Proceedings of the
47th IEEE Conference on Decision and Control
Cancun, Mexico, Dec. 9-11, 2008

Search Decisions for Teams of Automata

Dimitar Baronov and John Baillieul

***Abstract*— The dynamics of exploration vs exploitation decisions are explored in the context of robotic search problems. Building on prior work on robotic search together with our own work on reactive control laws for potential field mapping, we propose a new set of search protocols for teams of sensor-enabled mobile robots. The focus is on collaborative strategies for the search of potential fields that are possibly time varying. We pose the problem of quickly finding regions where the potential achieves or exceeds a certain threshold. The search protocol has two distinct components. In an "exploration phase", agents execute either a randomized or structured search, seeking places where the field achieves or exceeds the prescribed threshold. Once a threshold point is reached, the "exploitation" component is initialized and the agents deploy so as to rapidly map the evolving isoline associated with the given value of the field. Conservative strategies will emphasize refining the detailed knowledge of the field in a small neighborhood of the isoline, while aggressive strategies will emphasize wide-ranging exploration of neighboring territory. The main decision problem under study involves finding the optimally aggressive exploration strategy. Additionally, the problem of the allocation of the agents between "exploration" and "exploitation" is considered. A performance metric is developed to compare the proposed methods with standard approaches such as random search and distributed raster scans.**

I. Introduction

In various forms, studies of optimal search strategies have a long history. Formal search methods have been developed and incorporated into standard texts and monographs in engineering and computer science under subject headings such as heuristic search, branch-and-bound search, depth-first search, and so forth. Recently, a great deal of research has been focused on algorithms for distributed exploration by multiple autonomous vehicles [1], [2], [3]. The goal is to develop distributed control laws which are verifiably correct while allowing the closed-loop behavior of the system to depend on only local loop-closure and asynchronous operation of each vehicle. The present paper is part of a larger study of the dynamics of decision making in teams of agents that may include both humans and autonomous robots. A central focus is on operational decisions regarding whether to emphasize breadth or depth and how each agent's decisions affect the overall performance of the team.

Several aspects of this research set it apart from much of the literature to date dealing with robotic search. An important point is that we are studying decisions that are localized in both space and time. Each agent in the team problems we study will operate based on information acquired both from on-board sensors as well as from communication with other agents on the team. Agents will also change behavior depending on how they perceive they are contributing to the goals of the search.

This work was supported by ODDR&E MURI07 Program Grant Number FA9550-07-1-0528 and by the National Science Foundation ITR Program Grant Number DMI-0330171, all to Boston University.

D. Baronov and J. Baillieul are with the Department of Aerospace and Mechanical Engineering, Boston University, Boston, MA, 02215. {baronov,johnb}@bu.edu

Binary decisions are made by all levels of life forms – from bacteria that undergo rolling or tumbling motions based on chemotaxis [4] to humans where a growing body of literature seeks to understand the dependence of simple decisions on the neurodynamics of the brain [5]. Similarly in our problem formulation, a key decision that each agent must take at each time step is whether to examine an area or object of interest more closely or to go off to search unexplored territory. This is the so called "exploration" versus "exploitation" strategy trade-off. We shall study this choice using an information based utility metric. Although, inspired by the classical Shanon information theory, our present treatment of information is aimed at the specifics of the robotic search.

With the notion of function complexity developed in Sec. II, we are able to define the information content of an unknown environment in terms of what is known about its structure. We show how in this context a conceptual analog of mutual information can be applied to quantify the dynamics and the performance of the search process. Sec. III establishes the ability of an agent to discern between high and low information yield as a basis of the exploration versus exploitation decision.

We test our approach in a problem that has been a subject of extensive research in the last few years, a search of potential fields [6], [7], [3]. Building on our previous work on a reactive control law for level sets mapping [8], we are able to develop simple search strategies that enable a team of robotic agents to identify areas in which the field intensity exceeds a predefined threshold. Sec. IV shows the results of the explore-exploit strategy, and demonstrates the benefits of using our information metric to evaluate different search strategies.

II. Mathematical models for information-based search

We assume that for the problems being studied an unknown environment can be abstracted by a time dependent map $f : \mathbb{R}^m \times (0, \tau] \to \mathbb{R}^n$, on some compact and simply-connected domain $X \subset \mathbb{R}^m$, which is the area of interest. Although, in the current work we will look at scalar potential fields on the plane, that is $m = 2$ and $n = 1$, the notions developed in this paper can be easily extended to the general setup.

In a search process, a sensor-enabled agent accumulates information about the unknown environment by collecting multiple measurements in the domain of interest. The common thread in the search strategies being investigated is that members of a team of search agents must make decisions about where and at what level of detail to conduct the search. Assuming that the agents have some incentive to minimize time-to-complete, formal methods are needed to assess speed-versus-accuracy trade-offs. Our goal then is to define metrics in terms of which it is possible to make statements about how fast a particular search strategy is yielding new information.

A. The complexity of functions

To develop the information based approach to the search problem, we start by defining the concept of *function complexity*. We can think of a function $f : X \to Y$ as a communication channel that provides information in the *range* Y about the structure of the *domain* X. We can then think of the *complexity* of f as some measure of how well f "informs" Y about X.

Suppose X admits a measure μ which is nonsingular with respect to Lebesgue measure and $\mu(X) < \infty$. Suppose further that $\Delta = \{\delta_j : 1 \le j \le m\}$ is a finite partition of Y. Then, the domain X can be partitioned into the connected components of $\{f^{-1}(\delta_j)\}$, that is:

$$\mathcal{V}(\Delta) = \text{cc}\{f^{-1}(\delta_j) : 1 \le j \le m\}$$

where for any set $S \subset \mathbb{R}^2$, $\text{cc}\{S\}$ denotes the set of connected components of S.

The resulting domain partition expressed as $\mathcal{V}(\Delta) = \{V_i : 1 \le i \le n\}$, has cardinality satisfying $n \ge m$. Through its structure, we define the *complexity* of f with respect to Δ as

$$\mathcal{C}(f, \Delta) = H(\mathcal{V}) := -\sum_{j=1}^{n} \frac{\mu(V_i)}{\mu(X)} \log_2 \frac{\mu(V_i)}{\mu(X)},$$

where $H(\mathcal{V})$ is defined as the entropy of the domain partition.

Defined in this way, the entropy of the domain partition, $H(\mathcal{V})$, resembles Shannon's information theoretic definition of entropy. To pursue the analogy, $H(\mathcal{V})$ resembles the entropy of a random process which has as events the cells in the partition whose probability of occurrence is equal to the normalized Lebesgue measure.

The properties of the *function complexity* are a direct consequence of this analogy and are as follows:

1) If Y contains a single element (i.e. if f is a constant), then $\mathcal{C}(f, \Delta) = 0$.
2) If all elements V_i in the domain partition have identical measure $\mu(V_i)$, then $\mathcal{C}(f, \Delta) = \log_2 n$.
3) If $\mu(V_i) \neq \mu(V_j)$ for some pair of cells $V_i, V_j \in \mathcal{V}$, then $\mathcal{C}(f, \Delta) < \log_2 n$.

It is less straightforward to quantify the complexity of f if either Y or the measure of X are infinite. These cases can be treated by making problem-specific assumptions, but a detailed discussion is beyond the scope of this paper. Instead we show how the notion of complexity applies to scalar potential fields. There, given a range partition Δ, the domain will be partitioned into the connected components of the level sets. For example, if $S(\mathbf{r})$, $S : \mathbb{R}^2 \to \mathbb{R}$ is a particular potential function bounded by $S_{sup} = \sup\{S(\mathbf{r} : \mathbf{r} \in X)\}$ and $S_{inf} = \inf\{S(\mathbf{r}) : \mathbf{r} \in X)\}$, we can chose a range partition $\{s_j\}$, where $s_j = S_{inf} + j\delta$, $j \in [0, m]$ and $\delta = (S_{sup} - S_{inf})/m$. The associated domain partition than has cells

$$V_i \in cc\{s_{j-1} \le S(\mathbf{r}) \le s_j : \mathbf{r} \in X\}.$$

With this example, the symbolic similarity with notions of entropy in information theory is apparent. One can think of the potential surface as conveying a message wherein small segments of different mean potential value constitute the letters used to encode the message. With this analogy in mind, the complexity can be thought of as the entropy of the message. A flat potential field will have zero complexity for all choices of partition coarseness Δ.

B. Mutual information

We assume that in the search process, the range of the potential function abstracting the unknown environment has a constant, predefined partition Δ of its range. Then, from the viewpoint of the agents conducting the search, for any time, there will exist two separate perspectives of the state of the domain. Unknown to the searchers, there exists an *objective perspective* expressing ground truth. This perspective can be thought as a reconstructed contour map of the potential given no measurement noise and infinite amount of sensor information. There also exists a *subjective perspective* representing the accumulated information by the searchers, that is a collection of mapped level sets.

Both of these perspectives can depend on time. The objective perspectives, we denote by $\mathcal{V}_k^o = \{V_j^o : j \in L_k^o\}$. Given that the range partition is predefined and constant, the index set L_k^o and its associated cells will depend on time only if the potential function also depends on time, that is the potential function is $S(\mathbf{r}, t)$, $S : \mathbb{R}^2 \times (0, \tau] \to \mathbb{R}$.

The objective perspective, $\mathcal{V}_k^s = \{V_j^s : j \in L_k^s\}$, on the other hand, will evolve as the search yields more information. For example, given no information in the beginning, we can assume that there is a time t_1 such that for all $0 \le t \le t_1$, the index set $L_t = \{0\}$, where the set V_0^s is the whole space and $L_{t_1} = \{0, i_1\}$, when the first cell is mapped.

With more information successively added, the ability of an observer to deduce the state of the *objective perspective* from the state of the *subjective perspective* is increased. In other words, as the search progresses the *subjective perspective* should in some sense converge to the *objective perspective*. In this context, to quantify the properties of the search process, we need to establish a metric of how close the two perspectives are. This metric should capture the information gathering aspects of the process as well. That is to say, that the metric should be able to express the distance between the perspectives as function of both the acquired information and the total information content of the unknown environment.

As established above, the *function complexity* is well equipped to capture the information content of the unknown environment (a scalar potential function in our case) with respect to its chosen range partition Δ. Therefore, given that there is no prior knowledge about the environment, the entropy of the *objective perspective* will characterize how far away the two perspectives are in the beginning of the search process,.

We define the *joint entropy* between the two perspectives as

$$H(\mathcal{V}_k^o, \mathcal{V}_k^s) = -\sum_{i \in L_k^s} \sum_{j \in L_k^o} \frac{\mu\left(V_i^s \cap V_j^o\right)}{\mu(X)} \log_2 \left(\frac{\mu\left(V_i^s \cap V_j^o\right)}{\mu(X)} \right).$$

We note again that there are obvious similarities between the so defined *perspectives joint entropy* and the more classical notion of the joint entropy of two random variables in information theory. In this case, the $\frac{\mu\left(V_i^s \cap V_j^o\right)}{\mu(X)}$ elements are analogous to the joint probabilities of occurrence between the events V_i^s and V_j^o, where in reality they correspond to the cross section between respectively a believed shape of a particular level set and an actual level set.

The joint entropy has the following properties:

$$H(\mathcal{V}_k^o, \mathcal{V}_k^s) \leq H(\mathcal{V}_k^o) + H(\mathcal{V}_k^s) \tag{1}$$

and

$$H(\mathcal{V}_k^o, \mathcal{V}_k^s) = H(\mathcal{V}_k^o) = H(\mathcal{V}_k^s), \text{ if } \mathcal{V}_k^o \equiv \mathcal{V}_k^s,$$

where these properties can be proved as a direct consequence of the analogy with information theory. (See [9] for the properties of the random variables based joint entropy.)

Note, that (1) becomes equality when the two perspectives, in some sense, maximally differ. From a probability point of view, this is equivalent to the two variables being independent.

The joint entropy finally allows us to define as a convergence metric the *perspectives mutual information* given by:

$$I(\mathcal{V}_k^o; \mathcal{V}_k^s) = H(\mathcal{V}_k^o) + H(\mathcal{V}_k^s) - H(\mathcal{V}_k^o, \mathcal{V}_k^s). \tag{2}$$

Given that the search process starts with no prior information, the *subjective perspective* as discussed above has a single element $V_0 = X$, and it can be confirmed by (2) that in this case the mutual information is equal to 0. As $\mathcal{V}^s$ converges to $\mathcal{V}^o$, on the other hand, the mutual information becomes $I(\mathcal{V}_k^o; \mathcal{V}_k^s) = H(\mathcal{V}_k^o) = H(\mathcal{V}_k^s)$.

We will further show how the *perspectives mutual information* evolves under the search process and demonstrate its applicability in evaluating the efficiency of a given search strategy.

III. Search decisions and the information flow

The search process will have two clearly distinguished components. At any instant of time a search agent should decide between two clear choices: whether it should collect information from a particular area; or whether it should explore the space in search of other parts of the domain with richer information content. This problem is known as the *exploration* vs *exploitation* paradigm, since the decision concerns whether the individual should *exploit* its current state or *explore* the space for better possibilities. In the context of potential field search, *exploitation* is equivalent to the mapping of the level sets within a certain area. *Exploration*, on the other hand, is the behavior of moving from the mapping of the level sets in one area to their mapping in another.

In our work we want to investigate strategies which balance *exploration* and *exploitation* based on the information yield of a particular area. High information yield will force the search agent to engage with *exploitation*. Boredom or the lack of information, on the other hand, will trigger *exploration*.

We first start by considering a static environment, that is $S(\mathbf{r}, t) = S(\mathbf{r})$. In an exploitation mode, the robotic agent generates a map of the connected components of a certain level set. These connected components as cells are then submitted to a cumulative list that updates the *subjective perspective*. Lets t_{k-1} and t_k be two consecutive update times of the *subjective perspective*, that is to say that t_{k-1} and t_k correspond to times at which a given cell map is completed. Moreover, we assume that each time a new structure is mapped the index set L_k^s is increased with one element and that all mapping is done with infinite precision. The information gained between the two updates of the *subjective perspective* is given by:

$$\begin{aligned} \Delta I_k^+ &= I(\mathcal{V}^o, \mathcal{V}_k^s) - I(\mathcal{V}^o, \mathcal{V}_{k-1}^s) \\ &= H(\mathcal{V}_k^s) - H(\mathcal{V}_{k-1}^s) - H(\mathcal{V}^o, \mathcal{V}_k^s) + H(\mathcal{V}^o, \mathcal{V}_{k-1}^s). \end{aligned}$$

However, the assumption that cells are mapped with infinite precision allows us to calculate the information gain without any knowledge of the *objective perspective* as:

$$\begin{aligned} \Delta I_k^+ &= I(\mathcal{V}_k^s, \mathcal{V}_k^s) - I(\mathcal{V}_k^s, \mathcal{V}_{k-1}^s) \\ &= H(\mathcal{V}_k^s, \mathcal{V}_{k-1}^s) - H(\mathcal{V}_{k-1}^s). \end{aligned} \tag{3}$$

One can think of ΔI^+ as the information inflow of the search process originating from the search party. Under this inflow the mutual information will evolve in time as:

$$I_{k+1} = I_k + \Delta I_k^+,$$

where we have defined I_k as

$$I_k := I(\mathcal{V}_k^o, \mathcal{V}_k^s).$$

The evolution of the mutual information is not so straightforward to describe, however, when the search is conducted in time dependent environments. In this case, the search agent can map a particular structure $\{\mathcal{V}_k^o\}_j$ of the environment at time t_k, but at time t_{k+1}, its shape could have changed or even the whole structure could have disappeared. Therefore, during the search process there will not just be accumulation of information but also dissipation due to the changes of the underlying environment, $S(\mathbf{r}, t)$. Thus we need to express the dynamics of the mutual information as the balance between the inflow and the outflow of information or in other words:

$$I_{k+1} = I_k + \Delta I_k^+ - \Delta I_k^-.$$

In this general case, we define the *information inflow* and the *information outflow*—respectively ΔI_k^+ and ΔI_k^-—as

$$\Delta I_k^+ = I(\mathcal{V}_k^o, \mathcal{V}_k^s) - I(\mathcal{V}_k^o, \mathcal{V}_{k-1}^s) \tag{4}$$

$$\Delta I_k^- = I(\mathcal{V}_{k-1}^o, \mathcal{V}_{k-1}^s) - I(\mathcal{V}_k^o, \mathcal{V}_{k-1}^s). \tag{5}$$

Note that according to these definitions the *information inflow* is the difference between the mutual information immediately before and after the *subjective perspective* is updated, where the *information outflow* is the total information lost in the period between the two updates of the *subjective perspective*. Given a static environment, that is $\mathcal{V}_{k-1}^o \equiv \mathcal{V}_k^o$, there is no information dissipation and $\Delta I_k^- = 0$. At the other extreme, rapid changes in the *objective perspective* lead to the complete dissipation of the previously acquired information. In other words, the range of ΔI_k^- obeys

$$0 \leq \Delta I_k^- \leq I_{k-1}. \tag{6}$$

When the environment is static, we can explicitly evaluate the utility of exploiting a given region by evaluating the *information inflow* through (3). For the time dependent environment, however, the relationship between the evolution of the *subjective perspective* and the *information inflow* is not exact and becomes:

$$\Delta I_k^+ \leq H(\mathcal{V}_k^s, \mathcal{V}_{k-1}^s) - H(\mathcal{V}_{k-1}^s).$$

(See the appendix for proof of this inequality.)

We define the bound $c_k := H(\mathcal{V}_k^s, \mathcal{V}_{k-1}^s) - H(\mathcal{V}_{k-1}^s)$ as the *information capacity* of the current action. Given that it takes ΔT units of time between the time an agent initiates its exploitation phase and the time it produces a map of a cell, we can think of $c_k/\Delta T$ as the information rate originating from this agent. In what follows, we will give an example of a simple threshold search strategy, in which an autonomous agent decides whether to be engaged in exploitation based on the value of this quantity. As the *information rate* drops, boredom will trigger exploration. We note that more elaborate strategies closer to actual human behavior are possible, in which the robot takes into account both short and long term utility (*information rates* in the presented setup) similar to the models shown in [5]. However, though simplistic, the threshold strategy presented in the next section will illustrate the benefits of the information based approach to the search problem, and will validate a framework against which human decision making can be investigated

IV. Applying information-based criteria in potential field search

Potential fields can abstract the spatial distribution of such properties as temperature, radioactivity, contaminant concentration and etc. The challenge in monitoring and mapping such fields is that they can be sensed only locally and thus to accumulate information a search agent must move throughout the domain of interest. A standard technique to prescribe trajectories for sensing is the raster scan. Although universally applicable, the raster scan is generally not the most efficient technique. A mobile sensor moving on a raster scan trajectory will spend its time uniformly in the search domain, but if the information content of the environment is not uniformly distributed much of this time will be wasted. This motivates the development of strategies that, on one hand, provide the user with accurate descriptions of some "interesting" environmental features and, on the other, parsimoniously allocate resources for this type of information acquisition.

An example, in which the raster scan is not the most efficient tool for mapping, is the identification of areas in which the potential field exceeds a predefined threshold. In what follows, we demonstrate the ability of our information metric to capture the inefficiency of the raster scan technique compared with a simple exploration vs exploitation approach to problem.

A. Modeling consideration

We will consider a potential field consisting of a collection of uniform radial sources. The sources are arranged on two concentric circles, and time dependence of the potential field is achieved through the rotation of the sources around the origin with constant turning rate (Fig. 1).

Given that the scalar potential field is defined as $S(\mathbf{r}, t)$, and the domain of interest is $\mathbf{r} \in D \subset \mathbb{R}^2$, we partition the range based on the threshold S_d as $(-\infty, S_d]$ and $[S_d, \infty)$. This leads to the partition of the domain into cells

$$V_i^o \in \text{cc}\{\mathbf{r} \in D | S(\mathbf{r}, t) \geq S_d\}, \ i \geq 1,$$

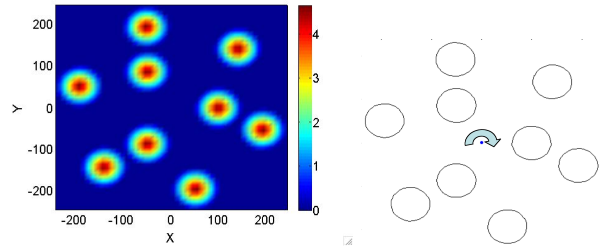

Fig. 1. The potential function of the example problem, and the resulting domain partition, $j \in [0, 9]$. Time dependence is achieved through the rotation of the sources around the origin with constant turning rate.

and

$$V_0^o = \{\mathbf{r} \in D | S(\mathbf{r}, t) \leq S_d\},$$

where we have implicitly assumed that the sets V_i^o, $i \geq 1$ are simply connected. It follows that the *objective perspective*, $\mathcal{V}_k^o$, is the collection of these cells for a given time $t = t_k$.

The *subjective perspective*, $\mathcal{V}_k^s = \{V_j^s : j \in L_k^s\}$, on the other hand, evolves as follows: there exist a time t_1 such that for all $0 \leq t < t_1$, the index set $L_t^s = \{0\}$, where the set V_0^s is the whole space; then for $t = t_1$, the first region exceeding the threshold potential is mapped and the index set becomes $L_{t_1}^s = \{0, i_1\}$. For every consecutive update t_k, $k \geq 2$, there are two possibilities; either a new structure is discovered and a new element is added to L_k^s or the map of a particular structure V_i^s is replaced with a more recent update.

B. Exploration and exploitation in terms of motion primitives

The agents conducting the potential field search are governed by a library of possible feedback control laws. We call these control laws *motion primitives* and they constitute the building blocks for higher level behaviors such as exploration and exploitation. These behaviors can be formally described by symbolic control languages, and therefore the notation that we will further use is consistent with the motion description languages MDL (See [10]).

We assume that the search team consists of identical physical plants:

$$\dot{\mathbf{r}}^i = f(\mathbf{r}^i, u) \tag{7}$$

where $\mathbf{r}^i \in \mathbb{R}^2$ is the position of the i^{th} agent in the plane and the control law u is a particular *motion primitive*.

Either of the components of the search process can be described through assigning the robots pairs of *motion primitives* and *trigger functions*, where the *trigger functions* are responsible for interrupting the given motion. The pair of a motion primitive and a trigger, that is $\{\xi, \mathbf{u}\}$, is called an atom, where the interrupt $\xi : \mathbb{R}^k \to \{0, 1\}$ maps the data from the k sensors of a given agent into a boolean number. In other words, $\{\xi, \mathbf{u}\}$ is treated as *execute* $\mathbf{u}$ *as long as* $\xi = 1$.

In general, a given exploration or exploitation behavior can be described as a sequence of atoms $b = a_1 a_2 ... a_n$, which will lead the robot to execute consecutively each one of these atoms until its respective interrupt function is activated. However, in the case we are currently investigating these behaviors will consist of single atoms.

We will use two possible motion primitives: $u^{RandomSearch}$ which will lead a search agent to conduct random motion through the domain of interest, and $u^{IsoFollow}$, which will lead the robot to track the boundary of an element V_j^o with $j \geq 1$. The random search will be achieved through motions in straight lines followed by bouncing off the walls of the search domain. The isoline following control, on the other hand, is the same as the one described in our previous work [8].

There are two interrupt functions defined as:

$$\xi^{Found} = \begin{cases} 1 & S(\mathbf{r},t) < S_d \\ 0 & else \end{cases} , \quad (8)$$

which can interrupt a motion primitive, when a region where the potential exceeds the given threshold value is discovered, and the interrupt

$$\xi^{mapped} = \begin{cases} 1 & mapping\ a\ boundary \\ 0 & boundary\ mapped \end{cases} , \quad (9)$$

which interrupts the mapping of a boundary after its map is completed. (It is assumed that each robotic agent maintains a knowledge of the isoline contour it has followed, and it knows when is retracing previously explored points.)

Based on these interrupts, we can specify the two behaviors as:

$$a^{explore} = (\xi^{Found}, u^{RandomSearch}) \quad (10)$$

that is the robot randomly traverses the domain of interest until it reaches a region where the potential exceeds the threshold and

$$a^{exploit} = (\xi^{Mapped}, u^{IsoFollow}), \quad (11)$$

which makes the robot complete a map of a given isoline.

The strategy we will use is as follows. The robots are initialized with an exploration behavior. When $a^{Explore}$ is interrupted the robot automatically switches to $a^{Exploit}$ and maps the discovered region. Once the map is completed the ξ^{Mapped} interrupt triggers an update of the *subjective perspective*. Then, through (15) the robot can calculate the *information capacity* of its exploitation behavior and it can decide whether to continue the exploitation or begin exploring based on the criteria

$$a^{explore} \quad \text{if} \quad \frac{c_k}{\Delta T_k} < \theta \quad (12)$$

$$a^{exploit} \quad \text{else} \quad (13)$$

where ΔT_k is the time it has taken the robot to complete the map of the given structure and θ is a predefined threshold characterizing the search strategy. While we are assuming that isoline following is done with infinite precision, there will be some tendency for the agent to continue to exploit (i.e. retrace) the isoline due to the time variance of the field.

C. Results

All the simulations that follow are for the vehicles moving with constant speed of 1.5 units per unit time. The domain is a square with dimension 500 by 500 units and the radius of the cell $V_{j\geq1}^o$ is 20 units (see Fig. 1).

By varying the threshold θ, we can achieve different search strategies. There exist values for the threshold, $\theta \leq \theta_{min}$, for which once the robot finds a given structure, it keeps exploiting it indefinitely, that is the search strategy can be described by the infinite sequence $a^{Explore}a^{Exploit}a^{Exploit}a^{Exploit}\ldots$

On the other extreme, if $\theta \geq \theta_{max}$ the agent will continuously alternate between exploration, and exploitation and the sequence describing the search strategy is

$$a^{Explore}a^{Exploit}a^{Explore}a^{Exploit}\ldots \quad (14)$$

Given a static environment, once a structure is mapped there is no information that can be gained by repeated mapping, therefore independent of the chosen value for the threshold, θ, the search strategy is represented by (14). Fig. 2 shows the evolution of the normalized mutual information defined as $I_k/H(\mathcal{V}_k^o)$, under a search conducted by 10 vehicles on raster scan trajectories in case the potential field is static. Two pixel sizes were used and one can clearly observe the speed verses accuracy trade-off. (Smaller pixel size corresponds to more time to complete the search but better accuracy of the resulting map.)

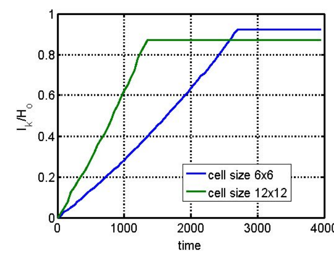

Fig. 2. The mutual information as function of time for two different cell sizes - 6×6 and 12×12, where the total size of the space is 500×500. The scans are compared for the time steps required to complete the 12 raster scan given that the vehicles move with speed 1.5units/per-time-step.

Fig. 3 compares the information acquisition for the exploration-exploitation strategy and the raster scan. It clearly establishes the raster scan as both slower and less accurate than the proposed search strategy.

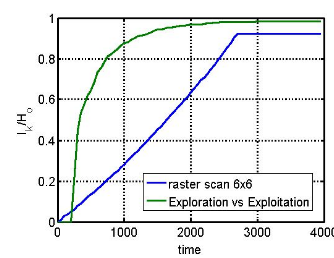

Fig. 3. The mutual information as function of time for search conducted by a raster scan and an exploration-exploitation decisions, given a static environment.

The information metric also illuminates the value of adding vehicles to the search team. Fig. 4 shows how the mutual information evolves under teams with different number of vehicles. It can be observed that after a certain point there is no significant benefit of assigning additional resources to the search process.

In the case of a dynamic environment, we can investigate the evolution of the mutual information under a range of threshold values θ. Fig. 5 shows the evolution of the information metric under $\theta < \theta_{min}$, $\theta_{min} < \theta < \theta_{max}$ and $\theta > \theta_{max}$. It can be observed that different boredom thresholds will lead to different steady state values of mutual

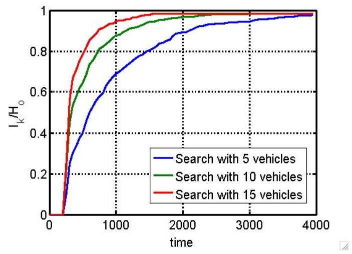

Fig. 4. The mutual information as function of time for search conducted by a motion primitives with different number of vehicles given a static environment

information. In fact, there will exist an optimal value for θ, which is established experimentally by Fig. 6.

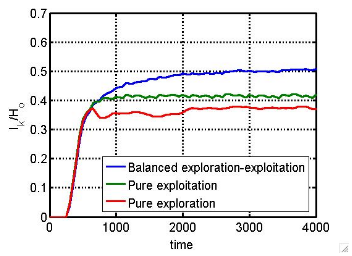

Fig. 5. The evolution of the mutual information under three different strategies - pure exploitation, pure exploration and balanced version between them, given dynamic environment.

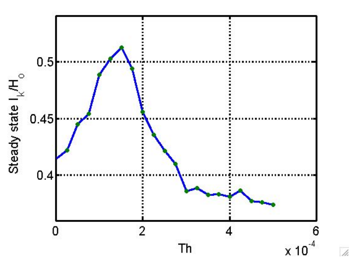

Fig. 6. The steady-state mutual information as function of the threshold.

V. Conclusions and Continuing Research

This paper has reported preliminary results on search strategies for teams of cooperating autonomous agents. Distinctive features of the strategies are: 1. they are information-seeking, and 2. they require on-the-fly decisions by each agent regarding wether to pursue depth (= an *exploitation* strategy) or breadth (= an *exploration* strategy) at each stage of the search. In order to discuss information-seeking search strategies, we have defined information metrics in Sections II and III. In Section IV, we have reported a crude comparison of raster-scan search with a hybrid breadth/depth search in which each agent employs a deterministic rule for switching between breadth and depth focused strategies. While the scope of this research on information-seeking search strategies is limited to the decision dynamics of teams of autonomous robots, an interesting extension to study the effects of human decision-making in terms of breadth versus depth could be pursued.

References

[1] A. L. Bertozzi, M. Kemp, and D. Marthaler, "Determining environmental boundaries: Asinchronous communiction and physical scales," in *Cooperative Control*, N. L. A. M. V. Kumar, Ed. Springer - Verlag Berlin Heidelberg, 2005, vol. LNCIS 309.

[2] W. Burgard, M. Moors, and F. Schneider, "Collaborative exploration of unknown environments with teams of mobile robots," in *Lecture Notes in Computer Science*, M. Beetz, M. G. J. Hertzberg, and M. Pollack, Eds. Springer Berlin / Heidelberg, 2002, vol. 2466 / 2002, pp. 52 – 70.

[3] N. E. Leonard, D. Paley, and F. Lekien, "Collective motion, sensor networks and ocean sampling," *Proceedings of the IEEE*, 2006.

[4] J. Adler, "Chemotaxis in bacteria," *Science*, vol. 153, no. 3737, pp. 708–716, August 1966.

[5] J. D. Cohen, S. M. McClure, and A. J. Yu, "Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration," *Philos Trans R Soc Lond B Biol Sci.*, vol. 362, pp. 933–942, May 2007.

[6] F. Zhang, E. Fiorelli, and N. E. Leonard, "Exploring scalar fields using multiple sensor platforms: Tracking level curves," in *Proceedings of the 46th IEEE Conference on Decision and Control*, December 2007, pp. 3579–3584.

[7] S. Susca, S. Martinez, and F. Bullo, "Monitoring environmental boundaries with a robotic sensor network," *ACC*, 2006.

[8] D. Baronov and J. Baillieul, "Reactvie exploration through following isolines in a potential field," in *Proceedings of the American Control Conference*, 2007, pp. 2141–2146.

[9] T. M. Cover and J. A. Thomas, *Elements of information theory*, 2nd, Ed. Hoboken, N.J.: Wiley-Interscience, 2006.

[10] V. Manikonda, P. S. Krishnaprasad, and J. Hendler, "Languages, behaviors, hybrid architechtures and motion control," in *Mathematical control theory*, J. Baillieul and J. C. Willems, Eds. New York, NY, USA: Springer-Verlag New York, Inc., 1998, pp. 200–226.

Appendix

Theorem 1: The information inflow defined by (4) satisfies

$$\Delta I^+k \leq H(\mathcal{V}_k^s, \mathcal{V}_{k-1}^s) - H(\mathcal{V}_{k-1}^s) := H(\mathcal{V}_k^s | \mathcal{V}_{k-1}^s). \quad (15)$$

The proof that follows is based on existing identities from the Shanon information theory. (See [9] for a reference of such identities.) The equivalence between these identities and the identities governing our definitions of mutual information and entropy will be establish elsewhere.

Proof: Substituting (2) into the definition of the information inflow, (4), yields

$$\Delta I_k^+ = H(\mathcal{V}_k^s) - H(\mathcal{V}_{k-1}^s) - H(\mathcal{V}_k^o, \mathcal{V}_k^s) + H(\mathcal{V}_k^o, \mathcal{V}_{k-1}^s). \quad (16)$$

On the other hand, the chain rule of entropies states:

$$H(\mathcal{V}_k^o, \mathcal{V}_k^s, \mathcal{V}_{k-1}^s) = H(\mathcal{V}_{k-1}^s | \mathcal{V}_k^s, \mathcal{V}_k^o) + H(\mathcal{V}_k^o, \mathcal{V}_k^s).$$

Through this equation, we can express $H(\mathcal{V}_k^o, \mathcal{V}_k^s)$ and substitute the resulting expression into (16) which yields

$$\Delta I_k^+ = H(\mathcal{V}_k^s) - H(\mathcal{V}_{k-1}^s) - \\ -\Big(H(\mathcal{V}_k^o, \mathcal{V}_k^s, \mathcal{V}_{k-1}^s) - H(\mathcal{V}_k^o, \mathcal{V}_{k-1}^s) - H(\mathcal{V}_{k-1}^s | \mathcal{V}_k^s, \mathcal{V}_k^o) \Big).$$

Now, we note that

$$H(\mathcal{V}_k^o, \mathcal{V}_k^s, \mathcal{V}_{k-1}^s) - H(\mathcal{V}_k^o, \mathcal{V}_{k-1}^s) \geq 0$$

and

$$H(\mathcal{V}_{k-1}^s | \mathcal{V}_k^s, \mathcal{V}_k^o) \leq H(\mathcal{V}_{k-1}^s | \mathcal{V}_k^s)$$

with both becoming equalities for $\mathcal{V}_k^o \equiv \mathcal{V}_k^s$. Then, it follows that:

$$\begin{aligned} \Delta I_k^+ &\leq H(\mathcal{V}_k^s) - H(\mathcal{V}_{k-1}^s) + H(\mathcal{V}_{k-1}^s | \mathcal{V}_k^s) = \quad (17) \\ &= H(\mathcal{V}_k^s, \mathcal{V}_{k-1}^s) - H(\mathcal{V}_{k-1}^s), \quad (18) \end{aligned}$$

which concludes the proof. ■