DISSERTATION

OPTIMAL PATH PLANNING FOR DETECTION AND CLASSIFICATION OF UNDERWATER TARGETS USING SONAR

Submitted by

Christopher P. Robbiano

Department of Electrical and Computer Engineering

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2021

Doctoral Committee:

Advisor: Dr. Edwin K. P. Chong
Co-Advisor: Dr. Mahmood R. Azimi-Sadjadi

Dr. Ali Pezeshki
Dr. Iuliana Oprea

ABSTRACT

OPTIMAL PATH PLANNING FOR DETECTION AND CLASSIFICATION OF UNDERWATER TARGETS USING SONAR

The work presented in this dissertation focuses on choosing an optimal path for performing sequential detection and classification state estimation to identify potential underwater targets using sonar imagery. The detection state estimation falls under the occupancy grid framework, modeling the relationship between occupancy state of grid cells and sensor measurements, and allows for the consideration of statistical dependence between the occupancy state of each grid cell in the map. This is in direct contrast to the classical formulations of occupancy grid frameworks, in which the occupancy state of each grid cell is considered statistically independent. The new method provides more accurate estimates, and occupancy grids estimated with this method typically converge with fewer measurements. The classification state estimation utilises a Dirichlet-Categorical model and a one-step classifier to perform efficient updating of the classification state estimate for each grid cell. To show the performance capabilities of the developed sequential state estimation methods, they are applied to sonar systems in littoral areas in which targets lay on the seafloor, could be proud, partially or fully buried.

Additionally, a new approach to the active perception problem, which seeks to select a series of sensing actions that provide the maximal amount of information to the system, is developed. This new approach leverages the aforementioned sequential state estimation techniques to develop a set of information-theoretic cost functions that can be used for optimal sensing action selection. A path planning cost function is developed, defined as the mutual information between the aforementioned state variables before and after a measurement. The cost function is expressed in closed form by considering the prior and posterior distributions of the state variables. Choice of the optimal

sensing actions is performed by modeling the path planning as a Markov decision problem, and solving it with the rollout algorithm.

This work, supported by the Office of Naval Research (ONR), is intended to develop a suite of interactive sensing algorithms to autonomously command an autonomous underwater vehicle (AUV) for the task of detection and classification of underwater mines, while choosing an optimal navigation route that increases the quality of the detection and classification state estimates.

ACKNOWLEDGEMENTS

I would like to thank my two doctoral advisors. Dr. Edwin K. P. Chong chose to admit me as his student at a time when I was dreamy eyed with the prospect of being a graduate student. He took a risk allowing me to be his student, and was always able to distill the most abstract ideas into manageable portions. He pushed me to understand that no problem is too hard, no understanding is unachievable. Dr. Mahmood R. Azimi-Sadjadi provided me with the opportunities to learn, fail, and more importantly explore the vast and rewarding world of statistical signal processing. He was always willing to listen to half-baked ideas, and provide guidance on how to develop those ideas into full fledged intellectual contributions. My life has been forever molded by the influence of these two men, and I am incredibly grateful, and especially lucky, to have crossed path with them. I will forever consider you mentors, and more importantly friends.

I would like to thank my committee members, Drs. Ali Pezeshki and Iluiana Oprea. They were both instrumental in my education and evolution throughout graduate school. Both saw me as a first-semester graduate student, lacking all appreciable knowledge and experience, yet were encouraging and supportive.

Dr. Louis Scharf gets a special mention. Being such an animated individual, and *the* leader in our field, I am delighted that I had the opportunity to study with him and experience first hand how true innovation occurs.

I would like to thank the Office of Naval Research (ONR) for providing the funding for this project. This project was funded by the ONR under contract N00014-18-1-2805. Without the funding from the ONR, this work would have never been completed. I would also like to thank Denton Woods at the Naval Surface Warfare Center – Panama City, FL for providing the logistical and data support needed for this project.

I would like to thank my colleagues that I have met throughout my tenure at Colorado State University. They have provided a great environment for discussing work and providing help when most needed. Thanks to Megan, Jack, Patrick, and Pranav.

Finally, I would like to thank my family for always believing in me, and never giving up when I experienced youthful indiscretion. Specifically, my dog Durby. Miss you, pup.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

1.1 Problem Statement

The material presented in this dissertation has been developed through research as part of a project funded through the Office of Naval Research (ONR). The project has set out to solve the active perception problem of simultaneously performing detection, classification, and path planning utilizing a sonar system in a littoral regions. Specifically, the research is focused on performing underwater mine counter-measures (MCM), namely identification and localization of mines, in an adaptive manner inspired by the biosensor systems of animals such as bats and dolphins.

This biomimetic MCM problem is possibly best described by the following description of a bat during its hunt for prey. As a bat hunts for food, it is constantly emitting audible chirps and listening for reflections of interest. In the case that the bat does not hear a reflection, the bat knows with a high probability that there is nothing in its immediate path. However, if the bat does hear a reflection indicating presence of a target, then the bat must simultaneously attempt to determine what type of object caused the reflection and how to navigate without encountering clutter objects, e.g., brick walls or trees, while trying to reach its prey.

If the bat does not have enough information to correctly identify the target with high probability, then it may choose a new path and/or modify the emitted signal while considering how each new location will increase the probability of correct identification. In this way, the bat is performing simultaneous path planning for detection and classification of its prey.

The problem addressed in this dissertation attempts to solve a similar problem faced by the bat, using autonomous platforms, such as autonomous underwater vehicles (AUVs) equipped with an array of sonar sensors used to survey the area under investigation. As the AUV navigates and collects sonar measurements, it needs to decide where the best position is to take the next

measurement. Much like the bat, the best position provides the most amount of information to both detect targets as well as simultaneously classify those targets.

At each time step, the AUV performs a sensing action in the form of selecting and moving to the next position to collect a measurement that is used to update the detection, localization, and classification estimates. This exploration problem, often referred to as the *active perception* problem [1], has no pre-planned vehicle path as there is no *a priori* information about object locations in the search area, and all initial sensing actions are regarded as providing the same amount of information. The motion of the vehicle, however, is restricted by some dynamical model, hence precluding arbitrary sequential sensing locations. Through the process of collecting measurements, a map is formed that contains information about where potential objects of interest exist and of what type they are. It is the contents of this map that enables sensing locations to be chosen.

This problem, and particularly the environment and sensors used, poses significant difficulty and faces some shortcomings. In particular, the sonar sensor has a wide beamwidth, and the emitted and reflected chirps must radiate through salinated water, a relatively lossy medium. This presents two issues in particular. The first being that a large volume of space is ensonified by a single chirp and the originating location of a reflection can be ambiguous. The second is that the lossy medium greatly attenuates reflected signals, and effectively raises the noise floor for potential targets further away from the AUV. The different potential positions of targets, such as being buried or proud of the bottom, introduces problems as well. For buried targets, the angle of incidence between the AUV and the target plays a large role in the ability to detect the target.

This dissertation effectively segments the problem into three components: sequential detection state estimation, sequential classification state estimation, and optimal path planning. The first two components are solved separately, and then fused together in an information-theoretic manner to aide in the solution of the third component.

1.2 Survey of Previous Work

The sequential detection state estimation problem described in the problem statement involves estimating an occupancy map given a set of observations. The occupancy grid estimation process was originally introduced by Elfes and Moravec in the mid 1980s [2]. Several subsequent papers explored alternative methods for performing sensor fusion for distributing sensor measurements over the occupancy grids [3], and for combining multiple occupancy grids [4, 5] from multiple independent sensors into a single grid. These methods make the assumption that the occupancy states of the grid cells are statistically independent by modeling the problem as a Markov Random Field [2,4,5]. This allows for the factorization of the joint occupancy distribution on the occupancy map into the product of occupancy distributions of individual grid cells. For our problem, potential targets overlap multiple grid cells. The factorization ignores this fact, and provides estimation techniques that are not utilizing all of the available information.

Thrun [6] provided a new occupancy grid formulation, using forward sensor models, that accounts for statistical dependence between the occupancy state of grid cells. The method assumes that the sensor provides range measurements from within an observation cone, and that the single range measurement comes from only a single source within the cone. This measurement is assumed to be produced by either a true positive (detection of object), a false positive (false alarm), a true negative (no object in range), or a false negative (missed detection). Each of these possible events is modeled with a distribution from the exponential family, and the process of identifying the occupancy grid becomes a most-likely-model selection process through the use of an expectation-maximization (EM) algorithm [6]. This formulation assumes that only a single range measurement is produced, whereas our sonar system is able to produce a vector of measurements associated with multiple ranges at each time step.

The recent work in [7, 8] used real antenna radiation patterns to better inform occupancy grid estimation and hence provide more realistic maps. In [9, 10], the authors model the dependence between grid cells using Gaussian processes. The use of Bayesian Occupancy Filters (BOF) [11]

for generating occupancy grids, assuming statistical independence between the grid cells, has been studied in [12–14].

The optimal path planning problem finds roots in many areas, such as the traveling salesman problem [15] and packet routing for digital communications [16]. In the problem presented in this dissertation, the optimal path is the one that provides the most information for providing confusion free detection and classification of underwater objects using sonar measurements. Methods for choosing an optimal path have seen success while using information-theoretic measures [17–21] in autonomous navigation and exploration problems. In the case of parameter estimation using measurements that are corrupted by Gaussian noise, maximizing the Shannon entropy (amount of information) of the error distribution is equivalent to minimizing the determinant of the parameter estimate covariances [17]. This provides a rule for selecting the sensing action that maximizes the predicted variance of the measurement produced after a sensing action is performed.

In [18], a subclass of the active perception problem is addressed, where an AUV is used to inspect the hull of a large ship and estimate its surface shape. Gaussian process function approximation is exploited to approximate a mutual information-based cost function. In this particular multi-hypothesis testing problem, *a priori* information is available that allows the entire set of sensing actions, and their outcomes, to be observed prior to visiting all sensing locations.

Information-theoretic cost functions, specifically utilizing *information gain*, have been previously developed [19–21], and used successfully, in the context of navigation using information from the occupancy grid estimation process [4], [6].

In [19], the positions of the AUV and potential targets are estimated for a given sensing action using an extended Kalman filter (EKF), and the mutual information is directly calculated following an update to the occupancy grid. In addition to the occupancy grid based information gain cost, the authors in [19] suggest formulating an additional information-theoretic cost function from the outputs of the simultaneous localization and mapping (SLAM) problem using an EKF to estimate the positions of the AUV and objects in the search area. Specifically, the cost function they choose is related to the determinant of the AUV and object position error covariance matrices, similar

to [17]. A convex combination of the two normalized cost functions is used in the sensing action selection, providing the ability to trade-off performance in localization (through SLAM) and detection (through occupancy grids) of objects.

In [20], the mutual information is directly calculated after each sensing action and subsequent measurement is taken *a priori*, and used to train the Gaussian process regression network for estimating the mutual information for future sensing actions. Bayesian optimization is then used in conjunction with the Gaussian process upper confidence bound to estimate the information gain for each point in an occupancy grid.

The formulation for explicitly calculating the predicted mutual information in [21] is developed using an occupancy grid framework under the assumption of statistical independence of measurements. Measurements are also assumed to be conditionally independent of the occupancy state of obscured grid cells, i.e. grid cells behind occupied cells in the perceptual range of the sensor, given an occupied grid cell.

1.3 Contribution of Present Work

The block diagram in Figure 1.1 shows the active perception sonar problem that is addressed in this work. The process starts, in the top right image of Figure 1.1, with an AUV producing wide-band chirps from a multi-element hydrophone array, sampling the local sound pressure as the chirp is reflected back towards the sonar sensor. The sampled sound pressure levels are used in two different ways, one for detection and one for classification.

The bottom right image of Figure 1.1 captures the detection and classification steps of the process. For detection, the sampled sound pressure levels are passed through an adaptive coherence estimate (ACE) detector [22], which produces detection statistics for each sample. The detection statistics are thresholded to produce a 1 or a 0, indicating that the sampled signal contains enough power in the signal subspace, or not, respectively. The thresholded detection statistics are used as the input to the sequential detection state estimation process.

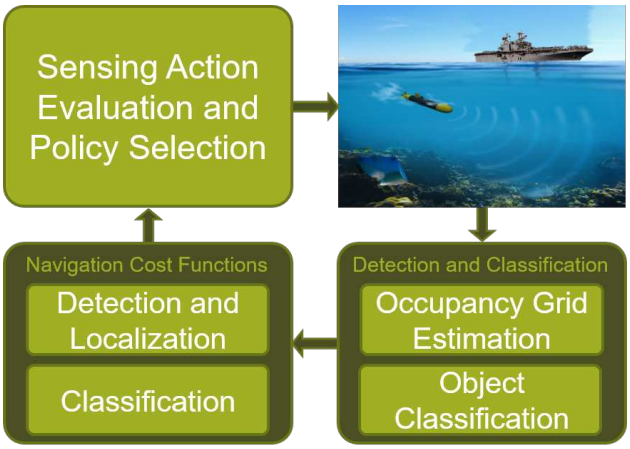


Figure 1.1: Block diagram of the current system implementation

For classification, the sampled sound pressure levels are converted into feature vectors of the form expected by the one-step classification process. This form can range from the raw, non-processed time series to some mapped version of the data, where the mapping has been designed to increase the discriminative capabilities of a one-step classifier. The outputs of this classifier is then used as the input to the sequential classification state estimation process.

The bottom left image of Figure 1.1 captures the calculation of the path planning cost function using the outputs of the detection and classification processes. The path planning cost function is estimated by solving for the mutual information [23] between prior and posterior distributions. Finally the top left image of Figure 1.1 shows the action selection process, which evaluates the path planning cost function for different possible outcomes.

The methods for sequential detection state estimation using occupancy grid estimation typically make the assumption that a sensor produces a range measurement, i.e., reporting the range of the closest detected object in the sensors' field of view (FoV). However, in our active perception problem depicted in Figure 1.1, a sensory system on a moving platform takes observations at all possible ranges within its beam length, producing a vector-valued measurement. Classical occupancy grid estimation does not deal with this situation. Additionally, the classical methods treat the occupancy state of each grid cell as being statistically independent. In a typical occupancy grid type problem, the area of a grid cell is much smaller than the area which an object occupies,

leading to objects within the interrogation field typically occupying multiple grid cells. Consequently, the occupancy state of a grid cell is inherently correlated with its neighboring cells, and thus the occupancy state of the grid cells should be jointly considered when performing occupancy grid estimation. These two departures from the classical methods motivated us to develop a new occupancy grid estimation method to circumvent these problems.

To this end, we present a framework for solving the joint and marginal distributions of grid cell occupancy states, while accounting for the statistical dependence between the occupancy states of grid cells, given vector-valued sensor measurements. We then present a method for solving this problem with a sensor model that exploits a network of binary asymmetric channels (BACs). This BAC network sensor model can be used to represent any real sensor, with examples given for an ideal ranging sensor, a sonar system, and a distributed network of pressure sensors. The main contributions of this work are as follows. Using BACs, we build a sensor model for the interaction between the occupancy state of each grid cell and each sensor measurement. This sensor model provides a tractable method for considering the statistical dependence between grid cell occupancy states. The use of BACs provides a method of modeling *any* physical sensor by considering its statistical performance (probability of false alarm and probability of detection), and does not rely on a presumed distribution for sensor measurements. We also show that the original formulation of Elfes [2,4], which assumed statistical independence between each grid cell occupancy state, can be considered as a special case of the methods proposed here.

In addition to the new occupancy grid estimation method, we propose a new approach to the problem of active perception using two information-theoretic cost functions described using information gain. The first cost function is associated with object detection and localization, and measures the mutual information between the occupancy state variable for a single grid cell and a binary measurement random variable. Solving for its closed-form representation relies on the measurement model and posterior occupancy distributions produced through the occupancy grid estimation process presented in [24]. The second cost function associated with the classification of detected objects measures the mutual information between a class state variable for a single

grid cell and random variable that is the parameter to the class state variable distribution. In this formulation, we choose to model the class state variable as a Categorical random variable, and its distribution parameter as a Dirichlet random variable [25]. The motivation for choosing this modeling scheme comes from the need to perform sequential updating of the class state variable distribution, akin to occupancy grid estimation process. This sequential updating process has a closed-form due to the conjugacy between the Dirichlet and Categorical distributions, and allows for fast tracking of the class state distribution as new measurement are drawn. A *one-step* classification process is used in our formulation, producing class labels used to sequentially update the class state variable distribution. Similar to [19], a convex combination weighting of two normalized cost functions for sensing action selection is also utilized here to provide a trade-off between exploration of the search area and localization and classification of detected targets.

The occupancy grid estimation process provides a probabilistic method for estimating whether any point within an environment is occupied, and produces a collection of probability distributions as its output. Similarly, the classification map estimation process provides a probabilistic method for estimating the class membership of each point within an environment, and produces a collection of probability distributions as its output. These distributions can be used in a natural manner to provide informed clues used for path planning through information-theoretic metrics used as path planning cost functions. By modeling the path planning problem as a Markov Decision Problem (MDP), optimal and near-optimal trajectory planning can be rigorously formulated, and the aforementioned information-theoretic metrics can be used in the optimization of the trajectory.

The aforementioned theoretical developments are tested in a series of experiments using sonar data. The occupancy grid developments are exemplified with a fixed path, showing the ability to discriminate multiple targets that are closely spaced. The performance of the entire system is evaluated with a simulated AUV conducting simultaneous detection, classification, and path planning.

1.4 Notation

Many mathematical ideas are presented throughout this dissertation. To make sure the reader is familiar with the notation that is used, and can follow along without confusion (at least due to the use of mathematical symbols), we present the general notation that is used in the remainder of the document. This is meant as reference material, to be reexamined by the reader as needed, as most of the ideas referenced below have not been introduced yet.

Scalar values are represented as an italic lowercase symbol x, vectors as lowercase italic bold symbols $\boldsymbol{x}$, and matrices and sets as uppercase italic bold symbols $\boldsymbol{X}$. The symbol notation for all probability measures p is annotated with the random variable it measures, i.e., the measure p taking argument $\boldsymbol{b}$ is represented as $p_{\boldsymbol{b}}(\boldsymbol{b})$.

The following is a summary of the notation used in discussion of the sequential estimation frameworks.

- B : The number of grid cells in a map.
- $\boldsymbol{b}$: A random vector taking values in $\{0,1\}^B$ representing cellular occupancies.
- $\tilde{\boldsymbol{b}} \in \{0,1\}^B$: The virtual cellular occupancies after the occupancy state of each grid cell has been passed through a binary asymmetric channel (BAC).
- $\boldsymbol{p} = [p_{b|\boldsymbol{J}}(b_1 = 1|\boldsymbol{J}_S), \cdots, p_{b|\boldsymbol{J}}(b_B = 1|\boldsymbol{J}_S)] \in [0,1]^B$: The vector of cellular posterior probabilities.
- $\mathbb{B}$: Set of all possible cellular occupancies $\boldsymbol{b}$. $|\mathbb{B}| = 2^B$.
- $\mathbb{B}(r,\beta) = \{\boldsymbol{b}|b_r = \beta\}$: Set of all cellular occupancies with $b_r = \beta$, $\beta \in \{0,1\}$.
- S : Current/latest time index.
- s : Time index with $1 \leq s \leq S$.
- $\boldsymbol{j}_s \in \{0,1\}^K$: Binary-valued measurement vector of dimension K taken from the output of a sensory system at time s.

- $\boldsymbol{J}_S = \left[\boldsymbol{j}_1, \cdots, \boldsymbol{j}_S\right]$: Collection of binary-valued measurement vectors up to time S.

1.5 Dissertation Organization

This dissertation is arranged in the following order. A detailed review of the sequential detection state estimation methods is presented in Chapter 2. The type of sonar data used in all experiments is presented in Chapter 3. The adaptive coherence estimator (ACE), which is used to preprocess the sonar data, is also presented in Chapter 3. Chapter 4 presents our extension to the sequential detection state estimation process, which can solve for the joint distribution of the occupancy states and account for vector-valued measurements. The sequential classification state estimation process is described in detail in Chapter 5. First a review of the Dirichlet and Categorical distributions is presented, along with some of their significant properties, followed by an algorithm for performing the state estimate updates. Chapter 6 presents the formulation of an information-theoretic cost function for path planning based on the outputs of the detection and classification state estimation processes. Experiments, using synthetically generated sonar data, illustrating the effectiveness of the developed occupancy grid formulation over classical methods, as well as experiments involving concurrent detection, classification, and path planning, are presented in Chapter 7. Final concluding remarks are made in Chapter 8.

Chapter 2

Sequential Detection State Estimation

2.1 Introduction

The sequential detection state estimation is posed as an *occupancy grid* (OG) estimation problem. A sequential process for estimation of the detection state at each point on a map is used to incorporate new measurements as the AUV navigates and measurements become available to the system. This process provides a confidence measure of the occupancy for each point in a map, and allows for a continued updating of this confidence measure as new measurements are taken. In this chapter, we will introduce and examine the traditional occupancy grid estimation methods. More specifically, two different classes of occupancy grid estimation methods are presented. The first class considers that each of the state variables representing the occupancy of each grid cell are statistically independent, allowing for a simpler estimation of the detection state variables to be carried out. The second class does not make the simplifying independence assumption and solves the estimation problem considering statistical dependence between the occupancy state variables. In all of the methods for occupancy grid estimation presented here, the sensor measurements are considered to be conditionally independent given the map. This assumption is valid for a *static world*, in which the only moving component is the sensor.

In this chapter, we present the general occupancy grid estimation process in Section 2.2. Estimation methods using independent occupancy states is presented in Section 2.3, while a method using dependent occupancy states is presented in Section 2.4.

2.2 Occupancy Grid Estimation – The General Problem

Occupancy grid estimation is a popular process for generating an occupancy map of an area given a set of measurements taken from that area [4, 6, 24]. The map is partitioned into a set of B disjoint *grid cells* $\{g_i\}_{i=1}^{B}$, all with the same shape and size. To each grid cell a binary occupancy

state indicator variable $b_i \in \{0, 1\}$ is attached with $b_i = 1$ indicating that a grid cell g_i is *occupied* (e.g., by a scatterer of radiation), and $b_i = 0$ indicating that g_i is *empty*. We call the set $\boldsymbol{b} = \{b_i\}_{i=1}^{B}$ the set of *cellular occupancies*, commonly referred to as a *map*. The map $\boldsymbol{b}$ can be any one of 2^B possible unique maps from the set of all possible maps $\mathbb{B}$.

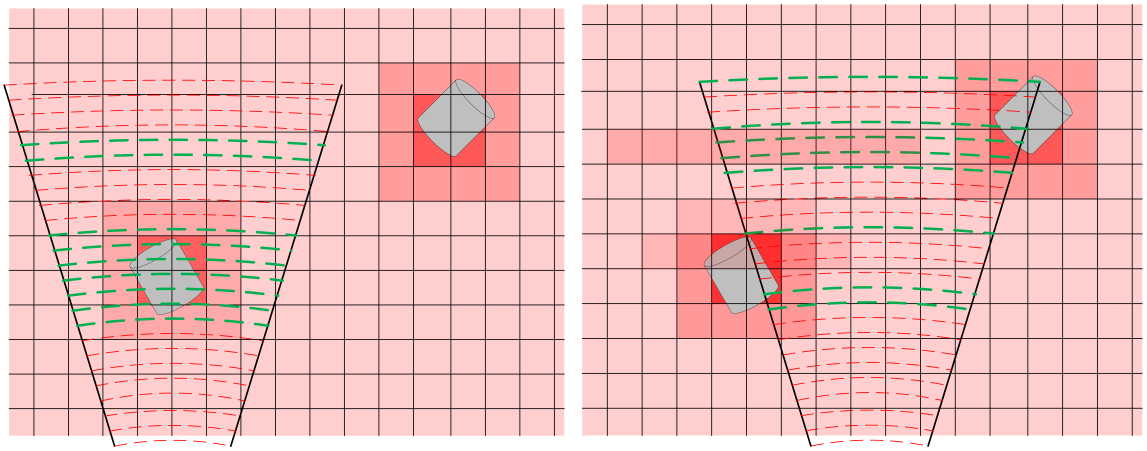
Now, given the measurement matrix $\boldsymbol{J}_S = \left[\boldsymbol{j}_1, \cdots, \boldsymbol{j}_s, \cdots, \boldsymbol{j}_S\right]$, consisting of a collection of measurement vectors (e.g., thresholded detection statistics) $\boldsymbol{j}_s = \left[j_{s,1}, \cdots, j_{s,K}\right] \in \mathcal{J}^K = \{0, 1\}^K$ for $s \in \{1, \cdots, S\}$ with K elements taken at time s, the estimation problem produces the set of marginal posterior probabilities, or occupancy grids (OGs) [3–5, 26], arranged as a vector

$$\boldsymbol{p} = \{p_{b|\boldsymbol{J}}(b_r = 1|\boldsymbol{J}_S)\}_{r=1}^{B}. \tag{2.1}$$

If only a single range measurement is produced at time s, then $K = 1$, j_s is a scalar, and $\boldsymbol{J}_s = \left[j_1, \cdots, j_s, \cdots, j_S\right]$.

Before we present the mathematical machinery of occupancy grid estimation, we explain the process with a simple figure mimicking a potential sonar setup. An environment to be interrogated is illustrated in Figure 2.1 by a large rectangular region which is partitioned into disjoint rectangular grid cells. The state of occupancy (occupied or empty) of each grid cell is represented by the respective elements of vector $\boldsymbol{b}$. There are two cylindrical objects depicted in grey, each occupying multiple grid cells. Hence, elements of $\boldsymbol{b}$ that correspond to those grid cells containing portions of a cylinder are equal to 1 while all the others are 0.

At each time (e.g., sonar ping) s, the moving sensor is effectively in a fixed position in its trajectory path with a conical field of view (FoV) that covers a subset of the grid cells within the occupancy map. In Figures 2.1a and 2.1b, the sensor cones are shown at slow times S and $S + 1$ with an axis perpendicular to the sensor path containing one or parts of both cylinders. In the case of sonar, the vertical axis captures range and the horizontal axis captures cross-range. For a sonar system, the width of the cone is directly related to the sonar beamwidth which is a function of the number of hydrophone array elements, aperture size, frequency, etc. The cone is divided



(a) Observation cone at time S, showing range locations of detections in green dashed lines and non-detections in red dashed lines. Note the two separate detections above the object relate to false alarm events.

(b) Observation cone at time $S + 1$, with updated occupancy grid from observations at time s.

Figure 2.1: An illustration of sequential Bayesian estimation of the occupancy grid for a sonar or radar application. Colored grid cells represent the occupancy grid, with the intensity of the color representing the probability of occupancy. Darker colors represent higher probabilities. Two cylindrical targets are depicted in grey. The presence of green dashed arcs in areas that do not overlap with the cylindrical object in the sonar cone indicate false positives (false alarms) and the presence of red dashed arcs in regions that do overlap with the cylindrical object indicate false negatives (missed detection).

into arc segments, shown in dashed lines, each intersecting multiple grid cells. Each arc segment corresponds to a range interval for which detection statistics are computed and thresholded to provide a desired false alarm rate p_{fa} and hence probability of detection p_{d}. A *range gate* comprises multiple range intervals.

The thresholded detection statistic, e.g., generated using ACE [22] for a linear phased sonar array, for each range interval produces a binary decision (random variable) $j_{s,k}$ for that range interval, where subscript s denotes the time at which the decision was made and subscript k is the range interval index. Each of the thresholded detection statistics is generated from one of four situations: true detection, missed detection, true miss, and false alarm. In Figures 2.1a and 2.1b, the red- and green-dashed arcs depict the range intervals for which the detector has produced a zero ($j_{s,k} = 0$) or a one ($j_{s,k} = 1$), respectively. Thus, the presence of a green-dashed arc in an area that does not overlap with the cylindrical object in the sonar cone indicates a false alarm and the presence of a red-dashed arc in a region that does overlap with the cylindrical object indicates a

missed detection. An example of a false alarm can be seen in the two isolated green-dashed arcs in top of the sensor cone in Figure 2.1a, while an example of a missed detection can be seen in the two isolated red-dashed arcs in the top of the sensor cone in Figure 2.1b.

As mentioned before, at each time s, the sensor produces a new vector $\boldsymbol{j}_s$ of binary decisions $j_{s,k}$, $k = 1, \ldots, K$ for the cone, where K is the number of range intervals that depends on the maximum detection range of the sensor. As the sensor moves along its path, the sequence of these collected vectors forms the matrix $\boldsymbol{J}_S = \left[\boldsymbol{j}_1, \cdots, \boldsymbol{j}_S\right]$ with columns that are all the past and present decision vectors.

The occupancy grid estimation process provides a sequential Bayesian framework for updating the posterior probability of the occupancy of each grid cell g_i with associated indicator random variable b_i, given the sequence of measurements in $\boldsymbol{J}_S$. Cellular posterior probabilities are updated after each time step. Referring to Figure 2.1a, the shade of pink in each grid cell represents our current estimate for the probability of occupancy of that grid cell, given the sequence of measurements $\boldsymbol{J}_S$. The darker the shade of pink in a cell, the higher the computed probability that the cell is occupied. Going from Figure 2.1a to Figure 2.1b, the sensor platform moves to a new location, and the measurement vector for the corresponding sensor cone is used to update of the occupancy probabilities. The shades of pink in Figure 2.1b represent the computed posterior probabilities of the occupancies of the grid cells, given all the measurements up to time $S + 1$. This is the general idea of occupancy grid estimation.

Recall that the mathematical notation used to presented the following ideas can be references in Section 1.4.

2.3 Classical Occupancy Grid Estimation Methods – Independent Occupancy States

The methods presented in this section [2–5, 26] assume that the occupancy state of each grid cell is statistically independent. This allows the joint density function of the collection of grid cell

occupancy state variables to be factored as the product of the individual cell occupancy density functions, which greatly simplifies the estimation problem.

2.3.1 The Original Method

The seminal work in OG estimation was performed by Hans P. Moravec and Alberto Elfes [2–5] during their time at Carnegie Mellon University in the early- to mid-eighties. Their research started as a method for mapping range measurements, taken from a Polaroid sensor range finder at known locations, to a 2-dimensional map that conveys the notion of occupancy for any particular observed point on the map.

The problem they wanted to solve was to find the map $\boldsymbol{b}^*$ that is the maximum *a posteriori* (MAP) estimate of all possible maps $\boldsymbol{b} \in \mathbb{B}$, given a collection of measurements taken from the area under investigation. Their initial approach used a set of heuristics to combine measurements that were taken sequentially. Elfes later rigorously formulated the problem as a sequential Bayesian update that is described below [4].

Given the set of observations up to time S, $\boldsymbol{J}_S = \left[\boldsymbol{j}_1, \cdots, \boldsymbol{j}_S\right]$, and the set of all possible maps $\mathbb{B}$ with $|\mathbb{B}| = 2^B$, the MAP estimation problem that can be formulated as:

$$\boldsymbol{b}^* = \arg\max_{\boldsymbol{b}\in\mathbb{B}} p_{\boldsymbol{b}|\boldsymbol{J}}(\boldsymbol{b}|\boldsymbol{J}_S). \tag{2.2}$$

In the original formulation, each j_s is considered to be a single scalar range measurement. The map $\boldsymbol{b}$ is modeled by a Markov Random Field of order 0 [27], implying that each of the grid cells b_r composing $\boldsymbol{b}$ can be considered statistically independent. This allows (2.2) to be rewritten as

$$\boldsymbol{b}^* = \arg\max_{\boldsymbol{b}\in\mathbb{B}} \prod_{r=1}^{B} p_{b|\boldsymbol{J}}(b_r|\boldsymbol{J}_S), \tag{2.3}$$

which greatly simplifies the problem to estimating the individual grid cell occupancy probabilities. Using Bayes' rule, and assuming that measurements j_s are conditionally independent given the map $\boldsymbol{b}$, $p(b_r|\boldsymbol{J}_S)$ (i.e., $p_{j|\boldsymbol{b}}(j_S|\boldsymbol{b}, \boldsymbol{J}_{S-1}) = p_{j|\boldsymbol{b}}(j_S|\boldsymbol{b})$) can be represented as

$$p(b_r|\boldsymbol{J}_S) = \frac{p_{j|b}(j_S|b_r)p_{b|\boldsymbol{J}}(b_r|\boldsymbol{J}_{S-1})}{p_{j|\boldsymbol{J}}(j_S|\boldsymbol{J}_{S-1})} = \eta \sum_{\boldsymbol{b}\in\mathbb{B}(r,\beta)} p_{j|\boldsymbol{b}}(j_S|\boldsymbol{b})p_{\boldsymbol{b}|\boldsymbol{J}}(\boldsymbol{b}|\boldsymbol{J}_{S-1}), \tag{2.4}$$

where $\mathbb{B}(r,\beta)$ is the set of all maps $\boldsymbol{b}$ such that the rth element is equal to β, i.e., $\mathbb{B}(r,\beta) = \left\{\boldsymbol{b} = [b_1,\cdots,b_r,\cdots,b_B]|b_r = \beta\right\}$, and η is a normalization constant that ensures $p_{b|\boldsymbol{J}}(b_r|\boldsymbol{J}_S)$ is a valid probability mass function.

The original sensor measurement model used, $p_{j|\boldsymbol{b}}(j_S|\boldsymbol{b})$, was one of two: an ideal sensor, and a Gaussian sensor. The ideal sensor produces a delta at the true range of a target, and the Gaussian sensor produces a measurement drawn from a Gaussian distribution with mean μ_S and constant variance σ^2 when conditioned on the map $\boldsymbol{b}$.

When a measurement $\boldsymbol{j}_S$ is recorded, an update to the occupancy grid estimate is made for all grid cells in an ε *range gate* around the range associated with $\boldsymbol{j}_S$, where ε is directly related to σ^2. In other words, all grid cells within the ε range gate centered at $\boldsymbol{j}_S$ are updated independently with a measurement model indicating that a reflection was made at that grid cell. Additionally, all grid cells at a range closer than the ε range gate are independently updated with a measurement model indicating that there was not a reflection made at that grid cell. This process is illustrated in Figure 2.2.

2.3.2 Log-Odds Method

The formulation for performing occupancy grid estimation with a log-odds expression [6] differs from that presented in Section 2.3.1 in that it removes the need to directly calculate some of the harder quantities, i.e., $p_{j|\boldsymbol{J}}(j_S|\boldsymbol{J}_{S-1})$. Again, the assumption is made that all measurements j_s are in the form of scalar measurements, and the measurement model can be described by either an ideal sensor or a Gaussian sensor, both described in Section 2.3.1. This assumption allows all measurements on a grid cell to be conditionally independent given that grid cell, which provides a simplification of the conditional probability mass functions.

The formulation of the posterior marginal probability that the rth grid cell is occupied given the set of measurements $\boldsymbol{J}_S$ can be written as:

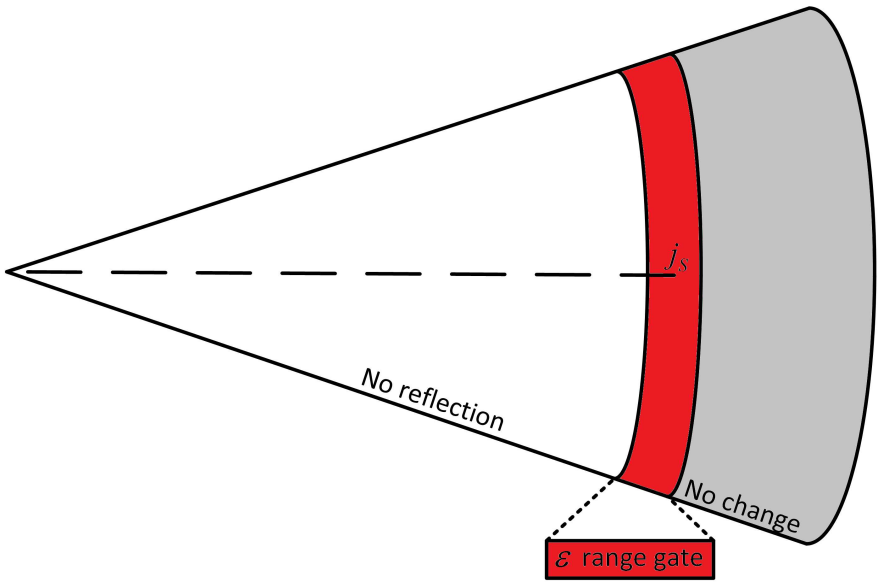


Figure 2.2: Model for selecting grid cells to update with each measurement j_S. Grid cells in grey region are not updated. Grid cells in white region are updated with an 'empty' measurement. Grid cells in red region are updated with an 'occupied' measurement.

$$p_{b|\boldsymbol{J}}(b_r = 1|\boldsymbol{J}_S) = \frac{p_{b,\boldsymbol{J}}(j_S, \boldsymbol{J}_{S-1}, b_r = 1)}{p_{\boldsymbol{J}}(j_S, \boldsymbol{J}_{S-1})} = \frac{p_{j|\boldsymbol{J},b}(j_S|\boldsymbol{J}_{S-1}, b_r = 1)p_{b|\boldsymbol{J}}(b_r = 1|\boldsymbol{J}_{S-1})}{p_{j|\boldsymbol{J}}(j_S|\boldsymbol{J}_{S-1})} \tag{2.5}$$

$$= \frac{p_{b|j}(b_r = 1|j_S)p_j(j_S)}{p_b(b_r = 1)} \frac{p_{b|\boldsymbol{J}}(b_r = 1|\boldsymbol{J}_{S-1})}{p_{j|\boldsymbol{J}}(j_S|\boldsymbol{J}_{S-1})}. \tag{2.6}$$

Writing a similar equation for $p_{b|\boldsymbol{J}}(b_r = 0|\boldsymbol{J}_S)$ and then dividing (2.5) by it gives:

$$\frac{p_{b|\boldsymbol{J}}(b_r = 1|\boldsymbol{J}_S)}{p_{b|\boldsymbol{J}}(b_r = 0|\boldsymbol{J}_S)} = \frac{p_{b|j}(b_r = 1|j_S)p_j(j_S)p_{b|\boldsymbol{J}}(b_r = 1|\boldsymbol{J}_{S-1})}{p_b(b_r = 1)p_{j|\boldsymbol{J}}(j_S|\boldsymbol{J}_{S-1})} \frac{p_b(b_r = 0)p_{j|\boldsymbol{J}}(j_S|\boldsymbol{J}_{S-1})}{p_{b|j}(b_r = 0|j_S)p_j(j_S)p_{b|\boldsymbol{J}}(b_r = 0|\boldsymbol{J}_{S-1})}$$

$$= \frac{p_{b|j}(b_r = 1|j_S)p_{b|\boldsymbol{J}}(b_r = 1|\boldsymbol{J}_{S-1})}{p_b(b_r = 1)} \frac{p_b(b_r = 0)}{p_{b|j}(b_r = 0|j_S)p_{b|\boldsymbol{J}}(b_r = 0|\boldsymbol{J}_{S-1})}. \tag{2.7}$$

Taking the natural log of (2.7) yields the log-odds representation:

$$l_r^n = \log \frac{p_{b|\boldsymbol{J}}(b_r = 1|\boldsymbol{J}_S)}{p_{b|\boldsymbol{J}}(b_r = 0|\boldsymbol{J}_S)}$$

$$= \log \frac{p_{b|j}(b_r = 1|j_S)}{1 - p_{b|j}(b_r = 1|j_S)} + \log \frac{1 - p_b(b_r = 1)}{p_b(b_r = 1)} + \log \frac{p_{b|\boldsymbol{J}}(b_r = 1|\boldsymbol{J}_{S-1})}{1 - p_{b|\boldsymbol{J}}(b_r = 1|\boldsymbol{J}_{S-1})}$$

$$= \log \frac{p_{b|j}(b_r = 1|j_S)}{1 - p_{b|j}(b_r = 1|j_S)} + \log \frac{1 - p_b(b_r = 1)}{p_b(b_r = 1)} + l_r^{n-1}. \tag{2.8}$$

The second term in (2.8) is the log-odds ratio of the prior on the occupancy for grid cell g_r. An uninformative prior ($p_b(b_r = 1) = p_b(b_r = 0) = 0.5$) is typically assumed, driving that term to 0. The marginal posterior probability can be recovered simply by the following:

$$\begin{aligned} l_r^n &= \log \frac{p_{b|\boldsymbol{J}}(b_r = 1|\boldsymbol{J}_S)}{p_{b|\boldsymbol{J}}(b_r = 0|\boldsymbol{J}_S)} = \log \frac{p_{b|\boldsymbol{J}}(b_r = 1|\boldsymbol{J}_S)}{1 - p_{b|\boldsymbol{J}}(b_r = 1|\boldsymbol{J}_S)} = \log \frac{1 - (1 - p_{b|\boldsymbol{J}}(b_r = 1|\boldsymbol{J}_S))}{1 - p_{b|\boldsymbol{J}}(b_r = 1|\boldsymbol{J}_S)} \\ &= \log \left[\frac{1}{1 - p_{b|\boldsymbol{J}}(b_r = 1|\boldsymbol{J}_S)} - 1 \right] \\ \Longrightarrow p_{b|\boldsymbol{J}}(b_r = 1|\boldsymbol{J}_S) &= 1 - \frac{1}{1 + e^{l_r^n}} \end{aligned} \tag{2.9}$$

At each time step, equation (2.8) is recursively updated using the sensor model and current measurement, and the associated grid cells in the occupancy grid is updated using (2.9). The process of updating the grid cells follows the same type of heuristic that is illustrated in Figure 2.2.

2.4 Classical Occupancy Grid Estimation – Dependent Occupancy States

The methods presented in this section [6, 28] assume that the occupancy state of each grid cell is statistically dependent. These methods provide a more accurate estimate of the true occupancy, even when there are inconsistent measurements.

2.4.1 Data Association Method

This approach was proposed by Thrun [6] as the first occupancy grid estimation method that considers statistical dependence between occupancy states. It is formulated by first considering that two classes of measurements, *random* and *non-random*, can come from a sensor cone. For the random class of measurements, the detector has a certain probability that it will produce a false alarm due to sensor noise modeled by p_{rand}. For the non-random class of measurements, each occupied cell in the cone has a probability of reflecting the incident signal, modeled as p_{hit}. This

method assumes that only one detection from within the measurement cone can be distinguished, and that the returned measurement is assumed to be a single range measurement per sensor cone.

Consider a sensor cone with K_S objects in it at time S. The ith object has an associated distance, $d_{S,i}$, with $d_{S,1} \leq d_{S,2} \leq \cdots \leq d_{S,K_S}$ from the AUV. Now also consider a binary vector of association variables, or correspondence variables, $c_S = [c_{S,*}, c_{S,0}, c_{S,1}, \cdots, c_{S,K_S}]$ that identify what class, random or non-random described previously, the measurement is from within the cone. At arbitrary time S, all elements in c_S are 0 except for a single element, indicating exactly what the true source of the detection is from. In other words, $\sum_k c_{S,k} = 1$. Element $c_{S,*}$ represents a detection from the random class, $c_{S,0}$ indicates no detection, and $c_{S,k}$ for $1 < k \leq K_S$ indicates a detection from the kth object, a non-random class detection.

Let $r_{\max}$ indicate the max range a sensor can observe, and r_S be a scalar measurement containing the first reported detection in a sensor cone. Then, the likelihood models are as follows [6]:

$$p_{r|\boldsymbol{b},c}(r_S|\boldsymbol{b}, c_{S,*} = 1) = \begin{cases} \frac{1}{r_{\max}} & 0 \leq r_S < r_{\max} \\ 0 & \text{otherwise} \end{cases} \tag{2.10}$$

$$p_{r|\boldsymbol{b},c}(r_S|\boldsymbol{b}, c_{S,0} = 1) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(r_S - r_{\max})^2}{2\sigma^2}\right) & 0 \leq r_S < r_{\max} \\ 0 & \text{otherwise} \end{cases} \tag{2.11}$$

$$p_{r|\boldsymbol{b},c}(r_S|\boldsymbol{b}, c_{S,k} = 1) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(r_S - d_{S,k})^2}{2\sigma^2}\right) & 0 \leq r_S < r_{\max} \\ 0 & \text{otherwise} \end{cases} \tag{2.12}$$

Equation (2.10) can be written in the form of a Gaussian allowing the combination of the above likelihoods into one likelihood function that is consistent with all three, i.e.

$$p_{r|\boldsymbol{b},c}(r_S|\boldsymbol{b}, c_S) = p_{r|\boldsymbol{b},c}(r_S|\boldsymbol{b}, c_{S,*} = 1)^{c_{S,*}} \prod_{k=0}^{K_S} p_{r|\boldsymbol{b},c}(r_S|\boldsymbol{b}, c_{S,k} = 1)^{c_{S,k}}. \tag{2.13}$$

The distribution of the correspondence variables when conditioned on the map, $p_{c|\boldsymbol{b}}(c_S|\boldsymbol{b})$, can be written as [6]

$$p_{c|\boldsymbol{b}}(c_S|\boldsymbol{b}) = p_c(c_S) = \begin{cases} p_{\text{rand}} & c_{S,*} = 1 \\ (1-p_{\text{rand}})(1-p_{\text{hit}})^{K_S} & c_{S,0} = 1 \\ (1-p_{\text{rand}})(1-p_{\text{hit}})^{k-1}p_{\text{hit}} & c_{S,k} = 1 \end{cases}. \tag{2.14}$$

Now writing the conditional joint between the measurement and association vector we get

$$p_{r|\boldsymbol{b}}(r_S|\boldsymbol{b}) = \sum_{c_S} p_{r,c|\boldsymbol{b}}(r_S, c_S|\boldsymbol{b}) = \sum_{c_S} p_{r|\boldsymbol{b},c}(r_S|\boldsymbol{b}, c_S)p_{c|\boldsymbol{b}}(c_S|\boldsymbol{b}). \tag{2.15}$$

Assuming targets are stationary and only the sensor moves, and letting $\boldsymbol{R}_S$ be the collection of range measurements r_s for $s = 1, \cdots, S$, the measurements are conditionally independent given the map $\boldsymbol{b}$, implying

$$p_{\boldsymbol{R}|\boldsymbol{b}}(\boldsymbol{R}_S|\boldsymbol{b}) = \prod_{s=1}^{S} \sum_{c_S} p_{r|\boldsymbol{b},c}(r_s|\boldsymbol{b}, c_S)p_{c|\boldsymbol{b}}(c_S|\boldsymbol{b}). \tag{2.16}$$

The treatment of applying the logarithm to (2.16) and then applying the expectation maximization (EM) algorithm [28] ends up solving for the most likely map $\boldsymbol{b}$ that produced the observed measurements.

There are a couple of short comings with this approach. The first is that the use of the EM algorithm ends up providing a map that is strictly 1s and 0s, and not a collection of occupancy distributions. Some heuristics are added to the final formulation in a quasi leave-one-out calculation for each of the marginal posterior occupancy probabilities. The second is the limitation to applications with scalar measurements only, which makes this method impractical for use in our sonar problem.

2.5 Conclusion

In this chapter, we reviewed three different methods for performing occupancy grid estimation. These methods differ through their modeling of statistical dependence between occupancy state

variables. The first two methods which assumed independent occupancy states are limited in applicability, as targets may occupy multiple grid cells but are not modeled as such in these methods. They are also derived by different means, the first through an optimal Bayesian estimation process and the second through the so-called log-odds formulation. The third method lifted the independence assumption, but still was not fully applicable to our problem. All three formulations tend to have fast implementation, but are only suited to applications with scalar range measurements. Additionally, these three methods show the evolution of occupancy grid estimation as the popularity of occupancy grids grew. Moreover, there has been evolution of the sensor models that are used in each of the formulations, typically with the sensor models being tailored to specific real-world sensors. This has allowed for various optimization and shortcuts when implementing the aforementioned methods. We will show in Chapter 4 that the occupancy grid estimation process can be generalized to estimate the marginal and joint probabilities of occupancy given a collection of measurement, while allowing for vector-valued measurements, dependent occupancy states, and using a generalized sensor model that is able to model many real-world sensors.

Chapter 3

Data Description and Preprocessing

3.1 Introduction

The data used to help develop and test the methods presented in this document is primarily synthetically generated data, emulating low frequency sonar data for various hydrophone arrays. The data used in our experiments was generated using aan underwater sonar simulator, namely the Personal Computer Shallow Water Acoustic Toolset (PC SWAT) developed at the Naval Surface Warfare Center Panama City (NSWC PCD) [29], in order to determine the success of the developed algorithms for underwater detection and classification for optimal vehicle path planning. PC SWAT allows for the experiment designer to place targets into the sonar field, whether in the water column, buried or proud of the bottom. Incident signals are projected from the hydrophone array into the sonar field and reflected back towards the hydrophone array if objects exist within the ensonification area of the incident signal. All of the data generated by PC SWAT is complex-valued numbers representing temporal samples of the in-phase (I) and quadrature (Q) channels from each hydrophone in the array.

Prior to being used with any of the sequential detection algorithms presented in this dissertation, the data is preprocessed by the adaptive coherence estimator (ACE) detector [22] in order to generate detection statistics for each of the sampled sonar returns. Each sampled sonar return is associated with some time delay from the time the incident signal was emitted. The processing through ACE provides invariance to scaling of the noise magnitude. This preprocessing provides a number of benefits, including the ability to perform adaptive beam steering and reducing the effect of the noise floor on detection performance. Preprocessing by ACE also generates a binary vector for each collected measurement required by the sensor model in our occupancy grid formulation presented in Chapter 4.

PC SWAT generated data is used in sequential detection state estimation (cf. Chapter 4), sequential classification state estimation (cf. Chapter 5), and the calculation of their distributions as presented in Chapter 6.

This chapter is organized as follows. Section 3.2 provides details of the PC SWAT generated data, and presents illustrations of said data. Section 3.3 presents the details of the ACE detector, and how the PC SWAT generated data is processed using ACE.

3.2 Simulated Sonar Data

Due to the difficulties (time, cost, etc.) of collecting real data using autonomous underwater vehicles (AUV), e.g., a Bluefin AUV equipped with various low and high frequency sonar, synthetic sonar data was used in all experiments presented in this dissertation. The PC SWAT a cutting-edge, physics-based sonar simulator that models scattering from the target by a combination of the Kirchhoff approximation [30] and the geometric theory of diffraction. Propagation of sound into a marine sediment with ripples is described by an application of Snell's law and second order perturbation theory in terms of Bragg scattering [29]. PC SWAT has been used to produce simulations providing *exemplar* template measurements that closely match real data generated by real shallow water sonar systems [30].

In addition to providing realistic sonar returns, PC SWAT allows for total control of the simulation environment, including: waveform type, environmental properties (scattering, sediment type, sound speed, etc.), AUV properties (heading, position, motion, etc.), and much more. To better display the type of data that is generated by PC SWAT, a collection of simulations are presented in Figure 3.1 below. The general sonar setup used by PC SWAT in our experiments includes:

1. targets of different shapes, sizes, and orientations buried under or proud on the bottom,

2. a single AUV placed some distance above the bottom and traveling with fixed velocity along a path (predetermined or not) taking measurements at fixed distance intervals,

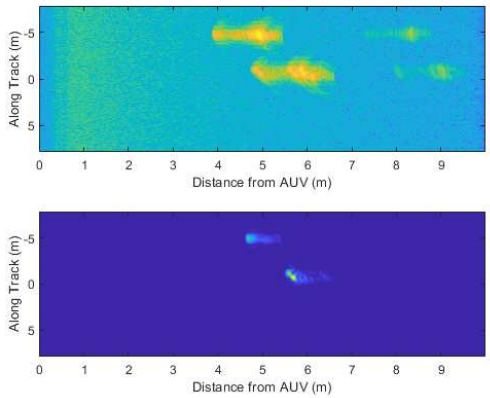
3. horizontal or vertical hydrophone array attached to the AUV with a positive depression angle (pointed downwards from the horizon) typically with a single projector and multiple receivers, and

4. bottom types varying from sand and mud to gravel and rock.

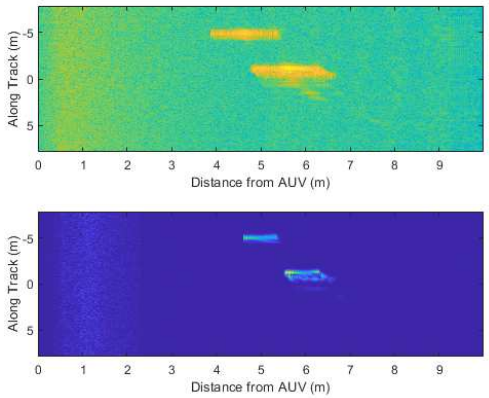
The upper image in each figure is the magnitude squared of the raw stave data, while the lower image is after applying a matched filter using the incident waveform. The AUV follows a straight path in each of the simulations except for the simulation in Figure 3.1e, which has a circular path. The hydrophone array spacing in all experiments is $1/2$ the wavelength of the center frequency of the incident signal, and the hydrophone array has a 20 degree downward depression angle. The simulations in Figure 3.1a and 3.1b each have 2 cylindrical targets sitting proud on the seafloor, with 5 hydrophones in a horizontal orientation. The simulations in Figure 3.1c and 3.1d have three targets all residing in a water column, and 3 hydrophones in a vertical orientation. The remaining simulation in Figure 3.1e has a single target in the water column, and 3 hydrophones in a horizontal orientation. All remaining details of the simulations (center frequency f_c, bandwidth BW, sediment type, chirp type, and number of hydrophone elements) are captured in the captions of their respective figures. The bottom matched filtered images of Figures 3.1c and 3.1d appear to be missing two targets that are present in each of their associated magnitude images. This is merely an artifact of the image scaling in the matched filtered image, as the targets that show up in the matched filtered image has a magnitude significantly greater than the other two targets. In reality, all three targets are seen in both matched filtered images.

3.3 Adaptive Coherence Estimator – Preprocessing

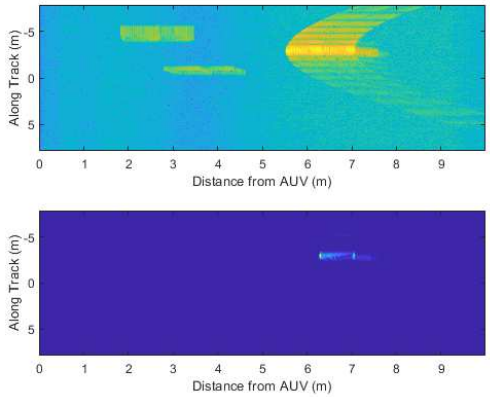
To generate detection statistics for the occupancy grid estimator (sequential detection state estimator presented in Chapter 4), the matched filtered data is applied to ACE. ACE is a special case of the adaptive constant false alarm rate (CFAR) matched subspace detector (MSD) [22] which uses an array steering vector as the subspace model and uses a sample covariance matrix to estimate the true noise covariance matrix in an adaptive manner. In general, the problem of detection a



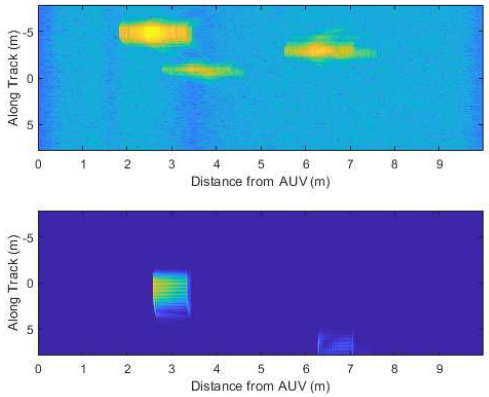
(a) f_c = $12kHz$, BW = $20kHz$, LFM chirp, 5 horizontal elements, sandy bottom



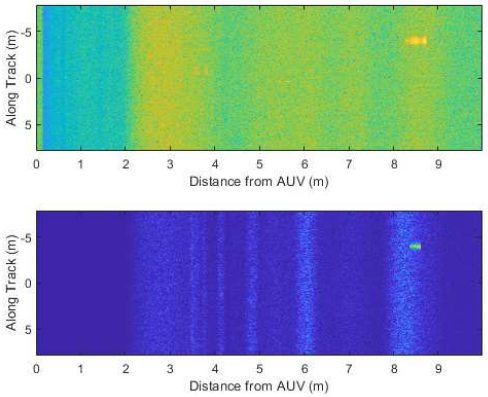
(b) f_c = $60kHz$, BW = $20kHz$, LFM chirp, 5 horizontal elements, sandy bottom



(c) f_c = $40kHz$, BW = $10kHz$, LFM chirp, 3 vertical elements, sandy bottom



(d) f_c = $20kHz$, BW = $10kHz$, LFM chirp, 3 vertical elements, gravely bottom



(e) f_c = $120kHz$, BW = $40kHz$, LFM chirp, 3 horizontal elements, muddy bottom

Figure 3.1: PC SWAT simulations using different configurations for center frequency f_c, bandwidth BW, bottom type, and orientation and number of hydrophones in the array.

subspace signal can be described as follows. The parameter $\boldsymbol{\theta}$ defines a signal $\boldsymbol{s} = \mu\boldsymbol{H}\boldsymbol{\theta}$ in the signal subspace $\langle \boldsymbol{H} \rangle$ with $\mu = 0$ or $\mu \neq 0$ in the absence or presence of the signal in the subspace, respectively. The signal is passed through a channel which adds noise $\sigma\boldsymbol{n}$ with covariance matrix $\boldsymbol{R}_0$ to the signal to produce a vector-valued measurement $\boldsymbol{y} = \boldsymbol{s} + \sigma\boldsymbol{n}$. Clearly the observation is complex normal, i.e., $\boldsymbol{y} \sim CN(\mu\boldsymbol{H}\boldsymbol{\theta}, \sigma^2\boldsymbol{R}_0)$. The CFAR MSD problem tests hypothesis $\mathcal{H}_0 : \mu = 0$ against hypothesis $\mathcal{H}_1 : \mu \neq 0$ when $\boldsymbol{H}$ and $\boldsymbol{R}_0$ are known but μ, σ^2, and $\boldsymbol{\theta}$ are unknown. The adaptive version of the CFAR MSD that is utilized in this document, ACE, replaces the known covariance $\boldsymbol{R} = \sigma^2\boldsymbol{R}_0$ by $\boldsymbol{S} = \boldsymbol{Y}\boldsymbol{Y}^H/(N-1)$, the unbiased estimator of covariance under the null hypothesis (sample covariance), where $\boldsymbol{Y}$ is the collection of reference measurements $\boldsymbol{y}$ taken under the null hypothesis and arranged into matrix form.

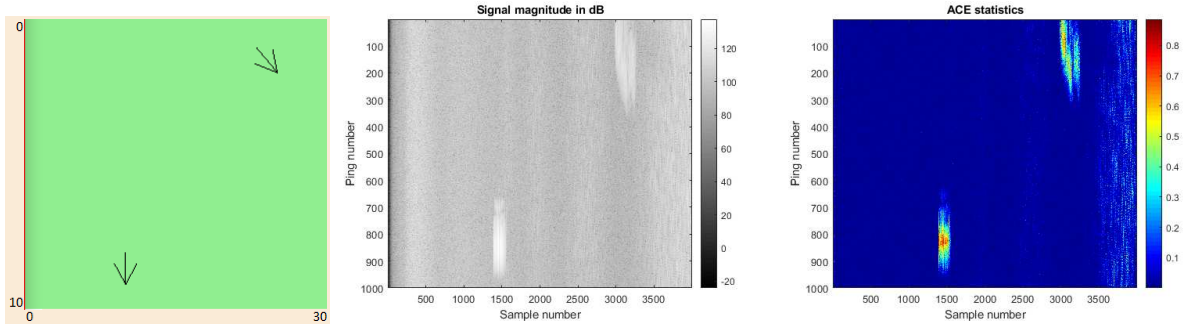
Given the observation vector $\boldsymbol{y}$, the ACE statistic [22, 31] takes the form of

$$\beta = \frac{\boldsymbol{z}^H\boldsymbol{P}_G\boldsymbol{z}}{\boldsymbol{z}^H\boldsymbol{z}} = \frac{\boldsymbol{y}^H\boldsymbol{S}^{-1}\boldsymbol{H}(\boldsymbol{H}^H\boldsymbol{S}^{-1}\boldsymbol{H})^{-1}\boldsymbol{H}^H\boldsymbol{S}^{-1}\boldsymbol{y}}{\boldsymbol{y}^H\boldsymbol{S}^{-1}\boldsymbol{y}}, \tag{3.1}$$

where H is the Hermetian operator, $\boldsymbol{z} = \boldsymbol{S}^{-1/2}\boldsymbol{y}$ is the whitened version of the measurement $\boldsymbol{y}$, i.e., $\boldsymbol{z} \sim CN(\mu\boldsymbol{S}^{-1/2}\boldsymbol{H}\boldsymbol{\theta}, I)$, $\boldsymbol{G} = \boldsymbol{S}^{-1/2}\boldsymbol{H}$, and $\boldsymbol{P}_G = \boldsymbol{G}(\boldsymbol{G}^H\boldsymbol{G})^{-1}\boldsymbol{G}^H$ is the projection matrix onto the subspace $\langle \boldsymbol{G} \rangle$. For ACE, the subspace is $\boldsymbol{H} = \boldsymbol{\psi} = \begin{bmatrix} 1, & e^{j\omega}, & \ldots, & e^{j\omega(M-1)} \end{bmatrix}^T$, i.e., the array steering vector typically used in sonar and radar applications. In this case, the ACE statistic which measures the squared cosine of the angle between $\boldsymbol{y}$ and the subspace $\langle \boldsymbol{\psi} \rangle$ in the adapted coordinate system becomes

$$\beta = \frac{|\boldsymbol{\psi}^H\boldsymbol{S}^{-1}\boldsymbol{y}|^2}{(\boldsymbol{y}^H\boldsymbol{S}^{-1}\boldsymbol{y})(\boldsymbol{\psi}^H\boldsymbol{S}^{-1}\boldsymbol{\psi})} \tag{3.2}$$

The ACE statistic exhibits many nice properties including invariance to scaling of the training data used to generate $\boldsymbol{S}$, and scaling of the measurement $\boldsymbol{y}$. This is particularly nice in radar, sonar, and communication problems where the magnitude of the measurement and training data may be significantly different. It is the maximum likelihood detector for detecting a subspace signal in noise of unknown scale [31], and it is uniformly most powerful, which is to say no other detector



(a) Target Layout **(b)** Stave Data **(c)** ACE Statistics

Figure 3.2: ACE example simulation

of equal or smaller false alarm probability than ACE can have a higher detection probability. Such optimality statements are rare in detection theory. The reader is referred to [22, 32] for more information regarding the CFAR MSD and ACE.

Figure 3.2 illustrates an example of applying ACE to PC SWAT data. The simulation used an $M = 10$ element uniform linear hydrophone array, $f_c = 20$ kHz center frequency, $BW = 10$ kHz, and a linearly frequency modulated (LFM) chirp sampled at 100 kHz.

In this simulation, two cylindrical mine-like targets were placed on the sandy seafloor with their orientation and placement illustrated in Figure 3.2a. The array elevation was 2 meters above the seafloor. The AUV carrying the array traveled a straight path over 10 meters, passing by both of the targets. Figure 3.2b shows the magnitude of the stave data in dB, and Figure 3.2c shows the calculated ACE statistics. The observations seen in the stave data shows the general position and orientation of each of the targets. The observations, once passed through the ACE detector, show a much better representation of the position, size, and orientation of the two targets, and notably, much of the background clutter has been removed when compared to the stave data image.

3.4 Conclusion

Sonar data generated by means of PC SWAT simulation is used as the main source of data in the experiments presented in this work. This synthetic sonar data provided a low-cost option for generating the necessary sequences of realistic sonar data for a multitude of sonar array configu-

rations and setups. The ACE detector is then used as a physical layer detector for preprocessing the raw sonar data. The output from the ACE detector is then applied to the proposed sequential detection algorithms which is described in the next chapter.

Chapter 4

A New Sequential Detection and Occupancy Grid Estimation

4.1 Introduction

In this chapter, we are concerned with the sequential detection state estimation portion of our problem, seen in the lower right image of Figure 1.1, and we extend the traditional occupancy grid (OG) framework for vector-valued measurements. The original formulation presented in Sections 2.3.1 and 2.3.2 solves the problem by assuming the occupancy state of each grid cell in an occupancy grid is statistically independent from one another, allowing for the factorization of the joint probability density into the product of individual densities. This allowed for a formulation that was able to be solved quickly while admitting an unrealistic model for the probability distribution of occupancy states given a set of measurements. This is, in part, because objects that reflect the transmitted waveform typically occupy more than a single grid cell. Furthermore, in the case of sonar, multiple objects can be the same distance from the sensor, making it ambiguous which grid cell the waveform was reflected from.

The formulation in Section 2.4.1 dealt with these problems by estimating the joint probability density by means of performing maximum *a posteriori* (MAP) estimation, but was developed with sensors that provide a single range reading per measurement. This formulation improves upon the original occupancy grid formulations, but does not provide a general solution for estimating the joint probability density for vector-valued measurements.

The formulation presented in this chapter provides a statement of, and a solution for solving, the original occupancy grid problem of estimating the joint probability density for the occupancy states of the grid cells on a map given a set of measurements taken on that map. This formulation explicitly shows how to calculate both the marginal posterior probabilities, also estimated by the

other methods, as well as the full joint probabilities. The estimation of the joint probabilities is provided through the use of a sensor model exploiting a network of binary asymmetric channels (BACs). The use of the BAC network allows for the modeling of almost any real-world sensor model. Additionally, this formulation is able to consider vector-valued measurements, making it suitable for use with all types of sensors.

This chapter is organized as follows. Section 4.2 introduces the new occupancy grid estimation method that we developed using a network of binary asymmetric channel models. The binary asymmetric channel model is discussed in Section 4.2.1, and describes the interactions between each measurement and the occupancy state of each grid cell. The sequential Bayesian update procedure is developed in Section 4.2.2, and introduces how to calculate the both the marginal and joint occupancy occupancy distributions. Section 4.2.3 describes how one may choose the BAC transition probabilities to model real-world sensory systems. Two special cases for reducing the computation complexity of the overall system are given in Section 4.2.4.

4.2 A New Occupancy Grid Estimation Using Binary Asymmetric Channel Sensor Model

Before introducing the following formulation, we invite the reader to refresh themselves with the notation, presented in Section 1.4. To produce the estimate of the vector of cellular posterior probabilities $\boldsymbol{p}$, most of the existing methods [2,4,5,7,12] factor the joint distribution of the cellular occupancies as $p_{\boldsymbol{b}}(\boldsymbol{b}) = \prod_i p_b(b_i)$. Similarly, when conditioned on the collection of measurement vectors in $\boldsymbol{J}_S$, the joint distribution is assumed to be factored as $p_{\boldsymbol{b}|\boldsymbol{J}}(\boldsymbol{b}|\boldsymbol{J}_S) = \prod_i p_{b|\boldsymbol{J}}(b_i|\boldsymbol{J}_S)$ which implies that the elements in the cellular occupancies are conditionally independent. As pointed out before, these simplifying assumptions limit the applicability of such methods in many practical situations. The general formulation (GF) presented in this section is intended to lift these restrictive assumptions.

Throughout the development of this formulation, we assume that the sensor position at each time is reasonably accurate, i.e., the odometry errors are small enough to ignore. This is a valid

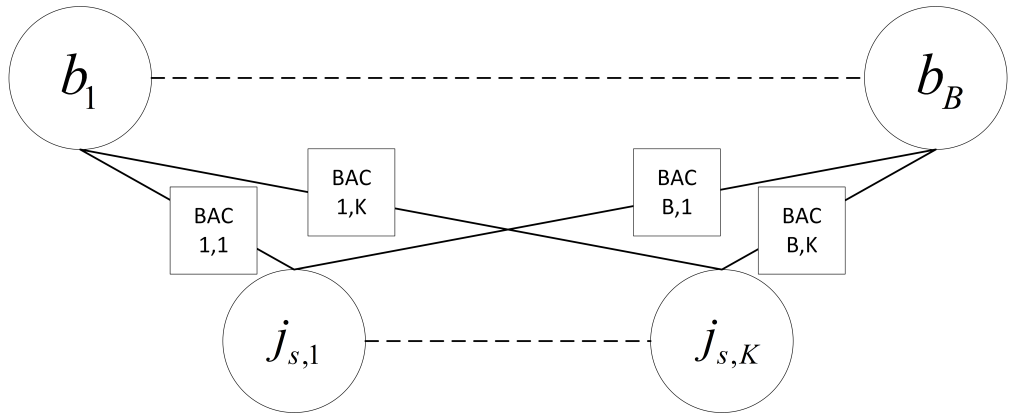
assumption for sensory system deployed on many autonomous platforms, where a low velocity and small measurement error from the odometry sensors satisfies the ego-motion estimation for the vehicle.

4.2.1 Model for Grid Cell and Measurement Interactions

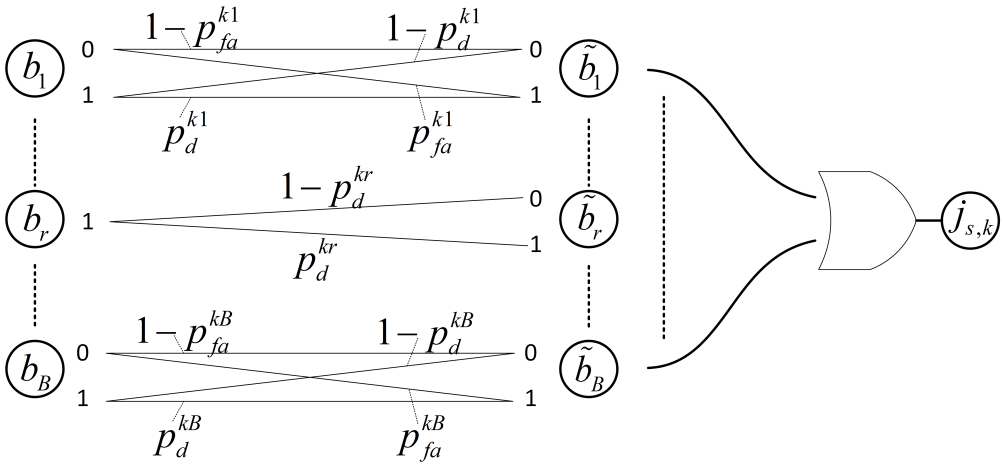
The sensor model presented in this section provides a way to probabilistically solve the data association problem of determining which grid cells are occupied given the measurements. A network of binary asymmetric channels (BACs) [23] are used to model the relationship between grid cell occupancy states and measurements. The BAC outputs, the so-called *virtual occupancies*, are latent variables, and are merely used to model the measurements.

The occupancy state b_i of each grid cell influences each $j_{s,k}$ in measurement vector $\boldsymbol{j}_s$. This relationship can be represented by a directed acyclic graph (DAG) of Figure 4.1a, where the information flow is from b_i to $j_{s,k}$. The influence that the occupancy state of each grid cell has on each measurement can be modeled by transition probabilities of a BAC. These transition probabilities can be functions of the distance between the grid cell and location of the measurement, or any other physical quantity depending on the employed sensory system, and are discussed further in Section 4.2.3. Recalling that $j_{s,k}$, $k = 1, \cdots, K$ are binary-valued decisions from the detector at time s, one can view a BAC as the device that models bit flips associated with missed detection and false alarm for each grid cell.

The binary occupancy information transmitted from each b_i through a BAC and received at each $j_{s,k}$ node are logical OR'd together to produce the measurement $j_{s,k}$ for all k, with different BAC transition probabilities for each (i, k) pair. That is, letting the virtual occupancy $\tilde{b}_i$ be the output of the BAC for each grid cell c_i, then $j_{s,k} = \sum_{i=1}^{B} \tilde{b}_i$ where the sum is Boolean. The justification for the choice of the OR gate model is the following. An occupied grid cell reflects a transmitted signal back to a detector, and if any of the signals received at the sensor has enough power to trigger the detector then it will be declared as detection. This behavior is analogous to many signals being transmitted over electrical wires and connected to the circuitry that composes an OR gate.



(a) Causal chain of interaction between cell occupancies b_r and measurements $j_{s,k}$.



(b) Binary asymmetric channel behavior for each measurement $j_{s,k}$, conditioned on $b_r = 1$.

Figure 4.1: Interaction and measurement model

Figure 4.1b illustrates the graphical representation of this model in which the indicator $b_r = 1$ is assigned whereas the remaining $b_i \neq b_r$ may be either a 1 or 0 before transmission through the BACs.

4.2.2 Sequential Bayes' Updating

In this section, BAC models are used to provide a sensor model for tractably estimating the joint probability of the grid cell occupancy states *without* making any of the restrictive independence assumptions stated earlier.

Given the collection of S sequential observations $\boldsymbol{J}_S$, we would like to estimate the occupancy of each grid cell. Using Bayes' rule, the update rule for the marginal probability of grid cell c_r

being occupied (or empty), i.e. $b_r = \beta \in \{0, 1\}$, given the entire set of measurements $\boldsymbol{J}_S$, is given as

$$\begin{aligned} p_{b|\boldsymbol{J}}(b_r = 1|\boldsymbol{J}_S) &= \frac{p_{b,\boldsymbol{J}}(b_r = 1, \boldsymbol{J}_S)}{p_{\boldsymbol{J}}(\boldsymbol{J}_S)} = \frac{p_{b,\boldsymbol{J}}(b_r = 1, \boldsymbol{j}_S, \boldsymbol{J}_{S-1})}{p_{\boldsymbol{J}}(\boldsymbol{j}_S, \boldsymbol{J}_{S-1})} \\ &= \frac{\left[\sum_{\boldsymbol{b}\in\mathbb{B}(r,1)} p_{\boldsymbol{b},\boldsymbol{J}}(\boldsymbol{b}, \boldsymbol{j}_S, \boldsymbol{J}_{S-1})\right]}{p_{\boldsymbol{J}}(\boldsymbol{j}_S, \boldsymbol{J}_{S-1})} \\ &= \frac{\left[\sum_{\boldsymbol{b}\in\mathbb{B}(r,1)} p_{\boldsymbol{j}|\boldsymbol{b}}(\boldsymbol{j}_S|\boldsymbol{b}, \boldsymbol{J}_{S-1}) p_{\boldsymbol{b}|\boldsymbol{J}}(\boldsymbol{b}|\boldsymbol{J}_{S-1})\right]}{p_{\boldsymbol{j}|\boldsymbol{J}}(\boldsymbol{j}_S|\boldsymbol{J}_{S-1})} \\ &= \eta \sum_{\boldsymbol{b}\in\mathbb{B}(r,1)} p_{\boldsymbol{j}|\boldsymbol{b}}(\boldsymbol{j}_S|\boldsymbol{b}) p_{\boldsymbol{b}|\boldsymbol{J}}(\boldsymbol{b}|\boldsymbol{J}_{S-1}), \end{aligned} \tag{4.1}$$

where η is a normalization term such that $p_{b|\boldsymbol{J}}(b_r = 1|\boldsymbol{J}_S) + p_{b|\boldsymbol{J}}(b_r = 0|\boldsymbol{J}_S) = 1$. The second to last line comes from the conditional independence of measurements. The second term on the right hand side of the last line, $p_{\boldsymbol{b}|\boldsymbol{J}}(\boldsymbol{b}|\boldsymbol{J}_{S-1})$, is the posterior probability from the previous time step $S - 1$. Likewise, the posterior probability $p_{\boldsymbol{b}|\boldsymbol{J}}(\boldsymbol{b}|\boldsymbol{J}_S)$ can be written using the same sensor model used in (4.1):

$$p_{\boldsymbol{b}|\boldsymbol{J}}(\boldsymbol{b}|\boldsymbol{J}_S) = \frac{p_{\boldsymbol{b},\boldsymbol{J}}(\boldsymbol{j}_S, \boldsymbol{b}, \boldsymbol{J}_{S-1})}{p_{\boldsymbol{J}}(\boldsymbol{j}_S, \boldsymbol{J}_{S-1})} = \mu p_{\boldsymbol{j}|\boldsymbol{b}}(\boldsymbol{j}_S|\boldsymbol{b}) p_{\boldsymbol{b}|\boldsymbol{J}}(\boldsymbol{b}|\boldsymbol{J}_{S-1}), \tag{4.2}$$

where μ is a normalization term such that $\sum_{\boldsymbol{\beta}\in\mathbb{B}} p_{\boldsymbol{b}|\boldsymbol{J}}(\boldsymbol{b} = \boldsymbol{\beta}|\boldsymbol{J}_S) = 1$. For any value of S, plugging (4.2) into (4.1) and expanding (4.2) shows the explicit dependence on all historical measurement information for the occupancy state update.

Given the map or occupancy state $\boldsymbol{b}$, the measurements $\{j_{s,k}\}_{k=1}^{K}$ in $\boldsymbol{j}_s$ are indeed conditionally independent. Thus, we can write

$$p_{\boldsymbol{j}|\boldsymbol{b}}(\boldsymbol{j}_s|\boldsymbol{b}) = \prod_k p_{j|\boldsymbol{b}}(j_{s,k}|\boldsymbol{b}). \tag{4.3}$$

Note that this is different than the commonly used assumption that the elements in the cellular occupancies are conditionally independent, i.e., $p_{\boldsymbol{b}|\boldsymbol{J}}(\boldsymbol{b}|\boldsymbol{J}_S) = \prod_i p_{b|\boldsymbol{J}}(b_i|\boldsymbol{J}_S)$.

To better describe the term $p_{j|\boldsymbol{b}}(j_{s,k}|\boldsymbol{b})$ in (4.3), and specifically $p_{j|\boldsymbol{b}}(j_{s,k}=0|\boldsymbol{b})$, let $\tilde{\boldsymbol{b}}=\{\tilde{b}_i\}_{i=1}^{B}$ be the collection of BAC outputs (cf. Section 4.2.1) which are the latent variables that capture the influence of grid cell c_i on each measurement $j_{s,k}$. More specifically, each $j_{s,k}$ is a Boolean function of virtual occupancies, i.e., $j_{s,k}=\sum_{i=1}^{B}\tilde{b}_i$. Thus, we can write

$$\begin{aligned} p_{j|\boldsymbol{b}}(j_{s,k}=0|\boldsymbol{b}) &= p_{\tilde{b}|b}(\tilde{b}_i=0\ \forall i|\boldsymbol{b})\ :\ \text{OR gate assumption} \\ &= \prod_i p_{\tilde{b}|b}(\tilde{b}_i=0|b_i)=\prod_i p_{ki}^{00}(1-b_i)+p_{ki}^{01}b_i, \end{aligned} \tag{4.4}$$

where $p_{\tilde{b}|b}(\tilde{b}_i=0|b_i)$ is implicitly parameterized by k, i.e., the index of the measurement. For each pair (k,i) of measurement index k and grid cell c_i, we define p_{d}^{ki} and p_{fa}^{ki} to be the probability of detection and false alarm, respectively. Then, we have

$$p_{ki}^{00}=p_{\tilde{b}|b}(\tilde{b}_i=0|b_i=0)=1-p_{\mathrm{fa}}^{ki}, \qquad p_{ki}^{01}=p_{\tilde{b}|b}(\tilde{b}_i=0|b_i=1)=1-p_{\mathrm{d}}^{ki},$$

where p_{ki}^{00} models the probability that $b_i=0$ is transmitted through the BAC and received correctly as $\tilde{b}_i=0$; while p_{ki}^{01} models the probability that $b_i=0$ is transmitted through the BAC and received incorrectly as $\tilde{b}_i=1$. The other probabilities $p_{ki}^{11}=p_{\mathrm{d}}^{ki}$ and $p_{ki}^{10}=p_{\mathrm{fa}}^{ki}$ are similarly defined.

Plugging all these into (4.1) and (4.2) gives the following closed-form expressions that can be used for computing the Bayesian updates:

$$\begin{aligned} p_{b|\boldsymbol{J}}(b_r=1|\boldsymbol{J}_S) &= \eta \sum_{\boldsymbol{b}\in\mathbb{B}(r,1)} p_{\boldsymbol{j}|\boldsymbol{b}}(\boldsymbol{j}_S|\boldsymbol{b})\,p_{\boldsymbol{b}|\boldsymbol{J}}(\boldsymbol{b}|\boldsymbol{J}_{S-1}) \\ &\propto \sum_{\boldsymbol{b}\in\mathbb{B}(r,1)} \prod_k \prod_i \Big[\big(p_{ki}^{00}(1-b_i)+p_{ki}^{01}b_i\big)(1-j_{S,k}) \qquad (4.5) \\ &\quad + \big(1-(p_{ki}^{00}(1-b_i)+p_{ki}^{01}b_i)\big)j_{S,k}\Big]\,p_{\boldsymbol{b}|\boldsymbol{J}}(\boldsymbol{b}|\boldsymbol{J}_{S-1}), \qquad (4.6) \end{aligned}$$

and,

$$p_{\boldsymbol{b}|\boldsymbol{J}}(\boldsymbol{b}|\boldsymbol{J}_S) = \mu \prod_k \prod_i \Big[\big(p_{ki}^{00}(1-b_i) + p_{ki}^{01} b_i\big)(1 - j_{S,k}) + \big(1 - (p_{ki}^{00}(1-b_i) + p_{ki}^{01} b_i)\big) j_{S,k} \Big] p_{\boldsymbol{b}|\boldsymbol{J}}(\boldsymbol{b}|\boldsymbol{J}_{S-1}). \tag{4.7}$$

This is what we refer to as the general formulation (GF) of the occupancy grid problem. The formulation in [4], which assumes independence between occupancy states of grid cells, can be viewed as a special case of this formulation. To see this let us consider the case when the set $\mathbb{B}(r,1)$ contains only a single element, indicating that only one grid cell is being processed at a time. The sum over $\mathbb{B}(r,1)$ and the product over i in (4.5) vanish. The product over k would also be restricted to the indices $k \in \boldsymbol{\kappa} = \{\kappa, \kappa+1, \cdots, \kappa+K'\}$, with $\kappa \geq 0$ and $\kappa + K' \leq K$, that coincide with the neighborhood around grid cell c_r. Thus, (4.5) in this case reduces to

$$p_{b|\boldsymbol{J}}(b_r = 1|\boldsymbol{J}_S) = \eta \prod_{k \in \boldsymbol{\kappa}} \Big(p_{ki}^{01}(1 - j_{S,k}) + \big(1 - p_{ki}^{01}\big) j_{S,k} \Big) p_{b|\boldsymbol{J}}(b_r = 1|\boldsymbol{J}_{S-1}), \tag{4.8}$$

which is equivalent to the formulation in [4] derived using our BAC model.

4.2.3 Choices of Transition Probabilities

The choice of appropriate transition probabilities for each BAC can be made using two different categories of approaches. The first category involves using heuristic factors where the designer imparts a belief into the choice of the transition probabilities. Such factors exploit distance-based or any other plausible measure. For example, a similar idea was adopted in the formulations presented in [7, 8] where real-world receiver antenna gain patterns were used as a way to choose the values of p_{ki}^{10} and p_{ki}^{11} for different grid cells. The second category of methods uses statistical estimation methods to learn the transition probabilities given a collection of training data. The success of these methods depends on the environments in which the senors are operating and amount of data used to estimate the transition probabilities. For example, if the environments are consistent from experiment to experiment, and we have sufficient data, then convergence of the transition probability estimates can be expected. However, if the sensor takes measurements

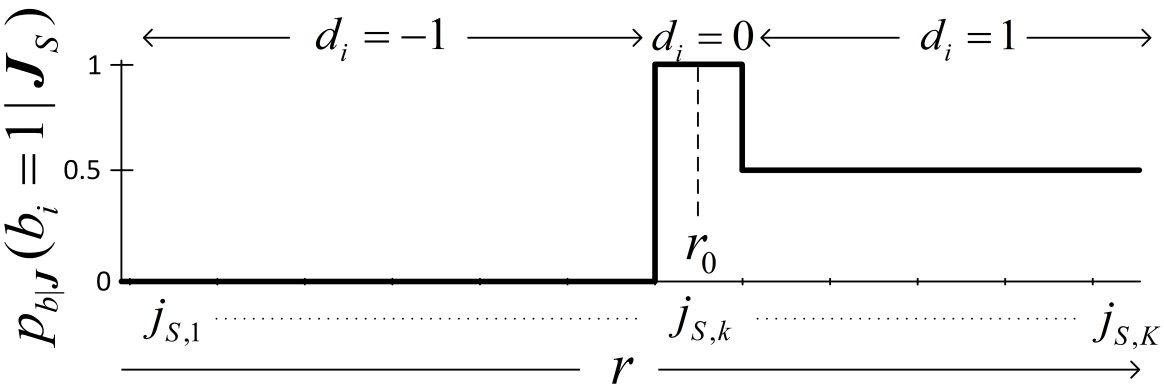


Figure 4.2: Occupancy probability profile for an ideal sensor.

in a different environment for each experiment, then the training data may not contain sufficient information for proper estimation of the transition probabilities.

We must point out that the occupancy probability profile (OPP) [4] for almost any sensor model may be represented at any time S using the proposed BAC model through proper selection of the transition probabilities. As an example, we show how to model the OPP for the ideal sensor, which returns a scalar value representing the closest distance at which a reflection from an object was seen. The ideal sensor OPP [4], illustrated in Figure 4.2, can be modeled using the BAC model through appropriate choice of transition probabilities. This OPP modeling is performed with the use of (4.8). Let $\eta' = \eta p_{b|\boldsymbol{J}}(b_r = 1|\boldsymbol{J}_{S-1})$. Then, (4.8) becomes $p_{b|\boldsymbol{J}}(b_r = 1|\boldsymbol{J}_S) = \eta' \prod_{k \in \boldsymbol{\kappa}} (p_{ki}^{01}(1 - j_{S,k}) + (1 - p_{ki}^{01}) j_{S,k})$.

Let the index k of $j_{S,k}$ be related to the range from the sensor, placing $j_{S,k} \in \boldsymbol{j}_S$ in ascending order by range from the sensor (i.e., $j_{S,k}$ further away from the sensor have a larger k). Also, let p_{ki}^{01} be a function of $j_{S,k}$ and k. Denote by r a scalar range measurement from the ideal sensor, and associate $r = r_0$ with the smallest index k such that $j_{S,k} = 1$. Figure 4.2 illustrates these details along with the OPP for the ideal sensor. The scalar range measurement can be converted to its associated vector of length K by assigning 0s at all indices of the vector, then placing a 1 at the index k encoding for the distance closest to $r = r_0$ and 0.5 for all indices afterwards.

Indicator $d_i = -1$ implies that grid cell c_i with occupancy state variable b_i is closer than $r_0 - \varepsilon$ units from the sensor, $d_i = 0$ indicate that it is $r_0 \pm \varepsilon$ units from the sensor, and $d_i = 1$ otherwise. The

ranges associated with d_i are annotated on Figure 4.2. The following set of rules can be applied to choose p_{ki}^{01} such that $p_{b|\boldsymbol{J}}(b_i = 1|\boldsymbol{J}_S)$ matches the OPP for the ideal sensor, where $|\boldsymbol{\kappa}|$ is the cardinality of the set $\boldsymbol{\kappa}$, and $\mathbb{I}(k) = \sum_{k\in\boldsymbol{\kappa}} 1 - j_{S,k}$ is the indicator function that counts the number of $j_{S,k} \in \boldsymbol{j}_S$ that are 0:

$$p_{ki}^{01} = \begin{cases} 0 & d_i = -1 \\ \frac{1-j_{S,k}}{\eta'}^{\frac{1}{\mathbb{I}(k)}} & d_i = 0 \quad \forall k \in \boldsymbol{\kappa}. \\ \frac{0.5}{\eta'}^{\frac{1}{|\boldsymbol{\kappa}|}} & d_i = 1 \end{cases}$$

4.2.4 Special Cases for Cone-like Sensor Models

Here, we present two special cases for sensors that can be modeled with a *cone-like* radiation pattern (e.g., radar, sonar, lidar, etc.). These special cases are introduced to overcome the computational issues, specifically the exponential scaling that occurs due to the cardinality of set $\mathbb{B}(r, 1)$ as the number of grid cells increases. Each special case improves computational feasibility by reducing the cardinality of $\mathbb{B}(r, 1)$ such that $|\mathbb{B}(r, 1)| \ll 2^{B-1}$, provided that there is a reduction in the amount of statistical dependence that is modeled between grid cells.

The environmental setup of the general sonar system used in our experiments is illustrated in Figure 4.3. The two special cases are illustrated in Figure 4.4a and Figure 4.4b. Similar special cases can be constructed for other types of sensors, such as a distributed network of stationary sensors. The following physical descriptors are displayed in the figure: position of the sensor and targets at time s, horizontal and vertical beamwidths θ and ψ, respectively, and sensor depression angle ϕ. Practically speaking, this information is required only for determining how to update each grid cell with each new measurement. The binary detection statistics $j_{s,k}$ are also shown, with red dashed lines indicating $j_{s,k} = 0$ and green indicating $j_{s,k} = 1$.

Each of the binary detection statistics $j_{s,k}$ is generated from one of four situations: a true positive (detection of object), a false positive (false alarm), a true negative (no object in range), or a false negative (missed detection). Figure 4.3 shows these different situations in the context of a sonar system. The red- and green-dashed arcs depict the locations at which the detector has

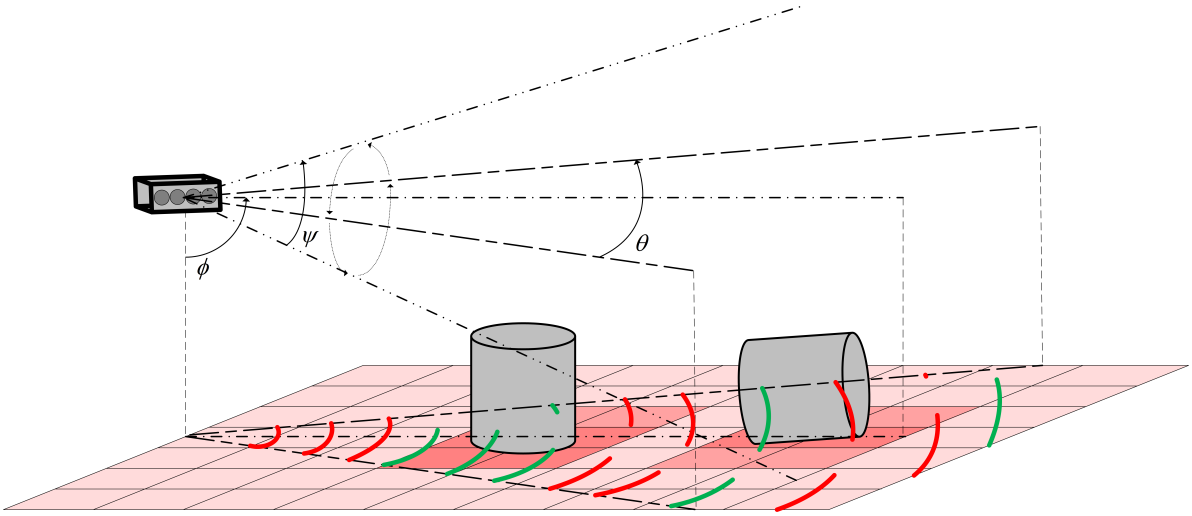


Figure 4.3: Sonar system and environmental setup for occupancy grid estimation at time s. Uniform Linear Array (ULA) with four hydrophone elements is shown with two targets sitting on the seafloor. The sensor horizontal θ and vertical ψ beamwidth, and sensor depression angle ϕ are shown. Colored grid cells represent the occupancy grid annotated on the seafloor, with the intensity of the color representing the probability of occupancy. Darker colors represent higher probabilities. Red dashed lines indicate $j_{s,k} = 0$ and green indicate $j_{s,k} = 1$.

produced a zero ($j_{s,k} = 0$) or a one ($j_{s,k} = 1$), respectively. Thus, the presence of a green-dashed arc in an area that does not overlap with the cylindrical object in the sensor cone indicates a false alarm and the presence of a red-dashed arc in a region that does overlap with the target indicates a missed detection. An example of a false alarm can be seen in the leftmost green-dashed arc closest to the sensor in Figure 4.3, while an example of a missed detection can be seen in the red-dashed arc intersecting the rightmost target in Figure 4.3.

The first special case, which is referred to as Cone Only (CO), assumes only cone-wise interaction between the grid cells within the same observation cone at a particular time s and the measurement vector $\boldsymbol{j}_s$. The second case, which is called Range Gate Only (RGO), makes the assumption that there is only local interaction between grid cells and measurements, restricted to within a particular range gate. Note that we refer to a *range gate* as that comprising multiple range intervals where each *range interval* is a constant distance from the sensor. A graphical interpretation of these special cases is illustrated in Figure 4.4, which helps to provide some physical intuition of each special case.

Case 1: Cone Only

Here, we assume that the interaction between the measurement vector $\boldsymbol{j}_s$ and the grid cells that lie outside the observation cone at time s is negligible, and hence both the probability of detection p_{ki}^{11} and false alarm p_{ki}^{10} are 0 for grid cells c_i outside of the observation cone.

To formally describe this, let us partition the grid indices $\{1, 2, \cdots, B\}$ into two parts, $\mathcal{I}$ and $\mathcal{O}$, such that $\mathcal{I} \bigcup \mathcal{O} = \{1, 2, \cdots, B\}$ and $\mathcal{I} \bigcap \mathcal{O} = \varnothing$. The set $\mathcal{I}$ captures the indices of grid cells that fall within the sensor cone, while the set $\mathcal{O}$ captures the indices of the grid cells that fall outside of the sensor cone.

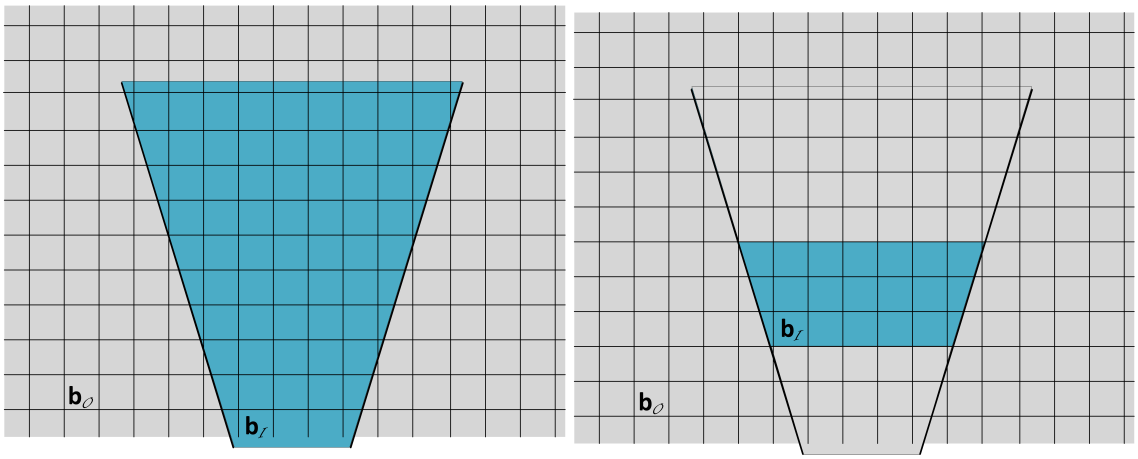
The indicator variables composing a particular vector $\boldsymbol{b}$ of cellular occupancies can then be partitioned into two disjoint sets, the set $\boldsymbol{b}_{\mathcal{I}} = \{b_i | i \in \mathcal{I}\}$ for grid cells *inside* the sensor cone and the set $\boldsymbol{b}_{\mathcal{O}} = \{b_o | o \in \mathcal{O}\}$ for grid cells *outside* such that $\boldsymbol{b} = \boldsymbol{b}_{\mathcal{I}} \bigcup \boldsymbol{b}_{\mathcal{O}}$. Evaluating (4.4) for this case gives

$$\begin{aligned} p_{j|\boldsymbol{b}}(j_{s,k} = 0|\boldsymbol{b}) &= \prod_o p_{\tilde{b}|b}(\tilde{b}_o = 0|b_o) \prod_i p_{\tilde{b}|b}(\tilde{b}_i = 0|b_i) \\ &= (1) \prod_i p_{\tilde{b}|b}(\tilde{b}_i = 0|b_i) = p_{j|\boldsymbol{b}}(j_{s,k} = 0|\boldsymbol{b}_{\mathcal{I}}), \end{aligned} \tag{4.9}$$

because the conditional probability of receiving a 0 through the BAC for each $\tilde{b}_o$ is one. By making this assumption, the cardinality of $\mathbb{B}(r, 1)$ is reduced to $|\mathbb{B}(r, 1)| = 2^{B-1-|\mathcal{I}|} \ll 2^{B-1}$. Although this provides faster computational time while maintaining the statistical dependence between all grid cells within the sensor cone, it has the side effect of ignoring information imparted by neighboring occupied cells if, for example, an obstacle occupies the inside and outside of the sensor cone.

Case 2: Range Gate Only

This special case accounts for only the interaction between measurements and grid cells within a specific *range gate* in the observation cone. A visualization of a range gate can be seen in the blue band of grid cells in the observation cone in Figure 4.4b. Similar to CO, this case also ignores



(a) CO: Grid cell occupancy outside of the sensor cone does not influence measurements inside of the sensor cone.

(b) RGO: Grid cell occupancy outside of the range gate in blue does not influence measurements inside of the range gate.

Figure 4.4: Physical interpretation of CO and RGO cell-measurement interactions. The set $\boldsymbol{b}_{\mathcal{I}}$ is depicted in blue, while the set $\boldsymbol{b}_{\mathcal{O}}$ is depicted in grey for both cases.

information imparted by neighboring grid cells if an obstacle occupies the inside and outside of a particular range gate. A potential way to help mitigate this is to allow for overlapping range gates.

In a similar manner as was done for CO, let $\boldsymbol{b} = \boldsymbol{b}_{\mathcal{I}} \bigcup \boldsymbol{b}_{\mathcal{O}}$, but for this special case $\boldsymbol{b}_{\mathcal{I}}$ is the set of grid cells within a range gate of an observation cone as in shown in Figure 4.4b. Although (4.9) can still be used for this case, a restriction is put on $\boldsymbol{j}_s$ such that the delay at which the kth measurement is taken must coincide with the selected ranges, i.e., only the $j_{s,k}$ falling within the range gate are considered. In this way, we calculate the posterior distributions of grid cell occupancies in each range gate separately. This further reduces the cardinality of $\mathbb{B}(r,1)$ when compared to the first special case.

Remarks:

The implementation of the proposed methods is expected to use the most general form that is tractable for the problem. For a small number of total grid cells, GF can be applied without much overhead. As the total number of grid cells increases, the cardinality of $\mathbb{B}(r,1)$ increases and thus so does the computation time for the marginalization in (4.1). For implementation purposes, the choice between GF, CO, and RGO should be driven by the desire to keep the cardinality of $\mathbb{B}(r,1)$ at a reasonable level that allows for a tractable amount of computation time for the system.

Consider segmenting the observation cone into many smaller disjoint cones and then using RGO to update the cellular posterior probabilities. It is not hard to imagine that in the limit of splitting the observation cone into smaller cones and applying RGO to small range gates, the updating of grid cells would happen independently from one another. In other words, there would only be a single grid cell being updated at any particular time.

These simplifications produce the same Bayesian update rule, presented in (4.8), that one would obtain by assuming independence between the occupancy state of grid cells, and independence between measurement errors. In this case, the computational complexity would no longer scale exponentially in B, at the expense of possible degradation in the performance.

4.3 Conclusion

The formulation of the general occupancy grid problem was presented in this chapter, where the solution for the joint distribution of occupancy state variables is proposed. A method for estimating the joint distribution is brought to fruition through the use of BAC networks. Multiple special cases of the proposed formulation are given, to provide different physical modeling of real-world systems, and to allow for different computational complexities. Although this framework provides a general solution to the occupancy grid estimation problem, the computational complexity is restrictive for applications with a large number of grid cells without the implementation of the special cases. This formulation provides a framework for modeling most real-world sensor configuration through the use of the BAC network sensor model and clever choice of BAC transition probabilities, a significant benefit over occupancy grid estimation techniques which need to be tuned for individual sensor models. The true benefits of this formulation over the formulations presented in Chapter 2 will be exhibited through a series of experiments in Chapter 7.

Chapter 5

Sequential Classification State Tracking

5.1 Introduction

The ultimate goal of the active perception problem, as laid out in Chapter 1 is to perform optimal path planning for simultaneous detection and classification of underwater targets. The idea of using a sequential detection state estimation technique that produced a set of distributions, such as the occupancy grid methods from Chapters 2 and 4, lends itself nicely to the information gain metric proposed to optimal path planning, briefly mentioned in Chapter 1 and to be discussed in detail in Chapter 6. We wanted to develop a sequential classification analogue to the occupancy grid estimation, specifically to use the classification distributions in the formulation of the information gain path planning metric along side distributions generated from the detection estimation process.

In this chapter, we are concerned with the sequential class state estimation portion of our problem, seen in the lower right image of Figure 1.1, and we present a new sequential classification method that allows tracking of the class state for each grid cell, formulated with the need of an information-theoretic classification cost function in mind. We desire a method that produces a set of distributions as its output, and the sequential updating of these distributions to be performed quickly, without the need to iterate an algorithm to convergence. This is required to aide in efficient calculation of the next sensing location from which to collect measurements. As such, we chose the Dirichlet-Categorical model (DCM) [25, 33] for representing the class state variable and its associated distribution random variable. The DCM provides an efficient method for combining sequential class labels that are produced as part of the navigation and interactive sensing process.

The idea behind this sequential updating process is to take a measurement from the sensor at time s and use it to update the class membership probabilities for the grid cell. The measurement is

first converted into a crude estimate of the class label l_s, and that label is merged with all previous labels to update the posterior predictive class distribution for the grid cell.

The formulation of our Dirichlet-Categorical classifier is presented in Section 5.2, and details the process of collecting a measurement, producing a crude class label estimate, merging that estimate with all previous measurements and estimates, and updating the distributions tracking the class membership for each grid cell.

5.2 Problem Formulation

Let c be the class state variable for grid cell g_i. Recall the arrangement of grid cells and measurements from Figure 2.1. At each time step s, a *one-step* classifier, shown in the second box from the left in Figure 5.2, is employed to assign a class label $l_s \in [1, L]$ to the most recent measurement from grid cell g_i. The one-step classifier in this system can be any commonly used classifier such as support vector machine (SVM), relevance vector machine (RVM) [34] or a deep neural network (DNN) [35]. The collection of sequential class labels l_s for grid cell g_i up to the current sensing time S, are formed into a set $\mathcal{L} = \Big[l_1, \cdots, l_S\Big]$. Now, the goal here is to generate the posterior predictive distribution of the class state variable c, $p_{c|\mathcal{L}}(c|\mathcal{L})$, given all the past and present labels in $\mathcal{L}$. The process of collecting measurements and updating the posterior predictive distribution of the class state variable is illustrated in Figure 5.1, where the yellow AUV collects three different measurements and updates the distribution over the available classes after each measurement. After collecting the third measurement, the probability of the class being a red star jumps dramatically.

To begin, we model c as a Categorical random variable [25] taking on L possible, non-orderable values. A Categorical random variable is used as each class label is arbitrary and their order has no significance. A random variable c is Categorically distributed if $p_{c|\boldsymbol{\lambda}}(c = l|\boldsymbol{\lambda}) = \lambda_l = P(c = l)$ for $l = 1, \cdots, L$, where $\boldsymbol{\lambda} = \Big[\lambda_1, \cdots, \lambda_L\Big]$ and $\sum_{l=1}^{L} \lambda_l = 1$, and can be expressed as $c|\boldsymbol{\lambda} \sim \text{Cat}(\boldsymbol{\lambda})$ [25]. The probability mass function of the Categorical distribution can be written as

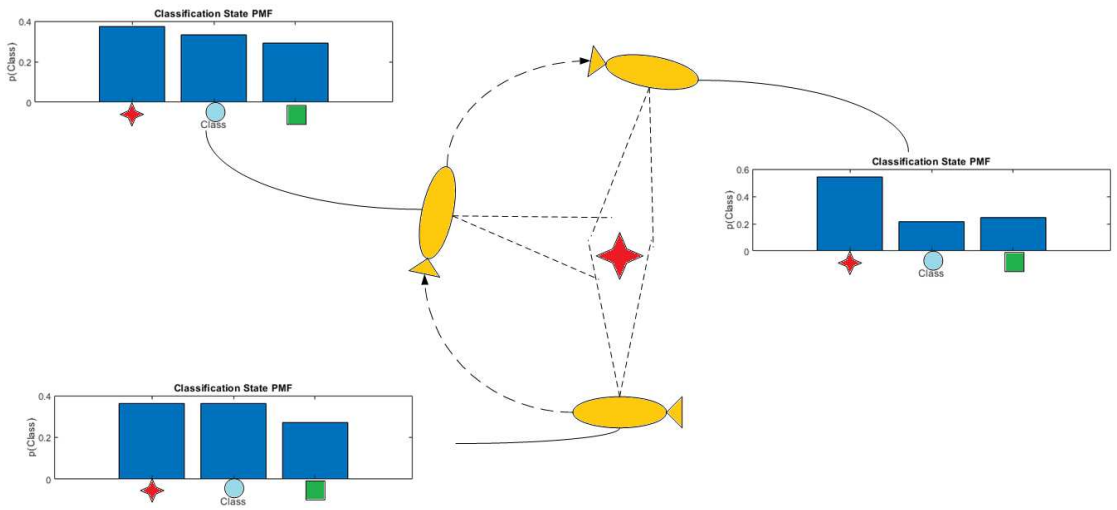


Figure 5.1: The yellow AUV traverses around a target collecting measurements. After each measurement the posterior predictive distribution tracking the classification state variable is updated. The probability of the true class, a red star, is increased after each measurement.

$$p_{c|\boldsymbol{\lambda}}(c|\boldsymbol{\lambda}) = \prod_{l=1}^{L} \lambda_l^{\delta_{cl}}, \qquad \delta_{cl} = \begin{cases} 1 & c = l \\ 0 & \text{otherwise} \end{cases}. \tag{5.1}$$

The Categorical distribution parameter $\boldsymbol{\lambda}$ is modeled as a Dirichlet distributed random variable with distribution parameter $\boldsymbol{\alpha}$, $\boldsymbol{\lambda} \sim \text{Dir}(\boldsymbol{\alpha})$. The probability density function of $\boldsymbol{\lambda}$ is defined as

$$p(\boldsymbol{\lambda}) = \frac{1}{B(\boldsymbol{\alpha})} \prod_{l=1}^{L} \lambda_l^{\alpha_l - 1}, \tag{5.2}$$

where $B(\boldsymbol{\alpha}) = \frac{\prod_{l=1}^{L} \Gamma(\alpha_l)}{\Gamma(\alpha_0)}$ is the multivariate beta function and $\alpha_0 = \sum_{l=1}^{L} \alpha_l$ [25]. The parameter vector $\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1, \cdots, \alpha_L \end{bmatrix}$ is non-random, with $\alpha_l > 0 \ \forall \ l$.

The Dirichlet distribution is the conjugate prior for the Categorical distribution, and thus the posterior distribution of $\boldsymbol{\lambda}|c$ is $\text{Dir}(\boldsymbol{\alpha}^\circ)$ where $\boldsymbol{\alpha}^\circ = \begin{bmatrix} \alpha_1^\circ, \cdots, \alpha_L^\circ \end{bmatrix}$ and $\alpha_l^\circ = \alpha_l - 1 + \delta_{cl}$ [25, 33]. That is, we can write $\boldsymbol{\lambda}|c = \boldsymbol{\lambda}^\circ \sim \text{Dir}(\boldsymbol{\alpha}^\circ)$. This shows that after getting a new class label l_s, the updated estimate of the distribution parameter $\boldsymbol{\lambda}$ is now Dirichlet distributed with parameter $\boldsymbol{\alpha}^\circ$. We call this updated distribution parameter $\boldsymbol{\lambda}^\circ \sim \text{Dir}(\boldsymbol{\alpha}^\circ)$.

The DCM provides an efficient closed-form equation for calculating the posterior predictive distribution [33] of the class state variable c given the label data in $\mathcal{L}$ using

$$p_{c|\mathcal{L}}(c|\mathcal{L}) = \int p_{c|\boldsymbol{\lambda}}(c|\boldsymbol{\lambda}) p_{\boldsymbol{\lambda}|\mathcal{L}}(\boldsymbol{\lambda}|\mathcal{L}) d\boldsymbol{\lambda} = \int \lambda_c \frac{1}{B(\boldsymbol{\alpha})} \prod_{l=1}^{L} \lambda_l^{\alpha_l - 1 + \sum_{l' \epsilon \mathcal{L}} \delta_{cl'}} d\boldsymbol{\lambda}$$
$$= \frac{B(\boldsymbol{\alpha}')}{B(\boldsymbol{\alpha})} \int \text{Dir}(\boldsymbol{\alpha}') d\boldsymbol{\lambda} = \frac{B(\boldsymbol{\alpha}')}{B(\boldsymbol{\alpha})},$$

where $\alpha_l' = \alpha_l - 1 + \sum_{l' \epsilon \mathcal{L}} \delta_{cl'}$. Thus, the posterior predictive distribution is also Categorically distributed as $c|\mathcal{L} \sim \text{Cat}(\boldsymbol{\lambda}')$ with $p_{c|\mathcal{L}}(c|\mathcal{L}) = \frac{B(\boldsymbol{\alpha}')}{B(\boldsymbol{\alpha})} = \lambda_c'$. Only one label l_s is added at each time s, thus using the recursive property of the Gamma function, $\Gamma(n+1) = n\Gamma(n)$, we see that

$$\lambda_c' = \frac{B(\boldsymbol{\alpha}')}{B(\boldsymbol{\alpha})} = \frac{\Gamma(\alpha_0)}{\prod_{l=1}^{L} \Gamma(\alpha_l)} \frac{\Gamma(\alpha_c + 1) \prod_{l=1 \neq c}^{L} \Gamma(\alpha_l)}{\Gamma(\alpha_0 + 1)} = \frac{\Gamma(\alpha_0)}{\prod_{l=1}^{L} \Gamma(\alpha_l)} \frac{\alpha_c' \prod_{l=1}^{L} \Gamma(\alpha_l)}{\alpha_0' \Gamma(\alpha_0)} = \frac{\alpha_c'}{\alpha_0'}. \tag{5.3}$$

Similar to occupancy grids, which capture the posterior marginal probabilities of occupancy for all grid cells, the posterior predictive class distribution generated by the sequential classification process is captured in classification maps (CMs). We represent a CM as a set of distributions denoted as

$$\boldsymbol{q} = \{p_{c|\mathcal{L}}(c_r|\mathcal{L})\}_{r=1}^{B}, \tag{5.4}$$

where c_r is the class state variable for the rth grid cell g_r. Figure 5.2 depicts the idea behind the proposed class state tracking process. The sequential class updating process takes each new class label l_s as well as the previous distribution parameters $\boldsymbol{\lambda}$ and $\boldsymbol{\alpha}$ to produce the new distribution parameters $\boldsymbol{\lambda}'$ and $\boldsymbol{\alpha}'$, hence allowing us to successively update the parameter estimates for every new measurement. The process depicted in Figure 5.2 is used to update the posterior predictive class distribution after each measurement as seen in Figure 5.1.

Recall that the goal is to perform optimal path planning for simultaneous detection and classification of underwater targets. Using the updated classification distributions at each time step, the

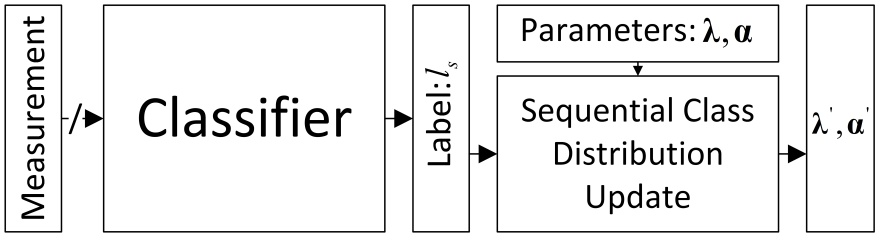


Figure 5.2: Block diagram of sequential classification state tracking. A measurement is converted into a class label l_s by means of a one-step classifier. The class label is used in conjunction with the previous distribution parameters $\boldsymbol{\lambda}$ and $\boldsymbol{\alpha}$ to produce a new distribution parameter $\boldsymbol{\lambda}'$ and $\boldsymbol{\alpha}'$.

information gain used for choosing the next sensing action can be calculated as outlined in the next chapter.

5.3 Conclusion

The sequential classification state tracking presented in this chapter exploits the relationship between the Dirichlet and Categorical distributions, and applies it to combine sequential classifications generated by a crude one-step classifier. This method allows the rigorous merging of sequential class labels to provide a distribution over all possible class labels. The DCM provides some benefits over other existing sequential classification method, such as a nice closed form that promotes fast computation. Most notably, the DCM does not require training data to learn an underlying statistical relationship between states. This distribution can be used for different estimation methods, such as maximum likelihood or maximum *a posteriori* estimation of the class label, or for estimation of the information gained from each new class label as presented in Chapter 6.

Chapter 6

Information-Theoretic-Based Optimal Path Planning

6.1 Introduction

We develop and propose a new approach to the problem of active perception using two information-theoretic cost functions based on information gain. The first cost function is associated with object detection and localization, and measures the mutual information between the occupancy state variable for a single grid cell and a binary measurement random variable. Solving for its closed-form representation relies on the measurement model and posterior occupancy distributions produced through the occupancy grid estimation process presented in [24]. The second cost function associated with the classification of detected objects measures the mutual information between a class state variable for a single grid cell and random variable that is the parameter to the class state variable distribution. In this formulation, we choose to model the class state variable as a Categorical random variable, and its distribution parameter as a Dirichlet random variable [25]. The motivation for choosing this modeling scheme comes from the need to perform sequential updating of the class state variable distribution, akin to occupancy grid estimation process. This sequential updating process has a closed-form due to the conjugacy between the Dirichlet and Categorical distributions, and allows for fast tracking of the class state distribution as new measurement are drawn. A *one-step* classification process is used in our formulation, producing class labels used to sequentially update the class state variable distribution. We call the entire process of using a one-step classifier for producing class labels and the Dirichlet-Categorical model for tracking the classification state the DCM. Similar to [19], a convex combination weighting of two normalized cost functions for sensing action selection is also utilized here.

The development of the DCM and one-step classifier provides a novel application for sequential classification and tracking of the classification state of each grid cell. The derivation of our proposed cost functions and their convex combination provides the theoretical estimates of the in-

formation gained from each sensing action, hence providing a framework for optimal navigation informed by detection and classification state estimates.

This chapter is organized as follows. Section 6.2 introduces the concept of information-theoretic cost functions. The detection and classification components of the total cost function are presented in Section 6.2.1 and 6.2.2, respectively. The total cost function that combines the detection and classification components of the cost function is presented in Section 6.2.3. Finally, Section 6.3 presents the use of the total cost function in the rollout algorithm.

6.2 Information-Theoretic Cost Functions

Here, we develop a novel information-theoretic cost function that is used in evaluating the utility of sensing actions during vehicle path planning. This cost function, called the *total information gain* at any time s $IG_T(s)$, is a convex combination of two individual information-theoretic cost functions, one for detection $IG_D(s)$ and one for classification $IG_C(s)$. We first present the detection cost function, derived with the sensor model and outputs from the occupancy grid estimation process covered in Section 4.2. The classification cost function is then presented, derived with the sequential classification process using the DCM described in Section 5.2.

Letting X and Y be two random variables, and $H(\cdot)$ the Shannon entropy, then the information gain is defined as [23]

$$I(X;Y) = H(X) - H(X|Y), \tag{6.1}$$

which can be thought of as the reduction in uncertainty of the distribution on X that the knowledge of random variable Y would bring. In the active perception problem, $H(X)$ can be thought of as the prior distribution on X, while $H(X|Y)$ can be viewed of as the posterior distribution on X after observing random variable Y.

In the context of our problem, both the detection and classification cost functions are closed-forms of the information gain, or mutual information, between a state variable X (occupancy state

or class state) and a random variable Y that capture information about the latest measurements. As such, they are both predictors of the information that is gained by taking a measurement from a grid cell given the current state of the grid cell. The detection and classification cost functions can be viewed as predicting the theoretical mutual information between the current state and the updated state after performing a sensing action, and aide in the selection of the next sensing action that provides the most information.

In the following, we will denote the set of grid cells observed at time s by $\mathcal{G}$.

6.2.1 Detection Cost Function

We define the detection information gain at time s, $IG_D(s)$, as the sum of the mutual information between the occupancy state and the measurement random variable for all observed grid cells in $\mathcal{G}$. This is stated mathematically as

$$IG_D(s) \triangleq \sum_{g_i \in \mathcal{G}} I(b_i; j_{s,k}). \tag{6.2}$$

Recall from Chapter 4 that b_i is the occupancy state variable for grid cell g_i and $j_{s,k}$ is the random variable representing a measurement at time s and range k. The information gain for grid cell g_i is given by,

$$\begin{aligned} I(b_i;j_{s,k}) &= H(b_i) - H(b_i|j_{s,k}) = -\mathbb{E}_{\mathcal{B}} \log p_b(b_i) - \mathbb{E}_{\mathcal{J}} \log p_{b|j}(b_i|j_{s,k}) \\ &= -\sum_{b \in \mathcal{B}} p_b(b_i) \log p_b(b_i) - \sum_{b \in \mathcal{B}} \sum_{j \in \mathcal{J}} p_{j|b}(j_{s,k}|b_i) p_b(b_i) \log \frac{p_{j|b}(j_{s,k}|b_i) p_b(b_i)}{p_j(j_{s,k})}. \end{aligned} \tag{6.3}$$

where $\mathcal{J} = \{0, 1\}$ is the set of possible values that realizations of $j_{s,k}$ can take, and $\mathcal{B} = \{0, 1\}$ is the set of possible values that realizations of b_i can take.

Using the occupancy grid estimation model in (4.4), the interaction between the occupancy state variable b_i and the measurement random variable $j_{s,k}$ can be represented by,

$$p_{j|b}(j_{s,k}|b_i) = [(1 - p_{\text{fa}})(1 - b_i) + (1 - p_{\text{d}})b_i](1 - j_{s,k}) + [p_{\text{fa}}(1 - b_i) + p_{\text{d}} b_i] j_{s,k}, \tag{6.4}$$

where p_{d} and p_{fa} are the probabilities of detection and false alarm, respectively, for the physical detector that produces the measurement vectors $\boldsymbol{j}_s$.

We treat the output of the occupancy grid estimation $p_{b|\boldsymbol{J}}(b_i|\boldsymbol{J}_{s-1})$, i.e., the *posterior estimate* from the previous time step, as the *prior* $p_b(b_i)$ for time s [36]. To simplify notation, we let $p_i = p_b(b_i = 1)$. Now, using this together with (6.4), and invoking the total probability, the marginal probability mass function for $j_{s,k}$ is,

$$\begin{aligned}
p_j(j_{s,k}) &= \sum_{\beta\in\mathcal{B}} p_{j|b}(j_{s,k}|b_i = \beta)p(b_i = \beta) \\
&= p_{j|b}(j_{s,k}|b_i = 1)p_i + p_{j|b}(j_{s,k}|b_i = 0)(1 - p_i) \\
&= \left[(1 - p_{\mathrm{d}})(1 - j_{s,k}) + p_{\mathrm{d}} j_{s,k}\right] p_i + \left[(1 - p_{\mathrm{fa}})(1 - j_{s,k}) + p_{\mathrm{fa}} j_{s,k}\right](1 - p_i) \\
&= \left[(1 - p_{\mathrm{fa}})(1 - p_i) + (1 - p_{\mathrm{d}})p_i\right](1 - j_{s,k}) + \left[p_{\mathrm{fa}}(1 - p_i) + p_{\mathrm{d}}p_i\right]j_{s,k}, \qquad (6.5)
\end{aligned}$$

Using this result, the prior and conditional entropy in (6.3), become

$$H(b_i) = -\left[p_i \log p_i + (1 - p_i)\log(1 - p_i)\right], \qquad (6.6)$$

and

$$\begin{aligned}
H(b_i|j_{s,k}) = -\Big[&(1 - p_{\mathrm{fa}})\big(1 - p_i\big)\log\frac{(1 - p_{\mathrm{fa}})\big(1 - p_i\big)}{(1 - p_{\mathrm{d}})p_i + (1 - p_{\mathrm{fa}})\big(1 - p_i\big)} \\
&+ p_{\mathrm{fa}}\big(1 - p_i\big)\log\frac{p_{\mathrm{fa}}\big(1 - p_i\big)}{p_{\mathrm{d}}p_i + p_{\mathrm{fa}}\big(1 - p_i\big)} \\
&+ (1 - p_{\mathrm{d}})p_i\log\frac{(1 - p_{\mathrm{d}})p_i}{(1 - p_{\mathrm{d}})p_i + (1 - p_{\mathrm{fa}})\big(1 - p_i\big)} \\
&+ p_{\mathrm{d}}p_i\log\frac{p_{\mathrm{d}}p_i}{p_{\mathrm{d}}p_i + p_{\mathrm{fa}}\big(1 - p_i\big)}\Big] \\
= (1 - p_{\mathrm{fa}})&\big(1 - p_i\big)\log\left[1 + (1 - p_{\mathrm{d}})p_i\right] + p_{\mathrm{fa}}\big(1 - p_i\big)\log\left[1 + p_{\mathrm{d}}p_i\right] \\
&+ (1 - p_{\mathrm{d}})p_i\log\left[1 + (1 - p_{\mathrm{fa}})\big(1 - p_i\big)\right] + p_{\mathrm{d}}p_i\log\left[1 + p_{\mathrm{fa}}\big(1 - p_i\big)\right], \qquad (6.7)
\end{aligned}$$

respectively.

Plugging (6.4), (6.5), (6.6), and (6.7) into (6.3) gives a closed-form expression for the detection information gain as

$$I(b_i; j_{s,k}) = p_i\Big[(1-p_{\text{d}})\log\big[1+(1-p_{\text{fa}})\big(1-p_i\big)\big]+p_{\text{d}}\log\big[1+p_{\text{fa}}\big(1-p_i\big)\big]-\log p_i\Big] + \big(1-p_i\big)\Big[(1-p_{\text{fa}})\log\big[1+(1-p_{\text{d}})p_i\big]+p_{\text{fa}}\log\big[1+p_{\text{d}}p_i\big]-\log\big[1-p_i\big]\Big]. \quad (6.8)$$

6.2.2 Classification Cost Function

In Chapter 5 we presented the class state variable c_i and the Dirichlet-Categorical model for sequential class state tracking. With that information. we define the classification information gain at time s $IG_C(s)$ as the sum of the mutual information between the class state variable, c_i, and the Dirichlet distributed parameter vector, $\boldsymbol{\lambda}$, for all observed grids $g_i \in \mathcal{G}$. This is stated mathematically as

$$IG_C(s) \triangleq \sum_{g_i \in \mathcal{G}} I(\boldsymbol{\lambda}; c_i). \quad (6.9)$$

The distribution parameter vector $\boldsymbol{\lambda}$ for the Categorical distribution on c_i in essence captures information about the latest measurements.

For a *mixed-pair* of discrete scalar random variable X and continuous random vector $\boldsymbol{Y}$, assuming they satisfy the sufficient conditions to be a *good mixed-pair* [37], their mutual information is,

$$\begin{aligned} I(\boldsymbol{Y}; X) &= h(\boldsymbol{Y}) - h(\boldsymbol{Y}|X) \\ &= -\int p_{\boldsymbol{y}}(\boldsymbol{y}) \log p_{\boldsymbol{y}}(\boldsymbol{y}) d\boldsymbol{y} + \sum_{x \in \mathcal{X}} \int p_{\boldsymbol{y},x}(\boldsymbol{y}, x) \log p_{\boldsymbol{y}|x}(\boldsymbol{y}|x) d\boldsymbol{y}, \end{aligned} \quad (6.10)$$

where $h(\cdot)$ is the differential entropy [23].

Applying (6.10) to (6.9), the mutual information $I(\boldsymbol{\lambda}; c)$ can be evaluated as

$$I(\boldsymbol{\lambda}; c_i) = h(\boldsymbol{\lambda}) - h(\boldsymbol{\lambda}|c_i) = h(\boldsymbol{\lambda}) + \sum_{c_i=1}^{L} \int_{\Delta^L} p_{\boldsymbol{\lambda},c_i}(\boldsymbol{\lambda}, c_i) \log p_{\boldsymbol{\lambda}|c_i}(\boldsymbol{\lambda}|c_i) d\boldsymbol{\lambda}. \tag{6.11}$$

The entropy of a Dirichlet distributed random vector is well-known [38] and can be written as

$$h(\boldsymbol{\lambda}) = \log B(\boldsymbol{\alpha}) + (\alpha_0 - L)\psi(\alpha_0) - \sum_{l=1}^{L} (\alpha_l - 1)\psi(\alpha_l), \tag{6.12}$$

where $\psi(x) = \frac{d}{dx} \log \Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}$ is the digamma function.

To evaluate the conditional entropy term $h(\boldsymbol{\lambda}|c)$ in (6.11), we use the fact that the Dirichlet distribution is the conjugate prior of the Categorical distribution. Thus, we can write

$$\begin{aligned}
h(\boldsymbol{\lambda}|c) &= -\sum_{c=1}^{L} \int_{\Delta^L} p_{\boldsymbol{\lambda},c}(\boldsymbol{\lambda}, c) \log p_{\boldsymbol{\lambda}|c}(\boldsymbol{\lambda}|c) d\boldsymbol{\lambda} \\
&= -\sum_{c=1}^{L} \int_{\Delta^L} p_{c|\boldsymbol{\lambda}}(c|\boldsymbol{\lambda}) p_{\boldsymbol{\lambda}}(\boldsymbol{\lambda}) \log p_{\boldsymbol{\lambda}|c}(\boldsymbol{\lambda}|c) d\boldsymbol{\lambda} \\
&= -\sum_{c=1}^{L} \frac{B(\boldsymbol{\alpha}')}{B(\boldsymbol{\alpha})} \int_{\Delta^L} \frac{1}{B(\boldsymbol{\alpha}')} \prod_{l=1}^{L} \lambda_l^{\alpha'-1} \log \frac{1}{B(\boldsymbol{\alpha}')} \prod_{l=1}^{L} \lambda_l^{\alpha'-1} d\boldsymbol{\lambda} \\
&= -\sum_{c=1}^{L} \frac{\alpha'_c}{\alpha'_0} E_{\boldsymbol{\lambda}'} \left[\log \frac{1}{B(\boldsymbol{\alpha}')} \prod_{l=1}^{L} \lambda_l^{\alpha'-1} \right] = \sum_{c=1}^{L} \frac{\alpha'_c}{\alpha'_0} h(\boldsymbol{\lambda}') \\
&= \sum_{c=1}^{L} \frac{\alpha'_c}{\alpha'_0} \left[\log B(\boldsymbol{\alpha}') + (\alpha'_0 - L)\psi(\alpha'_0) - \sum_{l=1}^{L} (\alpha'_l - 1)\psi(\alpha'_l) \right]
\end{aligned} \tag{6.13}$$

Combining (6.9), (6.12), and (6.13) provides the information gain for classification as

$$\begin{aligned}
IG_C(s) &= \sum_{g_i \in \mathcal{G}} I(\boldsymbol{\lambda}; c_i) \\
&= \sum_{g_i \in \mathcal{G}} \Big[\log B(\boldsymbol{\alpha}) + (\alpha_0 - L)\psi(\alpha_0) - \sum_{l=1}^{L} (\alpha_l - 1)\psi(\alpha_l) \\
&\qquad - \sum_{c_i=1}^{L} \frac{\alpha'_c}{\alpha'_0} \Big(\log B(\boldsymbol{\alpha}') + (\alpha'_0 - L)\psi(\alpha'_0) - \sum_{l=1}^{L} (\alpha'_l - 1)\psi(\alpha'_l) \Big) \Big].
\end{aligned} \tag{6.14}$$

6.2.3 Total Information Gain

As previously stated, the total information gain is defined as the convex combination of the detection and classification information metrics. To ensure that one information metric does not

dominate the other at all times s due to scaling, we normalize them by their respective maximal values,

$$IG_T(s) = w_D \frac{IG_D(s)}{IG_{D_{\max}}} + w_C \frac{IG_C(s)}{IG_{C_{\max}}}, \qquad w_D + w_C = 1. \tag{6.15}$$

The values for $IG_{D_{\max}}$ and $IG_{C_{\max}}$ are calculated as the sum of the maximum information gain available from a grid cell over all grid cells $g_i \in \mathcal{G}$ for the detection and classification states, respectively. The information gain for an individual grid cell is maximized when the prior distribution is uniform and the posterior distribution is a delta, implying that the maximum value is equal to the entropy of a uniform random variable. Hence, $IG_{D_{\max}} = |\mathcal{G}| \log(2)$ and $IG_{C_{\max}} = |\mathcal{G}| \log(L)$ where $|\mathcal{G}|$ is the cardinality of the set $\mathcal{G}$. The values for $IG_{D_{\max}}$ and $IG_{C_{\max}}$ are considered constant across all times s, as the distribution of the state variables could be uniform at any time s (whether or not they have been fluctuating at all times $< s$) and then completely determined at time $s + 1$ thus achieving the maximal information gain.

This convex weighting allows for different strategies to be employed by the system. For example, at the beginning of a sortie, there will likely be insufficient information to consider classification for choosing sensing actions. In this case, we can assign a higher weight (e.g., $w_D = 0.9$) for the detection while choosing a lower weight (e.g., $w_C = 1 - w_D = 0.1$) for the classification. In contrast, once most of the grid cells are observed, there may be little to no information left in performing target detection and localization, in which case we can choose the sensing actions primarily based on the classification criterion by choosing $w_C = 0.9$ and $w_D = 1 - w_C = 0.1$.

6.3 Optimal Trajectory Planning

In this section, we discuss three types of trajectory-planning policies for performing interactive sensing. Recall that the action at each time step is the selection of the next location for the sensor platform. For all policies, let $\boldsymbol{a}$ denote an action, and $\mathcal{A}_s$ be the set of all feasible actions the sensor can take under dynamical constraints at time step s. The three policies choose sensing locations

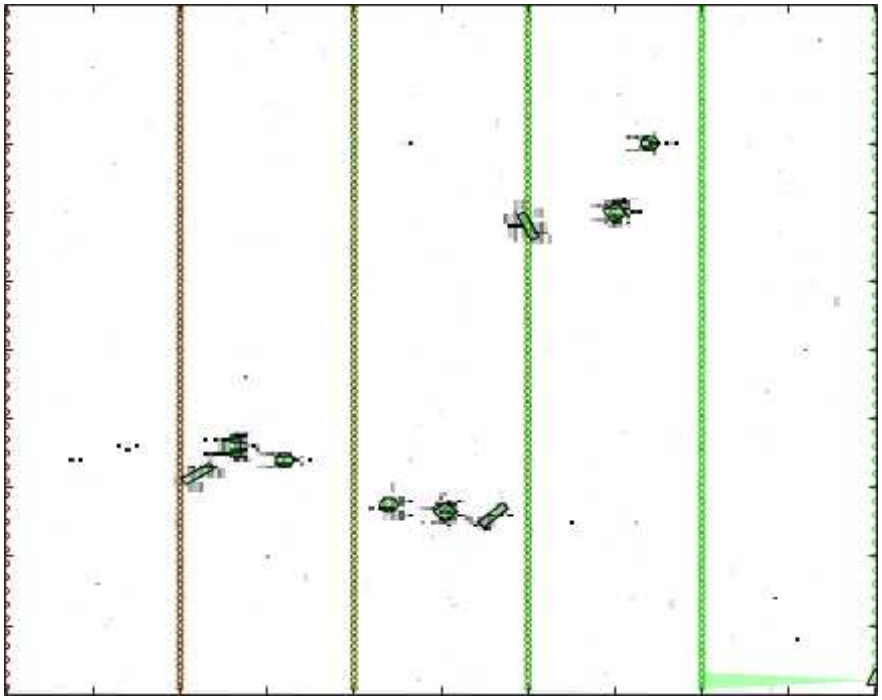


Figure 6.1: Lawn mower path used in traditional search sorties.

according to: a pre-determined lawn mower path; and maximizing a cost function for choosing $\boldsymbol{a} \in \mathcal{A}_s$ in a greedy, and non-greedy manner.

In many traditional underwater target detection and classification operations, a predetermined lawn mower path, as shown in Figure 6.1, is used. In the figure, the AUV travels in a north-south direction, facing either east or west. The AUVs path starts from the dark red circles and continues towards the light green circles. Green objects on the map outlined by black indicate the position and orientation of example targets that exist within the search area.

The sequence of prespecified actions $\boldsymbol{a}$ of the lawnmower path guarantees that each region in the search area is observed though at the cost of potentially poor detection and classification accuracy due to inadequate viewing angles.

The greedy policy for action selection chooses the action as the one that maximizes the one-step *reward*, or one-step navigation cost function, i.e.,

$$\boldsymbol{a}^*_{s+1} = \text{argmax}_{\boldsymbol{a}\in\mathcal{A}_s} R_s(\boldsymbol{b}_s, \boldsymbol{c}_s, \boldsymbol{a}), \tag{6.16}$$

where $R_s(\boldsymbol{b}_s, \boldsymbol{c}_s, \boldsymbol{a}) = IG_T(s)$ is the immediate reward for performing action $\boldsymbol{a}$ with the current state variable distributions at time s, where $\boldsymbol{b}_s$ and $\boldsymbol{c}_s$ are the occupancy state and the class state of the system at time s, respectively.

The non-greedy policy for action selection chooses the next sensing action that maximizes the one-step reward, while also considering the reward from future actions along a finite horizon of T future actions. This policy, in essence, chooses the next sensing action that it believes will generate the largest cumulative reward given the current information available. This type of problem is typically cast as a partially observable Markov decision problem (POMDP), which admits a number of solution methods [39–41]. In this document, the *rollout policy* method [40] is used to solve the problem given our choice of a heuristic navigation cost function $IG_T(s)$.

Mathematically speaking, this policy can be described as choosing the sensing action that maximizes the one-step reward plus the *expected reward-to-go* associated with a prespecified policy, typically a heuristic rule. The decision rule for action selection used by the rollout policy is defined as

$$\boldsymbol{a}^*_{s+1} = \text{argmax}_{\boldsymbol{a}\in\mathcal{A}_s} R_s(\boldsymbol{b}_s, \boldsymbol{c}_s, \boldsymbol{a}) + E_{s+1}, \tag{6.17}$$

where $R_s(\boldsymbol{b}_s, \boldsymbol{c}_s, \boldsymbol{a})$, $\boldsymbol{b}_s$, and $\boldsymbol{c}_s$ are the same as the greedy case, $E_{s+1} = \mathbb{E}\left[\sum_{i=s+1}^{T} R_i(\boldsymbol{b}_i, \boldsymbol{c}_i, \boldsymbol{a}_i) | \boldsymbol{b}_s, \boldsymbol{c}_s\right]$ is the expected reward-to-go, $R_i(\boldsymbol{b}_i, \boldsymbol{c}_i, \boldsymbol{a}_i) = IG_T(i)$ is the reward for performing action $\boldsymbol{a}_i$ with the current state variable distributions at time i, where $\boldsymbol{b}_i$ and $\boldsymbol{c}_i$ are the occupancy state and the class state of the system at time i, respectively, and T is the length of the finite horizon.

6.4 Conclusion

This chapter contains the methods by which decisions are chosen for the autonomous navigation component of our active perception problem. The mutual information is a convenient cost function that is able to exploit the outputs of the detection and classification components of the

system, specifically the sets of distributions that are generated through the estimation process. Intuitively, the proposed total cost function allows for a system to make decisions about how to choose actions given its confidence in what has been previously observed and what information it expects to get out of future observations.

The rollout algorithm provides a convenient method for choosing the next sensing action, by using the mutual information as the cost function. The choice of rollout provides a simple method to extend the path planning cost function in the event that additional components, such as movement and time costs, are to be considered in action selection.

The total total cost function presented here will be used with the rollout algorithm in Chapter 7 to illustrate how the sequential detection and classification state estimation methods, presented in Chapter 4 and Chapter 5, respectively, can be used together for optimal path planning.

Chapter 7

Experimental Results

7.1 Introduction

In this chapter, we present a collection of experimental results on optimal path planning for simultaneous detection and classification of underwater targets. The experimental results are split into two distinct segments. The first segment presents the detection abilities of the occupancy grid framework developed in Section 2.2. We compare the classical occupancy grid frameworks presented in Chapter 2 to our proposed framework presented in Chapter 4. The second segment presents experiments utilizing the combined sequential detection and classification components of Chapter 4 and Chapter 5, respectively, coupled with the developed total cost function proposed in Chapter 6, that examine the path planning abilities of the system to optimally find and classify underwater targets.

All of the experiments, with the exception of those conducted on a toy problem in Section 7.3.2, emulate a sonar system used to search littoral regions for underwater mine-like and other targets, such as unexploded ordinances. The experiments use synthetic sonar data generated by the Personal Computer Shallow Water Acoustic Toolset (PC SWAT) simulation tool [29] as described in Chapter 3. The sonar is a side-looking sonar (SLS) system that directs acoustic radiation to the starboard side of the AUV. This system is equipped with multiple hydrophones arranged in a uniformly spaced linear array (ULA), all pointing slightly downwards from horizontal (positive depression angle). The ULA has a relatively wide beamwidth in both the horizontal and vertical directions, and an interrogation range up to tens of meters. The sensor array is attached to an AUV, which is some distance above the field that is under interrogation. Further discussion of the sensor's physical configuration will be presented in the sections detailing the individual experiments.

For the detection components of the system, the sensor data generated by PC SWAT for the ULA elements at ping s is preprocessed with an adaptive coherence estimator (ACE) detector

[22, 31, 32] to produce a single *beamformed* measurement vector. This beamformed vector is thresholded at a predetermined value to produce a desired p_{fa} and probability of detection p_{d}. The thresholded detection statistics (see Figure 7.2a) are then used to produce single binary-valued measurement vectors $\boldsymbol{j}_s$ (cf. Section 4.2.1). The sensor data is not preprocessed for use in the classification components of the system.

For the experiments concerned with comparing occupancy grid frameworks, we include results for our general formulation (GF) defined in Section 4.2 and the two special cases for exploiting a cone-like radiation pattern as presented in Section 4.2.4. Our proposed formulations are benchmarked against the classical occupancy grid framework presented in Section 2.3, modified to accept vector-valued measurements, as well as a special case of GF to be described in Section 7.3. These experiments show that our proposed framework outperforms the competing algorithms in terms of producing occupancy grids that represent the ground truth.

The optimal path planning for detection and classification experiments utilize our proposed occupancy grid framework of Chapter 4 and the sequential classification framework of Chapter 5 in conjunction with the cost function described in Section 6.2.3 and one of many different policies for choosing the next sensing action described in Section 6.3 . Our proposed cost function provides the theoretical estimate of information gained from a new measurement while using our proposed sequential detection and classification frameworks. We benchmark our theoretically derived cost function to a data driven process, namely the Gaussian process regression (GPR) to be discussed in Section 7.4, for estimating the information gained from a new measurement. These experiments show that our proposed theoretical solution outperforms the data driven approach for estimating the information gained from a new measurement by producing more realistic maps, both for detection and classification purposes, in a short number of sensing actions.

This chapter is organized as follows. The metrics used for evaluating the detection and classification performance are presented in Section 7.2. These metrics provide a mechanism for quantifying the difference between probability distributions, as the outputs of both the sequential detection and classification state estimation processes are collections of random variables. Section

7.3 presents the experiments comparing occupancy grid formulations. The experiments on optimal path planning for detection and classification of underwater targets are presented in Section 7.4. Conclusions and observations on the results in this chapter are given in Section 7.5.

7.2 Performance Metrics

Both the sequential detection and classification state estimation processes produce a collection of probability distributions, and these distributions are arranged into maps for better interpretation by both human operator and machine alike. The generated map, either an occupancy grid or classification map, estimates the actual map or ground truth. As the maps comprise distributions for each grid cell it is natural to use similarity measures in comparisons to the ground truth.

To evaluate the performance for each of the experiments, two different metrics are used, as detailed below. For all metrics we evaluate the performance for detection (occupancy grids) and classification (classification maps) separately to better illustrate the strengths and weaknesses of each. In the following, $\boldsymbol{t}$ represents the true set of distributions, either true occupancy $\boldsymbol{b}$ or true classification $\boldsymbol{c}$, and $\boldsymbol{e}$ represents the estimated set of distributions, either occupancy grid $\boldsymbol{p}$ or classification map $\boldsymbol{q}$. Only like pairs are compared, i.e., $\boldsymbol{b}$ and $\boldsymbol{p}$ or $\boldsymbol{c}$ and $\boldsymbol{q}$. The true distributions are formed from delta functions (e.g., $p_{b|\boldsymbol{J}}(b_r|\boldsymbol{J}_S) = [1, 0]$ if a cell is occupied, and $p_{c|\mathcal{L}}(c_r|\mathcal{L}) = [0, 0, 1, 0]$ if a cell is of class 3).

1. Similarity between the true distribution $\boldsymbol{t}$ to that of the estimated distribution $\boldsymbol{e}$:

$$\rho = \frac{\langle \boldsymbol{t}, \boldsymbol{e} \rangle_F}{||\boldsymbol{t}||_F^2 ||\boldsymbol{e}||_F^2},$$

   where $||\cdot||_F$ is the Frobenius norm, and $\langle \cdot, \cdot \rangle_F$ is the Frobenius inner product. For calculating this metric, we form $\boldsymbol{t}$ and $\boldsymbol{e}$ into matrices by making each row the vectorized form of the distribution for each grid cell. Clearly, $0 \leq \rho \leq 1$, and $\rho = 1$ when $\boldsymbol{t} = \boldsymbol{e}$.

2. Sum of the Jensen-Shannon distance (SJSD) $D_{\text{JS}}(t_i||e_i)$ [42] over all $i = 1, \cdots, B$ grid cells:

$$
\begin{aligned}
\text{SJSD} &= \sum_r D_{\text{JS}}(t_i \| e_i) \\
&= \sum_i \frac{1}{2} D_{\text{KL}}(t_i \| m_i) + \frac{1}{2} D_{\text{KL}}(e_i \| m_i) \\
&= -\frac{1}{2} \sum_i \Big[\sum_{x \in \mathcal{X}} t_i(x) \log\Big(\frac{m_i(x)}{t_i(x)}\Big) + e_i(x) \log\Big(\frac{m_i(x)}{e_i(x)}\Big)\Big],
\end{aligned}
$$

where $m_i(x) = \frac{1}{2}\big(t_i(x)+e_i(x)\big)$, $t_i(x)$ and $e_i(x)$ are the distributions for grid cell g_i evaluated at point x, and $D_{\text{KL}}(\cdot,\cdot)$ is the Kullback-Leibler (KL) divergence [42]. The Jensen-Shannon distance is used in favor of the KL divergence as it is symmetric, positive, and always finite. The maximum value of SJSD is $\log(2) \times B$, with smaller values indicating that $\boldsymbol{t}$ and $\boldsymbol{e}$ are similar and $\boldsymbol{t} = \boldsymbol{e}$ when SJSD is 0.

In addition to the two metrics presented above, the path planning experiments in which an AUV searches for underwater targets also calculates the percentage of total grid cells observed during the experiment. The percentage of grid cells observed, when considered in conjunction with the above two metrics, gives a notion of the efficiency with which the search is carried out. If the results using two different methods exhibit similar performance for ρ and SJSD, then the method that observed the greater percentage of grid cells would obviously be better. It required less measurements taken from the same grid cell to extract the same amount of information, thus allowing it to observe new grid cells sooner.

7.3 Detection Results Using Occupancy Grid Estimation

In all experiments benchmarking the different occupancy grid frameworks we assume a stationary map and an array of sensors taking observations of the map. The sensor array produces measurements that are thresholded detection statistics in the form of binary-valued vectors. The detector operates with a known probability of detection p_{d} and probability of false alarm p_{fa}.

Two types of experiments were conducted with an increasing number of grid cells per experiment. The first type of experiment is considered as a toy experiment, where a small number of grid cells compose the entire map. This type of experiment is necessary for comparison between

the general formulation (GF), the cone-only (CO) special case, and range-gate only (RGO) special case, as there is a limit to the number of grid cells GF can be tractably applied to. This is due to the cardinality of the set of all possible binary maps scaling exponentially with the number of grid cells. We show through the series of the experiments that GF provides the best performance, and that each special case provides performance that is related to the number of grid cells that are updated concurrently. The second type of experiment uses simulated sonar data as previously described, and emulates the detection and localization component of our active perception problem.

All experiments are designed for situations where classical occupancy grid estimation performs poorly due to small gaps between objects that tend to be missed when a beam overlaps both objects and the gap at the same time [6]. The AUV path has been chosen prior to the start of the experiment, with the location of the targets being known *a priori*. These experiments are meant to show each occupancy grid frameworks ability, or lack thereof, to produce an occupancy grid that closely resembles the ground truth.

Any real sensor can be modeled, in terms of its statistical detection performance, by choosing an appropriate set of BAC transition probabilities. In the case of the sonar that is used in the following experiments, we choose the BAC transition probabilities to model an attenuation in the returned sonar signals due to the propagation losses that the wavefront experiences as it travels through the medium, i.e., salty water.

In each of the experiments, the BAC transition probabilities p_{ki}^{00} and p_{ki}^{01} were modeled as $p_{ki}^{00} = (1 - p_{\text{d}})/(1 + \text{dist}(b_i, j_{s,k}))^{\alpha}$ and $p_{ki}^{01} = (1 - p_{\text{fa}})/(1 + \text{dist}(b_i, j_{s,k}))^{\alpha}$, where $\text{dist}(b_i, j_{s,k})$ represents the Euclidean distance between the location of grid cell c_i and that at which sample $j_{s,k}$ was taken, and $\alpha \geq 1$. This particular modeling is used to emulate degraded detection performance due to attenuation in the sonar return signal strength as a function of distance.

7.3.1 Comparison with Other Occupancy Grid Methods

We must note that a direct comparison to the traditional occupancy grid frameworks of [4, 5] is not appropriate. This is due to the fact that those methods assume that the sensor produces a single

range measurement (identifies the first peak in the return above a predetermined threshold) for each observation and fit a heuristic model for sharing these range measurements over the observed grid cells at each time step. Additionally, a direct comparison with Thrun's method [6] that does not make the statistical independence of grid cell assumption is not possible, as the method also relies on single range measurements in its formulation. Therefore, we benchmark the performance of the proposed methods against two similar methods, both of which follow the traditional occupancy grid mapping algorithm. That is, they use a Bayesian update [43] for each grid cell assuming independence between the grid cells, and independence between measurements. The first method uses our measurement model and will henceforth be referred to as the *independence method* (IM). The second method uses the conventional occupancy grid sensor model [5, 6] and is adapted to allow for binary-valued vector measurements. We call the second method the *conventional method* (CM) throughout the remained of the paper.

The method for performing the Bayesian update on a 2-dimensional map, using IM, for experiments that use a cone shaped observation model is summarized as follows:

1. Find the global position of each grid cell observed within the observation cone at time s.

2. Determine the centerline of the cone, and project the position of each observed grid cell onto the centerline.

3. Associate each measurement $j_{s,k}$ with its distance along the centerline. Identify the set of all projected grid cells where projections equal this centerline distance.

4. Update the probability of each grid cell c_r, at location (x_r, y_r), being occupied given the set of measurements on that grid cell, $\{j_{s,k}\}_{k=\kappa}^{\kappa+K'}$, using (4.8).

Similarly, the Bayesian update on a 2-dimensional map using CM follows steps 1)-3) from above, and uses the sensor model proposed in [5, 6] to perform the update. The built-in occupancy grid estimation tools available in the navigation toolbox of Matlab R2019b were used to perform the probabilistic integration for CM.

7.3.2 Experimental Results — Toy Problem

The first set of experiments involves a toy problem that is sufficiently small such that all three proposed methods can be compared. The toy environment used a 2-dimensional map comprising 16 grid cells with an equal distance of 0.5 units between the centers of individual grid cells. Each grid cell had a total of 9 different measurements sampled uniformly throughout the area it covers. The measurement vector $\boldsymbol{j}_s$ comprises 144 measurements $\boldsymbol{j}_s = [j_{s,1}, \cdots, j_{s,144}]$, 9 samples for each of the 16 cells, taken at equal distance covering the same overall area as the grid cells.

This experiment can be thought of as modeling multiple application scenarios. One scenario entails using a traditional ego-vehicle that is observing an area and estimating the locations of occupied grid cells using transmit-receive sensing e.g., sonar, radar, laser, etc. In contrast, the second scenario involves using a distributed network of stationary sensors to capture simultaneous measurements. An example is the detection of objects in a room through a distributed network of pressure sensors embedded within the floor. The latter scenario clearly shows the flexibility of our formulation for a non-traditional use of occupancy grid estimation.

In this experiment, we compare each method of computing cellular posterior probabilities. All grid cells were observed at all times s. Special case CO is implemented on a neighborhood around grid cell locations instead of an observation cone while RGO takes the same neighborhood used for CO and splits it into two disjoint sections, updating each section separately.

The data used for the toy problem was synthesized by first choosing a $\boldsymbol{\beta}$ (see Figure 7.1a) and finding the ideal $\boldsymbol{j}_{\text{ideal}}$ measurement vector by letting $j_{\text{ideal},k} = 1$ if it corresponds to an occupied grid cell, and $j_{\text{ideal},k} = 0$ otherwise. A series of 15 observations $\boldsymbol{J}_{15} = [\boldsymbol{j}_1, \cdots, \boldsymbol{j}_{15}]$ was then generated by passing each element of $\boldsymbol{j}_{\text{ideal}}$ through a BAC with $p_{\text{d}} = 0.80$ and $p_{\text{fa}} = 0.08$. These values were chosen experimentally to match the performance from the detector used in the sonar experiments. An example of a single measurement $\boldsymbol{j}_s$ is shown in Figure 7.1b. Each of the measurement vectors $\boldsymbol{j}_s$ was sampled randomly according to the probability law associated with the true occupancy of that grid cell. If the grid cell ground truth is occupied, then the samples associated with that grid cell are chosen such that they are a 1 with probability p_{d} and a 0 with probability $1 - p_{\text{d}}$. Similarly,

if the grid cell ground truth is empty, then the samples associated with that grid cell are chosen such that they are a 1 with probability p_{fa} and a 0 with probability $1 - p_{\text{fa}}$.

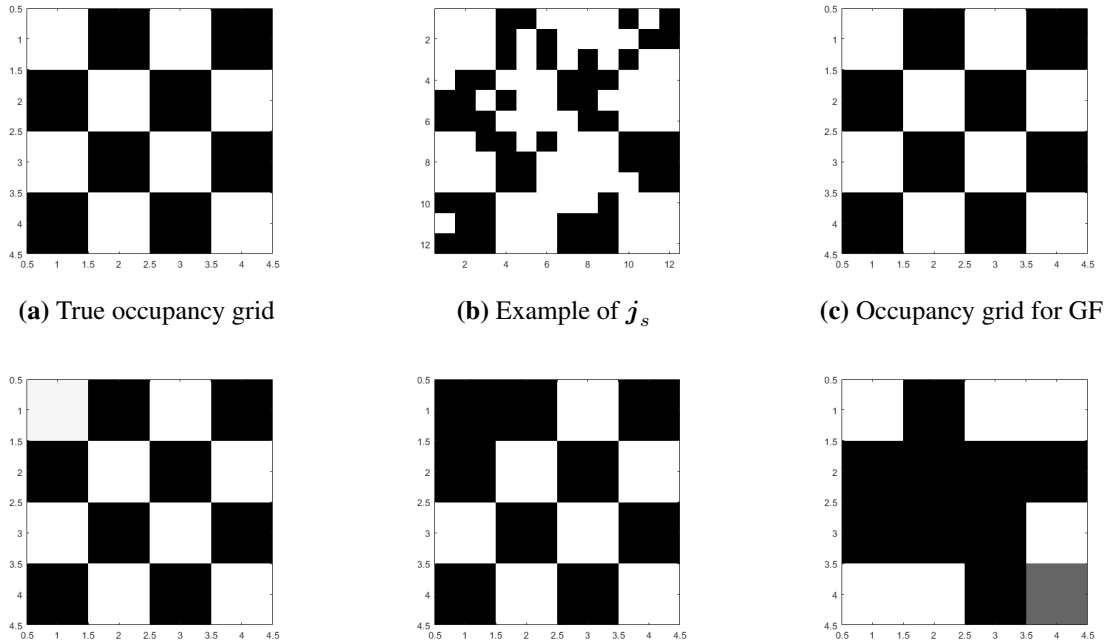
As mentioned earlier, the BAC transition probabilities p_{ki}^{00} and p_{ki}^{01} were functions of the distance between the grid cell c_i and the measurement location of $j_{s,k}$ with $\alpha = 5$. This choice of α provided the best overall results for this toy problem given that there is not a physical interpretation of distance.

Multiple experiments were conducted for each possible configuration of occupied and unoccupied grid cells (2^{16} configurations), with the average performance for each method (GF, CO, RGO, and IM) reported in Table 7.1. Results for CM are not presented, as they were essentially identical to those of IM for this problem. The resulting occupancy estimates for this toy example are presented in Figures 7.1(c)-(f) for the GF, CO, RGO, and IM methods, respectively. These figures illustrate the cellular posterior probabilities $\boldsymbol{p}$ where the probability of occupancy of a grid cell is represented by the gray level of the cell with a higher probability being associated with a darker cell.

As is evident from the visual results in Figure 7.1 as well as the performance results in Table 7.1, GF outperformed all other methods in all metrics while both CO and RGO special cases outperformed the IM method. The performance of the GF and CO are essentially identical. This is likely due to the neighborhoods of measurements used in CO being sufficiently large and including enough measurements on surrounding grid cells to accurately estimate underlying occupancy states. RGO gave up a little performance compared to GF and CO, due to the statistical dependence that it trades off in favor of reduction in computation complexity. It is interesting to note that the cardinality of $\mathbb{B}(r, 1)$ was reduced from 2^{16} to 2^8 for the CO and 2^4 for the RGO.

7.3.3 Experimental Results — Simulated Sonar Data

For the following experiments, the path of the AUV and placement of the targets are illustrated in Figure 7.2b and 7.4b. The targets are marked by a transparent green cylindrical shape, showing the size and orientation of each target. In some of the images, the black grid cells occlude the



(a) True occupancy grid **(b)** Example of $\boldsymbol{j}_s$ **(c)** Occupancy grid for GF

(d) Occupancy grid for CO **(e)** Occupancy grid for RGO **(f)** Occupancy grid for IM

Figure 7.1: Checkerboard Example - True occupancy grid and cellular posterior probabilities occupancy grids generated with all three proposed methods and the independent method.

green rectangles (e.g., in the ground truth images). The true occupancy grids in Figure 7.2b and 7.4b were generated by setting $\beta_i = 1$ for any grid cell that contained any significant part of a target, and a 0 otherwise. The path of the AUV follows the green-to-black gradient line. The last position of the AUV is marked in a black triangle, with the AUV traveling from the dark end of the line to the triangle. After the sonar returns at each ping were received, an update to the cellular posterior probabilities takes place.

Table 7.1: Comparison of cellular posterior probabilities to true occupancy grid for GF, CO, RGO, and IM. Performance (mean±std.) for the checkerboard (CB) and average (avg.) map.

	Metric	GF	CO	RGO	IM
CB	SJSD	**1.801e-3**	1.705e-2	6.939e-1	5.1252
	ρ	**1**	.9998	.94281	0.57783
Avg.	SJSD	$\mathbf{0.58 \pm 0.51}$	0.60 ± 0.51	0.70 ± 0.58	2.31 ± 1.14
	ρ	$\mathbf{0.92 \pm 0.07}$	0.92 ± 0.07	0.92 ± 0.08	0.70 ± 0.21

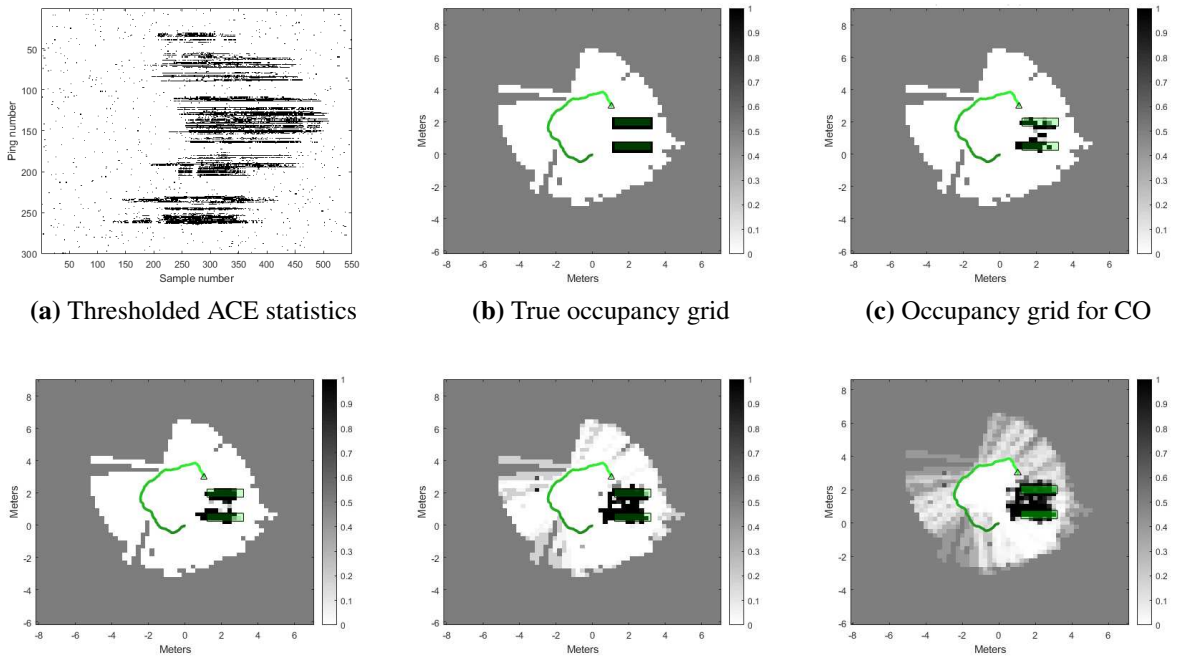
Two experiments are conducted here. The first experiment uses a short range, narrow beamwidth and coarse spatial gridding, which allows CO and RGO to be used while the second experiment uses a longer range, wider beamwidth and a finer spatial gridding, which allows only RGO to be used. The sonar data is generated with PC SWAT as described at the beginning of this chapter. Details about each experiment will be presented in their respective sections.

7.3.3.1 Experiment 1 - Short, skinny beam

In this experiment the AUV is at 10 meters above the seafloor with two cylindrical objects located 5 meters above the seafloor (i.e. in the water-column) in the sonar interrogation area. The targets were both 2 meters long along the major axis with a radius of 0.25 meters. The AUV uses a single sonar projector and an 11-hydrophone ULA with 3° horizontal beamwidth. Hydrophone elements were separated by a half-wavelength of the carrier frequency, with the geometry required to achieve the desired beamwidth being designed by PC SWAT. The transmit waveform was a linearly frequency modulated (LFM) chirp with center frequency f_c = 80 kHz, bandwidth BW = 20 kHz, and sampling frequency f_s = 60 kHz. A total of 200 pings were collected along the shown curved path, spaced at 0.01 meters apart. The output for all 200 pings forms $\boldsymbol{J}_{200} = [\boldsymbol{j}_1, \cdots, \boldsymbol{j}_{200}]$. Only CO, RGO and IM were used for this experiment owing to the very large size of $\mathbb{B}(r, 1)$ for this problem.

The choice of transition probabilities plays a significant role in the performance of each method. The BAC transition probabilities p_{ki}^{00} and p_{ki}^{01} used for CO were chosen according to the model given earlier with $\alpha = 2$. However, for RGO these probabilities were $p_{ki}^{00} = (1 - p_{\text{d}})/(1 + 0.96)^2$ and $p_{ki}^{01} = (1 - p_{\text{fa}})/(1 + 0.96)^2$, i.e., did not change as a function of distance. This constant scaling across all cells in a range gate provided better estimates for RGO.

The occupancy grids generated by the CO, RGO, and IM are illustrated in Figures 7.2c, 7.2d, and 7.2e, respectively. Six range gates were used for RGO. We can see that both CO and RGO visually perform in a similar manner, surrounding the true occupied grid cells with areas of high probability of occupancy while providing an area of low probability of occupancy between the two targets. IM and CM, on the other hand, could not resolve the two closely spaced targets



(a) Thresholded ACE statistics **(b)** True occupancy grid **(c)** Occupancy grid for CO

(d) Occupancy grid for RGO **(e)** Occupancy grid for IM **(f)** Occupancy grid for CM

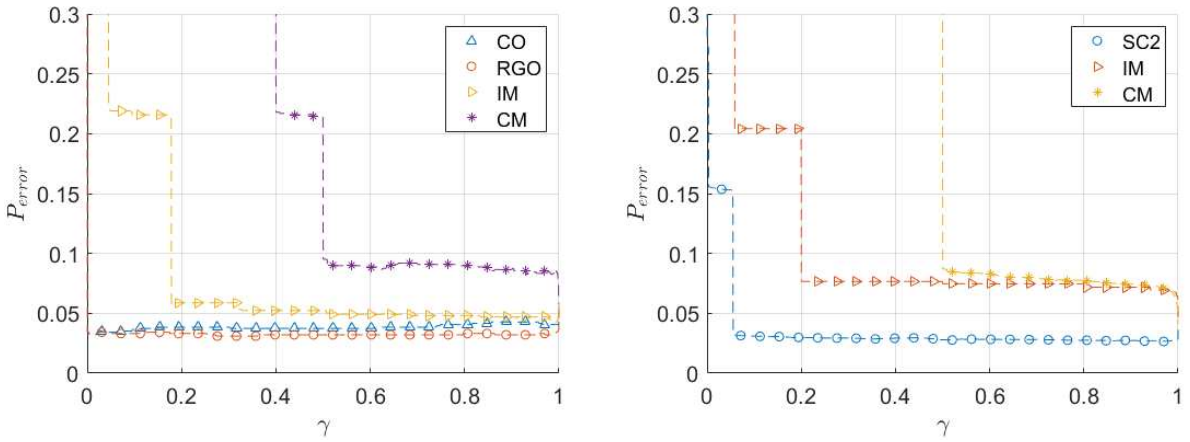
Figure 7.2: Experiment 1 - (a) The series of measurement vectors $\boldsymbol{j}_s$, $s = 1, \cdots, 200$. (b) The true occupancy grid. (c), (d), (e), (f) The occupancy grids generated by CO, RGO, IM, and CM, all with 0.25×0.25 meter grid cells. Green rectangles represent the true size, position, and rotation of the targets. Grey-scale value of pixels represent the computed posterior occupancy probability given the measurements.

and hence merged them together. When compared to CO, they had more grid cells that are in an uncertain state with their posterior occupancy probabilities remaining within a small range around 0.5, instead of being close to 0 or 1.

The performance for each of the methods was also evaluated using the SJSD and ρ measures and the results are presented in Table 7.2. As can be seen from these results, RGO slightly outperformed CO. This is, in part, attributed to a smaller number of hot spots on the upper target in Figure 7.2c, which is likely due to the larger mismatch in the collection of transition probabilities

Table 7.2: Experiment 1 - Comparison of cellular posterior probabilities occupancy grids to true occupancy for CO, RGO, IM, and CM

Metric	CO	RGO	IM	CM
SJSD	23.8306	**20.2548**	43.2365	119.99
ρ	0.9768	**0.9805**	0.9673	0.9127



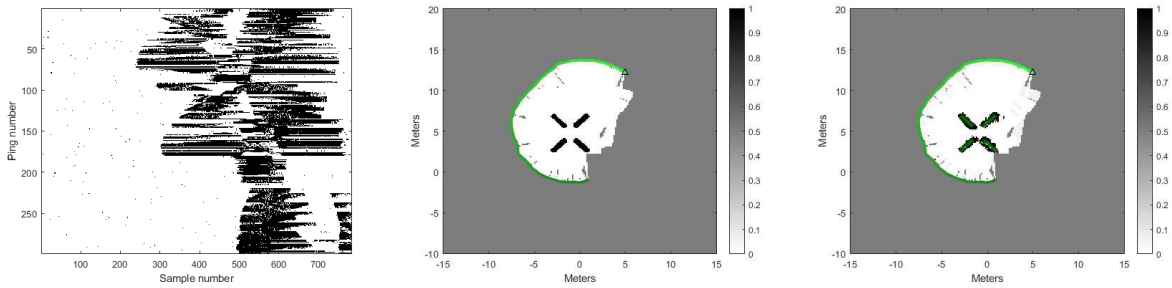
(a) Experiment 1 **(b)** Experiment 2

Figure 7.3: Probability of error as a function of threshold γ for different methods.

for CO resulting from poorly modeling the statistical dependence between grid cell occupancy states. Both IM and CM were outperformed by CO and RGO as they produced more false alarms, even though they had a greater number of hits on each target.

It is possible to generate an estimate of the underlying occupancy by applying a threshold $0 \leq \gamma \leq 1$ to the cellular posterior probabilities in order to produce a binary detection map that facilitates further analysis by human operator or autonomous navigation systems. An appropriate value for the threshold γ can be chosen, e.g., to minimize the probability of error. Figure 7.3a shows the probability of error as a function of the threshold γ for CO, RGO, IM, and CM. The results presented in Table 7.2 are echoed in Figure 7.3a, as RGO provides a lower probability of error for *all* thresholds, followed by CO, and finally IM and CM.

This result, along with the results from the previous toy experiment, suggest that using RGO provides similar performance to those of both GF and CO while outperforming the IM method. Moreover, RGO involves fewer computations than CO. Compared to IM, RGO incurs only a minor increase in computation time while providing considerably better occupancy grid estimation results. All methods outperform CM.



(a) Thresholded ACE statistics **(b)** True occupancy grid **(c)** Occupancy grid for RGO

(d) Occupancy grid for IM **(e)** Occupancy grid for CM

Figure 7.4: Experiment 2 - (a) The series of measurement vectors $\boldsymbol{j}_s$, $s = 1, \cdots, 200$. (b) The true occupancy grid. (c), (d), (e) The occupancy grids generated by RGO, IM, and CM, all with 0.2×0.2 meter grid cells. Green rectangles represent the true size, position, and rotation of the targets. Grey-scale value of pixels represent the computed posterior occupancy probability given the measurements.

7.3.3.2 Experiment 2 - Long, wider beam

In many ways, this experiment is similar to Experiment 1 in Section 7.3.3.1, with some exceptions discussed here. The AUV is 10 meters above the seafloor. Four cylindrical objects are partially buried and/or proud on the seafloor. The horizontal beamwidth is $10°$. A total of 300 pings were collected along a curved path, spaced at 0.1 meters apart between pings. The cardinality of $\mathbb{B}(r, 1)$ is greater than 2^{100} for an observation cone. Thus, this precludes the use of CO, as the computation time for a single ping becomes unreasonable. Therefore, only RGO and IM were used for this experiment.

The occupancy grids generated by the RGO, and IM are illustrated in Figures 7.4c and 7.4d, respectively. One hundred and thirty two overlapping range gates were used for RGO. The BAC transition probabilities used for RGO are the same as those in Experiment 1.

Table 7.3: Experiment 2 - Comparison of cellular posterior probabilities occupancy grids to true occupancy grid for RGO, IM, and CM

Metric	RGO	IM	CM
SJSD	**120.22**	367.01	1019.94
ρ	**0.9916**	0.9754	0.9384

From Figure 7.4c, it is seen that RGO clearly identified the four separate targets by correctly separating them while generating hot spots over the majority of each target. It does this without producing too many false alarms. The IM also identifies the four targets, but produced many false alarms that join the four targets together. It produced the correct target shapes, but offset the hot spots slightly when compared to the actual target locations. Unlike RGO and IM methods, the results of CM method shown in Figure 7.4e are unacceptable as it greatly overestimated the cellular posterior probabilities surrounding the targets.

The values for SJSD and ρ are recorded in Table 7.3, where RGO can be seen to outperform IM. The value for ρ is similar between the two methods, which is to be expected as both of the occupancy grids are similar to the truth.

As with Experiment 1, the probability of error as a function of the threshold γ was computed and illustrated in Figure 7.3b. As can be seen, the probability of error for RGO is lower than that of IM and CM for every choice of the threshold. This indicates that the addition of some complexity to the joint distribution model in the form of inter-cell statistical dependence, albeit small in the case of RGO, gives considerable improvements over IM.

7.4 Detection, Classification, and Path Planning Results using OG-DCM and a Total Information Gain Cost Function

In this section, we present the simulation results from autonomous navigation experiments utilizing the action selection policies described in Section 6.3 which use the total information gain as the reward function. The experiments conducted in this section expose each action selection policy's ability to simultaneously perform detection, localization, and classification of targets, while

exploring new areas. For the remainder of this section, we call our proposed method for calculating $IG_T(s)$ the occupancy grid Dirichlet Categorical model (OG-DCM).

The purpose of these experiments is to show the benefit of using an adaptive path planning regime to improve the overall performance of the system in terms of detection and classification of underwater targets. The historical method for path planning is allowing the AUV to traverse in a predetermined lawn mower path, seen as the green trace in Figure 7.5a. This path allows the AUV to survey as much area as possible in the least amount of sensing actions, but typically the detection and classification performance suffers as the AUV cannot view a target from the most informative angles, and does not get the option to view an object more than once. Allowing for an adaptive path, using our proposed total information gain cost function, lets the AUV to choose the next sensing action in a flexible manner. If more information is available from a recently observed region of the map, the AUV can choose to turn and re-examine that region. Conversely, if there is no more information available, then the AUV can choose to traverse to a new region.

We provide a set of experiments showing the path planning for detection and classification performance using three different methods for choosing the next sensing action. The first is a fully deterministic choice of sensing actions through the use of the lawn mower path. The second is estimating the total information gain using our theoretically derived OG-DCM, and the third is using a data driven function approximation method for estimating the total information gain cost function. In particular, we use a Gaussian process regression (GPR) for the data driven function approximation approach [18], to be described further in Section 7.4.1. The two non-deterministic methods for choosing sensing actions are iterated upon by means of the rollout algorithm [40] presented in Section 6.3. These experiments compare the abilities of our theoretical estimate of the total information and the data driven estimate of the total information gain to correctly estimate which sensing locations will provide the most information for detection and classification purposes. The rollout algorithm does not impact this calculation in any way, but rather allows for a less greedy approach in the choice of subsequent sensing actions.

It must be noted that the same sequential detection and classification methodologies, the occupancy grid framework of Chapter 4 and the Dirichlet-Categorical model of Chapter 5, respectively, are used regardless of the action selection policy. That is to say, the detection and classification state estimates used in both the theoretical and data driven estimates of the total information gain is generated from the same framework. The same state estimates are also used to produce the occupancy grids and classification maps for the lawn mower path, OG-DCM and GPR.

In the experiments conducted here, the one-step classifier is the modified version of the matched subspace classifier (MSC) [44] as the one-step classifier described in Section 5.2. The modified MSC (MMSC) was used owing to the proven success in classifying underwater objects [45]. Additionally, the MMSC has many desirable properties including the ability to use any learned subspace dictionaries, and incremental updating of the dictionary matrices when operating in new measurements [46].

At each time step, a sensing action in the form of selecting and moving to the next position to collect a measurement is taken. The measurement is collected and used to update the occupancy grid and classification map. The next location from which to take a measurement was chosen according to the greedy and non-greedy policies of (6.16) and (6.17), respectively. Each reward, $R(\boldsymbol{b}_s, \boldsymbol{c}_s, \boldsymbol{a}) = IG_T(s)$ for OG-DCM and $R(\boldsymbol{b}_s, \boldsymbol{c}_s, \boldsymbol{a}) = IG_{GPR}(s)$ for GPR, was calculated by first generating the ping from the new location with PC SWAT, then performing the occupancy grid estimation and classification map estimation, and finally calculating the reward for that action. The non-greedy policy was evaluated to time step $s + T$ for different finite horizon lengths of $T = 0$ (greedy), $T = 5$ and $T = 10$ for both GPR and OG-DCM methods. That is, each decision involves considering T time steps into the future. We denote the finite horizon policies as GPR-T and OG-DCM-T, i.e., GPR-0 and OG-DCM-0 are for $T = 0$, etc. Thus, including the lawn mower policy, a total of 7 different experiments were conducted. Each experiment uses a different action selection policy under different configuration, i.e., different estimation methods for the navigation cost function.

7.4.1 Comparison with Data Driven Estimation of Total Information Gain

As shown in Chapter 6, the theoretical representation for calculating the information gained from each measurement utilizes the sequential detection and classification state estimates. As such, there is no appropriate non-deterministic benchmark method other than data driven function approximation methods. Although, many function approximation methods such as neural networks and deep learning architectures exist, the Gaussian progress regression (GPR) has been shown to be successful in approximating the mutual information for the active perception problems [18, 20]. Thus, here we benchmark OG-DCM to that of a GPR for estimating the information gain expected from both the detection and classification components, $IG_D(s)$ and $IG_C(s)$, respectively. The ground truth outputs used to train the GPR were the detection and classification information gains, calculated as the difference between the entropy of prior and posterior state distributions for detection and classification, respectively [23].

The GPR estimates mean and associated covariance of an input data vector using an appropriate kernel function and pairs of training data and their corresponding outputs. In our problem, the current position and previous detection and classification mutual information compose the input vector. The output, estimated by the GPR, is a composite vector of the detection and classification mutual information calculated for the current sensing action. We trained the Gaussian process with a Matérn kernel function [20] and utilized K-D trees [18] to find the nearest 100 neighbors for forming the covariance matrix on a per-cell basis.

The total information gain is approximated by the Gaussian process regression following the Gaussian process upper bound confidence algorithm [20, 47], and is given by

$$IG_{GPR}(s) = w_D \frac{\mu_D(\boldsymbol{x}) + \beta\sigma_D(\boldsymbol{x})}{IG_{D_{\max}}} + w_C \frac{\mu_C(\boldsymbol{x}) + \beta\sigma_C(\boldsymbol{x})}{IG_{C_{\max}}},$$

where β is the tradeoff parameter between exploration and exploitation, $\mu_.(\boldsymbol{x})$ and $\sigma_.(\boldsymbol{x})$ are the predicted mean and variance for the detection and classification components, respectively, derived from the GPR, and $IG_{._{\max}}$ is described in Section 6.2.3.

7.4.2 Experimental Results — Simulated Sonar Data

The sonar data is generated with PC SWAT as described at the beginning of this chapter. A total of 500 pings (actions) with 1 meter ping separation were simulated for each experiment. Nine targets, in three clusters of three different targets (i.e., $L = 3$), are proud on the seafloor within the 50×50 meter search field. Medium sandy bottom was used in all the experiments to simulate bottom clutter. Each cluster contains three types of targets ($L = 3$), namely a 2-meter long cylindrical target with a radius of 0.25 meters, a 1-meter^3 cubic target, and a partially hollow sphere with 1 meter radius. The spatial orientation of the three clusters can be seen in the classification map (CM) of Figure 7.5b, where the cylinders are color-coded yellow, the cubes light blue, and the spheres red. The deep blue color indicates no target.

An AUV is equipped with multiple (11) hydrophones arranged in a uniform linear array (ULA), all pointing slightly downwards from horizontal (positive depression angle). The transmitted waveform was an LFM chirp with center frequency $f_c = 80$ kHz, bandwidth $BW = 20$ kHz, and sampling frequency $f_s = 60$ kHz. The ULA has a 7° beamwidth with an interrogation range up to tens of meters. The sensor is attached to a platform, which is 10 meters above the seafloor.

The true detection and classification distributions are illustrated in Figure 7.5a and 7.5b, respectively. In Figure 7.5a, occupied grid cells are black while empty grid cells are white. In Figure 7.5b, each grid cell is color-coded according to the non-zero component of its distribution (i.e., cylindrical class grid cells are yellow, etc.).

For each type of non-deterministic policy, i.e., greedy, and non-greedy, 20 different trials were conducted where the starting locations and headings of the AUV were randomly chosen in each trial. The lawn mower policy was only executed once. Table 7.4 gives the mean values of SJSD, ρ, and the percentage of grid cells observed after 500 sensing actions. Bold values in each column of the table represent the best performance for the metric associated with that column.

As seen from these results the OG-DCM non-greedy policy with a $T = 10$ step finite horizon outperformed all other non-deterministic policies in all metrics. The lawn mower policy outperformed OG-DCM-10 in the number of observed grid cells and the SJSD for detection. These two

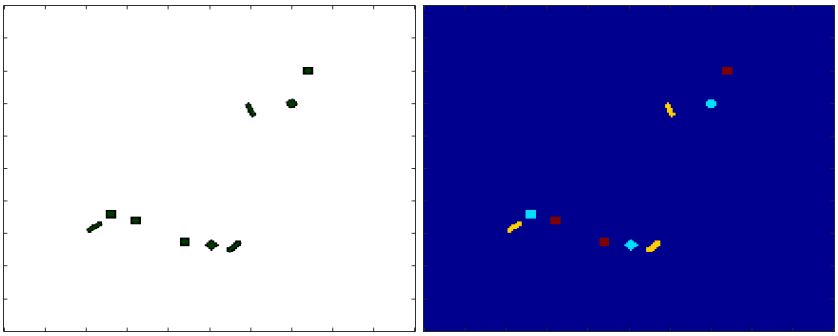
Table 7.4: SJSD and ρ for detection (Det.) and classification (Class.), and % of grid cells seen for different navigation policies after 500 sensing actions. **Bold** values indicate best performance per metric.

Policy	% Seen	SJSD Det.	SJSD Class.	ρ Det.	ρ Class.
Lawn Mower	**0.97**	**157.5**	69.2	**0.52**	0.58
GPR-0	0.10	764	66.5	0.22	0.61
GPR-5	0.45	542	68.6	0.26	0.58
GPR-10	0.41	558.1	70.83	0.15	0.57
OG-DCM-0	0.35	604.7	62.7	0.38	0.64
OG-DCM-5	0.62	442.1	57.7	0.47	0.67
OG-DCM-10	0.62	427.1	**48.7**	**0.52**	**0.75**

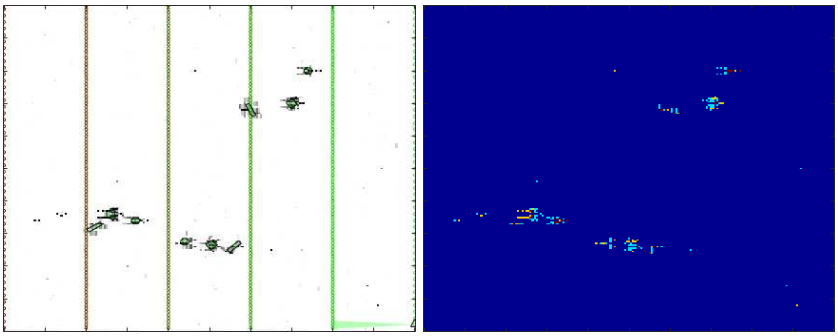
metrics can be thought of as measuring the same information, as an increase in the number of grid cells seen implies a relatively good estimate of occupancy for any reliable detector. Note that the lawn mower policy is designed to observe as many grid cells as possible, which is why it achieves an almost perfect score in the percent-seen metric, while matching the best non-deterministic policy in detection performance ρ. The accuracy of the occupancy grid, as measured by ρ for detection, is much higher for the percentage of grid cells observed when comparing the OG-DCM to the lawn mower policy, indicating that it extracts more information per grid cell it observes. In other words, because there is a tradeoff between ρ and the percentage of cells seen, showing that OG-DCM-10 outperforms the lawnmower scheme entails comparing *both*. The tradeoff clearly favors OG-DCM-10.

The OGs and CMs illustrated in Figures 7.5c-7.5h show one realization of the results for the lawn mower, and non-greedy policy experiments. The policies that use GPR wound up falling into local minima, i.e., over-observing a particular region without extracting any new detection or classification information, and thus reduced the overall efficiency of the system. Again, as can be observed the OG-DCM method generally provided much better classification results when compared with the GPR results.

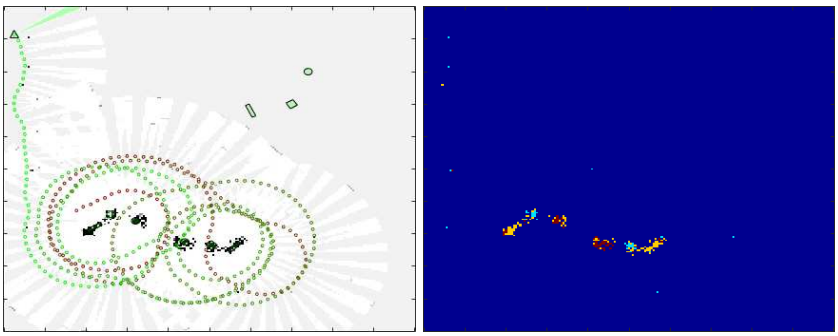
Finally, to compare the temporal evolution of these metrics during the navigation the mean value of each metric is plotted in Figures 7.6b-7.6a against the action number for each case. A major take away, illustrated in Figure 7.6c, 7.6d, and 7.6e, is that OG-DCM-10 not only outper-



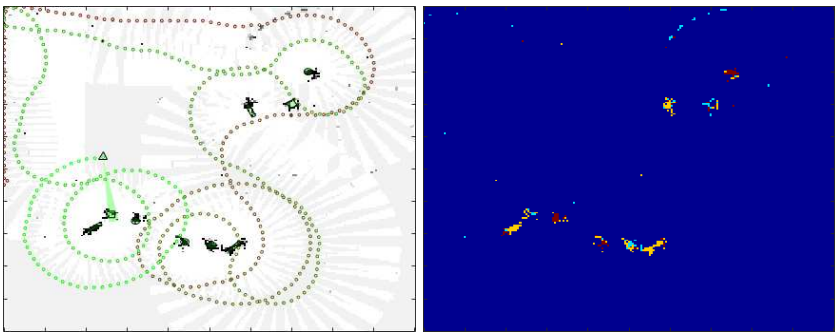
(a) Truth OG **(b)** Truth CM



(c) Lawn Mower OG **(d)** Lawn Mower CM

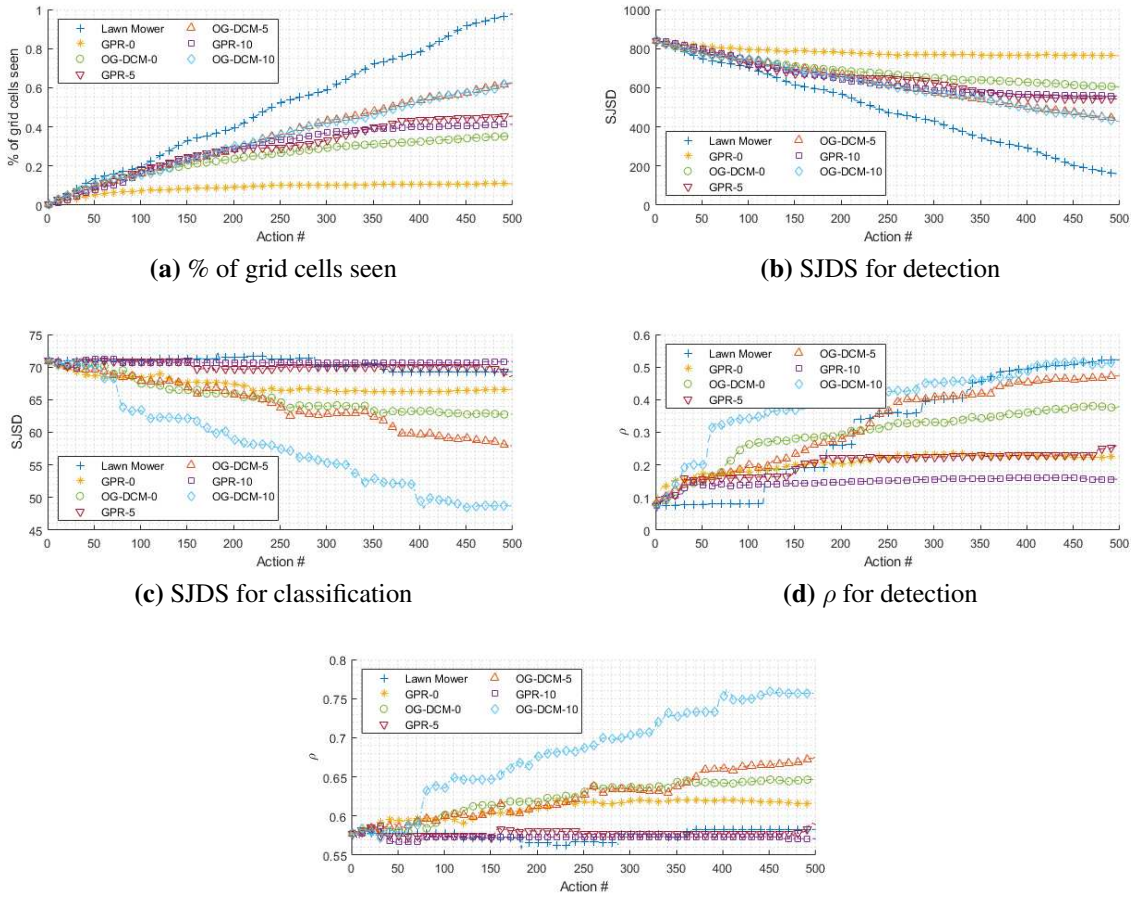


(e) GPR-5 OG **(f)** GPR-5 CM



(g) OG-DCM-5 OG **(h)** OG-DCM-5 CM

Figure 7.5: Occupancy grids (OG) on left and classification maps (CM) on right. OG and CM shown for: the underlying truth, lawn mower, GPR-5, and OG-DCM-5. Cylinders are colored yellow, cubes light blue, and spheres red. The deep blue color indicates no target. Each figure shows the same 50×50 meter area.



(a) % of grid cells seen

(b) SJDS for detection

(c) SJDS for classification

(d) ρ for detection

(e) ρ for classification

Figure 7.6: Performance plots for each of the navigation policies used in experiments.

forms the other policies, but does so in a relatively small number of sensing actions. From these results, one can conclude that as more sensing actions are taken, and more grid cells are observed, the OG-DCM-10 outperforms all other policies in all metrics.

7.5 Conclusion

This chapter presented experimental results on both the proposed occupancy grid framework alone, and the collective system comprising the proposed occupancy grid framework, Dirichlet-Categorical model and one-step classifier for sequential classification. The theoretical estimate of the total information gain cost function, and the application of all of these components with the sensing action selection using the rollout algorithm were also presented.

The experimental results on just the occupancy grid framework draw a number of conclusions. The most notable comes when comparing the classical formulation, adapted to accept binary (thresholded detection statistics in our problem) vector-valued measurements. This adapted classical formulation is a special case of our general formulation, specifically for when the occupancy state of each grid cell is statistically independent from one another. We can see from the results that this adapted classical formulation provided the worst occupancy estimates, and that adding even a small amount of statistical dependence between the occupancy states of grid cells, such as with the RGO and CO special cases, provided a much better estimate of the true occupancy. Another take away from these experiments is the flexibility that our BAC network sensor model provides. The toy experiment results, mimicking a distributed sensor network, along with the simulated sonar results showed that our occupancy grid formulation and associated sensor model can be used to model many different sensor types with success. Lastly, the computational complexity of the proposed framework must be mentioned. Most classical occupancy grid frameworks scale linearly in complexity with the number of grid cells observed at any one time. Ours scales exponentially. For any real detection environment, this exponential scaling becomes a potential issue. One major saving factor about our occupancy grid framework is that it can be implemented on a massively parallel scale, perfectly applicable for implementation on a graphical processing unit (GPU).

The results from the detection, classification, and optimal path planning showed that the combined system performs well at both exploring new areas and extracting detection and classification information about observed targets. The system was designed with biomimicry in mind, and attempts to accomplish the same decision making steps that a knowledgeable being might make in the event that they are uncertain about their next action. The notion of information gained from an observation, or series of observations, is thus a very natural way to consider choosing what to do next. The information gain cost function has shown success in many path planning applications, but can impose restrictions on the ability to find a closed form solution. The knowledge of the probability law driving a process is required, and furthermore the ability to calculate its entropy is

required. This is why many, if not most, methods for calculating an information gain cost function resort to a data driven function approximation method, such as GPR.

Our choice to model the detection state by means of an occupancy grid framework and the classification state by means of our proposed Dirichlet-Categorical model with one-step classifier provides a novel method to calculate the total information gain cost function in closed form, with no need for data driven approximations. The experimental results showed that our modeling choice allows for an excellent theoretical estimation of the information gain, providing sensing action choices that provide better detection and classification estimates compared to a deterministic method and a data driven approach. Future work should include a study on the effectiveness of different schemes for selecting the cost function component weights, w_D and w_C. Some type of adaptive method for choosing these weights would almost certainly provide better results, and could include choosing the weights based on the information gain from the previous step, i.e., choosing to weight detection more heavily if information gain from the detection component was higher in the previous time step than that of the classification component.

Chapter 8

Conclusions and Future Work

8.1 Conclusions

This dissertation principally discusses a novel framework to solving the active perception problem, in which actions are taken by an autonomous system to gain as much information about the surrounding environment as possible. The particular problem we address relates to identifying underwater targets through the use of sonar. Our approach is inspired from the biomimicry of wildlife, such as bats or dolphins when searching for and tracking prey. We consider the act of choosing how to observe the environment as the defining factor of our problem.

Like bats or dolphins, or many intelligent systems for that matter, a single observation of the surrounding environment can provide substantial information about that environment, but considering multiple sequential observations simultaneously can provide exemplary information. As a simple, albeit contrived, example consider a dolphin looking to have some fun. The dolphin is continuously sending out sound pressure clicks and examining any reflected clicks in an attempt to both detect and classify what objects may exist. Given a continuous reflection from an object near the surface, the dolphin must be able to identify if the return is originated from a boat or an orca whale. One can imagine that the reflections from the bow of a boat and the nose of an orca could be similar from some viewing angles, and thus the dolphin would need to observe a reflection from another part of the object before determining what the object truly is. This need guides the path of the dolphin, at least until it can collect enough information to determine if it is play time or, more importantly, time to escape.

Using our approach, we choose to segment the active perception problem in 4 parts:

(1) Generating a map of the scene through an occupancy grid estimation process, which produces a set of marginal posterior probabilities that any one point in an area is occupied.

(2) Classifying occupied regions through a sequential classification process, which produces a set of marginal posterior probabilities of class membership for each occupied region.

(3) Computing the (i) mutual information between the occupancy state of a grid cell and a random variable modeling a measurement on that grid cell and (ii) mutual information between the class state of a grid cell and a random variable modeling the class distribution parameter.

(4) Exploiting the information gain to select the sensing action that produces the best opportunity to detect, localize, and classify an object.

The motivation for this partitioning of steps, particularly the use of the proposed sequential detection and classification frameworks, is to provide a collection of probability distributions that can be used in the calculation of the total information gain cost function. The proposed total information gain cost function provides the analogue for choosing where the system is to take the next observation from.

Ultimately, there was a major choice to be made when choosing the detection and classification frameworks for our problem. The choice was between a classical theoretically driven approach and a data driven approach that tends to be more prevalent in recent years, likely due to advancements in many deep learning and neural network architectures. We chose the non-data driven approach for a couple of reasons. Primarily, we wanted our framework to be applicable across many fields without the need of an abundant amount of training data. Even in similar fields, a non-trivial amount of data is typically required to retrain the networks when moving from, e.g., environments with different target types.

Our advancements in the theory of occupancy grids, specifically the general consideration of statistical dependence between occupancy states of grid cells, were required due to the physical considerations of our sensing environment. The littoral regions considered by our problem are relatively sparse in the sense that only a small fraction of the total volume is populated with objects of interest. This leads to the intuition that the detection state of nearby grid cells should influence each other, driving the need for our innovations. The statistical dependence of grid cell occupancy

states has been discussed previously in literature, but we are the first to fully formulate the general problem, propose a sensor model that can provide a tractable solution while modeling almost any type of sensor, and at the same time allow for vector-valued measurements.

We chose the Dirichlet-Categorical model to provide a fast and efficient method for representing and updating the class state variable and its associated distribution random variable. This provides an efficient method for combining sequential class labels that are produced as part of the navigation and interactive sensing process. Although the Dirichlet-Categorical model is a well known model in the statistical communities, providing closed form posterior predictive distribution forms due to the conjugacy between the Dirichlet and Categorical distributions, our approach of including the one-step classifier combines well trained classification machines with a framework that can quickly update and track the classification state estimates of grid cells. Our proposed classification tracking framework allows for sequential noisy measurements to be combined in a rigorous manner to provide the best possible class estimates while generated the data required for optimal path planning.

The derivation of the total information gain cost function, which results in a convex combination of the cost functions for detection and classification, provides flexibility in scheduling for the entire system. By allowing for changing weights of the detection and classification components throughout the observation process, the system can be tuned to address different priorities. Our derivation provides the theoretical estimate of the information gain when using our proposed methods. Our experiments using the whole system show that any other method for choosing sensing actions provides inferior performance.

The accumulation of the proposed sequential detection framework, the proposed sequential classification framework, the derivation of the their combined total information gain cost function, and ultimately the application of these for the optimal path planning of detection and classification of underwater targets presents its self in this dissertation.

The results from the detection, classification, and optimal path planning showed that the combined system performs well at both exploring new areas and extracting detection and classification information about observed targets.

When observing a target, the AUV will continue to observe the target until it has extracted most of the information available for detection and classification. In our experiments, this behavior was manifested through the path that was optimally chosen using our total information gain cost function, typically an elliptical path around the object. For non-symmetric targets there are optimal viewing angles that provide greater amounts of information for both detection and classification. The elliptical path around the target exploits the ability to gather observations at these optimal viewing angles, and thus the elliptical path is intuitive for our AUV.

When not observing a target, the AUV chooses a path that provides the minimal amount of overlap between observed regions. In other words, the AUV is always trying to explore previously unobserved regions.

This type of behavior is both desirable and expected. One could imagine that a similar outcome could be produced by evaluating a list of conditional statements that were compiled by an expert. If a target is detected, then start along a 360 degree elliptical path to observe the target. If no target is detected, then search in some predetermined manner. Many questions arise about how to adapt this set of conditional statements. What if the AUV has extracted enough information to correctly classify the target after only a couple of observations and did not need the remaining portion of the elliptical path? What if a second target is identified while observing the first? Proceeding to following the remainder of the predetermined path could wind up just contributing to the waste of resources. Some type of heuristics can be incorporated to manage these situation, but would ostensibly require the list of conditional statements to grow and be maintained for different situations.

Our combined sensing system using the theoretical estimate of our total information cost function is able to produce better estimates of the truth than systems that utilize a data-driven ap-

proximation of the same total information cost function. It does so, in an adaptive manner while providing consistent and interpretable results.

8.2 Future Work

The research world is ever evolving and it is hard to say that any work is ever complete. As such, a number of future research topics related to the material presented in this dissertation immediately come to mind. These topics all originate from either work that has been partially accomplished but abandoned due to higher priorities or personal interest in exploring some other facet of mathematics. The topics include: a) multiple agent search, b) improvements to the sequential classification process, c) adaptive weighting policy for total cost function, d) extension of the sequential detection process, and e) testing on real data sets.

Some preliminary work was performed on incorporating multiple sensor platforms, or multiple AUVs in a cooperative search. The total cost function can be reformulated to include the information gained from all sensors, with the optimization process choosing the set of sensing actions for the collection of sensors that maximizes the information gain in the next iteration. The preliminary results showed that the emergent behavior of multiple sensors is to generally partition the available search area. In some experiments with two AUVs, the AUVs would simultaneously observe a target from opposing sides of the target eliminating the need for either AUV to fully observe the target. The estimates produced using multiple AUVs were slightly better when compared to the estimates from a single AUV and were generated in less than half the time.

The sequential classification process could still use some further investigation before it is declared as a complete work. The DCM provides a very convenient way to efficiently track the classification state of grid cells given a collection of sequential measurements. That said, there is nothing about DCM that inherently assumes statistical dependence of class states between between grid cells. We have looked at methods [25, 43, 48] for rigorously combining these class states to improve the classification accuracy, however, they rely on the sparse population of objects and the

ability to accurately cluster the grid cells belonging to those objects. A further evaluation of DCM can certainly incorporate the statistical dependence of class states in a natural manner.

There are two areas of personal interest that are worth exploring. The first is an adaptive scheme for choosing the weights for the detection and classification components of the total information gain cost function. If nothing else, it would be interesting to see the emergent behavior and performance of the system when utilizing different optimization schemes. Many reasonable heuristic policies for selecting the weights immediately come to mind.

One could scale the detection weight according to the percentage of total grid cells that have been observed, starting with a very high weight (w_D = 0.9) when no grid cells have been observed smoothly transitioning to a very low weight (w_D = 0.1) once all grid cells have been observed. This would place more importance on classifying targets towards the end of the search, ideally once they have all been detected.

Another could scale the detection and classification weights based on how much information was gained at the current time step for each individual cost function. By assigning a larger weight to the individual cost function that had a larger value, the system would automatically be tuned in favor of that component. For instance, if the detection cost function was very small compared to the classification cost function then it is implied that the system is confident an object exists but is not confident about its classification. The classification cost function weight would be nearly maximized, and the next action would likely be chosen in favor of re-observing the same object, thus, likely increasing the confidence in the classification estimate instead of observing a new region.

Given the collection of enough training data, a machine learning algorithm such as reinforcement learning could be used to learn a policy for selecting the weights that did not rely on heuristics.

The most interesting extension of our work would be to examine and reformulate our problem for the case when there are moving targets. This would likely require a formal redefinition of the occupancy grid problem to include conditioning on the estimated state of clusters that exist.

One approach could be the following. From each cluster in the occupancy grid a centroid and be estimated. At each time step, the clusters can be associated and an estimate of the cluster state $\hat{\boldsymbol{\theta}}$ can be calculated by, e.g., a Kalman filter or other Bayesian estimator. The posterior state estimate can be used in the calculation of the updated occupancy grid $p_{b|\boldsymbol{J},\hat{\boldsymbol{\theta}}}(b_i|\boldsymbol{J}_S, \hat{\boldsymbol{\theta}})$.

The ultimate consideration for future work is the extensive testing on real side-looking and forward-looking sonar data. A limitation that we faced during the development of our system is the lack of real data sets to test our system against. This is due to the adaptive nature of our system and the complexities of sonar data. The real data sets that were available to us during development comprise sonar data taken from sensors mounted to a moving platform with a predetermined path. There simply was no way to incorporate the adaptive path planning of our system into these experiments. As such, all of the experiments presented in this dissertation were conducted with synthetic data. This allowed us full control over the sensor and platform position and path. As realistic as the synthetic data is, there is no replacement for real-world implementation and testing. Our experiments have primarily been concerned with side-looking sonar, but could easily be used for down-looking and forward-looking sonar as well. To get real-time computational speed, an implementation would need to be done in C++ and potentially integrated with a GPU, as our algorithms are parallelizable. This implementation can be done for a relatively low hardware cost and then be utilized on a real autonomous underwater platform.

Bibliography

[1] R. Bajcsy, "Active perception," *Proc. IEEE*, vol. 76, no. 8, pp. 966–1005, 1988.

[2] H. Moravec and A. Elfes, "High resolution maps from wide angle sonar," in *Proc. 1985 IEEE Int. Conf. Robot. Autom.*, vol. 2. IEEE, 1985, pp. 116–121.

[3] H. P. Moravec, "Sensor fusion in certainty grids for mobile robots," in *Sensor Devices and Syst. Robot.* Springer, 1989, pp. 253–276.

[4] A. Elfes, "Using occupancy grids for mobile robot perception and navigation," *Comput.*, vol. 22, no. 6, pp. 46–57, 1989.

[5] A. Elfes *et al.*, "Occupancy grids: A stochastic spatial representation for active robot perception," in *Proc. 6th Conf. Uncertain. AI*, vol. 2929, 1990, p. 6.

[6] S. Thrun, "Learning occupancy grid maps with forward sensor models," *Autonomous Robots*, vol. 15, no. 2, pp. 111–127, 2003.

[7] B. Clarke, S. Worrall, G. Brooker, and E. Nebot, "Sensor modelling for radar-based occupancy mapping," in *2012 IEEE/RSJ Int. Conf. Intell. Robots Syst.* IEEE, 2012, pp. 3047–3054.

[8] A. Guerra, F. Guidi, J. Dall'Ara, and D. Dardari, "Occupancy grid mapping for personal radar applications," in *2018 IEEE Statistical Signal Process. Workshop (SSP)*. IEEE, 2018, pp. 766–770.

[9] S. Bai, J. Wang, F. Chen, and B. Englot, "Information-theoretic exploration with Bayesian optimization," in *2016 IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*. IEEE, 2016, pp. 1816–1822.

[10] S. T. O'Callaghan and F. T. Ramos, "Gaussian process occupancy maps," *Int. J. Robot. Research*, vol. 31, no. 1, pp. 42–62, 2012.

[11] M. Saval-Calvo, L. Medina-Valdés, J. Castillo-Secilla, S. Cuenca-Asensi, A. Martínez-Álvarez, and J. Villagrá, "A review of the Bayesian occupancy filter," *Sensors*, vol. 17, no. 2, p. 344, 2017.

[12] T. Gindele, S. Brechtel, J. Schroder, and R. Dillmann, "Bayesian occupancy grid filter for dynamic environments using prior map knowledge," in *2009 IEEE Intell. Veh. Symp.* IEEE, 2009, pp. 669–676.

[13] C. Coué, C. Pradalier, C. Laugier, T. Fraichard, and P. Bessière, "Bayesian occupancy filtering for multitarget tracking: an automotive application," *Int. J. Robot. Research*, vol. 25, no. 1, pp. 19–30, 2006.

[14] A. Nègre, L. Rummelhard, and C. Laugier, "Hybrid sampling Bayesian occupancy filter," in *2014 IEEE Intell. Veh. Symp. Proc.* IEEE, 2014, pp. 1307–1312.

[15] D. L. Applegate, R. E. Bixby, V. Chvatal, and W. J. Cook, *The traveling salesman problem: a computational study*. Princeton University Press, 2006.

[16] E. K. P. Chong and P. Ramadge, "Optimal load sharing in soft real-time systems using likelihood ratios," *J. Optim. Theory Appl.*, vol. 82, no. 1, pp. 23–48, 1994.

[17] P. Whaite and F. P. Ferrie, "Autonomous exploration: Driven by uncertainty," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 3, pp. 193–205, 1997.

[18] G. A. Hollinger, B. Englot, F. S. Hover, U. Mitra, and G. S. Sukhatme, "Active planning for underwater inspection and the benefit of adaptivity," *Int. J. Robot. Research*, vol. 32, no. 1, pp. 3–18, 2013.

[19] F. Bourgault, A. A. Makarenko, S. B. Williams, B. Grocholsky, and H. F. Durrant-Whyte, "Information based adaptive robotic exploration," in *IEEE/RSJ Int. Conf. Intell. Robots Syst.*, vol. 1. IEEE, 2002, pp. 540–545.

[20] S. Bai, J. Wang, K. Doherty, and B. Englot, "Inference-enabled information-theoretic exploration of continuous action spaces," in *Robot. Research*. Springer, 2018, pp. 419–433.

[21] B. J. Julian, S. Karaman, and D. Rus, "On mutual information-based control of range sensing robots for mapping applications," *Int. J. Robot. Research*, vol. 33, no. 10, pp. 1375–1392, 2014.

[22] L. L. Scharf and L. T. McWhorter, "Adaptive matched subspace detectors and adaptive coherence estimators," in *Conf. Record 13th Asilomar Conf. Signals, Syst. and Comput.* IEEE, 1996, pp. 1114–1117.

[23] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & Sons, 2012.

[24] C. Robbiano, E. K. P. Chong, M. R. Azimi-Sadjadi, L. L. Scharf, and A. Pezeshki, "Bayesian learning of occupancy grids," *IEEE Trans. Intell. Transport. Syst.*, pp. 1–12, 2020.

[25] K. P. Murphy, *Mach. Learn.: a probabilistic perspective*. MIT press, 2012.

[26] A. Elfes, "Robot navigation: Integrating perception, environmental constraints and task execution within a probabilistic framework," in *Int. Workshop Reasoning Uncertain. Robot.* Springer, 1995, pp. 91–130.

[27] H. Rue and L. Held, *Gaussian Markov random fields: Theory and applications*. CRC press, 2005.

[28] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robot.* MIT press, 2005.

[29] G. S. Sammelmann, "Propagation and scattering in very shallow water," in *MTS/IEEE Oceans 2001. An Ocean Odyssey. Conf. Proc. (IEEE Cat. No. 01CH37295)*, vol. 1. IEEE, 2001, pp. 337–344.

[30] R. Lim, "Data and processing tools for sonar classification of underwater UXO," *SERDP MR-2230*, 2015.

[31] S. Kraut, L. L. Scharf, and L. T. McWhorter, "Adaptive subspace detectors," *IEEE Trans. Signal Process.*, vol. 49, no. 1, pp. 1–16, 2001.

[32] S. Kraut, L. L. Scharf, and R. W. Butler, "The adaptive coherence estimator: A uniformly most-powerful-invariant adaptive detection statistic," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 427–438, 2005.

[33] S. Tu, "The dirichlet-multinomial and dirichlet-categorical models for Bayesian inference," *Comput. Sci. Division, UC Berkeley*, 2014.

[34] C. M. Bishop, *Pattern recognition and machine learning*. Springer, 2006.

[35] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learn.* MIT press, 2016.

[36] D. V. Lindley, "The philosophy of statistics," *J. Royal Statistical Soc. Series D (The Statistician)*, vol. 49, no. 3, pp. 293–337, 2000. [Online]. Available: http://www.jstor.org/stable/2681060

[37] C. Nair, B. Prabhakar, and D. Shah, "On entropy for mixtures of discrete and continuous variables," *arXiv preprint cs/0607075*, 2006, https://arxiv.org/abs/cs/0607075.

[38] N. Ebrahimi, E. S. Soofi, and S. Zhao, "Information measures of dirichlet distribution with appl." *Applied Stochastic Models in Business and Industry*, vol. 27, no. 2, pp. 131–150, 2011.

[39] D. P. Bertsekas and D. A. Castanon, "Rollout algorithms for stochastic scheduling problems," *J. Heuristics*, vol. 5, no. 1, pp. 89–108, 1999.

[40] E. K. P. Chong, C. M. Kreucher, and A. O. Hero, "Partially observable markov decision process approximations for adaptive sensing," *Discrete Event Dynamic Syst.*, vol. 19, no. 3, pp. 377–422, 2009.

[41] J. C. Goodson, B. W. Thomas, and J. W. Ohlmann, "A rollout algorithm framework for heuristic solutions to finite-horizon stochastic dynamic programs," *Eur. J. Operational Research*, vol. 258, no. 1, pp. 216–229, 2017.

[42] J. Lin, "Divergence measures based on the Shannon entropy," *IEEE Trans. Inf. Theory*, vol. 37, no. 1, pp. 145–151, 1991.

[43] R. R. Murphy, *Introduction to AI Robotics*. MIT press, 2019.

[44] A.-B. Salberg, A. Hanssen, and L. L. Scharf, "Robust multidimensional matched subspace classifiers based on weighted least-squares," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 873–880, 2007.

[45] J. J. Hall, M. R. Azimi-Sadjadi, S. G. Kargl, Y. Zhao, and K. L. Williams, "Underwater unexploded ordnance (UXO) classification using a matched subspace classifier with adaptive dictionaries," *IEEE J. Ocean. Eng.*, vol. 44, no. 3, pp. 739–752, 2019.

[46] P. Pakrooh, L. L. Scharf, and M. R. Azimi-Sadjadi, "Underwater target classification using a pose-invariant matched manifold classifier," in *2016 IEEE 26th Int. Workshop Mach. Learn. Signal Process. (MLSP)*. IEEE, 2016, pp. 1–5.

[47] N. Srinivas, A. Krause, S. Kakade, and M. Seeger, "Gaussian process optimization in the bandit setting: no regret and experimental design," in *Proc. 27th Int. Conf. Mach. Learn.*, 2010, pp. 1015–1022.

[48] D. Birant and A. Kut, "ST-DBSCAN: An algorithm for clustering spatial–temporal data," *Data & Knowl. Eng.*, vol. 60, no. 1, pp. 208–221, 2007.