

Information based control of robots in search and rescue missions with human prior knowledge

Rafal Krzysiak, Sachit Butail

Abstract—Multi-robot systems provide a scalable and robust solution for monitoring tasks. In time-intensive missions such as search and rescue, the inclusion of a human has the potential advantage of incorporating prior knowledge about the target location or dynamics. In this paper we develop a general information-theoretic framework to control multiple autonomous robots in search and rescue missions that include a human teleoperator. Human prior knowledge is modeled to capture target location and dynamics, and mutual information based control is formulated to let autonomous robots weight between two strategies: independent search or assisting the human by staying in proximity. The control actions optimize a weighted sum of normalized mutual information calculated using particle filtered estimates of the target and the reference robot. We implement the framework to simulate two widely different scenarios designed after search and rescue missions from literature and incorporate varying levels of accuracy in human prior knowledge. Our results indicate that mission performance depends on how robots weight between the two strategies, with the amount of the optimal control effort shared between strategies affected by prior knowledge and number of robots. Comparison with existing strategies point to benefits of an information-based control in situations where human prior knowledge is inaccurate. The proposed information-theoretic abstraction of human robot interaction can be implemented on a wide variety of scenarios and the results highlight the role of human prior knowledge towards effective robotic assistance in time-intensive missions.

Index Terms—mutual information, human robot interaction, particle filtering, search and rescue, teleoperation.

I. INTRODUCTION

Multi-robot systems are particularly well suited for missions that involve spanning large unstructured environments such as search and rescue (SAR) missions and environmental monitoring [1], [2], [3], [4]. A multi-robot solution typically entails using robots that are inexpensive and controlled in a distributed manner [5], [6], [7], [8], [9]. As opposed to coordinated strategies that may require centralized control [10], distributed control of robotic members scale robustly to larger problems. Controlling multiple robots to solve challenging problems, however, is an open area of research with questions related to navigation, detection, and security that often require high levels of autonomy [11], [12], [13], [14]. In this context, human-robot interaction (HRI), and in particular, human-swarm interaction (HSI) is an upcoming field concerned with solving complex problems where humans can participate directly as a member of the swarm, or indirectly as a remote operator [11], [15], [16]. The inclusion of humans

increases the capability of a robotic swarm by offloading the challenges associated with navigation, resource allocation, and security concerns. For humans, having robotic assistance in time-intensive missions leaves them with more time to focus on high level tasks such as navigation and resource planning.

We specifically focus on setups where the human can participate through a first-person view from a teleoperated robot, which in turn exerts control on nearby members of the swarm through shared information and movement [11]. The adoption of teleoperated robots during SAR missions in firefighting and on park trails has been motivated by demanding environments, the need for protective clothing that can severely increase physiological strain on the human, and the urgency to search a large area in the smallest possible time which cannot be done on foot [2], [17], [18], [19]. A teleoperated robot not only mitigates the physiological stress, but also reduces the chance of exposure to poisonous toxins, hazardous materials, and extreme temperature fluctuations, while giving the operator freedom and capability to explore large areas [17], [20], [21], [16]. A teleoperator may independently control a single or multiple robots [22], and interpreting their actions or beliefs entails programming additional mediation layers [23].

A significant cognitive benefit of human involvement during SAR missions is that humans are bound to carry some form of prior knowledge with an associated confidence level [24], [20]. This knowledge may be related to the environment such as building layouts [25], [26] or could be related to the target being rescued and based on experience such as typical locations to find a missing person in a firefighting scene [25], general behavior of how a person may move when lost in a park, or platform specific such as how to operate a robot [27], [19], [28], [20]. Despite the benefits of such prior knowledge, how it may be exploited by autonomous robots during a mission is an open question [29]. It is possible for example that the prior knowledge creates an internal belief map that initially drives the human searcher to steer towards a particular location within the search domain [24]. However, how this prior knowledge must be translated into an actual control strategy that can be assigned to robots is an open question [30]. Relatedly, the accuracy of this prior knowledge, which the human may be aware of, could play a significant role in how exactly each robot should assist the human.

Recent work on human robot interaction that explicitly accounts for human participation on an equal basis with the robots include mixed initiative systems, where robotic agents and humans collaborate while switching between independent and human-assigned tasks [31], [32], and policy blending setups where the control strategy involves calculating a blend

of human and autonomous input [33], [34]. While mixed initiative strategies demonstrate a clear advantage over exclusive roles in SAR and monitoring missions [31], [32] they are often designed to be task-specific and binary in collaboration decisions (the robots can either follow or lead), which makes it difficult to identify the optimal degree of collaboration as a function of task performance and human expertise unless a large number of experiments are performed. A recently proposed framework based on homotopy classes enables navigation decisions driven by resource sharing [35], however this approach is designed for complementary sharing tasks and may not account for situations where robotic assistance may be better utilized through working together. Policy blending approaches provide a relatively more flexible arrangement of combining human and autonomous robot inputs but its effective use requires connecting the blending strategy to human and robot interpretation of the situation [33], [34]. In this context, a control framework that (a) abstracts the mission in terms of low-level objectives, (b) is responsive to human knowledge of the domain, and (c) enables encoding multiple objectives in a single measure can serve to highlight optimal collaboration strategies in a wide variety of settings.

Information-based control, which relies on an information-theoretic interpretation of the mission goal, is an approach that has been found to be especially useful in abstracting complex scenarios into instantaneously measurable objectives [36], [37], [38]. This is because the control objective is encoded in the form of mutual information gain, as opposed to for example minimizing the distance to an object in an unknown location or covering a region [36], [37]. Information-theoretic approaches, however, control the robots as they maximize mutual information with respect to a single entity—the target—and therefore need to be expanded for inclusion of humans in the loop.

The contribution of this work is twofold. First we present a framework that builds on particle filtering and mutual information-based control to realistically capture human robot interaction in a search and rescue setting: the framework allows a human to possess and act on prior knowledge about the target position and the autonomous robot to preferentially assist the human by staying close and thus reducing target uncertainty in their proximity. Second we utilize this framework to analyze the dependence of mission performance on the degree of robotic assistance provided to the human and the accuracy of human prior knowledge in two realistic scenarios simulated with parameters informed from literature. While we generally expect to see a combined strategy outperform independent roles based on results from mixed-initiative research [32], we also expect to find optimal division of control effort for different types and accuracies of human prior knowledge. To provide context with respect to search strategies that also weight between competing objectives, we perform search within the same two scenarios with two other strategies used commonly in the literature: mixed initiative and linear blending of control. With respect to latter, the approach described here could be interpreted as linear blending of mutual information with respect to target uncertainty in different regions of the domain.

This paper is organized as follows: Section II formulates the problem of the generic SAR scenario for a human robot interaction setup; Section III provides a background on particle filtering and mutual information-based control. Next, in Section IV we present the framework for modeling human prior knowledge and a weighted mutual information function that incorporates varying degrees of human assistance by an autonomous robot. Section V describes and analyzes the results of a SAR mission within an indoor environment, and Section VI describes and evaluates a SAR mission in a large park. Section VII compares the results with largest number of robots with mixed initiative and linear blending strategies where a human is able to control the movement of all robots. We conclude in Section VIII with a summary of main results and a discussion of future directions.

II. A GENERIC SAR SETUP FOR A HUMAN ROBOT TEAM

Consider a human robot team consisting of N robots and a single target. All but one robots are autonomous. The human controls a single robot hereby called the reference robot. The state of a robot at a discrete time step k consists of two dimensional position (x_k, y_k) and orientation ψ_k and is denoted by $\mathbf{X}_k \in \mathbb{R}^3$. The target state is denoted by two-dimensional position θ_k within a two dimensional domain $\Theta \subset \mathbb{R}^2$. The kinematics of a robot i can be represented by a nonlinear motion model f^i as

$$\mathbf{X}_{k+1}^i = f^i(\mathbf{X}_k^i, \mathbf{u}_k^i, \omega^i), \quad (1)$$

where $\mathbf{u}_k^i \in \mathbb{R}^{\geq 1}$ is the control input, and ω^i is the disturbance, which captures our confidence in the motion model.

The target location, if it falls within the sensor range, can be sensed by both the autonomous and reference robots based on a known measurement model h^i that relates robot and target position to measurement $\mathbf{Z}_k^{i,\theta} \in \mathbb{R}^{\geq 1}$ by the robot as

$$\mathbf{Z}_k^{i,\theta} = h^i(\mathbf{X}_k^i, \theta_k) + \eta^i, \quad (2)$$

where η^i is additive noise. All robots are assumed to be able to communicate with each other throughout the search domain, although a reference robot may only be able to communicate its location.

Target kinematics are modeled using a constant velocity motion model as

$$\theta_k = \theta_{k-1} + \mathbf{v}\Delta t + \vartheta, \quad (3)$$

where $\mathbf{v} \in \mathbb{R}^2$ is the velocity, $\vartheta \in \mathbb{R}^2$ is the disturbance, and Δt is the length of the time step. The human may possess prior knowledge about target location or dynamics and is expected to search accordingly. The goal of the human-robot team is to find the missing target in minimum possible time. This is equivalent to minimizing the uncertainty in target location estimate.

III. BACKGROUND

In this section we provide an overview of the particle filtering methods and mutual information based control that we utilize in formulating an HSI control strategy for search and rescue.

A. Particle filtering

Particle filters belong to the class of sequential Monte Carlo methods that are used to perform Bayesian estimation [39]. Unlike Kalman filters which expect linearized dynamics and measurement models of the underlying process, and assume additive Gaussian disturbance and noise, particle filters make no such assumptions and can be used directly with nonlinear motion models and non-Gaussian noise distributions [39].

In a particle filter, a continuous probability density function (pdf), $p(\mathbf{X})$, of a random variable $\mathbf{X} \in \mathbb{R}^n$, is approximated by weights corresponding to N_p points in the sample space. These points are called particles. Accordingly, the probability of a point $\tilde{\mathbf{X}}$ is

$$p(\tilde{\mathbf{X}}) = \sum_{q=1}^{N_p} w_q \delta(\mathbf{X}_q - \tilde{\mathbf{X}}), \quad (4)$$

where w_q is the weight of the particle q , with $q = 1, 2, \dots, N_p$ as the particle index, and $\delta(\cdot)$ is the Dirac delta function. The Dirac delta function $\delta(\mathbf{X}_q - \tilde{\mathbf{X}})$ acts as a kernel in that it sets the probability of values that are exactly the same as one of the particles ($\tilde{\mathbf{X}} = \mathbf{X}_q$) as equal to the weight w_q ; at the same time, the probability of a value that is not part of the particle representation ($\tilde{\mathbf{X}} \neq \mathbf{X}_q$) is zero. Therefore, the only way to make this representation of a probability distribution more accurate is to use more particles.

A particle filter estimates a dynamic quantity $\mathbf{X}_k$ as it evolves according to a known dynamics and is sensed with a known measurement model. This is equivalent to maximizing the posterior pdf conditioned on all measurements up until the current time step k , $p(\mathbf{X}_k | \mathbf{Z}_{1:k})$, by recursively iterating through predict and update steps; here $\mathbf{Z}_{1:k}$ denotes all measurements up to k . The predict step evolves the state of each particle through a motion model such as (1). The update step involves applying the Bayesian rule to calculate the probability of a state conditioned on a measurement obtained through (2). In a particle filter, this is accomplished by resampling the particles according to weights calculated as [39]

$$w_{k,q} = p(\mathbf{Z}_k | \mathbf{X}_{k,q}), \quad (5)$$

where the likelihood function $p(\mathbf{Z}_k | \mathbf{X}_{k,q})$ captures the measurement model; for example if the measurement model (2) has Gaussian noise η with standard deviation σ_η , then $p(\mathbf{Z}_k | \mathbf{X}_{k,q}) = \mathcal{N}(\mathbf{Z}_k, h(\mathbf{X}_{k,q}, \boldsymbol{\theta}_{k,q}), \sigma_\eta)$, where σ_η is the standard deviation of Gaussian noise η .

Being amenable to nonlinear representations, a particle filtering framework is adequately suited to model human prior knowledge in the form of a pdf, nonlinear sensing with limited visual range, and measurement sharing between autonomous robots by combining likelihood functions.

B. Mutual information based control

Search and rescue operations entail finding a missing target. This amounts to identifying the control action out of a range of possible actions that actively maximize the posterior pdf of the target location $p(\boldsymbol{\theta}_k | \mathbf{Z}_{1:k}^\theta)$ [36], where the superscript θ on $\mathbf{Z}$ denotes the target measurement.

Information theory provides an intuitive framework to actively search for a missing target. In an information theoretic sense, maximizing the posterior pdf is equivalent to minimizing the uncertainty in the target estimate. Within this framework information entropy is used to quantify the uncertainty in an estimate. With respect to target location $\boldsymbol{\theta}_k$, information entropy of the pdf $p(\boldsymbol{\theta}_k)$ is defined as $H(\boldsymbol{\theta}_k) = -\int_{\Theta} p(\boldsymbol{\theta}_k) \log_2 p(\boldsymbol{\theta}_k) d\boldsymbol{\theta}_k$, and is measured in bits [40]. Maximizing the posterior pdf is equivalent to minimizing the entropy of target estimate conditioned on current measurement $H(\boldsymbol{\theta}_k | \mathbf{Z}_k^\theta)$.

Minimizing the uncertainty of current target estimate, however, does not enable an active control strategy, which must act future possibilities. Therefore, we consider mutual information between the current target prediction denoted by $\boldsymbol{\theta}_k^-$ and possible target measurements $\mathbf{Z}_k^{\theta,-} \in \mathcal{Z}$, within a support $\mathcal{Z}$, defined as [36]

$$I(\boldsymbol{\theta}_k^-; \mathbf{Z}_k^{\theta,-}) = H(\boldsymbol{\theta}_k^-) - H(\boldsymbol{\theta}_k^- | \mathbf{Z}_k^{\theta,-}) \quad (6a)$$

$$= H(\mathbf{Z}_k^{\theta,-}) - H(\mathbf{Z}_k^{\theta,-} | \boldsymbol{\theta}_k^-), \quad (6b)$$

where $H(\boldsymbol{\theta}_k^- | \mathbf{Z}_k^{\theta,-})$ for example is the entropy of $\boldsymbol{\theta}_k^-$ conditioned on knowing $\mathbf{Z}_k^{\theta,-}$. Based on this expression, we note that minimizing uncertainty $H(\boldsymbol{\theta}_k^- | \mathbf{Z}_k^{\theta,-})$ is equivalent to maximizing the mutual information $I(\boldsymbol{\theta}_k^-; \mathbf{Z}_k^{\theta,-})$.

Calculating mutual information is possible in a particle-filtering framework by computing the two quantities on the right hand side of equation (6b) as [36]

$$\begin{aligned} H(\mathbf{Z}_k^{\theta,-}) &= -\int_{\mathcal{Z}} p(\mathbf{Z}_k^{\theta,-}) \log_2 p(\mathbf{Z}_k^{\theta,-}) d\mathbf{Z}_k^{\theta,-} \\ &= -\int_{\mathcal{Z}} \left(\int_{\Theta} p(\mathbf{Z}_k^{\theta,-} | \boldsymbol{\theta}_k^-) p(\boldsymbol{\theta}_k^-) d\boldsymbol{\theta}_k^- \right) \cdot \\ &\quad \log_2 \left(\int_{\Theta} p(\mathbf{Z}_k^{\theta,-} | \boldsymbol{\theta}_k^-) p(\boldsymbol{\theta}_k^-) d\boldsymbol{\theta}_k^- \right) d\mathbf{Z}_k^{\theta,-} \\ &\approx -\sum_{l=1}^M \left[\left\{ \sum_{q=1}^{N_p} \left(p(\mathbf{Z}_{k,l}^{\theta,-} | \boldsymbol{\theta}_{k,q}^-) w_{k,q} \right) \right\} \cdot \right. \\ &\quad \left. \log_2 \left\{ \sum_{q=1}^{N_p} \left(p(\mathbf{Z}_{k,l}^{\theta,-} | \boldsymbol{\theta}_{k,q}^-) w_{k,q} \right) \right\} \right], \end{aligned} \quad (7)$$

and as

$$\begin{aligned} H(\boldsymbol{\theta}_k^- | \mathbf{Z}_k^{\theta,-}) &= -\int_{\mathcal{Z}} \int_{\Theta} p(\mathbf{Z}_k^{\theta,-}, \boldsymbol{\theta}_k^-) \log_2 p(\mathbf{Z}_k^{\theta,-} | \boldsymbol{\theta}_k^-) d\mathbf{Z}_k^{\theta,-} d\boldsymbol{\theta}_k^-, \\ &= -\int_{\mathcal{Z}} \int_{\Theta} p(\mathbf{Z}_k^{\theta,-} | \boldsymbol{\theta}_k^-) p(\boldsymbol{\theta}_k^-) \log_2 p(\mathbf{Z}_k^{\theta,-} | \boldsymbol{\theta}_k^-) d\mathbf{Z}_k^{\theta,-} d\boldsymbol{\theta}_k^-, \\ &\approx -\sum_{l=1}^M \left[\sum_{q=1}^{N_p} \left\{ p(\mathbf{Z}_{k,l}^{\theta,-} | \boldsymbol{\theta}_{k,q}^-) w_{k,q} \cdot \log_2 p(\mathbf{Z}_{k,l}^{\theta,-} | \boldsymbol{\theta}_{k,q}^-) \right\} \right], \end{aligned} \quad (8)$$

where $w_{k,q}$ is the weight of the q -th particle after resampling, and the support for measurement space $\mathcal{Z}$ is made of M distinct values. Note that the likelihood function $p(\mathbf{Z}_k^{\theta,-} | \boldsymbol{\theta}_k^-)$ represents the probability of a measurement conditioned on

a possible move by the target and the robot and therefore contributes to the effect of possible control input on the uncertainty of the target estimate. The control input is determined by solving the following optimization problem at every time step for a robot [36]

$$\mathbf{u}_k = \max_{\mathbf{u} \in U} I(\boldsymbol{\theta}_k^-; \mathbf{Z}_k^{\theta, -}), \quad (9)$$

where U represents the range of control inputs available to the robot. These range of inputs for example can be in the form of speeds and turn rates that the robot may apply at the next time step, or over a sequence of time steps [37].

Implementing this control on a target search scenario gives rise to several intuitive search strategies [36]: in the case of a bearings-only measurement model, multiple robots expectedly move towards the target in trying to minimize the uncertainty by reducing the distance between themselves and the target, and in the case of range-only measurements, multiple robots encircle the target to gain a wide range of perspectives.

IV. INFORMATION BASED SEARCH WITH PRIOR KNOWLEDGE

We seek to develop a robot control strategy which can assist a human with prior knowledge about target location or dynamics in a search and rescue mission. The human is expected to initiate their search according to this prior knowledge possibly favoring certain parts of the search domain; in contrast, all autonomous robots are expected to search through the domain without any location preference. To model robotic assistance in an information-theoretic sense, we seek to formulate a control strategy that weights between an independent exploratory search of the environment and improving human's search efficiency by reducing the uncertainty of target location within the neighborhood of the reference robot.

We assume that human prior knowledge can exist either in the form of a particular location within the search domain (for example, a location where witnesses say the victim was last seen prior to a fire) or in the form of dynamics (for example, a lost person will walk randomly from the position last seen with a speed of 1 m/s). We model human prior knowledge about location as initial target distribution $p(\boldsymbol{\theta}_0^h)$, where the superscript h , whenever present, is used to denote variables related to reference robot.

The form of $p(\boldsymbol{\theta}_0^h)$ depends on the type of knowledge, which may be about a particular location within the domain, in which case it can be modeled as a Gaussian distribution, or about a certain region within the search domain, in which case it can be modeled as a uniform distribution within specific bounds contained inside the search domain. In contrast, the initial target distribution $p(\boldsymbol{\theta}_0^i)$ of autonomous robots, who do not have such prior knowledge is uniform through the search domain. Human prior knowledge about target dynamics is modeled by assigning values to parameters within the constant velocity motion model in (3) such as velocity $\mathbf{v}$ or disturbance $\boldsymbol{\vartheta}$. Target search by the human is accomplished using the mutual information based control (9) so that the reference robot naturally explores a preferred target location.

For the human robot collaboration setup, we note that the human is not able to directly communicate with the autonomous robot beyond broadcasting its location. This for example represents a human controlled robot in a first-person view sense as in [11], [41]. Although it may be possible for a human to both teleoperate and communicate their beliefs to other robots, such a setup will likely put additional burden on the human as they continuously indicate a level of confidence they have regarding a particular location. Robotic assistance here entails a search strategy that reduces target location uncertainty near the reference robot.

In particular, an autonomous robot can reduce local uncertainty near the reference robot, which likely has a further limited sensor range, by being in the proximity of the reference robot itself. In terms of a mutual information based control strategy similar to (9) this can be achieved by taking actions that reduce the uncertainty of the reference robot location. A distance-dependent measurement model will automatically enable actions to reduce uncertainty near the reference robot with reduction in distance. An example of a distance-dependent measurement is that from a bearings-only sensor. Accordingly, the goal of the information theoretic control strategy for an autonomous robot that only assists will be to equivalently minimize the location uncertainty of the reference robot conditioned on possible distance-dependent measurements, $H(\mathbf{X}_k^{h, -} | \mathbf{Z}_k^{i, h, -})$. This in turn amounts to maximizing the mutual information $I(\mathbf{X}_k^{h, -}; \mathbf{Z}_k^{i, h, -})$, where $\mathbf{X}_k^h$ denotes the reference robot's two-dimensional position, and $\mathbf{Z}_k^{i, h}$ denotes a distance-dependent measurement by the autonomous robot.

The relative weighting between the two types of strategies can be captured with a parameter $0 \leq \alpha \leq 1$; $\alpha = 1$ for example should enable searching for a target only thus deploying a 'explore' only strategy by the robot and $\alpha = 0$ should keep the robot close to the reference robot thus increasing the effective field of view of the human in a 'exploit' only strategy with the human. Accordingly, the combined mutual information function is formulated as

$$\mathbf{u}_k = \max_{\mathbf{u} \in U} \left[\alpha \hat{I}(\boldsymbol{\theta}_k^-; \mathbf{Z}_k^{i, \theta, -}) + (1 - \alpha) \hat{I}(\mathbf{X}_k^{h, -}; \mathbf{Z}_k^{i, h, -}) \right], \quad (10)$$

where, for example

$$\hat{I}(\boldsymbol{\theta}_k^-; \mathbf{Z}_k^{i, \theta, -}) = \frac{I(\boldsymbol{\theta}_k^-; \mathbf{Z}_k^{i, \theta, -})}{H(\mathbf{Z}_k^{i, \theta, -})}$$

represents the normalized value of the mutual information of that target searched in the entire domain. Normalization of mutual information with respect to measurement uncertainty allows for meaningful comparison of strategies from two widely different sources of measurement: the target, who likely has random dynamics and whose location is unknown until it falls within a limited sensing range, and the human, who may have relatively more deterministic dynamics and whose location is known throughout the domain. (Fig. S1 in Supplementary information shows un-normalized mutual information values for comparison.) Algorithm 1 summarizes the mutual information based SAR implemented using particle filter in this paper.

Algorithm for HSI-SAR with mutual information

Input: Motion model, f^i , and measurement model, h^i for all robots, target dynamics, relative weighting α

Initialize: target estimates of human $p(\theta_0^h)$, and robot $p(\theta_0^i)$, search start location

Output: Mutual information based control at a time step k , $\mathbf{u}_k$

For each time step $k = 1, 2, \dots$

For each autonomous robot i

1. Update target estimate, θ_k , and reference robot estimate, $\mathbf{X}_k^h$, by resampling particle weights according to (5) and sensor model h^i
2. Calculate optimal control $\mathbf{u}_k$ according to (10) using f^i and likelihood functions $p(\mathbf{Z}_k|\theta_k)$ and $p(\mathbf{Z}_k|\mathbf{X}_k^h)$ in equations (8) and (7)
3. Apply control $\mathbf{u}_k$ to move robot according to model f^i and predict target and reference robot locations according to models (3) and f^i .

For the reference robot

1. Update target estimate, θ_k , by resampling particle weights according to (5) and sensor model h^i
2. Calculate optimal control $\mathbf{u}_k$ according to (9) using f^i and likelihood function $p(\mathbf{Z}_k|\theta_k)$ in equations (8) and (7)
3. Apply control $\mathbf{u}_k$ to move the reference robot according to model f^i and predict target location according to the model (3)

We note that even though the control (10) is designed for an autonomous robot to weight between two strategies, the formulation can be extended to a multi-strategy setup (such as for example following multiple members of a human team based on their level of expertise and prior knowledge) by attaching separate weights to each strategy. We analyze a two-strategy scenario in this work to allow an in-depth analysis of the dependence of task performance on relative weighting.

V. SCENARIO I: INDOOR ENVIRONMENT

We utilized the mutual information based control to first evaluate the dependence of time to find a missing stationary target within an indoor environment on the relative weighting parameter α and the accuracy of human prior knowledge. Human prior knowledge was assumed to exist in the form of a preferred location for initial search within the domain. We simulated differentially driven ground robots with sensor ranges and motion parameters informed by actual experiments from the literature [25], [37], [42], [21], which focus on search and rescue within indoor environments and firefighting operations. Specifically, we adopted the size of the domain from [25] and robot dynamics based on [37], [21]. The range of speeds and turn rates for the robot were based on the design from [25], [42], [37].

A. Setup

The indoor environment consisted of a 30×30 m obstacle free region with the target located in one of three different locations as shown in Fig. 1. These locations were picked so that they are at same distance of 15 m from the search start location at the corner of the region and the only difference in time to find would be due to searching in the wrong part of the region. Human prior $p(\theta_0^h)$ was modeled as a Gaussian distribution centered at one of the locations with standard deviation $\sigma_{\bar{\theta}}$.

Each autonomous robot maintained local estimates of location of the target and that of the reference robot by running an onboard sampling importance resampling (SIR) particle

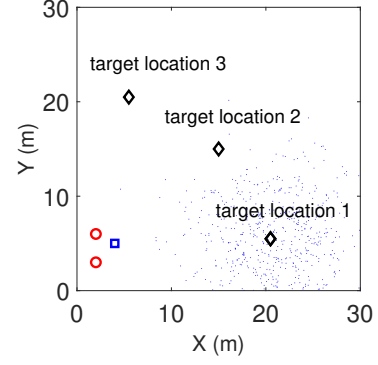


Fig. 1. Setup for SAR in Scenario I showing the reference robot starting position (blue square) and autonomous robots' starting positions (red circles). Three different target positions are shown with black diamonds; human prior knowledge was always in the form of Gaussian distribution of possible target locations (blue dots) centered on target location 1.

filter [39]; the reference robot maintained an estimate of target location only. All robots were modeled as differentially driven ground robots with the motion model

$$\begin{aligned} x_k^i &= x_{k-1}^i + v_k^i \cos \psi_{k-1}^i \Delta t + \omega_v^i \cos \psi_{k-1}^i \Delta t \\ y_k^i &= y_{k-1}^i + v_k^i \sin \psi_{k-1}^i \Delta t + \omega_v^i \sin \psi_{k-1}^i \Delta t \\ \psi_k^i &= \psi_{k-1}^i + \Omega_k^i \Delta t + \omega_\Omega^i \Delta t, \end{aligned} \quad (11)$$

where $\Omega_k^i \in \Psi$ denotes the turn rate and $v_k^i \in V$ denotes the speed, with Ψ and V denoting the range of possible turn rates and speeds; ω_v^i and ω_Ω^i denote zero-mean Gaussian disturbances with standard deviations σ_v and σ_Ω , in the robot velocity and turn rate respectively, and $\Delta t = 1$ second is the length of a time step.

All robots also had the same limited-range bearings-only measurement model for sensing the target, with the reference robot having a limited field of view of 70° based on field of view as seen in pictures of similar experiments [25], and the autonomous robot having a full 360° field of view. The likelihood function representing the sensing of the target by all robots is

$$p(Z_k^{i,\theta} | \mathbf{X}_k^i) = \begin{cases} \mathcal{N}(Z_k^{i,\theta}; h^i(\mathbf{X}_k^i, \theta_k), \sigma_\eta^i) & \text{if } r_k^{i,\theta} \in \rho^i \\ \mathbb{U}(\Theta) \setminus \rho^i & r_k^{i,\theta} \notin \rho^i, \end{cases} \quad (12)$$

where $Z_k^{i,\theta}$ denotes bearing measurement of the target, $\mathcal{N}(Z_k^{i,\theta}; h^i(\mathbf{X}_k^i, \theta_k), \sigma_\eta^i)$ denotes a normal distribution with mean $h^i(\mathbf{X}_k^i, \theta_k)$ and standard deviation σ_η^i sampled at $Z_k^{i,\theta}$, $r_k^{i,\theta}$ is the distance between the target and the robot, and $\mathbb{U}(\Theta) \setminus \rho^i$ denotes a uniform distribution sampled within the two-dimensional search domain that excludes a circular sensor range with radius ρ^i . The bearings-only measurement model h^i (Supplementary document Fig. S3) relates the robot and target state to target measurement as

$$Z_k^{i,\theta} = \Phi_k^{i,\theta} - \psi_k^i + \eta^i, \quad (13)$$

where ψ_k^i is the robot orientation, and $\Phi_k^{i,\theta}$ is the angular position of the target with respect to robot position, and η^i is the observation noise of the sensor.

A bearings-only sensor also serves as an effective distance-dependent measurement so that mutual information based control (10) may be computed at all times. We also note that the assumption of being able to sense the location of the telerobot throughout the domain is not a significant limitation in RF based communication in small domains [37]. The likelihood function representing the tracking of a reference robot by an autonomous robot is also set to bearings-only sensing as

$$p(Z_k^{i,h} | \mathbf{X}_k^i) = \mathcal{N} \left(Z_k^{i,h}; h^i(\mathbf{X}_k^i, \mathbf{X}_k^h), \sigma_\eta^i \right), \quad (14)$$

where $Z_k^{i,h}$ is the bearing measurement of the reference robot.

This setup assumes that estimates of the target and human robot location can be updated by sharing synchronous measurements and measurement models between autonomous robots within a communication range, κ^i , which spans the entire domain in this Scenario. Accordingly the combined likelihood function for an autonomous robot is

$$p(Z^{i,\theta}, Z^{i,h} | \mathbf{X}_k^i) = \prod_{j \in \{\kappa^i\}} \left[p(Z_k^{j,\theta} | \mathbf{X}_k^j) \cdot p(Z_k^{j,h} | \mathbf{X}_k^j) \right]. \quad (15)$$

where $\{\kappa^i\}$ is the set of all robots, including i , that are within κ^i distance of i .

Collision with other robots and the boundary of the environment is handled so that upon a collision, the robot is able to sense the side on which the collision takes place with respect to its heading. A collision on the left causes the robot to turn right and vice versa. Accordingly, upon a collision, the robots change their instantaneous turn rate to $\Omega_k^i = -K_c \gamma_k^i$, where the gain $K_c = 1.5$ determines the intensity of the turn, and γ_k^i is the angle that the point of collision makes with the robot heading.

B. Simulations

Simulations were run to investigate the dependence of mission performance quantified by time to find the target on (a) the relative weighting parameter $\alpha = \{0, 0.25, 0.5, 0.75, 1\}$, (b) the accuracy of human prior knowledge, and (c) the number of autonomous robots deployed in the search (0–2). We considered three levels of accuracy of human prior knowledge ranging from high accuracy so that the true target location aligned with the center of the Gaussian distribution which models the human prior knowledge about target location, to low accuracy where the target location is far from the center of the human prior (Fig. 1). In all cases the distance of the target from the start location was the same. The target was considered found when it fell within the sensor range $\rho^{i,\theta}$ of any robot (autonomous or reference). Imperfect sensing was implemented by setting a probability of detection, $p_d^h = 0.77$ for the reference robot [43] and $p_d^i = 0.66$ for the autonomous robot, assuming that it is equipped with an advanced convolutional neural network based target detection [44]. Table I lists the parameters used to simulate this scenario.

Twenty simulations were run according to the the algorithm in Section IV for each of the conditions and average time to search recorded in each case. The number of particles in the particle filter was determined on the basis of minimizing the variance of target estimate across ten runs in the same

TABLE I
SCENARIO I SIMULATION PARAMETERS

Parameter	Description	Value
Ψ	range of turn rates	$\{-0.25, 0, 0.25\}$ rad/s
V	range of speeds	$\{0, .415, .833\}$ m/s
κ^i	communication range between robots	30 m
ρ^i	sensor range for all robots	3 m
L	size of domain	30×30 m
$\mathbf{v}$	target velocity	0 m/s
σ_θ	standard dev. of target motion disturbance	0.05 m/s
σ_θ	standard dev. of human prior	4.5 m
σ_v	standard dev. in robot speed disturbance	0.1 m/s
σ_Ω	standard dev. in robot turn rate disturbance	0.01 rad/s
σ_η	standard dev. in target sensing noise	0.1 rad
N_p	number of particles in particle filter	600
p_d^h	detection probability for reference robot	0.77
p_d^i	detection probability for autonomous robot	0.66

domain. Specifically, we ran a particle filter with the number of particles ranging between 100 and 1600 to search for a stationary target. The number of particles N_p were selected as the value after which reduction in variance of position estimate was not significant [45] (Fig. S2 in Supplementary information).

Simulations were initialized with each robot placed at random location near the $(0, 0)$ location and oriented at an angle of between 0 and 45 degrees with the reference robot was always oriented at 20 degrees. Simulations that took more than 2000 seconds were stopped with the time to search recorded as the maximum value.

C. Results

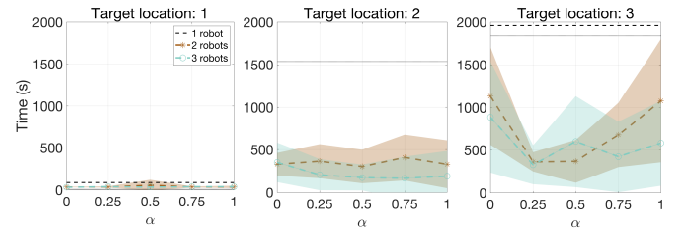


Fig. 2. Average time to find the target for each of the target locations (levels of accuracy of human prior knowledge) and values of the relative weighting parameter α . The number of robots always includes a reference robot and the dotted line marks the standard deviation for search with a single reference robot. Error envelopes denote $\pm$ standard deviation.

Figure 2 shows mission performance in terms of the time taken to find the target in terms of the relative weighting parameter α for three different levels of accuracy of human prior knowledge. We immediately note that the time to find the target for a single reference robot increases as the human prior knowledge becomes more inaccurate. With target location 3, the reference robot is never able to find the target within the limited time. Furthermore, if the human prior knowledge is accurate there appears to be no significant advantage afforded by the presence of autonomous robots. In all situations where human prior knowledge was inaccurate, the presence of an autonomous robot significantly improves performance.

As expected, when the human prior knowledge becomes inaccurate, we begin to see a dependence both on the relative weighting α as well as the number of robots. First, we note

that dependence on α is amplified as prior knowledge becomes inaccurate. Specifically, with respect to relative weighting, when the human is accompanied with one robot and the prior knowledge is inaccurate by about 10 m, the time to find the target does not appear to change significantly, but does so when the prior knowledge is inaccurate by 20 m, where we see a clear U shape in the curve attaining a minimum at $\alpha = 0.25$.

With respect to the number of robots, the time to search lowers in both scenarios where human knowledge was inaccurate. In the case of target location 2, the time to find at $\alpha = 0.25$ drops to less than two-thirds when two robots are assisting than when only one robot is assisting the human; in the case of target location 3, the time to find drops by nearly half at $\alpha = 1$ as the number of robots increase from two to three.

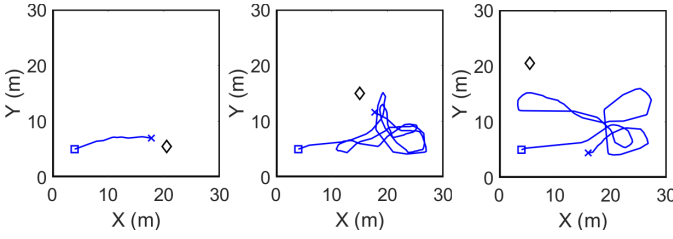


Fig. 3. Sample trajectories in blue of the reference robot with starting positions (blue square) and for three different target locations (black diamonds).

Figure 3 shows sample trajectories of the reference robot searching for the target at three different locations. These trajectories exemplify how a human teleoperator may initially search for the target according to their prior knowledge even if it is inaccurate.

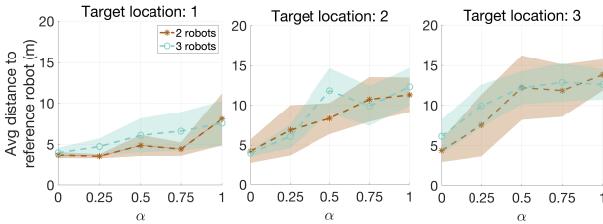


Fig. 4. Average distance of the autonomous robots to the reference robot for different target locations and a range of α values and number of robots. The number of robots in the legend denotes the total number of robots in the mission including the reference robot. Error envelopes denote $\pm$ standard deviation.

Figure 4 confirms the role of the relative weighting parameter α in terms of proximity to the reference robot. Specifically, on average autonomous robots tend to stay further from the reference robot with increase in α with for example the distance between a single autonomous robot and the reference robot increasing from 2.8 ± 0.5 m at $\alpha = 0$ to 7.5 ± 2.3 at $\alpha = 1$. This trend is maintained when the number of autonomous robots increase to two where they maintain approximately the same distance from the reference robot as α is increased. This reduction in distance to the reference robot as the weighting parameter decreases is an evidence of belief induced swarming where robots tend to come close to each other to reduce uncertainty.

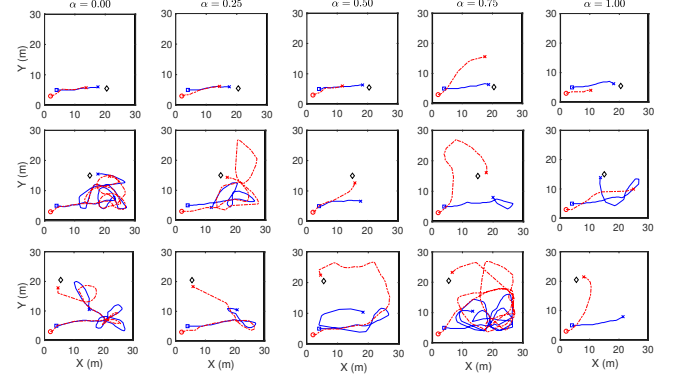


Fig. 5. Sample trajectories of the reference robot (solid blue) with a single autonomous robot (dot-dash red) with its starting position (red circle) for different target locations (individual rows).

Figure 5 shows sample trajectories of two-robot setup with one autonomous and one reference robot. We see that as the α parameter is increased from 0 to 1, the autonomous robot has an increased tendency search far from the reference robot. When the human prior knowledge is inaccurate (second and third rows from top), the benefit of partially weighting the control strategy with respect to the reference robot is evident as α is increased with the autonomous robot finding the target on its own (red cross appears near the target). (See sample trajectories of for all scenarios in Supplementary document.)

VI. SCENARIO II: MOVING TARGET IN A PARK

In this scenario we evaluated the effect of accuracy of human prior knowledge and relative weighting parameter on the time to find a missing person in a large park; human prior knowledge is assumed to exist the form of a probability of area (*POA*) in the search zone. The *POA* is a high-probability circular region centered at the position last seen (PLS) [46]. We simulated fixed-wing UAVs with sensor ranges and motion parameters informed by experiments from the literature [19], [42], [47], which focus on search and rescue in outdoor park environments. Specifically we utilized the same park environment as in [27], dynamics were based on the use of fixed-wing airplanes in [19], and the range of speeds and bank angles were based on [19], [48].

A. Setup

The park consisted of a 10×8 km region with the target last seen at a location marked 'X' (Fig. 6). Based on search and rescue strategies in the literature [46], [19], human prior knowledge, $p(\theta_0^h)$, was modeled in the form of a trail favoring distribution within a circular *POA* region centered on the PLS [46]. We note that lost person dynamics are complex and likely depend on a number of factors including age, motivation, energy levels, and topography [49], [28], [50], [46]. Here, we adopt a simple representation and assume that all these aspects can be captured in the form of disturbance ϑ of the target executing a random walk based on (3) that favors walking along a trail. In particular, the disturbance $\vartheta \in \mathcal{R}^2$ is sampled as from a two-dimensional Gaussian distribution $\mathcal{N}(0, \sigma_{tr})$,

where $\sigma_{tr} = \begin{bmatrix} \cos \zeta & -\sin \zeta \\ \sin \zeta & \cos \zeta \end{bmatrix} \begin{bmatrix} \sigma_v \\ \sigma_v \sqrt{1 - \exp(-d_{tr}^2)} \end{bmatrix}$, with ζ representing the local orientation of the trail, σ_v the speed of the random walker, and d_{tr} the distance from the trail in meters; the variable d_{tr} is used to vary the eccentricity of the covariance ellipse of the random variable so that a walker on the trail will tend to move along the trail, however once off the trail, the tendency to move along the trail will decrease exponentially. The angle ζ is computed by fitting a line to ten closest pixels on the trail.

Human prior knowledge distribution was created by weighting a uniform distribution with a blurred image of the map of the trails (in white) and rivers (in black). With a constant time between when the target was seen and the initiation of the search, this amounts to varying sizes of the *POA* (Fig. 6).

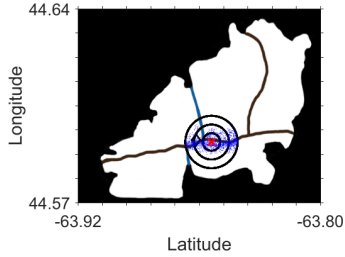


Fig. 6. Setup for SAR in scenario II with search area map adopted from [27]. Both reference and autonomous fixed-wing UAVs start from position (red X). Trails are shown in dark brown and rivers in blue. Search areas corresponding to three different sizes of *POA* are shown as circles with the target location shown as a black diamond; human prior knowledge distribution of target location estimate is shown for one of the *POAs*

All UAVs were modeled as fixed-wing Dubins airplanes adapted to fly at zero pitch so that [51]

$$\begin{aligned} x_k^i &= x_{k-1}^i + v_k^i \cos \psi_{k-1}^i \Delta t + \omega_v^i \cos \psi_{k-1}^i \Delta t \\ y_k^i &= y_{k-1}^i + v_k^i \sin \psi_{k-1}^i \Delta t + \omega_v^i \sin \psi_{k-1}^i \Delta t \\ \psi_k^i &= \psi_{k-1}^i + \frac{g}{v_k^i} \tan \phi_k^i \Delta t, \end{aligned} \quad (16)$$

where x_k^i , y_k^i , and ψ_k^i denote the position and orientation of the UAV, v_k^i denotes the speed, g is acceleration due to gravity, and ϕ_k^i is the commanded bank angle of the UAV at time step k that is constrained to be less than $\frac{\pi}{4}$ radians, the max allowable bank angle of the aircraft [48], and $\Delta t = 1$ second is the length of a time step. As in scenario I all robots also have the same bearings-only measurement model, with the reference UAV having a limited field of view of 70° , and the autonomous UAV having a full 360° field of view. All UAVs were set to localize themselves at every time step with GPS-like accuracy so that their estimate of their own position was updated to a Gaussian distribution centered on their true position with a standard deviation of 10 m.

Differently from scenario I, the UAVs do not avoid collision among themselves, as they can be made to fly at different heights. The UAVs are however steered to stay within the search boundary based on the bearing γ_k^i of the nearest point on the boundary within 400 m. Specifically, the instantaneous bank angle is set to $\phi_k^i = -K_c \gamma_k^i$, where $K_c = 1$ is the gain. Additionally, because RF range is limited for such large domains, we limit the communication range $\kappa_i = 500$ m [52].

B. Simulations

We investigated the dependence of mission performance quantified by time to find the target on (a) the weighting parameter $\alpha = \{0, 0.25, 0.5, 0.75, 1\}$, (b) the accuracy of human prior knowledge in terms of the three different *POAs*, and (c) the number of autonomous robots deployed in the search (0–7). The accuracy of human prior knowledge was varied according to the amount of disturbance in the target dynamics which in turn affected the size of the *POA*. In particular, we model the real target to move at a speed of 2 m/s from the PLS, which corresponds to the distance of 0.72 km ($\vartheta \sqrt{3600 \times 36}$) [53] moved by a random walker in thirty-six hours. We then set the values of $\sigma_\vartheta = \{1, 2, 3\}$ m/s. The corresponding *POAs* were set to uniform distributions within circular regions with radii 0.36 km, 0.72 km and 1.08 km centered on the PLS. The position of the target was arbitrarily set at 0.64 km away from the PLS towards the West. Accordingly, the *POA* = 0.36 km corresponded to searching inaccurately within too small a region, the *POA* = 0.72 km corresponded to an accurate and efficient search, and *POA* = 1.08 km, where the UAVs will search a region larger than needed, corresponds to an inefficient search.

Twenty simulations were run according to the algorithm described in Section IV for each of the conditions and average time to find the target recorded in each case. Simulations were run as follows: first the initial distribution of human prior was created by sampling from a uniform distribution within the *POA*. Next, the search was initialized with each robot placed at random location near the PLS and oriented towards East. The reference robot was constrained to stay within the *POA* by utilizing a boundary avoidance control similar to the one used to keep all robots within the search boundary. As in scenario I, the target was considered found when it came within the sensor ranges with probabilities of detection, $p_d^i = 0.77$ for the reference robot [43] and $p_d^i = 0.66$ for the autonomous robot [44]. Simulations were run for one hour (3600 seconds) at which point they were stopped and the maximum time to find recorded.

TABLE II
SCENARIO II SIMULATION PARAMETERS

Parameter	Description	Value
ϕ_k^i	range of commanded bank angles	$\{-\pi/4, 0, \pi/4\}$ rad
v_k^i	range of speeds	$\{20, 30\}$ m/s
κ^i	communication range for all robots	0.5 km
ρ^i	sensor range for all robots	80 m
L	size of domain	10×8 km
$\mathbf{v}$	target velocity	0 m/s
σ_ϑ	standard dev. of target motion disturbance	$\{1, 2, 3\}$ m/s
σ_v	standard dev. of robot speed disturbance	10 m/s
σ_η	standard dev. of target sensing noise	0.1 rad
N_p	number of particles in particle filter	1500
p_d^i	detection probability for reference robot	0.77
p_d^a	detection probability for autonomous robot	0.66

C. Results

Differently from scenario I, here the reference UAV explored a predetermined *POA* around the PLS. Therefore, we expected that unless the *POA* for reference UAV contained the target the searching with reference UAV only would never be

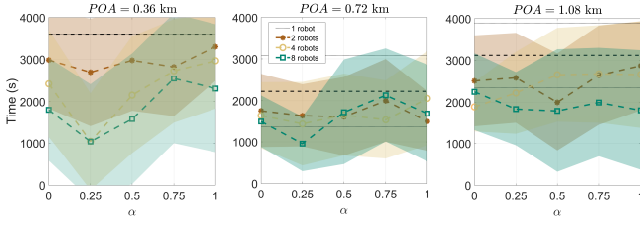


Fig. 7. Average time to find the target for the different sizes of POA as a function of the relative weighting parameter α . The dotted line marks the standard deviation for search with a single reference UAV. Error envelopes denote $\pm$ standard deviation.

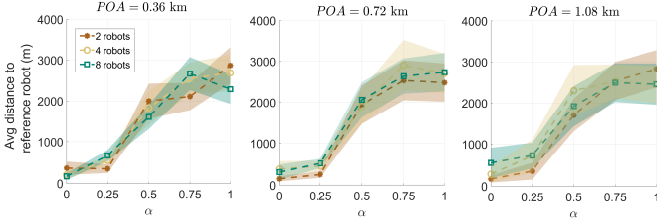


Fig. 8. Average distance of the autonomous UAVs to the reference UAV for different sizes of the POA as a function of the relative weighting parameter α . Error envelopes denote $\pm$ standard deviation.

successful. On the other hand, if the POA was larger than needed, searching with reference UAV only could succeed but would be inefficient. Therefore, the lowest time to find the target with a reference UAV only would occur when the ($POA = 0.72$ km) is just large enough to contain the target. We find this trend in our results (dashed lines in Fig. 7).

When autonomous UAVs were included in the search, we find a dependence on the relative weighting parameter α when the search region is inaccurate and the number of UAVs are more than equal to four. In particular, when $POA = 0.36$ km, the lowest time to find was found with $\alpha = 0.25$. In contrast when the search region was larger than needed, the only dependence on α is noted when four robots are used and they remain close to the reference robot (Fig. 8).

An increase in the number of UAVs amplifies the dependence on α when $POA = 0.36$ km, highlighting the role of a combined strategy in situations where human prior knowledge may not be accurate and the best utilization of robotic teams may involve a strategy that is aware of the accuracy in the prior knowledge. Interestingly, we find a dependence on α with eight UAVs even when the search region is accurate at $POA = 0.72$ km with the lowest time to find attained when $\alpha = 0.25$. This suggests that a large team is needed to effectively increase the sensor range of the human even resulting in higher performance even if the prior knowledge is accurate. (See sample trajectories with two, four, and eight UAVs in Supplementary document.)

VII. COMPARISON WITH EXISTING STRATEGIES

To provide context for these results within the larger literature of search and rescue missions with multiple robots, we simulated the two scenarios where multiple autonomous

robots were globally controlled using two different strategies: mixed initiative control [31] and linear blending [29].

A. Mixed initiative control

To simulate a mixed initiative type strategy, we let all the robots select between an independent sweep strategy of the full domain, and if unable to find the target in a single sweep, switch to searching in a contained region.

For Scenario I, a parallel sweep was performed by three robots from the same initial position. The spacing between robots and their trajectories were set up to avoid overlap of sensor ranges. Upon the switch, the robots move towards the perceived target location and conduct a parallel sweep within a bounded region of 10×10 m centered around target 1 location, corresponding to human prior knowledge. The robots performed the sweep with the same average speed $\tilde{v}$ as in the information-based control strategies presented in this paper. All other parameters including sensory range, detection probability, actual and perceived target locations were kept the same.

For Scenario II, the initial sweep was performed by eight robots as they start from the same initial position and move outward in an Archimedes spiral reaching up to the edge of the park along the river. Robots were placed in a grid-like arrangement ensuring no sensory overlap and large coverage. The width of the sweep was selected to ensure that robots achieve an approximately 60% success rate in a single sweep based on similar success rate with adaptive autonomy in [31]. Upon the switch the robots move to the corresponding POA and sweep back-and-forth in a smaller spiral.

Table III compares the results of twenty simulations of mixed initiative type control for each scenario with the weighted mutual-information based control with equal number of robots. For both scenarios, the mixed initiative either takes the same or more time to find the target than the minimum time attained by the weighted mutual-information based control. We also emphasize that the way a target is searched in a mutual information-based strategy versus a sweep are very different. In a mutual information-based search, control actions are optimized to improve target estimates over time and therefore once they are close to the target, a robot that optimizes mutual information with respect to future measurements spends more time near the target leading to a reinforcement of target measurements. In contrast, the mixed initiative approach sweeps by the target which implies that at lower detection rates such a strategy is likely to perform even worse.

B. Linear blending

For linear blending of control, we envision a scenario where a swarm of robots are controlled by a human, and the robots tune their control to blend operator input with random search. We accordingly defined a parameter $\beta = \{0, 0.25, 0.5, 0.75, 1\}$ which represents the degree to which the robots controlled directly; $\beta = 0$ corresponds to full control (robots go to specific locations) and $\beta = 1$ no control (robots perform a random search). All robots move with the speed $\tilde{v}$, with detection rates and sensor ranges as in the weighted mutual

information based control. For Scenario I, we accordingly implement the following control strategy at each timestep

$$\Omega_k = \beta \mathbb{U}(-0.25, 0.25) + 0.25(1 - \beta) \frac{\zeta_k^{i,\theta}}{\pi}, \quad (17)$$

where $\mathbb{U}(-0.25, 0.25)$ represents a uniformly distributed random variable between -0.25 and 0.25 , and $\zeta_k^{i,\theta}$ is the bearing to the target location 1. The normalization by π ensures that the turn rate does not exceed the range limits. For Scenario II, we implement the following control strategy at each step

$$\phi_k = \beta \mathbb{U}\left(-\frac{\pi}{4}, \frac{\pi}{4}\right) + \frac{\pi}{4}(1 - \beta) \frac{\zeta_k^{i,\theta}}{\pi} \quad (18)$$

where $\zeta_k^{i,\theta}$ is the bearing to the target location θ_{POA} , updated every twelve minutes, within the corresponding POA. The twelve-minute update would reflect human intent to call the swarm to search different locations and is informed by the time it would take to scatter search a region once a possible cue is detected [8].

Table III lists the average time to find the target over twenty simulations of linear blending control strategy for each value of β . In Scenario I, as expected, we note that when the robots are steered towards the correct target location (Target 1), the time to find is a few seconds more than the weighted information strategy; however, any actions steering toward a miscalculated target location are heavily penalized and the best times are only attained with a completely random search ignoring any human prior knowledge. This is expected because a blending approach directly converts incorrect knowledge to a control action. Such a strategy, therefore leaves no room for ignoring the control input based on current levels of uncertainty, a feature which occurs automatically in the weighted mutual information based control (10).

For Scenario II, we find that for high values of β (perform random search more compared to being steered to a location) linear blending outperforms weighted information strategy when $POA = 0.36$ km. This is because a moderate amount of steering to stay within the POA is able to constrain the area of search, ultimately resulting in a random search within a smaller region that contains the target. Contrastingly, for lower values of α (search close to the reference robot), linear blending takes more time to find the target than weighted information based strategy. This is because all the robots follow steering commands to move together within a few select regions resulting in largely unsuccessful search. Here too, as in the mixed initiative strategy, because robots pass by the target, lower detection rates will likely lead to worse times to find.

VIII. CONCLUSION

Large robot teams can robustly handle large scale complex missions, however they lack the prior knowledge that a human may possess through training and experience. In this context, a solution integrating humans within a multi-robot team should be able to exploit such knowledge. In this paper, a weighted information-theoretic control was formulated to evaluate the effectiveness of multiple robots as they vary their degree of assistance to human. A particle filtering framework allowed

TABLE III
COMPARISON OF MIXED INITIATIVE AND LINEAR BLENDING CONTROL WITH THE WEIGHTED MUTUAL INFORMATION STRATEGY. VALUES OF β AND α ARE IN PARENTHESES.

Strategy	Mixed initiative (s)	Linear blending (s)	Wtd. mutual information (s)
Scenario I			
Condition, speed (m/s)	3 autonomous	3 autonomous	2 autonomous & 1 reference
Location 1, $\bar{v} = 0.35$	46.2 ± 3.8	40.4 ± 0.88 (0)	33.6 ± 4.72 (0)
		41.2 ± 1.19 (0.25)	40.3 ± 9.5 (0.25)
		44.5 ± 3.8 (0.5)	35.6 ± 7.4 (0.5)
		54.2 ± 17.89 (0.75)	39.7 ± 4.9 (0.75)
		115.7 ± 103.38 (1)	39.9 ± 9.6 (1)
Location 2, $\bar{v} = 0.27$	154.3 ± 2.5	2000 ± 0 (0)	351.3 ± 229.2 (0)
		2000 ± 0.22 (0.25)	208 ± 181.5 (0.25)
		1490.2 ± 749.7 (0.5)	175.2 ± 150.7 (0.5)
		150.9 ± 116.85 (0.75)	167.5 ± 250.1 (0.75)
		174 ± 171.32 (1)	191.8 ± 291.57 (1)
Location 3, $\bar{v} = 0.18$	343.1 ± 1.3	2000 ± 0 (0)	887 ± 647.9 (0)
		2000 ± 0 (0.25)	328.5 ± 225.3 (0.25)
		$2000 \pm .22$ (0.5)	598.8 ± 530.1 (0.5)
		1999 ± 489 (0.75)	422.2 ± 413.3 (0.75)
		222 ± 190.5 (1)	580.2 ± 495.01 (1)
Scenario II			
	8 autonomous	8 autonomous	7 autonomous & 1 reference
$POA = 0.36$ km, $\bar{v} = 20.5$	2756 ± 1499.8	3210.7 ± 803.5 (0)	1805.3 ± 1225.5 (0)
		3385 ± 607.7 (0.25)	1037.3 ± 1375.9 (0.25)
		2321 ± 1394.8 (0.5)	1584.2 ± 1555.6 (0.5)
		359.9 ± 354 (0.75)	2547.6 ± 1549.6 (0.75)
		1383.8 ± 1300.7 (1)	2309 ± 1528.7 (1)
$POA = 0.72$ km, $\bar{v} = 20.8$	1150.1 ± 1056.5	2578.5 ± 1498.4 (0)	1499.4 ± 626.8 (0)
		2681.7 ± 1331.2 (0.25)	955.9 ± 662.1 (0.25)
		1776 ± 1533.3 (0.5)	1724.3 ± 1262.2 (0.5)
		1274.3 ± 1372.7 (0.75)	2129.6 ± 1122.1 (0.75)
		1265.3 ± 1332.7 (1)	1692 ± 1170.4 (1)
$POA = 1.08$ km, $\bar{v} = 20.4$	1699.4 ± 992.2	2751.3 ± 1514.8 (0)	2247.8 ± 919.2 (0)
		2293.3 ± 1410.8 (0.25)	1836.7 ± 880.9 (0.25)
		2035.6 ± 1531.2 (0.5)	1793 ± 1470.8 (0.5)
		1406.8 ± 1281.1 (0.75)	1992.8 ± 1304.9 (0.75)
		1957.5 ± 1429.9 (1)	1805.7 ± 1435 (1)

for a realistic representation and Bayesian update of human prior knowledge as well as limited-range sensor models while searching for a missing target. Simulations in two widely different scenarios with parameters informed from experiments in the literature revealed a dependence of mission performance on the weighting parameter as well as the accuracy of human prior knowledge. Specifically, when the human prior knowledge is inaccurate, both in terms of location or the size of the region to search, robotic assistance is effective if autonomous robots weigh between complete independence and human assistance. Comparison with existing strategies where robots are globally controlled with mixed initiative and linear blended inputs highlights the importance of weighing information in search and rescue missions with uncertainties.

We note that the first-person teleoperation is in contrast to many swarm-robotic approaches that have explored teleoperation at a global level. In these instances, the human typically extends a leader-like role to reference robots, and therefore, the research focus shifts to improving swarm performance through leader selection [54], relative positioning [55], directing through waypoints [56], or managing the collective state [56] of the swarm. Compared to global control of robotic swarm, a first-person view is able to directly utilize human knowledge albeit at the cost of limiting the degree of control on the rest of the swarm.

There are several directions where this work can be continued in the future including a better understanding of human integration as well as formulation of human-aware optimal control strategies. With respect to understanding the role of a human in SAR missions, the task-specific representation and update of prior knowledge requires a data-driven approach. Relatedly, we also assumed a static value of the relative weighting

parameter, however, it is possible that a time-dependent α may constitute an optimal search strategy that is also responsive to human awareness of the mission. Implementing such a strategy will require online estimation of human behavior and intent, a growing area of research within human-robot interaction [57], [58]. Another direction that can be pursued is with respect to tuning the interaction rules between autonomous robots towards a certain objective such as maintaining a particular formation [59] or by adaptively sharing information with select neighbors as has been shown to improve performance in particle swarm optimization [60]. Finally, a ground robotic test bed is being developed for a human-subjects study that will be used to collect data for modeling human bias in complex search and rescue missions. Additionally, the proposed formulation can be extended to evaluate combinations of strategies in complex scenarios where multiple humans with varying degrees of prior knowledge team with robots with different capabilities.

ACKNOWLEDGMENT

This research was supported by National Science Foundation under grant # IIS-2033918. Rafal Krzysiak acknowledges support from the Great Journeys Assistantship program.

REFERENCES

- [1] R. N. Darmanin and M. K. Bugeja, "A review on multi-robot systems categorised by application domain," *Proceedings of the Mediterranean Conference on Control and Automation*, pp. 701–706, 2017.
- [2] A. Hong, O. Igharoro, Y. Liu, F. Niroui, G. Nejat, and B. Benhabib, "Investigating Human-Robot Teams for Learning-Based Semi-autonomous Control in Urban Search and Rescue Environments," *Journal of Intelligent and Robotic Systems: Theory and Applications*, vol. 94, no. 3-4, pp. 669–686, 2019.
- [3] W. Burgard, M. Moors, D. Fox, R. Simmons, and S. Thrun, "Collaborative multi-robot exploration," *Proceedings of the IEEE International Conference on Robotics and Automation*, vol. 1, pp. 476–481, 2000.
- [4] W. K. Wong, S. Ye, H. Liu, and Y. Wang, "Effective Mobile Target Searching Using Robots," *Mobile Networks and Applications*, 2020.
- [5] A. Gautam and S. Mohan, "A review of research in multi-robot systems," in *IEEE International Conference on Industrial and Information Systems (ICIIS)*, 2012, pp. 1–5.
- [6] C. Wei, K. V. Hindriks, and C. M. Jonker, "Dynamic task allocation for multi-robot search and retrieval tasks," *Applied Intelligence*, vol. 45, no. 2, pp. 383–401, 2016.
- [7] W. Saab, P. Racioppo, and P. Ben-Tzvi, "A review of coupling mechanism designs for modular reconfigurable robots," *Robotica*, vol. 37, no. 2, pp. 378–403, 2019.
- [8] R. D. Arnold, H. Yamaguchi, and T. Tanaka, "Correction to: Search and rescue with autonomous flying robots through behavior-based cooperative intelligence," *Journal of International Humanitarian Action*, vol. 4, no. 1, 2019.
- [9] M. Duarte, J. Gomes, V. Costa, T. Rodrigues, F. Silva, V. Lobo, M. M. Marques, S. M. Oliveira, and A. L. Christensen, "Application of swarm robotics systems to marine environmental monitoring," *OCEANS*, pp. 1–8, 2016.
- [10] Y. Cao, W. Yu, and W. Ren, "An Overview of Recent Progress in the Study of Distributed Multi-Agent Coordination," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 1, pp. 427–438, 2013.
- [11] A. Kolling, P. Walker, N. Chakraborty, K. Sycara, and M. Lewis, "Human Interaction with Robot Swarms: A Survey," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 1, pp. 9–26, 2016.
- [12] C. Sampedro, A. Rodriguez-Ramos, H. Bavle, A. Carrio, P. de la Puente, and P. Campoy, "A Fully-Autonomous Aerial Robot for Search and Rescue Applications in Indoor Environments using Learning-Based Techniques," *Journal of Intelligent and Robotic Systems: Theory and Applications*, vol. 95, no. 2, pp. 601–627, 2019.
- [13] R. Jonschkowski and O. Brock, "Learning state representations with robotic priors," *Autonomous Robots*, vol. 39, no. 3, pp. 407–428, 2015.
- [14] S. A. Green, X. Q. Chen, M. Billingham, and J. G. Chase, "Human robot collaboration: A Literature Review and Augmented Reality Approach in," *Proceedings of the ASME International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, vol. 4, no. 1, pp. 117–126, 2008.
- [15] J. Hu, J. Xu, and L. Xie, "Cooperative Search and Exploration in Robotic Networks," *Unmanned Systems*, vol. 1, no. 1, pp. 121–142, 2013.
- [16] Y. Wang and F. Zhang, *Trends in Control and Decision-Making for Human-Robot Collaboration Systems*. Springer International Publishing, 2017.
- [17] J. Zhao, J. Gao, F. Zhao, and Y. Liu, "A search-and-rescue robot system for remotely sensing the underground coal mine environment," *Sensors*, vol. 17, no. 10, pp. 1–23, 2017.
- [18] A. Kolling, K. Sycara, S. Nunnally, and M. Lewis, "Human Swarm Interaction: An Experimental Study of Two Types of Interaction with Foraging Swarms," *Journal of Human-Robot Interaction*, vol. 2, no. 2, pp. 104–129, 2013.
- [19] T. Niedzielski, M. Jurecka, B. Miziński, J. Remisz, J. Śłopek, W. Spallek, M. Witek-Kasprzak, Ł. Kasprzak, and M. Świerczyńska-Chłasciak, "A real-time field experiment on search and rescue operations assisted by unmanned aerial vehicles," *Journal of Field Robotics*, vol. 35, no. 6, pp. 906–920, 2018.
- [20] R. R. Murphy, "Human-Robot Interaction in Rescue Robotics," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 34, no. 3, pp. 138–153, 2004.
- [21] C. Luo, A. P. Espinosa, D. Pranantha, and A. De Gloria, "Multi-robot search and rescue team," *IEEE International Symposium on Safety, Security, and Rescue Robotics*, pp. 296–301, 2011.
- [22] Y. Liu and G. Nejat, "Robotic urban search and rescue: A survey from the control perspective," *Journal of Intelligent & Robotic Systems*, vol. 72, no. 2, pp. 147–165, 2013.
- [23] R. Wegner and J. Anderson, "Agent-based support for balancing teleoperation and autonomy in urban search and rescue," *International Journal of Robotics and Automation*, vol. 21, no. 2, pp. 120–127, 2006.
- [24] J. L. Burke, R. R. Murphy, M. D. Coover, and D. L. Riddle, "Moonlight in Miami: A field study of human-robot interaction in the context of an urban search and rescue disaster response training exercise," *Human-Computer Interaction*, vol. 19, no. 1-2, pp. 85–116, 2004.
- [25] Y. Tamura, H. Amano, and J. Ota, "Analysis of firefighting skill with a teleoperated robot," *ROBOMECH Journal*, vol. 7, no. 1, 2020.
- [26] C. Wang, J. Cheng, J. Wang, X. Li, and M. Q. Meng, "Efficient object search with belief road map using mobile robot," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3081–3088, 2018.
- [27] K. A. Hill, "Cognition in the woods: Biases in probability judgments by search and rescue planners," *Judgment and Decision Making*, vol. 7, no. 4, pp. 488–498, 2012.
- [28] J. Alanis, M. M. Brown, J. Kitchens, J. Magaña, C. Velastegui, M. Thakur, L. Arriola, B. Espinoza, A. Murillo, M. Rodríguez-Messan, R. Koester, and C. Castillo-Garsow, "Topography and Behavior Based Movement Modeling for Missing Hikers in Land-Wilderness Settings," pp. 1–53.
- [29] S. Musić and S. Hirche, "Control sharing in human-robot team interaction," *Annual Reviews in Control*, vol. 44, pp. 342–354, 2017.
- [30] B. Hayes and B. Scassellati, "Challenges in Shared-Environment Human-Robot Collaboration," *Learning*, vol. 62, no. 1, pp. 453–458, 2013.
- [31] B. Hardin and M. A. Goodrich, "On using mixed-initiative control: A perspective for managing large-scale robotic teams," in *Proceedings of the ACM/IEEE International Conference on Human Robot Interaction*, 2009, pp. 165–172.
- [32] J. Y. Chen and M. J. Barnes, "Human - Agent teaming for multirobot control: A review of human factors issues," *IEEE Transactions on Human-Machine Systems*, vol. 44, no. 1, pp. 13–29, 2014.
- [33] A. D. Dragan and S. S. Srinivasa, "A policy-blending formalism for shared control," *International Journal of Robotics Research*, vol. 32, no. 7, pp. 790–805, 2013.
- [34] P. Trautman, "Assistive Planning in Complex, Dynamic Environments: A Probabilistic Approach," *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, pp. 3072–3078, 2016.
- [35] V. Govindarajan, S. Bhattacharya, and V. Kumar, "Human-robot collaborative topological exploration for search and rescue applications," *Springer Tracts in Advanced Robotics*, vol. 112, pp. 17–32, 2016.
- [36] G. M. Hoffmann and C. J. Tomlin, "Mobile sensor network control using mutual information methods and particle filters," *IEEE Transactions on Automatic Control*, vol. 55, no. 1, pp. 32–47, 2010.

- [37] B. Charrow, N. Michael, and V. Kumar, "Cooperative multi-robot estimation and control for radio source localization," *International Journal of Robotics Research*, vol. 33, no. 4, pp. 569–580, 2014.
- [38] K. P. Valavanis, "The Entropy Based Approach to Modeling and Evaluating Autonomy and Intelligence of Robotic Systems," *Journal of Intelligent and Robotic Systems: Theory and Applications*, vol. 91, no. 1, pp. 7–22, 2018.
- [39] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/nongaussian bayesian tracking," *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*, vol. 50, no. 2, pp. 723–737, 2007.
- [40] T. M. Cover, *Elements of Information Theory*. John Wiley & Sons., 1991, vol. 1.
- [41] J. W. Crandall, N. Anderson, C. Ashcraft, J. Grosh, J. Henderson, J. McClellan, A. Neupane, and M. A. Goodrich, "Human-swarm interaction as shared control: Achieving flexible fault-tolerant systems," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10275 LNAI, pp. 266–284, 2017.
- [42] K. S. Tseng and B. Mettler, "Analysis and augmentation of human performance on telerobotic search problems," *IEEE Access*, vol. 8, pp. 56 590–56 606, 2020.
- [43] R. J. Koester, K. B. Chiacchia, C. R. Twardy, D. C. Cooper, J. R. Frost, and R. Q. Robe, "Use of the visual range of detection to estimate effective sweep width for land search and rescue based on 10 detection experiments in North America," *Wilderness and Environmental Medicine*, vol. 25, no. 2, pp. 132–142, 2014.
- [44] S. Sambolek and M. Ivacic-Kos, "Automatic person detection in search and rescue operations using deep CNN detectors," *IEEE Access*, vol. 9, pp. 37 905–37 922, 2021.
- [45] D. Fox, "KLD-sampling: Adaptive particle filters and mobile robot localization," *Advances in Neural Information Processing Systems*, vol. 14, no. 1, pp. 26–32, 2001.
- [46] E. Sava, C. Twardy, R. Koester, and M. Sonwalkar, "Evaluating Lost Person Behavior Models," *Transactions in GIS*, vol. 20, no. 1, pp. 38–53, 2016.
- [47] L. Lin and M. A. Goodrich, "UAV intelligent path planning for wilderness search and rescue," *IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 0, no. 1, pp. 709–714, 2009.
- [48] F. A. Administration, "Basic Flight Maneuvers," *Helicopter Flying Handbook*, pp. 9–1—9–20, 2019.
- [49] 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.
- [50] A. Hashimoto and N. Abaid, "An Agent-Based Model of Lost Person Dynamics for Enabling Wilderness Search and Rescue," in *Dynamic Systems and Control Conference*, oct 2019.
- [51] R. W. Beard and T. W. McLain, "Implementing Dubins Airplane Paths on Fixed-wing UAVs," *Handbook for Unmanned Aerial Vehicles*, pp. 1–27, 2013.
- [52] T. D. Dinh, R. Pirmagomedov, V. D. Pham, A. A. Ahmed, R. Kirichek, R. Glushakov, and A. Vladyko, "Unmanned aerial system-assisted wilderness search and rescue mission," *International Journal of Distributed Sensor Networks*, vol. 15, no. 6, p. 1550147719850719, 2019.
- [53] A. Papoulis, *Probability, Random Variables and Stochastic Processes*. McGraw Hill, 1991.
- [54] P. Walker, S. Amirpour Amraii, N. Chakraborty, M. Lewis, and K. Sycara, "Human control of robot swarms with dynamic leaders," *IEEE International Conference on Intelligent Robots and Systems*, pp. 1108–1113, 2014.
- [55] R. Tiwari, P. Jain, S. Butail, S. P. Baliyarasimhuni, and M. A. Goodrich, "Effect of leader placement on robotic swarm control," *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS*, vol. 3, pp. 1387–1394, 2017.
- [56] D. S. Brown, M. A. Goodrich, S.-Y. Jung, and S. C. Kerman, "Two Invariants of Human Swarm Interaction," *Journal of Human-Robot Interaction*, vol. 5, no. 1, p. 1, 2015.
- [57] C. Nam, P. Walker, H. Li, K. Sycara, and M. Lewis, "Models of trust in human control of swarms with varied levels of autonomy," *IEEE Transactions on Human-Machine Systems*, no. 3, pp. 194–204, 2019.
- [58] V. Jain, R. Jena, H. Li, T. Gupta, D. Hughes, M. Lewis, and K. Sycara, "Predicting Human Strategies in Simulated Search and Rescue Task," *arXiv*, no. 2011.07656, 2020.
- [59] D. F. Spears, D. R. Thayer, and D. V. Zanzhitzky, "Foundations of swarm robotic chemical plume tracing from a fluid dynamics perspective," *International Journal of Intelligent Computing and Cybernetics*, vol. 2, no. 4, pp. 745–785, 2009.
- [60] R. Poli, J. Kennedy, and T. Blackwell, "Particle swarm optimization: An overview," *Swarm Intelligence*, vol. 1, no. 1, pp. 33–57, 2007.