the one-dimensional convolution, the processing load is a function of the number of voxels encompassed by each event. The actual processing load for each event depends on the returned range, but the computational expense is *only* $O(R^3)$ (the requirements could be *much* worse for a three-dimensional implementation). In fact, this still lets us place an upper bound on computations, since the maximum range is limited.

The normalized signal, g_ρ, represents a heuristic extrapolation of the received signal. The value ranges between 0 and 0.5 at ranges less than R, and between 0.5 and 1.0 at R. Here the normalized signal is interpreted as the probability that a given region of space contains a target. For ranges less than R, the signal is below some threshold (normalized to 0.5), and the corresponding volume of space

is probably unoccupied by any significant scatterers. For a value of 0.5, uncertainty is greatest and the odds are even that a voxel is empty.

The normalized signal parameter is defined with respect to the maximum range—the distance at which the signal is masked by ambient noise and no information can be extracted. At shorter ranges, the signal-to-noise (S/N) ratio is greater and the probability of detection increases. The approach is similar to that taken by *Moravec and Elfes* [1985], except that a S/N-dependent uncertainty is explicitly represented.

The depth map, $Z=f(x,y)$, is a single-valued function over a two-dimensional grid. The depth is estimated as the first moment of S on a vertical column-by-column basis, for values of S greater than an arbitrary threshold (see Section 4.2). With a downward-looking sonar—largely incapable of detecting overhangs or vertical scarps—this is a reasonable approach. For fully three-dimensional geometries, a more sophisticated technique would be needed to extract surface features from the probability distribution. I discuss this further in the next chapter, and use simulation results to clarify the process.

3.6.2 Continuous Model

In this next example, I look at an uncalibrated near-bottom-towed sidescan sonar (refer to Figure 3.4), which applies an "unknown" time-varying gain (TVG), and returns a digitized signal. The system operates in the medium-frequency band (say, 20 KHz) with a range of several kilometers. In this regime, ping cycles are on the order of 10 s and platform motions can degrade the received signal. In particular, I assume pitch and roll stability, but excessive yaw, and formulate a corrective gain. A two-dimensional model is used, and the planar-bottom assumption is applied for slant-range correction. As before, I assume a homogeneous medium, stationary white Gaussian noise, and a Kalman filter for state estimation. From the Sonar Equation:

$$S(\tau,x) = RL(\tau,x) - [0 + 2DI(\tau,\alpha) - 2TL(r)]$$

Like the previous example, a calibrated signal is unavailable. However, with several assumptions, the sonar output can be cast in a normalized form for modeling. The modeling vectors are:

$$\delta_m = \{\tau, x_m, \dot{x}_m, \alpha_m, \dot{\alpha}_m, \delta\}$$
$$\delta_e = \{\tau, x_e, \alpha_e, \delta, \sigma_K\}$$
$$\delta_k = \{x, \delta, \sigma_\delta\}, \qquad r \leq R_{max} + f(\sigma_R),\ \theta \leq \theta_c$$
$$S = \{x, S, \sigma_S\}$$
$$I = \{S\}$$

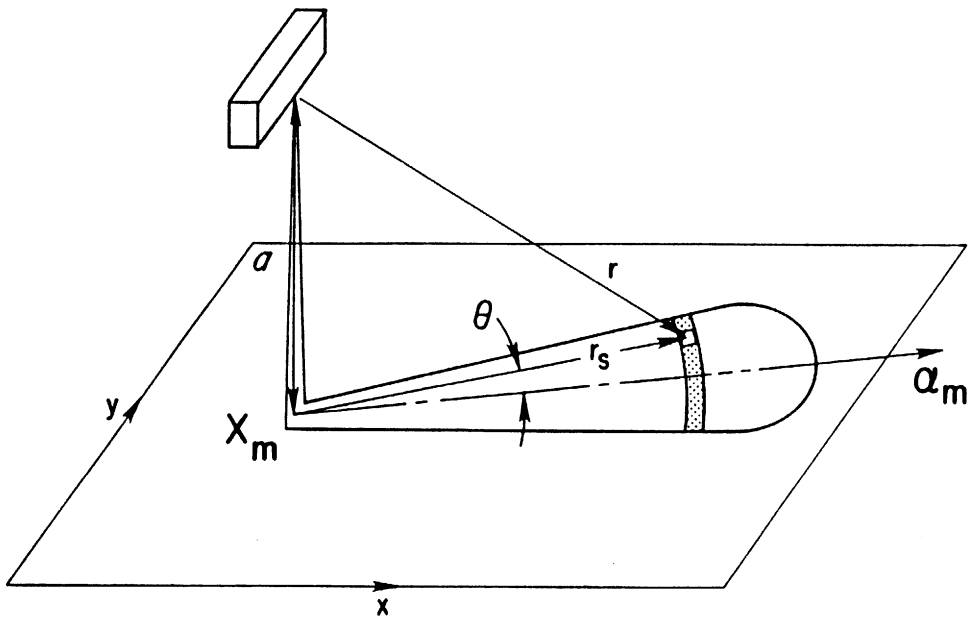

Figure 3.4: Geometry for two-dimensional modeling.

where: $\mathbf{x} = \{x,y\}$

$\alpha = \{\alpha,\beta\}$

$r = c\tau/2$

c = local speed of sound

δ = demodulated sonar signal

For a rectangular source, the along-track beam pattern can be modeled as a sinc function (sin(x)/x) parameterized by transducer length and acoustic wavelength [*Urick*, 1975]. As before, the detection-envelope half-angle is taken at the second null, and σ_K comes from the Kalman-filter covariance matrix. The feature vector, **I**, directly corresponds to the model, an uncorrected intensity map. The mapping functions are:

$$\mathbf{f_e} = \{h_K(x_m,\dot{x}_m), h_K(\alpha_m,\dot{\alpha}_m), \delta, \sigma_K\}$$

$$\mathbf{f_k} = \{f_x(r,\alpha), g_c(r,\theta), h_c(r,\theta)\}$$

$$f_x = x_t + \alpha[r^2-a^2]^{-\frac{1}{2}}$$

$$g_c = h_c(r,\Theta)*g_\delta(r)$$

$$h_c = G(\sigma_R)*G(\sigma_\alpha)*b(\Theta)$$

$$g_\delta = g_N(d\Omega)h_N(j,k)\delta_{jk}$$

$$h_{N_{jk}} = \eta_0/\eta_{jk}, \qquad 0 \le j\Delta r \le R_{max}$$

$$\eta_{jk} = \kappa\delta_{jk}+(1-\kappa)\delta_{j,k-1}, \qquad k\neq 0, \quad 0<\kappa<1$$

$$\eta_{j0} = \delta_{j0}$$

$$g_N = \frac{\Psi(0)\Theta_b}{\Psi(d\Omega)(\Theta_b-d\Omega)}, \qquad \Psi(d\Omega)\neq 0$$

$$\Psi(\Theta) := b(\Theta)*b(\Theta)$$

$$d\Omega = \tan^{-1}\left[\frac{\beta_t-\beta}{\alpha_t-\alpha}\right]$$

Δr = discrete range increment

j = range index

x_t = position at transmit time

α_t = heading at transmit time

a = altitude over bottom

η_0 = normalization constant

κ = filter constant

$$f_M = \{x, f_S(\delta_k, S_{k-1}), f_H(\sigma_\delta, \sigma_{S_{k-1}})\}$$

$$f_H = \sigma_\delta + \sigma_{S_{k-1}}$$

$$f_S = (\delta_k\sigma_\delta + S_{k-1})/\sigma_{S_k}$$

$$f_F = \{S\}$$

With a few exceptions, the development and notation are the same as in the previous example. The slant-range correction uses the standard assumption that the seafloor is generally planar, and at an average distance below the sonar equal to the measured altitude (derived from a separate sonar altimeter or bottom-detect time from the sidescan). This is not always a good assumption (discussed further in Chapter 6), but it allows some correction for the imaging geometry. I also drop the translational-error convolutions under the assumption that the relative positioning between neighboring ping locations is good. For towed imaging purposes, this relaxation is appropriate to the application.

Before mapping the sonar signal into model space, two corrections are applied. The function h_N is an adaptive normalization that helps compensate for attenuation and average grazing angle for each range bin. A static TVG, built into the system, is based on average conditions—towing altitude, acoustic attenuation, bottom scattering strength—and will be less than optimal for specific conditions. The "time" constant, κ, controls the rate at which a simple recursive filter estimates average signal strength at each range. This is inverted and multiplied by η_0, chosen to maintain the normalized signal in a convenient range for display or digital recording.

The second correction comes from a simplification of the bidynamic formulation of Section 3.4. With one transmit/receive transducer and no sensor motion, the received signal is weighted by the product of the beam-pattern sensitivities integrated over a range surface—for a finite pulse width, this defines a surface area of the bottom. For a short pulse length, though, the same bottom area (for a given range surface) is ensonified regardless of the transducer's motion after the pulse is transmitted. Assuming a small angular displacement, the received signal is attenuated because (1) the intersection of detection envelopes (bottom scattering area) is smaller, and (2) the integral of the beam-pattern product is reduced.

The reduction in scattering area is linearly proportional to the angular displacement. The attenuation caused by the beam patterns is proportional to the ratio of the beam-pattern integral at the displaced angle to that at a zero displacement. Considering the integrals at all possible displacements, this is simply the convolution of the beam pattern with itself. Then the correction function g_N can be precalculated numerically and indexed by the relative yaw, $d\Omega$, and applied to the received signal. An implicit assumption is that the sensor translation is small with respect to a range arc, which is reasonable for such a long-range sonar. The signal is then backprojected onto the plane into the area defined by the transmit detection envelope.

3.7 SUMMARY OF IMPORTANT POINTS

In this chapter I have presented a broad concept of modeling underwater, provided an analytical framework for considering different sensor models, and described two hypothetical examples to clarify the ideas. The basic philosophy that guided this development is outlined in Sections 1.5 and 3.1, and I do not reiterate it here. However, to summarize a few important ideas emphasized in this chapter:

- Stochastic backprojection comprises two competing forces: sharpening introduced by accumulating constraints, and blurring from the explicit representation of uncertainty.

- Cumulative constraints come from the redundant information available from modern high-bandwidth sensors, and can be combined with an incremental summation algorithm.

- The explicit representation of uncertainty is important to avoid unjustified assertions about the world, particularly for an intelligent autonomous system.

- By combining physical and sensory data at a low level of representation, high-level processes can avoid information overload: process then threshold.

- A low-level model can serve as an intermediate representation that decouples high-bandwidth sensory processors from more asynchronous information extractors and consumers.

- Tradeoffs of fidelity against real-time performance are inevitable and acceptable for high-resolution three-dimensional modeling.

- Model size and processing load can be bounded, and processing resource requirements can be forecast. The actual processing load is on the order of the range cubed.

- By partitioning algorithms into preprocessing and real-time components, computational performance can be enhanced to suit cost-effective field systems.

Chapter 4

COMPUTER SIMULATIONS

Several computer simulations were developed early in this research to scrutinize first assumptions and to focus on important issues. Simple representations of active sensors with random noise components were used to generate measurement vectors for processing by an event modeler. Most simulations are of sonars, though the effect of a narrow-beam sensor (laser) was examined in a few cases. Figure 4.1 shows a data-flow diagram for the modeling system. Except for the positioning component inside the dashed lines, the same process was also tested on field data (the results are given in Chapters 5 and 6).

The three-dimensional cone in Figure 4.2 is a simple model vector that shows the probability distribution for a binary sonar event. The higher probabilities denote a greater scattering strength, corresponding to a strong sonar return at that range. Because of the uncertainty discussed earlier, the range surface is smeared over a volume. The lower probabilities represent a region of space through which the sonar pulse has passed without detecting any targets. The scattering probability decreases toward the axis of the beam and closer to the sensor because the sonar's signal-to-noise ratio is higher there.

The next sections show the results of several different kinds of simulations. First, simple two-dimensional geometries are used to demonstrate basic properties of the backprojection approach. In Section 4.2, several three-dimensional binary simulations show the effects of modeling natural terrains and regular geometric features over different ranges of uncertainty. The next section demonstrates terrain-relative positioning using a stochastic model. The chapter concludes with a summary of simulation results and important points.

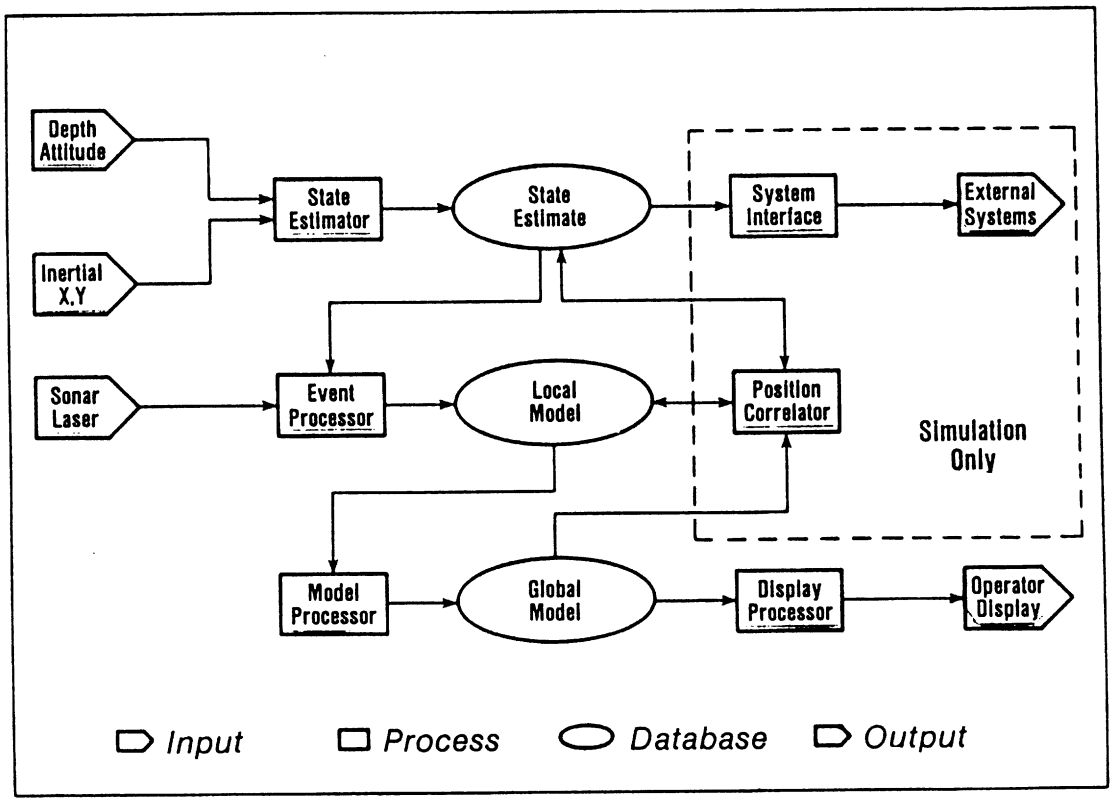

Figure 4.1: Data-flow diagram.

4.1 SENSOR MODELS

Figure 4.3 shows the geometry for a generic reconstruction simulation in a two-dimensional setting. In this scenario, there is a single point target at the center of a circle with a radius of 600 units (the units are nondimensional but were mapped to frame-buffer pixels for convenience). An active sensor moves around the circle and illuminates the central region from different look angles. Here, I assume perfect compensation for attenuation, no position or attitude error, and model the CLF as separable functions of axial and angular uncertainty so that:

$$h_c = h_\theta(\theta) * h_r(r)$$

$$h_\theta = \frac{\cos(\theta)-\cos(\theta_c)}{1-\cos(\theta_c)}, \qquad \theta \le \theta_c$$

$$h_r = |\sigma_r - r| / \sigma_r, \qquad r \le \sigma_r$$

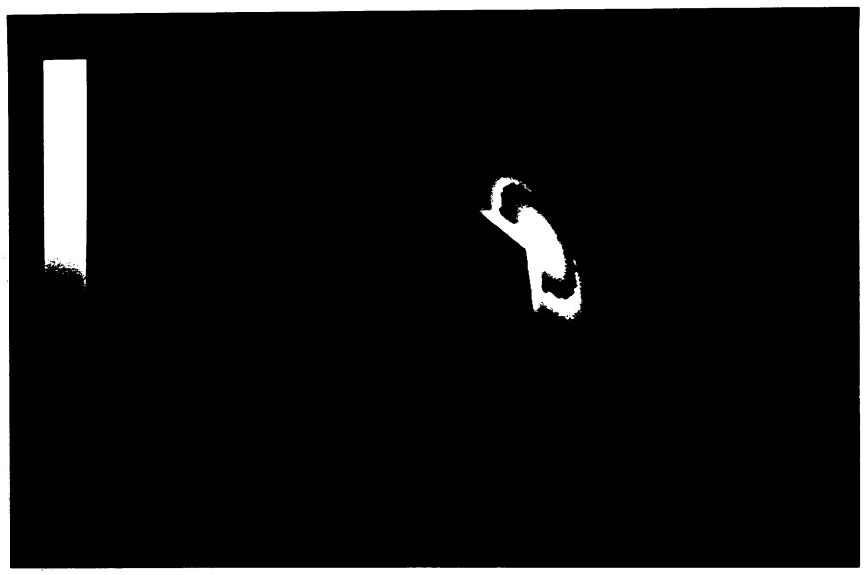

Figure 4.2: Probability distribution for a sonar event.

$\Theta_c = 12°$

$\sigma_R = 2.5$ units

The two functions are chosen purely for computational convenience. The formulation for h_Θ produces a cosine-like "beam pattern" 24° wide. The ideal, normalized signal received by the sonar and the composite modeling function are specified as:

$$g_\rho = \begin{cases} 0, & r < R \\ 1, & r = R \\ 0, & R < r < R_{max} \end{cases}$$

$$g_c = h_c(r,\Theta) * g_\rho(r)$$

The mapping functions, f_k and f_M, are identical with those of the examples in Sections 3.6.1 and 3.6.2.

First, consider the usual approach. At position x_1 (refer to Figure 4.3), the sensor emits a pulse of energy in the direction α_1, offset from the center where the point target lies. A threshold is applied to the received signal, and the target is detected at range R. Though the return may have come from any point on the range arc shown in the figure, a simple deterministic assumption fixes the position at x_{s1}, directly on the sensor look axis. At a later time, the sensor sends a pulse from x_2 in the direction α_2, and another target is assumed at x_{s2}.

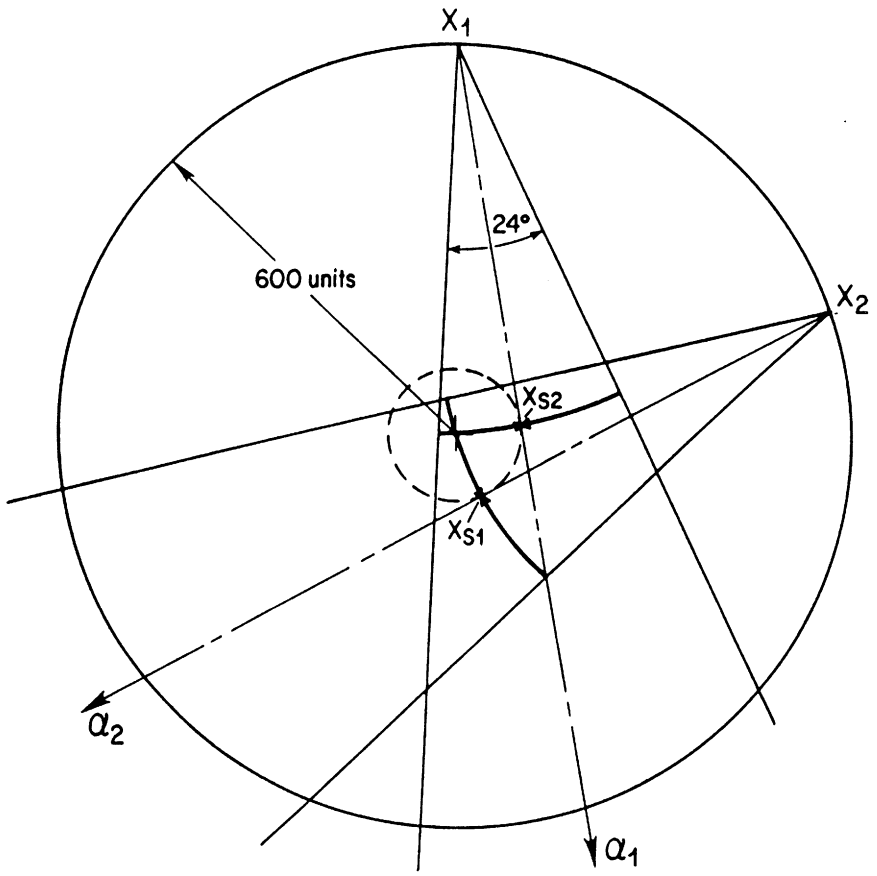

Figure 4.3: Omnidirectional reconstruction: scanning geometry.

If the sensor traced a path around the circle, and emitted regularly spaced signals, the locus of all assumed target positions would form a small circle (dashed line) around the target location. I have assumed a regular scanning pattern and a fixed angular offset from the center, but a similar result would occur for a less regular geometry—an irregular closed shape, or a scattered cloud of assumed locations. My point here is that a simple deterministic approach (threshold then process) can easily lead to erroneous conclusions. A higher-level process that attempts to make sense of the target distribution, with no model of the sensor, might assume either one large target or several small targets.

Figure 4.4 shows the stochastic model after the first event at x_1 has been backprojected (only the central region—800 X 800 pixels—is shown). Here, the global model is identical with the model of that first event. The target-localization probability is distributed along a range arc, and smeared over a finite area. But there is other information now in the model—the region through which the signal has

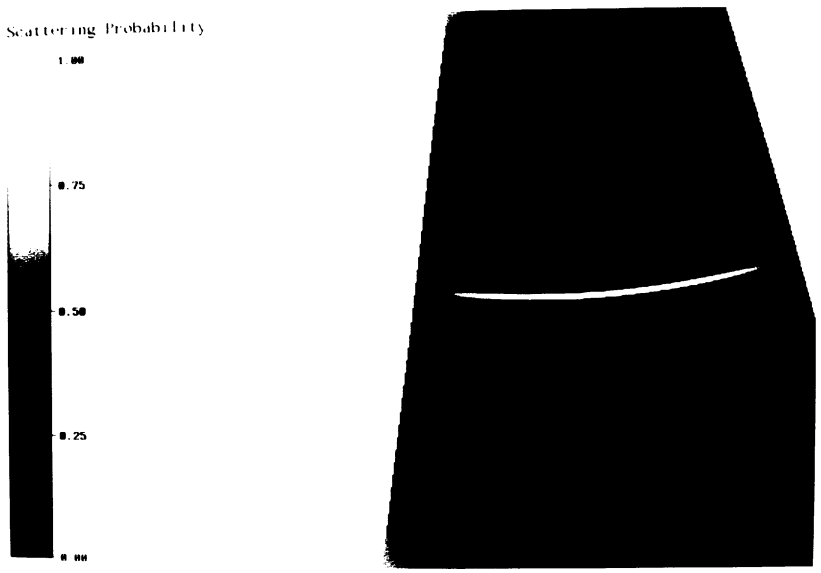

Figure 4.4: Omnidirectional reconstruction: one event.

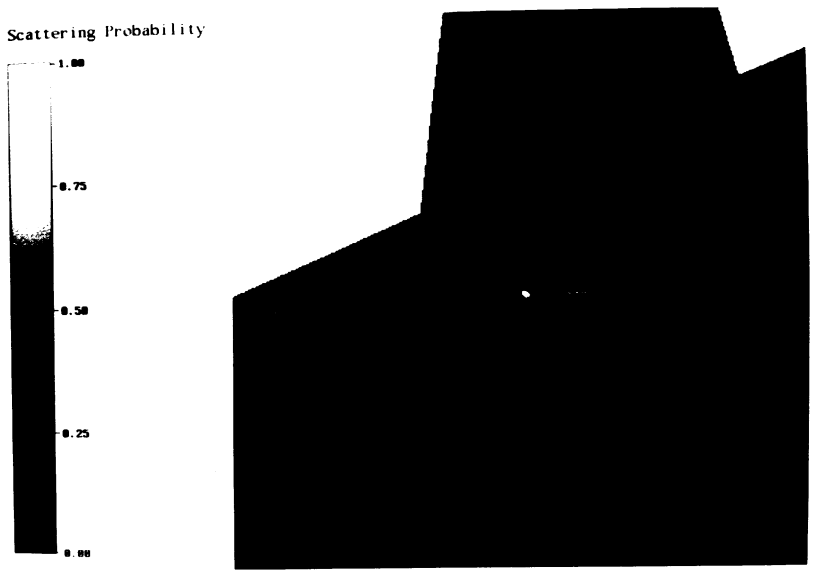

Figure 4.5: Omnidirectional reconstruction: two events.

passed is indicated as probably empty, with a degree of probability that depends on the detectivity at each grid cell. For a free-swimming vehicle, this kind of information is equally important, for example, to plan a path through the region. In this model, the information is explicit, and does not require a consideration of all objects in a database to establish that indirectly.

The model also represents our lack of knowledge about the area. For example, the region neighboring the range arc has transitions from low probabilities to high. Here, cells have values near 0.5, and uncertainty is greatest. With an insider's view of the simulation, it is easy to see for those pixels that uncertainty is associated with the axial and angular error. However, the area outside the detection envelope is an *unknown* region, simply because the sensor has not sampled there. The importance of modeling such ignorance explicitly is that it can be used to guide the development of strategies to fill in information gaps. I mention this again briefly in Chapter 7.

In Figure 4.5, an event at x_2 has been backprojected and combined with the first. Compared with Figure 4.4, the two localization distributions along the arcs have been decayed, and the intersection, where the target lies, has been relatively enhanced. The star-shaped artifact of the summation method (described in Section 2.5.1) is analogous to the crossed target-ray sums at this step. However, these rays represent a finite probability that other scatterers may still be located along the arcs. Real objects, with complex scattering properties, may become effectively invisible with only a small change in look direction.

In a sense, the model of Figure 4.5 gives the graphical equivalent of a navigator's two-point fix. Such an analytical solution can also contain an error estimate, though the geometry might be troublesome. As more returns accumulate, however, the analytical model grows in size and complexity, and an exact solution is unlikely to be maintained. Similarly, object- or feature-based models expand as more data arrives. In theory, the eventual size is unbounded as more objects are found and detail increases. Computation can grow exponentially if new features must be compared with or fitted to a rising number of those already part of the model.

In this, at least, the stochastic model is deterministic. For a given world space, the model's size is fixed by the desired level of detail, and is independent of the information it eventually comprises. Computationally, processing time is bounded and can be forecast accurately for a specified sensor and resolution. Though there is variation from event to event—the number of voxels in a sonar envelope is $O(R^3)$—real sensors have practical limitations. Worst-case limits can be formulated directly, and average processing load can be established for most scenarios.

Figure 4.6 shows the convergent model after a full circular scan at 2° increments. The target has been localized accurately and the remaining area characterized as empty. In this example, the

Figure 4.6: Omnidirectional reconstruction: convergent.

sharpness of the probability peak at the target is conditioned mainly by σ_r, the range uncertainty. Here, the model represents the point-spread function, or impulse response, of a two-dimensional stochastic backprojection for this *particular* scanning geometry and composite localization function. The response is no longer proportional to $1/r^2$ (see Section 2.5.1), but shows a ringing caused by the finite-length ray sums along the range arcs. The effect is a direct result of convolving the beam pattern with each ray. This is not a consequence of the stochastic formulation; rather, it is a reflection of physical reality—the sensor itself only samples over a finite surface.

Though best results are achieved with scans from many look angles, this is impractical (or impossible) for most applications. Figure 4.7 shows the more plausible case of a platform moving along a straight-line trajectory with a fixed, downward-looking sensor. The geometry is similar to that of a profiling sonar aboard a ship or towed platform as it is used to conduct a bathymetric survey. The "bottom" represented here is intended to show the ramp response for this geometry.

The model in Figure 4.8 shows the result of a simulation using the same parameters from the previous example, except that $\Theta_c = 6°$ (12° "beam pattern") and g_p is undefined for all $r>R$ (the signal beyond the shadow boundary is ignored). The sensor is at a fixed altitude 380 units above the

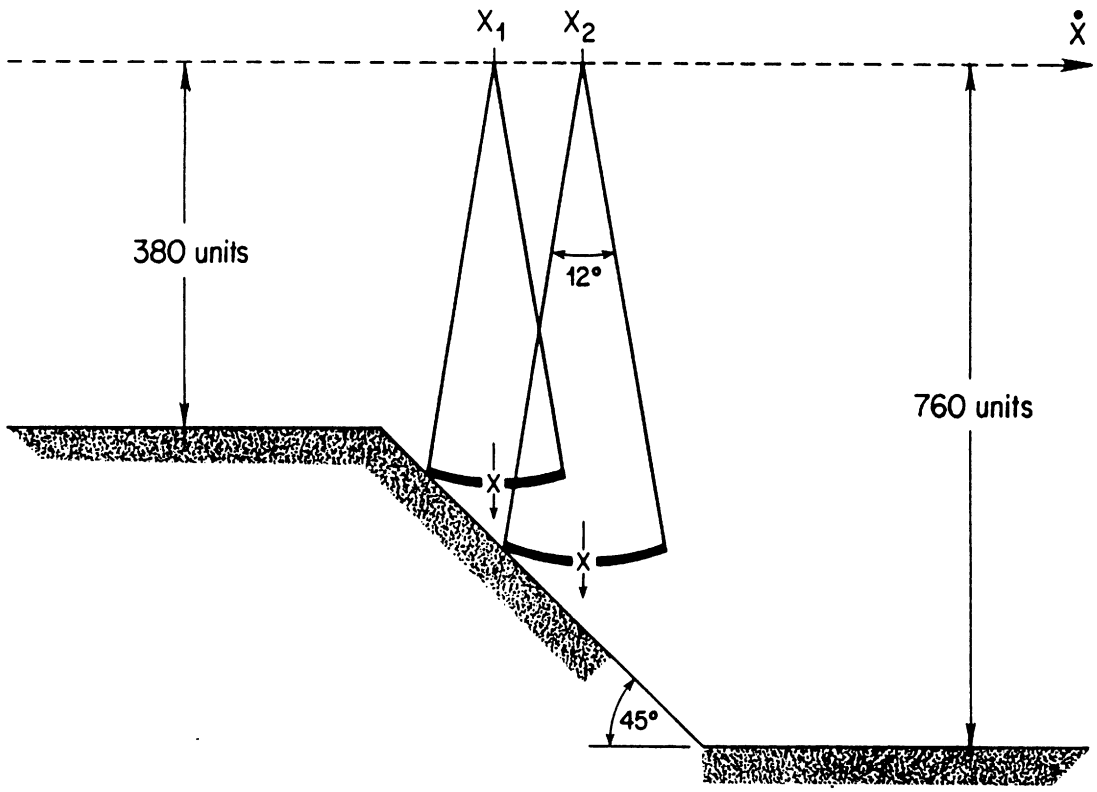

Figure 4.7: Down-look reconstruction: scanning geometry.

upper plateau and 760 units above the bottom floor, and samples at a spacing of 4 units. Only the lower part (800 X 400 pixels) of the model is shown.

Figure 4.9 compares the stochastic surface estimate with a simple deterministic estimate and with the actual surface. A surface profile derived from the model is estimated as the first moment, as described in section 3.6.1; the deterministic estimate is taken as the locus of all points on the look axis at thresholded ranges, as described in the previous example. Figures 4.10 and 4.11 show a similar comparison of the impulse responses. In this simulation, the sensor is at an altitude of 780 units above the floor, and 460 units above a single point target.

In the two simulations, both estimates are biased because of the sensor beam pattern. The flat bottoms are determined correctly, but along the ramp and near the point target a return comes from off axis. However, the stochastic estimate shows less error, since overlapping detection envelopes incrementally constrain target positions. In this simple example, the stochastic estimate gives a

Figure 4.8: Down-look reconstruction: ramp response.

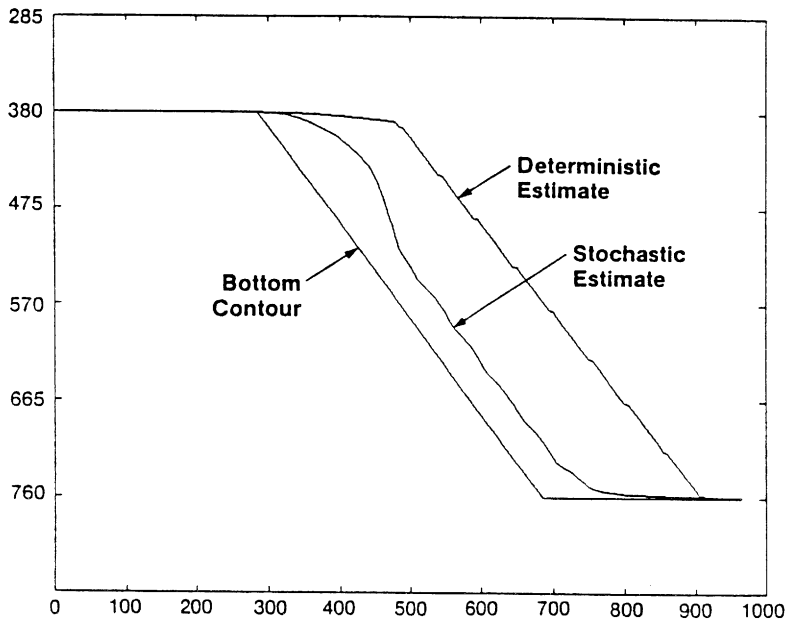

Figure 4.9: Ramp response: surface estimates.

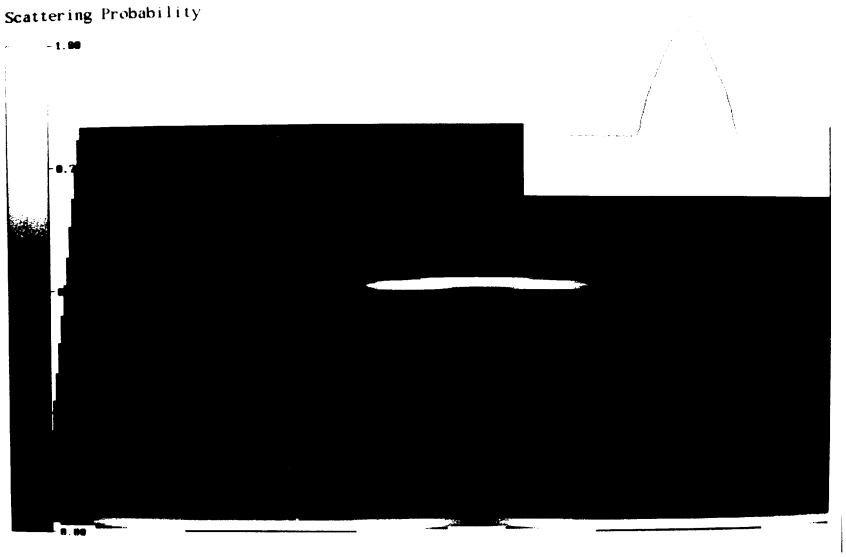

Figure 4.10: Down-look reconstruction: impulse response.

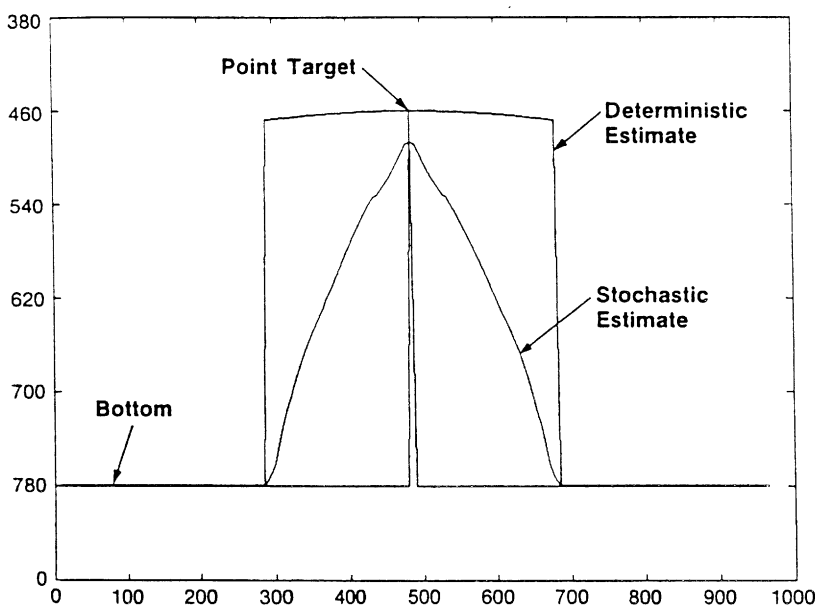

Figure 4.11: Impulse response: surface estimates.

continuous bottom with a broadened "pulse" around the target. Other estimators could recognize a range discontinuity and segregate target and bottom. This point is discussed briefly in Chapter 7.

The three examples just given range from a best-case reconstruction to a worst-case. Where the sensor is able to view a target from many directions, information is maximized and the convergent model approaches a point estimate. In the ramp simulation, the surface is at an oblique angle to the look axis, and there is an overlap between successive detection envelopes at different ranges. The impulse response for this geometry shows the least favorable case. There is only one look direction and little overlap at different ranges.

4.2 OPEN-LOOP MODELING

In the previous section, the examples are contrived to show basic properties of the backprojection formulation without much regard to real-world considerations. The next examples focus on modeling a three-dimensional environment with more realistic constraints. In a typical simulation, I model a free-swimming underwater vehicle with a down-looking sensor. To maintain a perspective of independent events and to avoid any bias from a regularized platform trajectory, positions and attitudes are generated randomly, but are constrained to lie in a volume above the terrain, with a directional axis falling inside a downward-pointing cone.

Relief maps—regular two-dimensional arrays of depth—give the simulator a basis for generating sensor returns. Figure 4.12 shows a perspective view of a relief map used for the first simulations. Here, the intent is to model a natural "underwater" terrain, though in others a regular geometry is introduced. For all simulated events, a model vector is generated by "growing" a detection envelope around the directional axis outward from the sensor. Specular effects are ignored and the intersection of the depth map with any voxel on a range surface increments a counter corresponding to the time axis at that range. The first counter to exceed a threshold of "hits" or scatterers is returned as range to target. Noise-corrupted position, attitude, and range values constitute a simulated measurement vector passed to the modeler.

Simulations used in early thesis research incorporate a variant of the *Moravec-Elfes* [1985] technique, which is extended to three dimensions, augmented by explicit consideration of positioning and pointing error, and given a S/N dependence on range. The binary formulation is similar to that of Section 3.6.1, except that the model vectors are:

$$\rho_k = \{\mathbf{x}, 1, \sigma_\rho\}$$
$$\mathbf{S} = \{\mathbf{x}, 1, \sigma_S\}$$

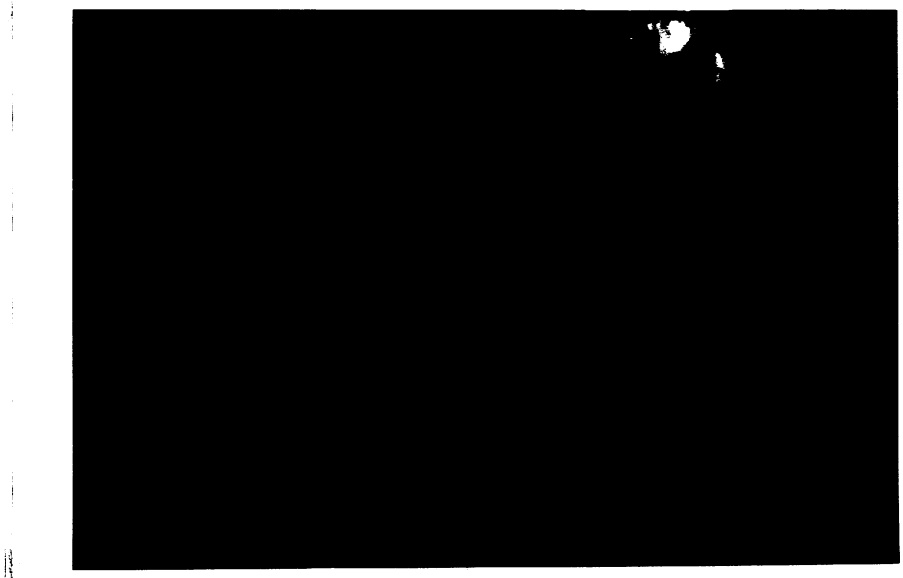

Figure 4.12: Mount St. Helens terrain database.

and the mapping functions are:

$$f_k = \{f_x(r,\alpha),1,g_c(r,\theta)\}$$

$$f_M = \{x,1,f_H(\sigma_\rho,\sigma_{S_{k-1}})\}$$

$$f_H = \begin{cases} \sigma_\rho + \sigma_{S_{k-1}} - \sigma_\rho \sigma_{S_{k-1}}, & \sigma_\rho, \sigma_{S_{k-1}} \geq 0 \\ \sigma_\rho + \sigma_{S_{k-1}}, & \sigma_\rho \sigma_{S_{k-1}} < 0 \\ \sigma_\rho + \sigma_{S_{k-1}} + \sigma_\rho \sigma_{S_{k-1}}, & \sigma_\rho, \sigma_{S_{k-1}} < 0 \end{cases}$$

The 1's above indicate that the normalized scattering strength is implicit in the model, which comprises a single array of probability values. In practice, the probability values are maintained as unsigned bytes, and scaled to the range $-1 < \sigma_S < +1$ for combining.

The combining function, f_H, is a variation of the MYCIN formula [*Shortliffe*, 1976], except that the term $1 - \min[|\sigma_\rho|,|\sigma_{S_{k-1}}|]$ (see Section 2.4) is omitted for probabilities with different signs. my original reason for differing with the Moravec-Elfes approach was mainly to avoid the extra memory requirements for maintaining separate arrays of *Occupied* and *Empty* probabilities. This seemed an artificial distinction, since the two are related to the same sonar signal by a threshold value. For like

Figure 4.13: Terrain modeling simulation with low error.

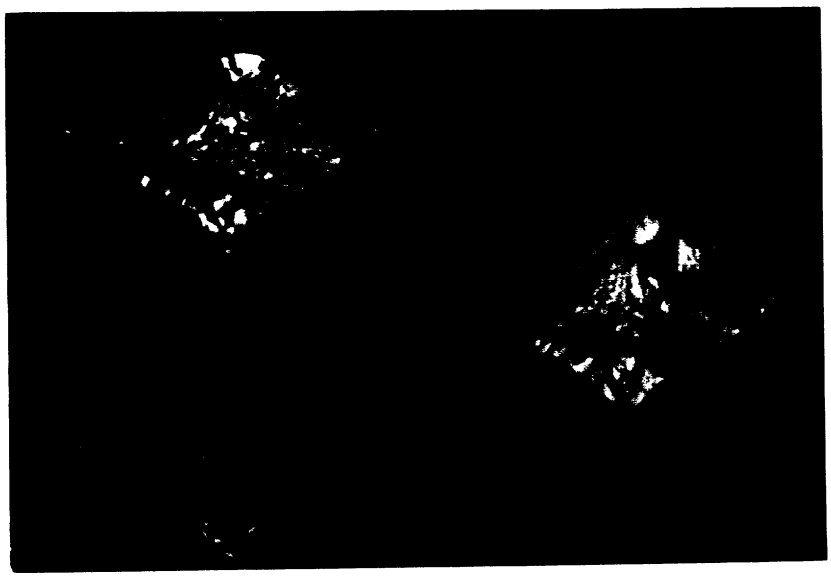

Figure 4.14: Terrain modeling simulation with high error.

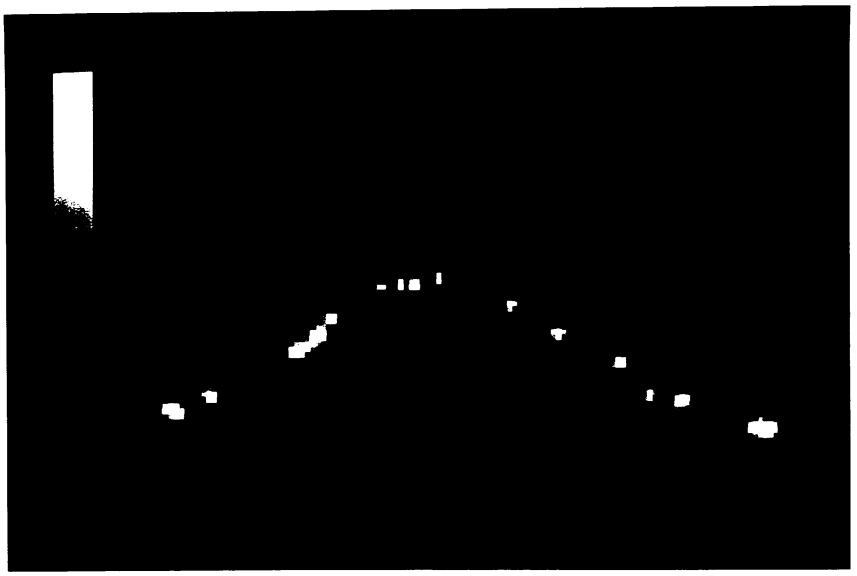

Figure 4.15: Probability plane: sparse data.

reasons, I omit intermediate steps used by Moravec and Elfes that treat the values asymmetrically (*Moravec* [1987a, b] has evolved the approach in several ways and, for example, now uses a Bayesian formulation).

Figure 4.13 shows the result of simulating an active sensor with $\Theta_b = 3°$, $\sigma_\Theta = 1°$, $\sigma_R = 1.5$, $\sigma_x = \{0,0,0\}$, using a model size $\{128,128,32\}$. Here, there is no translational error, and moderate range and pointing error. Measurement noise is simulated by calling a random-number generator, scaling the returned value according to the selected uncertainty value, then adding that error to the exact "measurement" value from which the simulated range return is derived.

The output of the simulation (lower left) bears a faithful resemblance to the original (upper right), with only a small loss in sharpness because of poor sensor resolution. Figure 4.14 shows the result of a similar run with the same parameters, except $\sigma_x = \{3,3,1\}$. The unfiltered rendering (lower left) has been degraded significantly by the large horizontal positioning error, equivalent to about $\pm 10\%$ of the terrain's full vertical relief. However, a spatially-filtered version (upper left) shows that the low-pass representation still corresponds well with the original (right). Experimentation with different ranges of error verifies a graceful degradation with increasing σ.

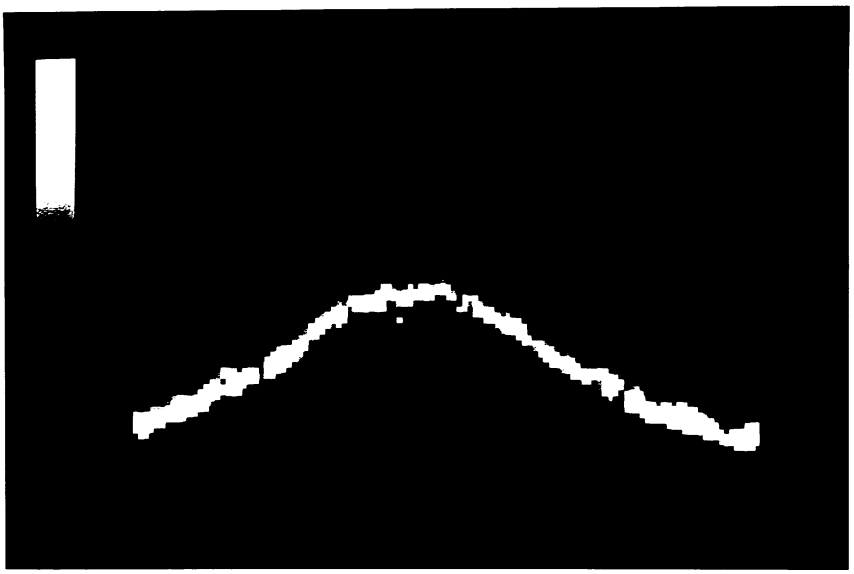

Figure 4.16: Probability plane: 6° sonar.

Figure 4.17: Probability plane: scanning laser.

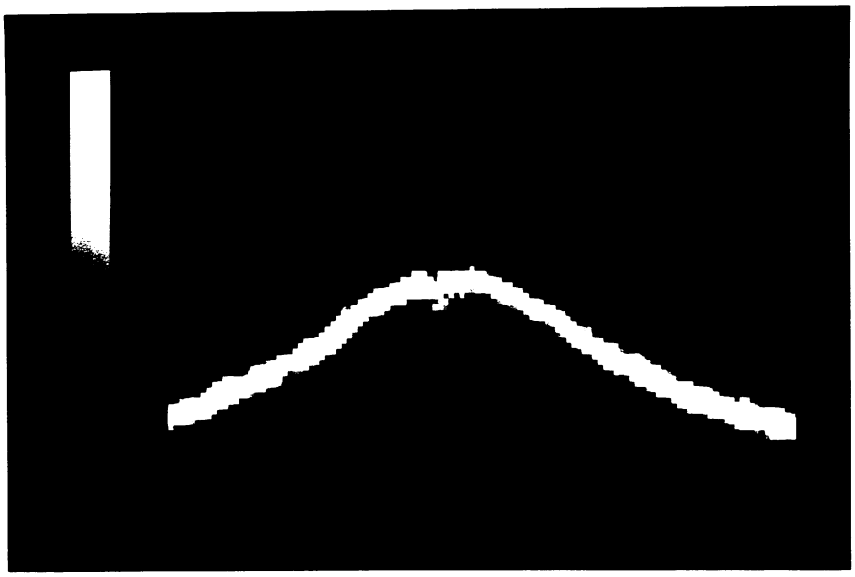

Figure 4.18: Probability plane: convergent.

Planar cross sections through the probability distribution ($\sigma_S(X,y,z)$, for example) show a "fuzzy" surface boundary, whose thickness depends on error and sensor resolution. The overall effect on the stochastic surface estimate is similar to that of a three-dimensional low-pass filter, a result that may be anticipated from the recursive formulation. Figure 4.15 shows one such cross section taken from an early stage of the low-error simulation. Here, the model comprises 1000 events (if we assume 1 unit = 1 m, then a sonar working at scaled ranges would have been scanning for less than a minute).

In Figure 4.16, the model is at an intermediate stage of development (10,000 events), and the surface distribution is more complete. For comparison, Figure 4.17 shows a model at the same stage, but a scanning laser is simulated ($\Theta_b=0$). As discussed earlier, the upper region where $\sigma_S=0.5$ is *unknown*, because the sensor has not sampled there; the lower region of mid-probabilities is *unknowable*, because this sensor cannot see beyond the shadow boundary. If the signal were considered after the strong surface return, then the corresponding volume would be modeled incorrectly as a low-scattering region, and probably empty. The consequences would be undesirable for a path planner trying to find a route through the mountain's "hollow" interior.

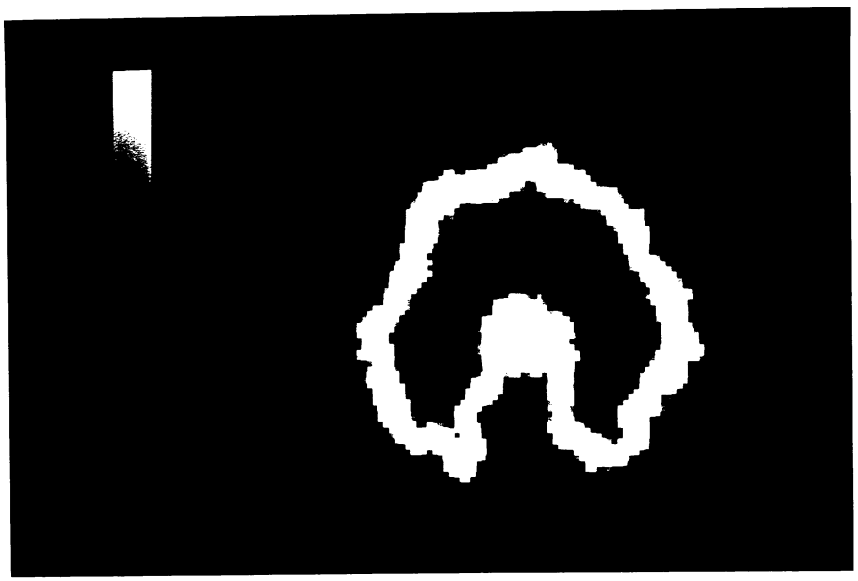

Figure 4.19: Probability plane: horizontal.

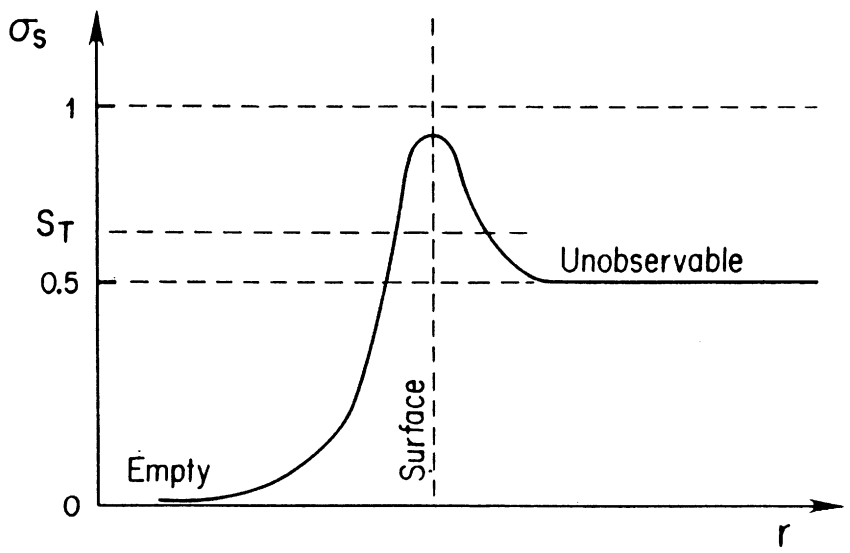

Figure 4.20: Probability profile (generic).

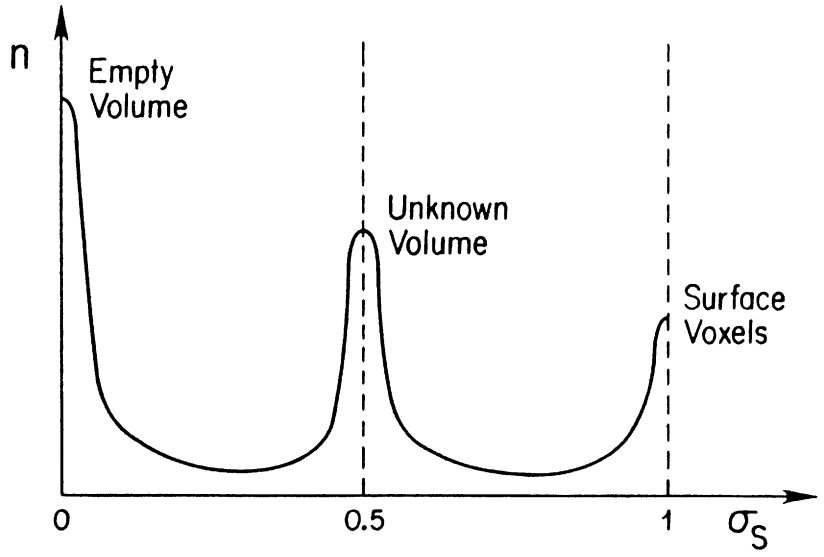

Figure 4.21: Convergent probability histogram (generic).

In Figure 4.18, the model comprises 32,000 events and has converged to a probability distribution that changes little if more events are considered. Figure 4.19 shows a horizontal plane through the convergent model, taken at floor level in the volcanic crater. If we consider a line through this plane, starting at a model boundary and ending at the center, a probability profile is generated as in Figure 4.20. The stochastic surface estimate is associated with the peak in the profile and might be taken as the mode or the first moment. In practice, the profile is often more ragged because of coarse voxel size and random error, so I use the first moment. The scattering threshold (see Section 3.6.1), S_T, is chosen empirically to compensate for asymmetry in the profile when an estimate is made.

If we consider a generic histogram of probability values for such a convergent model, it appears as in Figure 4.21. With no events, the histogram consists of a single impulse at $\sigma_S=0.5$. As the model evolves, peaks begin to form at 0 and 1, and sharpen as more information accumulates. The relative sizes of the peaks depend on the environment and on the choice of reference frame. For example, if the vertical origin were fixed far below average surface level, the model would encompass a large unobservable region and the central mode would dominate. An environment with much detail and surface area would show a larger peak near $\sigma_S=1$. In chapter 7, I discuss how such features might be used to characterize the model and the environment, and to measure convergence.

Figure 4.22 shows another horizontal plane through the model, taken closer to the base of the mountain. The high-probability area is enlarged because the surface slope is closer to being parallel

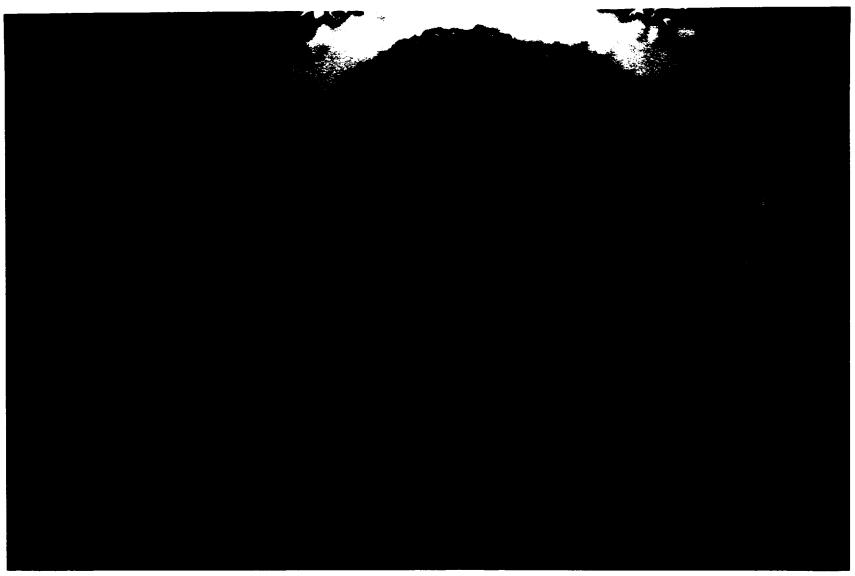

Figure 4.22: Probability plane: interpolated and bit-clipped.

with the plane. The plane is bilinearly interpolated and the image uses a bit-clipped (iterated) color scale to help make two observations. First, the continuous fringes between high- and low-probability areas indicate a high probability gradient, which makes a distinction between the two easier to make computationally. Second, the fringe continuity, at a scale much finer than the coarse voxel size (compare with Figure 4.19), suggests that surface estimates may be valid (certainly they can be made) at subvoxel resolution.

The artificial relief map shown in Figure 4.23 introduces a regular geometry so modeling effects can be seen more readily. Figure 4.24 shows the result of simulating a towed mapping platform with a downward-looking sensor. The track spacing is 4-units, range measurements are taken at 4-unit intervals, and the platform maintains a (smoothed) altitude of about 20 units over the bottom. The modeling parameters are: $\Theta_b = 1.5°$, $\sigma_\Theta = 1°$, $\sigma_R = 1.5$, $\sigma_x = \{0,0,0\}$, model size $\{256,256,64\}$. Results are good, with little loss in fidelity other than high-frequency detail. Figure 4.25 shows a second simulation with the same configuration, except $\Theta_b = 6°$ and $\sigma_x = \{1,1,1\}$. The main effect here is caused by the larger beam width. A comparison with Figure 4.23 shows that the features have been broadened horizontally, similar to the impulse response of Figure 4.10.

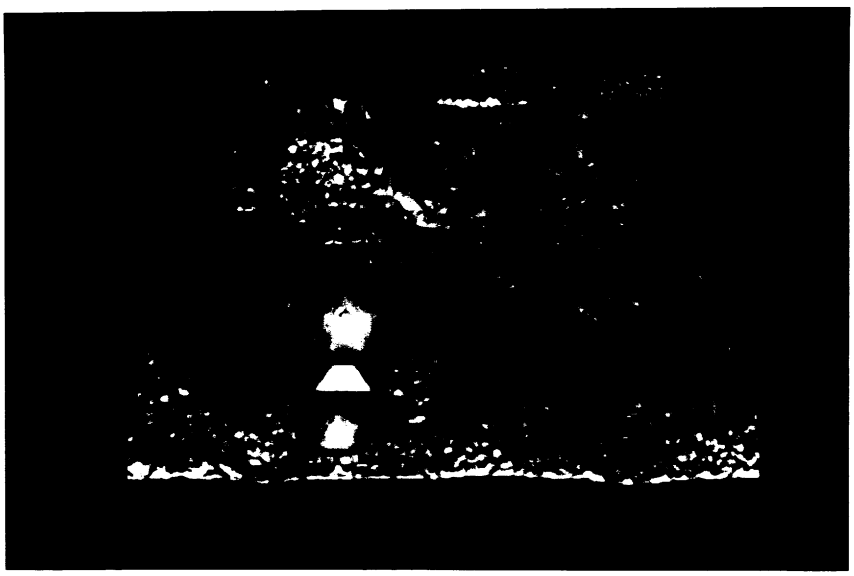

Figure 4.23: Perspective view of seascape relief map.

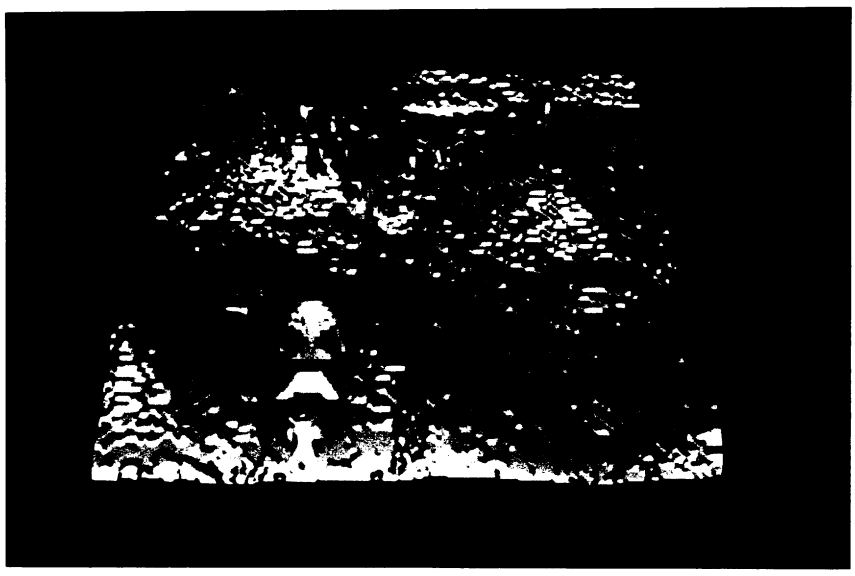

Figure 4.24: Low-error, 3°-beamwidth seascape model.

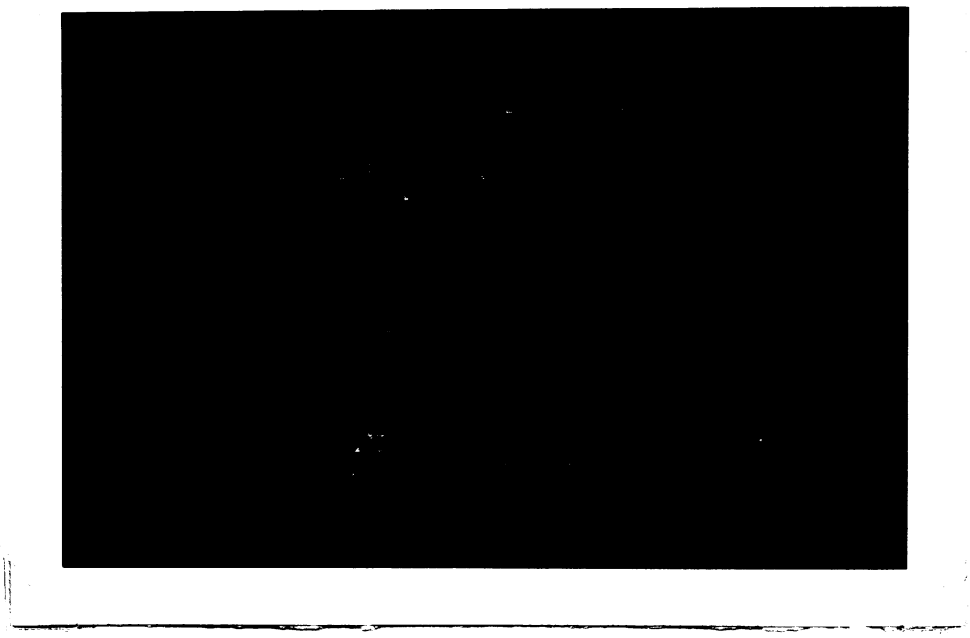

Figure 4.25: Moderate-error, 12°-beamwidth seascape model.

4.3 CLOSING THE LOOP

The simulations just described, and the real data sets treated in Chapters 5 and 6, all investigate model formation and evolution. This thesis focuses on that aspect, since it lays a foundation for higher-level processes that extract information from a model. I have alluded to such uses as obstacle avoidance and path planning, but see an important application in terrain-relative (geophysical) navigation. In this section, I describe the results of two series of simulations that use a stochastic model directly for positioning. The simulations were conducted early in this thesis research to verify that such an approach is feasible.

The first simulation models a free-swimming vehicle that moves about the artificial seascape of the last two examples. As a basis for fixing its position, it uses a prior probability model from the second simulation, where $\Theta_b = 6°$ and $\sigma_x = \{1,1,1\}$. The free-swimming vehicle follows the same track as the towed vehicle and samples at the same horizontal spacing, but flies at a (smoothed) altitude of about 10 units. Here, though, the sensor scans horizontally at angular intervals of 6°, with $\Theta_b = 1.5°$, $\sigma_\Theta = 1°$, and $\sigma_R = 1.5$ units.

For each circular scan, a local model, $\mathbf{L_{k,n}}$, is constructed from events in the scan; the subscripts indicate that the local model comprises n events (n=60) starting with the k^{th}. The events are not combined, but maintained as a list of voxels, where M is the number of voxels in the list. To estimate the vehicle's position after each scan, the local model is correlated with the global model over a range of ± 2 units such that:

$$C_{ij} = \sum_{m=0}^{M-1} \sigma_{\rho_m}(x+i,y+j)\sigma_S(x,y), \qquad \sigma_\rho > 0.5, \quad -2 \leq i,j \leq +2$$

$$\Delta x := \{i,j\} \text{ where } C_\Delta = \max[\{C_{ij}\}]$$

To calculate the sum at each offset (i,j), the list is traversed and, if $\sigma_\rho > 0.5$ (in other words, it corresponds to a probable target), the probability product is accumulated; otherwise, the next voxel in the list is considered. The estimated position correction, Δx, is taken as the offset at which the sum of the products is greatest.

For this simulation, the scan correlations, $C_\Delta(x)$, are shown in Figure 4.26, where the color table is given in scaled integer units ($0 < \sigma_\rho, \sigma_S < 255$)). A comparison with Figure 4.27 shows that the highest correlations are associated with areas of lower relief. There are no surprises here—it shows only that more targets are detected while the vehicle flies at greater depths. As the vehicle moves over the higher-relief features, there are fewer surrounding objects at that scan altitude.

Figure 4.28 shows the absolute position estimation error, $|\Delta x|$, for the same run. There is less dependence on the terrain, which suggests that only a few returns are needed to get reasonable position fixes. However, there is a trend toward higher error along the edges of the model and in the corners. At these locations, there are fewer targets and they generally lie in a narrow sector toward the center. The geometry is less favorable than for an omnidirectional target distribution, and such results may be anticipated. In the central quadrant ($64 \leq x \leq 191$), the root-mean-square (rms) error is computed as 1.6 units. If we use the scale 1 unit = 20 cm, then the rms positioning accuracy is about 30 cm for a region 50 m on a side.

A second series of simulations uses the same vehicle, model, and scanning parameters (including noise), but generates a random walk in the x, y, and z directions. Starting at the center of the model, $x = \{128,128,32\}$, a scan is taken, a position difference is calculated using the same correlation technique, and the perceived location is estimated as:

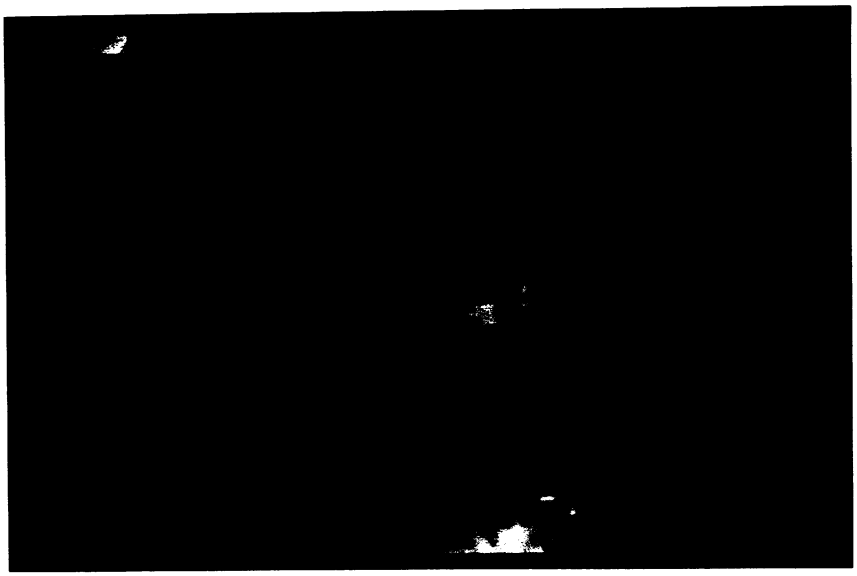

Figure 4.26: Scan correlations for stochastic positioning.

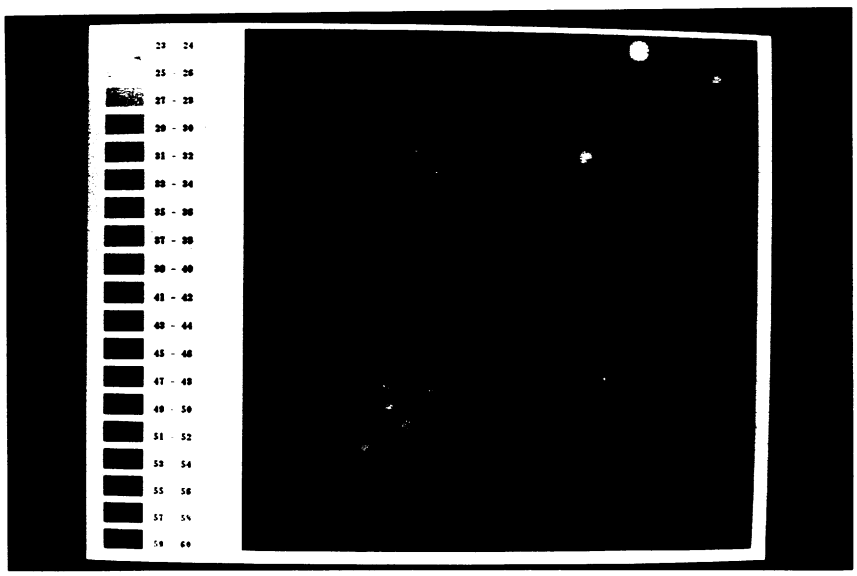

Figure 4.27: Depth map of artificial seascape.

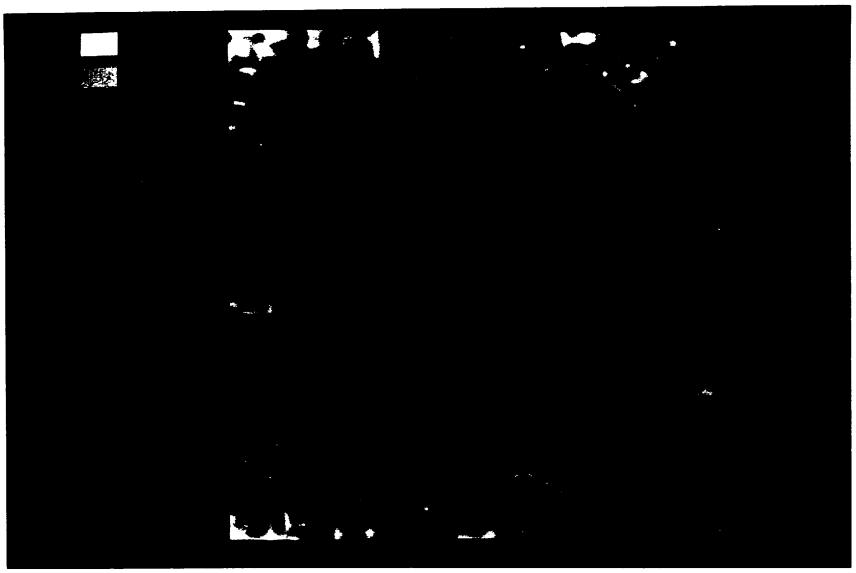

Figure 4.28: Navigation error for stochastic positioning.

$$C_{ijk} = \sum_{m=0}^{M-1} \sigma_{\rho_m}(x_{k-1}+i, y_{k-1}+j)\sigma_S(x,y), \qquad \sigma_\rho > 0.5, \quad -4 \leq i,j \leq +4$$

$$\Delta x_k := \{i,j\} \text{ where } C_{\Delta_k} = \max[\{C_{ijk}\}]$$

$$x_k = \Delta x_k + x_{k-1}$$

$$x_{a_k} = f_\sigma(\sigma_W) + x_{a_{k-1}}$$

where: x_k = perceived position

x_{a_k} = actual position

$\sigma_W = \{4,4,1\}$

The vehicle is assumed to have a sensor for measuring depth, and only horizontal position estimates are considered. The random positioning offsets are generated as described earlier for noise. After each scan, the new position estimate is simply the sum of the last estimate and the estimated correction. Actual positions are used to generate the scan measurements, and are unknown to the correlator and estimator. For the parameters above, the simulation runs for several hours with absolute position error less than 9 units. With more noise or larger σ_W, the position error increases without bound after a few iterations (the simulation stops if x_k exceeds the model boundary).

As in the previous example, performance is sensitive to noise and other uncertainty, and depends on the terrain itself. This simulation is intended to show that such terrain-relative positioning is plausible using the stochastic model directly. However, in one sense this is a worst-case scenario, since a real system would likely be equipped with inertial or velocity sensors, and would have a model of platform dynamics to help estimate position changes. In another sense, it is an ideal case because a prior (stochastic) model is available. However, a concurrent approach to modeling and positioning can be reformulated with the correlation technique described above. I will describe an unimplemented approach that uses a variation of *fading* as it is applied to the HILARE robot [*Giralt et al.*, 1979].

As before, the local model, $\mathbf{L}_{k,n}$, comprises a list of the n most recent events. The global model, $\mathbf{S}_{k-1}$, has not yet been updated by the newer events $\{\rho_k, \rho_{k+1}, \ldots, \rho_{k+n-1}\}$. At time t_{k+n}, the next event is added to the local model, and the position correction Δx_{k+n} is estimated. The new correction is propagated backward through $\mathbf{L}$, and corresponding estimates are updated as:

$$\Delta x'_{k+j} = \kappa_j \Delta x_k + (1-\kappa_j)\Delta x_{k+j}, \qquad 0 \leq j \leq n+1$$

$$\kappa_j = \frac{j+1}{n+1}$$

The correction progressively diminishes as it approaches t_k, reflecting the cumulative nature of dead-reckoning error. Here I use a simple linear fading, but more sophisticated approaches are conceivable (a Kalman filter, perhaps). As described, the process is a recursive estimator, and $\Delta x'_k$ is refined over n iterations. As a final step, ρ_k is removed from $\mathbf{L}$ and merged with the global model. A brief discussion on starting with a blank slate is reserved for Chapter 7.

4.4 SUMMARY OF IMPORTANT POINTS

In this chapter I have presented the results of several computer simulations, most of which were conducted in the early stages of thesis research. My intent has been to show (1) basic properties of stochastic backprojection; (2) the method is computational tractable; (3) an appropriate fidelity is realizable; and (4) stochastic modeling establishes an appropriate foundation for higher-level processes. Some important points mentioned are:

1. Deterministic estimates may lead to erroneous assertions about the world when thresholds are applied to measurements without consideration to a sensor model.

2. Other information made explicit in the model—emptiness and ignorance, for example—can be used directly by higher-level processes.

3. Analytical and feature- or object-based approaches may lead to models of unbounded size, as increasing detail is added.

4. A deterministic model size and processing load allow computational hardware and timing considerations to be forecast for an application.

5. More look angles generally mean more information, and improved modeling results.

6. With enough events, models converge to a low-pass, "fuzzy" surface distribution, which degrades gracefully as more uncertainty is introduced.

7. Surface estimates of subvoxel resolution can be extracted, and may offer a path to greater efficiency by allowing a coarser spatial partitioning.

8. Terrain-relative positioning can be implemented directly with a stochastic model, and may be extended to concurrent modeling and positioning applications.

Chapter 5

APPLICATION DATA SETS

The simulations described in the last chapter were helpful in developing a computational framework and in focusing research efforts on important issues. However, an important test is the processing of real-world information, which usually does not fit all our simulated preconceptions. In this chapter, I describe four applications tested on data acquired in the field. With one exception, these are from large-area mapping systems.

5.1 MULTIBEAM BATHYMETRIC: SEA BEAM

Sea Beam is a 12-kHz bathymetric profiling system that generates a swath map of the ocean floor using sonar arrays mounted to a ship's hull [*Renard and Allenou*, 1979; *Farr*, 1980]. The transmitting array generates a pitch-stabilized beam pattern of 54° across-track and 2-2/3° along-track. The receiving array forms 16 beams, uncompensated for motion, at 2-2/3° intervals, each with a beam pattern of 20° along the track and 2-2/3° across the track. The composite beam patterns, then, are roughly 2-2/3° by 2-2/3°, evenly spaced over a total field of view 20° athwartships (see Figure 5.1). This produces an effective swath width about 73% of water depth.

The echo-processing electronics comprise 16 receivers, corresponding to the ensemble of receiving beams. For each beam, a bottom-tracking gate determines the interval over which the received signal is digitized (to a resolution of about 2.5 m of slant range). The slant range is determined from the center of mass of all signal components above a predetermined threshold level. For each valid detected depth, simple geometry is applied to generate a slant-range corrected depth (referenced to 1500 m/s speed of sound) and a track offset corrected for refraction and nominal speed of sound. The depth/offset values are recorded on magnetic tape and plotted on a strip-chart recorder.

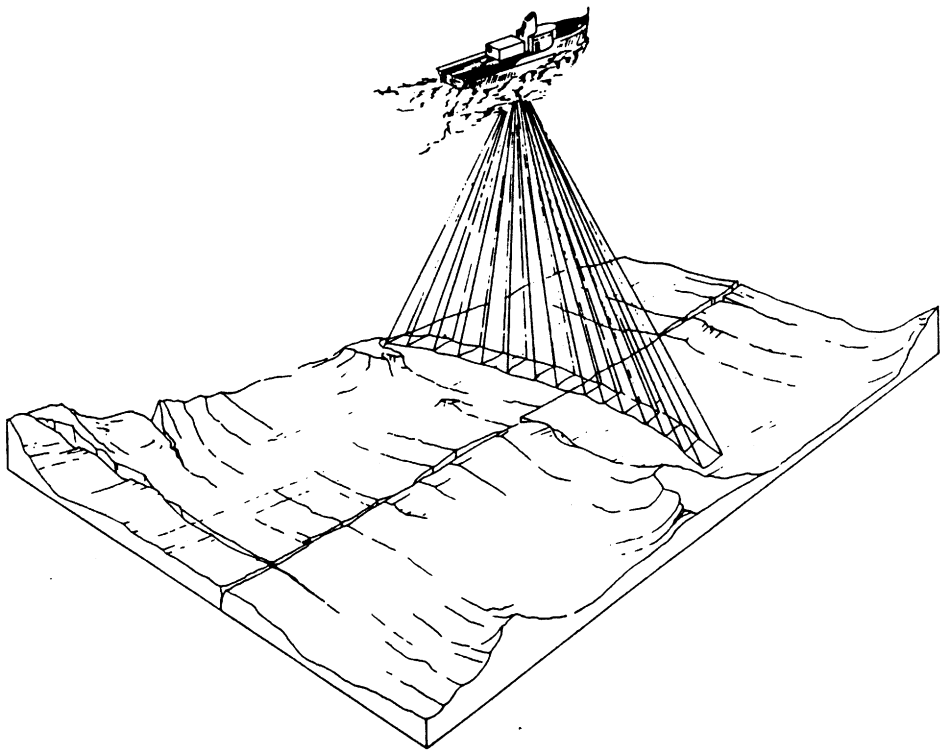

Figure 5.1: Sea Beam profiling geometry.

De Moustier and Kleinrock [1986] discuss several depth artifacts introduced by the Sea Beam processing that are related to bottom tracking, gating, and thresholding.

To test the binary model on real-world data, a Sea Beam tape with merged navigation was acquired from the NECOR facility at the University of Rhode Island, by courtesy of Dr. Robert Embley of NOAA [*Embley et al.*, 1988]. The tape contains raw Sea Beam data without multi-ping averaging, usually performed before the data are gridded in postprocessing applications. The ship's navigation data are derived from intermittent Transit satellite fixes supplemented by long periods of dead reckoning. The navigation had not been corrected with track-crossing correlations, usually applied before gridding to improve results. Positioning errors on the order of several hundred meters are common in the data.

To generate a model, the data are read sequentially as though the processing is being carried out in real time aboard ship. For each ping cycle, depth and track offset are used to generate a slant

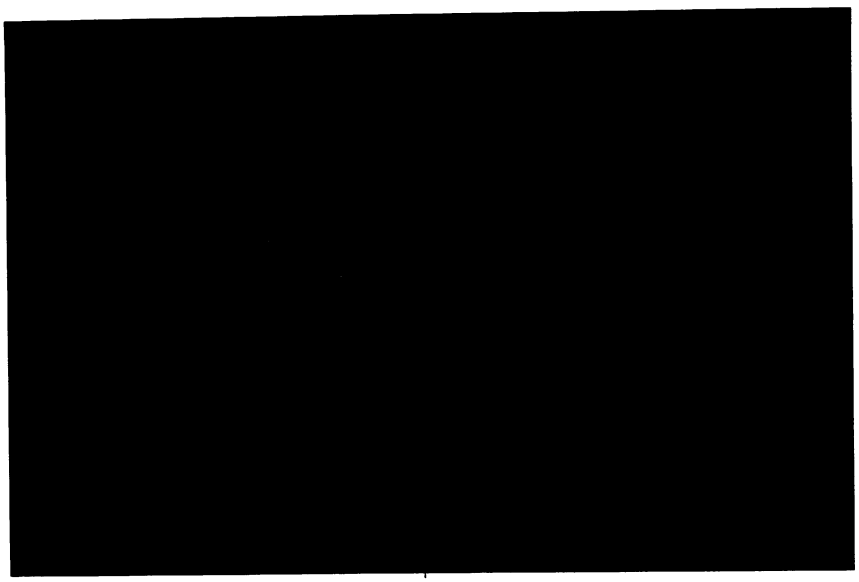

Figure 5.2: Sea Beam bathymetric tracks.

range, assumed to lie on the axis of each corresponding beam. Because the depth/offset pairs are interpolated and refraction-corrected, an effective cross-track look angle is also calculated for each "beam." Ship's heading and position are read from the file and augmented by the calculated slant ranges and look angles to constitute a maximum of 15 event vectors (the number depends on the valid data pairs on the tape). The vectors and mapping functions are the same as in the binary modeling simulation of Section 4.2, except that:

$$f_H = K_\rho \sigma_\rho + (1 - K_\rho)\sigma_{S_{k-1}} \qquad 0 < K_\rho < 1$$

where: K_ρ = filter constant

The formulation given here is an early (relative to thesis research) approach to modeling as a purely recursive filter. It was motivated by observations of large registration errors in the Sea Beam data, and accommodates the error by controlling the "memory" in the model. Higher values for K_ρ take advantage of the relative accuracy in dead-reckoned navigation among neighboring ping cycles, so the shape and crispness of bottom features are preserved. Lower values tend to blur the model and, in effect, result in a three-dimensional low-pass filtering of the estimated surface. This is consistent with the global uncertainty introduced by navigation error.

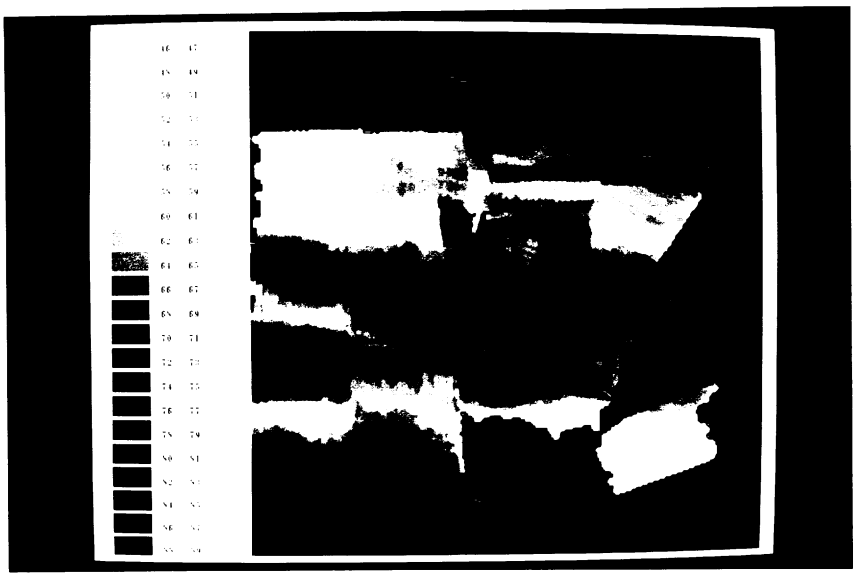

Figure 5.3: Sea Beam high-error tracks.

Figure 5.2 shows a section of the surface derived from two short tracks, with about a 1.8-km swath width at these depths (2200-2800 m). The area is 4-km on a side and part of the Galapagos Rift, a crustal spreading zone characterized by the terrain shown here. The upper track parallels a tectonic ridge, and the lower track diagonally crosses a second ridge, separated from the first by a rift valley. The modeling parameters are: $\Theta_b = 3°$, $\sigma_\Theta = 1°$, $\sigma_R = 1.5$ voxels, $\sigma_x = \{4,4,4\}$ voxels, $K_\rho = 0.5$, and the model size is $\{128,128,32\}$ (the model is confined to a region encompassing the bottom, so the large water column is not considered). The coarse voxel size corresponds to a resolution of about 30 m in all three coordinates, but depth estimates of sub-voxel precision are extracted.

Figure 5.3 demonstrates the registration error that appears in the uncorrected navigation data. In this model, all track segments come from the same day's survey. Figure 5.4 shows the result of modeling with data collected over a 3-week period (not all at this location). The data include several days of random swinging on station, since a sequence of ALVIN dives was conducted at the site. All data are considered, not just the straight, even-velocity tracks normally culled and correlated to make a good fit. The low-pass nature of the surface estimate is plainly evident as the result of including much redundant data with large positioning error ($K_\rho = 0.20$).

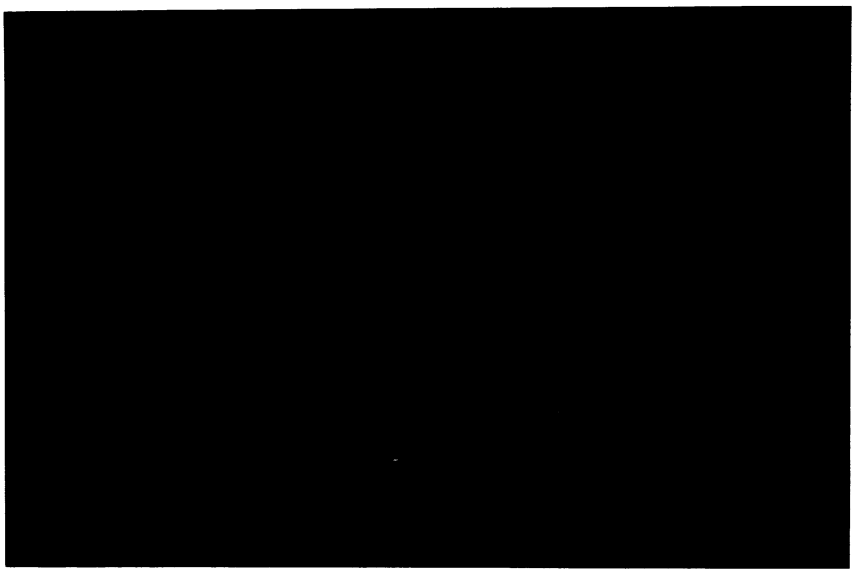

Figure 5.4: Sea Beam bathymetric surface.

In all the examples just described, a "real-time" display is generated concurrently with the incremental modeling. The display is not just written over by new data, as for Sea Beam shipboard plotting. For each ping cycle, the beams are modeled and backprojected, and any modified voxels are flagged. For those voxels, a new surface estimate is extracted and filtered locally, and the display is updated only if the change propagates that far. In all cases, modeling and display is accomplished at speeds greater than or equal to real-time data acquisition rates, using a 16-MHz 68020 workstation with Weitek floating-point chips.

In fact, real-time displays are the most computationally intensive process with the current hardware configuration. In a "batch" mode with no display, modeling alone is carried out at about six times the data acquisition rate; modeling with a real-time color-contour display at three times the data rate; and modeling with a three-dimensional shaded perspective display at a rate slightly faster than real time. In the latter case, the visual effect is that of "painting" a three-dimensional seascape as the ship moves along its track. These performance characterizations are offered only as qualitative assessments to show the process is feasible for real-time shipboard processing with modest computational assets. The modeling and display routines are not optimized for performance, and an increased efficiency may be achievable.

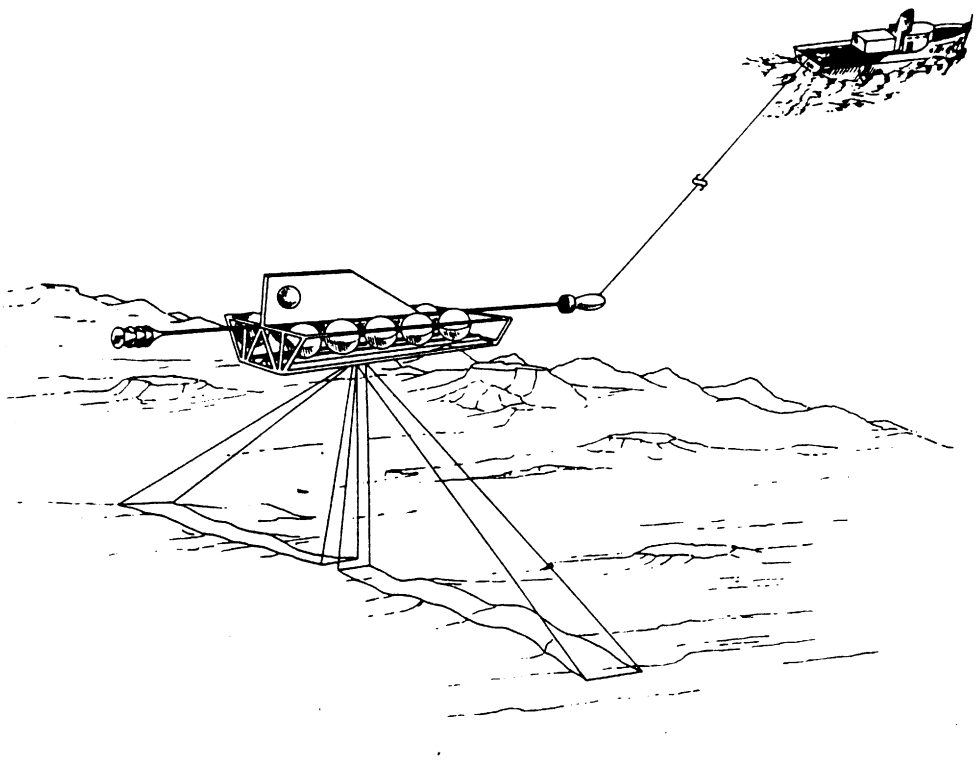

Figure 5.5: Sea MARC I scanning geometry.

5.2 TOWED SIDESCAN: SEA MARC I

Sea MARC I is a relatively long-range 27/30 kHz (port and starboard) sidescan sonar that creates an intensity map of the ocean floor [*Kosalos and Chayes*, 1983]. The system uses a linear array to generate a beam pattern of 1.7° horizontally and 50° vertically (refer to Figure 5.5). As the sensor is towed near the bottom, the acoustic signal reflected back from the seafloor is amplified, demodulated, filtered, and digitized. Sea MARC I uses a variable-increment sampling scheme that incorporates a slant-range correction based on altitude over the bottom. On board ship, the data are recorded on magnetic tape and an image is created line-by-line on a gray-scale paper recorder. Each line represents the one-dimensional signals received back from outgoing pulses on each side of the towed sonar fish.

Figures 5.6-5.9 are intensity maps generated from Sea MARC I data contributed by the Lamont-Doherty Geological Observatory (LDGO), courtesy of Drs. Kim Kastens and Bill Ryan [*Kastens et al.*, 1986]. These are 5-km-swath-width records with a 2.5-m cross-track resolution. The terrain is typical of the crustal spreading, fracturing, and volcanic activity of the Clipperton Transform Zone in the eastern Pacific. The merged navigation records are derived from LDGO processing, which draws from satellite fixes, dead reckoning, and short-baseline fish tracking with depth-contour matching. The navigation is interpolated with cubic splines to generate a smooth, finished track. Heading and altitude are measured by Sea MARC sensors for each ping cycle, which occur at about 3-s intervals.

Figure 5.6 shows a linear display of unprocessed data, similar to the gray-scale paper recording produced aboard ship. The black center stripe in the image is a 200-m data gap windowed by the system. Figure 5.7 shows a map linearly compressed by averaging to a scale approximating the true geometry, roughly the equivalent of a speed-corrected paper recording.

In reality, though, the fish does not follow a straight, constant-speed track that suits the linear recorder. Figure 5.8 is a map created by a postprocessing technique that corrects each line for position and heading using a grid of averaged intensity bins. In these first three images, the wide, dark and light bands along the track are caused by the sonar's fixed TVG. The narrow, dark lines across the track are "dropouts" where there are missing or corrupted data, probably caused by yawing motions of the fish.

Figure 5.9 shows the result of two-dimensional backprojection applied to the data in a sequential, real-time manner. The modeling vectors are as given in Section 3.6.2, except that yaw rate is unavailable and only one heading measurement is recorded for each event. The mapping functions are the same also, except that:

$$g_\delta = g_N(k) h_N(j,k) \delta_{jk}$$

$$g_N = \xi_k / \delta_{ak}$$

$$\xi_k = \kappa_\xi \delta_{ak} + (1-\kappa_\xi)\xi_{k-1}, \quad k \neq 0, \quad 0 < \kappa_\xi < 1$$

$$\xi_0 := \delta_{a0}$$

$$\delta_{ak} = \frac{1}{J} \sum_{j=0}^{J-1} \delta_{jk}$$

where:

- j = range index
- J = number of range bins
- κ_ξ = filter constant

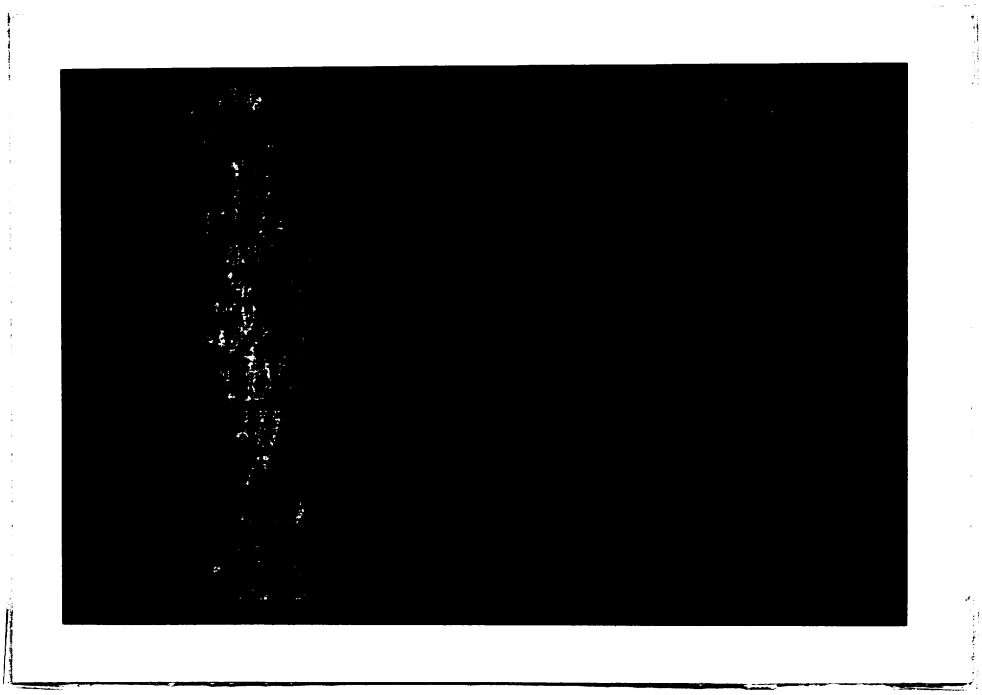

Figure 5.6: Sea MARC I: raw linear map.

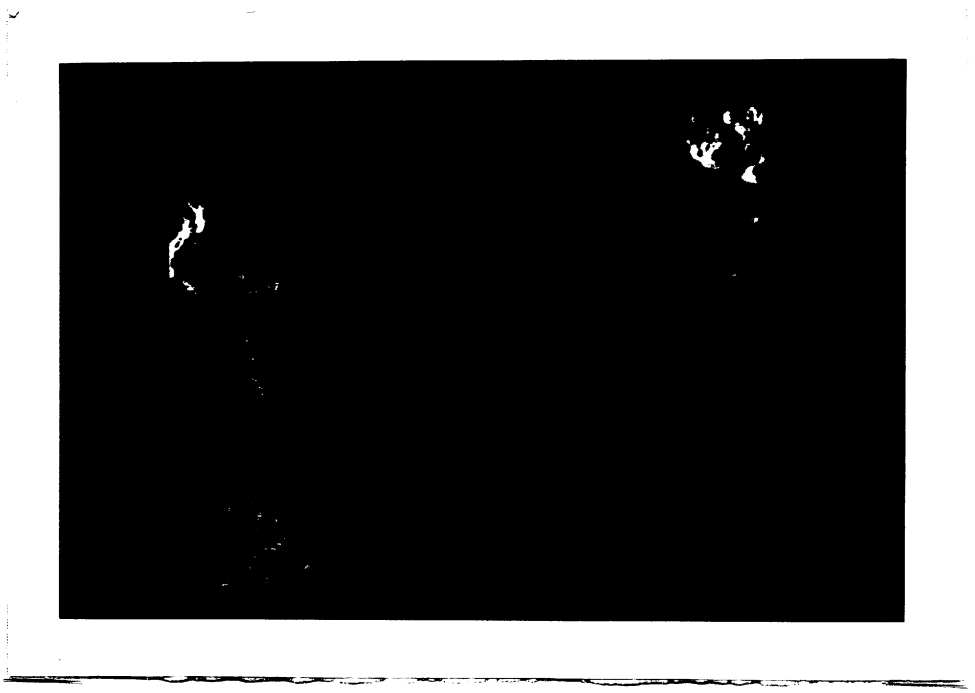

Figure 5.7: Sea MARC I: averaged linear map.

Figure 5.8: Sea MARC I: grid-averaged map.

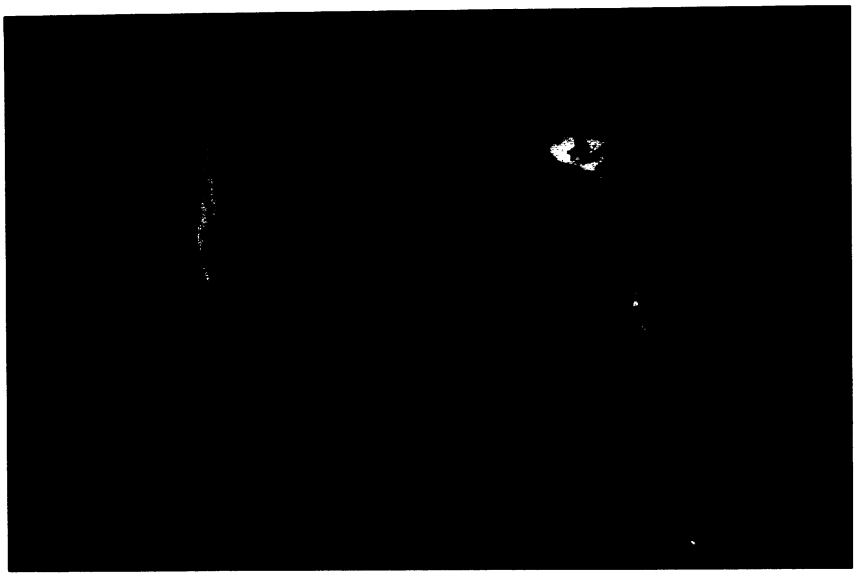

Figure 5.9: Sea MARC I: backprojected and corrected map.

and Θ_c is computed as in Section 4.2. The striping gain, g_N, is based on the idea of yaw compensation developed in Section 3.6.2. However, since heading is not measured for the entire event, the correction is computed with the average line intensity. The rationale is that for local neighborhoods the average intensity should vary slowly and smoothly.

For the model in Figure 5.9, $\kappa_\xi = 0.05$ and the other modeling parameters are $\Theta_c = 2°$, $\sigma_R = 2.5$ pixels, $\sigma_x = \{0,0\}$ pixels, $\kappa_\delta = 0.05$, and the model size is $\{1152,900\}$ pixels. Again, the model is mapped to the frame buffer for convenience. In comparison with Figure 5.8, it is apparent that g_N is effective in reducing the striping noise, and h_N enhances the visual information by compensating for the fixed TVG. As an overall result, the image quality is improved over simple intensity averaging.

An important consideration, though, is the incremental nature of the process. Similar to the Sea Beam example, the model is grown along the track, on a line-by-line basis, as though processing is being conducted in real time aboard ship. Though such smoothed navigation was not available at the time the data were collected, near-real-time navigation is feasible in many of today's applications. Processing time at this resolution is about eight times the data acquisition rate, including a real-time display.

A preferred approach is to start with the raw digitized signal (no slant-range correction), more complete attitude information with a higher update rate, and process three-dimensionally (discussed in Chapter 6). Rather than try to extract all the assumptions built into the data set as it was received (altitude, slant range), I adopted a minimal two-dimensional formulation. The obvious advantage, though, is increased processing speed at a higher resolution. As it is, the method is suitable for real-time model building and display aboard ship, an attractive option for search, survey, and geological mapping and interpretation.

5.3 TOWED SIDESCAN: SEA MARC II

Sea MARC II also is a towed sidescan sonar that evolved from Sea MARC I, and the two share many subsystem components [*Blackington et al.*, 1983; *Hussong and Fryer*, 1983]. However, Sea MARC II is a dual-receiver sonar, and uses the phase angle between the signals from two closely-spaced transducers to estimate bathymetric relief. In practice, the system is towed over a relatively flat bottom, and statistics are compiled that allow a conversion from electrical angle to look angle. With this technique, a ray-bending correction is automatically included as part of the lookup-table conversion.

Unlike the deep-towed Sea MARC I, the system operates near the surface, typically just below the thermocline at depths of about 100 m. The sonar also uses lower frequencies (11/12 kHz) and can map swaths up to 10 km wide with a cross-track resolution of 5 m. The beam pattern is shaped by a linear array that produces an angular spread of 2° horizontally and 40° vertically. The received signal is processed and digitized with a built-in slant-range correction as for Sea MARC I.

On board ship, the phase angles are accumulated in 75-m range bins (for the 10-km swath width), and a single phase estimate is taken from the distribution. The slant range and phase lookup are used to generate depth/offset pairs, with a bathymetric accuracy about 3% of water depth. The raw amplitude and phase data are recorded on magnetic tape, along with the computed depth/offset pairs. Shipboard presentation uses linear paper recorders for gray-scale sidescan intensity and for color range-bin plots of bathymetry.

Figures 5.10-5.13 show intensity and bathymetric maps generated from backprojected Sea MARC II models. The data are from the Siqueiros Transform in the eastern Pacific, and cover a region about 200 km in east-west extent. The Sea MARC II was operated by the Hawaii Institute of Geophysics (HIG), and the data were supplied by courtesy of Dr. Dan Fornari of the Lamont-Doherty Geological Observatory and Dr. Dave Gallo of the Woods Hole Oceanographic Institution. Ship's navigation is derived from GPS, Transit, and dead reckoning. The position fixes were manually adjusted aboard ship, but are reported to be highly consistent.

5.3.1 Intensity Model

With a few exceptions, the intensity modeling is similar to that of Sea MARC I: (1) A yaw correction, g_N, is not implemented; however, the data are of good quality and final results do not suffer appreciably. (2) The cross-track normalization, h_N, is computed as a fixed value (for each range bin) from a preliminary pass through all data. (3) The beam pattern is modeled with uniform intensity over a narrower angular extent; in effect, this is similar to a directivity factor [*Urick*, 1975]. (4) Scaled-integer processing is used rather than a more precise floating-point implementation. For the larger-scale models, the effect is negligible; at higher resolution, though, aliasing and round-off errors degrade results.

A large-scale model of the entire survey area is shown in Figure 5.10, covering 2° of longitude and 1½° of latitude (about 222 km by 172 km). The representation is not a *mosaic*, in the usual sense, because the modeling is accomplished continuously and incrementally. Figure 5.12 shows a higher-resolution segment from the southeast corner of the survey area, covering 40' of longitude and

Figure 5.10: Sea MARC II: sidescan intensity (large-scale).

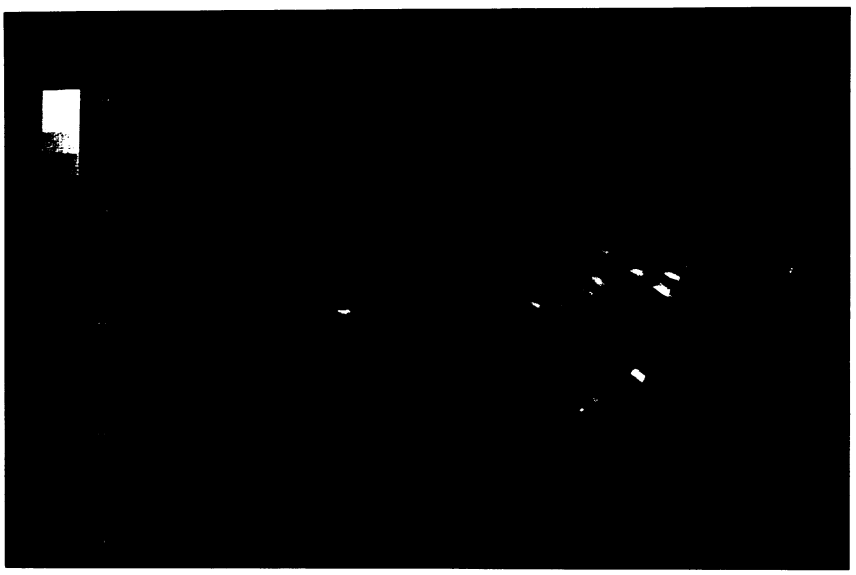

Figure 5.11: Sea MARC II: bathymetry (large-scale).

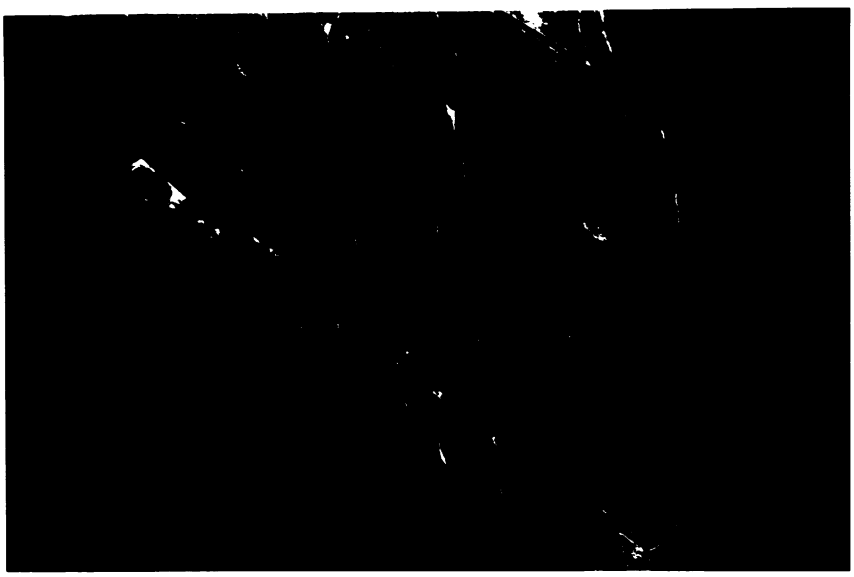

Figure 5.12: Sea MARC II: sidescan intensity (medium-scale).

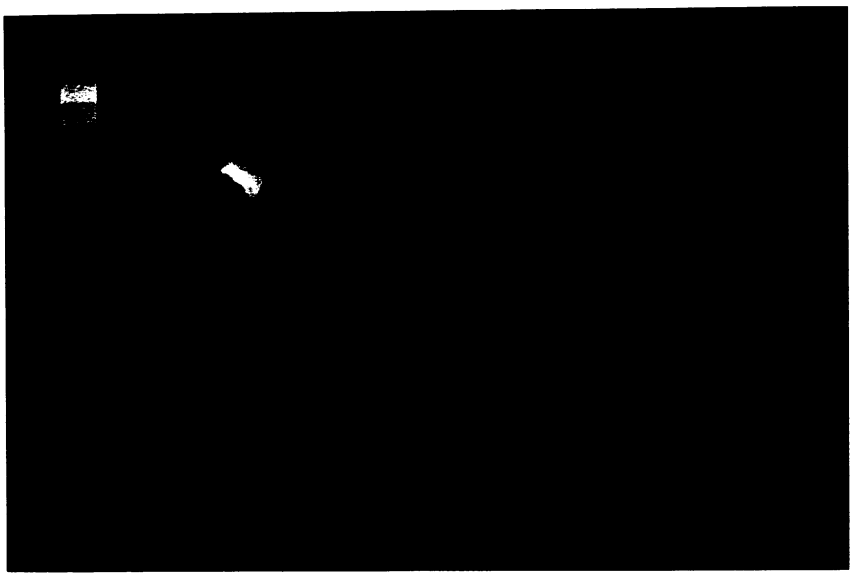

Figure 5.13: Sea MARC II: bathymetry (medium-scale).

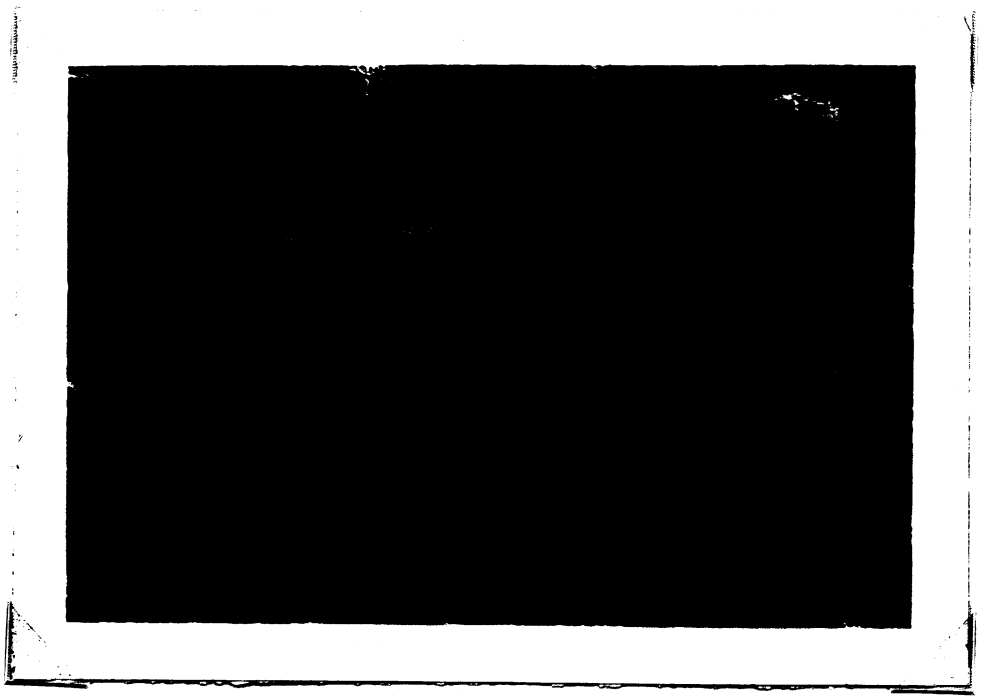

Figure 5.14: Sea MARC II: sidescan intensity (small-scale).

30' of latitude (about 74 km by 56 km). Note that the continuity of linear features is preserved through the turns.

Figure 5.14 shows a higher-resolution model segment from the center of the survey area, covering 20' of longitude and 15' of latitude (about 37 km by 28 km). As for Sea MARC I, the black center stripe is gated by the system. The shorter strips across the track are data gaps between the original tapes. In this survey there is about 5% track overlap on average. No special techniques are used in these overlapped areas; in effect, they show an average intensity from opposing look angles. Geometric distortion from the planar-bottom assumption also causes misregistration of features that have any significant relief.

Processing time at the resolution of Figure 5.10 is just under one hour (for about 3½ days of survey data). At the scale used for Figure 5.12 (one third that of Figure 5.10, or nine times as many pixels) processing takes about 1½ hours for the entire data set. The increase in processing time is not linear with resolution because there is significant overhead in file access and initialization for each event. The relative speedup over Sea MARC I processing is partly because of the difference in data acquisition rates (about 3:1). The remainder is accounted for by the simplified Sea MARC II implementation.

5.3.2 Bathymetric Model

The bathymetric processing uses HIG-generated depth/offset values. These are interpolated to a resolution that matched the sidescan intensity data, and processed as a "two-dimensional" event. Figures 5.11 and 5.13 show examples of bathymetric maps that correspond to the first two intensity models. In Figure 5.13, the high-relief feature in the upper left is an undersea volcano (the crater is just visible as a small "dimple" near the center). Between the two lower tracks, there is a noisy strip where they overlap. This is a consistent defect in the port-side data (apparent also in other locations) where the depth is underestimated at limiting ranges. This artifact is not introduced by the modeling implementation; it is present on the raw data tapes.

A preferred approach to modeling Sea MARC II events is to start with the raw phase data and process it three dimensionally. Phase error can be modeled as a probability distribution over a range surface bounded by the vertical beam pattern. However, the real potential for the Sea MARC II system comes from the joint availability of intensity and bathymetric information. With the bottom surface shape defined, the intensity map can be corrected geometrically. More sophisticated approaches are possible, as I discuss in the next chapter.

5.4 PROFILE SCANNING SONAR: MESOTECH

The most thorough modeling treatment to date was undertaken as part of a *USS Monitor* survey conducted by NOAA and the U.S. Navy [*Arnold et al.*, 1988]. A downward-looking, mechanically-scanned profiling sonar (Mesotech 971) was mounted on the Navy's *Deep Drone*—a free-swimming ROV—which was fitted with a good quality attitude measurement package (see Figure 5.15). In profile mode, the 675 kHz, 1.5°-beam-width sonar returned a thresholded range along with scan position. Most surveys were conducted using the 5-m range scale, and occasionally 10-m. A real-time processor collected data from the sonar, from an external long-baseline system, and from the attitude package. The measurements were filtered and buffered, then passed to a separate processor for modeling and display.

From statistics collected by the survey contractor, the positioning accuracy of the navigation system was estimated as 0.5 m (three standard deviations) with the vehicle in motion. In practice, though, there was acoustic shadowing by the wreck, and self-shadowing at the responder location on the vehicle. This resulted in frequent, long periods without navigation (up to 26 s), and overall accuracy was degraded. Depth measurements are from a pressure transducer with about a 5-cm resolution. To compensate for tidal variations (about 1 m), the vehicle was positioned on the bottom at

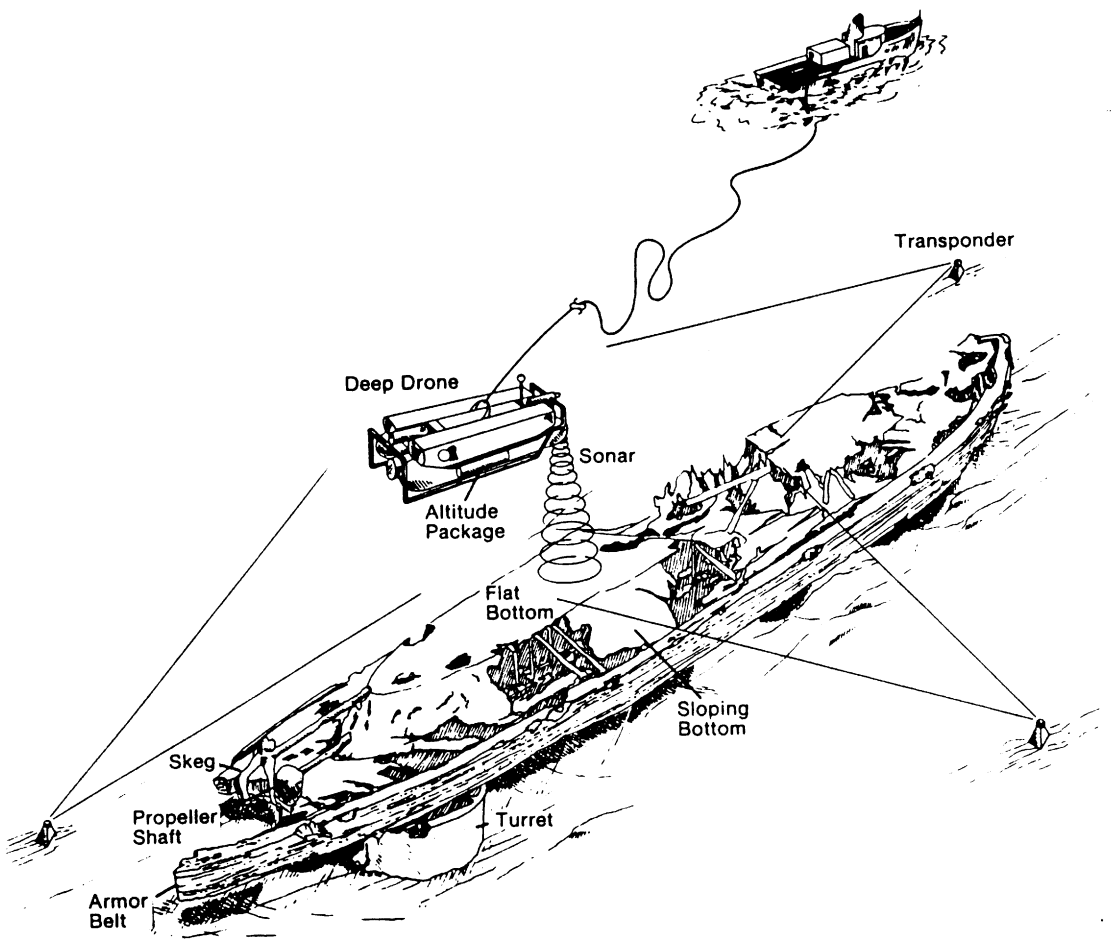

Figure 5.15: Deep Drone profiling geometry.

the same location every hour, and an average depth was calculated for a 1-min interval. Corrections are interpolated between readings and applied to the pressure measurement.

A discrete, steady-state Kalman filter was used to estimate position and attitude for real-time processing. Though smoothing can produce better results, all post-cruise modeling uses the same technique to simulate real-time performance. Since a model of the vehicle dynamics was unavailable, the simple filter uses a constant velocity assumption for all parameters. This is reasonable for heading and for vehicle translations since the survey consists mainly of straight-line tracks. However, for oscillatory motions in pitch and roll, there is some overshoot in the estimates.

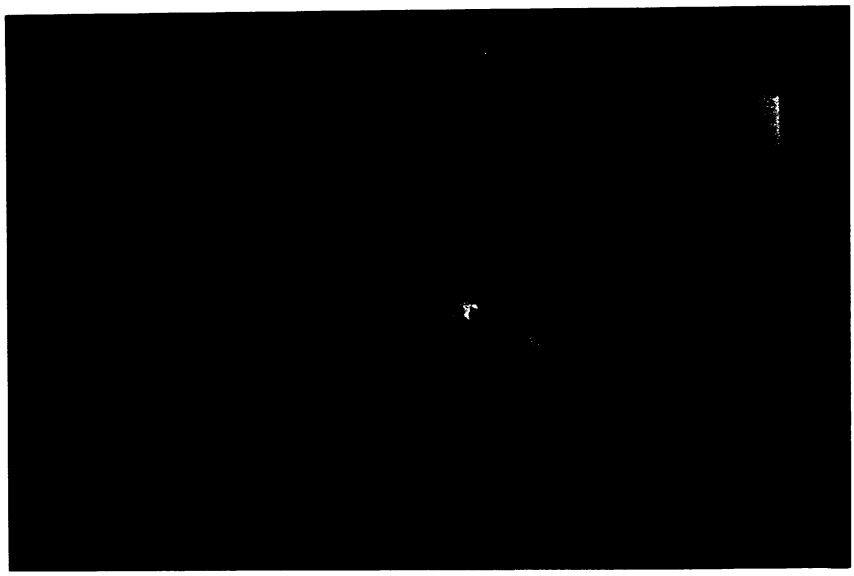

Figure 5.16: Depth map of *USS Monitor*: transverse tracks.

The binary formulation is the same as that given in Section 3.6.1, except that the beam pattern is modeled as the simple cone of Section 4.2. The modeling parameters are: $\Theta_b = 2.5°$, $\sigma_\Theta = 0.7°$, $\sigma_R = 2$ cm, $\sigma_x = \{20,20,8\}$ cm, $R_{max} = 10$ m, and the model size is $\{512,380,75\}$. In this model, the voxel edge size is about 14 cm for the survey area considered (70 m by 52 m). To produce a good representation of the wreck, estimates for horizontal positioning error, σ_{xy}, are relaxed to the values given above, so that $\Theta_c \cong 4.4°$.

Figure 5.16 shows a depth map from a model built with transverse survey lines, running diagonally northeast by southwest. The zigzag pattern is caused by the cross-track sonar scans as the vehicle moves along each line. A similar model, built from a longitudinal survey, is shown in Figure 5.17 (in these two figures $\sigma_x = \{70,70,8\}$ cm and $\Theta_c \cong 7.2°$; all other models use the lower values given above). A composite model, built from the transverse survey and two separate longitudinal surveys, is shown in Figure 5.18. This model is similar to the one produced aboard ship at the wreck site.

In Figure 5.19, the composite map is extrapolated with an iterative dilation algorithm (new pixels values are just the average of their neighbors). Noisy patches over and around the wreck are mainly caused by the abundant schools of fish that inhabit the "artificial reef" (swim bladders are good sonar reflectors). To produce the image in Figure 5.20, the extrapolated map is segmented into the wreck

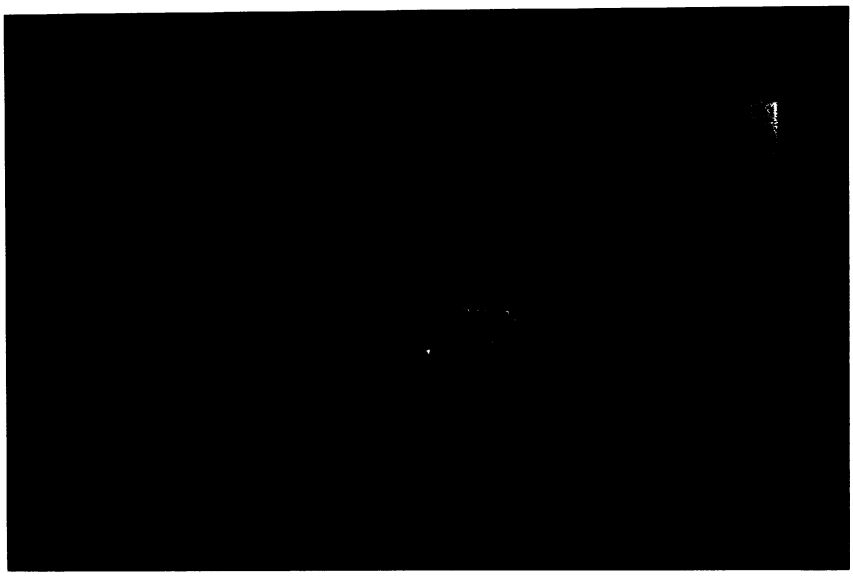

Figure 5.17: Depth map of *USS Monitor*: longitudinal tracks.

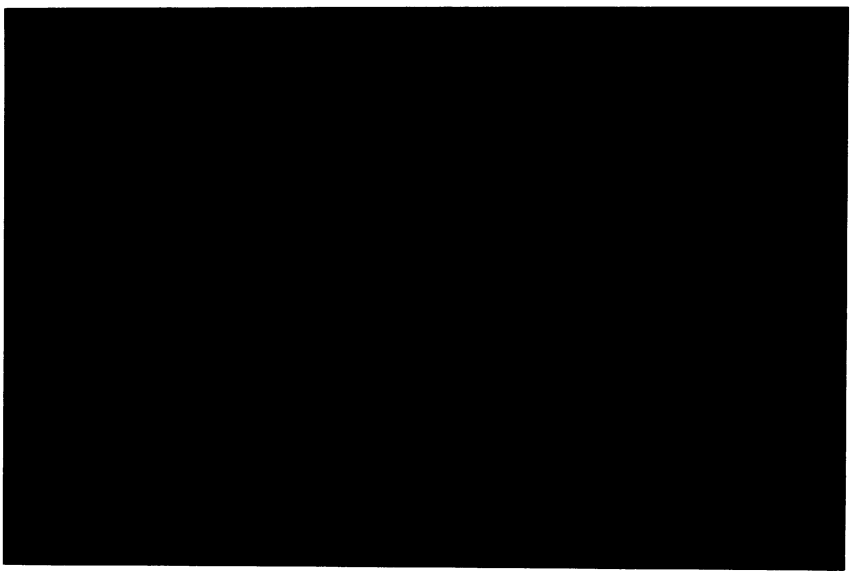

Figure 5.18: Depth map of *USS Monitor*: composite tracks.

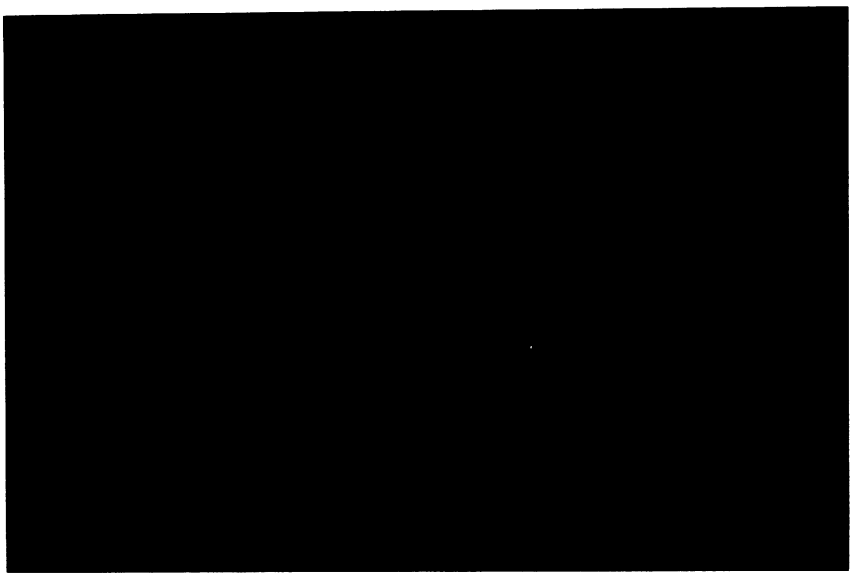

Figure 5.19: Depth map of *USS Monitor*: extrapolated.

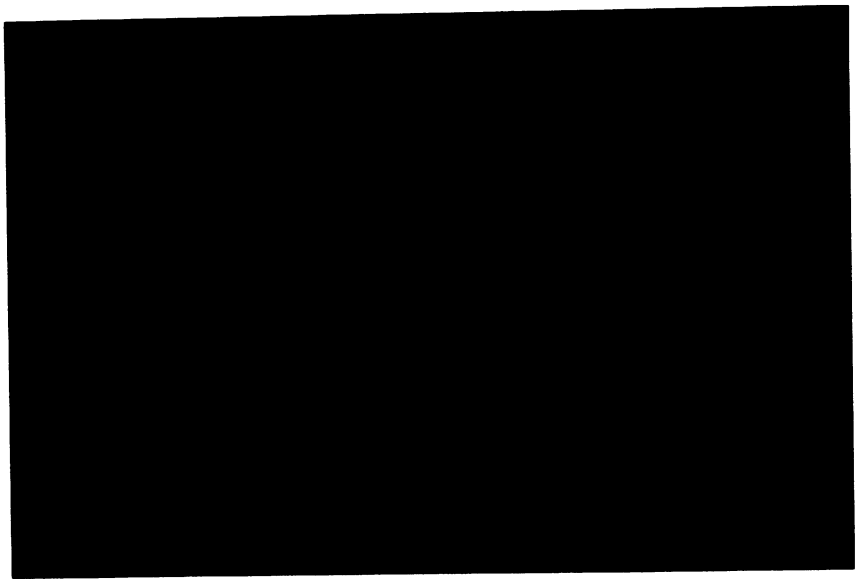

Figure 5.20: Depth map of *USS Monitor*: filtered and scaled.

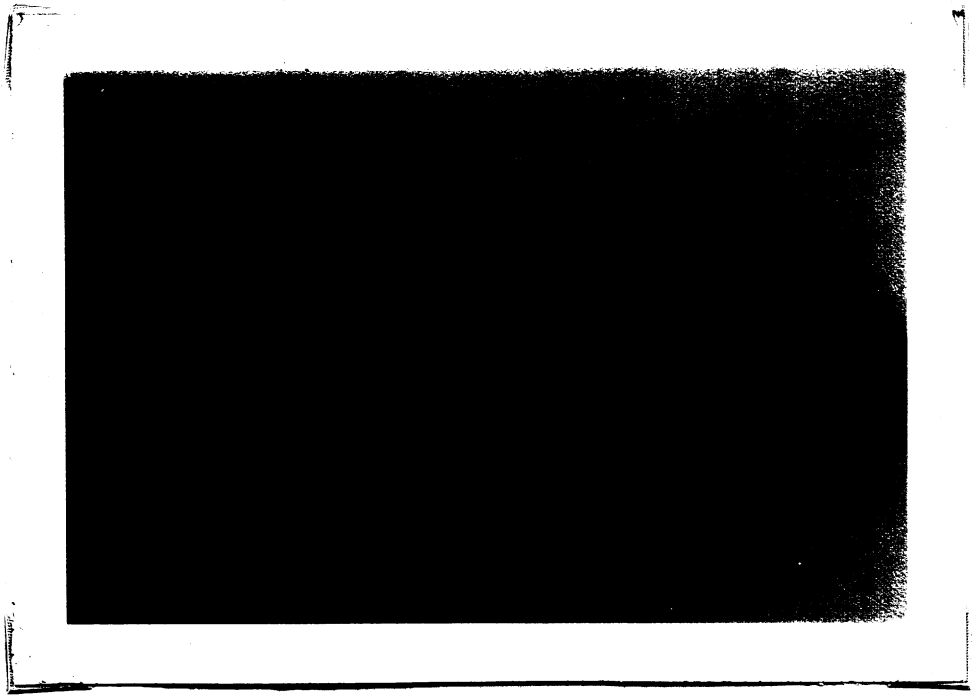

Figure 5.21: *USS Monitor*: Photo mosaic of wreck site.

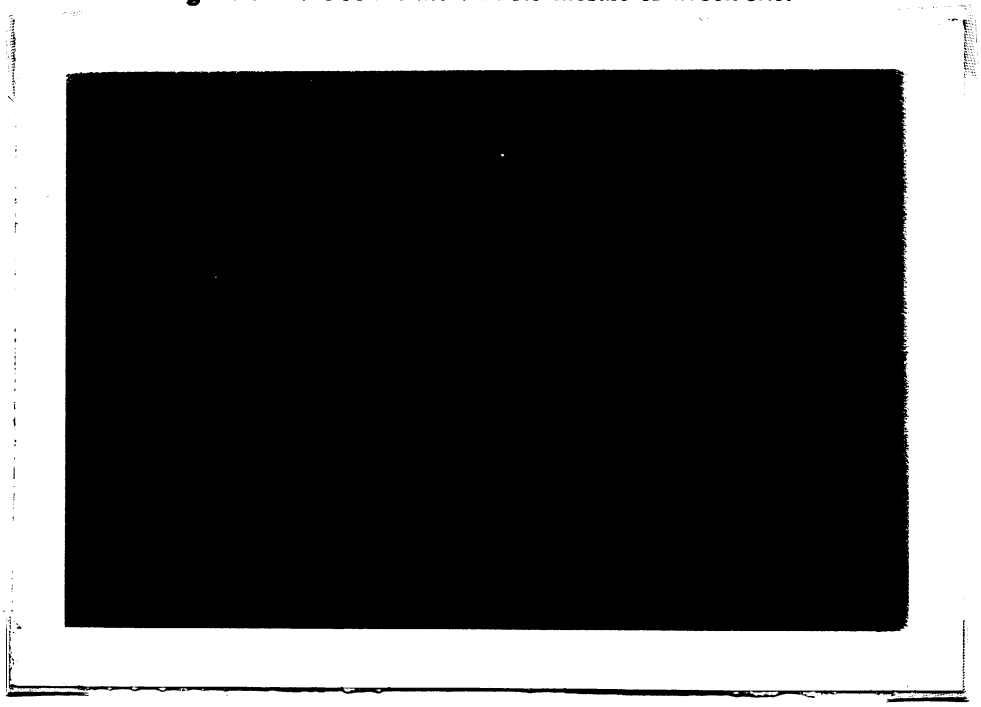

Figure 5.22: *USS Monitor*: Transverse section through turret.

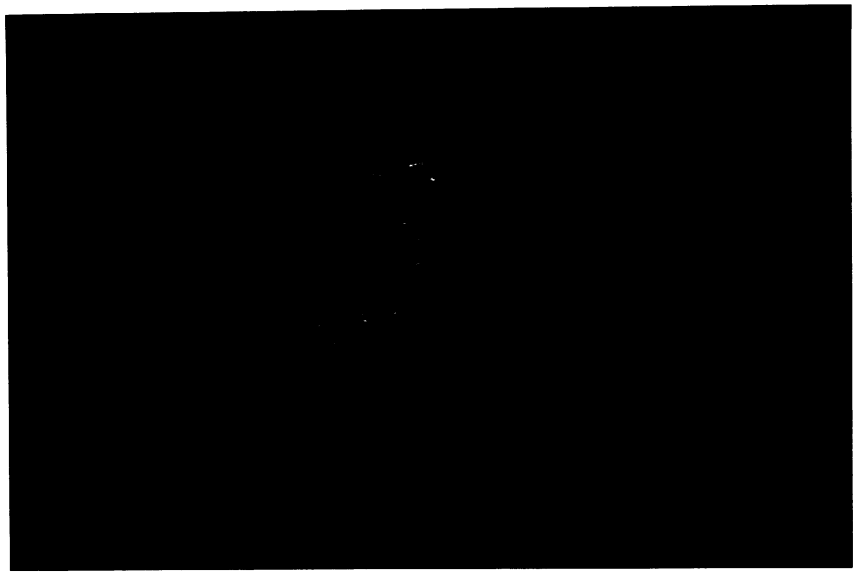

Figure 5.23: Perspective view of *USS Monitor* from the east.

and seafloor using a depth threshold; suspected fish noise is removed with localized median filters; and the two segments are filtered with a convolution mask before being remerged.

Unlike the modeling examples in previous sections, there is some "ground-truth" information for the *Monitor* survey (other comparisons are given in Section 6.6). Figure 5.22 shows a cross section through the *Monitor's* hull taken from the original drawings [*Peterkin*, 1985]. Notable features are the gun turret, the flat bottom, the sloping bottom, and the armor belt—a massive, 3-ft-thick structure around the periphery. Figure 5.21 is a photo mosaic of the wreck site from a 1974 survey (courtesy of NOAA, *USS Monitor* Marine Sanctuary Program).

The ship—capsized and sunk in a storm—is lying upside down with the broken stern quarter resting on the gun turret (hemisphere, upper right; the bow is to the west). From corrosion and probable depth charging during World War II, much of the structure is deteriorated. Seen in Figures 5.20 and 5.21, the band along the northern perimeter is the armor belt, which retains its shape though most parts of the hull have collapsed below it. The south side of the wreck is partly buried in the sand; prevailing currents are from the southwest, and an associated sand bar—with scouring around bow and stern—can be identified in the sonar maps.

The large, smooth patch in the aft section is a remaining portion of the vessel's flat bottom; pieces of the sloping bottom that connect the armor belt can also be seen. Just forward of midships are several patches in the wreck at the same level as the bottom. These can be seen in the depth maps, but are not identifiable in the photo mosaic. Details of about 20 cm are also recognizable in the sonar models, such as the skeg and propeller shaft (linear features, lower right). Though the poor navigation should preclude such detail, the redundant information in overlapping events enhances the resolution and accuracy of the model.

Figure 5.23 shows a perspective view from the east created from the model of Figure 5.20. This illustrates the drawbacks of the simple technique used to extract a depth map from the three-dimensional probability model. Though the armor belt appears to be contiguous to the bottom (right side of image), it is supported by the turret, and separated from the seafloor by about two meters. Similarly, the skeg and propeller shaft do not appear as distinct three-dimensional features, though cross sections through the probability distribution show that their shapes are preserved. I discuss this problem briefly in Chapter 7.

5.5 SUMMARY OF IMPORTANT POINTS

In this chapter I have shown that stochastic modeling can be applied to real sonar data sets with good results. Despite the limitations of noise and positioning inaccuracy, the incremental, backprojected models produce deterministic maps comparable to those created with traditional methods. Some important points that have been raised are:

- Though postprocessing techniques—which can invert all data—should produce the highest fidelity, incremental methods can approach that level of quality in many applications. Where real-time performance is needed, the tradeoff can be acceptable.

- Stochastic modeling is computationally tractable over different scales of range and resolution, and the performance and quality of results is appropriate to the applications considered in this chapter.

- Processing efficiency is such that the approach is realistic for practical, cost-effective field systems.

- Simple sensor models can produce acceptable results without the performance degradation of more sophisticated representations.

- In the context of stochastic modeling, large positioning error has the effect of a low-pass filter. For large-scale mapping applications, the limiting parameter is registration.

- In the production of seafloor intensity models, the use of sidescan amplitude data alone introduces geometric and radiometric distortion.

- For fully three-dimensional modeling, a better approach to feature estimation is needed.

Chapter 6

MULTISENSOR MODELING

The two previous chapters treat the modeling of a single environmental feature—surface shape or acoustic scattering intensity. These are *multidimensional* models in the sense that they are characterized by spatial distributions over two or three directions. If we consider model evolution, or allow for a dynamic environment, they might be categorized as four-dimensional representations. They are *multisensor* models because they comprise information from navigation systems and attitude measurements, and from a *primary* sensor, such as sonar.

In this chapter, I discuss more sophisticated models derived from two or more primary sensors. The first section presents some fundamental issues as a basis for elaboration. Following that is a brief overview of current techniques, where I begin to focus on a specific problem—using shape and intensity information to model surface scattering. The next three sections give a brief overview of seafloor backscattering in the undersea acoustic domain, and a discussion of its ramifications in sidescan sonar modeling. With that as background, I present an approach to combining Sea Beam bathymetry with Sea MARC I intensity data. The last section summarizes the important points raised in this chapter.

6.1 ISSUES

In the first chapter, I motivated the discussion of an approach to multisensor modeling with a general statement of the limitations of individual sensors. To be more specific, multiple sensors are needed mainly because: (1) many features must be characterized to more fully understand an environment; (2) different sensors have practical limitations—in range and resolution, for example; (3) similar features, detected with different sensing modalities, may be characterized more accurately with redundant information; and (4) sensors do not always measure distinct physical properties.

First, a simple characterization, such as by surface shape alone, may not suffice in many undertakings. An AUV may need more complete information about texture, reflectance, color, hardness, and so on to adequately discriminate among different objects or regions. Scientists also need more complete descriptions to understand subsea processes. For example, a study of seafloor spreading centers could be enhanced with information about temperature, optical transmissivity, and chemical constitution of the overlying water mass, as well as the shape and composition of the bottom surface (or subsurface). Such comprehensive description exceeds the capabilities of single-sensor surveys.

Second, there is often a tradeoff among different sensory capabilities. Video or still photography can supply detailed information about the seafloor, often enough to visually study the structure and distribution of geological or biological features. For fine-scale topography, a scanning laser can survey more quickly and at a higher resolution than a scanning sonar. For longer ranges, acoustic methods provide the only real alternative. Here also, different sonars offer competing options in range, resolution, speed, and signal analysis (sidescan intensity versus bathymetric profile, for example).

Third, in much of this thesis I stress the importance of redundant information in enhancing the accuracy and certainty of a model. Redundancy can also be derived from different sensing modalities. To determine surface shape, for example, data may be integrated from a sonar, a laser rangefinder, stereo photography, or even the touch of a manipulator. However, there is a danger here of confusing apples with oranges. For example, consider a rocky bottom with a layer of soft mud, all overgrown with a thick grass or algae. A laser or camera might detect the outer surface of the plants; to a high-frequency sonar, these would be transparent, but the sound might not penetrate deeply into the mud; depending on the sensitivity of its touch sensor, a probe on the arm could pass through to the rock before anything is detected. In such a scenario, the optical, acoustical, and tactile surfaces do not coincide; a single surface representation would be inaccurate and misleading.

Fourth, a camera is a good example of a device that can characterize surface features only indirectly. The optical intensity measured at the sensor is a function of the surface reflectance, of the surface normals with respect to the camera, and of the lighting power and geometry. Even with a perfect sensor model, it is impossible to deduce any one of these parameters without prior knowledge or simplifying assumptions about the others. To extract three-dimensional features from two-dimensional images, shape-from-shading techniques [*Ballard and Brown*, 1982] require a known camera and lighting geometry, and assume a uniform surface albedo (or impose it with a can of spray paint). Even then, lateral dimensions are only specified relative to an unknown distance along the look axis. When complemented with laser range information, though, more complete and more accurate information can be acquired from the two sensors, and with less effort than for an exhaustive analysis of the camera data alone.

The first three issues have been addressed earlier. Where the feature information is distinct and complementary (apples and oranges), it can be maintained as a stochastic feature vector, with different features individually or jointly accessible to higher-level processes (Chapter 7 has other discussion of multidimensional representations). Redundant multisensory information (small apples and large apples) can be combined in the same model if attention is paid to accuracy and relative certainty; as I have shown with computer simulations and real data sets, the modeling approach developed in this thesis is largely independent of scale and resolution.

In the example above, where the detected surfaces may be inconsistent (red and green apples?), the best approach is less easily defined. The real issue is: at what level of representation should the different data be combined? The answer is: it depends on the particular sensors, environment, and application. For most cases, though, I contend that the preferred approach is to combine distinct sensor data in separate low-level components of a model. The basis for this statement—process *then* threshold—has already been elaborated and supported by earlier examples. For the optical, acoustic, and tactile features, an agreement of surface estimates is self-supporting; a disagreement should be resolved with guidance from a higher level, perhaps by a call for more sensor data. Also, by maintaining separate low-level model components, the environment can be characterized by multisensor signatures.

The last issue is more interesting. The camera intensities cannot be merged directly into a simple, data-level model if they are taken from different positions or look directions—the camera does not measure distinct *physical* properties that can be represented in a *three-dimensional* spatial distribution. The same is true of the sidescan examples in the last chapter. In fact, all the sonar models I've described so far are based on the original assumption of directional independence—uniform scattering in all directions.

For the binary models, where a scattering threshold is used to infer only the probable presence of some target, such an assumption is useful; it reduces the computational load and results in reasonable surface estimates. For continuous models, if the look directions of neighboring events are similar (individual sidescan tracks), results are also useful, though the model is view-dependent. However, where there is a significant variability in sensor attitude, the assumption breaks down. Such is the case for overlapping sidescan tracks; a simple data-level combination of sidescan events from opposing look directions tends to produce a mid-intensity blur.

In the context of the numerical modeling approach developed here, there are two possible solutions. First, the model could be expanded as a function of look direction as well as position. In other words, rather than a three-dimensional representation, a five-dimensional model could aggregate events according to the incident angle of energy and location in 3-space. If a separate transmitter and

receiver are considered along with the bidirectional scattering properties of real surfaces, then a seven-dimensional representation is needed. Though some simplification might be made, memory and processing requirements go beyond the practical bounds of today's computational technology.

The second approach is to make use of information from other sensors that measure important parameters directly. For cameras and sidescan sonars, the information needed to build a model of reflectance or scattering properties is surface shape. Such information is sometimes available from a single sensor package—a laser that measures range and intensity [*Nitzan et al.*, 1977], or a sidescan that measures phase and amplitude (Sea MARC II). Separate sensors can also be used, leaving open many more avenues to multisensor approaches. In both cases, though, the main advantages are that complementary sensors can be used to improve modeling accuracy and efficiency, and to derive "new" information that is, in a sense, greater than the sum of its parts.

6.2 CURRENT TECHNIQUES

A growing interest in multisensor techniques is creating a body of literature too broad to review here. Extensive references and a summary of a workshop on multisensor integration for manufacturing automation are given by *Henderson et al.* [1987]. A collection of recent papers from a similar workshop on spatial reasoning and multisensor fusion can be found in *Kak and Chen* [1987]. Other references to relevant techniques are given in Chapter 2. Most researchers approach sensor integration from, at least, the physical level, and mainly relegate the combination to higher-level processes.

This rest of this chapter treats physical-level modeling by mapping from data-level events with multisensor information. In the first part of this section I briefly discuss three representative terrestrial applications that combine optical intensity with shape information, since the problem is similar to sonar modeling underwater. The second part gives an overview of techniques used by the underwater community for large-scale acoustic intensity mapping.

6.2.1 Optical Intensity Modeling

An early investigation of combining optical reflectance and range data is that of *Nitzan et al.* [1977]. For their experiments, the authors use a calibrated laser that returns range and amplitude for each scan position. The range data are used to correct optical intensities for inverse square-law losses, resulting in a value that is the product of the cosine of incident angle and the diffuse surface reflectance (except at near-normal incidences where specular reflection dominates). The range-image jump boundaries define the occluding contours of planar surfaces in the test scenes. After segmenting these

planar objects, the incident angle for each range-corrected intensity datum is used to calculate the coefficient of diffuse reflection. The resulting model of the scene is view- and lighting-independent, and represents derived physical properties.

Liebowitz and Casasent [1987] describe a more complex approach intended for object recognition on DARPA's *Autonomous Land Vehicle*. Two dimensional range images are segmented into background/foreground pixels by scan-line thresholds. Different algorithms are then used to segregate foreground pixels into *Regions Of Interest* (ROI's). All operations are carried out in the two-dimensional plane for convenience, but may be transformed to rectangular coordinates after processing. The authors focus on range-data processing, but discuss how the approach is to be implemented with complementary color and intensity data for resolving the ROI-ambiguity of contiguous objects.

Whiteside et al. [1987] describe an application of mapping aerial photography onto a three-dimensional surface derived from a separate digital terrain-elevation database. Each point in a photo mosaic is projected onto the surface with a perspective transformation, which provides a basis for "hidden-pixel" removal. From the composite scene, stereo pairs or animated sequences of simulated overflights are generated. Though the horizontal resolution of the elevation data is much less than that of the photographs, the higher-frequency camera data combined with the coarse surface description provides a visual realism better than either representation alone (an animated video tape may be obtained from the authors' company).

6.2.2 Acoustic Intensity Modeling

Sidescan sonars are the most common sensor for acoustic intensity mapping underwater. In most applications, a fixed TVG and a planar-bottom assumption are applied, as discussed earlier. The resultant two-dimensional data sets are then treated as two-dimensional sidescan "images." These are sometimes enhanced with image-processing techniques, or merged to form sidescan mosaics.

Perhaps the most sophisticated sidescan processing system is MIPS (Mini Image Processing System), developed by the U.S. Geological Survey [*Chavez*, 1986; *Chavez et al.*, 1987]. MIPS is mainly devoted to processing data from GLORIA, the long-range sidescan sonar being used to map the Exclusive Economic Zone, but has also been applied to Sea MARC II data [*Twitchell*, personal communication]. In operation, the water column is removed from linear sidescan image segments and a slant-range correction is calculated. Pixel interpolation compensates for aspect-ratio differences along and across the track, line averaging is used to correct for velocity variations, and a radiometric calculation normalizes the intensities across the track. Two filtering steps compensate for speckle and striping noise, and the images are enhanced with sharpening filters and contrast stretching. A two-

dimensional translation and rotation then registers the images for mosaicking. Manually-entered control points are used for two-dimensional warping as a final registration adjustment of overlapping segments. Some work is also being done in extracting stereo pairs [*Chavez et al.*, 1987] and in the projection of GLORIA images onto bathymetric surfaces [*Twitchell*, 1988].

Farre and Ryan [1985, 1987] have combined Sea MARC I and Sea Beam data sets. In separate preprocessing steps, contoured bathymetric data are interpolated and gridded to match the resolution of geometrically-corrected (for position and heading) digital sidescan images. Because the sidescan data are slant-range corrected with the planar-bottom assumption, the combination is a kind of orthographic projection, which does not take advantage of the bathymetry to remove relief-dependent distortion in the sidescan data. To generate stereo pairs, each sidescan pixel is shifted as a function of the associated depth value to generate left- and right-eye views. Though no intensity corrections are used to compensate for bottom-slope effects, the method of presentation allows a geologist to more readily make the association between bathymetry and intensity variations.

De Moustier [1986] investigates quantitative assessment of bottom backscattering with Sea Beam. Using an ensemble of successive pings, the first-order statistics of the demodulated signals from near-specular beams form an envelope pdf. These are fitted to a Rician distribution, whose shape is related to signal coherence, and an estimate of surface roughness is derived (see also *Stanton*, 1983). In analyzing data from three test areas of different bottom types (with supporting ground-truth information), good comparative results are demonstrated. To complement this information, the angular dependence of different beams is considered for the test areas, but results are less compelling. Later work [*de Moustier*, 1988] uses the signal amplitudes from different beams in a sidescan-like presentation, which provides a qualitative indication of roughness within each beam footprint.

6.3 SEAFLOOR BACKSCATTERING

In developing the backprojection approach in Chapter 3, I assume the environment comprises point targets that scatter uniformly in all directions. The examples of the previous chapters show this assumption gives good results under most circumstances. However, the scattering function of real surfaces is more complex and has a strong directional dependence. Also, I have used the term *scattering* to describe a process that comprises reflection (coherent) and scattering (incoherent) components. A full treatment of the subject is beyond the scope of this thesis, but more complete discussion and references may be found in *Horton* [1959]; *Urick* [1975, 1979]; *Clay and Medwin*, [1977]; and *Ogilvy*, [1987].

Backscattering from the ocean floor is particularly complex. The bottom is often layered with materials that have different acoustic properties and, because the sound can penetrate below the surface where it is reflected, scattered, or refracted and reradiated, the received waveform is a composite of returned signals from different scatterers. The seafloor also shows much variation over even small regions, vertically and horizontally; it may comprise different materials—mud, sand, or rock, for example—with different absorption and reflectance properties, and a range of surface roughnesses as well. Angular and frequency variations are highly variable with bottom type.

Probably the first reported work on scattering from rough surfaces was by Rayleigh [*Ogilvy*, 1987], who discussed the effects of normally incident sound on corrugated surfaces separating two acoustic media. Assuming a locally smooth and planar surface, the ratio of incident to reflected intensity, μ, is described by the Rayleigh formula [*Urick*, 1975] as:

$$\mu_1 = \left[\frac{m \sin(\phi) - (n^2 - \cos^2(\phi))^{1/2}}{m \sin(\phi) - (n^2 + \cos^2(\phi))^{1/2}}\right]^2$$

$$\mu_2 = 1 - \mu_1$$

where: $m = \rho_1/\rho_2$

$n = c_1/c_2$

and: ρ = material density

c = speed of sound

ϕ = grazing angle

The subscript *1* identifies the material in which the incident and reflected rays are traveling, and *2* refers to the reflective material, into which part of the energy (μ_2) is transmitted.

In considering roughness, Rayleigh assumed that the scattering could be written as the sum of plane waves travelling away from the surface. A measure of surface roughness is given by the Rayleigh parameter, R, as [*Urick*, 1979]:

$$R = 2kh \sin(\phi)$$

where: k = acoustic wave number $:= 2\pi/\lambda$

λ = acoustic wave length

h = rms roughness height

for scattering in a two-dimensional plane. The Rayleigh parameter is found empirically to be a good indicator of roughness [*Urick*, 1979]. When $R \ll 1$ the surface is considered *smooth*; when $R \gg 1$ the surface is *rough*. Note that the characterization is frequency dependent, since scattering varies with the relative sizes of acoustic wavelength and surface irregularities. The intensity ratio of returned sound to incident sound for a rough surface may be expressed as [*Urick*, 1979]:

$$\mu_R(\phi) = \mu_s(\phi)e^{-R^2(\phi)}$$

where μ_s is the reflection coefficient that would exist if the surface were smooth. The approach has also been extended by other authors to non-normal incidence and to random rough surfaces [*Ogilvy*, 1987].

The perturbation method (also known as the Rayleigh-Rice method) models the rough surface as a smooth plane that deviates locally due to irregularities. The approach calculates the deviation from the planar scattering coefficient caused by these roughness perturbations. The constraints imposed by the method are that roughness heights are everywhere small compared with the wavelength of the incident wave, and that the surface gradient is small compared with unity. The use of this method has been limited because of these strict conditions, and because the range of actual surfaces to which the method applies are just those for which the effects of roughness are small [*Ogilvy*, 1987].

The most prevalent approach to scattering [*Ogilvy*, 1987] is the tangent-plane method, or Kirchhoff method, first applied to acoustic wave scattering by *Eckart* [1953]. The method gives an approximation to the scattered field on the surface, in terms of the incident field and plane-wave reflection coefficients. The scattering surface is assumed to be everywhere smooth enough that the reflection properties at each point on the surface can be modeled by a plane parallel to the local tangent. In most work, the reflection coefficient over the surface is assumed constant, usually unity [*Ogilvy*, 1987]. The method requires no constraints on the magnitude or gradient of local surface height, but restricts the rate of change of the gradient (radius of curvature). Surfaces for which this restriction cannot be satisfied may be considered using the Rayleigh method [*Ogilvy*, 1987].

Both the conventional Kirchhoff method and the perturbation approach assume that all points on the surface are ensonified by the incoming wave and that multiple scattering can be neglected. Particularly at low grazing angles these assumptions are invalid, and various attempts to incorporate the effects in a theoretical framework have been attempted [*Ogilvy*, 1987]. Other works consider the effects of different scales of roughness, generally using the Kirchhoff method for the large-scale roughness and perturbation theory for the small scale [*Bass and Fuks*, 1979; *Ogilvy*, 1987]. Other such composite models are extended to account for volume scattering, which may dominate at low frequencies or at intermediate grazing angles in soft-sediment areas [*Jackson et al.*, 1986].

Experimental results of backscattering and bistatic measurements are profuse [*Horton*,1959; *Urick*, 1975, 1979; *Clay and Medwin*, 1977; *Ogilvy*, 1987] and show that the received signal can be regarded as comprising a coherent component, strong near specular angles, and a diffuse component of more constant intensity over other scattered angles. The relative strengths of these components (and the shape of the scattering function) depends strongly on the Rayleigh parameter [*Ogilvy*, 1987]. Ogilvy concludes that the different theories are valid over certain regimes of roughness only, that no complete treatment of the problem exists, and that this is likely to remain the case, as attempts to enhance the validity of any theory necessarily introduce greater mathematical complexity.

Stanton and Clay [1986] offer a practical perspective on the issues concerning sonar as a remote-sensing tool. For classification of the seafloor (and marine organisms in the water column) there is an engineering tradeoff between (1) high-resolution systems and direct analytical techniques, and (2) lower-resolution systems and indirect analytical techniques. Very high resolution is required for adequate classification; when such resolution is not available, indirect or "inverse" methods are employed. With low-resolution systems, high-volume coverage is obtained with a single beam; but interpretation of the data may be involved and indirect. High-resolution sonars give a more direct mapping of the environment but, because of low-volume coverage, multibeams are usually required and data rates are high. According to the authors, the engineering challenge is to process and display the data in a useful form, and it is crucial to produce meaningful, quantitative images in real time so that sampling decisions can be made.

6.4 THE SIDESCAN PROBLEM

Sidescan sonars depend on variations in bottom backscatter, as a function of material properties and surface shape, to produce an intensity map of the seafloor. Because of this joint dependency, though, it is impossible to distinguish between the effects of large-scale (relative to acoustic wavelength) surface relief and surface scattering due to roughness and material properties. In theory, a perfectly smooth bottom with a spatial distribution of different materials can return a sonar signal identical with one from a surface that has uniform scattering properties but an appropriate relief.

Acoustic shadows also offer a roughly quantitative measure of relief where the bottom is mostly smooth and level (and altitude is known), and a qualitative measure where the planar-bottom assumption does not hold. However, shadowing and intensity variations can introduce into two-dimensional sidescan images an artificial structure that does not correspond to the physical structure. This can make correct interpretation difficult, if not impossible, for even a trained human expert.

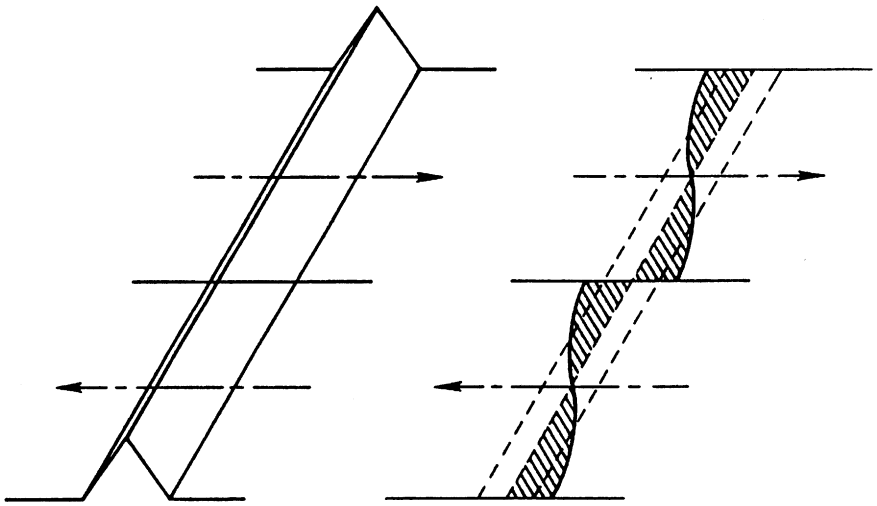

Figure 6.1: Sidescan shadow transitions.

Figure 6.2: Sea MARC II: sidescan map showing intensity transitions.

Figure 6.1 represents a sidescan sonar moving along tracks indicated by the arrows. A regular, triangular feature with positive relief lies obliquely across the track, as seen in Figure 6.1a. In Figure

Figure 6.3: Sea MARC II: relief map for intensity transitions.

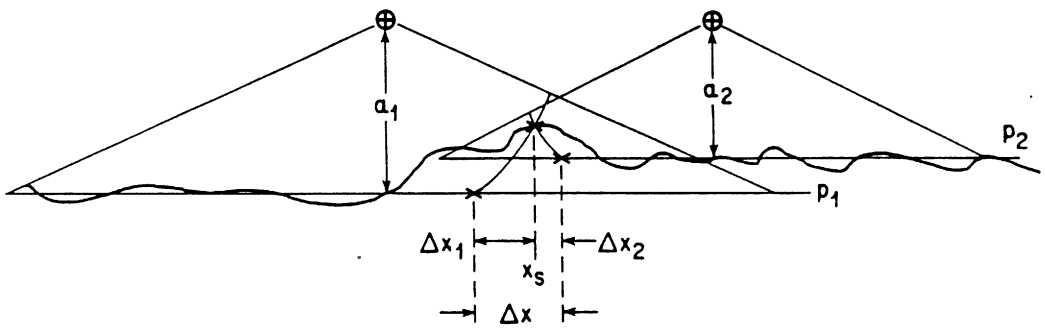

Figure 6.4: Sidescan geometry artifacts.

6.1b. the approximate pattern of acoustic shadows that would be generated for such a case is shown schematically. As the sonar nears the feature, the shadow-length decreases and goes to zero as the sensor flies overhead; as the sonar passes and draws away, the shadow transitions to the opposite side and lengthens smoothly. At the juncture between two swaths, there is a step change in the shadow position caused by the opposing look angles.

In reality, the intensity variations of the ensonified areas are complicated by the nonlinear dependence on grazing angle and by the more irregular three-dimensional geometry. Figure 6.2 shows several such S-shaped artifacts associated with the deeper rifts as well as the ridges seen in Figure 6.3. The effect is analogous over negative-relief areas as the brighter intensities transition from one wall to the other. At the overlaps between tracks, juxtapositions of high and low intensities are evident (see also Figures 5.12 and 5.14).

In Figure 6.3, the curvature of fractures (trending north-northwest) between the transform segments (trending east-northeast) is seen in the bathymetry as a true characteristic of the fault region caused by the plates sliding past each other as they separate [*Fornari and Gallo*, personal communication]. However, the artificial structure imposed by the intensity transitions is not readily distinguished from reality using the evidence in Figure 6.2 alone; and many such artifacts are present at a smaller, more subtle scale. For a geologist trained in sidescan interpretation, such analysis may be possible; for a machine intelligence at today's level of development, the view-dependent intensity representation is intractable.

In typical sidescan processing, geometric distortions are also caused by slant-range corrections based on the planar-bottom assumption. Figure 6.4 is a profile view of two sonar tracks normal to the page (crossed circles represent the sensor location). The two tracks are flown at the same depth but, because of the variation in altitudes (a_1 and a_2), the assumed planes (p_1 and p_2) do not correspond. A point on the surface at x_s, appearing in the two sidescan images, is projected onto each plane at the intersection with the corresponding range arc. For the first pass, the lateral distortion is indicated as Δx_1, and for the second as Δx_2. If the two-dimensional images are mosaicked according to the sensor track, the total misregistration of the point is Δx.

The total horizontal distortion shown in this figure is significant—about 25% of maximum range for each side—but not unreasonable for high-relief terrain. In a real mosaic, this geometric distortion would be compounded by the superposition of shadows and high-intensity regions of images generated from opposing look angles at x_s. Even if a human operator could correctly identify x_s at corresponding control points in the separate images, and if a two-dimensional warp were applied to bring the points into registration, the intensity differences would still exist; and other points in the images could suffer more geometric distortion from the warp itself. For automatic processing, the correspondence problem is nontrivial and computationally expensive; it is further complicated by the indistinct features of natural underwater terrain. Regardless of whether it can be accomplished, the same problems exist as for the human-supervised process.

6.5 LAMBERTIAN SCATTERING

To establish the true scattering geometry and to help resolve ambiguity in a sidescan intensity signal, bottom-shape information is needed. With position and attitude information, geometric correction is straightforward; intensity manipulations are more complex. Because of the wide variation in bottom characteristics, a comprehensive approach to scattering estimation must be treated as spatially nonstationary. Not only does the acoustic albedo vary with material type, but the angular dependence—the shape of a normalized scattering function—also depends on bottom characteristics (and on sonar frequency).

The function called Lambert's law is a type of angular variation for rough surfaces often satisfied in both optical [*Ballard and Brown*, 1982] and acoustic [*Urick*, 1975] domains. If we consider a small surface, dA, then the power intercepted by that area is $I_i\sin(\phi)dA$, where I_i and ϕ are the incident intensity and grazing angle, as before. By Lambert's law, this power is assumed to be scattered in proportion to the sine of the grazing angle. Then the intensity at a unit distance in the direction ψ is:

$$I_L = \mu_L I_i \sin(\phi)\sin(\psi)dA$$

where μ_L is a proportionality constant, or surface albedo. Then for a unit surface area and scattering in the backward direction, for which $\psi=\pi-\phi$, the intensity ratio is:

$$\mu = \mu_L \sin^2(\phi)$$

In other words, the angular dependence exhibited by such Lambertian surfaces can be modeled as the sine-squared of the grazing angle for a monostatic active sensor.

Though no surface satisfies Lambert's law exactly, very rough surfaces fit the model reasonably well. *Urick* [1975] reports that Lambert's law appears to be a good scattering description for many deep-water bottoms. At grazing angles less than 45°, there is little frequency dependence in the range 1-30 kHz. However, Urick cautions that Lambertian scattering probably applies only to *type-III* bottoms—those characterized by heavy dissection and underwater ridges. For a *type-I* bottom—abyssal plains with little roughness—angular and frequency variations are high. A *type-II* bottom, termed "hill regions," shows an intermediate behavior with angle and frequency.

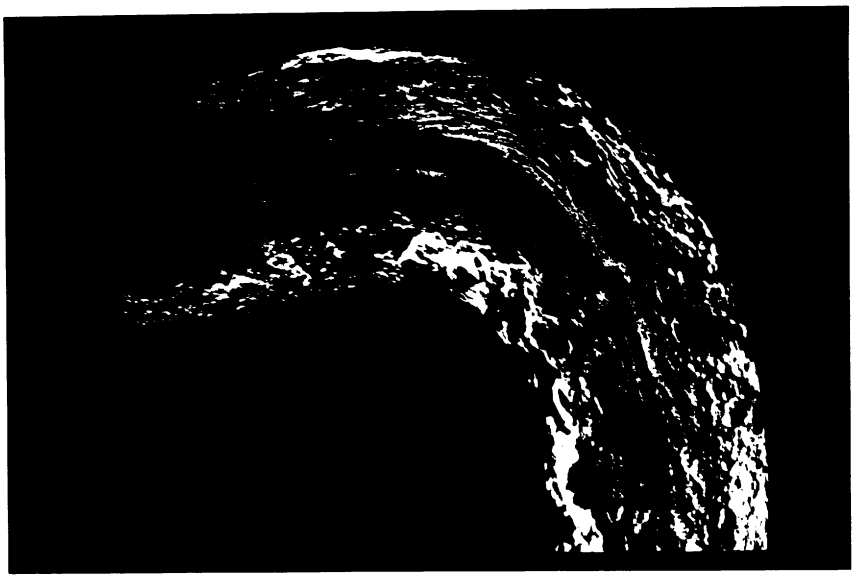

Figure 6.5: Sea MARC I intensity map of Clipperton area.

6.6 SEA MARC/SEA BEAM MODEL

In this section I describe an approach to modeling Sea MARC I intensity data where a Sea Beam bathymetric model provides the basis for correcting the geometry, and for partially removing the view-dependent effects of bottom shape. The data come from the Clipperton Transform Zone—a site of crustal spreading, fracturing, and volcanic construction—which offers a likely fit to Urick's *type-III* bottom characterization [*Gallo et al.*, 1986; *Kastens et al.*, 1986]. Then for a sidescan sonar operating in the appropriate frequency range (1-30 kHz), the scattering properties of the region can be reasonably characterized as Lambertian. Since the Sea MARC I operates at 27/30 kHz, such a scattering model is used in the discussion that follows. The simple Lambertian characterization is not required for this development, and a more accurate scattering model (perhaps derived from local data) could be substituted directly.

Figure 6.5 shows a Sea MARC I sidescan intensity map of the Clipperton area where the Sea MARC/Sea Beam model is constructed. The data set is the same as in Section 5.2, where only the first part of the track had been processed at a higher resolution. The modeling is identical, except that a fixed cross-track normalization gain is computed from a preliminary pass through the data (as for Sea MARC II processing in Section 5.3). For comparison, an intensity map from the same data set—

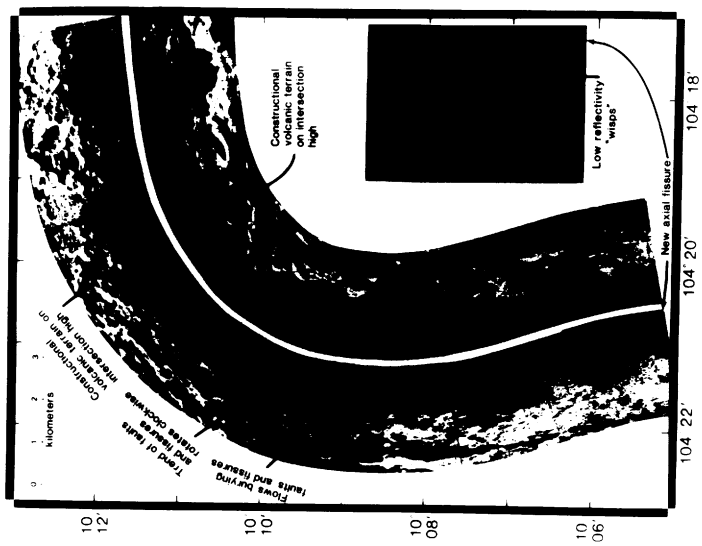

Figure 6.6: Sea MARC I intensity map of Clipperton area (from *Kastens et al.*).

processed at Lamont-Doherty Geological Observatory—is shown in Figure 6.6; the photograph is from *Kastens et al.* [1986].

Figures 6.7-6.9 show the bathymetry from the same region of the Clipperton Zone. The data used for stochastic modeling were supplied by NECOR at the University of Rhode Island, by courtesy of Dr. Dave Gallo of the Woods Hole Oceanographic Institution. The full data set comprises several surveys conducted over a 3-year period. Unlike the Sea Beam data described in Section 5.1, the navigation had been corrected with track correlations, and the depth/offset pairs had been averaged over five pings. The processing is as described in Section 5.1, except that the modeling functions of Section 3.6.1 are used:

$$f_M = \{x, f_S(\rho_k, S_{k-1}), f_H(\sigma_\rho, \sigma_{S_{k-1}})\}$$

$$f_H = \sigma_\rho + \sigma_{S_{k-1}}$$

$$f_S = (\rho_k \sigma_\rho + S_{k-1}) / \sigma_{S_k}$$

Because the 5-ping averaging produces a staircase effect in the model, a convolution mask is applied to the raw depth map to smooth out the artificial texture; the low-pass filtered depth map is

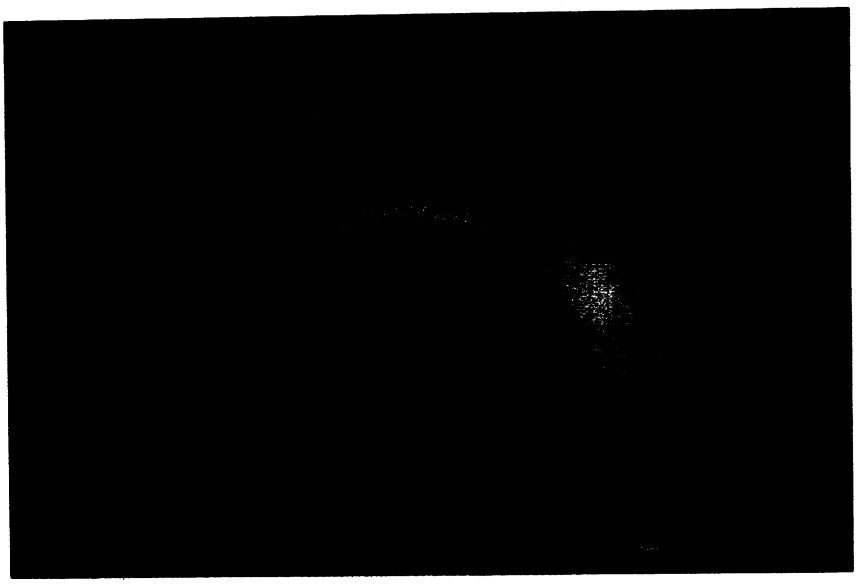

Figure 6.7: Sea Beam contour map of Clipperton area.

shown in Figure 6.7. For comparison, a contour map from *Gallo et al.* [1986] is shown in Figure 6.8, with the slightly smaller modeling area delineated. A perspective view of the area is shown in Figure 6.9 (no vertical exaggeration).

Because of positioning error and the large Sea Beam footprint at these depths (up to 160 m), the bathymetric maps cannot give the detail seen in the sidescan images. The two data sets are complementary: Sea Beam has good vertical resolution but poor horizontal resolution; Sea MARC I is incapable of resolving depth but furnishes high-frequency information horizontally. For geometric correction of sidescan data, the Sea Beam model is adequate. In the area considered here, for an assumed altitude of 200 m, at a maximum Sea MARC I horizontal range of 2500 m, looking down the slope where relief varies up to about 1000 m, the lateral distortion introduced by a planar-bottom assumption is greater than 300 m, or about 12% of maximum range. Small errors from the low-pass Sea Beam model do not affect the correction significantly.

However, the match is not ideal for mapping intensity data to a physical scattering distribution. The Sea Beam surface is too smooth to estimate surface normals that match the high-resolution Sea MARC I data. Even if this were not so, registration errors between the two data sets preclude such fine-scale correction. The approach I take uses the Sea Beam surface to compensate for gross variations in intensity.

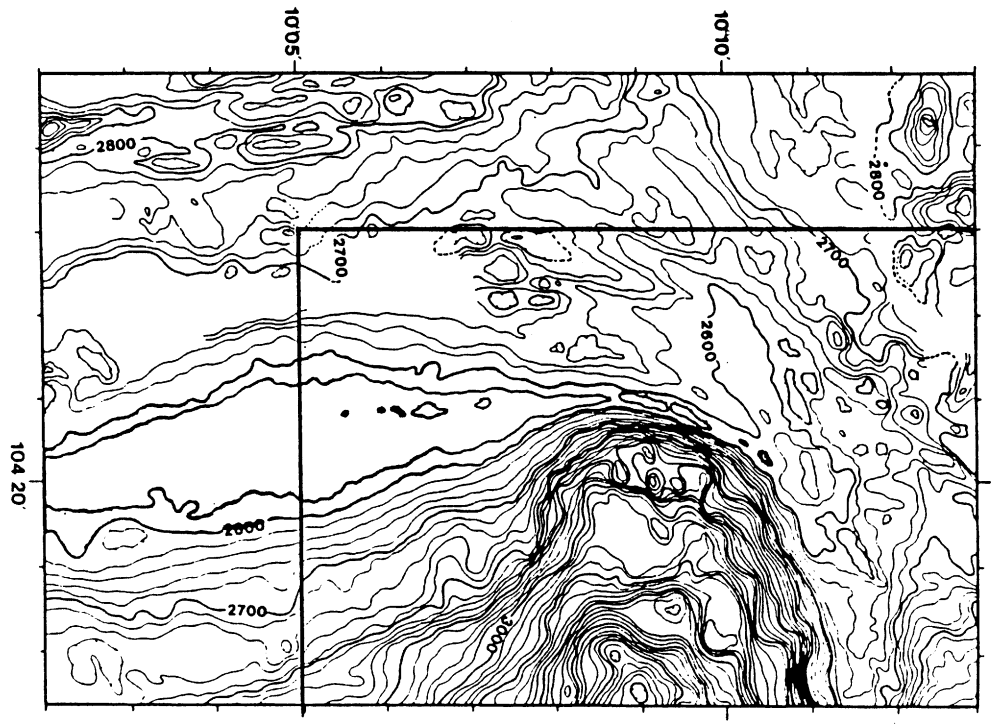

Figure 6.8: Sea Beam contour map of Clipperton area (from *Gallo et al.*).

If we assume a flat bottom and consider the digitized Sea MARC I signal for a specific range bin (indicated by the subscript *j*), the received signal may be expressed as:

$$I_j = I_0 \iint_A \upsilon(r) b^2(\theta) \mu(r,\theta) \sin^2(\phi) dA$$

where: I_0 = transmitted intensity

υ = transmission loss (absorption and spreading)

b = beam pattern

The expression is similar to the one described above for Lambertian scattering, except that the intensity losses are modeled for propagation through the medium and for the beam pattern.

The area of the bottom intersected is a range annulus defined by the beam pattern and the sampling interval, which corresponds to a fixed horizontal distance, Δr, because of the Sea MARC sampling scheme. Since the sampling increment is small, any variation with range over Δr is negligible. With these assumptions, the expression can be rewritten as:

$$I_j = I_0 \upsilon_j \sin^2(\phi) \Delta r \int_\theta b^2(\theta) \mu(\theta) d\theta$$

Figure 6.9: Sea Beam perspective view of Clipperton area.

The remaining integral can be considered as a weighted sampling of the acoustic scattering albedo over the width of the sonar beam, or:

$$\int_{\Theta} b^2(\Theta)\mu(\Theta)d\Theta = \mu_a\Theta_b$$

then: $I_j = I_0 \upsilon_j \sin^2(\phi)\Delta r \mu_a \Theta_b$

$= \Xi_j \mu_a \sin^2(\phi)$

so that all constants and range dependent parameters are merged in one factor, Ξ_j, for each range bin.

If the fish flew at a constant altitude over this planar bottom, then an appropriate TVG, T(j), could be calculated as:

$$T_j = \frac{K}{\Xi_j / N \sum_{k=0}^{N-1} \mu_{a_k} \sin^2(\phi_j)}$$

$$= \frac{K}{\Xi_j \mu_0 \sin^2(\phi_j)}$$

where: $\mu_0 = 1/N \sum_{k=0}^{N-1} \mu_{a_k}$

and: $\phi_j = \tan^{-1}\left[\frac{j\Delta r}{a}\right]$

μ_0 = average albedo

a = altitude

In fact, the Sea MARC I TVG is determined by towing the system over a relatively flat bottom at the normal operational altitude and averaging the received signal over many pings, as expressed above. The constant K is chosen so that the TVG-compensated signal normally falls within an amplitude range suited to the system's analog-to-digital converters (switch settings on the console allow manual adjustment for varied conditions).

If we now consider a real bottom, where the bottom shape and scattering properties vary, then the signal output by the system can be expressed as:

$$\delta_j = \Upsilon_j \Xi_j \mu_a \sin^2(\phi)$$

or: $$\delta_j = \frac{(\mu_0 + \Delta\mu)\sin^2(\phi)}{\mu_0 \sin^2(\phi_j)}$$

where: $\mu_a = \mu_0 + \Delta\mu$

and I assume that the signal is normalized to fall within the range $0 \leq \delta \leq 1$. If I define the *normalized scattering residual* as $\xi = \Delta\mu/\mu_0$, then:

$$\xi_j = \frac{\delta_j \sin^2(\phi_j)}{\sin^2(\phi)} - 1$$

Using surface normals computed from the Sea Beam depth map and the altitude, position, and heading from Sea MARC I data, ξ is computed directly from the sidescan signal. The residuals are modeled in the same manner as intensity data with the same rationale: since μ_a is an average weighted by the beam pattern, the backprojection reflects this same shading.

For the examples that follow, I use a variation that limits the compensation to a factor of two. The residuals are calculated as:

$$\xi_j = \delta_j[a - be^{-c \sin^2(\phi_j)/\sin^2(\phi)}] - 1$$

where: a = 2.0

Figure 6.10: Sea MARC/Sea Beam synthetic sidescan maps.

b = 1.5

c = 0.405

The exponential form is chosen as a convenient way to prevent large differences between the surface normals derived from Sea Beam and the real (unknown) surface normals from causing large fluctuations in the residuals. The motivation is to limit the dynamic range for display purposes.

Figure 6.10 shows a sequence of models (clockwise from upper left) that reflect the processing components. The first image is of an intensity model similar to that of Figure 6.5. The next image (upper right) is from a synthetic sidescan model generated from the Sea MARC track and the Sea Beam surface normals, where:

$$\delta_s = [\alpha \cdot \mathbf{n}]^2$$

In other words, the model is built from a synthetic sonar signal, δ_s, taken as the dot product squared of each surface normal, **n**, with the unit vector from each surface grid element to the sensor location, α.

In the lower right, the synthetic sidescan model is normalized by the sine-squared of the grazing angle for a flat plane at the altitude recorded in the Sea MARC I data, so that:

Figure 6.11: Normalized Sea MARC/Sea Beam intensity map.

$$\delta_s = \frac{[\alpha \cdot n]^2}{\sin^2(\phi_j)}$$

which corresponds to the inverse of the compensation factor used to compute the residuals.

These two images are not exactly analogous to the analytical formulation though they were created by a computer implementation that processed intermediate terms as a synthetic model. However, they allow a more intuitive appreciation for what is taking place. The first is like an intensity map that results from a sonar signal without a TVG (in an area of uniform albedo). The high-intensity region just below the sonar returns a strong signal at near-normal look angles, and the signal decays with increasing range and decreasing grazing angle.

The second "TVG-corrected" model shows the effects of surface slope. In the brighter regions the actual grazing angle, ϕ, is less than the assumed planar grazing angle, ϕ_j. In such areas the signal can exceed the dynamic range of the sonar system, and saturate or clip at maximum intensity. Such a problem is evident in Figures 6.5 and 6.6 on the inside curve where the look angle is near-normal to the high-relief areas downslope. The converse is true for the darker regions where $\phi \gg \phi_j$; the signal can fall below the smallest quantization level and detail is lost.

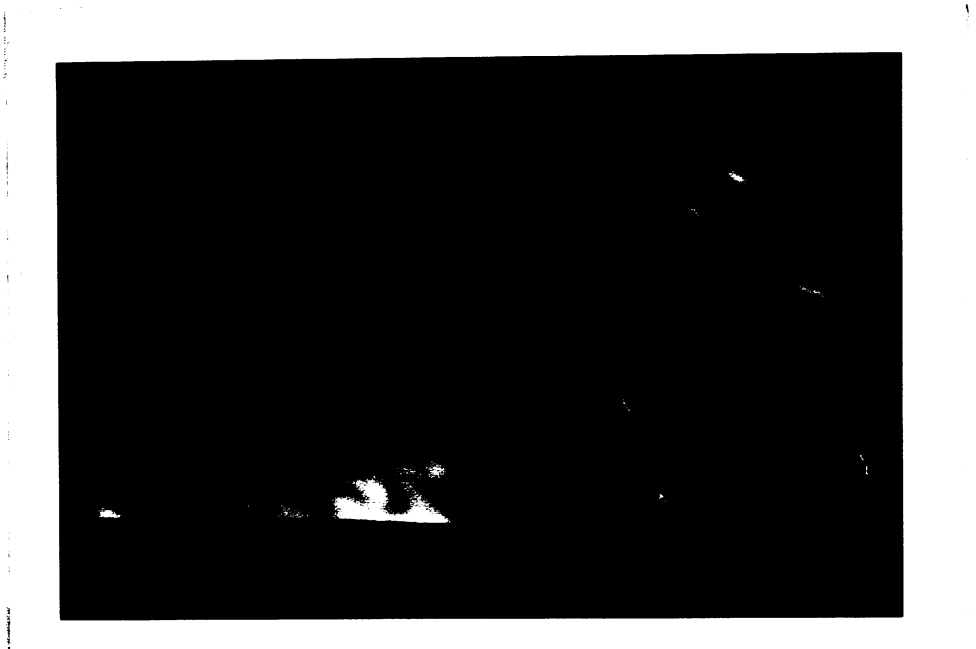

Figure 6.12: Sea MARC/Sea Beam multisensor model.

The image at the lower left in Figure 6.10 shows the model of residuals, computed as described above. In effect, the result is similar to dividing the original image at the upper left by the synthetic model at the lower right (compressed by the exponential formulation). Because the bathymetric map is a low-pass representation of the real surface, the residuals show the high-frequency component of the intensity signal and include variations caused by fine-scale relief; the model shows the effect of removing gradual variations caused by large-scale surface features without destroying all the useful information in the signal. The overall effect is to "normalize" the final sidescan map so that more detail is apparent in the extremely dark or extremely light areas of the raw intensity image. An enlarged view is shown in Figure 6.11.

In practice, the processing is not carried out with a sequence of models, but includes the intensity compensation and geometric correction on an event-by-event basis. Figure 6.12 shows the output of an implementation that creates a perspective display, mapped to the Sea Beam surface as each event is processed. As before, all modeling is incremental, using the sidescan data in a sequential, real-time manner (the Sea Beam model is preprocessed). At the scale shown above (the two-dimensional model size is {384,288}), the processing speed is about four times faster than the Sea

MARC I data rate, and the bulk of the computational load is associated with display (Gouraud-shaded polygons with hidden surface removal).

The results of this series of models are mixed. Because of the gross mismatch in resolution of the two data sets (and some misregistration), the intensities in some areas are over- or undercompensated, and appear as "hot" and "cold" spots in the images. As a postprocessing step, an adaptive filter may produce superior results. For real-time applications, though, the implementation shows the potential for combining multiple sensors to produce physical models of the environment with "nonphysical" data. The synthetic sidescan representation is a *forward* model derived from assumed shape and scattering properties. The stochastic backprojection produces an *inverse* description that uses the forward model to map intensity data to the physical level.

6.7 SUMMARY OF IMPORTANT POINTS

This preliminary sortie into multisensor modeling is sparked by a conviction that such approaches are becoming more feasible, and will become essential as more sophisticated underwater problems are addressed. Though a Sea Beam/Sea MARC I model does not offer an ideal framework for multisensor research, it serves here as a reasonable (and readily available) testbed for preliminary investigation. However, a system such as Sea MARC II, which provides phase and amplitude together, has a greater potential for building more sophisticated scattering models, and would facilitate an estimate of the angular dependence of scattering with grazing angle, as well as the acoustic albedo. Other such multisensor approaches are foreseeable, and I mention several in the next chapter. To summarize the important points in this chapter:

- Multisensor models are needed to: fully characterize an environment; overcome practical limitations of individual sensors; add redundancy; and furnish information to resolve physical properties.

- Cameras and sidescan sonars measure intensity, and cannot resolve the underlying physical properties without other information or simplifying assumptions.

- Intensity and other view-dependent (nonphysical) data cannot be combined in a simple three-dimensional representation.

- With complementary information from multiple sensors, more complete and more accurate models can be built, with less effort than for an exhaustive analysis of single-sensor data.

- Acoustic backscatter from the seafloor is complex, and cannot be characterized fully with sidescan intensity mapping alone.

- Simple sidescan image processing cannot eliminate geometric distortion introduced by the planar-bottom assumption or intensity artifacts caused by the interaction of surface shape and look angle.

- Though Sea Beam and Sea MARC I are not an ideal match for multisensor modeling, they offer complementary information that can be used to advantage.

- The stochastic backprojection approach provides a consistent computational framework in which real-time, multisensor data can be managed.

Chapter 7

CONCLUSIONS

In this final chapter I highlight important points discussed earlier, describe the results and limitations of thesis research, and point to opportunities for future investigations and to possible applications. The first section briefly restates the guiding motivations and general approach, and summarizes results. Next, the limitations of thesis research and of the method itself are described. This provides a background for discussion of potential research areas that may resolve unanswered questions and expand on the basic framework. Finally, I point to several areas that can benefit from a stochastic modeling approach, and describe other multisensor applications that appear promising.

7.1 SUMMARY OF RESULTS

My approach to building a model of the underwater environment is motivated partly by a need to give operators of remotely operated vehicles (and manned submersibles) more cues to enhance efficient piloting. Techniques that begin to take on part of the human's load can also facilitate the transition to more intelligent systems, and serve as building blocks for autonomous underwater vehicles. Other applications that can benefit from real-time feedback—exploration, surveying, and mapping, for example—are also candidates for stochastic modeling. The salient characteristics of such applications are: real-time constraints, high-bandwidth sensors with redundant information, lack of prior knowledge about the environment, and inherent inaccuracy or uncertainty in sensing and interpretation.

A low-level numerical approach, made practical by today's computational technology, can lead to more complete and more accurate models, often with less processing effort. Rather than perform an exhaustive analysis of sparse data, far beyond the point of diminishing returns, we can make use of the redundancy in overlapping data combined with complementary information from multiple sensors. By

explicitly representing uncertainty, a stochastic model lessens the danger of making unjustified assertions about the world, and offers a measure of information quality.

The method I formulate to satisfy these criteria—stochastic backprojection—is a synthesis of two "competing" processes. By convolving a spatial model of each event with the corresponding localization functions, our knowledge of that event is blurred according to the error or uncertainty in the application. At the same time, an incremental approach to reconstruction sharpens the model as more constraints are added. The effect is that of a recursive estimator, so that the detail and certainty of the model are enhanced as more information accumulates. In the absence of any inaccuracy or ambiguity, the method becomes an incremental reconstruction by backprojection and summation.

With enough events, the model converges to a "fuzzy" surface distribution, which degrades gracefully as error and uncertainty increase; the effect is that of a three-dimensional low-pass filter. From this distribution, estimates can be made to subvoxel resolution, which may permit the use of coarser, more economical models. Even simple sensor models can produce acceptable results without the performance degradation of more sophisticated representations.

Other information made explicit in the model—emptiness and ignorance, for example—can be used directly by higher-level processes. The low-level approach can also reduce the processing burden at higher levels, and serve as an intermediate representation that decouples these more asynchronous information sinks from the high-bandwidth data sources. By partitioning algorithms into preprocessing and real-time components, modeling efficiency can be further enhanced. Unlike analytical or object- and feature-based methods, the processing load and model size are deterministic—computational hardware and timing requirements can be forecast.

With the computer simulations and application results, I have demonstrated a consistent framework for real-time modeling. Because the approach is largely independent of scale, resolution, and sensing modality, it can be effective over a range of applications. For those data sets described in Chapters 5 and 6, the quality of modeling results and computational efficiency are appropriate to each real-world system. Even the unoptimized computer implementations developed for thesis research are fast enough with standard hardware to accommodate real data rates. The implication is that cost-effective field systems are well within reach.

7.2 LIMITATIONS OF CURRENT IMPLEMENTATION

The approach I have taken to modeling research relies mainly on a qualitative, visual assessment of results. On the one hand, this is important for man-in-the-loop applications that are subject to the

same criteria of relevance and utility. On the other hand, vision provides the highest bandwidth of all our senses, and offers a practical way to digest the large volume of information that a model contains. Such an approach has allowed me to quickly define the "envelope" of stochastic modeling—to look at the big picture and spot important determinants of performance.

For other applications, though, more quantitative measures of modeling fidelity and certainty are needed. Purely analytical formulations are intractable for all but the simplest combinations of geometry, sensor characteristics, environment, noise, inaccuracy, and ambiguity. The advantage of a numerical model is that these parameters can be accommodated more readily and more accurately without analytical oversimplification. Simulations offer another alternative for research validation, but need a sophisticated forward model for good results. Even then, the outcome is subject to built-in preconceptions and limitations to our understanding of real-world systems.

Further assessment and refinement require good ground-truth information. This is a problem not only for this research, but for other investigations in remote sensing underwater. Unlike terrestrial or satellite applications, the seafloor is largely inaccessible, and the provision of a realistic benchmark environment mandates resources unavailable for this work. The supplementary data for the *Monitor* work (photographs and drawings) is a step in this direction but falls short of a fully quantitative basis for evaluation. The Sea Beam and Sea MARC I comparisons in Chapter 6 are also useful, but only as a qualitative comparison with other data-processing methods.

Similar to the lack of baseline information about the modeled environments, is the lack of complete sensor models. In practice, calibrated data are rarely available from standard systems; normalized data, partly inverted with the system's built-in model, are sufficient or desirable for most applications. For research purposes, however, the aggregate of assumptions—the unknown approximations of the system designer as well as my own—complicates the evaluation of distinct modeling parameters. In one sense, though, it is a testament to the generality of stochastic modeling that such data can be treated consistently with good results. The sensor's built-in approximations are just another form of uncertainty that can be modeled directly.

Regardless of these limitations, the field data have been useful for testing under "real-world" conditions. However, an appropriate data set was not available for validating the stochastic positioning component of this thesis. As for ground-truthing, the resources needed to conduct such a field program were not at my disposal. Though the simulations show that such an application is computationally tractable with a stochastic model and suggest that good positioning may be achievable, my faith in pure simulation is bounded.

In practice, the effectiveness of terrain-relative navigation will be determined largely by the spatial bandwidth and distinctness of environmental landmarks. A smooth, featureless bottom offers no basis for horizontal positioning; such repetitive features as sand bars or large ripples may offer good resolution but are ambiguous. An interesting possibility for small-scale operations, though, is to seed such an area with distinctive markers—the Hansel-and-Gretel approach. For example, passive sonar reflectors could be inexpensive and disposable (even biodegradable), and serve the same purpose as a transponder network. An operational strategy might start with a survey of the passive net using techniques similar to those for long-baseline acoustic navigation.

Perhaps the strongest objection to this work that might be raised is that other techniques produce equivalent or better results (in large-scale mapping for example). I do not consider this a limitation if the comparison is made with postprocessing methods. My philosophy is that postprocessing is always capable (potentially) of producing better results, but perhaps only marginally so. A posterior analysis can usually bring more resources to bear on the data (more time and computing horsepower), can smooth out noise and cull the best samples, and can invert all the data in aggregate. This is not the problem I address; my approach is constrained by the needs of practical exploratory systems. As such, the assumptions and approximations I introduce are conditioned by: (1) lack of prior knowledge; (2) real-time, incremental response; and (3) cost-effective technology.

First, the constraint is inherent to the class of applications with which I am concerned. A machine intelligence exploring an unknown environment must start with a blank slate and incrementally "learn" about its surroundings. However, as more information accumulates, there arises a potential for using this "prior" knowledge to incorporate new data in the model. I have not addressed such a scenario so far, but discuss it briefly in the next section.

Second, a practical system must respond to the environment in a timely manner, preferably, using all information at its disposal. People behave analogously, making judgments and taking actions according to immediate needs. Though long deliberation or later reflection might reveal a better solution, a less-optimal but faster response is often warranted. There are two components here. If there are computational resources available, later "deliberation" or background processing might be applied to improve modeling results (see **Iterative Techniques** in the next section). Also, data smoothing rather than filtering or estimating could be used before incorporating new information in the model; the tradeoff is in the delay imposed by a causal approach.

Finally, I have made several approximations for the sake of computational efficiency. The assumption of stationary error, for example, facilitates a decomposition into preprocessing and real-time components. With more computational resources, such constraints could be relaxed and more accurate models could result. The assumption of a homogeneous medium is appropriate for short-range sonars,

but longer-range systems would have to deal with refraction and multipath errors. With today's technology, a ray-tracing solution, for example, is impractical for each event. Theoretically, though, the detectivity and localization functions can be computed with similar techniques. My point is that a numerical model offers a powerful framework for today's applications, and the scope of such an approach can only expand as computational technology evolves.

7.3 ISSUES TO EXPLORE

My goals in this thesis have been to elaborate a philosophy of modeling and to present a practical formulation with substantive results. I feel the surface has only been scratched, though, and much remains to be explored. Part of this will be to refine and further validate the modeling framework, as I discuss in the preceding section. Other research is needed to build on this work toward more complete applications for intelligent, autonomous systems. Also, there are several interesting possibilities that occurred to me as this thesis evolved. In this section I highlight a few of the more promising areas.

7.3.1 Iterative Techniques

The backprojection and summation method is the simplest approach to reconstruction, and I adopted it in this work for the sake of computational efficiency. However, the more advanced reconstruction applications in other fields almost exclusively use such techniques as ART and its variants. It seems probable that an iterative formulation of stochastic backprojection can also offer significant advantages in underwater applications. Though I am referring mainly to an approach that redistributes the signal over each range surface—analogous to the redistribution of a ray sum over the corresponding ray—iterative positioning (as in Section 4.3) or a combination of the two seems plausible.

For postprocessing use, where computational resources are not at such a premium as in real-time applications, iterative modeling would probably be most attractive for now. However, as more computing power moves to smaller packages at lower cost, a threshold of practicality will inevitably be crossed. As I mentioned in the last section, a background-approach, which would allocate resources to iterative model refinement while real-time demands are low, might be the first step. Eventually, as size and performance considerations allowed, an iterative numerical technique would probably become the preferred approach to merging most events. At the earliest stages of model evolution, however, iteration is impractical and could even magnify error.

7.3.2 Bootstrap Modeling and Positioning

Concurrent modeling and positioning is also an iterative process in which information flows bidirectionally between the two functions. To merge new events, an estimate of the sensor's location is needed; at the same time, the global model serves as a reference against which local models are compared and a position estimate extracted. In such an approach is a threshold of error beyond which the "solution" diverges. Certainly human beings get lost, so we cannot look to ourselves as an existence proof of the perfect relative-positioning system.

The problem is particularly acute when starting from a clean slate, without a model for navigational reference. I sidestepped this difficulty in the positioning simulations (Section 4.3) by using a prior model from a previous simulation, but expect to do more research in this area. It seems clear, though, that the viability and performance of bootstrap modeling depends on the algorithms, sensor characteristics, sampling rate, and the environment itself. Where the surroundings are rich in detail, the sampling rate is high, and the sensor platform moves slowly, no special considerations may be needed. In other circumstances, some kind of strategy to quickly establish reference points will lead to best results.

7.3.3 Strategies

An intelligent exploratory probe should have some notion of any holes in its knowledge base and of the most fruitful approaches to filling them in accurately and economically. Recovery from a power failure, loss of a sensor, and bootstrap navigation motivate further investigation of choosing the right sensor, range/resolution, pattern of exploration, and so forth. Because the models I have discussed also represent the system's ignorance at any moment, they offer a basis for developing plans to remove knowledge gaps. As I mentioned in the last section, a preliminary strategy to establish navigation might call for straight, even-velocity tracks so that distinctive landmarks are surveyed as a long-baseline reference frame. A spiral trajectory to incrementally expand the modeled area is another possibility.

For a multisensor platform, different sensing modalities or different combinations of range and resolution would be suited to different phases of model development, for example, low-resolution, long-range scans followed by more detailed, close-up investigation—a coarse-to-fine strategy. It is interesting to note that the backprojection approach can also be applied to omnidirectional sensors. Sometimes such a sensor would be preferred to a narrow-beam type, for example, localization of one or more targets in a large volume. With a narrow-beam sonar, the time required to scan

omnidirectionally can be great; yet, a sensor with omnidirectional sensitivity can localize over four events (from four locations if one does not lie in the same plane).

7.3.4 Virtual Models

In all the implementations developed for thesis research, a fixed modeling region is used for convenience. Widely-ranging systems, though, are not likely to maintain the necessarily large model in memory, especially if many features are considered or high resolution is desirable. This does not seem to be a significant drawback if auxiliary storage can be used. For example, model segments could be swapped to a hard disk as the sensor moved out of the corresponding region, and new segments swapped into memory. This is just a virtual memory scheme, and a clever implementation could take advantage of memory-management hardware. A straightforward extension could use a model-segment cache, and a "smart" look-ahead-cache manager with a knowledge of the platform trajectory and sensor envelopes.

7.3.5 Dynamic Modeling

Though I have only considered static environments so far, it seem plausible to extend numerical modeling to dynamic environments, where there may be objects in motion—other vehicles, for example. Such an application, the target-association problem, motivates the approach to image reconstruction developed by *Das and Boerner* [1978] and by *Rockmore et al.* [1979], and is proposed for passive tracking by *Rockmore* [1981]. *Moravec* [1987b] also mentions the possibility of detecting and tracking objections across the time dimension using "snapshots" of certainty grids.

Some accounting for moving objects probably should be made to avoid smearing the model with multiple target tracks; however, with repetitive sensing in a stochastic model they would eventually be decayed as "noise." Exactly how to represent such a dynamic model is an open question, though. Saving a complete three- or higher-dimensional model at regular intervals is an inefficient and impractical approach. Some interesting possibilities are run-length encoding along the time axis or a sequence of differencing, where only newly modified elements are recorded at each time step. This leads directly to "fingerprinting" applications for change detection in monitored environments.

7.3.6 Multifeature Modeling

I have purposely maintained a generality with a mind to future applications that hold my interest. As I have mentioned, a successful approach to modeling for more intelligent probes must accommodate

multiple redundant sensors with complementary characteristics. For example, such range sensors as sonar or lasers can create a good three-dimensional representation for modeling and interpretive processes that analyze data from imaging sonars, video, digital still cameras, and other sensors. My perspective is that there is no one paradigm to encompass all multisensor techniques. Rather, it calls for a flexible framework that can be tailored to different applications and that accommodates multiple, concurrent, cooperating processes, each knowledgeable in its domain.

7.3.7 Other Representations

I already use several representations including 2-D arrays, 3-D arrays, 3-D arrays of vectors, and vector lists of different kinds. The arrays can be tiled and passed to graphics routines as lists of vertices, polygons, surface normals, and so on. The structures are compatible with such architectures as pipelines, vector and array processors, or massively parallel systems. Though performance has been adequate with existing hardware, vector lists could be "streamed" to a digital signal processor or other pipelined architecture for more demanding applications. And I see no problem in formulating parallel decompositions that would open the door to more powerful numeric modeling.

For future work, though, I am inclined toward octree, hextree, and n-tree representations [*Jackins and Tanimoto*, 1980; *Meagher*, 1980, 1982; *Gillespie and Davis*, 1981]. Several researchers in the field have demonstrated a suite of octree tools for image processing [*Gillespie and Davis*, 1981], display [*Doctor and Torborg*, 1981], pattern recognition [*Chaudhuri*, 1985], world modeling [*Connolly*, 1984; *Shneier et al.*, 1984; *Jain and Grosky*, 1987], volumetric medical imaging [*Meagher*, 1985], obstacle avoidance [*Faverjon*, 1984], and representation of moving objects [*Nash and Ahuja*, 1983]. The main advantages I see are: a spatial decomposition that facilitates description of unstructured environments; a hierarchical representation suited to multidimensional data of varying resolution; and an economy of representation for an exploratory probe following an unpredictable trajectory. A hextree, for example, could also represent the time dimension directly and economically.

7.3.8 Multidimensional Feature Extraction

In Section 5.4 I point out the limitations of the simple technique I use for extracting three-dimensional surfaces; for more sophisticated applications other methods must be found. I see no insurmountable difficulty in formulating such surface estimators, but the issues to resolve are efficiency and accuracy. Simple candidates include: search (look for local probability peaks and link with neighboring peaks); an extension of two-dimensional convolution edge detectors to three-dimensional surface detectors (first and second difference, gradient); and morphological operations (stochastic

thinning or skeletonizing). Tougher problems must be addressed in dynamic or multifeature modeling. However, the pattern-analysis community has compiled a rich body of literature in this area, and it seems likely that many techniques can be applied directly or modified to suit a nondeterministic context.

7.3.9 Bookkeeping

There is a justifiably strong interest in multisensor techniques these days with an eye to current and future applications. My conviction is that some accounting system is needed in such work to avoid a degeneration into pure heuristics. The energy balance I use in the acoustic domain (the sonar equation) is a reasonable approach for a single sensor, but more subtle issues arise when disciplinary boundaries are crossed. Quantitative measures of performance, fidelity, convergence, and confidence should guide our applications and experimentation.

My apologies go to Dr. Shannon for loose usage of the word *information*. I feel, though, that Information Theory [*Shannon*, 1949] may offer a common token for knowledge exchange in a multidimensional, probabilistic context. As a simple example, the average information in a model can be characterized by its entropy, H, as:

$$H(\sigma_S) = -\sum_{j=0}^{J-1} \sigma_S \log_2(\sigma_S)$$

where σ_S is the scattering probability used earlier and J is the number of voxels in the model. As the information in a model increases (scattering probabilities approach 0 or 1) entropy decreases. If you consider the probability histogram in Section 4.2, it is easy to see that the entropy will converge similarly toward some constant value. Though that value depends on the different factors I discussed, the rate of change in entropy, dH/dt, might serve as an indicator of model convergence.

Individual events can also be characterized by their information content with respect to the model, and different combinations of range, resolution, beam pattern, and S/N ratio would have distinctive measures of entropy. Such measures offer a basis for formulating scanning strategies, comparing the results of different modeling algorithms, or characterizing the terrain. To cast an active sensor with transmitter and receiver in a communications framework, the channel is analogous to the volume of space being interrogated. Here we know the information potential in the signal we transmit and some measure of what we receive, but want to infer something about the channel through which it passed—the medium is the message?

7.4 MODELING APPLICATIONS

My research so far has focused on modeling and computational issues to build a foundation for future efforts. My immediate goal now is a full implementation of concurrent modeling and positioning for more autonomous systems. However, the results of thesis research suggest to me that the level of development is already sufficient to serve several underwater applications. These fall loosely into three broad categories: (1) mapping and survey; (2) piloting aids for manned systems; (3) world modeling for autonomous systems.

First, the approach offers the benefit of real-time feedback for survey and exploration, particularly for such professionals as geologists or archaeologists requiring more complete dimensional information than is typically available. In this group I include applications of shipboard mapping (Sea Beam), towed vehicles (Sea MARC), as well as ROV's and manned submersibles for small-scale, high-resolution survey. Of course, good navigation is the determining factor for real- or near-real-time shipboard modeling. For large-scale applications using satellite data, most navigation can be turned around within a few hours on today's surveys; as GPS coverage increases, the situation can only improve. For smaller-scale operations using acoustic positioning or shore stations, smoothed navigation is generated with little delay.

Incremental modeling and display give immediate feedback on the quality of the data and the coverage obtained, and allow an investigator to modify his survey plan according to results. Because the processing is faster than the data-collection rates for the applications I have considered, computational resources could also be devoted to interactive manipulation of the model. Perspective views from different vantage points, color-table manipulation to highlight different features, feature location and dimensioning with a mouse and pointer, and depth profiles from bathymetry have all been demonstrated with stochastic models on an interactive basis. It is also possible to finish the cruise with a hard-copy product reflecting the full data set. Though this would not eliminate shore-based postprocessing for best results, it offers researchers more productive time at sea.

Second, for man-in-the-loop systems—remotely operated vehicles or manned submersibles—real-time modeling can augment the operator's tunnel-vision perspective by presenting a global view of the surroundings. In the laboratory, the model has been used to generate a dynamic display with a representation of the vehicle superimposed on the image. Depth and range information can be provided using color contour maps or a shaded perspective view. Such an auxiliary display gives the pilot a more easily assimilable representation of his surroundings, and as the model's certainty increased, could be used directly for low-visibility piloting.

This can improve operator efficiency, lessen fatigue, decrease performance time for many tasks, enhance the safety of working underwater, and reduce the risk of damage or loss of the vehicle. Significant economic benefits would also accrue. Day rates for sophisticated ROV operations including crew, surface support, and a vessel, may exceed ten thousand dollars. More significantly, on an oil production platform idled while an ROV completes emergency repairs, lost revenue is measured in the hundreds of thousands of dollars a day. A reduction in task-completion times of only a few percent would show substantial saving.

Third, real-time stochastic modeling offers a powerful tool for autonomous vehicles. This thesis has mainly addressed the needs of such intelligent systems and I will not reiterate the advantages here. For those applications in which external navigation is available, the current implementation offers a context in which higher-level processes can function. *Moravec* [1987a, b] points to implementations of obstacle avoidance and path planning in a two-dimensional certainty grid that might also be extended to a three-dimensional stochastic model. *Yoerger and Slotine* [1987] also formulate a control methodology that allows an underwater vehicle to track an environmental model's potential field; this could be applied to a "probability field" in a similar manner.

Beyond these immediate applications, other multisensor implementations seem attractive for marine science research and for underwater robotics. Multifrequency sonars are appearing and, if used with a bathymetric model as in Chapter 6, could lead to more accurate and informative acoustic modeling applications. Optical imagery also could benefit from three-dimensional range information. For example, a simple surface projection of the high-frequency camera data would visually complement a coarser shape derived from a scanning laser. A better approach would use the range data, a lighting model, and an optical propagation model to compensate for attenuation and noise in color imagery. The interpretation of subbottom profiles might benefit from a composite presentation of registered bathymetric or sidescan models. The spatial distribution of magnetic anomalies, gravity, temperature, and other ocean features may offer new insights when considered with bottom shape and surface or subsurface properties.

We may be at the threshold of a new era for the exploration and understanding of the undersea environment. The remote sensors and vehicles, the computational technology and algorithms, and the tools and techniques for visualization—all are evolving apace. Physically, analytically, and conceptually, they extend our reach. Applied at sea, they allow more timely and more complete feedback, reducing cost and delay in the postprocessing tedium. And with the new information technologies, interesting results and techniques will be communicated more quickly, widely, and effectively.

REFERENCES

Albus, J. S. and D. R. Blidberg, A Control System Architecture for Multiple Autonomous Undersea Vehicles, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 444-466, University of New Hampshire, Marine Systems Laboratory, June 1987.

Amblard, F. G., D. B. Cooper, and B. Cernuschi-Frias, Estimation by Multiple Views of Outdoor Terrain Modeled by Stochastic Processes, in *Proceedings of the SPIE Conference on Intelligent Robots and Computer Vision*, vol. 726, pp. 36-43, Cambridge, MA, October 1986.

Arnold, J. B. III, J. F. Jenkins, E. M. Miller, E. W. Peterkin, C. E. Peterson, and W. K. Stewart, U.S.S. MONITOR Project: Preliminary Report on 1987 Field Work, in *Proceedings of the Conference on Historical Archaeology*, Society for Historical Archaeology, Reno, NV, January 1988.

Baggeroer, A. B., Sonar Signal Processing, in *Applications of Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1978.

Bajcsy, R., Three-Dimensional Scene Analysis, in *Proceedings of the IEEE Conference on Pattern Recognition*, pp. 1064-1074, 1980.

Ballard, D. H. and C. M. Brown, *Computer Vision*, Prentice-Hall, Englewood Cliffs, NJ, 1982.

Bane, G. L. and J. Ferguson, The Evolutionary Development of the Military Autonomous Vehicle, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 60-88, University of New Hampshire, Marine Systems Laboratory, June 1987.

Bass, F. G. and I. M. Fuks, *Wave Scattering from Statistically Rough Surfaces*, Pergammon Press, Oxford, UK, 1979.

Belcher, E., Acoustic Lens Imaging System (ALIS I): Surveillance Phase I, Unpublished Report, Applied Physics Laboratory, University of Washington, September 1987.

Belcher, E., Acoustic Lens Imaging System (ALIS I): General Description, Unpublished Report, Applied Physics Laboratory, University of Washington, November 1987.

Besl, P. and R. Jain, Range Image Understanding, in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 430-449, San Francisco, CA, June 1985.

Bixler, J. P. and D. P. Miller, A Sensory Input System for Autonomous Mobile Robots, in *Proceedings of the Workshop on Spatial Reasoning and Multi-Sensor Fusion*, pp. 211-219, American Association for Artificial Intelligence, St. Charles, IL, October 1987.

Blackington, J. G., D. M. Hussong, and J. G. Kosalos, First Results from a Combination Side-Scan Sonar and Seafloor Mapping System (Sea MARC II), in *Proceedings of the Offshore Technology Conference*, pp. 307-314, Houston, TX, May 1983.

Blidberg, D. R., E. E. Allmendinger, and N. Sideris, The Development of an Unmanned, Self-Controlled, Free-Swimming Vehicle, in *Proceedings of the Offshore Technology Conference*, Houston, TX, 1978.

Blidberg, D. R., Time-Ordered Architecture for Knowledge Based Guidance of an Unmanned Untethered Submersible, in *Proceedings of the MTS/IEEE Oceans '84 Conference*, 1984.

Blidberg, D. R., Guidance Control Architecture for the EAVE Vehicle, *IEEE Journal of Oceanic Engineering*, October 1986.

Borot, P., J. L. Michel, and H. LeRoux, EPAULARD Operational Utilizational Development of the Unmanned Untethered Submersible Systems, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 4-12, University of New Hampshire, Marine Systems Laboratory, June 1983.

Briot, M., J. C. Talou, and G. Bauzil, A Navigation Sub-System Using Ultrasonic Sensors for the Mobile Robot HILARE, in *Proceedings of the Conference on Robot Vision and Sensory Control*, Stratford-Upon-Avon, UK, 1981.

Brooks, R. A. and G. di Chiro, Principles of Computer Assisted Tomography (CAT) in Radiographic and Radioisotopic Imaging, *Physics in Medicine and Biology*, vol. 21, no. 5, pp. 689-732, 1976.

Brooks, R. A., Visual Map Making for a Mobile Robot, in *Proceedings of the IEEE Conference on Robotics and Automation*, pp. 824-829, St. Louis, MO, March 1985.

Brooks, R. A., A Robust Layered Control System for a Mobile Robot, Technical Memo, Massachusetts Institute of Technology, Artificial Intelligence Laboratory, September 1985.

Brooks, R. A., Achieving Artificial Intelligence Through Building Robots, ***

Budinger, T. F. and G. T. Gullberg, Three-Dimensional Reconstruction in Nuclear Medicine Emission Imaging, *IEEE Transactions on Nuclear Science*, vol. NS-21, pp. 2-20, June 1974.

Busby, F., *Undersea Vehicles Directory*, Busby and Associates, Arlington, VA, 1981.

Butler, B. and S. Maryka, Evolution of the Dolphin Multi-Vehicle Control System, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 23-32, University of New Hampshire, Marine Systems Laboratory, June 1987.

Carnap, R., *Logical Foundations of Probability*, University of Chicago Press, Chicago, IL, 1950.

CG&A Staff, Visualization in Scientific Computing—A Synopsis, *IEEE Computer Graphics and Applications*, pp. 61-70, July 1987.

Chandé, A. M. and K. M. Noon, Obstacle Avoidance and Navigational Sensing for an Autonomous Underwater Vehicle, in *Proceedings of the SPIE Conference on Mobile Robots*, vol. 727, pp. 296-303, Cambridge, MA, October 1986.

Chappell, S. G., A Blackboard Based System for Context Sensitive Mission Planning in an Autonomous Vehicle, in *Proceedings of the Symposium on Unmanned Untethered Submersible*

Technology, pp. 467-476, University of New Hampshire, Marine Systems Laboratory, June 1987.

Chatila, R. and J.-P. Laumond, Position Referencing and Consistent World Modeling for Mobile Robots, in *Proceedings of the IEEE Conference on Robotics and Automation*, pp. 138-145, St. Louis, MO, March 1985.

Chaudhuri, B. B., Applications of Quadtree, Octree, and Binary Tree Decomposition Techniques to Shape Analysis and Pattern Recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-7, no. 6, pp. 652-661, November 1985.

Chavez, P. S. Jr., Processing Techniques for Digital Sonar Images from GLORIA, *Photogrammetric Engineering and Remote Sensing*, vol. 52, no. 8, pp. 1133-1145, August 1986.

Chavez, P. S. Jr., J. A. Anderson, and J. W. Schoonmaker, Underwater Mapping Using GLORIA and MIPS, in *Proceedings of the MTS/IEEE Oceans '87 Conference*, pp. 1202-1205, 1987.

Cheeseman, P., In Defense of Probability, in *Proceedings of the International Joint Conference on Artificial Intelligence*, vol. 2, pp. 1002-1009, August 1985.

Cho, Z. H., General Views on 3-D Image Reconstruction and Computerized Transverse Axial Tomography, *IEEE Transactions on Nuclear Science*, vol. NS-21, pp. 44-71, June 1974.

Clay, C. S. and H. Medwin, *Acoustical Oceanography: Principles and Applications*, Wiley-Interscience, New York, NY, 1977.

Coles, S., B. Raphael, R. O. Duda, C. Rosen, T. D. Garvey, R. Yates, and J. Munson, Application of Intelligent Automata to Reconnaissance, Technical Report, Stanford Research Institute, Menlo Park, CA, 1969.

Collins, J. S., Advanced Marine Robotics as a Strategic Technology for Canada and Observations of a Related Large Scale Project in Japan, in *Proceedings of the MTS/IEEE Oceans '87 Conference*, pp. 1246-1253, April 1987.

Connolly, C. I., Cumulative Generation of Octree Models from Range Data, in *Proceedings of the IEEE Robotics Conference*, pp. 25-32, Atlanta, GA, March 1984.

Connolly, C. I., The Determination of Next Best Views, in *Proceedings of the IEEE Conference on Robotics and Automation*, pp. 432-435, St. Louis, MO, March 1985.

Cornsweet, T. N., *Visual Perception*, Academic Press, New York, NY, 1970.

Crowley, J. L., Dynamic World Modeling for an Intelligent Mobile Robot, in *Proceedings of the IEEE Conference on Robotics and Automation*, pp. 207-210, St. Louis, MO, March 1985.

Crowley, J. M., Navigation for an Intelligent Mobile Robot, *IEEE Journal of Robotics and Automation*, vol. RA-1, no. 1, pp. 31-41, March 1985.

Cuschieri, J. M. and M. Hebert, Sonar Applications for Underwater Vision, in *Proceedings of the ASME Conference on Energy Sources Technology*, OED-13, pp. 5-11, New Orleans, LA, January 1988.

Cyr, R., A Review of Obstacle Avoidance/Search Sonars Suitable for Submersible Applications, *Marine Technology Society Journal*, vol. 20, no. 4, pp. 47-57, 1987.

Das, Y. and W. M. Boerner, On Radar Target Shape Estimation Using Algorithms for Reconstruction from Projections, *IEEE Transactions on Antennas and Propagation*, vol. AP-26, no. 2, pp. 274-279, March 1978.

Davis, E. E., R. G. Currie, B. S. Sawyer, and J. G. Kosalos, The Use of Swath Bathymetric and Acoustic Image Mapping Tools in Marine Geoscience, *Marine Technology Society Journal*, vol. 20, no. 4, pp. 17-27, December 1987.

Dean, T., High-Level Planning and Low-Level Control, in *Proceedings of the SPIE Conference on Intelligent Robots and Computer Vision*, vol. 848, pp. 496-501, Cambridge, MA, November 1987.

Dempster, A. P., Upper and Lower Probabilities Induced by a Multivalued Mapping, *Annals of Mathematics and Statistics* **38**, pp. 325-339, 1967.

Denton, R. V., A. J. Rockmore, B. Friedlander, and J. S. Malin, An Image Reconstruction Approach to Target Association, in *Proceedings of the IEEE Asilomar Conference on Circuits, Systems, and Computers*, pp. 304-308, Pacific Grove, CA, November 1978.

Dixon, T. H., T. J. Pivirotto, R. F. Chapman, and R. C. Tyce, A Range-Gated Laser System for Ocean Floor Imaging, *Marine Technology Society Journal*, vol. 17, no. 3, pp. 5-12, September 1983.

Doctor, L. J. and J. G. Torborg, Display Techniques for Octree-Encoded Objects, *IEEE Computer Graphics and Applications*, pp. 29-38, July 1981.

Doob, J. L., *Stochastic Processes*, John Wiley and Sons, New York, NY, 1953.

Drumheller, M., Mobile Robot Localization Using Sonar, A. I. Memo 826, Massachusetts Institute of Technology, Artificial Intelligence Laboratory, January 1985.

Duntley, S. Q., Light in the Sea, *Journal of the Optical Society of America*, vol. 53, no. 2, pp. 214-233, 1963.

Durham, J., P. Heckman, D. Bryan, and R. Reich, EAVE-WEST: A Testbed for Plan Execution, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 33-43, University of New Hampshire, Marine Systems Laboratory, June 1987.

Durrant-White, H. F., Uncertain Geometry in Robotics, *IEEE Journal of Robotics and Automation*, vol. 4, no. 1, pp. 23-31, February 1988.

Eckart, C., The Scattering of Sound from the Sea Surface, *Journal of the Acoustical Society of America* **25**, pp. 566-570, 1953.

Elfes, A., A Sonar-Based Mapping and Navigation System, in *Proceedings of the IEEE Conference on Robotics and Automation*, San Francisco, CA, April 1986.

Elfes, A., Sonar-Based Real-World Mapping and Navigation, Submitted to *IEEE Transactions on Robotics and Automation*, June 1986.

Embley, R. W., I. R. Jonasson, M. R. Perfit, J. M. Franklin, M. A. Tivey, A. Malahoff, M. F. Smith, and T. J. G. Francis, Submersible Investigation of an Extended Hydrothermal System on the Galapagos Ridge: Sulfide Mounds, Stockwork Zones, and Differentiated Lavas, Special Issue of *Canadian Mineralogist: Recent Hydrothermal Mineralogy at Seafloor Spreading Centres*, in press, 1988.

Eppig, S. H., Autonomous Vehicles for Underwater Search and Survey, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 46-60, University of New Hampshire, Marine Systems Laboratory, June 1985.

Farr, H. K., Multibeam Bathymetric Sonar: Sea Beam and Hydrochart, *Marine Geology* **4**, pp. 77-93, 1980.

Farre, J. A. and W. B. Ryan, 3-D View of Erosional Scars on the U.S. Mid-Atlantic Continental Margin, *The American Association of Petroleum Geologists Bulletin*, vol. 69, no. 6, pp. 923-932, June 1985.

Farre, J. A. and W. B. Ryan, Superficial Geology of the Continental Margin Offshore New Jersey in the Vicinity of Deep Sea Drilling Project Sites 612 and 613, in *Initial Reports of the Deep Sea Drilling Project*, vol. XCW, pp. 725-759, eds. C. W. Poang, A. B. Watts, et al., U. S. Government Printing Office, Washington, DC, 1987.

Farreny, H. and H. Prade, Tackling Uncertainty and Imprecision in Robotics, in *Robotics Research: The Third International Symposium*, eds. O. D. Faugeras and G. Giralt, pp. 85-91, MIT Press, Cambridge, MA, 1986.

Faugeras, O. D., New Steps Toward a Flexible 3-D Vision System for Robotics, in *Proceedings of the International Joint Conference on Pattern Recognition*, pp. 796-805, July 1984.

Faugeras, O. D., Artificial 3D Vision, in *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 1169-1171, Milan, Italy, August 1987.

Faverjon, B., Obstacle Avoidance Using an Octree in the Configuration Space of a Manipulator, in *Proceedings of the IEEE Robotics Conference*, pp. 504-512, Atlanta, GA, March 1984.

Feller, W., *An Introduction to Probability Theory and Its Applications*, John Wiley and Sons, New York, NY, 1950.

Ferguson, J. and E. Jackson, Design and Development of a Diesel Powered Semi-Submersible ROV, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 39-51, University of New Hampshire, Marine Systems Laboratory, June 1983.

Ferrari, L. A., ed., Pattern Recognition and Acoustical Imaging, *Proceedings of the SPIE*, vol. 768, Newport Beach, CA, February 1987.

Flynn, A. M., Redundant Sensors for Mobile Robot Navigation, Technical Report 859, Massachusetts Institute of Technology, Artificial Intelligence Laboratory, September 1985.

Gaines, B. R., Foundations of Fuzzy Reasoning, *International Journal of Man-Machine Studies* **8**, pp. 623-668, 1976.

Gallo, D. G., J. Fox, and K. C. Macdonald, A Sea Beam Investigation of the Clipperton Transform Fault: The Morphotectonic Expression of a Fast Slipping Transform Boundary, *Journal of Geophysical Research*, vol. 91, no. B3, pp. 3455-3467, March 1986.

Gelb, A., ed., *Applied Optimal Estimation*, MIT Press, Cambridge, MA, 1974.

Geyer, E. M., P. M. Creamer, J. A. D'Appolito, and R. G. Rains, Characteristics and Capabilities of Navigation Systems for Unmanned Untethered Submersibles, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 320-347, University of New Hampshire, Marine Systems Laboratory, June 1987.

Gillespie, R. and W. A. Davis, Tree Data Structures for Graphics and Image Processing, in *Proceedings of the Canadian Man-Computer Communications Conference*, pp. 155-162, Waterloo, Ontario, June 1981.

Giralt, G., R. Sobek, and R. Chatila, A Multi-Level Planning and Navigation System for a Mobile Robot; A First Approach to HILARE, in *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 335-337, Tokyo, Japan, August 1979.

Giralt, G., R. Chatila, and M. Vaisset, An Integrated Navigation and Motion Control System for Autonomous Multisensory Mobile Robots, in *Robotics Research: First International Symposium*, eds. M. Brady and R. Paul, pp. 191-214, MIT Press, Cambridge, MA, 1983.

Glynn, J. M., Considerations for Acquisition and Processing of Side Scan Sonar Data by an Autonomous Submersible, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 109-113, University of New Hampshire, Marine Systems Laboratory, June 1985.

Gordon, J. and E. H. Shortliffe, A Method for Managing Evidential Reasoning in a Hierarchical Hypothesis Space, *Artificial Intelligence* **26**, pp. 323-357, 1985.

Gordon, R. and G. T. Herman, Three-Dimensional Reconstruction from Projections: A review of Algorithms, in *International Review of Cytology*, eds. G. H. Bourne, J. F. Danielli, and K. W. Jeon, vol. 38, pp. 111-151, Academic Press, New York, NY, 1974.

Gordon, R., G. T. Herman, and S. A. Johnson, Image Reconstruction from Projections, *Scientific American*, vol. 233, no. 4, pp. 56-68, October 1975.

Grimson, W. E. L., *From Images to Surfaces*, MIT Press, Cambridge, MA, 1981.

Grimson, W. E. L. and T. Lozano-Perez, Model-Based Recognition and Localization from Sparse Range or Tactile Data, A.I. MEMO 738, Massachusetts Institute of Technology, Artificial Intelligence Laboratory, August 1983.

Hahn, D. A., G. N. Williams, M. Wilcox, and P. Wilcox, A Computerized High Resolution Underwater Ultrasound Triangulation Mapping System, in *Proceedings of the MTS/IEEE Oceans '85 Conference*, pp. 104-109, San Diego, CA, November 1985.

Hanson, K. M., Bayesian and Related Methods in Image Reconstruction from Incomplete Data, in *Image Recovery: Theory and Application*, ed. H. Stark, pp. 79-125, Academic Press, New York, NY, 1987.

Harris, S. E., R. H. Squires, and E. M. Bergeron, Underwater Imagery Using an Electronic Still Camera, in *Proceedings of the MTS/IEEE Oceans '87 Conference*, pp. 1242-1245, April 1987.

Havlice, J. F. and J. C. Taenzer, Medical Ultrasonic Imaging: An Overview of Principles and Instrumentation, *Proceedings of the IEEE*, vol. 67, pp. 620-641, April 1979.

Heckman, P. J. Jr., Free Swimming Submersible Testbed (EAVE WEST), Technical Report NOSC-TR-622, Naval Ocean Systems Center, San Diego, CA, September 1980.

Henderson, A. R., Light and Lasers Underwater, in *Proceedings of the MTS ROV '88 Conference*, pp. 554-564, Bergen, Norway, April 1988.

Henderson, T. C., P. Allen, I. Cox, A. Mitiche, H. Durrant-White, and W. Snyder, Workshop on Multisensor Integration in Manufacturing Automation, Technical Report UUCS-87-006, University of Utah, Department of Computer Science, Salt Lake City, UT, February 1987.

Herman, G. T., A. Lent, and S. W. Rowland, ART: Mathematics and Applications; A Report on the Mathematical Foundations and on the Applicability to Real Data of the Algebraic Reconstruction Techniques, *Journal of Theoretical Biology* **42**, pp. 1-32, 1973.

Hong, T. H. and M. O. Shneier, Describing a Robot's Workspace Using A Sequence of Views from a Moving Camera, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1985.

Horn, B. K. P., Density Reconstruction Using Arbitrary Ray-Sampling Schemes, *Proceedings of the of the IEEE*, vol. 66, no. 5, pp. 551-562, May 1978.

Horton, J. W., *Fundamentals of Sonar*, United States Naval Institute, Annapolis, MD, 1959.

Holmes, L., Solid-State Lasers: New Technology and Higher Performance Standards, *Laser Focus/Electro-Optics*, pp. 58-68, November 1986.

Horvitz, E. J., D. F. Heckerman, and C. P. Langlotz, A Framework for Comparing Alternative Formalisms for Plausible Reasoning, in *Proceedings of the AAAI '86 Conference*, pp. 210-214, Philadelphia, PA, August 1986.

Hussong, D. M. and P. Fryer, Back-Arc Seamounts and the Sea MARC II Seafloor Mapping System, *EOS Transactions of The American Geophysical Union*, vol. 64, no. 45, pp. 627-632, November 1983.

Jackins, C. L. and S. L. Tanimoto, Oct-Trees and Their Use in Representing Three-Dimensional Objects, *Computer Vision, Graphics, and Image Processing* **14**, pp. 249-270, 1980.

Jackson, D. R., D. P. Winebrenner, and A. Ishimaru, Application of the Composite Roughness Model to High-Frequency Bottom Backscattering, *Journal of the Acoustical Society of America*, vol. 79, no. 5, May 1986.

Jackson, E., Autonomous Remotely Controlled Submersible - "ARCS", in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 77-88, University of New Hampshire, Marine Systems Laboratory, June 1983.

Jaffe, J. S. and P. M. Cassereau, Multibeam Imaging Using Spatially Variant Insonification, *Journal of the Acoustical Society of America*, vol. 83, no. 4, April 1988.

Jain, R. and W. I. Grosky, Hyper-Pyramids for Integration of Spatial Information, in *Proceedings of the Workshop on Spatial Reasoning and Multi-Sensor Fusion*, pp. 72-81, American Association for Artificial Intelligence, St. Charles, IL, October 1987.

Jarry, J. and J. L. Michel, Unmanned Untethered Submersible Systems, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 2-9, University of New Hampshire, Marine Systems Laboratory, June 1985.

Jarvis, R. A., A perspective on Range Finding Techniques for Computer Vision, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-5, no. 2, pp. 122-139, March 1983.

Johnson, D. and S. H. Eppig, Aided Inertial Navigation Systems for Underwater Vehicles, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 265-282, University of New Hampshire, Marine Systems Laboratory, June 1987.

Kak, A. and S. Chen, eds., *Proceedings of the Workshop on Spatial Reasoning and Multi-Sensor Fusion*, American Association for Artificial Intelligence, St. Charles, IL, October 1987.

Kastens, K. A. and W. B. F. Ryan, Structural and Volcanic Expression of a Fast Slipping Ridge-Transform-Ridge-Plate Boundary: Sea MARC I and Photographic Surveys at the Clipperton Transform Fault, *Journal of Geophysical Research*, vol. 91, no. B3, pp. 3469-3488, March 1986.

Kino, G. S., Acoustic Imaging for Nondestructive Evaluation, *Proceedings of the IEEE*, vol. 67, pp. 510-525, April 1979.

Klepaker, R. A., K. Vestgård, J. O. Hallset, and F. T. Knudsen, A Freeswimming ROV, in *Proceedings of the MTS ROV '86 Conference*, pp. 273-282, Bergen, Norway, April 1986.

Klepaker, R. A., H. J. Alker, and F. T. Knudsen, A Freeswimming Intelligent Robot, in *Proceedings of the MTS ROV '88/Intervention '88 Conference*, pp. 567-576, Bergen, Norway, April 1988.

Klepsvik, J. O., M. Haandlykken, and H. O. Torsen, Subsea Lasers for Real-Time 3D Inspection and Surveying: Principles and Applications, in *Proceedings of the MTS ROV '87 Conference* (addendum), San Diego, CA, March 1987.

Kosalos, J. G. and D. Chayes, A Portable System for Ocean Bottom Imaging and Charting, in *Proceedings of the MTS/IEEE Oceans '83 Conference*, pp. 649-656, 1983.

Krabach, M., Potential AUV Applications to Nuclear Power Plant Inspection, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 226-230, University of New Hampshire, Marine Systems Laboratory, June 1983.

Kuipers, B. J. and Y. T. Byun, A Qualitative Approach to Robot Exploration and Map-Learning, in *Proceedings of the Workshop on Spatial Reasoning and Multi-Sensor Fusion*, pp. 390-404, American Association for Artificial Intelligence, St. Charles, IL, October 1987.

Lee, H. and G. Wade, eds., Modern Acoustical Imaging, IEEE Press, New York, NY, 1986.

Levitt, T. S., D. T. Lawton, D. M. Chelberg, P. C. Nelson, and J. W. Dye, Visual Memory Structure for a Mobile Robot, in *Proceedings of the Workshop on Spatial Reasoning and Multi-Sensor Fusion*, pp. 92-106, American Association for Artificial Intelligence, St. Charles, IL, October 1987.

Liebowitz, S. and D. Casasent, Multi-Sensor Processing: Object Detection and Identification, in *Proceedings of the SPIE Conference on Mobile Robots II*, vol. 852, Cambridge, MA, November 1987.

Magee, M. and M. Nathan, Spatial Reasoning, Sensor Repositioning and Disambiguation in 3D Model Based Recognition, in *Proceedings of the Workshop on Spatial Reasoning and Multi-Sensor Fusion*, pp. 262-271, American Association for Artificial Intelligence, St. Charles, IL, October 1987.

Marce, L. and M. Julliere, Dynamic Localization of A Mobile Robot Through Range Measurements, in *Proceedings of the SPIE Conference on Mobile Robots*, vol. 727, pp. 274-277, Cambridge, MA, October 1986.

Marr, D., *Vision*, W. H. Freeman and Company, New York, NY, 1982.

Martin, W. N. and J. K. Aggarwal, Volumetric Description of Objects from Multiple Views, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-5, no. 2, pp. 150-158, March 1983.

Meagher, D. J., Octree Encoding: A New Technique for the Representation, Manipulation, and Display of Arbitrary 3-D Objects by Computer, Technical Report IPL-TR-80-111, Rennselaer Polytechnic Institute, Image Processing Laboratory, Troy, NY, October 1980.

Meagher, D. J., Geometric Modeling Using Octree Encoding, *Computer Graphics and Image Processing* 19, pp. 129-147, 1982.

Meagher, D. J., The Application of Octree Techniques to 3-D Medical Imaging, in *Proceedings of the Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 612-615, Chicago, IL, September 1985.

Mersereau, R. M. and A. V. Oppenheim, Digital Reconstruction of Multidimensional Signals from Their Projections, *Proceedings of the of the IEEE*, vol. 62, no. 10, pp. 1319-1338, October 1974.

Michel, J. L., EPAULARD: Operational Developments, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 14-17, University of New Hampshire, Marine Systems Laboratory, June 1987.

Miller, D. P., Two Dimensional Mobile Robot Positioning Using Onboard Sonar, in *Proceedings of the Ninth IEEE William T. Pecora Memorial Remote Sensing Symposium*, pp. 362-369, Sioux Falls, SD, October 1984.

Miller, D. P., A Spatial Representation System for Mobile Robots, in *Proceedings of the IEEE Conference on Robotics and Automation*, pp. 122-127, St. Louis, MO, March 1985.

Miller, D. P. and M. G. Slack, Efficient Navigation Through Dynamic Domains, in *Proceedings of the Workshop on Spatial Reasoning and Multi-Sensor Fusion*, pp. 230-241, American Association for Artificial Intelligence, St. Charles, IL, October 1987.

Moravec, H. P., Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover, Ph.D. Thesis, Stanford University, Stanford, CA, May 1980.

Moravec, H. P., The Stanford Card and The CMU Rover, *Proceedings of the of the IEEE*, vol. 71, no. 7, pp. 872-884, July 1983.

Moravec, H. P. and A. Elfes, High Resolution Maps from Wide Angle Sonar, in *Proceedings of the IEEE Conference on Robotics and Automation*, pp. 116-121, St. Louis, MO, March 1985.

Moravec, H. P., Certainty Grids for Mobile Robots, Presented at the JPL/NASA Space Telerobotics Workshop, Pasadena, CA, January 1987.

Moravec, H. P., Sensor Fusion in Certainty Grids for Mobile Robots, Presented at the NATO Workshop on Sensor Devices and Systems for Robotics, Costa Brava, Spain, October 1987.

de Moustier, C., Beyond Bathymetry: Mapping Acoustic Backscattering from the Seafloor with Sea Beam, *Journal of the Acoustical Society of America*, vol. 79, no. 2, pp. 316-331, February 1986.

de Moustier, C. and M. C. Kleinrock, Bathymetric Artifacts in Sea Beam Data: How to Recognize Them and What Causes Them, *Journal of Geophysical Research*, vol. 91, no. B3, pp. 3407-3424, March 1986.

de Moustier, C., State of the Art in Swath Bathymetry Survey Systems, in *Current Practices and New Technology in Ocean Engineering—OED-vol. 11*, eds. G. K. Wolfe and P. Y. Chang, pp. 29-38, ASME, New York, NY, 1988.

Mueller, R. K., M. Kaveh, and G. Wade, Reconstructive Tomography and Applications to Ultrasonics, *Proceedings of the IEEE*, vol. 67, pp. 567-587, April 1979.

Munk, W. and C. Wunsch, Ocean Acoustic Tomography: A Scheme for Large Scale Monitoring, *Deep-Sea Research*, vol. 26A, pp. 123-161, 1979.

Nash, C and N. Ahuja, Octree Representations of Moving Objects, in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 380-381, Washington, DC, June 1983.

Nichols, B. and K. Jensen, Autonomously Recorded Side Scan Sonar on an Untethered Submersible (ARSUS), in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 114-123, University of New Hampshire, Marine Systems Laboratory, June 1985.

Nilsson, N. J., A Mobile Automaton: An Application of Artificial Intelligence Techniques, in *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 509-520, Washington, DC, May 1969.

Nishimura, C. E. and D. W. Forsyth, Improvements in Navigation Using Sea Beam Crossing Errors, Submitted to *Marine Geophysical Researches*, July 1987.

Nitzan, D., A. E. Brain, and R. O. Duda, The Measurement and Use of Registered Reflectance and Range Data in Scene Analysis, *Proceedings of the of the IEEE*, vol. 65, no. 2, pp. 206-220, February 1977.

Norton, S. J. and M. Linzer, Ultrasonic Reflectivity Tomography: Reconstruction with Circular Transducer Arrays, in *Ultrasonic Imaging 1*, pp. 154-184, Academic Press, New York, NY, 1979.

Norton, S. J. and M. Linzer, Ultrasonic Reflectivity Imaging in Three Dimensions: Reconstruction with Spherical Transducer Arrays, in *Ultrasonic Imaging 1*, pp. 210-231, Academic Press, New York, NY, 1979.

Nutter, J. T., Uncertainty and Probability, in *Proceedings of the International Joint Conference on Artificial Intelligence*, pp. 373-379, Milan, Italy, August 1987.

Ogilvy, J. A., Wave Scattering from Rough Surfaces, *Reports on Progress in Physics* **50**, pp. 1553-1608, 1987.

Ol'shevskii, V. V., *Statistical Methods in Sonar*, Consultants Bureau, New York, NY, 1978.

Orser, D. J. and M. Roche, The Extraction of Topographic Features in Support of Autonomous Underwater Vehicle Navigation, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 502-514, University of New Hampshire, Marine Systems Laboratory, June 1987.

Peterkin, E. W., *U.S.S. Monitor* Historical Report Series, vol. 1, no. 1, National Oceanic and Atmospheric Administration, National Ocean Service, Washington, DC, December 1985.

Polvani, D. G., Magnetic Guidance of Autonomous Vehicles, Part 2, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 257-264, University of New Hampshire, Marine Systems Laboratory, June 1987.

Posdamer, J. L., Computer Geometric Modeling for Machine Perception of Three-Dimensional Solids, in *Proceedings of the SPIE Conference on 3-D Machine Perception*, vol. 283, pp. 7-14, Washington, DC, April 1981.

Pratt, W. K., *Digital Image Processing*, John Wiley and Sons, New York, NY, 1978.

Rao, S. V. N., S. S. Iyengar, C. C. Jorgensen, and C. R. Weisbin, Concurrent Algorithms for Autonomous Robot Navigation in an Unexplored Terrain, in *Proceedings of the IEEE Conference on Robotics and Automation*, pp. 1137-1144, San Francisco, CA, April 1986.

Renard, V. and J. P. Allenou, SEA BEAM Multi-Beam Echo-Sounding in *Jean Charcot*; Description, Evaluation, and First Results, *International Hydrographic Review*, vol. 56, no. 1, pp. 35-67, 1979.

Rockmore, A. J., R. V. Denton, and B. Friedlander, Direct Three-Dimensional Image Reconstruction, *IEEE Transactions on Antennas and Propagation*, vol. AP-27, no. 2, pp. 239-241, March 1979.

Rockmore, A. J., Multi-Array Surveillance Processing Using Image Reconstruction, in *Proceedings of the IEEE Asilomar Conference on Circuits, Systems, and Computers*, pp. 119-241, Pacific Grove, CA, November 1981.

Rosen, C. and N. J. Nilsson, Application of Intelligent Automata to Reconnaissance, Technical Report No. RADC-TR-67-657, Rome Air Development Center, Air Force Systems Command, January 1968.

Rosen, C., Research on Intelligent Automata, Technical Report No. RADC-TR-69-333, Rome Air Development Center, Air Force Systems Command, April 1970.

Russell, G. T., R. M. Dunbar, and R.T. Holmes, A review of the Free Swimming Submersible Research Programme 1981/83, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 13-38, University of New Hampshire, Marine Systems Laboratory, June 1983.

Russell, G. T. and D. M. Lane, An Intelligent Guidance System Utilizing Sector Scan Sonar Image Interpretation, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 257-271, University of New Hampshire, Marine Systems Laboratory, June 1985.

Russell, G. T., A Knowledge-Based System Framework for Environmental Perception in a Subsea Robotics Context, *IEEE Journal of Oceanic Engineering*, vol. OE-11, no. 3, pp. 401-412, July 1986.

Salmon, W. C., *The Foundations of Scientific Inference*, University of Pittsburgh Press, Pittsburgh, PA, 1966.

Salmon, W. C., Confirmation, *Scientific American*, vol. 228, no. 5, pp. 75-83, May 1973.

Saridis, G. N. and K. P. Valavanis, On the Theory Of Intelligent Controls, in *Proceedings of the SPIE Conference on Intelligent Robots and Computer Vision*, vol. 848, pp. 488-495, Cambridge, MA, November 1987.

Scudder, H. J., Introduction to Computer Aided Tomography, *Proceedings of the of the IEEE*, vol. 66, no. 6, pp. 628-637, June 1978.

Serey, B. and L. Mathies, Obstacle Avoidance Using 1-D Stereo Vision, Technical Report, Carnegie-Mellon University, Robotics Institute, Pittsburgh, PA, November 1986.

Shafer, G., *A Mathematical Theory of Evidence*, Princeton University Press, Princeton, NJ, 1976.

Shannon, C. E., *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, IL, 1949.

Sheridan, T. B., Supervisory Control: Problems, Theory, and Experiment for Application to Human-Computer Interaction in Undersea Remote Systems, Technical Report, Massachusetts Institute of Technology, Man-Machine Systems Laboratory, Cambridge, MA, March 1982.

Shneier, M. O., E. W. Kent, and P. Mansbach, Representing Workspace and Model Knowledge for a Robot with Mobile Sensors, in *Proceedings of the International Conference on Pattern Recognition*, pp. 199-202, Montreal, Quebec, July 1984.

Shneier, M. O., R. Lumia, and E. W. Kent, Model-Based Strategies for High-Level Robot Vision, *Computer Vision, Graphics, and Image Processing* **33**, pp. 293-306, 1986.

Shortliffe, E. H., *Computer-Based Medical Consultations: MYCIN*, Elsivier, New York, NY, 1976.

Smith, R., M. Self, and P. Cheeseman, A Stochastic Map for Uncertain Spatial Relationships, in *Proceedings of the Workshop on Spatial Reasoning and Multi-Sensor Fusion* (addendum), American Association for Artificial Intelligence, St. Charles, IL, October 1987.

Srihari, S. N., Representation of Three-Dimensional Images, *ACM Computing Surveys*, vol. 13, no. 4, pp. 399-424, December 1981.

Stanton, T. K., Sonar Estimates of Seafloor Roughness, *Journal of the Acoustical Society of America*, vol. 75, no. 3, pp. 809-818, March 1983.

Stanton, T. K. and C. S. Clay, Sonar Echo Statistics as a Remote-Sensing Tool: Volume and Seafloor, *IEEE Journal of Oceanic Engineering*, vol. OE-11, no. 1, pp. 79-95, January 1986.

Stenovec, G. M., Systems Overview of Unmanned, Untethered Submersibles, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 10-22, University of New Hampshire, Marine Systems Laboratory, June 1985.

Stephanou, H. E. and A. P. Sage, Perspectives on Imperfect Information Processing, *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-17, no. 5, pp. 780-798, September/October 1987.

Stewart, W. K., A Non-Deterministic Approach to 3-D Modeling Underwater, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, vol. 1, pp. 283-309, University of New Hampshire, Marine Systems Laboratory, June 1987.

Stewart, W. K., Computer Modeling and Imaging Underwater, *Computers in Science*, vol. 1, no. 3, pp. 22-32, November/December 1987.

Stewart, W. K., A Model-Based Approach to 3-D Imaging and Mapping Underwater, in *Proceedings of the ASME Conference on Offshore Mechanics and Arctic Engineering*, vol. 6, pp. 61-71, Houston, TX, February 1988.

Stoffel, J. C., ed., *Graphical and Binary Image Processing and Applications*, Artech House, Dedham, MA, 1981;

Strat, T. M. and G. B. Smith, The Management of Spatial Information in a Mobile Robot, in *Proceedings of the Workshop on Spatial Reasoning and Multi-Sensor Fusion*, pp. 240-249, American Association for Artificial Intelligence, St. Charles, IL, October 1987.

Sutton, J. L., Underwater Acoustic Imaging, *Proceedings of the IEEE*, vol. 67, pp. 554-566, April 1979.

Thomas, B., Potential Applications of Underwater Autonomous Vehicles, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 223-225, University of New Hampshire, Marine Systems Laboratory, June 1983.

Thomas, C. C., Continuing Program of the DOLPHIN Semi Submersible, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 223-229, University of New Hampshire, Marine Systems Laboratory, June 1985.

Thorpe, C., Path Relaxation: Path Planning for a Mobile Robot, Technical Report CMU-RI-TR-84-5, Carnegie-Mellon University, Robotics Institute, Pittsburgh, PA, April 1984.

Togai, M. and H. Watanabe, Expert System on a Chip: An Engine for Real-Time Approximate Reasoning, *IEEE Expert*, pp. 55-62, Fall 1986.

Trimble, G. M., A Multiprocessor System for AUV Applications, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 208-219, University of New Hampshire, Marine Systems Laboratory, June 1987.

Tusting, R. F. and F. M. Caimi, Inertially Aided Navigation of Undersea Vehicles, in *Proceedings of the MTS/IEEE Oceans '87 Conference*, pp. 1278-1283, April 1987.

Twitchell, D. C. Jr., Erosion of the Florida Escarpment: Eastern Gulf of Mexico, Ph.D. Thesis, University of Rhode Island, July 1988.

Tyce, R. C., Deep Seafloor Mapping Systems—A Review, *Marine Technology Society Journal*, vol. 20, no. 4, pp. 4-16, December 1987.

Tyrén, C., Magnetic Terrain Navigation, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 245-256, University of New Hampshire, Marine Systems Laboratory, June 1987.

Urick, R. J., *Principles of Underwater Sound*, McGraw-Hill, New York, NY, 1975.

Urick, R. J., *Sound Propagation in the Sea*, U. S. Government Printing Office, Washington, DC, 1979.

Veenstra, J. and N. Ahuja, Octree Generation from Silhouette Views of an Object, in *Proceedings of the IEEE Conference on Robotics and Automation*, pp. 843-848, St. Louis, MO, March 1985.

von der Heydt, K., G. L. Duckworth, A. B. Baggeroer, Acoustic Array Sensor Tracking System, in *Proceedings of the MTS/IEEE Oceans '85 Conference*, San Diego, CA, November 1985.

Wang, F., Offshore Oil and Gas Development and Applications of Underwater Autonomous Vehicles, in *Proceedings of the Symposium on Unmanned Untethered Submersible Technology*, pp. 231-241, University of New Hampshire, Marine Systems Laboratory, June 1983.

Whiteside, A., M. Ellis, and B. Haskell, Digital Generation of Stereoscopic Perspective Scenes, in *Proceedings of the SPIE Conference on True Three-Dimensional Imaging Techniques and Display Technologies*, vol. 761, pp. 172-178, Los Angeles, CA, January 1987.

Winston, P. H., *Artificial Intelligence*, Addison-Wesley, Reading, MA, 1984.

Yoerger, D. R. and J.-J. E. Slotine, Task Resolved Motion Control of Vehicle-Manipulator Systems, *International Journal of Robotics and Automation*, vol. 2, no. 3, 1987.

Zadeh, L. A., Fuzzy Sets, *Information and Control* 8, pp. 338-353, 1965.

Zadeh, L. A., A Simple View of the Dempster-Shafer Theory of Evidence and Its Implication for the Rule of Combination, *AI Magazine*, vol. 7, no. 2, pp. 85-90, Summer 1986.

This blank page was inserted to preserve pagination.

CS-TR Scanning Project
Document Control Form

Date : 7/20/95

Report # AI-TR-1143

Each of the following should be identified by a checkmark:
Originating Department:

☒ Artificial Intellegence Laboratory (AI)
☐ Laboratory for Computer Science (LCS)

Document Type:

☒ Technical Report (TR) ☐ Technical Memo (TM)
☐ Other:_____

Document Information

Number of pages: 171 (179-IMAGES)
Not to include DOD forms, printer instructions, etc... original pages only.

Originals are:
☒ Single-sided or
☐ Double-sided

Intended to be printed as :
☐ Single-sided or
☒ Double-sided

Print type:
☐ Typewriter ☐ Offset Press ☒ Laser Print
☐ InkJet Printer ☐ Unknown ☐ Other:_____

Check each if included with document:

☒ DOD Form (2) ☐ Funding Agent Form ☒ Cover Page
☒ Spine ☐ Printers Notes ☐ Photo negatives
☐ Other: _____

Page Data:

Blank Pages (by page number): _____

Photographs/Tonal Material (by page number): 76, 78, 80, 82, 83, 85-90, 92-94, 96, 97, 102-104, 107, 108, 111-113, 116-120, 132, 133, 136-138, 140, 142-144,

Other (note description/page number):

Description :	Page Number:
(A) IMAGE MAP!	(1-171)
PAGE 2, UNNED ACKNOWLEDGEMENTS,	3-157, NO 158, 159-172
	(172-176)
SCANCONTROL, COVER, SPINE, DOD (2)	
	(177-179)
TARGET'S (3)	
(B) CUT & PASTE FIGS ON PAGES:	76, 78, 80, 82, 83, 85-90, 92-94, 96, 97, 102-104, 107, 108, 111-113, 116-120, 132, 133, 136-138, 140, 142-144,

Scanning Agent Signoff:

Date Received: 7/20/95 Date Scanned: 7/25/95 Date Returned: 7/27/95

Scanning Agent Signature: Michael W. Cook

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AI-TR 1143	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): Multisensor Modeling Underwater with Uncertain Information	5. TYPE OF REPORT & PERIOD COVERED: technical report
6. PERFORMING ORG. REPORT NUMBER	
7. AUTHOR(s): W. Kenneth Stewart, Jr.	8. CONTRACT OR GRANT NUMBER(s): N00014-86-K-0685
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Artificial Intelligence Laboratory, 545 Technology Square, Cambridge, MA 02139	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Advanced Research Projects Agency, 1400 Wilson Blvd., Arlington, VA 22209	12. REPORT DATE
	13. NUMBER OF PAGES: 172
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office): Office of Naval Research, Information Systems, Arlington, VA 22217	15. SECURITY CLASS. (of this report): UNCLASSIFIED
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution is unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

None

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

autonomous robot
object recognition
multisensor fusion
navigation
visual modeling

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This thesis develops an approach to the construction of multidimensional stochastic models for intelligent systems exploring an underwater environment. The important characteristics shared by such applications are: real-time constraints; unstructured, three-dimensional terrain; high-bandwidth sensors providing redundant, overlapping coverage; lack of prior knowledge about the environment; and inherent inaccuracy or ambiguity in sensing and interpretation. The models are cast as a three-dimensional spatial decomposition of stochastic, multisensor feature vectors that describe an underwater environment. Such models serve as intermediate descriptions that decouple low-level, high-bandwidth sensing from the higher-level, more asynchronous processes that extract information.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE S/N 0102-014-6601

UNCLASSIFIED (Cont. on back)

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Block 20 cont.

A numerical approach to incorporating new sensor information—*stochastic backprojection*—is derived from an incremental adaptation of the summation method for image reconstruction. Error and ambiguity are accounted for by blurring a spatial projection of remote-sensor data before combining it stochastically with the model. By exploiting the redundancy in high-bandwidth sensing, model certainty and resolution are enhanced as more data accumulate. In the case of three-dimensional profiling, the model converges to a "fuzzy" surface distribution from which a deterministic surface map is extracted.

Computer simulations demonstrate the properties of stochastic backprojection and stochastic models. Other simulations show that the stochastic model can be used directly for terrain-relative navigation. The method is applied to real sonar data sets from multibeam bathymetric surveying (Sea Beam), towed sidescan bathymetry (Sea MARC II), towed sidescan acoustic imagery (Sea MARC I & II), and high-resolution scanning sonar aboard a remotely operated vehicle. A multisensor application combines Sea Beam bathymetry and Sea MARC I intensity models. Targeted real-time applications include shipboard mapping and survey, a piloting aid for remotely operated vehicles and manned submersibles, and world modeling for autonomous vehicles.

Scanning Agent Identification Target

Scanning of this document was supported in part by the **Corporation for National Research Initiatives,** using funds from the **Advanced Research Projects Agency** of the **United states Government** under Grant: **MDA972-92-J1029.**

The scanning agent for this project was the **Document Services** department of the **M.I.T Libraries.** Technical support for this project was also provided by the **M.I.T. Laboratory for Computer Sciences.**

Scanned

Date: 7/25/1995

M.I.T. Libraries
Document Services