2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 25-29, 2020, Las Vegas, NV, USA (Virtual)

Expedited Multi-Target Search with Guaranteed Performance via Multi-fidelity Gaussian Processes

Lai Wei, Xiaobo Tan, and Vaibhav Srivastava

Abstract— **We consider a scenario in which an autonomous vehicle equipped with a downward facing camera operates in a 3D environment and is tasked with searching for an unknown number of stationary targets on the 2D floor of the environment. The key challenge is to minimize the search time while ensuring a high detection accuracy. We model the sensing field using a multi-fidelity Gaussian process that systematically describes the sensing information available at different altitudes from the floor. Based on the sensing model, we design a novel algorithm called Expedited Multi-Target Search (EMTS) that (i) addresses the coverage-accuracy trade-off: sampling at locations farther from the floor provides wider field of view but less accurate measurements, (ii) computes an occupancy map of the floor within a prescribed accuracy and quickly eliminates unoccupied regions from the search space, and (iii) travels efficiently to collect the required samples for target detection. We rigorously analyze the algorithm and establish formal guarantees on the target detection accuracy and the detection time. We illustrate the algorithm using a simulated multi-target search scenario.**

I. Introduction

Autonomous multi-target search requires an autonomous agent to quickly and accurately locate multiple targets of interest in an unknown and uncertain environment. Examples include search and rescue missions, mineral exploration, and tracking natural phenomena. A key challenge in a multi-target search task is to balance several trade-offs including explore-vs-exploit: detecting a target with high accuracy versus finding new targets, and speed-vs-accuracy: quickly versus accurately deciding on the presence of a target. The latter includes fidelity-vs-coverage trade-off: sampling at locations farther from the floor provides a wider field of view but less accurate measurements.

In this paper, we design and analyze a multi-target search algorithm that addresses these trade-offs. In particular, our algorithm leverages multi-fidelity Gaussian processes to capture the fidelity-coverage trade-off. Information-theoretic techniques are employed to efficiently explore the environment, and Bayesian techniques are used to accurately identify targets and construct an occupancy map.

Search and persistent monitoring problems have been studied extensively in the literature. Informative path planning is subclass of these problems in which robot trajectories are designed to maximize the information collected along the way-points while ensuring that the distance traveled is within a prescribed budget. Such informative path planning problems are studied in [1–5].

This work was supported by NSF Award IIS-1734272

The authors are with the Department of Electrical and Computer Engineering. Michigan State University, East Lansing, MI 48823 USA. e-mail: {weilai1, xbtan, vaibhav}@msu.edu

Gaussian processes (GPs) are widely used models for capturing spatiotemporal sensing fields in robotics [6, 7]. While GP-based approaches have been used extensively, most of them rely on single-fidelity measurements, i.e., the sensing model does not consider different altitudes at which the measurements can be collected. GP models have also been used extensively to plan informative trajectories for the robots [5, 8–11]. However, most of these works focus on maximizing the reduction in uncertainty of the estimates.

In the context of target search, the trajectory should be designed to balance the explore-exploit tension—the robot should spend more time at target locations, while learning target locations. There have been some efforts to address such explore-exploit tension within the context of informative path planning [11–22].

Hollinger *et al.* [17] study an inspection problem in which the robot needs to classify the underwater surface. They use a combination of GP-implicit surface modeling and sequential hypothesis testing to classify surfaces. Meera *et al.* [21] study informative path planning for a target search problem. They model target occupancy as a GP and design a heuristic algorithm for target detection that handles trade-offs among information gain, field coverage, sensor performance, and collision avoidance. They illustrate the performance of their algorithm using numerical simulations. Sung *et al.* [22] study the hot-spot identification problem in an environment within the framework of GP multiarmed bandits [23, 24]. The multi-target search can be viewed as a hot-spot identification problem in which, instead of global maximum of the field, all locations with value greater than a threshold need to be identified. Such problems have been studied in the multiarmed bandit literature [25, 26]; however, we are not aware of any such studies in the GP setting. Furthermore, all these works focus on single fidelity measurements, while we focus on multiple fidelities of measurements induced by the altitudes relative to the 2D floor at which the measurements are collected.

In this paper, we design an algorithm for expedited search of unknown number of targets located at the 2D floor of an unknown and uncertain 3D environment. We use autoregressive multi-fidelity GPs [27, 28] to model the likelihood of the presence of a target at a location as computed by a computer vision algorithm using the sample collected at that location at a given altitude. Here, fidelity corresponds to the altitude at which the samples are collected. A high altitude (low fidelity) sample provide more global but less accurate information compared with a low altitude (high fidelity) sample. The low fidelity information can be used

978-1-7281-6211-9/20/$31.00 ©2020 IEEE

to quickly find easy-to-detect targets and this enables the robot to focus on high-fidelity information, possibly only in small regions in the environment and consequently, expedite the search. The proposed EMTS algorithm comprises three main modules (i) a sampling and fidelity planner, (ii) a classification and region-elimination algorithm to construct occupancy map of the floor and eliminate unoccupied regions from search space, and (iii) a path planner that allows the vehicle to travel efficiently to collect required samples. The major contributions of this work are:

- We extend the classical informative path planning approach for single-fidelity GPs to multi-fidelity GPs. This novel extension allows for jointly planning for sampling locations and associated fidelity-levels, and thus, addresses the fidelity-coverage trade-off.
- We augment the sampling and fidelity planner with a Bayesian classification and region-elimination algorithm that ensures the targets are identified with desired accuracy, as well as a Traveling Sales Person (TSP) path planner that enables travel-efficient sampling.
- We rigorously analyze the interaction of above components and establish formal guarantees of the target detection accuracy and detection time.

The remainder of the paper is organized as the following. We present a mathematical formulation of our problem in Section II. In Section III, we present the EMTS algorithm and illustrate it using an underwater victim search scenario in Section IV. We analyze the performance of EMTS in Section V and conclude this work in Section VI.

II. Problem Description

We consider an autonomous vehicle that moves in a 3D environment, e.g., an aerial or an underwater vehicle. We assume that the vehicle either moves with unit speed or hovers at a location. The vehicle is tasked with searching for multiple targets on the 2D floor of the environment. Let $D \subset \mathbb{R}^2$ be the area of the floor in which the targets may be present. The vehicle is equipped with a fixed camera that points towards the floor. The vehicle travels across the environment and collects images/videos of the floor (samples) from different sampling points. These sampling points may be located at different altitudes relative to the floor of the environment. We assume that no sample is collected during the movement between sampling points to avoid misleading low-quality sensing information. The collected samples are processed with a computer vision algorithm that outputs a score, which corresponds to the likelihood of a target being present, for each frame. An example of such computer vision algorithm is the state of art deep neural network YOLOv3 [29]. The score will be used to update the estimate of the sensing output, i.e., the estimated score function $f : D \to [0, 1]$ which will be used to determine the location of the targets. The stochastic model for f is introduced below.

A. Multi-fidelity Sensing Model

GPs are widely used models for spatially distributed sensing outputs. In [21], a GP is used to model the target detection output of a computer vision algorithm. While target presence is a binary event, the computer vision algorithms such as YOLOv3 yield a score which is a function of the saliency and location of the target in the image. GPs are appropriate models for such score functions. So far in the literature, GPs have been used in the context of single-fidelity measurements. To characterized the inherent fidelity-coverage trade-off in sensing the floor scene by an autonomous vehicle operating in 3D space, we employ a novel multi-fidelity GP model. The two key physical sensing characteristics the model seeks to capture are: (i) there is some information that can only be accessed at lower altitudes, (ii) the sensing outputs are more spatially correlated at higher altitudes, since the fields of view at neighboring locations have higher overlaps in their field of views.

We assume that the vehicle can collect samples of the floor from M possible heights from the floor $z_1 > z_2 > \cdots > z_M$. We refer to these heights as the fidelity level of the measurement, with M (resp. 1) corresponding to the highest (resp. lowest) level of fidelity. Let the score function $g_m : D \to [0, 1]$ be defined by the output of the computer vision algorithm for an ideal noise-free image collected at fidelity level $m \in \{1, \ldots, M\}$ with the field of view of the camera centered at $\boldsymbol{x} \in D$. We assume that the score functions for a location $\boldsymbol{x}$ obtained from different altitudes (fidelity levels) are related to each other in an autoregressive manner as follows

$$g^m(\boldsymbol{x}) = a_{m-1} g^{m-1}(\boldsymbol{x}) + b^m(\boldsymbol{x}), \tag{1}$$

where a_{m-1} is a scale parameter and b^m is the bias term that captures the information that can be only be accessed at fidelities levels greater than m. Let $f^m(\boldsymbol{x}) = \left(\prod_{i=m}^{M-1} a_i\right) g^m(\boldsymbol{x})$ and $h^m(\boldsymbol{x}) = \left(\prod_{i=m}^{M-1} a_i\right) b^m(\boldsymbol{x})$. Then, equation (1) reduces to

$$f^m(\boldsymbol{x}) = f^{m-1}(\boldsymbol{x}) + h^m(\boldsymbol{x}), \tag{2}$$

where $f^0(\boldsymbol{x}) = 0$ and $f(\boldsymbol{x}) := f^M(\boldsymbol{x})$ is the score function at the highest fidelity level which we treat as ground truth. We model the influence of systemic errors in sample collection and environmental uncertainty on the output of the computer vision algorithm for an input at fidelity level m through an additive zero mean Gaussian random variable ϵ_m with variance s_m^2, i.e., $\epsilon_m \sim N(0, s_m^2)$. Consequently, the (scaled) score obtained by collecting a sample at location $\boldsymbol{x}$ is a random variable $y = f_m(\boldsymbol{x}) + \epsilon_m$.

We assume that each h_m is a realization of a Gaussian process with a constant mean μ_m and a squared exponential kernel function $k^m(\boldsymbol{x}, \boldsymbol{x}')$ expressed as

$$k^m(\boldsymbol{x}, \boldsymbol{x}') = v_m^2 \exp\left(-\frac{\left\|\boldsymbol{x} - \boldsymbol{x}'\right\|^2}{2l_m^2}\right), \tag{3}$$

where l_m is the length scale parameter, and v_m is the variance parameter that satisfies $v_1 > v_2 > \cdots > v_M$. This kernel function describes the spatial correlation of score function at neighboring locations at each fidelity level. Since the fields

of view are more overlapped at lower fidelity levels, it results in $l_1 > l_2 > \cdots > l_M$.

We make the following assumptions about the highest-fidelity sample. If the target is not in the field of view at $(\boldsymbol{x}, z_M)$, the mean score of the computer vision algorithm $f(\boldsymbol{x})$ is smaller than a threshold $\mathtt{th}$. If a target is at the center of image collected at $(\boldsymbol{x}, z_M)$, $f(\boldsymbol{x}) \geq \mathtt{th} + \Delta$, for some constant $\Delta > 0$. Here, $1/\Delta$ can be viewed as a measure of detection difficulty that depends both on the quality of computer vision algorithm and the environment complexity.

B. Objective of the Search Algorithm

Our objective is to design an algorithm for sequentially determining sampling points that lead to expedited detection and localization of targets within desired misclassification rate $\delta \in (0, 1/2)$. In particular, the algorithm should estimate the region containing targets $D_t \subseteq D$ such that (i) $\forall \boldsymbol{x} \in D_t : \mathbb{P}\left(f(\boldsymbol{x}) < \mathtt{th}\right) \leq \delta$ and (ii) $\forall \boldsymbol{x} \in D \setminus D_t : \mathbb{P}\left(f(\boldsymbol{x}) \geq \mathtt{th} + \Delta\right) \leq \delta$. The requirements about both false alarm and mis-detection rate are set by above two conditions.

Let $t(\Delta, \delta)$ be the total (traveling and sampling) time to finish the search task with misclassification rate smaller than δ. Then, the objective of the algorithm is to determine the sequence of sampling points that minimize $t(\Delta, \delta)$.

III. Expedited Multi-target Search Algorithm

The proposed EMTS algorithm is illustrated in Fig. 1. It operates using an epoch-based structure. In each epoch, sampling and fidelity planner computes a set of sampling points and the path planner optimizes a TSP tour going through those points. The vehicle follows the TSP tour to collect measurements at sampling points and the inference algorithm uses these measurements to update the estimate the score function f. Then, the Bayesian classification uses these estimates to compute an occupancy map of the floor and the region elimination module removes regions with no target with sufficiently high probability from the search space. In the following, we describe each of these modules in detail.

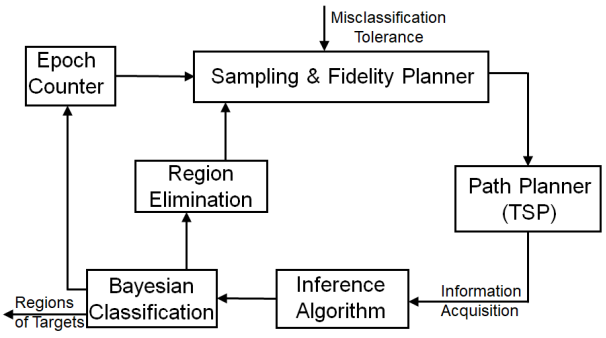


Fig. 1: Architecture of EMTS

A. Inference Algorithm for Multi-fidelity GPs

The Bayesian inference method for multi-fidelity GPs discussed in this section is an extension of the inference procedure in [27] for the case of no sampling noise. Let the set of sampling location-score-fidelity tuples after n observations be $\mathcal{P}_n = \{(\boldsymbol{x}_i, y_i, m_i) \mid i \in \{1, \ldots, n\}\}$. For each fidelity m, define a subset of $\mathcal{P}_n$,

$$P_n^m = \{(\boldsymbol{x}_i, y_i, m_i) \in \mathcal{P}_n \mid m_i = m\},$$

and $|P_n^m|$ denote the cardinality of P_n^m. Recall that $k^i(\boldsymbol{x}, \boldsymbol{x}')$ is the kernel function for the GP h_i at i-th fidelity level. Let $\boldsymbol{K}_0^i\big(P_n^m, P_n^{m'}\big)$ be a $|P_n^m| \times |P_n^{m'}|$ matrix with entries $k^i(\boldsymbol{x}, \boldsymbol{x}')$, $\boldsymbol{x} \in P_n^m$, $\boldsymbol{x}' \in P_n^{m'}$ and $\boldsymbol{K}_0^i(P_n^m, \boldsymbol{x})$ be a $|P_n^m|$ dimensional vector with entries $k_0^i(\boldsymbol{x}', \boldsymbol{x})$, $\boldsymbol{x}' \in P_n^m$. Let $\boldsymbol{K}$ be a $M \times M$ block matrix with $\big(m, m'\big)$ block submatrix

$$\boldsymbol{K}_{m,m'} = \sum_{i=1}^{\min(m,m')} \boldsymbol{K}_i\big(P_n^{(m)}, P_n^{(m')}\big).$$

Let $\boldsymbol{k}(\boldsymbol{x})$ be a $|\mathcal{P}_n|$ dimensional vector constructed by concatenating M sub-vectors $\boldsymbol{k}(\boldsymbol{x}) = \big(\boldsymbol{k}^1(\boldsymbol{x}), \ldots, \boldsymbol{k}^M(\boldsymbol{x})\big)$, where

$$\boldsymbol{k}^m(\boldsymbol{x}) = \sum_{i=1}^{m} \boldsymbol{K}_i(P_n^m, \boldsymbol{x}), \quad \forall m \in \{1, \ldots, M\}. \tag{4}$$

Denoted by $\boldsymbol{\Theta}$ is the $M \times M$ diagonal matrix with variance of sampling noise at diagonal entries

$$\boldsymbol{\Theta} = \operatorname{diag}\left\{s_m^2 \boldsymbol{I}_{|P_n^m|}\right\}_{m=\{1,\ldots,M\}}.$$

Let $\boldsymbol{\nu}_n = [\nu_1, \ldots, \nu_n]$ be the a priori mean of the sample $\boldsymbol{y}_n = (y_1, \ldots, y_n)$. In particular, if y_j is a sample at fidelity m, then $\nu_j = \sum_{i=1}^m \mu_i$. The a priori covariance of $\boldsymbol{y}_n$ is $\boldsymbol{K} + \boldsymbol{\Theta}$. In the training process with training dataset $\mathcal{P}_n$, the hyperparameters $\{\mu_m, v_m, l_m, s_m\}_{m=1}^M$ and $\{a_m\}_{m=1}^{M-1}$ in the multi-fidelity GP can be learned by maximizing a log marginal likelihood function $-\frac{1}{2}\log\Big(\det\big(2\pi\left(\boldsymbol{K} + \boldsymbol{\Theta}\right)\big)\Big) - \frac{1}{2}\left(\boldsymbol{y} - \boldsymbol{\nu}_n\right)^T\left(\boldsymbol{K} + \boldsymbol{\Theta}\right)^{-1}\left(\boldsymbol{y} - \boldsymbol{\nu}_n\right)$. Such training can be performed using the GP toolbox [30].

Due to the multi-fidelity structure described in (1) and (2), the prior mean and covariance of f are

$$\mu_0(\boldsymbol{x}) = \sum_{m=1}^{M} \mu_m, \quad k_0(\boldsymbol{x}, \boldsymbol{x}') = \sum_{m=1}^{M} k^m(\boldsymbol{x}, \boldsymbol{x}').$$

When running EMTS with learned hyperparameters, it can be shown that the posterior mean and covariance functions of f after n measurements are

$$\begin{aligned} \mu_n(\boldsymbol{x}) &= \mu_0(\boldsymbol{x}) + \boldsymbol{k}^T(\boldsymbol{x})\left(\boldsymbol{K} + \boldsymbol{\Theta}\right)^{-1}\left(\boldsymbol{y} - \boldsymbol{\nu}_n\right) \\ k_n\left(\boldsymbol{x}, \boldsymbol{x}'\right) &= k_0\left(\boldsymbol{x}, \boldsymbol{x}'\right) - \boldsymbol{k}^T(\boldsymbol{x})\left(\boldsymbol{K} + \boldsymbol{\Theta}\right)^{-1}\boldsymbol{k}(\boldsymbol{x}'). \end{aligned} \tag{5}$$

Note that the posterior variance $\sigma_n^2(\boldsymbol{x}) = k_n(\boldsymbol{x}, \boldsymbol{x})$ is a measure of uncertainty that will be utilized to classify $\boldsymbol{x}$. It should be noted that the measurements collected at different fidelity levels are appropriately scaled in inference (5).

B. Multi-fidelity Sampling & Path Planning

For each epoch j, we seek to design an efficient sampling tour through sampling locations $\{(\boldsymbol{x}_{n_j+1}, z_{n_j+1}), \ldots, (\boldsymbol{x}_{n_{j+1}}, z_{n_{j+1}})\}$ to ensure

$$\max_{\boldsymbol{x} \in D} \sigma_{n_{j+1}}(\boldsymbol{x}) \big/ \max_{\boldsymbol{x} \in D} \sigma_{n_j}(\boldsymbol{x}) \leq \alpha,$$

where n_j is the number of samples collected before the beginning of the j-th epoch and the selection of uncertainty reduction threshold α is discussed in Section III-C.

Notice that the posterior variance update in (5) depends only on the location of the observations $\boldsymbol{y}_n$, but not on the realized value of $\boldsymbol{y}_n$. Therefore, the sequence of sampling location-fidelity tuples can be computed before physically visiting the locations. Such deterministic evolution of the variance has also been leveraged within the context of single-fidelity GP planning to design efficient sampling tours [31].

1) Sampling Point Selection: The vehicle follows a greedy sampling policy at each fidelity level, i.e., at each sampling round the vehicle selects the most uncertain point as the next sampling point

$$\boldsymbol{x}_n = \arg\max_{\boldsymbol{x}\in D} \sigma_{n-1}(\boldsymbol{x}). \tag{6}$$

In the information theoretic view [5], the greedy policy is near optimal in terms of maximizing an appropriate measure of uncertainty reduction.

2) Fidelity Selection: For each sampling point $\boldsymbol{x}_n$, a fidelity level (or sampling altitude) needs to be assigned. We let the vehicle start at fidelity level 1 and successively visit all fidelity levels from the lowest to the highest. Since sampling f^m is not able to reduce the uncertainty about f introduced by the subsequent bias terms $h^{m+1}, \ldots, h^M$, we define the *inaccessible uncertainty* at fidelity level m as $\xi_m = \sum_{i=m+1}^{M} v_i^2$. Accordingly, we define the *accessible uncertainty* about f at fidelity level m by $r_n^m = \max_{\boldsymbol{x}\in D} \sigma_n^2(\boldsymbol{x}) - \xi_m$. The assigned fidelity level to sample point $\boldsymbol{x}_n$ is designed to change from fidelity m to $m+1$ when

$$r_n^m \leq v_{m+1}^2 l_{m+1}^2 / l_m^2.$$

Notice that before the vehicle begins to sample at fidelity level m, $r_n^m \geq v_m^2 \geq v_{m+1}^2 l_{m+1}^2 / l_m^2$, where the second inequality is due to the assumption that $v_m > v_{m+1}$ and $l_m > l_{m+1}$. This ensures that all fidelity levels are visited from the lowest to the highest successively.

3) Path Planning: Since the order of sampling locations does not influence the eventual posterior mean and variance, the path going through the sampling location can be optimized by computing an approximate TSP tour using packages, such as Concorde [32]. Such a tour-based sampling policy allows for energy and time-efficient operation of the vehicle. If all measurements within epoch j are collected at the same fidelity level, the vehicle traverses the TSP tour TSP$(\boldsymbol{x}_{n_j+1}, \ldots, \boldsymbol{x}_{n_{j+1}})$ to collect measurements from sampling points and update posterior distribution of f. Otherwise, a TSP tour each is designed at every fidelity level.

C. Classification and Region Elimination

The classification and elimination of regions follows a confidence-bound-based rule, which has been widely used in pure exploration multi-armed bandit algorithms [33] and robotic source seeking [34]. We extend these ideas to the case of multi-fidelity GP setting considered in this paper.

Conditioned on $\mathcal{P}_n$, the distribution of $f(\boldsymbol{x})$ is Gaussian with mean function $\mu_n(\boldsymbol{x})$ and variance $\sigma_n^2(\boldsymbol{x})$. Let $(L_n(\boldsymbol{x},\varepsilon), U_n(\boldsymbol{x},\varepsilon))$ be the Bayesian confidence interval containing $f(\boldsymbol{x})$ with probability greater than $(1-2\varepsilon)$. Here, the lower confidence bound L_n and upper confidence bound U_n are defined by $L_n(\boldsymbol{x},\varepsilon) = \mu_n(\boldsymbol{x}) - c(\varepsilon)\sigma_n(\boldsymbol{x})$, $U_n(\boldsymbol{x},\varepsilon) = \mu_n(\boldsymbol{x}) + c(\varepsilon)\sigma_n(\boldsymbol{x})$, with $c(\varepsilon) = \sqrt{2\ln\left(1/(2\varepsilon)\right)}$.

Given the desired maximum misclassification rate δ, at the end of epoch j, a location $\boldsymbol{x}$ is classified as *target*, if $L_{n_j}\left(\boldsymbol{x}, \delta/2^j\right) \geq \texttt{th}$, and is added to D_t; while it is classified as *empty*, if $U_{n_j}\left(\boldsymbol{x}, \delta/2^j\right) < \texttt{th}$, and is added to the set D_e. Note that the confidence parameter $\varepsilon = \delta/2^j$ defining the lower and upper bounds is decreased exponentially with epochs, and we will show that it ensures a misclassification rate smaller than δ. The locations in the set D_e are removed from sampling space D at the end of each epoch. EMTS is terminated if $\max_{\boldsymbol{x}\in D} 2\sigma_{n_j}(\boldsymbol{x}) \leq \Delta / c(\delta/2^j)$.

The selection of α depends on the balance between the efficiency of TSP path planer and region elimination. TSP path planer is more effective with smaller α since each exploration tour includes more sample points. While region elimination favors bigger α so that regions not likely to contain targets are removed more frequently.

IV. An Illustrative Example

In this section, we illustrate EMTS using the Unmanned Underwater Vehicle Simulator [35], which is a ROS package designed for Gazebo robot simulation environment. We integrate it with YOLOv3 [29] for image classification and Concorde solver [32] to compute TSP tours. We use 2 fidelity levels situated at 11m and 5m from the water floor, respectively. Fig. 2 shows our simulation setup, where 3 victims are located at different locations on a $40\text{m} \times 40\text{m}$ water floor. At each sampling point, the vehicle take 20 images and YOLOv3 returns an average score about the confidence level of the existence of victims in the view.

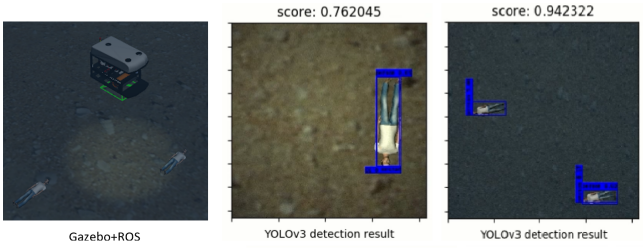
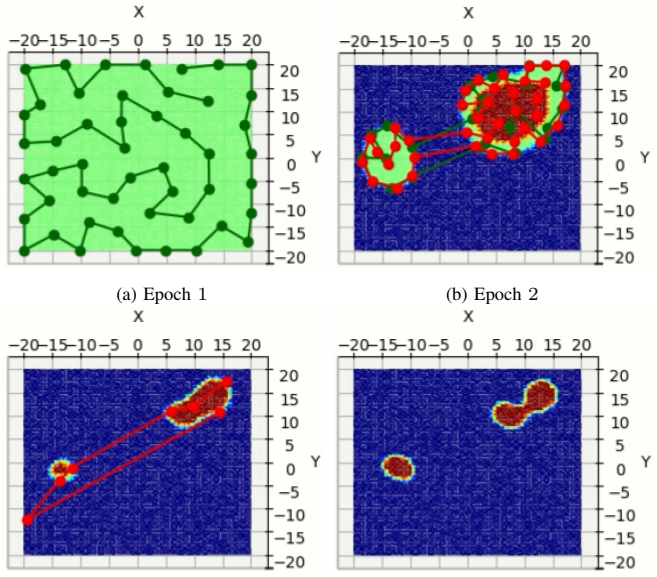


Fig. 2: Test Environment: An underwater vehicle is equipped with a downward camera and a flash light to facilitate the searching task in dark underwater environment. Middle figure and right figure are detection result with YOLOv3 at a high fidelity level and a low fidelity level, respectively.

The first three subplots of Fig. 3 shows the classification of regions before each epoch, the sampling points selected by the greedy policy and the planned path. Classifications of the environment are represented by 3 colors: red means target exist, blue means no target, and green means uncertain. The dark green points and lines are the planned sampling locations and paths at the low fidelity level and red points and lines are sampling locations and paths at the high fidelity



(a) Epoch 1 (b) Epoch 2

(c) Epoch 3 (d) Final result

Fig. 3: Performance of EMTS. (i) The green dots and lines are sampled locations and the path traversed by the vehicle at the low fidelity level and the red ones are for the high fidelity level. (ii) Classification results of the environment are represented by 3 colors: red means target exist, blue means no target, and green means uncertain. (iii) The vehicle switches to high fidelity level at epoch 2.

level. At the beginning of epoch 1, all regions are classified as uncertain. After each epoch, the region of targets is narrowed down. The search task is terminated after three epochs. Notice that the vehicle switches to the high fidelity level at epoch 2. The tours at low and high fidelity levels are plotted using two different colors. The vehicles do not sample in blue regions since they have been classified as empty. In the final result, the regions with target are successfully found. A video of the simulation is available online[1].

In Fig. 4a, we show the heat map of posterior variance for the whole region. The regions classified as empty have larger posterior variance since they have been eliminated from sampling space. This shows that EMTS is able to put more focus on areas likely to contain victims. The uncertainty reduction, i.e. the decrease in maximum posterior variance, for multi-fidelity greedy sampling and single-fidelity greedy sampling, are compared in Fig. 4b. It shows that greedy multi-fidelity sampling can reduce uncertainty much faster at the beginning stage, which will enable EMTS to eliminate unoccupied regions quickly, and hence, accelerate target search.

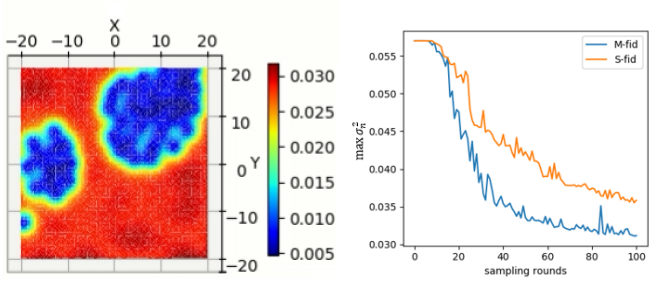
V. Analysis of EMTS

In this section, we present analyze accuracy and efficiency of EMTS. Detailed proofs are presented in the extended version of this paper [36].

A. Analysis of the classification algorithm

The Bayesian confidence interval classification rule is able to provide the following accuracy guarantee about EMTS.

[1]https://mediaspace.msu.edu/media/EMTS/1_phbul7ui



(a) Final posterior variance (b) Convergence of σ_n^2

Fig. 4: Uncertainty reduction results. (a) shows spatial nature of uncertainty reduction with EMTS, i.e., the posterior variance is low only at areas that likely contain a target. (b) shows the temporal nature of uncertainty reduction by comparing the decreasing speed of posterior variance with multi-fidelity greedy sampling and single fidelity greedy sampling.

*Theorem 1 (**Misclassification Rate**):* The EMTS algorithm detects regions $D_t \subseteq D$ with both false alarm and miss-detection rate upper bounded by δ.

B. Analysis of Detection Time

To analyze target search time, we first provide an upper bound on the number of samples $n(\Delta, \delta)$ required by EMTS to meet with the misclassification tolerance δ.

*Lemma 2 (**Sample complexity for EMTS**):* For a given misclassification rate δ, the number of samples satisfies

$$n(\Delta, \delta) \in O\left(d^2 \varphi(\Delta, \delta) \left(\ln \varphi(\Delta, \delta)\right)^3\right),$$

where $\varphi(\Delta, \delta) = \frac{\sigma_0^2}{\Delta^2} \ln\left(\frac{3\sigma_0}{\delta\Delta}\right)$ and d is the diameter of D.

Note this upper bound is similar to the sample complexity [37] in a pure-exploration multi-armed bandit problem. For kernels characterizing less correlations, e.g. Martén kernels, more sampling rounds are expected.

The above sample complexity result considers only the constant time required to collect the sample once the robot is present at the location. The target search time includes the sampling time as well as the travel time. Since EMTS requires the vehicle to search from low fidelity level to high fidelity level, the total number of altitude switches is no greater than $M - 1$. Besides, there exist a constant number of epochs before EMTS is terminated. As presented in [38], for n points in $[0, 1]^2$, the length of the shortest TSP Tour is upper bounded by $0.984\sqrt{2n} + 11$. Therefore, the traveling time belongs to $O\big(d\sqrt{n(\Delta, \delta)}\big)$. Accordingly, we get the following upper bound on target search time for EMTS.

*Theorem 3 (**Target search time for EMTS**):* For a given misclassification tolerance δ, the target search time satisfies

$$t(\Delta, \delta) \in O\left(d^2 \varphi(\Delta, \delta) \left(\ln \varphi(\Delta, \delta)\right)^3\right),$$

This upper bound has a natural implication that the target search time increases with the detection difficulty $1/\Delta$ and the desired classification accuracy $1 - \delta$.

VI. Conclusions and Future Directions

In this paper, we extended the classical informative path planning approach for single-fidelity GPs to multi-fidelity GPs. This novel extension allowed for jointly planning for

sampling locations and associated fidelity-levels, and thus, addressed the fidelity-coverage trade-off. We proposed and analyzed the EMTS algorithm for multi-target search that yields sampling points that the robot should visit and the fidelity level with which the robot should collect the information at these points. We illustrated our algorithm in an underwater victim search scenario using the Unmanned Underwater Vehicle Simulator. We rigorously analyzed the algorithm in terms of its accuracy in classifying the locations in the environment as empty or occupied by a target, as well as the time the robot takes to detect targets.

Future research includes the extension to cooperative multi-robot search scenarios and implementation of the proposed algorithm in our underwater multi-target search testbed.

REFERENCES

[1] N. E. Leonard, D. A. Paley, F. Lekien, R. Sepulchre, D. M. Fratantoni, and R. E. Davis, "Collective motion, sensor networks, and ocean sampling," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 48–74, 2007.

[2] S. L. Smith, M. Schwager, and D. Rus, "Persistent robotic tasks: Monitoring and sweeping in changing environments," *IEEE Transactions on Robotics*, vol. 28, no. 2, pp. 410–426, 2012.

[3] C. G. Cassandras, X. Lin, and X. Ding, "An optimal control approach to the multi-agent persistent monitoring problem," *IEEE Transactions on Automatic Control*, vol. 58, no. 4, pp. 947–961, 2013.

[4] R. N. Smith, M. Schwager, S. L. Smith, B. H. Jones, D. Rus, and G. S. Sukhatme, "Persistent ocean monitoring with underwater gliders: Adapting sampling resolution," *Journal of Field Robotics*, vol. 28, no. 5, pp. 714–741, 2011.

[5] A. Krause and C. E. Guestrin, "Near-optimal nonmyopic value of information in graphical models," in *Proceedings of the Twenty-First Conference Conference on Uncertainty in Artificial Intelligence*, Edinburgh, Scotland, Jul. 2005, pp. 324–331.

[6] C. K. Williams and C. E. Rasmussen, *Gaussian processes for Machine Learning*. MIT press Cambridge, MA, 2006, vol. 2, no. 3.

[7] S. Vasudevan, F. Ramos, E. Nettleton, and H. Durrant-Whyte, "Gaussian process modeling of large-scale terrain," *Journal of Field Robotics*, vol. 26, no. 10, pp. 812–840, 2009.

[8] A. Singh, A. Krause, C. Guestrin, and W. J. Kaiser, "Efficient informative sensing using multiple robots," *Journal of Artificial Intelligence Research*, vol. 34, no. 2, p. 707, 2009.

[9] A. Krause, A. Singh, and C. Guestrin, "Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies," *Journal of Machine Learning Research*, vol. 9, no. Feb, pp. 235–284, 2008.

[10] J. L. Ny and G. J. Pappas, "On trajectory optimization for active sensing in Gaussian process models," in *IEEE Conf on Decision and Control and Chinese Control Conference*, Shanghai, China, Dec. 2009, pp. 6286–6292.

[11] X. Lan and M. Schwager, "Planning periodic persistent monitoring trajectories for sensing robots in Gaussian random fields," in *IEEE Int Conf on Robotics and Automation*, Karlsruhe, Germany, May 2013, pp. 2415–2420.

[12] D. E. Soltero, M. Schwager, and D. Rus, "Generating informative paths for persistent sensing in unknown environments," in *IEEE/RSJ Int Conf on Intelligent Robots and Systems*, Vilamoura, Algarve, Portugal, Oct. 2012, pp. 2172–2179.

[13] J. Yu, M. Schwager, and D. Rus, "Correlated orienteering problem and its application to informative path planning for persistent monitoring tasks," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2014, pp. 342–349.

[14] V. Srivastava, F. Pasqualetti, and F. Bullo, "Stochastic surveillance strategies for spatial quickest detection," *The International Journal of Robotics Research*, vol. 32, no. 12, pp. 1438–1458, 2013.

[15] V. Srivastava, P. Reverdy, and N. E. Leonard, "Surveillance in an abruptly changing world via multiarmed bandits," in *IEEE Conference on Decision and Control*, 2014, pp. 692–697.

[16] G. A. Hollinger and G. S. Sukhatme, "Sampling-based robotic information gathering algorithms," *The International Journal of Robotics Research*, vol. 33, no. 9, pp. 1271–1287, 2014.

[17] G. A. Hollinger, B. Englot, F. S. Hover, U. Mitra, and G. S. Sukhatme, "Active planning for underwater inspection and the benefit of adaptivity," *The International Journal of Robotics Research*, vol. 32, no. 1, pp. 3–18, 2013.

[18] G. Hitz, E. Galceran, M.-È. Garneau, F. Pomerleau, and R. Siegwart, "Adaptive continuous-space informative path planning for online environmental monitoring," *Journal of Field Robotics*, vol. 34, no. 8, pp. 1427–1449, 2017.

[19] G. Hitz, A. Gotovos, M.-É. Garneau, C. Pradalier, A. Krause, R. Y. Siegwart *et al.*, "Fully autonomous focused exploration for robotic environmental monitoring," in *IEEE International Conference on Robotics and Automation*, 2014, pp. 2658–2664.

[20] N. Atanasov, J. Le Ny, K. Daniilidis, and G. J. Pappas, "Information acquisition with sensing robots: Algorithms and error bounds," in *IEEE International Conference on Robotics and Automation*, 2014, pp. 6447–6454.

[21] A. A. Meera, M. Popović, A. Millane, and R. Siegwart, "Obstacle-aware adaptive informative path planning for UAV-based target search," in *International Conference on Robotics and Automation*, 2019, pp. 718–724.

[22] Y. Sung, D. Dixit, and P. Tokekar, "Environmental hotspot identification in limited time with a uav equipped with a downward-facing camera," *arXiv preprint arXiv:1909.08483*, 2019.

[23] N. Srinivas, A. Krause, S. M. Kakade, and M. W. Seeger, "Information-theoretic regret bounds for gaussian process optimization in the bandit setting," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 3250–3265, 2012.

[24] P. Reverdy, V. Srivastava, and N. E. Leonard, "Modeling human decision making in generalized Gaussian multiarmed bandits," *Proceedings of the IEEE*, vol. 102, no. 4, pp. 544–571, 2014.

[25] S. Chen, T. Lin, I. King, M. R. Lyu, and W. Chen, "Combinatorial pure exploration of multi-armed bandits," in *Advances in Neural Information Processing Systems*, 2014, pp. 379–387.

[26] P. Reverdy, V. Srivastava, and N. E. Leonard, "Satisficing in multi-armed bandit problems," *IEEE Transactions on Automatic Control*, vol. 62, no. 8, pp. 3788 – 3803, 2017.

[27] M. C. Kennedy and A. O'Hagan, "Predicting the output from a complex computer code when fast approximations are available," *Biometrika*, vol. 87, no. 1, pp. 1–13, 2000.

[28] K. Kandasamy, G. Dasarathy, J. B. Oliva, J. Schneider, and B. Póczos, "Gaussian process bandit optimisation with multi-fidelity evaluations," in *Advances in Neural Information Processing Systems*, 2016, pp. 992–1000.

[29] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.

[30] P. Perdikaris, "Gaussian processes a hands-on tutorial," 2017. [Online]. Available: https://github.com/paraklas/GPTutorial

[31] S. Kemna, J. G. Rogers, C. Nieto-Granda, S. Young, and G. S. Sukhatme, "Multi-robot coordination through dynamic Voronoi partitioning for informative adaptive sampling in communication-constrained environments," in *IEEE International Conference on Robotics and Automation*, 2017, pp. 2124–2130.

[32] D. Applegate, R. Bixby, V. Chvatal, and W. Cook, "Concorde TSP solver," 2006.

[33] J.-Y. Audibert and S. Bubeck, "Best arm identification in multi-armed bandits," in *Conference on Learning Theory*, Haifa, Israel, June 2010, pp. 13–p.

[34] E. Rolf, D. Fridovich-Keil, M. Simchowitz, B. Recht, and C. Tomlin, "A successive-elimination approach to adaptive robotic source seeking," *arXiv preprint arXiv:1809.10611*, 2018.

[35] M. M. M. Manhães, S. A. Scherer, M. Voss, L. R. Douat, and T. Rauschenbach, "UUV simulator: A Gazebo-based package for underwater intervention and multi-robot simulation," in *MTS/IEEE Oceans*, Monterey, CA, 2016, pp. 1–8.

[36] L. Wei, X. Tan, and V. Srivastava, "Expedited multi-target search with guaranteed performance via multi-fidelity Gaussian processes," *arXiv preprint arXiv:2005.08434*, 2020.

[37] S. Mannor and J. N. Tsitsiklis, "The sample complexity of exploration in the multi-armed bandit problem," *Journal of Machine Learning Research*, vol. 5, no. Jun, pp. 623–648, 2004.

[38] H. J. Karloff, "How long can a Euclidean traveling salesman tour be?" *SIAM Journal on Discrete Mathematics*, vol. 2, no. 1, pp. 91–99, 1989.