Towards Joint Human-Robotic Solutions to Surveillance Problems

Christopher Reardon, Jonathan Fink
U.S. Army Research Laboratory
2800 Powder Mill Road, Adelphi, MD 20783
{christopher.m.reardon3, jonathan.r.fink3}.civ@mail.mil

Abstract

Resource-constrained continuous surveillance tasks represent a promising domain for autonomous robotic systems. However, solutions to these problems must operate within a high-dimensional space and have an accurate model of the underlying target distribution to perform well. We present an integrated human, autonomous planning system that iteratively explores the space of surveillance solutions based on limited human interactions in an attempt to maximize both quantitative task-performance and qualitative operator satisfaction measures. This paper includes a description of our approach, implementation, and a demonstration that illustrates representative planning.

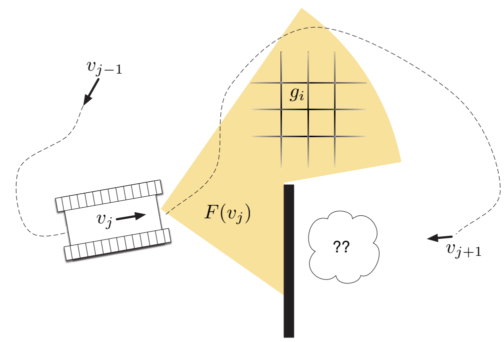


Figure 1: The task is to find a set of viewpoints v_j that maximize the expected target-detection rate based on sensor footprints $F(v_j)$ such that a path can be driven to visit all viewpoints within a cost budget B. A human interacts with the system by adjusting its prior belief on target locations, e.g., the cloud in this figure, to achieve information-gathering tours that are acceptable and high-performing.

Introduction

Mobile robotic systems offer significant advantages in long-duration search and surveillance tasks – continuous attention, precise location-awareness, and expendability in dangerous environments are but a few. While the use of autonomous robotic systems for these tasks seems inevitable, it remains an open question how these systems will interact with their human operators or teammates. The design of interactions between robots and humans has implications ranging from system performance to trust and acceptance.

In this paper, we consider a resource-constrained continuous surveillance task where a robot must traverse the environment to maximize its probability of detecting a target as depicted in Figure 1. We focus on the interaction between a human and the autonomous agent planning the observation route for the robot. Rather than assuming human interaction to be a one-shot effort, we consider interaction to be an iterative process between the human and autonomous system concluding when the human is satisfied with the solution.

Information-theoretic approaches to robotic map-exploration have received considerable attention as they provide mathematically well-founded information-gain functions that can be used for active control and planning (Charrow et al. 2015) and has recently been applied to target-detection and tracking problems (Charrow, Michael, and Kumar 2015).

The class of resource-constrained continuous surveillance tasks we are interested in have long been considered by the Operations Research (OR) community and referred to as the orienteering or selective traveling salesperson problems (Laporte and Martello 1990). In these problems, the task is to find a subset of locations such that the tour, i.e., solution to the traveling salesperson problem, maximizes reward for visiting each location while keeping the total solution cost under some budget. Recent work has further defined the correlated orienteering problem as an extension where the reward for visiting each location is correlated with the set of other locations visited, making the problem more amenable to planning informative tours in the context of persistent monitoring (Yu, Schwager, and Rus 2014). Efficient approximate algorithms have been shown to solve the correlated orienteering problem at speeds that make it reasonable to use in an online robotic setting (Arora and Scherer 2016).

While there have been many recent efforts addressing human interaction with autonomous planning systems (Reardon and Fink 2016), two are particularly related to the informative-tour-planning problem we are considering and offer contrasting approaches. One approach is to represent human interaction as a set of constraints within which the robot must plan for an optimally informative path (Yi,

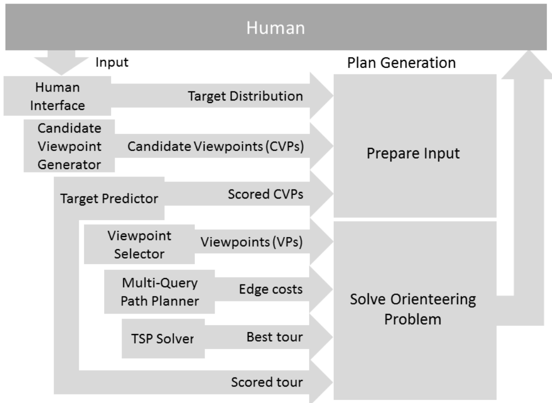


Figure 2: System overview. Progress flows downward, beginning with human input, as the orienteering problem solution is generated and presented back to the human operator.

Goodrich, and Seppi 2014) while the contrasting view is to have the human shape the information function used to make decisions (Lin and Goodrich 2010). Because correlated orienteering problems are already heavily constrained and additional constraints could easily lead to infeasibility, we adopt the concept of shaping the information function rather than applying human-derived constraints. We note that there is a strong connection between our work at that of reward shaping in the reinforcement learning community; in particular with the interactive approach described in (Raza, Johnston, and Williams 2015). However, while reward shaping is used to more quickly guide the system towards the discovery of solutions that maximize an underlying reward function, we are more interested in the case where the underlying reward function may not be entirely known to the system.

We propose an interactive human-autonomous agent planning system that solves the correlated orienteering problem with location rewards determined by an information-theoretic function to optimize the probability of target detection within the desired cost budget. Human interaction is manifested through the specification and updates to a prior belief on target locations. By iteratively examining the information-gathering tour and updating the prior target belief, the human operator is able to guide the system to realize plans that both perform well in terms of quantitative target-detection measures as well as qualitative operator-acceptance measures. This update process is achieved through minimal interactions that present a greatly reduced burden compared to, e.g., specifying a complete prior target belief. While our approach marries state-of-the-art methods for information-theoretic control (Charrow, Michael, and Kumar 2015) and correlated orienteering problems (Arora and Scherer 2016), the contribution of this work lies in the interaction between a human and the autonomous planning system.

Problem Statement

The statement of our resource-constrained continuous surveillance problem is as follows. Given a single mobile robot that is equipped with a visibility-based sensor, e.g., a camera or laser range-finder and a map of the two-dimensional environment, construct a cyclical tour of the environment so that the robot will optimally detect targets that appear in the environment. We assume that the robot has access to an uninformed prior on the probability of a target appearing at any point in the environment, i.e., the *target belief prior*. A human operator has access to a higher-fidelity belief distribution based on their experiences and cues (e.g., visual or auditory) in the environment, but it is intractable to fully specify this distribution for the robot.

We represent the *target belief prior* as an occupancy grid $\mathbf{g}$ which is a set of G cells $\{g_1, \ldots, g_G\}$ such that the probability of there being a target in cell i is $p(g_i = 1)$. We assume that the set of $\{g_i\}$ is independent.

A viewpoint that can be achieved by the robot to make an observation is $v_j \in SE(2)$ with $v_j = [x, y, \theta]$. The visibility-based sensor on the robot has a *sensor footprint* that can be found by raycasting on the map of the environment and is given by a set of cells $g_i \in F(v_j)$ as depicted in Figure 1.

Letting q_i^j be the measurement made of cell g_i from viewpoint v_j, we adopt the target detection model from (Charrow, Michael, and Kumar 2015),

$$\begin{aligned} p(q_i^j = 1 | g_i = 1) &= \gamma \qquad & p(q_i^j = 0 | g_i = 1) &= 1 - \gamma \\ p(q_i^j = 1 | g_i = 0) &= 0 \qquad & p(q_i^j = 0 | g_i = 0) &= 1. \end{aligned} \tag{1}$$

Note that this model assumes no false-positive measurements and a true-positive rate of γ.

We can compute the "reward" for visiting a single viewpoint $R(v_j)$ as the expected number of target detections,

$$R(v_j) = \sum_{g_i \in F(v_j)} p(g_i) \cdot p(q_i^j | g_i). \tag{2}$$

While the probability of a target in each cell is independent, the probability for a set of measurements given $\mathbf{g}$, e.g., $p(\mathbf{q}|\mathbf{g})$, is not. However, for a binary sensor with high true-positive rate γ, we can closely approximate by only considering the first observation of each cell g_i. So, for a set of observations $\mathbf{v} = \{\mathbf{v}_j\}$ we can write the reward $R(\mathbf{v})$ as

$$R(\mathbf{v}) = \sum_{g_i \in G_{\mathbf{v}}} p(g_i) \cdot p(q_i^j | g_i) \tag{3}$$

where $G_{\mathbf{v}} = \{F(v_1) \cup F(v_2) \cdots \cup F(v_j)\}$.

We are interested in cyclical information-gathering tours for continuous surveillance, i.e., the orienteering problem. A solution to this problem is a sequence of viewpoints $\mathbf{v} = [v_1, \ldots, v_N]$ where the cost of traversal between two viewpoints be $C(v_i, v_{i+1})$ so that the total cost for a route is $C(\mathbf{v}) = \sum_{i \in 1, \ldots, N, 1} C(v_i, v_{i+1})$. Given a set of possible viewpoints V, we can define the orienteering problem as

$$\begin{aligned} \underset{\mathbf{v} \subset V}{\operatorname{argmax}} \quad & R(\mathbf{v}) \\ \text{subject to} \quad & C(\mathbf{v}) \leq B \\ & \mathbf{v} = [v_s, \ldots]. \end{aligned} \tag{4}$$

The choice of cost budget B, controls the frequency of observations in our continuous surveillance setting. We also constrain the problem by pre-defining the starting viewpoint v_s to be the current pose of the robot.

From (3), it is clear that the sum of independent rewards $R(v_j)$ is an upper bound for the actual reward $R(\mathbf{v})$, i.e.,

$$\sum_{v_j \in \mathbf{v}} R(v_j) \geq R(\mathbf{v}), \tag{5}$$

thus forcing us to consider (4) as a correlated orienteering problem (Yu, Schwager, and Rus 2014).

Approach

Our approach revolves around the idea that the target belief prior $\mathbf{g}$ has an overwhelming impact on the solutions to (4) but in many applications it is unavailable or only partially defined. While machine learning could offer environment-awareness and a prediction of $\mathbf{g}$, the amount of training data required would be large and difficult to collect. Furthermore, transference of a learned model to a new environment-type remains an open question in the machine learning community. Instead, we propose that a human teammate iteratively construct a series of target belief priors $\mathbf{g}_k$ based on his or her knowledge of the environment and the previous solution $\mathbf{v}^{k-1}$ to the correlated orienteering problem, i.e., (4).

A high-level view of our human-integrated solution to the continuous surveillance problem is depicted in Figure 2. Given an uninformed target belief prior $\mathbf{g}^0$, we choose a set of candidate viewpoints and solve the correlated orienteering problem (4). This solution $\mathbf{v}^0$ is graphically presented to the human teammate as a set of viewpoints, the under-budget path that tours these viewpoints, and a visualization of the observed regions of the environment. The human then modifies the target belief prior to create $\mathbf{g}^1$, review the generated plan $\mathbf{v}^1$, and continue.

The set of candidate viewpoints $V = \{v_j = [x, y, \theta]\}$ are sampled from a regular grid over unoccupied space in the environment. The reward for each candidate viewpoint, $R(v_j)$, is scored individually based on target detection rate from the *Target Predictor* module in Figure 2 as defined in (2). Because the actual detection rates of novel targets are dependent on other viewpoints, these scores serve as an upper bound on the reward for visiting viewpoint v_j.

While the correlated information-gain based reward function is sub-modular, typically leading to efficient optimization solutions, the addition of a traveling budget constraint for the correlated orienteering problem makes this problem non-submodular (Arora and Scherer 2016). We draw heavily from (Arora and Scherer 2016), splitting the correlated orienteering problem into a combination of a Constraint Satisfaction and Traveling Salesperson problems to implement the so-called Random Orienteering (RO) algorithm with some modifications.

The RO algorithm addresses (4) by iteratively exploring subsets of candidate viewpoints $\mathbf{v} \subset V$, i.e., the Constraint Satisfaction Problem, and then checking to see if there exists a tour within the cost budget, i.e., the Traveling Salesperson Problem (TSP). We note that the construction of the edge weights for a TSP in a realistic robotics problem can be computationally expensive in it's own right and involves motion planning with respect to complicated environments and differential constraints. We address this problem by (a) doing a lazy evaluation of edge costs for only the subset of viewpoints being considered, (b) caching path queries, and (c) leveraging algorithms specifically designed for the "multi-query" setting, e.g., the Probabilistic Roadmap Method (Kavraki et al. 1996). This allows us to spend some precomputation effort based on the map of the environment to speed up later calculations of the cost to traverse from one viewpoint to another, $C(v_i, v_j)$. We assume costs between viewpoints to be symmetric such that $C(v_i, v_j) = C(v_j, v_i)$.

Our modification to RO is based on the observation that the rate of convergence to a solution of the RO algorithm is influenced by the initial chosen set of viewpoints. We overcome this by performing weighted sampling of candidate viewpoints to initially explore several disparate solutions with high reward upper bounds and continuing with the one that is under budget and maximizes reward. The reward $R(\mathbf{v})$ for a candidate viewpoint selection is scored based on (3) giving an accurate representation of the target detection rate for a set of viewpoints that may overlap in their coverage of the environment. The RO algorithm continues by randomly sampling candidate viewpoints $v_j \in V$ updating the active solution $\mathbf{v}$, and evaluating with respect to the current-best solution tour per (Arora and Scherer 2016).

Results

We are interested in studying the interaction between a human and autonomous planning system to solve the continuous-surveillance robotics task described above in terms of both task performance, i.e., target detection, and human satisfaction. While rigorous experimental results towards either of these goals is beyond the scope of this work, in the following section we seek to detail our implementation of the system described above and provide results demonstrating the types of interactions that can take place between a human and this system.

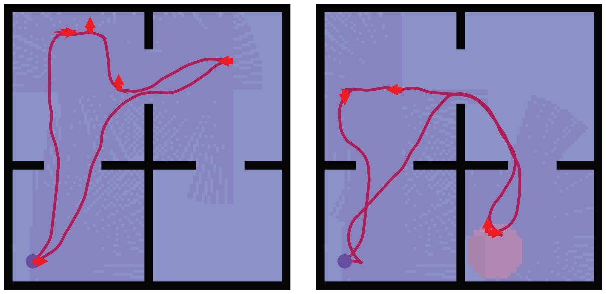
We implement the system depicted in Figure 2 with a suite of C++ and Python software modules, using ROS[1] to provide messaging, interprocess communication, and common robotics libraries. A multi-query path planner for computing edge costs to construct the TSP problem is implemented based on the Open Motion Planning Library[2] and we solve instances of the TSP problem with the Concorde library[3].

Visualization of the environment, current target belief prior $\mathbf{g}^k$, current tour solution, and observed regions of the environment is through the ROS RViz tool. Custom plugins for RViz allow the user to edit the target belief prior $\mathbf{g}$ by "painting" regions of higher or lower target probability with the mouse pointer. The autonomous system generates a new solution to the information-gathering tour after each user interaction produces a new prior $\mathbf{g}^k$. Note that in our examples to follow, the cost-budget parameter, B, is tuned by hand for

[1]http://ros.org

[2]http://ompl.kavrakilab.org

[3]http://www.math.uwaterloo.ca/tsp/concorde.html



(a) (b)

Figure 3: Example of the effect of human-specified target distribution information. (a) a route generated from a uniform target distribution probability, indicated by the blue background shading. (b) a route generated with a single human input (depicted by the red circle in the lower right room) to alter the target probability distribution. Purple circle indicates tour start/end, red arrows oriented viewpoints, red lines the path computed, and gray overlay sensor FOV.

each environment such that a tour whose observations fully cover the entire environment is not possible. In practice, this constraint would come from mission specifications based on the expected target arrival rate, desired update rate, or physical constraints such as platform fuel or battery capacity.

We present a simple, four-room environment as depicted in Figure 3. Figure 3a depicts an example tour solution generated from a uninformed uniform prior $\mathbf{g}^0$ with no human-provided knowledge of the target distribution. Because of the cost budget threshold, B, the map is not fully covered, but the two rooms nearest to the starting point, and part of the third, are well-covered. Figure 3b shows the same map after a single human interaction indicated a higher target belief probability (depicted in the figure and the user interface as a red circle in the bottom right room). As can be seen, the solution tour generated nearly completely covers the designated area of higher target belief. Additional, lower-cost viewpoints are found by the RO algorithm that are nearly co-linear with the path from the starting point to the high target probability area. Importantly, in this case the quantitative "reward" of the solution in Figure 3a is similar to Figure 3b with respect to the uniform $\mathbf{g}^0$, but Figure 3b qualitatively provides better coverage of the environment.

Discussion

In this paper, we describe a system whereby a human can work with an autonomous planning system to find solutions to a resource-constrained continuous surveillance task for a robot in complex environments with a realistic sensor model. This amounts to (1) an efficient solution to the correlated orienteering problem based on the most recent target belief prior which affects the reward surface for viewpoints in the environment and (2) an interface by which the human can analyze the current information-tour and modify the target belief prior to guide solutions to the surveillance task.

Our initial results, while qualitative, do provide a promising outlook on this style of interaction. In particular, we have shown in Figure 3 how a single human interaction can distinguish between two cost-constrained solutions that offer similar performance with respect to a uniform reward function but drastically different performance with respect to a "semantic" reward function, e.g., visiting all rooms.

Care has been taken in the development of this work to build a solution based on capabilities that are readily available in experimental robotic systems, e.g., occupancy grids from SLAM algorithms, sample-based motion planning, and probabilistic sensor models, so that this work can serve as a gateway to in-depth studies on human interaction with robotic systems capable of non-trivial autonomous tasks.

References

Arora, S., and Scherer, S. 2016. Rapidly Exploring Random Orienteering. http://www.frc.ri.cmu.edu/~sankalp/publications/rro.pdf.

Charrow, B.; Liu, S.; Kumar, V.; and Michael, N. 2015. Information-Theoretic Mapping Using Cauchy-Schwarz Quadratic Mutual Information. In *2015 IEEE International Conference on Robotics and Automation*, 4791–4798.

Charrow, B.; Michael, N.; and Kumar, V. 2015. Active control strategies for discovering and localizing devices with range-only sensors. In *Algorithmic Foundations of Robotics XI*. Springer. 55—-71.

Kavraki, L. E.; Svestka, P.; Latombe, J.-C.; and Overmars, M. H. 1996. Probabilistic roadmaps for path planning in high-dimensional configuration spaces. *IEEE transactions on Robotics and Automation* 12(4):566–580.

Laporte, G., and Martello, S. 1990. The selective travelling salesman problem. *Discrete Applied Mathematics* 26(2-3):193–207.

Lin, L., and Goodrich, M. A. 2010. A Bayesian approach to modeling lost person behaviors based on terrain features in Wilderness Search and Rescue. *Computational and Mathematical Organization Theory* 16(3):300–323.

Raza, S. A.; Johnston, B.; and Williams, M.-a. 2015. Reward from Demonstration in Interactive Reinforcement Learning. In *Twenty-Ninth International Florida Artificial Intelligence Research Society Conference*, 414–417.

Reardon, C., and Fink, J. 2016. Air-ground robot team surveillance of complex 3d environments. In *IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, 320–327. IEEE.

Yi, D.; Goodrich, M. A.; and Seppi, K. D. 2014. Informative path planning with a human path constraint. In *IEEE International Conference on Systems, Man and Cybernetics*, volume January, 1752–1758.

Yu, J.; Schwager, M.; and Rus, D. 2014. Correlated Orienteering Problem and its application to informative path planning for persistent monitoring tasks. In *IEEE International Conference on Intelligent Robots and Systems*, 342–349.