# A Decision-Making Framework for Control Strategies in Probabilistic Search

Timothy H. Chung and Joel W. Burdick

*California Institute of Technology, Pasadena, CA 91125, USA*

{*timothyc, jwb*}*@caltech.edu*

*Abstract*— This paper presents the search problem formulated as a decision problem, where the searcher decides whether the target is present in the search region, and if so, where it is located. Such decision-based search tasks are relevant to many research areas, including mobile robot missions, visual search and attention, and event detection in sensor networks. The effect of control strategies in search problems on decision-making quantities, namely time-to-decision, is investigated in this work. We present a Bayesian framework in which the objective is to improve the decision, rather than the sensing, using different control policies. Furthermore, derivations of closed-form expressions governing the evolution of the belief function are also presented. As this framework enables the study and comparison of the role of control for decision-making applications, the derived theoretical results provide greater insight into the sequential processing of decisions. Numerical studies are presented to verify and demonstrate these results.

## I. INTRODUCTION

The goal in a search problem is to generate the search paths in uncertain environments that best enable the searcher to locate a target (perhaps among other objects) using one or more mobile sensor platforms, possibly under resource constraints. Inspired by the seminal works of B. Koopman [1] and L.D. Stone [2], this *physical search* theory has been extensively developed.

Beyond the search for targets using mobile sensors, however, the search problem can be generalized to represent the class of problems where choice of observations is controlled to best search for an object or outcome. For instance, the scheduling of individual sensor nodes in a wireless sensor network or the selection of the focus-of-attention in visual systems can also be formulated as search problems, where a control policy is generated to improve the information obtained by observations.

This paper presents a formal framework which casts the search problem as a decision between hypotheses about the decision-maker's current knowledge. In the specific case typically considered in the physical search theory literature, the decision reflects a belief of whether or not the target of interest is present in the search region. However, our general formulation allows for analysis of the belief evolution as more observations are taken, as well as the definition of measures of the quality of the decision, such as performance and robustness of the decision-making process.

Given its relevance to a variety of autonomous, mobile sensor applications, search theory has been studied extensively in the literature. In addition to the defining works mentioned previously, [3] discusses various components of the search problem, including classes of search paths, models

for detection, and approaches for detecting mobile targets, e.g., amidst false targets. Further works, such as [4] and [5] address issues of dynamic coverage, e.g. using (possibly multiple) mobile robots, for optimally navigating in a region to gather sensory information regarding entities such as changing environmental parameters or the location of land mines. Relevant research presented in [6] and the references therein also present a Bayesian construction for the problem of searching for lost targets, using the probability of detection of these targets as the objective function for optimal search trajectory generation. As developed in [6], the Bayesian filtering approach offers an advantage over other methods for maintaining and updating all information relevant to the search (e.g. the target probability density function), and is thus the subject of much study.

The contributions of this paper include the formulation of the search control problem as a *decision-making* problem rather than a sensing task, where measures associated with decisions, e.g. confidence and robustness of the decision or time until the decision is made, are used to design an appropriate control policy. Our formulation of the search problem as a detection problem also allows us to include practical sensor artifacts (such as false alarms and missed detections) which have not be completely considered in other formulations [6] of the search problem. Furthermore, we derive closed-form formulas for representing the current belief in many practical cases. These formulas offer both significant insights into the search problem, and enable efficient computations for implementation of various search strategies, including ones relevant to other types of search. For example, in "lookahead search" (see below), we are able to practically compute a lookahead window of more than one step, whereas prior work [6] was limited to one step due to a lack of such efficient formulas. Next, we formalize two new search strategies (which we term the "saccadic search" and "*Drosophila*-inspired search") which have not previously appeared in the robotics literature. Numerical studies compare the different searching strategies in test cases. Finally, the analytic framework presented in this paper provides a foundation for investigation of other decision-based planning tasks.

Section II provides the formulation and definition of the search problem. Section III analyzes the search problem in the context of decision-making. Analytic expressions for the evolution of the decision-making process are also derived. The problem of control for search is framed as a control for decision-making task in Section IV, where several choices

of relevant decision-based objective functions are proposed. Results and insights gained from this examination are also discussed in this section, followed by concluding remarks and avenues for additional research in Section V.

## II. SEARCH PROBLEM STATEMENT

### A. Problem Setup

Consider the initial task of searching for a single, stationary target, denoted $x_T$, in an environment, $\mathcal{A}$, which is discretized into $|\mathcal{A}|$ cells. This discretization may reflect partitions of the environment due to system requirements, such as limited range of the detector, or alternatively physical representations of the environment, e.g. rooms in a building. The formulation presented here can be extended to a non-discretized search area. Discrete approaches also lend themselves to the practical implementation of the algorithms in physical systems. Hereafter, the expression $x_T \in \mathcal{A}$ will denote the presence of the target, $T$, in region $\mathcal{A}$, while $x_T \notin \mathcal{A}$ will otherwise denote the fact that the target is not present in $\mathcal{A}$. Similarly, $x_T = a$ and $x_T \neq a$ denote the presence or absence of the target in the $a^{th}$ cell ($a \in \mathbb{Z}^+ \leq |\mathcal{A}|$), respectively.

### B. Search as a Decision Problem

The primary question at hand is to determine whether or not the target is present in the region $\mathcal{A}$, and if the target is present, in which cell it is located. The former task can be easily constructed as a decision between the two hypotheses, by defining the binary random variable $H$, such that
$$H = \begin{cases} 0, & \text{if } x_T \notin \mathcal{A}, \\ 1, & \text{if } x_T \in \mathcal{A}. \end{cases}$$
The objective is now to determine the probability that the affirmative hypothesis is true, i.e. $Pr(H = 1)$. This probability is a measure of the *aggregate belief* (ranging between 0 and 1) that the target is present somewhere in the region $\mathcal{A}$. In a similar manner, the individual cell belief probability, $Pr(x_T = a)$, represents the probability that the target is specifically located in cell $a$. This latter probability expression addresses the secondary question of identifying the location of the target within $\mathcal{A}$. It can be immediately seen (by the Law of Disjoint Probabilities) that

$$Pr(H=1) = Pr(x_T=1 \vee \cdots \vee x_T=|\mathcal{A}|) = \sum_{a=1}^{|\mathcal{A}|} Pr(x_T=a).$$

We take a decision theoretic approach because we would like to admit detection sensors models that allow for false alarms and missed detections. In that case, a decision theoretic approach is needed, as detection of a target can only be guaranteed in a probabilistic sense. While our basic formulation allows for multiple searchers, the concrete formulas and simulations of this paper focus on the case of a single searcher.

### C. Target Detection

Practically speaking, measurements are generally taken in the presence of noise. Given that these discrete measurements are drawn from a continuous interval of possible values, consider a detection binary random variable, denoted $d$, which can take values of either 0 or 1. This detection variable

represents the decision of whether or not the detection of a specific target in a given cell has been made. Note that no assumptions are made to restrict forms of the distribution from which the measurement noises are sampled.

To augment our notation for the purposes of the search problem, define $d_{a_l}^t$ to be the detection measurement at discrete time step $t$ taken in cell $a$. This subscript term, $a_l$, is included when referring to the $l^{th}$ detection measurement (allowing for multiple measurements) of the specified cell $a$, and will be omitted otherwise when there is no ambiguity.

In this manner, we can construct the following detection model for the imperfect detection measurement, given the presence or absence of the target in the cell:
$$Pr(d_a^t|x_T) : \begin{cases} Pr(d_a^t = 0|x_T = a) & = \beta, \\ Pr(d_a^t = 1|x_T = a) & = 1 - \beta, \\ Pr(d_a^t = 0|x_T \neq a) & = 1 - \alpha, \\ Pr(d_a^t = 1|x_T \neq a) & = \alpha, \end{cases}$$
where $\alpha$ and $\beta$ are the detection error probabilities for false alarms and missed detections, respectively. These error rates quantify the noise characteristics of the sensor, and for a given sensor ($\alpha$, $\beta$) can be determined experimentally or by sensor specifications. In the parlance of decision theory, the detector model is nothing more than the likelihood function of receiving a measurement given the hypothesis, and the error probabilities ($\alpha,\beta$) represent the power of the decision test for the sensor.

Before the start of the search and any measurements are taken, the prior belief that the target is in $\mathcal{A}$ is defined to be $Pr(H = 1) = \delta$, for $0 \leq \delta \leq 1$. For example, a prior belief value of $\delta = 1$ indicates that the target is certainly present in the region, and the search problem reduces to simply localizing the target, once detected, within the region. Non-unity values of $\delta$, however, allow for the possibility that the target isn't in the region to begin with. The confidence in this prior belief is reflected in the evolution of the belief as a function of observations, and is one of the advantages of the framework presented in this paper.

**The search problem.** The statement of the search task as defined above is as follows: *Given the detector model ($\alpha$, $\beta$) and the prior belief $\delta$, determine the evolution of the belief that the target is present in region $\mathcal{A}$ as a function of the observations made until time $t$, i.e. $Pr(H = 1|D^t)$, where $D^t = \{d^1, \ldots, d^t\}$.* This belief evolution ultimately governs the decision of whether (and if so, where) the target is located in $\mathcal{A}$, thereby completing the search.

## III. DISCRETE BAYESIAN FORMULATION

The use of Bayesian filters for autonomous applications is ubiquitous in the robotics community. The primary reasons for their popularity are their applicability to general (i.e. non-Gaussian) probability density functions, as well as their inherent recursive formulation, such that arbitrary probability density functions (PDF's) can be maintained and updated in a relatively efficient recursion algorithm.

### A. Recursive Computation of the Cell Belief Probabilities

Introduced in the previous section, the probability distribution of interest is the belief function, $\mathbf{B}(t)$, defined by

$$\mathbf{B}(t) \triangleq Pr(H=1|D^t) = \sum_{a=1}^{|\mathcal{A}|} Pr(x_T = a|D^t). \quad (1)$$

This expression shows that in order to compute the aggregate belief probability, given a sequence of $t$ observations, one must compute the individual cell belief probabilities. Note that $Pr(x_T)$ represents the PDF of the target location, $x_T$.

Computation of the cell belief probability is done using Bayesian filtering [7], which incorporates the propagation of the target PDF with an update step after an observation is received. The prediction component of the discrete filter utilizes the discrete analog of the Chapman-Kolmogorov equation (e.g. see [6]), which includes the convolution of the process model of the target with the target PDF of the previous time step.

For simplicity, the results presented in this paper assume a stationary target, but note that extensions to search for a dynamic target are a straight-forward application of this prediction step, as the motion of the target can be composed with the results given below to yield the desired result.

Given that we are interested in how observations affect the belief function, we focus instead on how the observation update is computed for each cell probability in Equation (1):

$$Pr(x_T = a|D^t) = \frac{Pr(d_k^t|x_T = a, D^{t-1})}{Pr(d_k^t|D^{t-1})} Pr(x_T = a|D^{t-1}),$$

where one can recognize the numerator term to be the detector model (i.e. likelihood function), and $Pr(x_T = a|D^{t-1})$ is the belief function for the previous time step, which provides the recursion in the filter. The term in the denominator, $Pr(d_k^t|D^{t-1})$, is the marginalization of the measurement, and can be computed in closed-form by

$$Pr(d_k^t|D^{t-1}) = \sum_{H=\{0,1\}} Pr(d_k^t|H, D^{t-1}) Pr(H|D^{t-1}).$$

Making use of the fact that

$$Pr(d_k^t|H=1, D^{t-1})$$
$$= \frac{\sum_{b=1}^{|\mathcal{A}|} Pr(d_k^t|x_T = b, D^{t-1}) \, Pr(x_T = b|D^{t-1})}{\sum_{b=1}^{|\mathcal{A}|} Pr(x_T = b|D^{t-1})},$$

and that

$$Pr(d_k^t|H=0, D^{t-1}) \Rightarrow Pr(d_k^t|x_T \neq b, D^{t-1}), \; \forall b,$$

we arrive after algebraic manipulations at the final expression for the update of the individual cell belief probabilities:

$$Pr(x_T = a|D^t) = \frac{Pr(d_k^t|x_T = a, D^{t-1}) \, Pr(x_T = a|D^{t-1})}{\left[\begin{array}{l} Pr(d_k^t|x_T = k, D^{t-1}) \, Pr(x_T = k|D^{t-1}) \\ + Pr(d_k^t|H=0, D^{t-1}) \left(1 - Pr(x_T = k|D^{t-1})\right) \end{array}\right]}. \quad (2)$$

This recursive expression provides a compact and efficient way to update the belief function at every time step as the searcher observes a sequence of unexplored and/or previously visited cells. The benefit of its simplicity is seen both in analytic extensions and in algorithmic implementations demonstrated in the following sections.

## B. Closed-Form Expressions for Uniform PDF's

Simplification of the update expressions can be made under the assumption that the initial belief probability density function is uniformly distributed. Given no prior knowledge of the distribution on the target's location, one can use this uniform distribution to represent maximal uncertainty. In this case where each of the initial (i.e. prior to any observations being taken) cell belief probabilities are uniform (constant), we can generate several closed-form expressions for the evolution of the belief function parameterized by the number of measurements taken. These functions provide insight into the search process, as well as form the foundation for computationally efficient algorithms.

For a sequence of null detections ($D^t = 0$) and for uniform prior probabilities, $Pr(x_T = a) = \frac{\delta}{|\mathcal{A}|}, \forall \, a$, we find

$$Pr(H=1|D^t=0) = \frac{t\beta\delta + (1-\alpha)(|\mathcal{A}| - t)\delta}{t\beta\delta + (1-\alpha)(|\mathcal{A}| - t\delta)}. \quad (3)$$

Figure 1 plots, for different values of the false alarm rate $\alpha$, the evolution of the belief function for a searcher with a constant $10\%$ missed detection rate while it searches a $10 \times 10$ array of cells. Such graphs enabled by our closed-form formulas give insight into how the belief function evolves as a function of time and sensor characteristics.
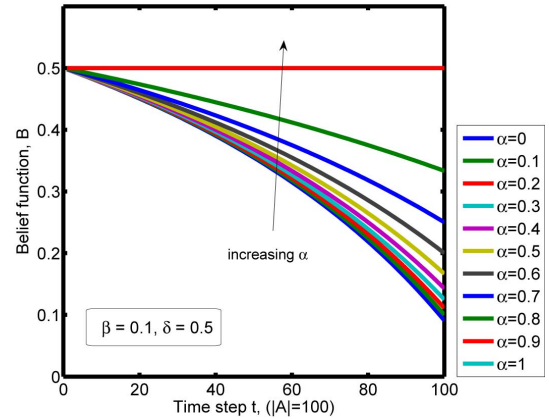


Fig. 1. Evolution of $\mathbf{B}(t)$, given a sequence of null observations, $D^t = 0$, for varying $\alpha$, constant $\beta = 0.1$, and $\delta = 0.5$

In the case of uniform priors, we can explicitly take the derivative of the above expression to determine the rate-of-change of the belief as a function of time step $t$:

$$\frac{\partial}{\partial t} Pr(H=1|D^t=0) = \frac{|\mathcal{A}|(1-\alpha)(\alpha + \beta - 1)(1-\delta)\delta}{[|\mathcal{A}|(1-\alpha) + t\delta(\alpha + \beta - 1)]^2}.$$

The belief evolution for a consecutive sequence of null measurements followed by a non-null measurement is often of interest (e.g. [7], [8]). This belief is given succinctly by

$$Pr(H=1|d^t=1, D^{t-1}=0)$$
$$= \frac{\alpha\beta(t-1)\delta + (1-\beta)(1-\alpha)\delta + \alpha(1-\alpha)(|\mathcal{A}| - t)\delta}{\alpha\beta(t-1)\delta + (1-\beta)(1-\alpha)\delta + \alpha(1-\alpha_D)(|\mathcal{A}| - t\delta)}, \quad (4)$$

where we again see the advantage of such a closed-form

expression in that we can easily compute its derivative:

$$\frac{\partial}{\partial t} Pr(H = 1 | d^t = 1, D^{t-1} = 0)$$

$$= \frac{|\mathcal{A}|\alpha^2(1-\alpha)(\alpha+\beta-1)(1-\delta)\delta}{[|\mathcal{A}|(1-\alpha)+(1-t\alpha)\delta(\alpha+\beta-1)]^2}.$$

Figure 2 illustrates the belief evolution where a non-null observation is received by the detector during the course of the search (at $t = 65$), where the searcher sequentially enters a new cell at each step. As can be seen, the perfect detector (i.e. $\alpha = \beta = 0$) can immediately decide the target is present upon arrival of a non-null observation, and thus the belief in this case jumps to unity, which ends the search task. In contrast, since false alarms are possible with an imperfect detector, only an incremental increase in the belief is registered with a positive detection.
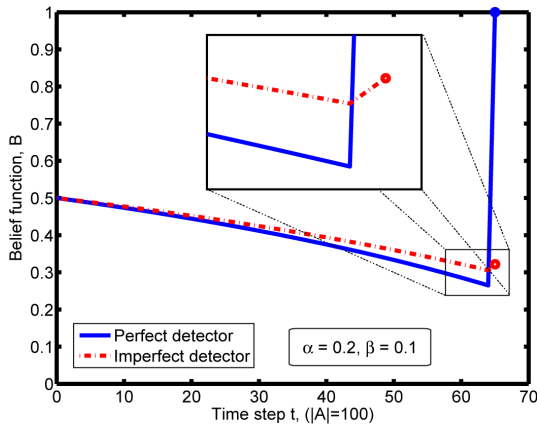


Fig. 2. Effect of receiving a non-null observation at $t = 65$ on the evolution of the belief for the perfect and imperfect detector.

The special case of uniform prior distributions is presented to simply demonstrate the structure of the Bayesian framework, but we stress the fact that the general case of non-uniform prior belief probabilities are easily and efficiently incorporated into the decision-based formulation, as will be shown in the following section.

## IV. CONTROL FOR DECISION-MAKING

Having generated a method for computing the belief evolution for a sequence of imperfect observations, we can now investigate how to appropriately choose the sequence of cells to observe, called the *search control policy*, in order to improve the decision-making performance, as defined by some metric. The task of generating this policy defines the search path control problem. In this section, we examine several intuitive and known strategies and also propose novel ones which are facilitated by the decision-making framework presented in this paper. In addition to being relevant to the target search problem studied here, these novel search strategies also offer insight into other search problems that arise in visual search and sensor networks.

### A. Search Strategies

This section presents an examination of the different search strategies. Of importance is the specification of the search termination criteria. An additional advantage of the decision framework proposed in this paper is that it intrinsically defines the conditions for terminating the search. More specifically, the search continues as long as the value of the belief function, as defined previously by Equation (1), lies within threshold values, $A$ and $B$ ($0 \leq B < A \leq 1$). A belief threshold value of $A = 1$ ($B = 0$) requires that the search terminate only when the searcher is absolutely certain that the target is present (absent). These decision thresholds offer additional degrees of freedom in the search problem definition, and can be used in the specification of the search problem. Our construction of the search problem has the form of the sequential hypothesis-testing methodology, or the Sequential Probability Ratio Test (SPRT), developed by Wald and Wolfowitz [9]. In the SPRT framework, observations are gathered as long as there is ambiguity in the acceptance or rejection of the hypothesis, which, in our case, is the determination of the presence or absence of the target within region $\mathcal{A}$ according to the thresholds $A, B$.

Further, in formulating the search problem as a decision to be made by the searcher, we can choose from a multitude of measures relevant to decision-making, such as time till decision, asymptotic confidence thresholds, rate-of-change of the belief function, and so on [10].

In this paper, we focus only on the amount of time expended until a decision is made regarding the presence or absence of the target as the measure of performance of the decision. This time till decision is denoted $t_d$, and is given by

$$t_d = \min \{ \arg(\mathbf{B}(t) = A), \arg(\mathbf{B}(t) = B) \}.$$

In the context of search and detection problems, this metric of time-to-decision is akin to the time-to-detection measure, as seen in the search theory literature [2].

Each search control strategy described below specifies which cell, $k^t$, to visit next at time $t$, given the sequence of previous observations up to time $t - 1$, $D^{t-1}$. Where relevant, dynamic constraints on the searcher may limit the reachability of cells and require that the next cell be within some distance $\kappa$ from the current cell, such that $\Delta k \triangleq \|k^t - k^{t-1}\| < \kappa$ for choice of $k^t$, where $\| \cdot \|$ denotes an appropriate distance metric. Further, since, by construction, the entire probability distribution function, $Pr(x_T | D^{t-1})$, at time, $t - 1$, is known, the cell containing the maximal cell belief value, $k_{max}^{t-1}$ at time, $t-1$, can easily be determined and used to guide the selection of $k^t$.

**Random Search:** Serving as a baseline strategy for comparison, a searcher employing a random walk provides a lower bound on the time to decision for any complete search, which is one that guarantees visiting all cells at least once [2]. The control signal for this strategy is:

$$k^t = \operatorname*{rand}_{\Delta k < \kappa}(k) \qquad \text{(Strategy I)}$$

where the rand function denotes the random selection of the next cell, given that it is a reachable cell. Note that this strategy does not require knowledge of the belief function. Variations on this strategy include biased random walk methods, where, on the average, the searcher tends toward

the peak of the distribution. Such an approach may be useful for low computation, local sensing of the belief gradient, which may be relevant in foraging or source (e.g. odor, chemical plume) localization.

**Sweeping Search:** The sweeping approach, where the searcher moves from one cell to the next in an incremental fashion, represents the classical Boustrophedan searching strategy used in classical robotic coverage algorithms [5]. The sweeping search also does not use the belief function in selecting the next cell to visit.

$$k^t = k^{t-1} + 1 \qquad \text{(Strategy II)}$$

One advantage of this approach is that the coverage of the region is done in linear time and minimizes the number of revisits to a cell (i.e. a cell will be revisited only after all other cells have been visited). However, in the presence of noisy detections, multiple measurements in individual cells may be beneficial for reducing the uncertainty present in those cells.

**Optimal Lookahead Search:** In optimal path-planning, the goal is to generate the search trajectory along which an objective function is maximized. As is well known, the solution to such problems is often computationally expensive, if not intractable [11]. E.g., the problem of search on a discrete grid to minimize the time to detection is $NP$-hard [12]. Choosing probability of detection as the objective function to be optimized, as in, e.g. [7], is even more difficult, being provably $NP$-complete. Hence, we must seek suboptimal solutions to solve the search path-planning problem, resorting to approximation or heuristic methods.

A common approach is to use a "lookahead" window of $w$ time steps, over which the optimization of the path is done. Also known in optimal control as *receding horizon control* [13], this approach generates a sequence of actions, or policy, for the next $w$ steps, denoted $\pi = \{k_1, \ldots k_w\}$, and finds the optimal policy, $\pi^*$, which maximizes the objective function along the path over the horizon window, i.e.

$$\pi^* = \arg \max_{\pi = \{k_1, \ldots k_w\}} \sum_j^w Pr(x_T = k_j | D^t, d_{k_1}, \ldots, d_{k_{j-1}}).$$

The first step of this optimal policy is then selected to be the next cell to be visited in the next time step. Implementation of the lookahead search algorithm in pseudo-code is described in Table I.

At the next time step, the optimal trajectory over the window is recomputed. Note that for $w = 1$, the one-step lookahead strategy is exactly the discrete gradient-approximation of the belief, which is simply the greedy algorithm for maximally increasing the objective function. More generally, the planning window size $w$ can be chosen to provide an approximation to the optimal solution while maintaining computational feasibility. In this manner, the closed-form expressions derived in this paper offer substantial implementation advantages by enabling search path plans over longer planning horizons.

**"Saccadic" Search:** Here we formalize a strategy that derives its name from the way the eye in the human visual

TABLE I
PSEUDO-CODE FOR THE LOOKAHEAD ALGORITHM

```
function Lookahead(current cell k^{t-1}, belief B_{t-1}, window size w)
    returns: optimal policy, π^{t-1} and utility, U(π^{t-1})

    Define set of allowable actions (e.g. list of reachable cells)
    For each action, k^t
        Take a detection measurement, d^t, in cell k^t
        Compute belief function, B_t, over all cells (Equation 2)
        If not at end of window
            Lookahead(k^t, B_t, w − 1)     % depth-first optimization
        Else
            Store belief in k^t as utility value
    end

    Identify the utility-optimizing action, k̄^t
    Return augmented policy π^t and utility, U(π^t)
```

system *saccades*, or executes a rapid jump in its focus of attention, from one salient feature in its visual field to another [14]. This strategy focuses the search on the cell containing the maximal belief probability at every time step, executing a "saccade"-like jump from peak to peak.

$$k^t = k_{max}^{t-1} \triangleq \arg \max_k (Pr(x_T = k | D^{t-1})) \quad \text{(Strategy IV)}$$

As alluded to by its namesake, this search strategy may offer some insight into or provide a model for understanding the mechanisms governing human visual search, in which information is likely gathered sequentially using foveated glimpses of patches of the visual field, which can be modeled as observations of cells in the presented framework. Other relevant uses include, as an example, sensor networks for event detection applications, where sensor nodes (i.e. cells) are triggered by possible events (such as intrusion). In these applications, a belief-dependent decision or action is executed in the context of search. The framework presented in this paper enables the examination of these additional classes of the search problem beyond the canonical example of autonomous search of a physical environment using mobile sensor platforms, and is one of the main contributions of this work.

**"Drosophila-inspired" Search:** One of the limitations of the previous proposed search strategy is that "saccading" to the cell with the highest belief value in the next step may not be dynamically feasible, e.g. for a mobile sensor platform with finite velocity. However, the computational simplicity of determining the maximum belief value offers an advantage to more computationally expensive, planning-oriented approaches. We propose this strategy as a low-computation-expense, high-information-gain alternative method. Inspired by observations of the search trajectories of individual fruit flies (*Drosophila melanogaster*), which are walking on a surface in search of food using visual sensory feedback, this strategy mimics the insects' behavior of finding a peak, pursuing a straight-line trajectory (with minimal deviation) towards this peak, and upon arrival, finding another peak and following another straight-line path to this next peak [15]. In this algorithm, the searcher determines and stores the cell with the highest belief value, $k_{max}$. Recall that $\kappa$ represents

a constraint on how far the searcher can travel in a single time step, and can either be given by the searcher dynamics or specified by the desired resolution of the search. Then the next cell to observe is determined by:

$$k^t = \arg \min_{\Delta k < \kappa} \left( \|k_{max} - k^{t-1}\| \right) \qquad \text{(Strategy V)}$$

For fruit flies walking on a flat surface (i.e. in continuous $\mathbb{R}^2$ space), the use of the 2-norm (i.e. Euclidean distance) offers a reasonable approximation to their behavior, and $\| \cdot \|$ is also used and shown in [6] to address scenarios of premature search termination due to sensitivities in the optimization routine.

*B. Numerical Studies and Discussion*

This section investigates the performance of each search control strategy in simulation. Because we allow for uncertainty in the sensor detection process, the belief evolution is studied in a statistical manner to deduce relevant parameters and conclusions for the decision-making process. In particular, we are interested in the time steps – equivalently, the expected number of observations – until decision.

For illustrative purposes, our simulations are based on a $10 \times 10$ grid, with the prior belief probability density function modeled as a discrete approximation to a Gaussian distribution (see Figure 3), centered about the cell $(i, j) = (1, 3)$ with covariance values $\sigma_i = 7$ and $\sigma_j = 15$, and appropriately normalized such that $\sum_{a=1}^{|\mathcal{A}|} Pr(x_T = a) = \delta$. However, it is important to note that any form for the initial belief PDF can be utilized due to the Bayesian formulation of this framework.
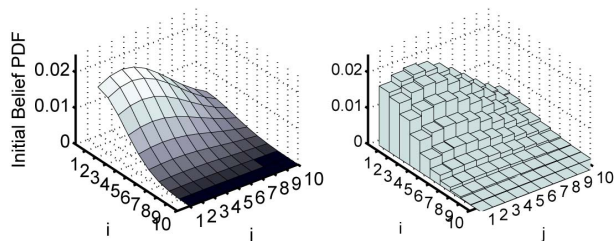


Fig. 3. The model of a Gaussian distribution and its discrete approximation for the initial prior belief probability density function.

Consider a single, stationary target located in cell $(4, 3)$ in the grid of $|\mathcal{A}| = 100$ cells. Note that in this example the initial belief (i.e. prior to taking any observations) allocates belief probability erroneously by attributing greater likelihood of finding the target in cells which do not contain the target, and vice versa. The robustness of the decision-based belief framework presented in this paper is demonstrated by the eventual correction of this "bad guess" by sequential accumulation of information.

Simulation runs of 5000 trials of each of the different search strategies were executed for constant values of initial prior belief $\delta = 0.5$, and false alarm and missed detection error rates, $\alpha = 0.2$ and $\beta = 0.1$, respectively. Random noise is injected into the observation process according to these error rates thereby introducing randomness in the iterations on the evolution of the belief function. A sample run of

the simulation utilizing the "saccadic" search strategy is depicted in Figure 4, for which the belief evolution is shown in Figure 5.
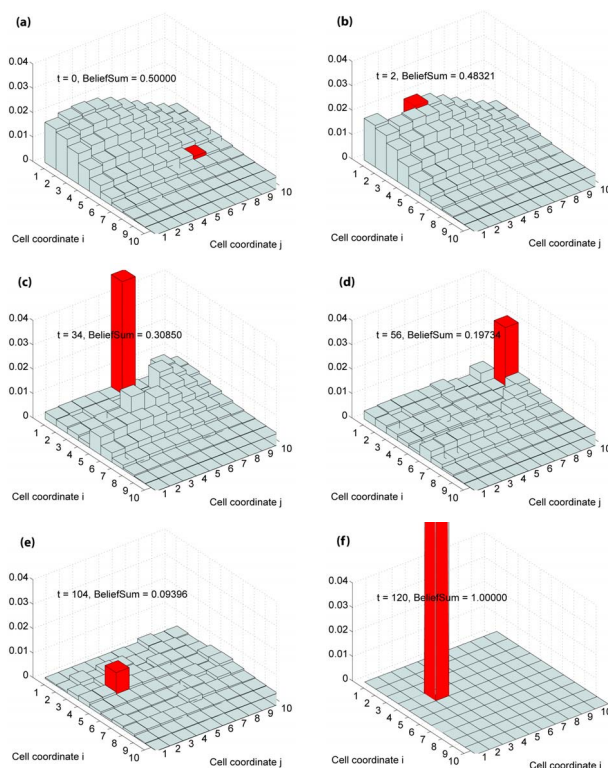


Fig. 4. Sample simulation run implementing the "saccadic" search strategy. The highlighted cell represents the one currently being observed.
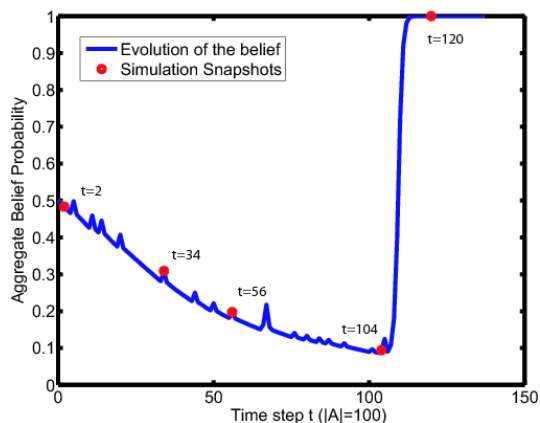


Fig. 5. Plot of the belief evolution over the duration of the example simulation run. The search terminates when the belief reaches 1, corresponding to the decision that the target is present in $\mathcal{A}$.

**Random walk and sweeping strategy.** As mentioned previously, the random walk and sweeping strategies generate search paths that ignore any available information about the belief of the target's presence. As such, these approaches do not offer much in the way of performance, as the time until decision are orders of magnitude larger when compared to the lookahead, "saccadic," or "*Drosophila*-inspired" strategies. Instead, the random walk and sweeping methods can serve as lower bounds on performance for

strategies that do take the current belief into account in determining search trajectories.

**Look-ahead strategy.** Figure 6 illustrates the histogram plot of the decision times for those trials employing the lookahead strategy for varying window sizes, $w = \{1, 2, 3\}$.
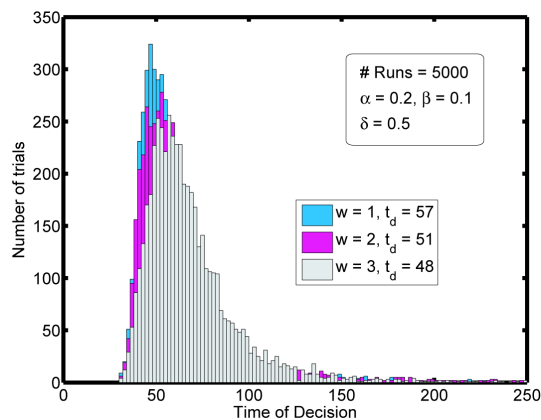


Fig. 6. Histogram plots of the lookahead search strategy for window sizes, $w = \{1, 2, 3\}$, with modes $t_d = \{57, 51, 48\}$, respectively.

From this plot, we can easily identify the mode (i.e., the most frequent value of the histogram) of the decision time distributions for each of the window sizes to be:

$$\text{decision time, } t_d : \begin{cases} 57, & w = 1 \\ 51, & w = 2 \\ 48, & w = 3 \end{cases}$$

As expected, there is a slight improvement in performance (i.e. a shorter time to decision) for increasing window sizes, which is due to the fact that, for longer planning horizons, additional information is taken into account in generating the search path. However, given the fact that the performance of the lookahead algorithm is greatly dependent on initial conditions (e.g. initial location of the searcher relative to the target, error in the initial prior belief probability distribution, size of the environment, etc.), the cost of computational burden seems to outweigh the improvement in performance. Instead, observing that the random-walk and sweeping strategies are extremely inefficient in their searches, a one-step lookahead control law (i.e. steepest gradient ascent), albeit suboptimal, may serve as a reasonable compromise between performance and complexity.

**"Saccadic" and "*Drosophila*-inspired" strategy.** Figures 7 and 8 illustrate the average behavior of the "saccadic" and "*Drosophila*-inspired" strategies. The "saccadic" approach yields a distribution of decision times with the mode at $t_d = 61$, where, as might be expected, the "*Drosophila*-inspired" approach requires additional observations, making a decision at $t_d = 89$ on the average. This relationship between these two strategies is intuitive, considering that the "saccadic" approach accrues the maximal amount of belief information at every time step, while, in the "*Drosophila*-inspired" strategy, the transit time to get to this maximal cell must be integrated. In this manner, this latter approach can be thought of as "saccadic" search with inertia or damping.
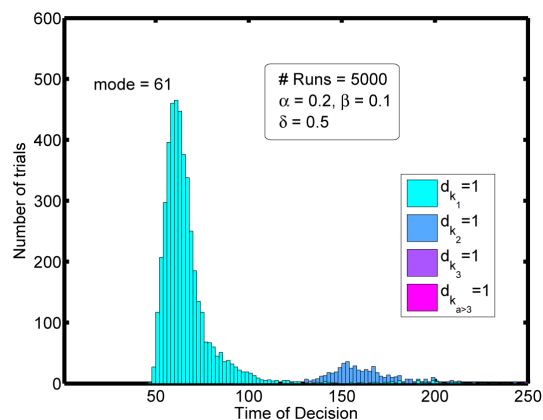


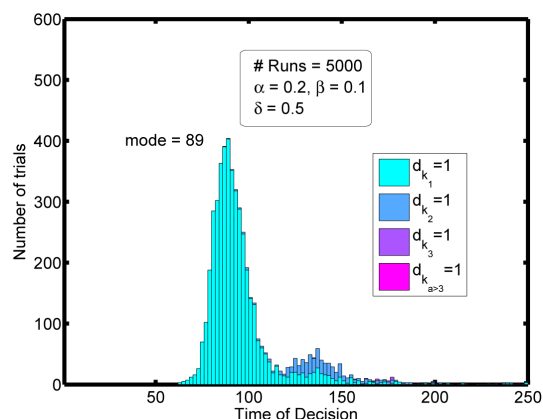Fig. 7. Histogram of the "saccadic" search strategy, with mode $t_d = 61$.



Fig. 8. Histogram of the "*Drosophila*-inspired" search strategy, with mode $t_d = 89$.

An interesting observation can be made regarding the appearance of additional smaller peaks in the histogram plots of these strategies. These secondary peaks correspond to distinct "modes" in the search evolution, which depend on the sequence of observations of the cell containing the target. As depicted in Figures 7 and 8, the primary peak corresponds to the trials where the first visit to the target cell resulted in a positive detections, and similarly for the subsequent peaks and corresponding revisits of the target cell.

These two strategies provide very computationally cheap approaches to search, which serve well in a host of applications where computation power is limited, e.g. sensors embedded on micro-aerial vehicles or simple nodes in a sensor network. Specifically, while the lookahead strategies offer slightly better performance, a more comprehensive metric of performance which accounts for computing time of search paths would immediately illuminate the immense advantages of the "saccadic" and "*Drosophila*-inspired" strategies. In the case of search using mobile sensors, we see that in these scenarios it may be more prudent to have simpler, fast-moving detectors rather than necessitating computationally powerful (which usually translates to more hardware and less dynamic capability) sensor platforms. Furthermore, unlike the optimization-based lookahead schemes, these strategies have the added benefit that their performance is largely

TABLE II

COMPARISON OF TIME-TO-DECISION PERFORMANCE OF VARIOUS

BELIEF-BASED SEARCH CONTROL STRATEGIES

| Strategy | # of time steps |
|---|---|
| Lookahead (w=1) | 57 |
| Lookahead (w=2) | 51 |
| Lookahead (w=3) | 48 |
| "Saccadic" | 61 |
| *Drosophila*-inspired | 89 |

independent of the initial conditions, and they also avoid the issues of getting stuck in local minima of the objective function which plague most optimization routines.

A side-by-side comparison of the search control strategies which use the belief function in determining the search path is provided in Table II.

## V. CONCLUSIONS AND FUTURE WORK

The theory of search is relevant to many interesting applications, and as such, has engendered a rich and active field of research, particularly in the study of searching for targets using mobile sensor platforms. In many of these contexts, the problem of interest is the generation of search paths that will improve the detection of the target(s) by the optimization of some objective function.

The research presented here offers a Bayesian framework that views search as a decision problem. The approach allows for imperfect sensors, which have not been well addressed in the literature on robotics coverage, and also allows for additional theoretical results and insights. Specifically, we formulated and investigated the belief evolution of the target's presence in the search region. We derived analytic closed-form expressions that govern the belief function, which serve two main purposes, both offering a more general understanding of the search problem (e.g. notions of rate-of-change of belief) and facilitating the easier implementation of search strategies in physical search systems.

Additionally, we examined and proposed several different strategies for generating search paths and found that the decision framework enables the use of a variety of measures of the search performance, including the time until a decision is made about the presence or absence of the target. We evaluated the different approaches using this metric in simulation studies, observing that slight increases in performance are often not worth the computational burden, which is a relevant and/or necessary consideration for many applications of interest.

Further, the two approaches proposed – the "saccadic" and "*Drosophila*-inspired" strategies – highlight the relevance and general notion of the search problem to other classes of search-related tasks. The decision-based formulation discussed in this paper offers a common framework for examining these types of problems, ranging from human visual search to efficient event detection in sensor networks to the search ethology exhibited by insects, and perhaps more interestingly, may provide a method for uncovering the mechanisms which govern the behavior of systems in these search tasks.

There are many avenues for future research. Of immediate interest is the investigation of other measures of performance, including notions of achievable confidence in the decision for limitations on the number of observations as well as the role of (possibly changing) error rates $\alpha$ and $\beta$ on the evolution of the belief. These concepts can be best studied by definition and derivation of the expected value of the belief function, due to the randomness introduced by uncertainty in the detections. Utilizing tools such as the SPRT may offer avenues for theoretical developments.

In the context of search with mobile sensors, a wide range of interesting problems exist, including the use of the decision framework for examining target identification, switching task allocation (e.g. whether to continue the search vs. perform the rescue), and coordination of a team of sensors which may fuse decisions rather than use sensor-based estimates. Another particularly intriguing extension to this work is the search for antagonistic targets, which maneuver to hinder the progress of the search. Game-theoretic notions, such as *levels of rationality* (which are related to lookahead strategies), are certainly relevant as applied to search [16] and pursuit-evasion [8], and thus merit further research into these areas.

## REFERENCES

[1] B. O. Koopman, "Search and its optimization," *The American Mathematical Monthly*, vol. 86, no. 7, pp. 527–540, Aug-Sept 1979.

[2] L. D. Stone, *Theory of Optimal Search*, 2nd ed., J. Kettelle, Ed. Academic Press, 1989.

[3] A. R. Washburn, *Search and Detection*, 4th ed., ser. Topics in Operations Research Series. INFORMS, 2002.

[4] J. Cortes, S. Martinez, T. Karatas, and F. Bullo, "Coverage control for mobile sensing networks," *IEEE Transactions on Robotics and Automation*, vol. 20, no. 2, pp. 243–255, April 2004.

[5] E. U. Acar, H. Choset, Y. Zhang, and M. Schervish, "Path planning for robotic demining: Robust sensor-based coverage of unstructured environments and probabilistic methods," *The International Jounal of Robotics Research*, vol. 22, no. 7-8, pp. 441–466, July-August 2003.

[6] E.-M. Wong, F. Bourgault, and T. Furukawa, "Multi-vehicle bayesian search for multiple lost targets," in *Proc. of Intl. Conf. on Robotics and Automation*, April 2005, pp. 3169– 3174.

[7] S. Thrun, W. Burgard, , and D. Fox, *Probabilistic Robotics*. MIT Press, 2005.

[8] R. Vidal, O. Shakernia, H. Kim, D. Shim, and S. Sastry, "Probabilistic pursuit-evasion games: Theory, implementation and experimental evaluation," *IEEE Transactions on Robotics and Automation*, vol. 18, no. 5, pp. 662–669, Oct 2002.

[9] A. Wald and J. Wolfowitz, "Bayes solutions of sequential decision problems," *PNAS*, vol. 35, pp. 99–102, 1949.

[10] A. Wald, *Sequential Analysis*. New York, NY: John Wiley and Sons, Inc., 1947.

[11] D. Bertsekas and S. Shreve, *Stochastic Optimal Control: The Discrete Time Case*. Academic Press, 1978.

[12] K. E. Trummel and J. R. Weisinger, "The complexity of the optimal searcher path problem," *Operations Research*, vol. 34, no. 2, pp. 324–327, Mar. - Apr. 1986.

[13] D. Mayne, J. Rawlings, C. Rao, and P. Scokaert, "Constrained model predictive control: Stability and optimality," *Automatica*, vol. 36, pp. 789–814, 2000.

[14] L. Itti and C. Koch, "A saliency-based search mechanism for overt and covert shifts of visual attention," *Vision Research*, vol. 40, pp. 1489–1506, 2000.

[15] A. A. Robie and M. H. Dickinson, "Walking *Drosophila* demonstrate shape preference during local search but not global search." Society for Neuroscience, November 2005.

[16] S. Gal, *Search Games*. Academic Press New York, 1980.