
# NAVAL POSTGRADUATE SCHOOL

**MONTEREY, CALIFORNIA**

# THESIS

**BOLTZMANN ENTROPY FOR UNDERSEA TERRAIN AIDED NAVIGATION**

by

Jason Cash

June 2020

| | |
|---|---|
| Thesis Advisor: | Douglas P. Horner |
| Second Reader: | Sean P. Kragelund |

THIS PAGE INTENTIONALLY LEFT BLANK

| REPORT DOCUMENTATION PAGE | | | Form Approved OMB No. 0704-0188 |
|---|---|---|---|



| **1. AGENCY USE ONLY** *(Leave blank)* | **2. REPORT DATE** June 2020 | **3. REPORT TYPE AND DATES COVERED** Master's thesis |
|---|---|---|

| **4. TITLE AND SUBTITLE** BOLTZMANN ENTROPY FOR UNDERSEA TERRAIN AIDED NAVIGATION | **5. FUNDING NUMBERS** |
|---|---|
| **6. AUTHOR(S)** Jason Cash | |
| **7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)** Naval Postgraduate School Monterey, CA 93943-5000 | **8. PERFORMING ORGANIZATION REPORT NUMBER** |
| **9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)** N/A | **10. SPONSORING / MONITORING AGENCY REPORT NUMBER** |

**11. SUPPLEMENTARY NOTES** The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

| **12a. DISTRIBUTION / AVAILABILITY STATEMENT** Approved for public release. Distribution is unlimited. | **12b. DISTRIBUTION CODE** A |
|---|---|

**13. ABSTRACT (maximum 200 words)**

The ability to navigate accurately without Global Positioning Systems (GPS) or beacon systems remains an elusive yet important capability for unmanned systems. Reliance on external beacon systems, like GPS and long baseline, create system vulnerabilities if those beacon systems are damaged or removed. This research explores the ability to improve Undersea Terrain Aided Navigation (UTAN) with respect to coverage path planning and creates a robust coverage path planning algorithm utilizing a new paradigm in configuration entropy. It highlights a classic challenge in the path planning problem between exploration and exploitation. Exploration ensures complete coverage of the region, while exploitation utilizes terrain features to reduce positional uncertainty of the Autonomous Underwater Vehicle (AUV). Using stochastic estimates of the vehicle location and map accuracy, it is possible to combine the variances of each together with exploration and exploitation policies to develop a coverage plan for UTAN. To do so requires development of a terrain entropy measure to influence the path planning algorithm. The thesis presents a new entropy measure that is a key architectural component for an information-theoretic framework to develop near-optimal paths for the AUV coverage problem.

| **14. SUBJECT TERMS** Undersea Terrain Aided Navigation, mapping, entropy | **15. NUMBER OF PAGES** 75 |
|---|---|
| | **16. PRICE CODE** |

| **17. SECURITY CLASSIFICATION OF REPORT** | **18. SECURITY CLASSIFICATION OF THIS PAGE** | **19. SECURITY CLASSIFICATION OF ABSTRACT** | **20. LIMITATION OF ABSTRACT** |
|---|---|---|---|
| Unclassified | Unclassified | Unclassified | UU |

NSN 7540-01-280-5500 Standard Form 298 (Rev. 2-89) Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

**Approved for public release. Distribution is unlimited.**

**BOLTZMANN ENTROPY FOR UNDERSEA TERRAIN AIDED NAVIGATION**

Jason Cash
Lieutenant, United States Navy
BSET, Old Dominion University, 2013

Submitted in partial fulfillment of the
requirements for the degree of

**MASTER OF SCIENCE IN MECHANICAL ENGINEERING**

from the

**NAVAL POSTGRADUATE SCHOOL**
**June 2020**

Approved by: Douglas P. Horner
Advisor

Sean P. Kragelund
Second Reader

Garth V. Hobson
Chair, Department of Mechanical and Aerospace Engineering

THIS PAGE INTENTIONALLY LEFT BLANK

# ABSTRACT

The ability to navigate accurately without Global Positioning Systems (GPS) or beacon systems remains an elusive yet important capability for unmanned systems. Reliance on external beacon systems, like GPS and long baseline, create system vulnerabilities if those beacon systems are damaged or removed. This research explores the ability to improve Undersea Terrain Aided Navigation (UTAN) with respect to coverage path planning and creates a robust coverage path planning algorithm utilizing a new paradigm in configuration entropy. It highlights a classic challenge in the path planning problem between exploration and exploitation. Exploration ensures complete coverage of the region, while exploitation utilizes terrain features to reduce positional uncertainty of the Autonomous Underwater Vehicle (AUV). Using stochastic estimates of the vehicle location and map accuracy, it is possible to combine the variances of each together with exploration and exploitation policies to develop a coverage plan for UTAN. To do so requires development of a terrain entropy measure to influence the path planning algorithm. The thesis presents a new entropy measure that is a key architectural component for an information-theoretic framework to develop near-optimal paths for the AUV coverage problem.

THIS PAGE INTENTIONALLY LEFT BLANK

# Table of Contents

# List of Figures

# List of Tables

THIS PAGE INTENTIONALLY LEFT BLANK

# List of Acronyms and Abbreviations

| | |
|---|---|
| **AUV** | Autonomous Underwater Vehicle |
| **ADCP** | Acoustic Doppler Current Profiler |
| **CAVR** | Center for Autonomous Vehicle Research |
| **CSUMB** | California State University Monterey Bay |
| **CTE** | Cross-Track Error |
| **DEM** | Digital Elevation Model |
| **DTM** | Digital Terrain Model |
| **DVL** | Doppler Velocity Log |
| **EKF** | Extended Kalman Filter |
| **FOV** | Field of View |
| **GPM** | Gaussian Process Model |
| **GPS** | Global Positioning System |
| **GIS** | Geographic Information Systems |
| **LBL** | Long Baseline |
| **LIDAR** | Light Detection and Ranging |
| **NPS** | Naval Postgraduate School |
| **PCA** | Principle Component Analysis |
| **PMF** | Probability Mass Function |
| **RRT*** | Rapidly-exploring Random Tree Star |

| | |
|---|---|
| **Sonar** | Sound Navigation and Ranging |
| **TAN** | Terrain-Aided Navigation |
| **TRI** | Topographic Ruggedness Index |
| **UAV** | Unmanned Aerial Vehicle |
| **UTAN** | Undersea Terrain-Aided Navigation |

# Acknowledgments

I would like to thank my advisor, Dr. Douglas Horner, for his insight, assistance, and direction provided throughout my research. Dr. Horner's guidance and expertise in this subject area were invaluable in refining my research and focusing my efforts to produce a novel and significant contribution to the field.

I would also like to thank Dr. Sean Kragelund for his feedback and assistance on my thesis. His advice was crucial in helping me clarify my writing and connect disparate concepts. My work with the Hamming supercomputer at NPS would not have been possible without Dr. Kragelund's guidance.

I would also like to express my appreciation to Ms. Teresa Cash for her feedback on my writing and providing a perspective from outside the field of Science, Technology, Engineering, and Mathematics (STEM).

I would also like to thank my wife, Rebekah Cash, for her continuous support throughout this process. I can say without hesitation that her steadfast devotion to our family is the reason I succeed.

Finally, I would like to thank my kids, Caleb, Isaac, Faith, Annabel and Emma, for making this thesis so hard to write. Because you are all such a joy to be with, I played more freeze tag than studying and kicked a soccer ball instead of reading. Thank you for reminding me of what's important in life.

THIS PAGE INTENTIONALLY LEFT BLANK

# CHAPTER 1:
# Introduction

## 1.1 Motivation

The ability to navigate autonomously without GPS or beacon systems remains an elusive yet important capability for the military [1]. Reliance on exterior beacon systems, like GPS and Long Baseline (LBL), create system vulnerabilities if those beacon systems are damaged or removed. Nowhere is this more apparent than in the undersea domain. This is because an Autonomous Underwater Vehicle (AUV) is unable to utilize GPS once submerged and beacon systems greatly restrict the operational capability and flexibility that are needed for the military.

Techniques in Undersea Terrain-Aided Navigation (UTAN) seek to accurately navigate without beacon references. It can be accomplished by correlating sensor data to a known map. This process allows the vehicle to accurately localize position relative to the map and reduce position error. Maintaining accurate position is important since the AUV error adds to onboard sensor error. There are three principle components to UTAN:

1. Sensor measurements and processing
2. AUV position estimation filter
3. Spatial estimation filter (Bathymetry map building)

Common UTAN sensors found on AUVs include Acoustic Doppler Current Profiler (ADCP) / Doppler Velocity Log (DVL), optical cameras, Light Detection and Ranging (LIDAR), and Sound Navigation and Ranging (Sonar). Sonar, however, is most popular because the optical and electromagnetic spectrum is greatly attenuated underwater. Before comparing the sensor data to the map, two steps are frequently necessary. First, the sonar data must be processed to obtain a set of discrete data points that are angular distance estimates of the ocean floor relative to the AUV's altitude. Collectively, this is known as a point cloud. Typically, image or signal processing techniques are used to generate point clouds from raw sensor data. The second step is to add an estimate of the AUV's depth. This is made more difficult due to surface waves. Figure 1.1, from [2], illustrates how sonar measurements

from a moving vehicle are collected to build bathymetry maps of an operating area. Figure 1.2, , from [2], shows an example of a BlueView 2250 micro bathymetry sample image overlaid with point clouds detected by computer vision image processing techniques.

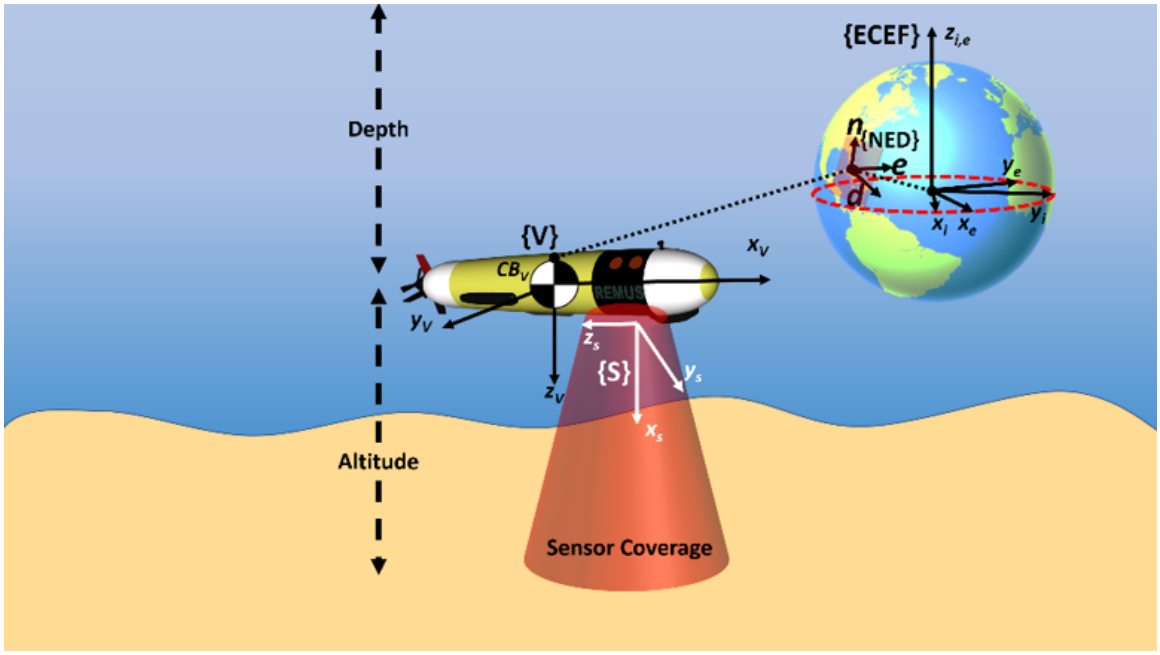



Figure 1.1. AUV data collection process

Optimal estimation theory provides a stochastic method for estimating the true value of a desired parameter from noisy sensor measurements. Optimal estimators typically compute the mean value and minimal variance estimate of the parameter. Historically, Terrain-Aided Navigation (TAN) applications used an Extended Kalman Filter (EKF) to estimate vehicle position. The EKF is a non-linear version of the Kalman filter that uses a Taylor series expansion for linearization. It assumes a Gaussian distribution of the vehicle position. More recently, it has been shown that TAN is a non-linear problem that produces non-Gaussian distributions. As a result, using an EKF for this task may produce a non-optimal result [3].

Particle filters are an established position estimation technique that are appropriate for non-linear, non-Gaussian filtering [4]. To establish a new position estimate, the following three items are provided to the particle filter:

1. A randomly distributed set of candidate poses.
2. The bathymetry map

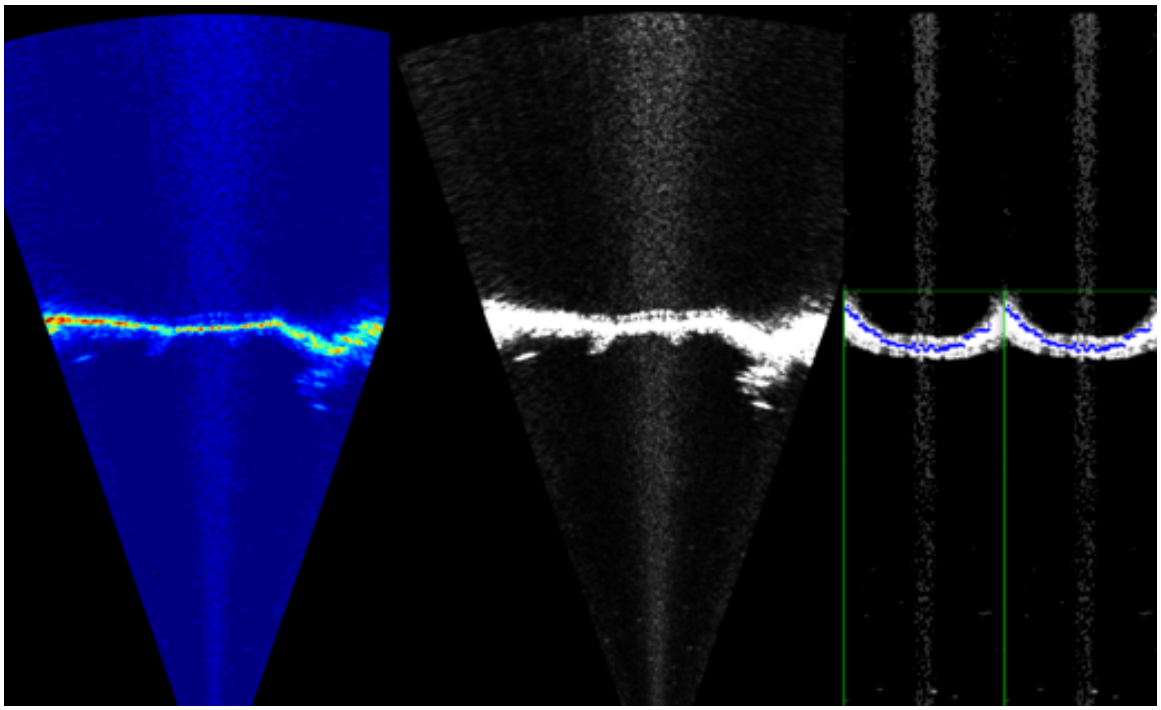


Figure 1.2. BlueView 2250 micro-bathymetry sample

3. Current prior position estimates

Each candidate pose is then used to produce a candidate sonar swath that is compared to the bathymetry map in order to determine the probability of correlation. As a result, the candidate position estimate with the highest probability of correlation is used as the updated position estimate for the AUV.

The third component of TAN is spatial estimation. This is the process of building and maintaining the map. Optimal Spatial Estimation is a process for estimating the terrain whereby the bounded search area is represented by a random field subdivided into gridded cells. Each cell has a gaussian mean and variance. This is known as a Gaussian Process Model (GPM) [5]. Calculation for the mean and variance is given by:

$$\bar{f}_* = k_*^T(k + \gamma_n^2 I)y \tag{1.1}$$

$$V[f_*] = k(x_*, x) - k_*^T (K + \sigma_n^2 I)^{-1} k \tag{1.2}$$

The first equation is the mean estimate for a cell. It is based on scaling the measurements (y) used in the cell estimation by a kernel function that represents the dissimilarity of the data as a function of distance. This is known as the semi-variance function.

The second equation is the variance calculation. It is based on the nugget of the semi-variance function, which is the dissimilarity of the data at distance zero, minus the scaled value between the point of estimation and the measurement by the semi-variance function.

Using a GPM is useful, in part, because the variance estimates of the cells provide a confidence evaluation of the map accuracy. Potential trajectories can be evaluated based on the reduction in the variance obtained by revisiting a given cell. Revisiting a cell is less advantageous when the reduction in variance decreases.

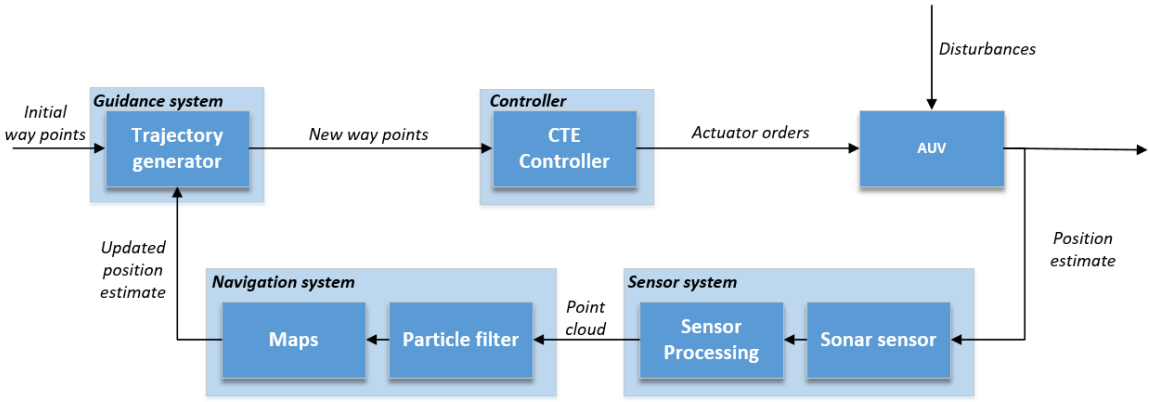



Figure 1.3. Process flow diagram representing the three principle components to UTAN

Figure 1.3 illustrates how the three principle components of UTAN are combined in a motion planning and localization framework to achieve complete sensor coverage of a region of interest. An initial set of waypoints are the default set of navigation objectives to fully cover a region. These points are iteratively replaced by new waypoints from the trajectory generator. A Cross-Track Error (CTE) controller computes a set of actuator commands for the rudder, propeller, and dive plane actuators to drive the AUV to the next waypoint. As the AUV moves through the medium, water currents and other disturbances induce error

into the vehicle's position estimate. Next, the sonar data is collected and processed. AUV position estimate and sensor data is then sent to the navigation system's particle filter to update the AUV position and bathymetry map. The updated map and AUV position are used by the guidance system to generate an updated trajectory for the AUV.

A common planning objective for coverage problems with uncertainty must balance between *exploration* and *exploitation*. *Exploration* is an AUV's ability to seek out unexplored regions in the map. This policy considers potential trajectories over a limited horizon, and evaluates the reward obtained for each of these trajectories. Over time (energy, distance), the goal of exploration is the complete coverage of a bounded region.

*Exploitation* is the ability to identify and utilize distinguishing features within the underwater terrain for the purpose of position estimation. This policy effectively manages the vehicle error by providing a stronger incentive (or reward) as the vehicle position error grows. Weighting functions can emphasize exploration or exploitation to effectively define a policy for determining the preferred trajectory that is dependent on the state of the AUV. If the AUV has a good position estimate (with small error), it can emphasize exploration in the planning process. Conversely, if the position estimate error exceeds a predetermined threshold, exploitation policy is emphasized and a new trajectory is calculated.

Key to this process is the ability to identify features that can be used by the AUV to reduce positional uncertainty through correlation with the particle filtering process. This thesis focuses on the ability of the vehicle to exploit the terrain. It introduces a new paradigm based on entropy as a measure of terrain uncertainty for determining features which permit the AUV to potentially navigate more accurately. The greater the feature's entropy, the greater the probability that the sonar data will successfully correlate to the feature in the bathymetry map. More specifically, we will introduce Boltzmann entropy. Traditionally, Boltzmann entropy has been used as a way of measuring uncertainty within gas particles in the field of thermodynamics [6]. It relates entropy to the number of potential variations within a sample. For UTAN, we will show how increased variation within the terrain results in higher entropy values that can be used for exploitation.

## 1.2 Background

Technical advances in UTAN have been numerous over the past decade. They have primarily been driven by significant increases in computer processing speeds and primary data storage. As a result, computationally expensive methods, like particle filters, were introduced as methods well-suited to handle non-linear processes. In the mid 2000s, they were introduced in conjunction with Kalman filters as an approach capable of producing good results [7]. This allowed the particle filters to handle the non-linear components and the Kalman filter to handle the linear components of the estimation problem. Several other published papers employ particle filters for non-linear processes instead of Kalman-based methods, namely [8], [9], and [10].

In 2006, Anonsen and Hallingstand [11] revealed concerns in map error due to surface waves inducing uncertainty into the AUV depth estimation. The authors introduced 2-dimensional and 3-dimensional solutions using point mass filters and Bayesian bootstrap particle filters that are well-suited for UTAN. A significant issue revealed, but not addressed, was the notion of a terrain's "usefulness". A suitable terrain will exhibit changes in contouring which allows for the AUV to more easily correlate it's position. The authors were able to demonstrate that the vehicle's positional uncertainty was minimized in suitable terrain. However, unsuitable terrain, which shows no changes in contouring, produced poor results which lead to filter divergence.

In 2012, Shane Dektor and Professor Stephen Rock of Stanford University addressed the limitations of an AUV covering flat terrain. The authors suggested that current position estimation techniques were the cause for false fixes when operating in flat terrain. As a result, an improvement in filter measurement weighting proved to be a successful approach in reducing error for both flat and dynamic environments [12].

In a paper focusing on landscape ecology, Samuel Cushman provided an excellent method for quantifying the information within a terrain map using Boltzmann entropy [13]. The author made parallels between the dynamics of a landscape and the mixing of particles associated with an ideal gas. In 2017, Boltzmann entropy was used to calculate the configurational entropy in several landscapes [14]. The author noted that an acceptable macrostate is challenging to define, and that the associated microstates are computationally difficult, even if the macrostate is well-defined.

## 1.3 Problem Statement

A common problem in coverage planning problems is finding a balance between exploration and exploitation [15]. Optimal solutions to a finite horizon coverage planning problem involve a sequence of policies $(\pi_1^*, \pi_2^* ... \pi_H^*)$ such that decisions are made to maximize the expected sum of rewards. The optimal solution can be characterized by the following equations,

$$Q_t(b,u) = \rho(b,u) + \gamma \int_{z' \in Z} \eta(z^{'}|b,u) V_{t-1}^*(\tau(b,u,z^{'}))dz^{'} \tag{1.3}$$

$$V_t^*(b) = \sup_{u \in U} Q_r(b,u) \tag{1.4}$$

$$\pi_t^*(b) = \arg\sup_{u \in U} Q_t(b,u) \tag{1.5}$$

where $u$ is an action, $b$ is a belief state, the function $\tau$ is the belief update function, $\eta$ is the prior probability of observing an observation $(z^{'})$, $V^*$ is a value function which calculates the expected rewards for acting optimally, $\gamma \in [0,1]$ is a discounted reward function and $\rho(b,u)$ is an initial condition at $t = 1$. [16].

Policies are defined by mapping belief states to actions. For the AUV, actions are defined in the coverage problem in terms of vehicle dynamics, sensor characteristics, and the environment. The vehicle's position, orientation, and velocity are quantified within the state vector (see Equation 1.6) and are commonly referred to as the vehicle's pose. Vehicle states include:

1. Position (x,y,z)
2. Velocity (u,v,w)
3. Orientation $(\phi, \theta, \psi)$
4. Angular velocity (p,q,r)

and evolve in time according to the vehicle's dynamic equations of motion [17]. In a stochastic system, the state is commonly represented by a mean and variance of each

variable. The covariance matrix (Equation 1.7) is the covariance between the $i^{th}$ and $j^{th}$ elements of the state vector. It plays a key role in path planning because different candidate trajectories can be evaluated based on the anticipated increase (or decrease) in positional uncertainty. Mathematically, this can be assessed by the sum of the covariance along the diagonal.

$$\bar{x} = \begin{bmatrix} x & y & z & u & v & w & \phi & \theta & \psi & p & q & r \end{bmatrix}' \tag{1.6}$$

$$cov = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} & \dots & \sigma_{xr} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} & \dots & \sigma_{yr} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \sigma_{rx} & \sigma_{ry} & \sigma_{rz} & \dots & \sigma_{rr} \end{bmatrix} \tag{1.7}$$

Sonar system characteristics (beamwidth, field of view, frequency, and range) also constrain the coverage problem. By increasing the altitude of the vehicle, the ensonified area of the sea floor will grow. This results in a reduction of sensor resolution and quality. Conversely, if the altitude is too low, the Field of View (FOV) narrows and is not optimal. Sonar characteristics will define the proper combination of altitude and resolution to ensure proper coverage.

The environment can play a major role in sonar performance and vehicle state. This is of particular interest because the environment is not a constant. The consequences of this observation are that sonar and vehicle performance will vary along a trajectory. The varying of both the sensor and vehicle state bring an increased uncertainty about the variance of the region being mapped. We seek to reduce this uncertainty by planning a trajectory which can exploit available information in the terrain.

Energy and time are constraints for the optimal coverage problem. This practical limitation is directly related to the distance the vehicle can travel, and bounds the size of the region that can be mapped. Although not addressed in this thesis, optimizing with respect to energy and/or time for UTAN coverage problems is a recommended area for future work.

While it is assumed, for this work, there exists an accurate bathymetry map, the framework we are developing will allow for the degradation of this assumption in the future. This map

is defined by both resolution and location. Ultimately, we are describing an optimization problem that minimizes total vehicle uncertainty over a finite amount of time or energy given limitations in sensing and vehicle dynamics.

## 1.4 System Description

The UTAN architecture is being designed for the Naval Postgraduate School (NPS) Center for Autonomous Vehicle Research (CAVR) REMUS 100 (Pictured in figure 1.4, from [18]). This vehicle is designed to operate in coastal areas at depths no greater than 100 meters. The AUV weights approximately 68 kg and has a diameter of 19 cm. Because of the vehicle's modular design, it can be configured to support a wide array of missions. The REMUS 100 is configured with a modular BlueView MB2250 downward looking sonar. It provides ultra-high resolution imaging that operates at 2250 KHz. Table 1.1 provides a summary of the BlueView MB2250 characteristics. Post processed sonar data runs at an approximate rate of 2 Hz.


Figure 1.4. REMUS 100 AUV onboard CAVR's SeaFox surface vessel

Table 1.1. BlueView MB2250 sonar specifications

| **Attribute** | **Value** |
|---|---|
| Field of View | 45° x 1° |
| Minimum Range | 0.5 meters |
| Maximum Range | 10 meters |
| Beam Width | 1° x 1° |
| Number of Beams | 256 |
| Beam Spacing | 0.18 meters |
| Time Resolution | 0.01 meters |
| Max Update Rate | 40 Hz |
| Frequency | 2.25 MHz |

## 1.5 Overview

The remainder of the thesis is organized as follows: In Chapter 2 probability theory is briefly introduced and past entropy measures are introduced and shown why these measures are insufficient for terrain analysis. In Chapter 3, Boltzmann entropy is introduced as it relates to landscape ecology. For Chapter 4 the Boltzmann entropy approach is demonstrated with respect to AUV path planning. Finally in the concluding chapter, open areas of research are presented and a summarize of results.

# CHAPTER 2:
# Review Of Terrain Metrics

The last chapter introduced a framework for conducting TAN that is capable of real-time path planning which potentially reduce a priori requirements for accurate maps. It involves stochastic processes for position estimation and map building. We described how an underwater survey mission can be formulated as a coverage problem that must balance exploration of unknown areas with exploitation of known features via TAN. The relative weighting between these two goals can be adjusted to develop motion planning policies which balance navigational accuracy and sensor coverage.

This thesis focuses on exploitation. In order to reduce vehicle localization error, we seek a tool to inform the planning process about terrain features that can be exploited by the particle filter correlation process to localize the vehicle. We call this process finding "information in the terrain". It is the ability to evaluate the bathymetry for regions where successful sensor measurements-to-map correlations are more likely to occur. In this chapter, different information metrics are presented that have historically been used for TAN and for related tasks in computer vision and landscape ecology. Central to the discussion is entropy. This chapter will briefly introduce this concept, and demonstrate why common measures of entropy are considered inadequate for this task.

## 2.1 Terrain Metrics

Topographic Ruggedness Index (TRI) (Riley et al [19]) is commonly used by Geographic Information Systems (GIS). It expresses the amount of elevation difference between adjacent cells in a Digital Elevation Model (DEM). It calculates the difference in elevation values by taking the difference between a center cell and its eight surrounding neighbor cells. Next, these eight differences are squared and averaged together. Finally, the TRI is computed as the square root of the average. The computational expense for this method is low. However, through averaging the eight differences, configurational data is lost. As a result, many terrain variations can produce the same TRI

Another measure called *rugosity index* quantifies terrain complexity as the ratio between

the area of a contoured surface (A) and the orthogonal projection of the surface (A'). The orthogonal projection is calculated using a tesselated triangular mesh of connected vertices to approximate the three-dimensional surface. Selecting the correct plane for the surface to project on is important, since it separates the rugosity index from slope. To accomplish this, Principle Component Analysis (PCA) is used to maximize the variance of the projected data [20].

Higher values of rugosity index indicate more dynamic terrain. It has traditionally been measured *in-situ* with divers in shallow water (less than 30-meters) using a methodology known as chain-tape [21]. Today, these measurements are being performed by AUVs with high resolution stereo cameras. In [22], the authors extrapolate two other metrics from within rugosity: slope and aspect. Slope is the angle between the horizontal plane of the ocean floor and the plane of best fit that resides at the peak of the terrain feature. Aspect projects vector arrows to indicate slope direction when viewing the feature from a particular face. Although these techniques provide good results in underwater terrain survey applications, the rugosity index is ill-suited as a navigation metric because it does not provide any details about the configuration of terrain features. As a result, terrain features with different characteristics can produce the same rugosity index value.

## 2.2 Entropy

In 1948, Claude E. Shannon introduced a revolutionary concept in his seminal paper "The Mathematical Theory of Communication" [23]. It provided a mathematical framework for quantifying information in a radio signal. This quantification revolved around the concept of uncertainty, which is based on complexity in a signal. Shannon called this uncertainty-based measure of information "entropy" [24]. Entropy is a way of measuring the variations within a sample. Its applications span a multitude of disciplines and are mathematically defined in various ways. Some applications include biodiversity over animal populations [25], communication signal processing [23], and gas particle kinematics [26]. Shannon entropy is defined by Equation 2.1,

$$H = -\sum_{i}^{N} p_i log_2(p_i) \tag{2.1}$$

(a) H=7.3655 (b) H=7.6546

Figure 2.1. Comparison of entropy values

where $p_i$ is the probability of a specific data value in a stream or batch of data.

There are two related fields that each use entropy measures. The first is for computer vision. It starts with an imaging system, typically a camera, that captures an image through a lens. The image is comprised of numerous pixels that store values. In the case of a grey-scaled image, 8-bit pixel values range between zero and 255; zero for white and 255 for black. Each pixel has a probability distribution based on the total number of pixels that contain the same value within the image. If the probability distribution is equal across all pixel values, then the uncertainty about that image is greatest. Conversely, probability distributions that have increased variation also contain greater certainty. Figure 2.1 displays a self portrait used in computer vision calculations. Entropy (H) for the two images was calculated using the entropy function from MATLAB's Image Processing Toolbox [27].

Figure 2.1(a) is a darker representation of the image. This results in less variation between the pixels and ultimately a lower entropy value. Conversely, Figure 2.1(b) has not been darkened and has increased variation. This results in a greater entropy value.

In [28], this approach was used to determine the likelihood of a mudslide occurring. By capturing images with an Unmanned Aerial Vehicle (UAV), the authors were able to characterize high information with sloped terrain and low information with flat terrain.

Razlighi et al [29] conducted a comparative analysis of spatial entropy measures used in image processing. The purpose was to identify how different spatial entropy measures mitigated the high dimensionality of the Probability Mass Function (PMF) used to calculate Shannon entropy (Equation 2.1). One the compared entropy measures is *generalized Ising entropy*. It is defined by the following equation

$$H_I = -\sum_{x_{i,j} \in x} p(x_{i,j}/N_{i,j}) log(p(x_{i,j}/N_{i,j})) \tag{2.2}$$

where $H_I$ is the Ising entropy value, $p$ is the joint probability, $x_{i,j}$ is an internal value and $N_{i,j}$ are the four neighbors associated with the internal value. This method limits the size of the PMF by only evaluating four neighbors surrounding the internal value. However, this method is computationally challenging because of the complexity involved in solving for $p(x_{i,j})$.

Alfréd Rényi generalized Shannon's original definition for entropy using the following equation:

$$H_\propto = \frac{ln(\sum_i P_i^\propto)}{1 - \propto} \tag{2.3}$$

where $P$ is the probability, and $\propto$ is a dynamic variable that is changed to achieve the different entropy measures examined in this chapter [30].

This thesis will examine three entropy measures derived from Rényi's generalized entropy measure: Shannon, Simpson, and Berger Parker. Specifically, Shannon is derived when $\propto \rightarrow 1$, Simpson is derived when $\propto \rightarrow 2$, and Berger Parker is derived when $\propto \rightarrow \infty$) [30]. In Sections 2.2.1, 2.2.2 and 2.2.3, each entropy measure will be examined to understand its limitations, and why a new paradigm in Chapter 3 is introduced.

### 2.2.1 Shannon Entropy

Shannon Entropy can be simply demonstrated by posing a question that has four possible outcomes (A, B, C, and D). If the probability of each outcome is 25%, then the Shannon entropy is calculated as follows:

$$H = -(0.25 * log_2(0.25) + 0.25 * log_2(0.25) + 0.25 * log_2(0.25) + 0.25 * log_2(0.25)) = 2$$

However, note that by changing the probability distribution of the four possible outcomes to 0.2, 0.4, 0.2, and 0.2, respectively, decreases the entropy.

$$H = -(0.2 * log_2(0.2) + 0.2 * log_2(0.2) + 0.4 * log_2(0.4) + 0.2 * log_2(0.2)) = 1.92$$

This results from reduced uncertainty about potential outcomes.

This same principle has been applied to many other disciplines. Shannon applied it to radio signals. In this thesis, entropy is used to quantify the information contained in bathymetry maps of the seafloor. To gain a better understanding of how the quantity of information varies within a small patch of terrain, consider three examples. These examples represent *flat*, *sloped*, and *fully unique* landscapes, respectively. Flat terrain is defined by having no variation in elevation. Sloped terrain is defined by having uniformly increasing or decreasing elevation values. And fully unique terrain has no repeated elevation values in the entire landscape. The following matrices will represent these three different landscapes:

$$Flat\ Example = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$Sloped\ Example = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix}$$

$$Fully\ Unique\ Example = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

Starting with the *flat* example, it is easy to see that the probability of achieving a 1 is 100%. As a result, $H = 0$. This is because no information can be extracted from the landscape. Once a single value in the matrix is known, there is nothing more that can be learned about the landscape. Next, the probability of achieving a 1, 2, or 3 in the *sloped* example is $\frac{1}{3}$. As a result,

$$H = -(\frac{1}{3} * log_2(\frac{1}{3}) + \frac{1}{3} * log_2(\frac{1}{3}) + \frac{1}{3} * log_2(\frac{1}{3})) = 1.585$$

This indicates that more information can be extracted from the sloped example than from the *flat* example. Finally, the *fully unique* example has a $\frac{1}{9}$ probability of achieving any number. As a result,

$$H = -9 * (\frac{1}{9} * log_2(\frac{1}{9})) = 3.17$$

Once again, the amount of information available in fully unique terrain has increased. These examples imply that there is a correlation between the uniqueness of a landscape and the amount of information it contains.

Unfortunately, many different arrangements of elevation data in a landscape produce the same entropy value. Therefore, Shannon entropy does not capture the uniqueness of a given landscape and is not a useful metric for UTAN. For example, the elevation values in the *sloped* example can be rearranged to any other configuration and this entropy measure will remain the same, $H = 1.585$.

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix} \equiv \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix} \equiv \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 2 & 1 & 3 \end{bmatrix} \equiv \begin{bmatrix} 1 & 3 & 3 \\ 2 & 2 & 3 \\ 1 & 1 & 2 \end{bmatrix} \equiv \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 3 \end{bmatrix}$$

This is a concern when trying to determine the dynamics of a landscape, since Shannon entropy does not account for all of the different configurations associated with $H = 1.585$.

As a result, landscapes can be misrepresented and maps that are inaccurate can be produced. Over the decades since Shannon wrote his seminar paper, several other probability-based measures of entropy have been introduced.

### 2.2.2 Simpson Entropy

Another entropy measure used to assess the diversity of an area was discovered by Simpson. Simpson's measure of entropy is commonly used to produce diversity profiles in the natural world. This measure quantifies the probability of finding particular animals in an area. Simpson Entropy is calculated using empirical methods and a robust theoretical framework [25]. It is also known as a dominant index because it puts a greater emphasis on common species, while rare species can have little to no effect on diversity. This is of concern not only for zoologists studying biodiversity, but also for our study where limited key features can be masked and go undetected. Equation 2.4 is known as Simpson's entropy.

$$H = \sum_{i=1}^{s} p_i^2 \tag{2.4}$$

First, we relate it to biodiversity where $s$ is the number of variations in the sample, and $p_i$ is the fraction of individuals ($n$) divided by the total number of individuals ($N$). This is best explained by an example using the biodiversity of birds as listed in Table 2.1.

Table 2.1. Example data for Simpson entropy

| | Robins | Cardinals | Blue Jays |
|---|---|---|---|
| individuals (n) | 10 | 2 | 8 |

Applying the values from Table 2.1 into Equation 2.4, we get the following result:

$$H = (\frac{10\ robins}{20\ total})^2 + (\frac{2\ cardinals}{20\ total})^2 + (\frac{8\ bluejays}{20\ total})^2 = 0.9$$

Another common practice in biodiversity is to express these entropy values in the Simpson Index form. This is done by taking the inverse of the entropy value. The reason for doing this is quite practical; it aligns the dynamics of the entropy value with the index value so

that as diversity increases so does the index [31]. As a result, the Simpson Index value in our example is calculated to be 1.11 using equation 2.5.

$$H_{index} = \frac{1}{H} \tag{2.5}$$

Now, to relate this to landscape ecology, we need only recognize the similarities between an individual species and an individual bathymetric sensor measurement. With that in mind, $n$ can be related to the number of bathymetry measurements with the same value, and $N$ can be related to the total number of bathymetric measurements in the region of interest. Once each $p_i$ value is determined, they will be summed together to solve for entropy. In addition, the inverse will be taken to show that higher entropy values correlate to greater uncertainty in the landscape. Next we will conduct an example using the same three landscapes as in Section 2.2.1.

$$Flat\ Example = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$Sloped\ Example = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix}$$

$$Fully\ Unique\ Example = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

Starting with the *flat* example, $n = 1$ and $N = 1$. As a result, we can clearly see that $H = 1$ and $H_{index} = 1$. This is because there is no diversity of landscape in the *flat* example. Next, the *sloped* example provides more diversity than the *flat* example.

$$H = (\frac{3}{9})^2 + (\frac{3}{9})^2 + (\frac{3}{9})^2 = \frac{1}{3}$$

$$H_{index} = \frac{1}{\frac{1}{3}} = 3$$

Comparing the *flat* example and the *sloped* example we can begin to see the impact that landscape diversity has on the Simpson Index value. Next, we calculate entropy and index values for the *fully unique* example.

$$H = (\frac{1}{9})^2 + (\frac{1}{9})^2 + (\frac{1}{9})^2 + (\frac{1}{9})^2 + (\frac{1}{9})^2 + (\frac{1}{9})^2 + (\frac{1}{9})^2 + (\frac{1}{9})^2 + (\frac{1}{9})^2 = \frac{1}{9}$$

$$H_{index} = \frac{1}{\frac{1}{9}} = 9$$

From these three examples we can see how the diversity within a landscape impacts its entropy and index values. Higher index values could be used to indicate key features in a landscape. However, these values are based in probability and exhibit the same concerns as Shannon entropy. Furthermore, key landscape features can get overshadowed by a majority of low diversity features. To illustrate this, consider the example of a landscape with one prominent key feature.

$$Key\ Feature\ Example = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 9 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$H = (\frac{8}{9})^2 + (\frac{1}{9})^2 = \frac{65}{81}$$

$$H_{index} = \frac{1}{\frac{65}{81}} = 1.25$$

We can see from the *key feature* example that the index value is very low even though a significant landscape characteristic exists. We therefore conclude that Simpson entropy is not a suitable method for determining key features in landscape ecology. The following

section will consider another entropy measure which uses a different approach.

### 2.2.3 Berger Parker Entropy

The Berger Parker entropy measure is another way of calculating the diversity of an area. Specifically, this measure is used to quantify biodiversity and only addresses the most common species in an area. This narrow approach means that certain information will be excluded from calculations. Recall that Equation 2.3 defines the foundation for the Berger Parker entropy measure. However, letting $\propto \rightarrow \infty$ produces the new approximation in Equation 2.6,

$$H_{\infty} = \frac{n_{max}}{N} \tag{2.6}$$

where $n_{max}$ is the most prevalent species in an area, and $N$ is the total number of species. This approach shares many similarities with Simpson's entropy, but it only accounts for the most dominant species [32]. It is natural to question the utility of an entropy measure that is less informative about landscape dynamics, especially since it neglects potentially significant features. How could this measure point to valuable information? Can it help identify areas of low information? We will again demonstrate this problem using our *flat*, *sloped*, and *fully unique* landscapes as examples.

$$Flat\ Example = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$Sloped\ Example = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix}$$

$$Fully\ Unique\ Example = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

Starting with the *flat* example, we determine that $n_{max} = 9$ and $N = 9$. Also, by using

equation 2.6, we find that $H_\infty = 1$. Now, let's look at the *sloped* example, where $n_{max} = 3$ and $N = 9$. This results in an entropy value of $\frac{1}{3}$. If we compare this to the flat example, we can see that the entropy value is decreasing with variability. Next, the *fully unique* example has $n_{max} = 1$ and $N = 9$. We can clearly see that $H_\infty = \frac{1}{9}$ and that the decreasing trend continues to hold true. This entropy measure is certainly attributing higher values of entropy to lower levels of information. Because of this inverse relationship, many biodiversity researchers use the Berger Parker Index to determine variability. The index is calculated by taking the inverse of the entropy value. As a result, lower index values indicate lower variability in an area.

From these three examples we can see how the lack of diversity in a landscape impacts the entropy and index values. Lower index values could be used to indicate areas of low information in a landscape. Like Shannon and Simpson entropy, Berger Parker entropy is also based in probability and raises similar concerns for use in UTAN. Whereas Shannon and Simpson entropy attempt to quantify all aspects of a landscape, however, Berger Parker entropy addresses areas of low variability. This may be useful for UTAN by potentially indicating areas to avoid.

## 2.3 Summary

After evaluating three separate entropy measures that address variability and diversity, it is clear that probability-based entropy measures are an insufficient metric for UTAN. However, with the increase of computing capabilities in recent years, a new paradigm in the use of configurational-based entropy may be possible. Chapter 3 will address this new paradigm and assess the advantages and disadvantages of this approach.

THIS PAGE INTENTIONALLY LEFT BLANK

# CHAPTER 3:
# Configurational Entropy of Underwater Terrain

The last chapter provided different examples of metrics commonly used for quantifying information. These metrics were shown to be insufficient for UTAN. In this chapter, we introduce configurational entropy as a way of quantifying the information within a landscape. In 1877, Ludwig Boltzmann discovered a relationship between entropy and the total number of variations in a sample [33]. Boltzmann also discovered that his approach was general enough to be used in a number of disciplines. Initially, this approach was used in thermodynamics as a way of measuring the variability in a sample of molecules. To accomplish this, Boltzmann made two assumptions about the sample:

1. The sample is in thermodynamic equilibrium.
2. There is no knowledge about the potential arrangements of molecules within a sample.

First, the term *thermodynamic equilibrium* means the sample under consideration does not change over time. This means that no new molecules are entering the sample and no current molecules are departing. Second, by assuming that no knowledge exists about the potential arrangements in the sample, we conclude that the probability distribution for each molecule is the same. As a result, each potential arrangement of the molecules in the sample is equally likely.

These assumptions place limits on how Boltzmann entropy can be applied to UTAN. The first assumption is only valid when the terrain being mapped is not changing during the coverage problem. This could pose a problem if the region being mapped contains terrain that is shifting due to underwater volcanic activity, underwater mudslides, or dredging operations. The second assumption implies that all potential bathymetry configurations within a patch of terrain are equally possible. In this thesis, we assume that this is the case; consideration of different probability distributions is an area for future research.

## 3.1 Macro- and Microstates

A macrostate is a set of three parameters that define the boundaries of the microstate. Specifically, they represent the sum and variance of the bathymetry measurements within the candidate terrain. These numerical values include:

1. The sum of all bathymetry values.
2. The minimum bathymetry value.
3. The maximum bathymetry value.

For example, if a received sonar swath, $S$, contains bathymetry values, $S = [1, 2, 3, 4]$, then the macrostate of $S$ is evaluated to be $\sum S_i = 10$, $min(S) = 1$ and $max(S) = 4$. After determining the macrostate, we can now determine all of the multisets. A *multiset* is a set of bathymetry values whose sum is equal to the sum of the original set of bathymetry values. In addition, the values within the multiset can not exceed the minimum and maximum values of the original set of bathymetry values. The microstate is made up of all the possible permutations within a multiset.

To determine the total number of permutations in a multiset with four variables, the following possibilities exist:

$$M_i = \begin{cases} 1 & \text{if } U = 1 \\ 4!/3! & \text{if } U = 2, B = 3 \\ 4!/(2! * 2!) & \text{if } U = 2, B \neq 3 \\ 4!/2! & \text{if } U = 3 \\ 4! & \text{if } U = 4 \end{cases} \tag{3.1}$$

Where $M$ is the number of permutations, $U$ is the number of unique values in the multiset, and $B$ refers to the case where each multiset has the same number of unique values, but their arrangements are different. As an example, both $\{1, \ 1, \ 3, \ 3\}$ and $\{1, \ 3, \ 3, \ 3\}$ have $U = 2$. The first set has a uniqueness broken up by 2/2 and the second set by 1/3. This results in a $B$ value of 2 and 3, respectively. The distinction is important because the number of permutations cannot be determined solely by the quantity of distinct values. Table 3.1 illustrates the idea of macrostates, multisets and microstates.

Table 3.1. Example of a macrostate, multisets, and microstates of an original set of bathymetry values

| **Original bathymetry values** | **Macrostate** | **Multiset** | **Microstates (Permutations)** | | |
|---|---|---|---|---|---|
| [1,2,3,4] | *sum* = 10 | {1,3,3,3} | (1,3,3,3) | (3,1,3,3) | |
| | | | (3,3,1,3) | (3,3,3,1) | |
| | *min* = 1 | | | | |
| | *max* = 4 | {2,2,3,3} | (2,2,3,3) | (2,3,2,3) | (2,3,3,2) |
| | | | (3,2,3,2) | (3,3,2,2) | (3,2,2,3) |

Next, we sum the total quantity of permutations, $M$, for each multiset. This results in the total number of permutations for all multisets in the terrain, and allows for the quantification of all configurational uncertainty. Specifically, this is the key value used in building our Boltzmann entropy map. It is given by the following formula:

$$W_u = \sum_{i=1}^{k} M_i \tag{3.2}$$

Lastly, if we desire to quantify the entropy within the entire map, the product of the summed multisets is calculated.

$$W = \prod_{i=1}^{k} W_{u,i} \tag{3.3}$$

Now, we introduce Boltzmann entropy in Equation 3.4,

$$S = k_b log(W) \tag{3.4}$$

where $k_b$ is known as the Boltzmann constant, and $W$ is the number of possible microstates within a macrostate. The Boltzmann constant relates the average kinetic energy of particles in a gas with the temperature of that gas. In [13], the Boltzmann constant is said to be a

value that "scales the relationship to the ideal gas laws." Because the process of calculating entropy in a landscape does not require accounting for an ideal gas, the need to utilize the Boltzmann constant is not required [13]. As a result, equation 3.4 can be simplified to:

$$S_{land} = log(W) \tag{3.5}$$

Where $S_{land}$ is the entropy for the landscape and $W$ is the total number of microstates in a landscape.

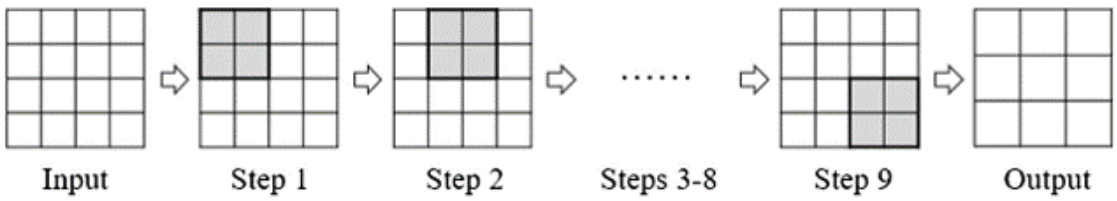
## 3.2 Map resolution

The values associated within a terrain are topographical and provide contour dynamics. To produce a more abstract representation of the terrain, the bathymetry map is broken into a grid-based Digital Terrain Model (DTM) [34]. Once the grids are formed, their resolution can be reduced. In [14], the reduction is accomplished by taking a sliding $2X2$ window and averaging each $2X2$ section one row at a time until each column is filtered. This reduces the size of the matrix by one row and one column. In this thesis, we apply the same approach to building lower-resolution maps. The following is an example of this process:

Given the simple $4X4$ array,

$$Example\ Map = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 8 & 7 & 6 & 5 \\ 4 & 3 & 2 & 1 \end{bmatrix}$$

it can be decomposed into nine smaller $2X2$ matrices by applying a sliding window.



These nine smaller matrices can then be averaged together.

$$Example\ Map = \left|\begin{array}{ccc} \frac{14}{4} & \frac{18}{4} & \frac{22}{4} \\ \hdashline \frac{26}{4} & \frac{26}{4} & \frac{26}{4} \\ \hdashline \frac{22}{4} & \frac{18}{4} & \frac{14}{4} \end{array}\right|$$

The final result is a single $3X3$ matrix that has lower resolution.

$$Example\ Map = \begin{bmatrix} 3.5 & 4.5 & 5.5 \\ 6.5 & 6.5 & 6.5 \\ 5.5 & 4.5 & 3.5 \end{bmatrix}$$

Figure 3.1 shows the resolution reduction process being applied to a bathymetry dataset from Monterey Bay.

## 3.3 The Experiment

In this thesis, we use bathymetry data provided by California State University Monterey Bay (CSUMB), (http://seafloor.otterlabs.org/SFMLwebDATA_mb.htm). Our goal in acquiring this map is to understand how configurational entropy changes based on underwater terrain. To accomplish this, we extracted three separate 256 $m^2$ maps from within the Monterey Bay each having different terrain complexity. We then calculate the entropies of each map and conduct a comparative analysis to better understand Boltzmann entropy.

### 3.3.1 Map building

Figure 3.2 provides the geographical locations of the three maps comprised of one within the Monterey Bay. The first map (Figure 3.3) has relatively flat bathymetry throughout the landscape. The second map (Figure 3.4) contains both sloped and flat bathymetry. The third map (Figure 3.5) has bathymetry characteristics that are continuously sloped throughout the map. Once the data point locations and values for each map are determined, different resolutions are created as described in Section 3.2.

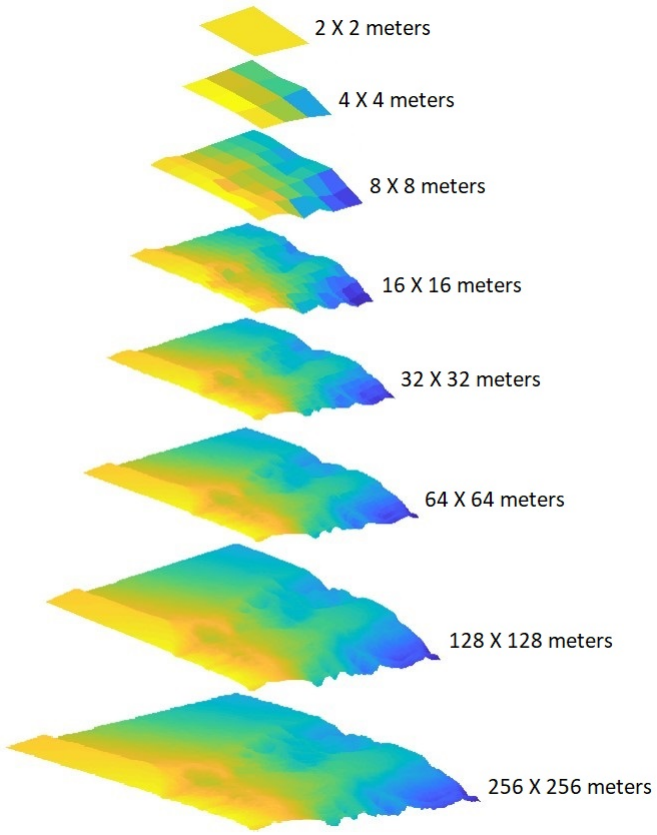



Figure 3.1. Example of different resolution levels in a landscape

### 3.3.2 Absolute and Relative Entropy

The amount of information contained in a map is a function of its resolution. When the resolution of a map is degraded, some information is lost. As a result, the entropy of a map can be expressed in two ways. This concept is addressed in [14], which separates these two paradigms into *relative* and *absolute* entropy. Relative entropy expresses the entropy within a single map resolution, whereas absolute entropy looks at the total entropy across all resolutions. Equation 3.6 shows how absolute entropy is calculated,

$$S_A = \sum_{1}^{t} = S_R(i) \tag{3.6}$$

where $S_A$ is the absolute entropy, $t$ is the number of resolutions, and $S_R$ is the relative entropy calculated from Equation 3.5. Table 3.2 lists the relative entropy computed at each

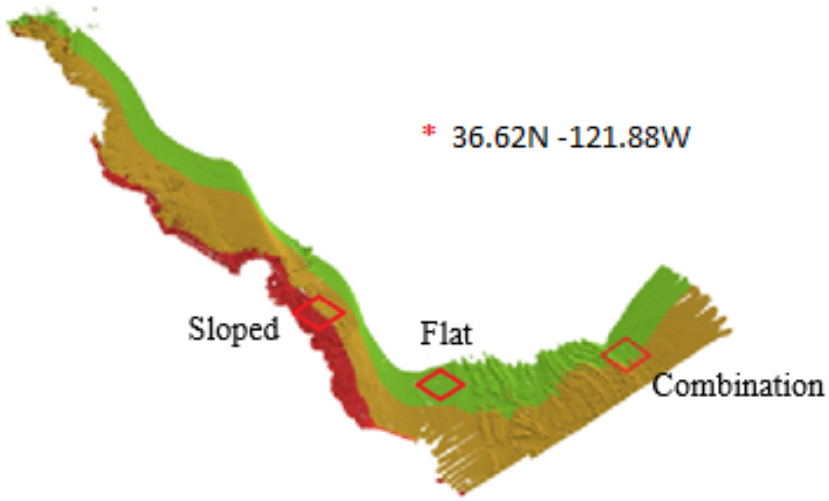



Figure 3.2. Locations of flat, combination, and sloped maps within the Monterey Bay

resolution level for our three landscapes under consideration.

Table 3.2. Relative entropy at each resolution level

| *Resolution Level* | *Flat* | *Combination* | *Sloped* |
|---|---|---|---|
| 1 | 413829 | 756268 | 1203563 |
| 2 | 420049 | 773357 | 1274166 |
| 3 | 425790 | 810400 | 1360855 |
| 4 | 431480 | 836321 | 1451909 |
| 5 | 438161 | 877515 | 1576521 |
| 6 | 445592 | 929438 | 1736963 |
| 7 | 452390 | 992291 | 1924242 |
| 8 | 462867 | 1073802 | 2211307 |
| 9 | 472810 | 1204231 | 2649966 |
| 10 | 497940 | 1468970 | 3557285 |

The rate at which relative entropy grows for different landscape typesare plotted in Figure 3.6. As the distance between data points decreases (map resolution increases), a flat landscape shows negligible entropy growth, the combination landscape shows positive linear entropy growth, and the sloped landscape shows positive parabolic entropy growth.

Each relative entropy growth rate is unique because the number of microstates is directly related to the uniqueness of the original bathymetry value (see Figure 3.1). We can also

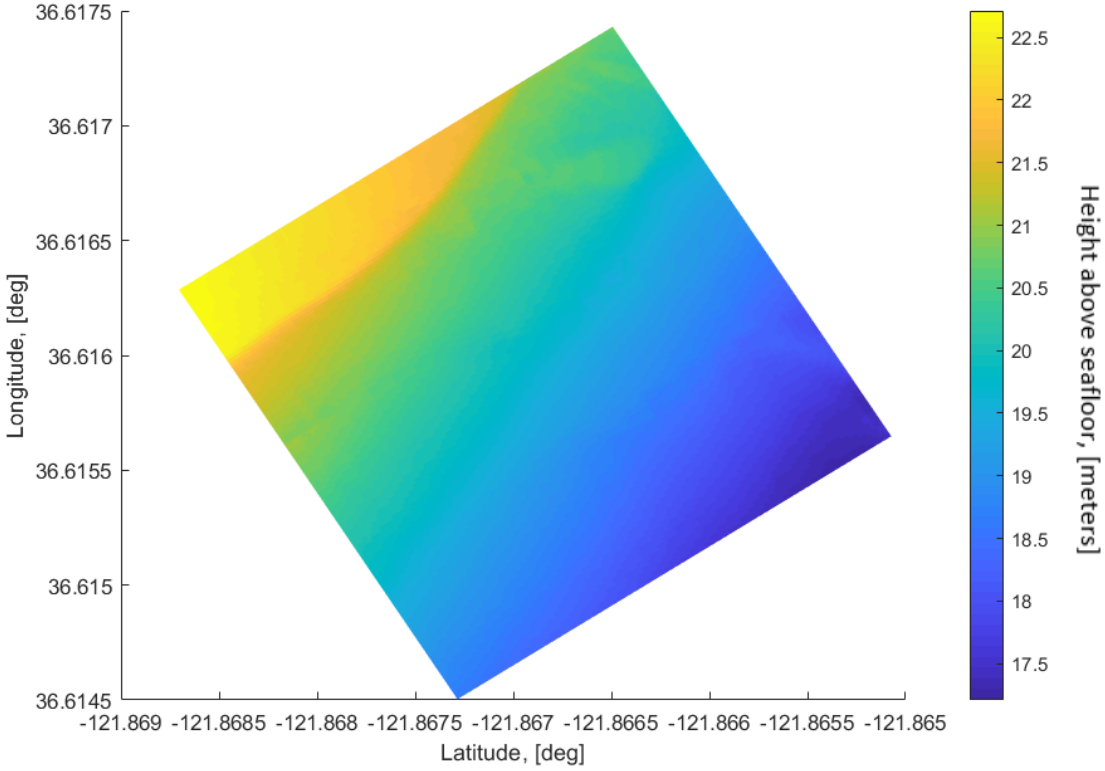



Figure 3.3. Flat landscape used in comparison calculations

observe, by comparing the terrain types at each resolution, that the differences in entropy become less distinct as the resolution decreases. This implies that navigation based on entropy should not be considered below a certain resolution because distinctions between terrain types are not viable. This is of significant importance since vehicles carry sensors with different resolutions. As a result, UTAN techniques may fail on these vehicles when it becomes impossible to distinguish terrain at this resolution.

## 3.4 Summary

In summary, we have introduced Boltzmann entropy as a new metric for quantifying the information available within terrain, and shown how changes in entropy are related to map resolution and bathymetry characteristics. In particular, the resolution of a map can determine if features are distinguishable using this metric. This implies that the most

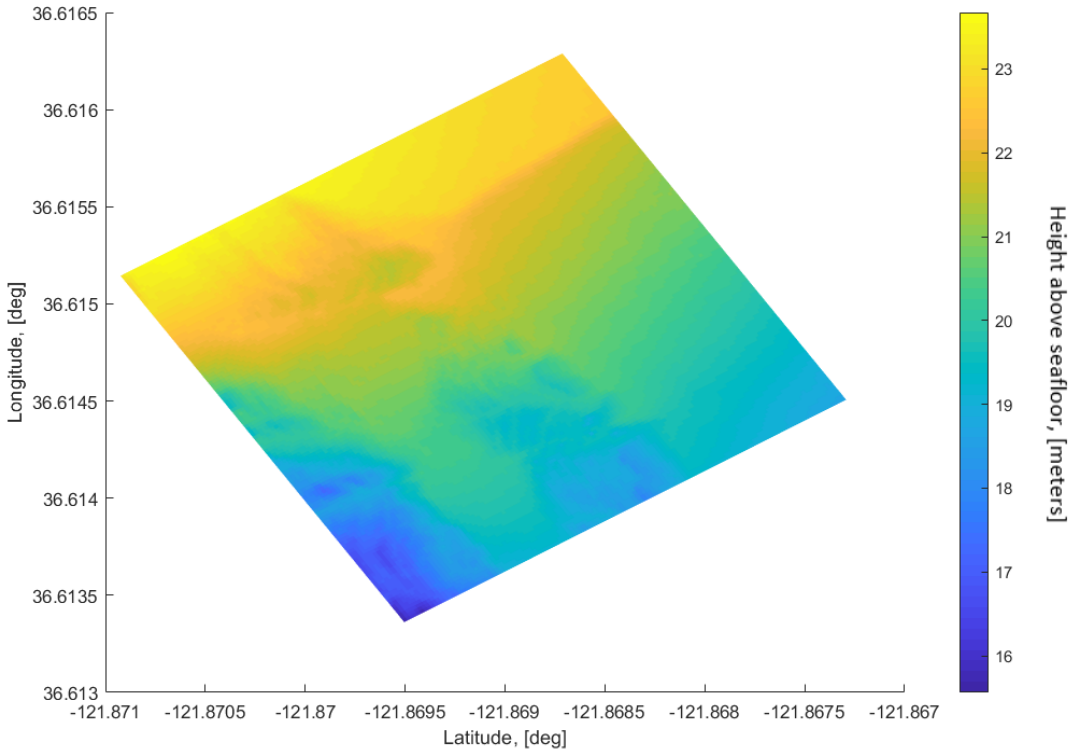



Figure 3.4. Combination landscape used in comparison calculations

valuable terrain features for use in UTAN are those features which appear unique in the greatest number of Boltzmann maps. As a result, these features are considered *key* to establishing an exploitation policy. This Chapter has utilized Boltzmann entropy within a terrain map to identify key bathymetric features. Next, we apply the results from this chapter to implement informative AUV path planning in chapter 4.

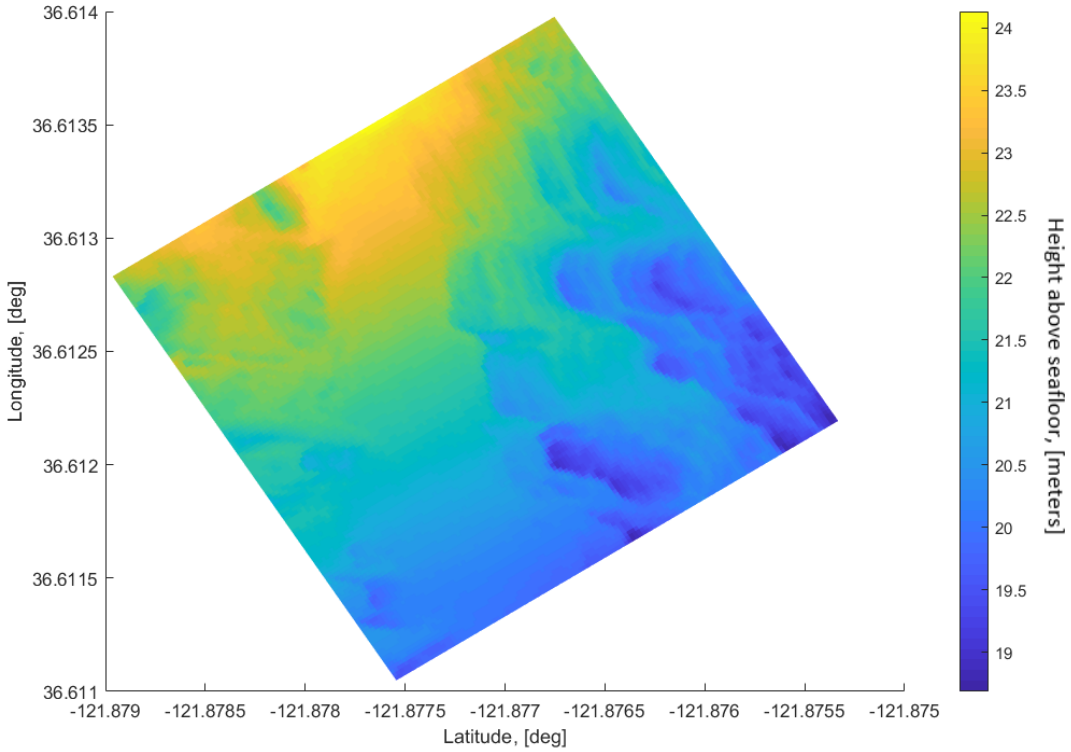



Figure 3.5. Sloped landscape used in comparison calculations

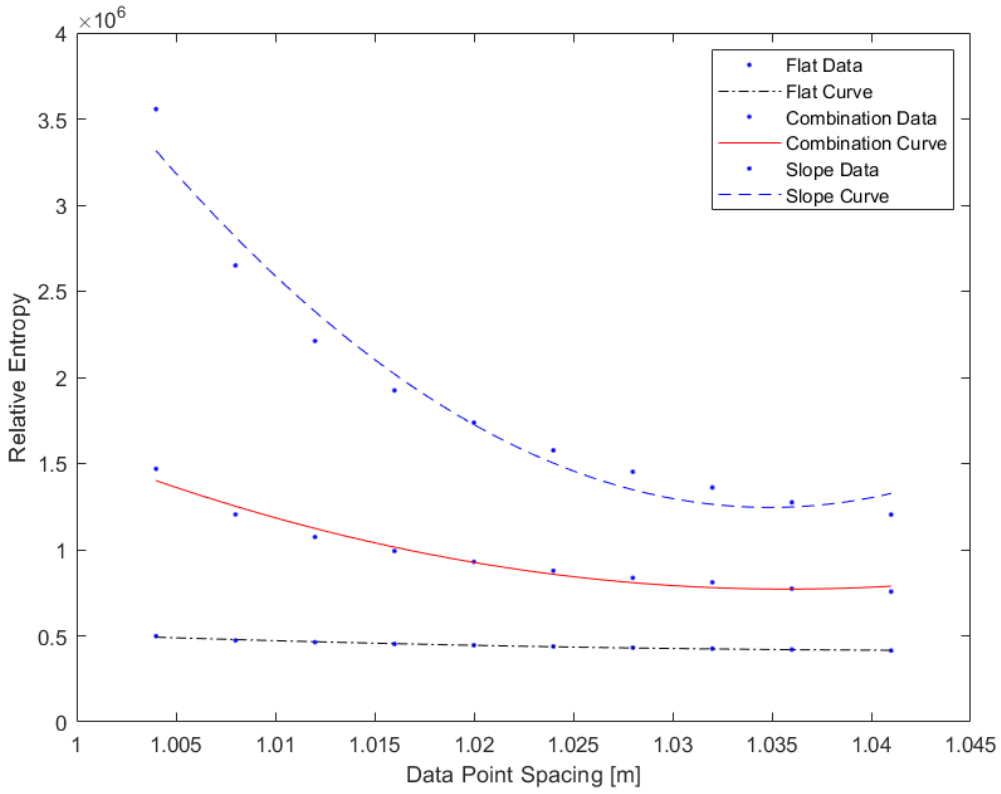



Figure 3.6. Curve fitted relative entropy growth rate of different landscapes

THIS PAGE INTENTIONALLY LEFT BLANK

# CHAPTER 4:
# Coverage Path Planning

Coverage path planning is the ability to determine a spatial path that allows a robot to completely cover a region, e.g., with a given sensor [35]. A coverage task is commonly applied in many of today's robots, such as vacuum cleaning robots [36], and painting robots [37]. In this chapter, we utilize the developed technique from Chapter 3 to implement coverage path planning. As described earlier, this can be accomplished by striking a balance between exploration and exploitation policies. In our coverage planning application, an AUV seeks to fully ensonify a mapped region with downward looking sonar while maintaining its own position estimate as accurately as possible. In this chapter, we demonstrate this with two standard coverage techniques: Boustrophedon and Voronoi partitions.

For testing these coverage approaches, a simple heuristic is used for path planning. If the AUV's position estimation error is less than the threshold, it continues to navigate along the prior planned path. If this error exceeds the threshold, the AUV plans a path to visit the nearest navigational feature. The feature is selected by analyzing the Boltzman map to identify maximal Boltzmann features that exceed a predetermined threshold. We assume that if the AUV is able to navigate over this feature, it is able to successfully correlate the sonar observation with the bathymetric map. This chapter will show the unintuitive results that even though the vehicle travels a longer distance, the Voronoi partitioning coverage planning technique provides better overall position estimation than the more traditional Boustrophedon approach. This is offset by the fact that paths planned using the Voronoi approach require more time and energy than Boustrophedon paths.

## 4.1 Problem Description

The three following figures give visual context to the operational area. Figure 4.1 shows the entirety of the bathymetry map provided by CSUMB and identifies our operational region with a red square. Figure 4.2 shows the geographical relationship between this bounded region and NPS when overlaid on Google Earth. Figure 4.3 gives a closeup view of this region overlaid on Google Earth.

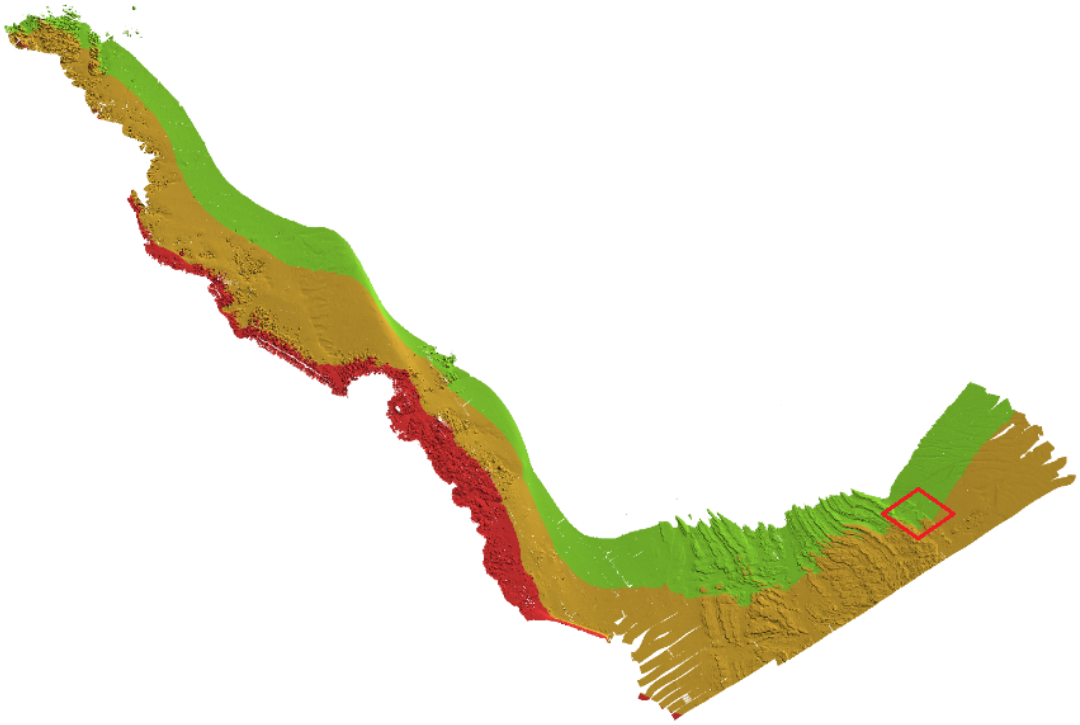


Figure 4.1. Bathymetry map of the south coast of Monterey Bay, adapted from CSUMB data

## 4.2 Informative Path Planning

Path planning is typically used as a means for a robot to reach a physical location while avoiding obstacles. In contrast, informative path planning enables a robot to understand the environment and collect information about a bounded region [38]. In [39], terrain-aided navigation is broken into two main categories:

1. **Bathymetric** terrain navigation. This is when bathymetric measurements are used as a means of comparison with an a priori bathymetry map to localize the vehicle.
2. **Feature-based** terrain navigation. This is when key features within the bathymetric map are used to compare and localize the vehicle.

This primary focus of this thesis is bathymetric terrain navigation. However, to simplify

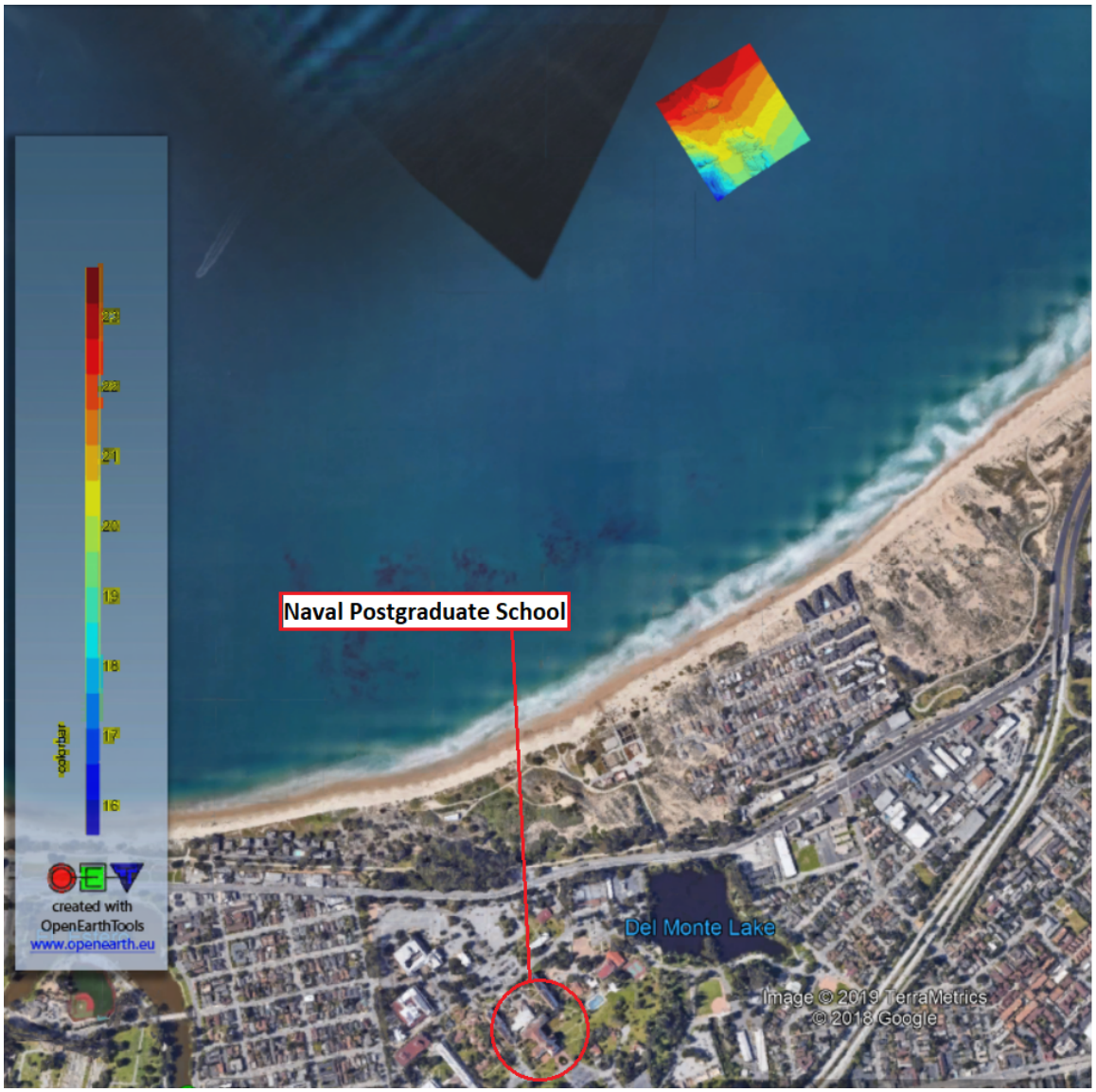



Figure 4.2. The bounded region in relation to NPS on Google Earth

discussion, this chapter assumes that the AUV conducts *feature-based* terrain navigation. For this implementation, two different maps are used that work together to achieve full and optimal coverage of a region. The *Boltzmann map* shown in Figure 4.5 is a map that provides the configurational entropy values corresponding to terrain features. The *Geographical* map shown in Figure 4.4 is used as the basis for calculating the Boltzmann map and for determining the vehicle's initial path.

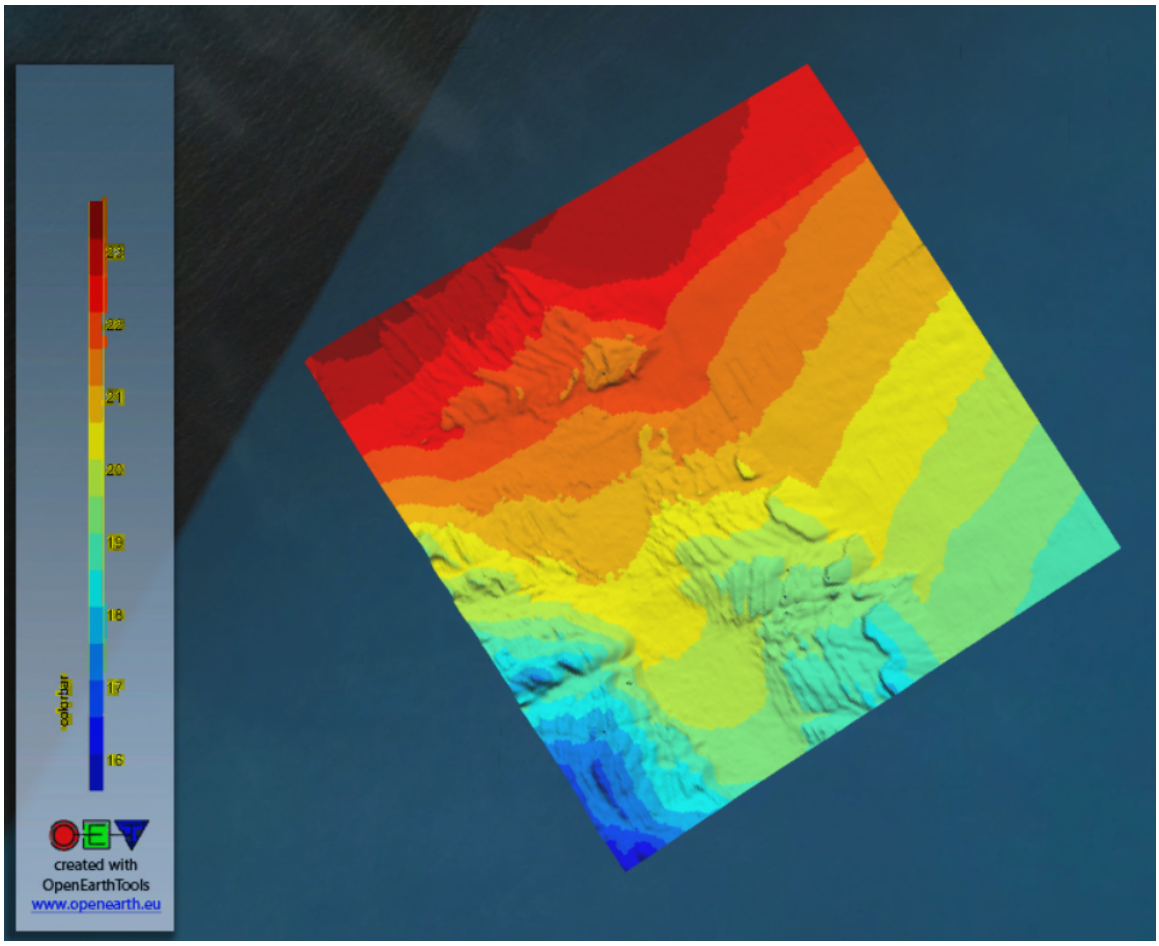



Figure 4.3. The bounded region overlaid on Google Earth

While using these maps, we implement the framework outlined in Chapter 1 to strike a balance between exploration and exploitation. To simplify our implementation, we assume the following:

1. An accurate *a priori* bathymetry map exists
2. The AUV has an accurate initial position from GPS on the surface
3. The AUV navigation model continuously introduces error over distance traveled
4. A Boltzmann map that can be used to identify key features exists
5. The AUV localizes its position when it navigates to a key feature and the sonar sensor observes it
6. Approximately 5-percent of the highest Boltzmann entropy values reduce the induced positional error by half.

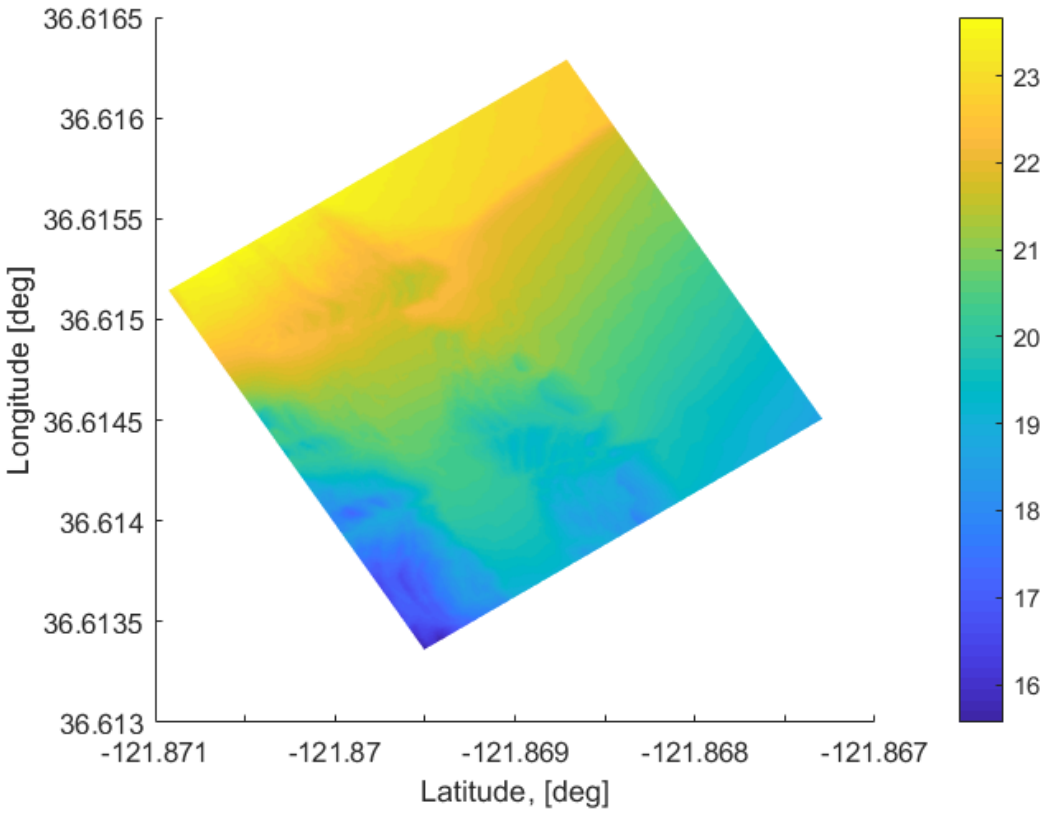



Figure 4.4. Geographical map used in experiment

7. Approximately 0.5-percent of the highest Boltzmann entropy values reduce the AUV's positional uncertainty to zero.

The goal in Section 4.2 is to utilize both maps to ensure complete and efficient coverage of a region. However, we need to identify a way of quantifying complete coverage. The method used in this thesis is known as *exact cellular decomposition*. This method breaks the coverage region into non-overlapping cells where coverage is complete once the AUV has ensonified each cell [40]. Because the downward looking sonar ensonifies a greater area than the physical AUV, the the optimal spacing between path segments must be calculated for the sonar's swath width. Sonar coverage is computed using the trigonometric relationship between the AUV's height above the seafloor and the sonar sytem's FOV. For this experiment, we use the BlueView MB2250 specifications listed in Table 1.1 and assume the vehicle maintains a height above the seafloor of eight meters. Figure 4.6 illustrates how

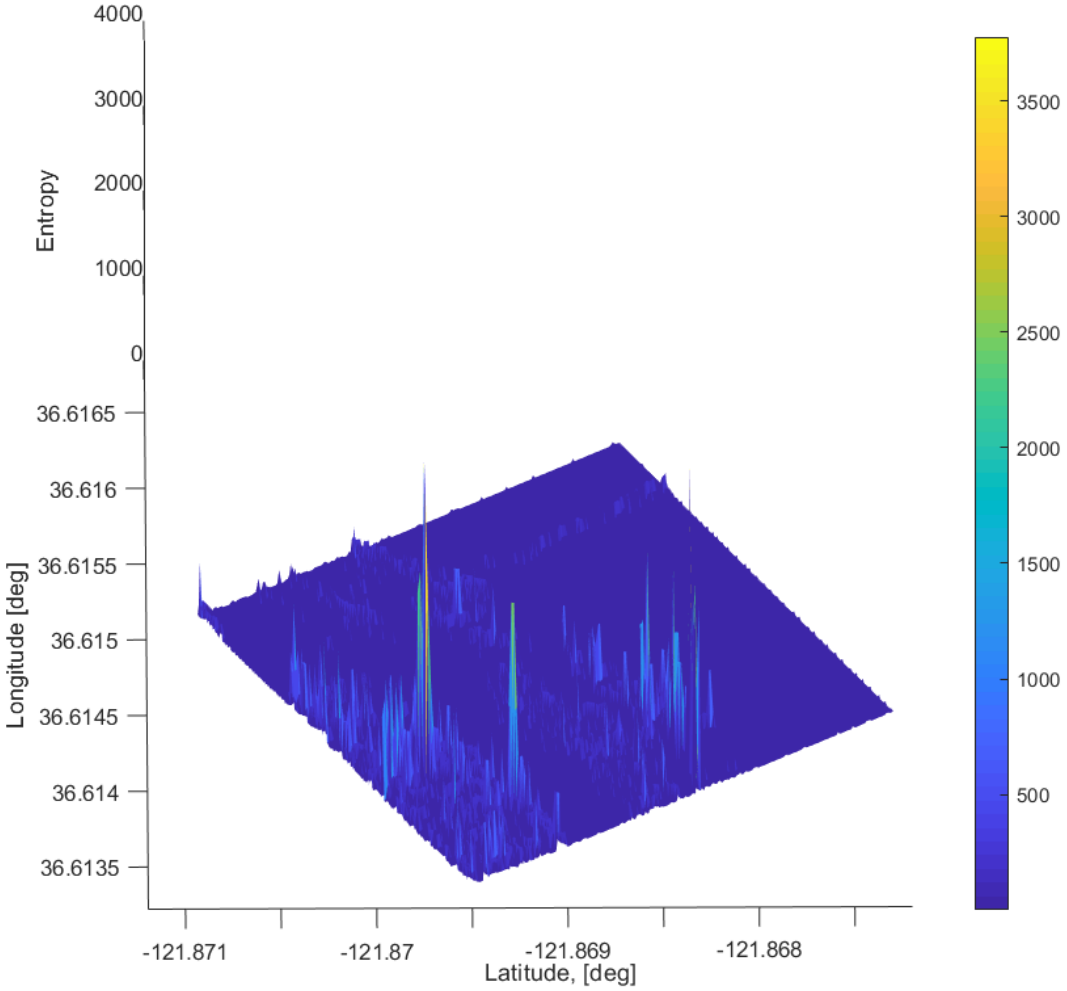



Figure 4.5. Boltzmann map used in experiment

these values are used to compute this sonar system's swath width in Equation 4.1.

$$Coverage = 2 * z * tan(0.5 * FOV) \tag{4.1}$$

$$Coverage = 2 * 8tan(0.5 * 45^{\circ}) = 6.63m$$

It is also good practice to have overlapping path segments when conducting a sweep. This helps the AUV ensure complete coverage of the region. In our example problems, we

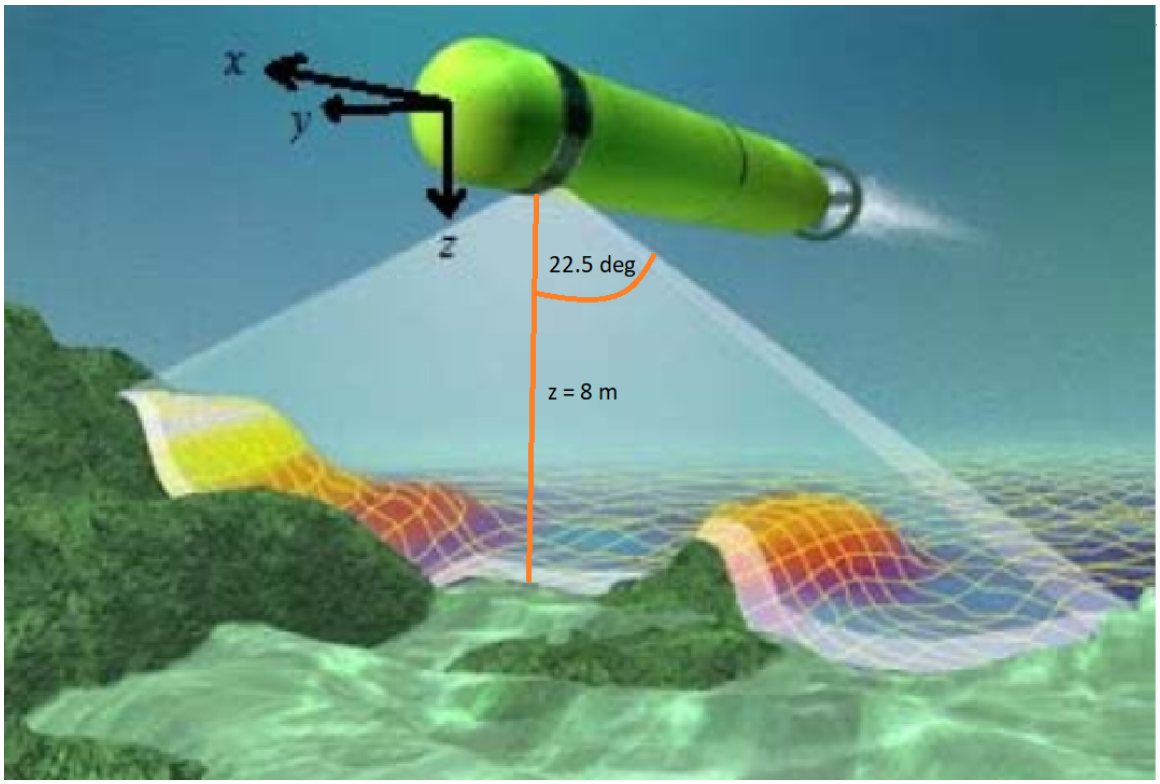



Figure 4.6. Visualization of FOV calculations

therefore assume a 6 meter spacing between path segments, producing a 0.63 meter overlap between sweeps.

## 4.3 Boustrophedon

The boustrophedon coverage method is more commonly referred to as "mowing the lawn". Its name was first used in 1699 and literally translates to English as "the way of the ox" [41]. This method, shown in Figure 4.7 (source: [41]), has historically been the most efficient way of exhaustively covering a region. However, the limitations imposed by the underwater environment force us to question the efficiency of this coverage method.

There are different metrics for determining the efficiency of a coverage plan. For this project we will look at 1) ideal vs. actual distance traveled, and 2) total uncertainty accumulated. We now compare these metrics for two different Boustrophedon patterns for covering a rectangular area: one with vertical path segments that sweep from left to right (e.g., Figure

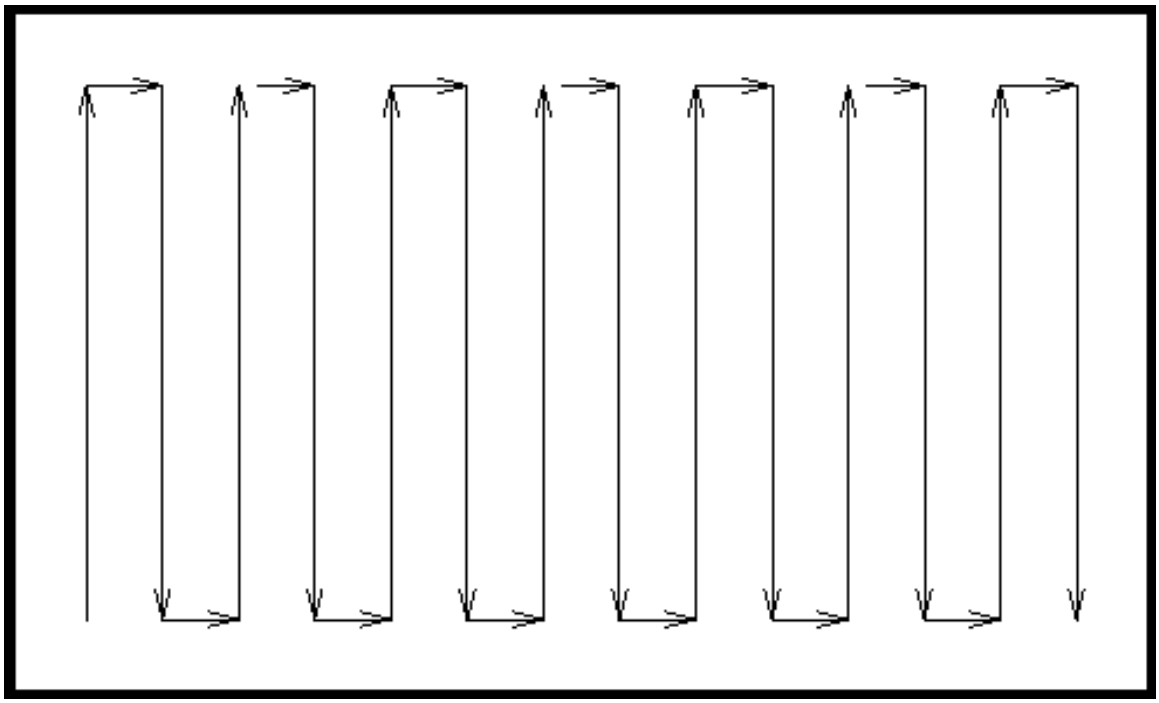


Figure 4.7. Boustrophedon coverage method

4.7), and one with horizontal path segments that sweep from bottom to top (e.g., Figure 4.8). In both cases, the AUV deviates from the ideal path. This occurs when its position uncertainty exceeds a specified threshold. As a result, the AUV must visit the nearest feature to fix its position error; it then returns to the point it left the ideal path to resume its mission.

The first metric compares the ideal distance traveled versus the actual distance traveled. The *ideal distance* is calculated using the dimensions of the mapped region (256 meters by 256 meters) and the sonar coverage swath width calculated from Equation 4.1. Next, we sum the distance required to completely cover the map. The *actual distance* includes the distance traveled to and from a nearby feature whenever it became necessary to fix accumulated position errors. The values for both sweep directions are enumerated in Table 4.1.

Table 4.1. Comparison of ideal and actual distance traveled

| *Sweep Direction* | *Ideal Distance* | *Actual Distance* |
|---|---|---|
| Bottom to Top | 9,885 m | 20,420 m |
| Left to Right | 9,885 m | 18,510 m |

The second metric tracks the position error accumulated as the vehicle covers the map.

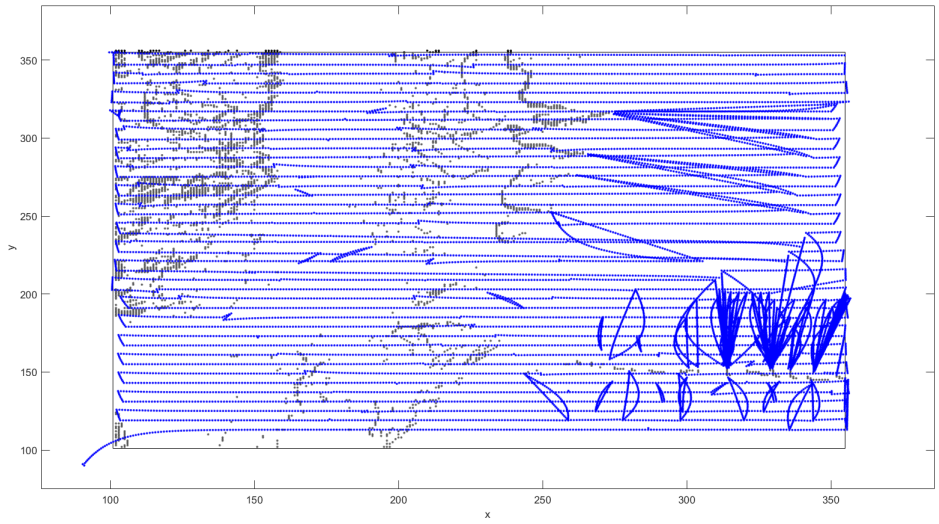



Figure 4.8. Boustrophedon with sweep direction from bottom to top

We limit the maximum position error by setting a threshold value which triggers a feature exploitation policy. This policy forces the AUV to visit the nearest key feature for a position fix before resuming its coverage pattern. Figures 4.9 and 4.10 illustrate how position error grows during path execution for both Boustrophedon sweep directions. Note that this error decreases when the AUV travels over key features. Threshold values for triggering the feature exploitation policy should consider map orientation and the distance a vehicle must travel to visit its localization features. For this reason, the error threshold was set to 8 meters for the bottom-to-top sweep direction, but only 5 meters for the left-to-right sweep direction.

Finally, we computed the total error by summing the trace of the covariance matrix. Total error for both Boustrophedon sweep directions are listed in Table 4.2 and shown in Figure 4.11.

Table 4.2. Comparison of total error accumulated

| *Sweep Direction* | *Total Error* |
|---|---|
| Bottom to Top | 74,776 m |
| Left to Right | 57,629 m |

An analysis of these metrics for both sweep directions suggests that partitioning the map into

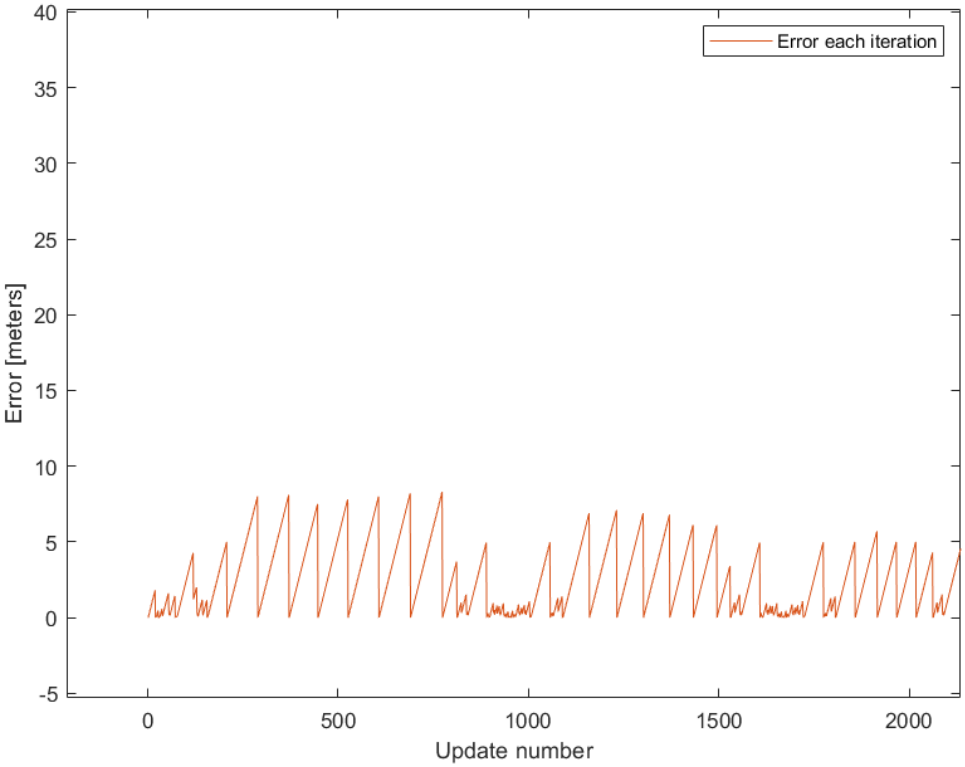



Figure 4.9. Accumulation of error using bottom to top sweep direction

different regions could reduce total accumulated error. A particular partitioning scheme might allow the AUV to maximize exploration in a feature rich region, while maximizing exploitation in another. Section 4.4 will explore one such method of dividing the map into regions: Voronoi Partitions.

## 4.4 Voronoi Partitions

Voronoi Partitions are part of a larger category of Grid Based Methods. Moravec and Elfes introduced this method in [42] for mapping a room with a ring-mounted sonar. The grids or Voronoi *cells* are determined by a given set of points, called *sites*, where the boundaries between two sites are equally spaced. It is defined by the following equation:

$$R_k = x \in X | d(x, P_k) \leq d(x, P_j) \; for \; all j \neq k \tag{4.2}$$

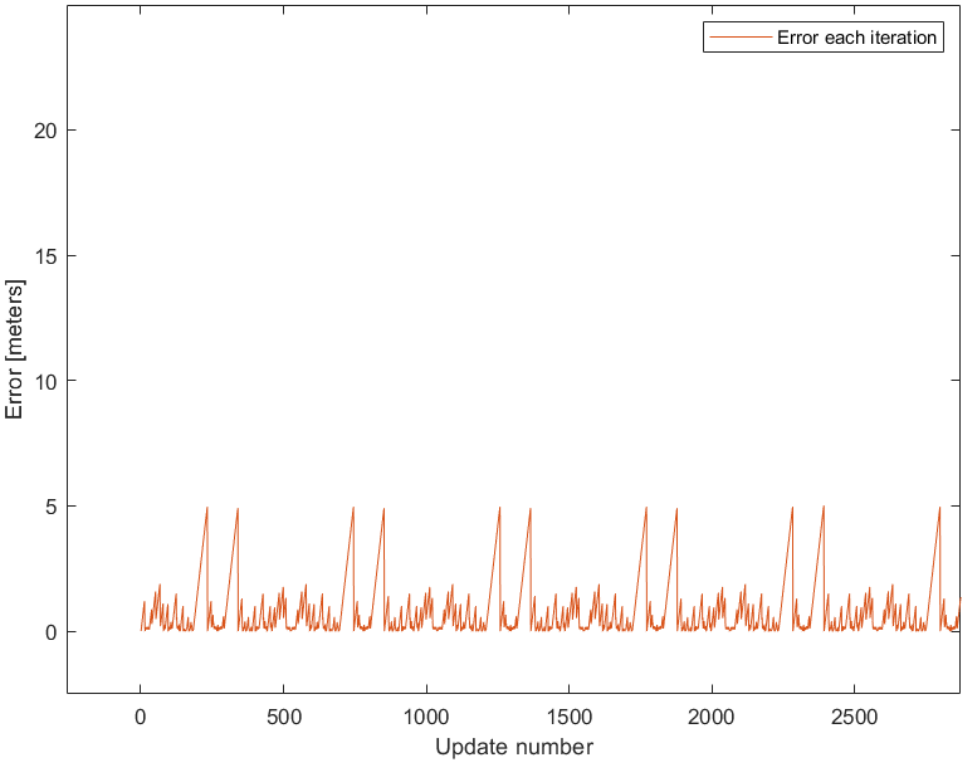



Figure 4.10. Accumulation of error using left to right sweep direction

where $R_k$ is a list of Voronoi regions, $x$ is a location inside the map $X$, and d is the distance from point $x$ to the site $P_k$ or $P_j$.

In autonomous navigation, Voronoi Partitions are typically used as a means of obstacle avoidance, where the sites are used to mark the obstacles. However, this method could pose problems if there are not enough high entropy features within the map or if those features are not properly spaced. As a result, specific high entropy values on our Boltzmann map are selected as the sites used to define the Voronoi cells. Furthermore, we can guarantee that a key feature is present within each partitioned region. Figure 4.12 shows the coverage map divided into Voronoi cells, each of which is covered by a Boustrophedon sweep pattern. Next, the distance traveled and accumulated error metrics from Section 4.3 are computed for this Voronoi coverage pattern. Table 4.3 compares the actual vs. ideal distance traveled. As expected, the Voronoi coverage pattern requires a longer path than

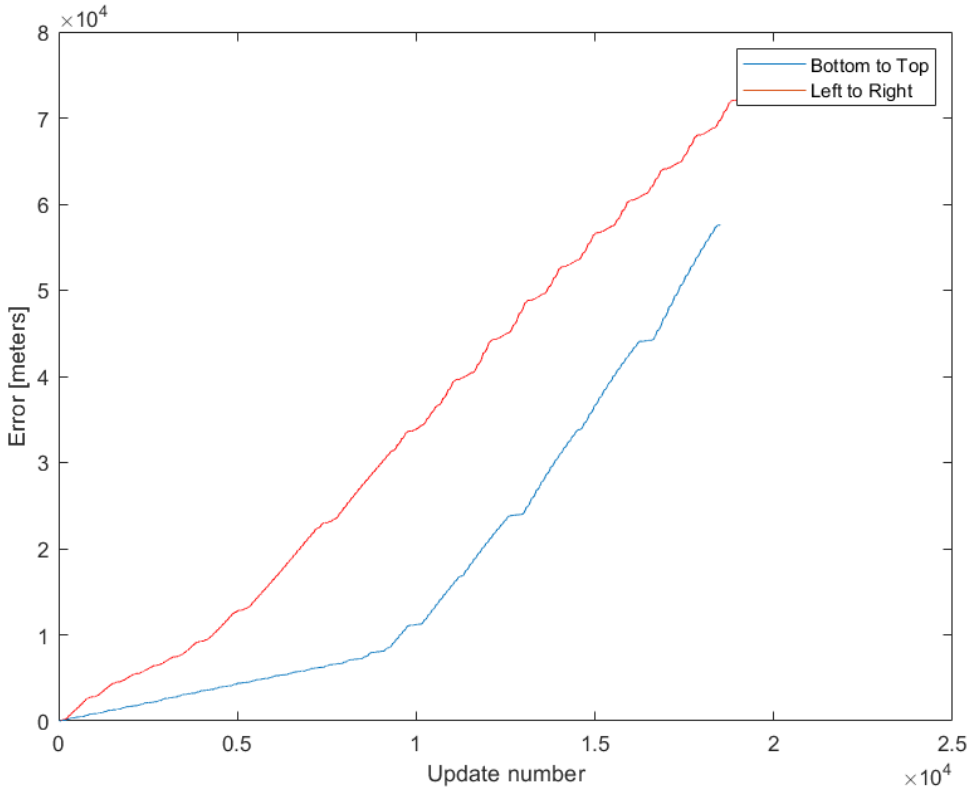



Figure 4.11. Comparison of total error for both sweep directions

either of the Boustrophedon coverage patterns. However, its error growth (Figure 4.13) and total accumulated error (Table 4.4) are reduced.

Table 4.3. Voronoi Partition comparison of ideal and actual distance traveled

| *Ideal Distance* | *Actual Distance* |
|---|---|
| 9,885 m | 24,569 m |

Figure 4.13 shows a sample of the total error growth and the error through each iteration. Table 4.3 shows a comparison between the idea distance traveled and actual distance traveled.

Table 4.4. Voronoi partitions total error accumulated

| *Search Type* | *Total Error* |
|---|---|
| Voronoi partitions | 27,591 m |

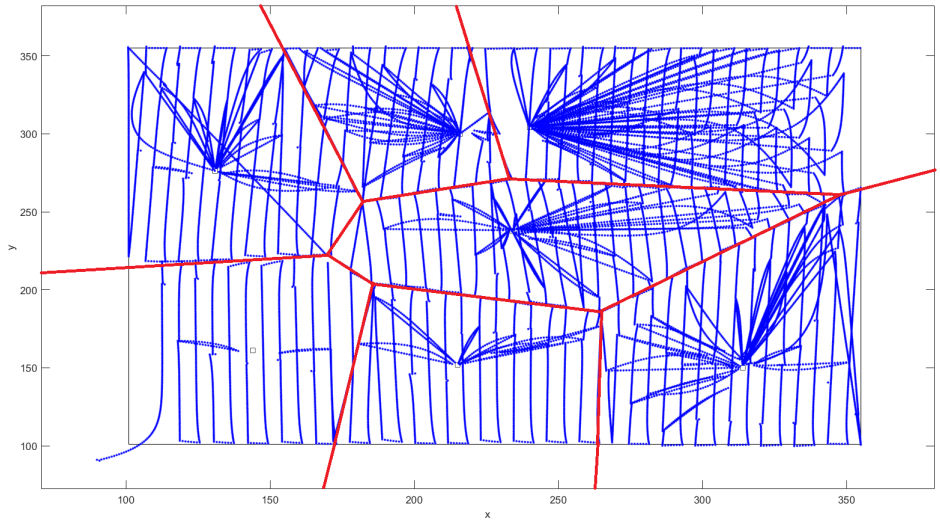



Figure 4.12. Image of AUV navigating the map divided by Voronoi partitions

## 4.5 Summary

In summary, the Voronoi Partition method maintained a lower positional error throughout the coverage problem. However, the distance traveled was notably larger than the Boustrophedon algorithms. This difference is a result of the increased number of turns required by the partitioned regions. Figure 4.14 compares the total positional error accumulated while executing each of the coverage patterns described above.

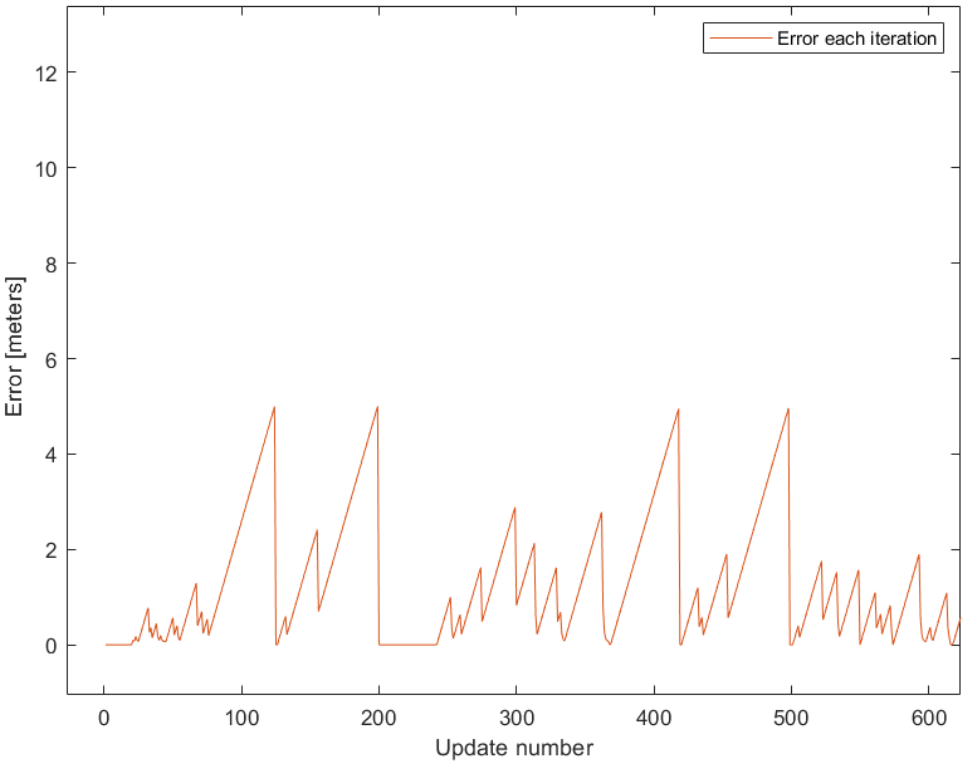



Figure 4.13. Accumulation of error using Voronoi partitions

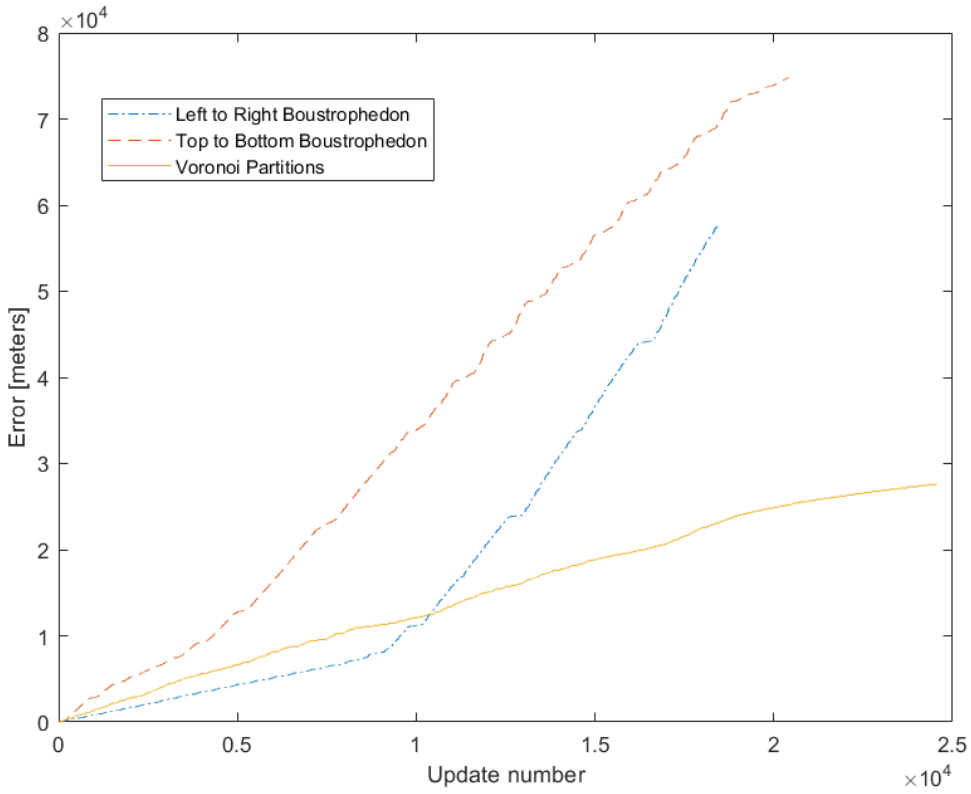



Figure 4.14. Comparison of total accumulated error

THIS PAGE INTENTIONALLY LEFT BLANK

# CHAPTER 5:
# Conclusion

UTAN is a potentially useful method for AUV navigation when GPS is degraded or coming to the surface risks vehicle safety or mission success. Current approaches have limitations that could misrepresent landscapes and produce inaccurate maps due to poor navigation. This thesis investigated a new paradigm to more accurately represent the information within a terrain. In this concluding chapter we summarize our findings (Section 5.1) and provide an overview of future extensions to this thesis (Section 5.2).

## 5.1 Contributions

This thesis addressed a wide range of topics. Significantly, we revealed concerns over conventional probability based entropy measures. We then introduced a new paradigm in landscape ecology which uses a configurational based entropy measure, namely Boltzmann Entropy. We concluded that this new paradigm, while more computationally expensive, removes ambiguity with the terrain and provides a metric for applying an exploitation policy.

We investigated different coverage planning techniques that utilized a Boltzman entropy map and a geographical map working in concert. We then compared different error metrics to determine the relative effectiveness of these techniques.

There are numerous potential applications for this research. The Information Theoretic Framework developed in this thesis can be applied to a number of scenarios, including sunken vessel recovery operations, search and rescue, or mine localization.

## 5.2 Future Work

Future work should focus on improving the robustness of our path planning algorithm. In Chapter 4 we outlined an approach to implement adaptive path planning to increase the efficiency of mapping a coverage region. In order to make our planning algorithm more resistant to failure, we seek to degrade the assumption of a prior map.

As currently implemented, the AUV is capable of switching between exploration and

exploitation policies in order to localize its position before continuing its mission. However, this requires a priori knowledge about the operating area. Implementing a modified Rapidly-exploring Random Tree Star (RRT*) planning algorithm would relax this requirement, as the planner could look over a limited horizon to make path planning decisions. New policies also need to be designed for scenarios in which the AUV is unable to localize with respect to the terrain.

In addition, the ability to utilize relative entropy as a means of making global path planning decisions should be explored. Visual matching algorithms would also make a useful addition addition to this planning framework. The ability to utilize camera systems in conjunction with sonar would increase robustness and widen operational applications.

In conclusion, the ability to navigate autonomously without GPS or beacon systems remains an elusive yet important capability for the military [1]. Configurational entropy as a means of quantifying the value of underwater features for AUV localization shows promising results. This is because underwater features that are not suitable for localization are not misconstrued as valuable. Continued research in this field could greatly increase the operational capability and flexibility that are needed for the military.

# List of References


[1] Sam Tangredi. Running silent and algorithmic. *Naval War College Review*, 72(2):129,165, 2019-04-01.

[2] Caliph Lebrun, Chris Price, and Brianna Kaufmann. Microbathymetry implementation, Mar-19.

[3] F Ababsa, M Mallem, and D Roussel. Comparison between particle filter approach and kalman filter-based technique for head tracking in augmented reality systems. In *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004*, volume 1, pages 1021,1026 Vol.1, 2004.

[4] G. T Donovan. Position error correction for an autonomous underwater vehicle inertial navigation system (ins) using a particle filter. *IEEE Journal of Oceanic Engineering*, 37(3):431,445, 2012-07.

[5] Carl Edward Rasmussen. Gaussian processes in machine learning. In *Summer School on Machine Learning*, pages 63–71, 2003.

[6] Richard Edwin Sonntag, Claus Borgnakke, Gordon John Van Wylen, and Steve Van Wyk. *Fundamentals of thermodynamics*, volume 6. 1998.

[7] P.-J Nordlund and F Gustafsson. Marginalized particle filter for accurate and reliable terrain-aided navigation. *IEEE Transactions on Aerospace and Electronic Systems*, 45(4):1385,1399, 2009-10.

[8] F Gustafsson, F Gunnarsson, N Bergman, U Forssell, J Jansson, R Karlsson, and P.-J Nordlund. Particle filters for positioning, navigation, and tracking. *IEEE Transactions on Signal Processing*, 50(2):425,437, 2002-02.

[9] P.-J Nordlund and F Gustafsson. Sequential monte carlo filtering techniques applied to integrated navigation systems. In *Proceedings of the 2001 American Control Conference. (Cat. No.01CH37148)*, volume 6, pages 4375,4380 vol.6. IEEE, 2001.

[10] Francisco Curado Teixeira, António Pascoal, and Pramod Maurya. A novel particle filter formulation with application to terrain-aided navigation. *IFAC Proceedings Volumes*, 45(5):132,139, 2012.

[11] Kjetil Bergh Anonsen and Oddvar Hallingstad. Terrain aided underwater navigation using point mass and particle filters. In *2006 IEEE/ION Position, Location, And Navigation Symposium*, pages 1027–1035. IEEE, 2006.

[12] S Dektor and S Rock. Improving robustness of terrain-relative navigation for auvs in regions with flat terrain. In *2012 IEEE/OES Autonomous Underwater Vehicles (AUV)*, pages 1,7, 2012-09.

[13] Samuel Cushman. Calculating the configurational entropy of a landscape mosaic. *Landscape Ecology*, 31(3):481,489, 2016-03.

[14] Peichao Gao, Hong Zhang, and Zhilin Li. A hierarchy-based solution to calculate the configurational entropy of landscape gradients.(report). *Landscape Ecology*, 32(6):1133–1146, 2017-06-01.

[15] Tim Schulte and Thomas Keller. Balancing exploration and exploitation in classical planning. In *Seventh Annual Symposium on Combinatorial Search*, 2014.

[16] Mikko Lauri and Risto Ritala. Planning for robotic exploration based on forward simulation. *Robotics and Autonomous Systems*, 83:15–31, 2016.

[17] Seung-Woo Byun, Joon-Young Kim, Seung-Keon Lee, Jong-Chun Park, and Kyung-Sung Kim. Development and experiment of a hovering auv tested-bed for underwater exploration. In *2007 Symposium on Underwater Technology and Workshop on Scientific Use of Submarine Cables and Related Technologies*, pages 285–291. IEEE, 2007.

[18] Jacob T. Juriga. Terrain aided navigation for remus autonomous underwater vehicle, June 2014.

[19] Shawn J Riley, Stephen D DeGloria, and Robert Elliot. Index that quantifies topographic heterogeneity. *intermountain Journal of sciences*, 5(1-4):23–27, 1999.

[20] Ariell Friedman, Oscar Pizarro, Stefan B Williams, and Matthew Johnson-Roberson. Multi-scale measures of rugosity, slope and aspect from benthic stereo image reconstructions. 7(12):e50440, 2012-12-12.

[21] Mark I McCormick. Comparison of field methods for measuring surface topography and their associations with a tropical reef fish assemblage. *Marine ecology progress series. Oldendorf*, 112(1):87–96, 1994.

[22] Ariell Friedman, Oscar Pizarro, and Stefan B Williams. Rugosity, slope and aspect from bathymetric stereo image reconstructions. In *OCEANS'10 IEEE SYDNEY*, pages 1–9, 2010.

[23] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379,423, July 1948-07.

[24] Information theory, 2011.

[25] Michael Grabchak, Eric Marcon, Gabriel Lang, and Zhiyi Zhang. The generalized simpson's entropy is a measure of biodiversity. *PLoS ONE*, 12(3):e0173305, 2017-03-07.

[26] Boris V Alekseev. Physical principles of the generalized boltzmann kinetic theory of gases. *Physics-Uspekhi*, 43(6):601–629, jun 2000.

[27] Matlab. online at https://www.mathworks.com/help/images/ref/entropy.html, June 2019.

[28] Shiuan Wan. Entropy-based particle swarm optimization with clustering analysis on landslide susceptibility mapping. *Environmental Earth Sciences*, 68(5):1349,1366, 2013-03.

[29] Qolamreza Razlighi and N Kehtarnavaz. A comparison study of image spatial entropy. volume 7257, 01 2009.

[30] L Masisi, V Nelwamondo, and Tshilidzi Marwala. The use of entropy to measure structural diversity. In *2008 IEEE International Conference on Computational Cybernetics*, pages 41–45, 2008.

[31] Isabelle Vranken, Jacques Baudry, Marc Aubinet, Marjolein Visser, and Jan Bogaert. A review on the use of entropy in landscape ecology: heterogeneity, unpredictability, scale dependence and their links with thermodynamics. *Landscape Ecology*, 30(1):51,65, 2015-01.

[32] Tancredi Caruso, Gaia Pigino, Fabio Bernini, Roberto Bargagli, and Massimo Migliorini. The berger–parker index as an effective tool for monitoring the biodiversity of disturbed soils: a case study on mediterranean oribatid (acari: Oribatida), journal =.

[33] Ludwig Boltzmann. *Über die Beziehung zwischen dem zweiten Hauptsatze des mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung, respective den Sätzen über das Wärmegleichgewicht*. 1877.

[34] Z. Li, C. Zhu, and C. Gold. *Digital Terrain Modeling: Principles and Methodology*. CRC Press, 2004.

[35] Howie Choset. Coverage of known spaces: The boustrophedon cellular decomposition. *Autonomous Robots*, 9(3):247,253, 2000-12.

[36] J Colegrave and A Branch. A case study of autonomous household vacuum cleaner. *AIAA/NASA CIRFFSS*, 107, 1994.

[37] Weihua Sheng, Ning Xi, Mumin Song, Yifan Chen, and P. MacNeille. Automated cad-guided robot path planning for spray painting of compound surfaces. volume 3, pages 1918–1923 vol.3, Oct 2000.

[38] Zhan Wei Lim, David Hsu, and Wee Sun Lee. Adaptive informative path planning in metric spaces. *The International Journal of Robotics Research*, 35(5):585,598, 2016-04.

[39] Kjetil Bergh Ånonsen. Advances in terrain aided navigation for underwater vehicles, 2010-10-18.

[40] Enric Galceran and Marc Carreras. A survey on coverage path planning for robotics. *Robotics and Autonomous Systems*, 61(12):1258,1276, 2013-12.

[41] Howie Choset. Coverage of known spaces: The boustrophedon cellular decomposition. *Autonomous Robots*, 9(3):247,253, 2000-12.

[42] H Moravec and A Elfes. High resolution maps from wide angle sonar. In *Proceedings. 1985 IEEE International Conference on Robotics and Automation*, volume 2, pages 116,121, 1985.

# Initial Distribution List

1. Defense Technical Information Center
   Fort Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California