Optimizing Autonomous Surveillance Route Solutions from Minimal Human-Robot Interaction

Christopher Reardon[1], Fei Han[2], Hao Zhang[2], and Jonathan Fink[1]

***Abstract*— Resource-constrained surveillance tasks represent a promising domain for autonomous robotic systems in a variety of real-world applications. In particular, we consider tasks where the system must maximize the probability of detecting a target while traversing an environment subject to resource constraints that make full coverage infeasible. In order to perform well, accurate knowledge of the underlying distribution of the surveillance targets is essential for practical use, but this is typically not available to robots. To successfully address surveillance route planning in human-robot teams, the design and optimization of human-robot interaction is critical. Further, in human-robot teaming, the human often possesses essential knowledge of the mission, environment, or other agents. In this paper, we introduce a new approach named Human-robot Autonomous Route Planning (HARP) that explores the space of surveillance solutions to maximize task-performance using information provided through interactions with humans. Experimental results have shown that with minimal interaction, we can successfully leverage human knowledge to create more successful surveillance routes under resource constraints.**

I. Introduction

Autonomous search and surveillance by robotic systems is an area of active research with strong potential advantages in applications such as security, defense, or search and rescue. Successful realization of this research will result in systems capable of continuously monitoring a complex environment over long durations and precisely localizing themselves and their targets of observation, that are low-cost and expendable in dangerous situations. Because the successful deployment of these systems alongside humans depends largely upon the interaction between robot and human teammates, the design and optimization of human-robot interaction in a mixed team is critical. In such a teaming scenario, the human teammate often, if not always, has essential information (e.g., about the mission, prior knowledge of the environment, or the presence of other agents) that will affect the performance of the task, which a robot team member may not often possess. In fielded teams (e.g., for search and rescue, disaster monitoring, and other time and resource-limited situations), the capability to convey this information from the human to the robot in an expedient and efficient manner is thus highly valuable.

An important problem in robot-assisted search and surveillance is to autonomously plan surveillance route solutions where the robot must maximize its probability of detecting a target while traversing an environment. Solving this problem is particularly challenging in the context of planning surveillance routes that maximize target detection subject to limited resource constraints, e.g., time, energy, or effort, that are imposed by real-world applications. Recently, several methods were proposed to address parts of this problem. Information-theoretic methods were used to provide a mathematical basis for autonomously optimizing target-detection trajectories [1]. It is clear that resource-constrained surveillance task belong to the class of problems referred to as the selective traveling salesperson problem or orienteering problem [2]. Information gathering tasks break the assumption that reward for visiting each site is independent.

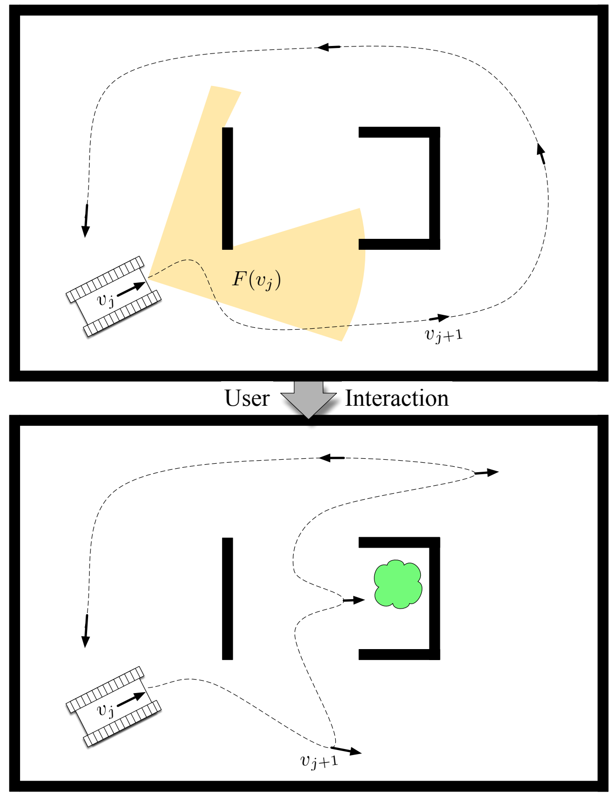


Fig. 1: The problem of resource-constrained surveillance is to find a set of viewpoints v_j that maximize the expected target-detection rate based on sensor footprints $F(v_j)$ such that a path can be driven to visit all viewpoints by a mobile robot within a cost budget B. The contribution focuses on the novel formulation and approach of solving this problem in a human-robot teaming scenario, in which a human interacts with the robotic system by adjusting its prior belief on target locations (e.g., the cloud) to achieve information-gathering tours that are high-performing.

[1]U.S. Army Research Laboratory, Adelphi, MD 20783 USA {christopher.m.reardon3,jonathan.r.fink3}.civ@mail.mil

[2]Department of Computer Science, Colorado School of Mines, Golden, CO 80401 USA {fhan,hzhang}@mines.edu

This motivates the correlated orienteering problem, along with candidate solution algorithms [3], [4]. However, these methods assume a fully-specified problem definition. We believe that with minimal interactions, we can leverage a human teammate's knowledge to dramatically improve the performance of autonomous route planning when addressing real-world tasks with uncertain or partial problem specifications.

We propose a novel approach, named *HARP* (standing for *H*uman-robot *A*utonomous *R*oute *P*lanning), to autonomously generate and optimize surveillance route solutions from minimal human-robot interactions under resource constraints, which addresses the correlated orienteering problem in a human-robot teaming scenario. Specifically, we present an interactive human-robot planning approach based on an information-theoretic function to automatically update viewpoint rewards from human input and optimize the target detection probability within the desired cost budget. A human teammate, with knowledge of the likely target locations, interacts with the robot by providing updates to a prior belief on the target locations. Through iterative examination of the surveillance tour in response to these updates to the target belief prior, the human teammate can guide the system to generate plans that are increasingly higher-performing. In an effort to minimize operator load while maximizing the speed to a solution, we propose methods for minimizing the number of human interactions. By autonomously generating route solutions with minimal human input, we can attain performance approaching a solution generated with complete situational knowledge.

The contribution of this paper is threefold:

- We propose a new problem of human-robot autonomous planning to construct resource-constrained surveillance route solutions in human-robot teaming scenarios, providing a novel formulation based upon constrained optimization that considers detection probability, resource constraints, and human-robot interaction.
- We introduce a new HARP method to solve the formulated constrained surveillance problem.
- We simulate extensive human-interactions to investigate the relationship between the number of interactions and quantitative route quality. We also conduct experiments with a physical robot in a realistic environment to demonstrate the effectiveness of HARP in autonomously generating resource-constrained surveillance routes.

The remainder of this paper is organized as follows. Section II provides a review of the related research. Our problem is formally defined in Section III. We detail our approach in Section IV. Experimental results are presented in Section V, before finally concluding with Section VI.

II. Related Work

In this section, we provide a review of the recent work on orienteering problems as well as the human interaction with autonomous planning systems.

A. Planning For Robotic Information Gathering

Information-theoretic methods for 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 [5], [6], [7]. Indeed, information theoretic-based control has also recently been applied to target-detection and tracking problems [8], [1].

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 problem (OP) or selective traveling salesperson problem [2]. 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. Being an NP-hard problem, most algorithms addressing the OP rely on approximations; the development of practical solution algorithms continues to be an active area of research [9], [10]. Recent work, however, has leveraged modern solutions to the OP to support solutions of information-gathering routes for hybrid aerial-ground systems [11].

When employing state-of-the-art information-gain metrics as the *reward* in an OP context, it becomes clear that in most practical scenarios, the rewards for visiting sites are not independent. This has led recent work to further define 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 [3]. 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 [4].

While there do exist promising solutions to address orienteering problems for realistic applications, these methods assume a fully specified and correct problem definition.

B. Human Interaction with Autonomous Planner

While several recent studies address human interaction with autonomous planning systems [12], [13], [14], two contrasting categories of methods are particularly related to the informative tour planning problem in the context of constrained orienteering we are considering. The first category of approaches represents human interaction as a collection of constraints within which the robot must plan for an optimally informative path [15]. The other category of approaches allows the human to shape the information function that is applied for decision making [16]. In this work, we adopt the strategy of shaping the information function rather than applying human-derived constraints due to the fact that correlated orienteering problems are already heavily constrained and additional constraints could easily lead to infeasibility, which has been studied experimentally in our early work [17]. Different from previous work, this paper introduces a new problem formulation of the human-in-the-loop constrained surveillance planning problem in human-robot teaming scenarios, and proposes an optimization-based solution that can jointly model detection probability, resource constrains, and human-robot interaction.

In addition, we notice that there is a potential connection between our work at that of reward shaping in the reinforcement learning community, in particular the interactive approach described in [18], [19], which however reshapes the reward function instead of the information distribution. While reward shaping can be applied to more quickly guide the system towards the discovery of solutions that maximize an underlying reward function, the methods based on reward shaping cannot be directly applied in our problem domain where the underlying reward function may not be entirely known to the system.

III. Problem Formulation

Given a mobile robot that is equipped with a visibility-based sensor (e.g., a camera or laser range-finder) and a map of an environment, the problem to *plan resource-constrained surveillance routes* is defined as the construction of 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 a 2D environment, i.e., the target belief prior. A human team member 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.

Mathematically, we represent the target belief prior as an occupancy grid $\mathbf{g}$ consisting of a set of G independent cells $\{g_1, \ldots, g_G\}$ such that the probability of there being a target in cell i is $p(g_i = 1)$, and denote a viewpoint as $v_j \in SE(2)$ with $v_j = \left[x, y, \theta\right]$, which is obtained by the robot to make an observation. 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 Fig. 1. Let q_i^j be the measurement made of cell g_i from viewpoint v_j, then the target detection model [1] can be denoted as:

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

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 the occupancy grid (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. This means that 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 $G_{\mathbf{v}} = \{F(v_1) \cup F(v_2) \cdots \cup F(v_j)\}$.

Then, the resource-constrained surveillance route planning can be defined as an optimization problem to find a sequence of viewpoints $\mathbf{v} = \left[v_1, \ldots, v_N\right]$, where the cost of traversal between two viewpoints be $C(v_i, v_{i+1}) > 0$ 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 $\mathcal{V}$, we can solve this problem[1] by

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

The choice of the 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.

IV. The HARP Approach

Our approach, *H*uman-robot *A*utonomous *R*oute *P*lanning, or *HARP*, begins with the observation that the belief prior for the target, $\mathbf{g}$, is fundamental to solving Eq. (4); however, we wish to address situations where this information is not completely available to a robotic system. While one approach could be to leverage machine learning to predict $\mathbf{g}$, the amount of training data required would be cumbersome, and generalizability of any learned model remains an open question. Instead, we choose to treat this circumstance as one where a human teammate possesses knowledge sufficient to provide $\mathbf{g}$, but a mechanism for fully defining the target belief prior would be intractable.

Therefore, we propose to examine the problem as formulated in Eq. (4) using iterative interactions between the human and the robot, where in interaction round k, the human teammate updates the series of target belief priors $\mathbf{g}^k$ using knowledge of the target distribution and the previous solution $\mathbf{v}^{k-1}$ to the correlated orienteering problem. We introduce a modification of Eq. (4) that simultaneously maximizes the reward and minimizes the cost in the objective function to reduce the amount of budget used to achieve the reward. Then, the problem in Eq. (4) can be rewritten as:

$$\begin{aligned} \underset{\mathbf{v} \subset \mathcal{V}}{\operatorname{argmax}} \quad & R(\mathbf{v}) - \lambda C(\mathbf{v}) \\ \text{subject to} \quad & C(\mathbf{v}) \leq B \\ & \mathbf{v} = \left[v_s, \ldots\right]. \end{aligned} \tag{5}$$

where λ is a trade-off parameter controling the cost effect.

The formulation in Eq. (5) is particularly useful in iterative, online applications. Without this formulation, when solution $\mathbf{v}^k$ is generated using $\mathbf{v}^{k-1}$ as input to Algorithm 1, the budget B would likely be expended, and generation of $\mathbf{v}^k$ would be overly dependent upon delete events (line 10), the frequency of which is the size of the current solution $|v^*|$ relative to the entire viewpoint space $|V|$. Instead, by introducing the trade-off parameter λ, we allow our the initial

[1] From Eq. (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., $\sum_{v_j \in \mathbf{v}} R(v_j) \geq R(\mathbf{v})$, thus Eq. (4) can be considered as a correlated orienteering problem [3].

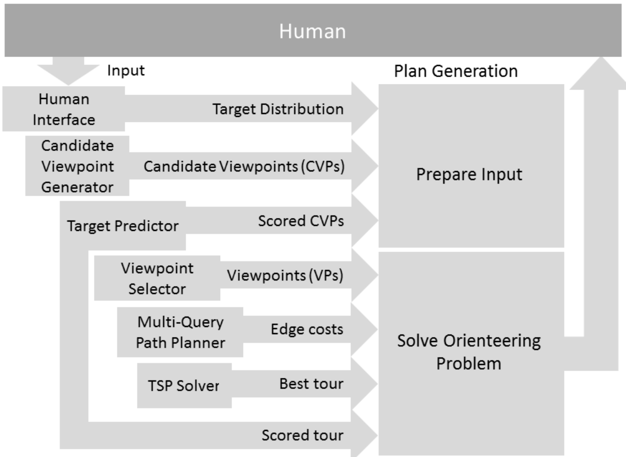


Fig. 2: HARP overview. Progress flows downward, beginning with human input and generated solutions will be presented back to the human teammate.

solution $\mathbf{v}^0$ to have unexpended budget so that we can find solutions sufficiently below budget B to allow for variability in execution.

The overall information flow in the HARP solution to the resource-constrained surveillance problem is depicted in Fig. 2. First, given an uninformed (e.g., uniform) target belief prior $\mathbf{g}^0$, a set of candidate viewpoints is selected, and a solution to the correlated orienteering problem Eq. (4) is constructed. This initial solution $\mathbf{v}^0$ is presented graphically as an under-budget surveillance tour of the environment. Then, the human is able to modify the target belief prior with a single interaction to create $\mathbf{g}^1$. This new $\mathbf{g}^1$ represents a better-informed prior, and a solution to the correlated orienteering problem is generated using this new information and $\mathbf{v}^0$. The human teammate can then view and provide successive interactions to shape the surveillance route generated.

As graphically illustrated in Fig. 2, the set of candidate viewpoints $V = \{v_j = [x, y, \theta]\}$ are generated by sampling over unoccupied space within the environment. Then, using the *Target Predictor* module, each candidate viewpoint's reward $R(v_j)$ is scored individually based upon the target belief as defined in Eq. (2). Because the actual detection rates of novel targets are dependent on observations from other viewpoints, these scores serve as an upper bound on the reward for visiting each viewpoint v_j.

While the correlated information-gain based reward function is sub-modular and therefore efficient optimization solutions can be devised, with the addition of a traveling budget constraint, this problem becomes non-submodular [4]. To address this issue, after the *Prepare Input* step is complete, we split the correlated orienteering problem into a combination of Constraint Satisfaction and Traveling Salesperson problems. Then, we implement a new approach by adapting and modifying the Random Orienteering (RO) algorithm [4], as shown in Algorithm 1.

In the new Algorithm 1, we address Eq. (5) by iteratively

Algorithm 1: Modified RO based on [4] for solving decoupled constraint satisfaction and TSPs

Input : $G = [V, E], v_s, v_{k-1}, B, m$
Output: The best route found in $3|V|$ steps

if $v_{k-1} == Null$ `// First iteration`
then
  $\mathbf{v}, C(\mathbf{v}) = SampleRandomSolutions(V, v_s, m)$
  $\mathbf{v}^* = \mathbf{v}$ `// Init weighted random solution`
else
  $\mathbf{v}^* = v_{k-1}$ `// Init previous solution`
end
for $i = 1 : 3|V|$ **do**
  $v_{new} = Sample(V)$ `// Sample a view`
  if $IsInRoute(\mathbf{v}, v_{new})$ **then**
    $\mathbf{v} = DeleteFromRoute(\mathbf{v}, v_{new})$
  else
    $\mathbf{v} = AddToRoute(\mathbf{v}, v_{new})$
  end
  $\mathbf{v}, C(\mathbf{v}) = TSP(\mathbf{v})$ `// TSP returns ordered v and cost`
  if $C(\mathbf{v}) <= B \wedge R(\mathbf{v}) > R(\mathbf{v}^*)$ **then**
    $\mathbf{v}^* = \mathbf{v}$
  end
end
$\mathbf{v}^* = GreedyLocalSearch(V, \mathbf{v}^*, B)$
return $\mathbf{v}^*$

Algorithm 2: Sample weighted random solutions

Input : V, v_s, m
Output: A selection of vertices weighted by reward

$\mathbf{v}^* = \{\}$
for $i = 1 : m$ **do**
  $v = v_s \cup WeightedRandomSample(V)$ `// select viewpoints`
  $\mathbf{v}, C(\mathbf{v}) = TSP(\mathbf{v})$
  if $C(\mathbf{v}) <= B \wedge R(\mathbf{v}) > R(\mathbf{v}^*)$ **then**
    $\mathbf{v}^* = \mathbf{v}$
  end
end
return $\mathbf{v}^*$

exploring subsets of candidate viewpoints $\mathbf{v} \subset V$, i.e., the Constraint Satisfaction Problem, and then checking for a tour within the cost budget B, which is the Traveling Salesperson Problem (TSP), using a small cost-effect value for λ. We note that the construction of the edge weights for a TSP in a realistic robotics application can be computationally expensive in its own right and involves motion planning with respect to complicated environments and differential constraints. Thus, we address this problem in three ways: (1) evaluating edge costs for only the subset of viewpoints being considered, (2) caching path queries, and (3) leveraging algorithms shown successful in the "multi-query" setting, e.g., the probabilistic roadmap method [20]. In this way, we spend some precomputation effort to speed up later calculations of the cost to traverse from one viewpoint to another, $C(v_i, v_j)$. We assume the costs between viewpoints to be symmetric such that $C(v_i, v_j) = C(v_j, v_i)$.

One challenge of implementing and applying Algorithm

Algorithm 3: Modified greedy local search

Input : $V, \mathbf{v}^*, B$
Output: $\mathbf{v}^*$ with greedily-selected neighbors

1 $v_{eligible} = \{\}$
2 **for** $i = 0 : |V|,\ j = 0 : |\mathbf{v}^*|$ **do**
3   **if** $Distance(v_i, v_j) < \delta \wedge v_i \notin \mathbf{v}^*$ **then**
4         $v_{eligible} = v_{eligible} \cup v_i$
5   **end**
6 **end**
7 **for** $e = 0 : |v_{eligible}|$ **do**
8   **if** $R(\mathbf{v}^* \cup v_e) > R(\mathbf{v}^*) + c \wedge C(\mathbf{v}^* \cup v_e) < B$ **then**
9         $\mathbf{v}^* = \mathbf{v}^* \cup v_e$
10   **end**
11 **end**
12 **return** $\mathbf{v}^*$

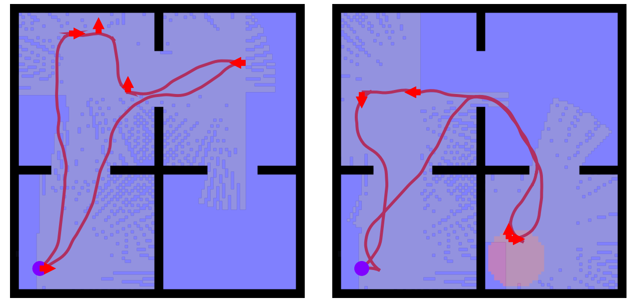
1 is that the rate of convergence to a solution is influenced by the initial chosen set of viewpoints. We overcome this by introducing an initialization method, shown in Algorithm 2, which performs sampling of m candidate viewpoints, weighted by reward, to initially explore several disparate solutions with high reward upper bounds and continuing with the one that is under budget and maximizes reward. We score the actual reward $R(\mathbf{v})$ for a candidate viewpoint selection based upon Eq. (3), which accounts for coverage overlapping between viewpoints and provides an accurate representation of the target detection rate.

Algorithm 1 continues by randomly sampling candidate viewpoints $v_j \in V$ (line 9), updating the active solution $\mathbf{v}$ (lines 10-15), and evaluating with respect to the current-best solution tour $\mathbf{v}^*$ (line 16). If the viewpoint selected is currently in the tour $\mathbf{v}$, it is removed (line 11); if it is not in the tour, then it is added (line 13). If the cost of $\mathbf{v}$ is under budget and it improves the reward over $\mathbf{v}^*$, it is kept as the current best tour (line 17).

After $3|V|$ iterations (per the standard probabilistic constraint satisfaction problem algorithm from [21] modified in [4]), a modified version of the Greedy Local Search from [4] is performed (Algorithm 3, where δ is a distance threshold and c is a reward threshold) to improve anytime performance by incorporating nodes in the neighborhood of the chosen route that increase the reward over a threshold value, c, while remaining under cost. First, a list of eligible candidates within a distance threshold δ of existing tour viewpoints is constructed (lines 1-6). Then, if the addition of any of those candidates increases the tour reward $R(\mathbf{v}^*)$ while being under budget B (lines 7-11), the viewpoint is added to the tour. Finally, after the greedy local search is complete, the $\mathbf{v}^*$ tour is returned.

V. Experimental Results

We evaluate the performance of the HARP solution with two sets of experiments in both simulated and real-world environments to address the resource-constrained autonomous surveillance route problem. We compare the performance of HARP with varying numbers of human interactions against a baseline method in 210 simulated surveillance trials, and demonstrate the HARP solution in 36 surveillance trials in a real-world environment with a physical robot.



(a) (b)

Fig. 3: Illustration of the effect of the human-specified target distribution information. (a) shows a route generated from a uniform target distribution probability, indicated by the blue background shading. (b) shows a route generated with a single human input (depicted by the red circle in the lower right room) to alter the target probability distribution. The purple circle indicates tour start/end. Red arrows represent oriented viewpoints and red lines represent the path computed for the tour. The gray overlay indicates the sensor fields of view.

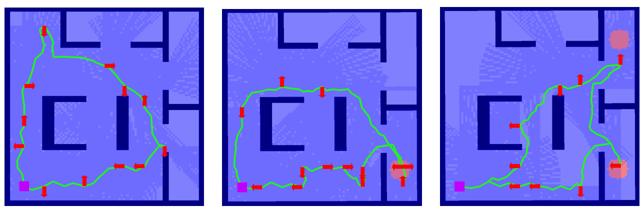
A. HARP Implementation

Our HARP system is implemented as a suite of C++ and Python software modules, leveraging ROS [22] for messaging, interprocess communication, and common robotics libraries. In addition, we implement a multi-query path planner based on the Open Motion Planning Library (OMPL) [23] for computing edge costs necessary for TSP problem formulation, and we adopt the Concorde library [24] to implement solutions to TSP instances.

To visualize the environment, as well as the current target belief prior $\mathbf{g}^k$, current tour solution, and regions observable by the robot, we leverage the ROS RViz tool. The user is able to edit the target belief prior $\mathbf{g}$ using custom plugins we have developed for RViz by "painting" regions of higher target probability with the mouse pointer. After each such interaction, the HARP system regenerates a new surveillance tour solution v_k using the new prior $\mathbf{g}^k$. In the experiments, the budget B and the small cost-effect parameter λ are tuned by hand for each environment to produce a resource-constrained scenario where a tour providing full observation coverage of the environment is not possible. This is necessary to examine the resource-constrained surveillance problem; in practice, this resource constraint would come from a higher-level mission specification, such as expected target arrival rate, desired surveillance update rate, or physical energy constraints of the robotic system.

B. Automated Surveillance in Simulation

To illustrate the effect of a human interaction, we perform a simple experiment in an orthogonal four-room environment as demonstrated in Fig. 3. Fig. 3a shows a tour solution autonomously generated from a uninformed uniform prior $\mathbf{g}^0$



(a) (b) (c)

Fig. 4: Example of the effect of interaction on route surveillance generation in a complex environment. (a) shows the baseline case, where a route is generated autonomously with a uniform target distribution. (b) shows the impact of a single interaction, where a route after the target distribution prior is modified to elevate an area of higher target probability, shown as the red circle. (c) shows the autonomous route generated after two interactions.

with no human-provided knowledge of the target distribution, which we consider a baseline. We constrain the resources available (i.e., the budget, B) so that full coverage of the map is not possible, to examine the impact that can result from an interaction in our system. As a result, in Fig. 3a the map is not fully covered, but the two rooms nearest to the starting point, and part of the third, are well-covered. Fig. 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). The informed solution tour covers the designated area of higher target belief. Additional, lower-cost viewpoints are found by HARP that are near the path and within budget. Importantly, in this case the quantitative "reward" of the solution in Fig. 3a is similar to Fig. 3b with respect to the uniform $\mathbf{g}^0$, but Fig. 3b qualitatively provides better coverage of the environment.

For the next experiment, we construct autonomous surveillance routes in a more realistic environment, as illustrated in Fig. 4. In the baseline case shown in Fig. 4a, a circular route generated from a uniform prior target distribution $\mathbf{g}_0$ achieves good coverage of the environment, subject to the budget B. Fig. 4b shows the route generated from a single interaction $\mathbf{g}_1$ that places a higher probability of target presence in the lower righthand room. The new route sacrifices coverage of the upper left area to fully cover the lower right room, a decision that was made without explicit instruction by the human teammate. In Fig. 4c, we see the increasing shaping effect the human's interaction has, as a second area of higher target presence likelihood is placed in the upper right room for $\mathbf{g}_2$. Now, because of the constrained budget, the surveillance route skews strongly to cover the higher likelihood areas at the expense of other areas. These results illustrate the power of this novel solution. Without directly instructing the robot or even complete communication of target distribution, the human teammate is able to guide the generation of an autonomous surveillance route.

We examine the performance increase for such human-to-robot interaction versus a baseline method with no interaction by measuring the target detection rate over a large sample set for our complex simulated environment. This experiment evaluates the contribution of such interactions to resource-constrained surveillance route planning (formulated in Eq. 5). We also show several budget B values to illustrate the impact of resource availability. The idea behind these experiments is to evaluate the effect of interaction on target detection performance. Rather than conduct a large-scale user study, we simulate human input. Given a hidden underlying ground-truth target distribution, our human-simulator generates a set of interactions ordered based on the size and probability value of each area of elevated probability in the ground truth, then communicates each interaction via the same mechanism as the human interface (i.e., ROS topics).

Fig. 5 shows box plots of 210 experiments, 10 per plot, with varying budget values. Each box represents a number of interactions, from baseline (no interactions, shown in red) to n interactions, where n is the number of elevated areas of target presence. Figs. 5d and 5b show the ground-truth target distribution probabilities for each set of experiments. These results show that, compared to baseline, routes generated from interactions in general reduce the variance of the target detection performance, and increase overall target detection. It is interesting to note that in some cases there is a point of diminishing returns with respect to interactions. In situations with sufficient budget B, it may not be necessary to provide a full target belief prior to the robot, as most autonomous surveillance routes will provide sufficient coverage from only a partially updated belief prior. See, for example, Fig. 5a for $B = 45$ between Interactions 2 and 3.

C. *Real-world Experiments for Urban Surveillance*

In these experiments, we evaluate our HARP method in an automated surveillance application in a real-world setting. A Clearpath Robotics Jackal robot (Fig. 6a) was used for these experiments, equipped with a Velodyne VLP-16 LiDAR and an ASUS Xtion camera.

Our experiments take place in a realistic field environment consisting of multiple concrete buildings and a street arranged and staged as a cluttered village marketplace (Fig. 6b). To simulate targets, 10 AprilTags[2] were placed throughout the environment by a researcher not otherwise involved in the experiment. To demonstrate the robust execution of our feedback system, as in Sec. V-B, we attempted baseline tours (0 interactions), and tours after 1 and 2 interactions, 12 each for 36 surveillance tours total. Of these, 35 tours completed successfully; one set was aborted during the $n = 1$ interactions run when the robot failed.

For these experiments, we constructed tours to stay within 20m of the middle of the marketplace, and used a distance budget B of 150m, with initial viewpoint count $V = 87$ and $\lambda = 35$. All tours began at the same location. The HARP system generated a tour route, under budget B, after an interaction (initially no interaction), and the robot conducted

[2]https://april.eecs.umich.edu/software/apriltag.html

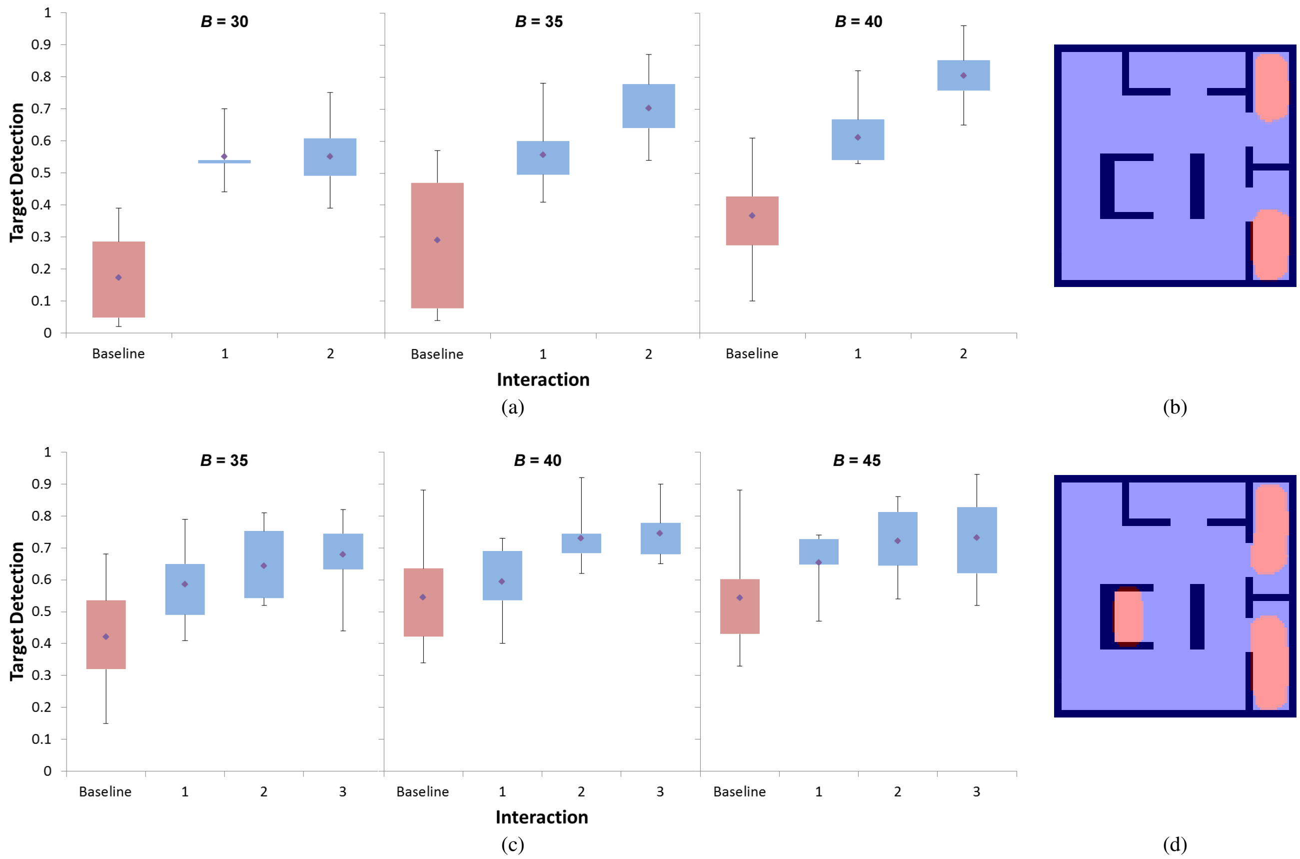


Fig. 5: Results comparing baseline interaction (red box) to different numbers of human interactions (blue boxes) across budget parameters. (a) shows target detection performance where there are two areas of high target likelihood shown in the map in (b), with budget $B = (30, 35, 40)$. (c) shows performance for three areas of high target likelihood shown in the map in (d), with $B = (35, 40, 45)$. The top and bottom of the boxes represent the interquartile range (middle 50% of samples), the purple diamond is the mean, and the whiskers represent the overall value range.

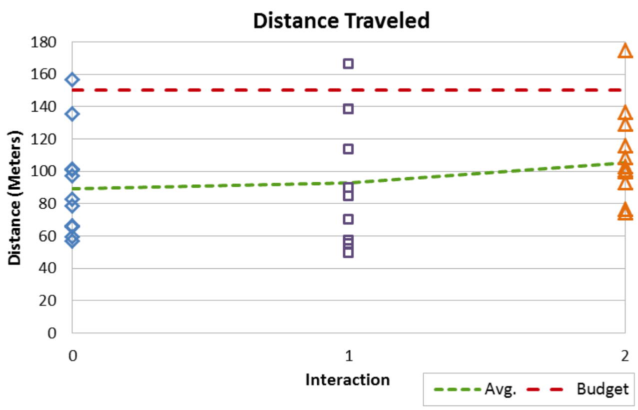
Fig. 6: Mobile robot (a) and realistic environment (b) used in the real-world experiments.

autonomous navigation using a kinematically feasible motion planner. Tour generation took approximately 5-15 seconds on a computer with an Intel Core i7 2.90GHz Quad Core mobile processor.

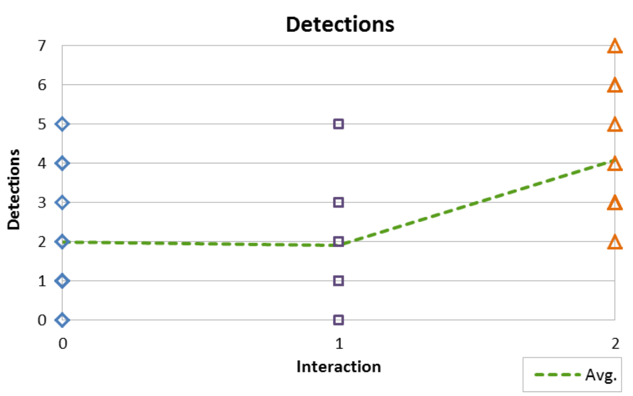
The actual distance traveled and the number of unique AprilTag detections were calculated for each complete tour, and the results are shown in Fig. 7. We observe that while there is a high amount of randomness in the executed path length, probably as a result of both the stochastic nature of our solution and kinematic planning over a real world environment, all but two (i.e., 94%) of the planned paths when executed stayed within the budget B (Fig. 7a), giving validity to our use of the trade-off parameter λ. We plan on examining the distribution of the randomness in surveillance tour length and its impact on performance in future work. There was also noise in the target detection, which may be due to variability in the orientations achieved by the kinematic motion planner; however, we also observe an increase in the mean number of target detections between 1 and 2 interactions (Fig. 7b), and we believe this outcome can be further refined in future experiments.

VI. Conclusion

In this paper, we present a novel Human-robot Autonomous Route Planning (HARP) method for optimizing autonomous surveillance route performance for robots using minimal human interaction in resource-constrained surveillance applications, i.e., where full coverage of the environment is not possible. The HARP approach uses a target belief prior to autonomously construct a surveillance tour for a robotic system that maximizes reward based on that prior. A human teammate, with knowledge of more accurate



(a)



(b)

Fig. 7: Experimental results from 34 real-world experiments. (a) shows distance traveled and (b) shows target detections out of 10 possible targets.

target distributions, interacts concisely with the robot by providing updates to the belief prior to shape the generation of surveillance tours to optimize detection performance. To evaluate the performance of the HARP method, experiments using both simulations and a real robot are performed for a surveillance task in an urban environment. Our results validate that interactions outperform a baseline case, and with successive minimal interactions, target detection performance increases.

REFERENCES

[1] B. Charrow, N. Michael, and V. Kumar, "Active control strategies for discovering and localizing devices with range-only sensors," in *Algorithmic Foundations of Robotics XI*, 2015.

[2] G. Laporte and S. Martello, "The selective travelling salesman problem," *Discrete Applied Mathematics*, vol. 26, no. 2-3, pp. 193–207, 1990.

[3] J. Yu, M. Schwager, and D. Rus, "Correlated orienteering problem and its application to informative path planning for persistent monitoring tasks," in *IEEE International Conference on Intelligent Robots and Systems*, 2014.

[4] S. Arora and S. Scherer, "Rapidly exploring random orienteering," 2016. [Online]. Available: http://www.frc.ri.cmu.edu/{~}sankalp/publications/rro.pdf

[5] B. Charrow, S. Liu, V. Kumar, and N. Michael, "Information-theoretic mapping using Cauchy-Schwarz quadratic mutual information," in *IEEE International Conference on Robotics and Automation*, 2015.

[6] T. Kollar and N. Roy, "Efficient optimization of information-theoretic exploration in SLAM," in *AAAI Conference on Artificial Intelligence*, 2008.

[7] B. J. Julian, S. Karaman, and D. Rus, "On mutual information-based control of range sensing robots for mapping applications," in *IEEE International Conference on Intelligent Robots and Systems*, 2013.

[8] P. Dames, P. Tokekar, and V. Kumar, "Detecting, localizing, and tracking an unknown number of moving targets using a team of mobile robots," in *International Symposium of Robotics Research*, 2015.

[9] A. Blum, S. Chawla, D. R. Karger, T. Lane, A. Meyerson, and M. Minkoff, "Approximation Algorithms for Orienteering and Discounted-Reward TSP," *SIAM Journal on Computing*, vol. 37, no. 2, pp. 653–670, 2007.

[10] P. Vansteenwegen, W. Souffriau, and D. Van Oudheusden, "The orienteering problem: A survey," *European Journal of Operational Research*, vol. 209, no. 1, pp. 1–10, 2011.

[11] P. Tokekar, J. Vander Hook, D. Mulla, and V. Isler, "Sensor planning for a symbiotic UAV and UGV system for precision agriculture," *IEEE Transactions on Robotics*, vol. 32, no. 6, pp. 1498–1511, 2016.

[12] J. Crossman, R. Marinier, and E. B. Olson, "A hands-off, multi-robot display for communicating situation awareness to operators," in *IEEE International Conference on Collaboration Technologies and Systems*, 2012, pp. 109–116.

[13] S. Dawson, C. Crawford, E. Dillon, and M. Anderson, "Affecting operator trust in intelligent multirobot surveillance systems," in *IEEE International Conference on Robotics and Automation*, 2015, pp. 3298–3304.

[14] J. Alonso-Mora, S. H. Lohaus, P. Leemann, R. Siegwart, and P. Beardsley, "Gesture based human-multi-robot swarm interaction and its application to an interactive display," in *IEEE International Conference on Robotics and Automation*, 2015, pp. 5948–5953.

[15] D. Yi, M. A. Goodrich, and K. D. Seppi, "Informative path planning with a human path constraint," in *IEEE International Conference on Systems, Man and Cybernetics*, 2014.

[16] L. Lin and M. A. Goodrich, "A Bayesian approach to modeling lost person behaviors based on terrain features in wilderness search and rescue," *Computational and Mathematical Organization Theory*, vol. 16, no. 3, pp. 300–323, 2010.

[17] C. Reardon and J. Fink, "Towards joint human-robot solutions to surveillance problems," in *International Florida Artificial Intelligence Research Society Conference*, 2017.

[18] S. A. Raza, B. Johnston, and M.-a. Williams, "Reward from demonstration in interactive reinforcement learning," in *International Florida Artificial Intelligence Research Society Conference*, 2015.

[19] K. Judah, A. P. Fern, P. Tadepalli, and R. Goetschalckx, "Imitation learning with demonstrations and shaping rewards." in *AAAI*, 2014.

[20] L. E. Kavraki, P. Svestka, J.-C. Latombe, and M. H. Overmars, "Probabilistic roadmaps for path planning in high-dimensional configuration spaces," *IEEE Transactions on Robotics and Automation*, vol. 12, no. 4, pp. 566–580, 1996.

[21] T. Schoning, "A probabilistic algorithm for k-sat and constraint satisfaction problems," in *Foundations of Computer Science, 1999. 40th Annual Symposium on*. IEEE, 1999, pp. 410–414.

[22] M. Quigley, K. Conley, B. P. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, "ROS: an open-source robot operating system," in *Workshop of International Conference on Robotics and Automation*, 2009.

[23] I. A. Şucan, M. Moll, and L. E. Kavraki, "The Open Motion Planning Library," *IEEE Robotics & Automation Magazine*, vol. 19, no. 4, pp. 72–82, December 2012.

[24] D. L. Applegate, R. E. Bixby, V. Chvatal, and W. J. Cook, *The traveling salesman problem: a computational study*. Princeton University Press, 2011.