Obstacle-aware Adaptive Informative Path Planning for UAV-based Target Search

Ajith Anil Meera, Marija Popović, Alexander Millane and Roland Siegwart

Abstract—The autonomous target search problem for Unmanned Aerial Vehicles (UAV) in urban environments requires solving a 3D path planning problem for maximal information gain, given a restricted flight duration. In this paper, we propose a general, Obstacle-aware Adaptive Informative Path Planning (OA-IPP) algorithm for the target search problem which uses active perception. The main contribution is the layered optimization approach that balances the exploration-exploitation trade-off through a Bayesian Optimization (BO) framework and simultaneously optimizes the 3D path using a standard optimizer. The planner simultaneously trades off between information gain, field coverage, altitude-dependent sensor performance, collision avoidance, target re-observation and Field of View (FoV) while planning. Through experiments in a simulated environment, we show that the proposed approach outperforms a pure exploratory IPP planner, a coverage planner, and a random sampling planner by demonstrating the fastest decrease in error related to target position estimates. Furthermore, we demonstrate the planner in simulations of varying complexity and obstacle density, demonstrating its applicability to a range of environments. Finally, we combine the proposed planning approach with an existing human detection pipeline, and demonstrate its efficacy in locating human victims in a realistic simulated environment.

I. INTRODUCTION

Autonomous target search in a cluttered environment is a challenging problem that has the potential to solve a wide range of problems. Example applications include finding victims in search and rescue operations [1], detecting weeds in precision agriculture [2], patrolling borders for military, and automated search for endangered species. Recently, UAVs have emerged as an attractive robotic platform for this purpose because they provide the capability for rapid observation of large areas, are highly maneuverable and can provide measurements with high spatial and temporal resolution.

The deployment of a UAV for target search imposes certain challenges including limited battery life, obstacle collision, deterioration of sensor performance with altitude, occlusion of the ground by buildings etc. The existing research gap for an algorithm that can simultaneously handle these challenges, motivates our work. We treat the target search as an optimization process that solves the IPP problem by seeking maximal information gain with minimal flight time and obstacle collision. Our IPP framework consists of three major obstacle-aware components: 1) modelling of the environment, sensor and ground field, 2) mapping, and 3) planning. The core contributions of this work are:

1) A target search algorithm for UAV which couples informative planning with obstacle awareness.
2) Introduction of a layered optimization approach that binds the BO based exploration-exploitation strategy and the altitude-dependent sensor performance into an objective function that can be optimized for a continuous 3D path.
3) Extensive evaluation of our framework in simulation and its validation in a realistic urban search and rescue scenario.

II. RELATED WORK

Target search problem has been used widely to solve real-world problems [3, 4]. Due to the information seeking nature of IPP, it is widely used to solve the target search problem. The problem can also be formulated as a Partially Observable Markov Decision Process (POMDP) and can be solved using general POMDP solvers through policy learning [1]. In general, these methods are computationally expensive for real-life scenarios.

IPP literature can broadly be classified based on three properties: adaptivity, myopicity and continuity. Adaptive IPP [5] differs from non-adaptive IPP [6] in terms of the role of measurements on the planning process. The former plans the path online, while the later plans the path a priori without considering the measurement updates. Myopic methods [7] differ from non-myopic ones [2] in terms of the greedy information search. The former greedily plans the path, while the later plans with a finite look ahead. Discrete IPP [6] considers distinct measurement locations in the environment and solves the problem using combinatorial methods like branch and bound technique [8], while continuous IPP solves the problem as a continuous path planning problem. Discrete methods have poor scalability and are limited by resolution. Our algorithm is adaptive, non-myopic and continuous.

Sampling based algorithms like RIG [9] that are based on randomly exploring random trees have also shown success in target search. Recent methods rely on optimizing a polynomial trajectory for maximal information gain using standard optimizers like the CMA-ES [2, 5] and BO [10]. Our algorithm falls into this category. We use a sampling based initial myopic solution to optimize the continuous trajectory that finally yields an adaptive, non-myopic and continuous polynomial trajectory.

Accounting for sensor uncertainty and facilitating target re-observation is crucial in estimating an error free target map. Despite a large body of prior work, there is a lack of an obstacle-aware algorithm that accounts for these factors.
Although Dang et al. [11] and Marchant and Ramos [12] plan for target re-observation, they do not account for altitude dependent sensor performance. Popovic et al. [2] accounts for the sensor uncertainties, but they do not consider obstacles. This work aims to fill this research gap.

III. PROBLEM STATEMENT

The target search problem is formulated as an IPP problem that aims at simultaneously maximizing the information gain \( O_{info}(T) \) and minimizing the collision cost \( C_{coll}(T) \) and flight time \( t_{flight}(T) \), where \( T \) is the trajectory. We evaluate an optimal continuous path \( T^* \) out of all possible paths \( \psi \) as:

\[
T^* = \arg \max_{T \in \psi} \frac{k_1 O_{info}(T) - k_2 C_{coll}(T)}{t_{flight}(T)},
\]

where \( k_1 \) and \( k_2 \) are non-negative constants. Setting \( k_2 \) high prioritize collision avoidance. The UAV uses the resultant trajectory \( T^* \) for taking discrete measurements at a measurement frequency.

IV. MODELLING

This section describes the modelling approaches we use to evaluate the objective function given in Equation 1: \( C_{coll}(T) \) through environment modelling and \( O_{info}(T) \) through field and sensor modelling.

A. Environment modelling

An urban environment comprises of buildings that are treated as obstacles with standard geometric shapes inside the planner. We evaluate the Euclidean Signed Distance Function (ESDF map) [13], which is an environment representation to check for possible obstacle collisions at any given point in the map. We exploit the hard constraint artificial potential field given by Equation 2 to define the penalty for obstacle collision given by Equation 3, where \( x_{UAV} \) is the radius of the smallest sphere that can contain the UAV and \( x \) is any point on the trajectory \( T \).

\[
C_{hard}(x) = \begin{cases} 0 & \text{if } ESDF(x) \geq \frac{x_{UAV}}{2}, \\ 1 & \text{otherwise}. \end{cases}
\]

\[
C_{coll}(T) = \sum_{x \in T} C_{hard}(x)
\]

\( C_{coll}(T) \) is non-zero if any point \( x \in T \) is inside the obstacle, thereby incurring a penalty to the objective function.

B. Modelling the field

We model the ground field of target occupancy using a Gaussian Process (GP) which can be used to model spatial correlations in the map in a probabilistic and non-parametric way [14]. The target map is assumed to be a continuous function in 2D space \( \zeta : \mathbb{R} \rightarrow \mathbb{R} \), where \( \mathbb{R} \subset \mathbb{R}^2 \) is a point on the map \( \zeta \) under consideration. GP is fully characterized by the mean \( \mu = E[\zeta] \) and the covariance \( P = E[(\zeta - \mu)(\zeta^T - \mu^T)] \). The continuous 2D function of the environment being mapped can therefore be expressed as \( \zeta \sim GP(\mu, P) \).

A constant low initial prior is set for the field that is discretized at certain resolution with \( n \) training locations \( X \in \varepsilon \), assuming a sparse distribution of targets in the field. The isotropic Matern 3/2 kernel function [14] common in geostatistical analysis and spatial statistics is used to describe the field. Each element of the kernel \( K(X, X) \) is given by Equation 4, where \( l \) and \( \sigma_n^2 \) are the hyperparameters representing the lengthscale and signal variance respectively, and \( d \) is the Euclidean distance between inputs locations \( x \) and \( x' \).

\[
k_{Mat3}(x, x') = \sigma_f^2 \left( 1 + \sqrt{3}d/l \right) \exp \left( -\sqrt{3}d/l \right), \quad (4)
\]

The covariance is then evaluated using the classic GP regression equation:

\[
P = K(X', X') - K(X', X)[K(X, X) + \sigma_n^2 I]^{-1} K(X', X)^T, \quad (5)
\]

where \( K(X, X') \) denotes the \( n \times n' \) matrix of the covariances evaluated at all pairs of training and test points, \( P \) is the posterior covariance, and \( \sigma_n^2 \) is a hyperparameter representing the noise variance. The diagonal elements of covariance matrix are the variances of each point on the map. Therefore, trace of the covariance matrix \( P \) represents the total variance or uncertainty of the map. With new measurement updates, the total GP variance is expected to decrease.

C. Sensor modelling

State-of-the-art target detectors employ complex neural networks for their internal architecture. We use You Look Only Once (YOLO) [15], the general object detector as a human target detector. The performance of the classifier was analyzed on images captured at different altitudes by the UAV in a simulated Gazebo environment. Figure 1 shows the altitude-dependent performance of YOLO in detecting a human victim on the field by using F1 score as the performance metric. It can be concluded that the sensor prefers certain altitude \( h_{opt} \) for maximal performance, while the classifier exhibits poor performance beyond a saturation height \( h_{sat} \). However, maximum altitude is preferred for collision avoidance. Therefore, a trade-off between sensor
performance and collision avoidance should be embedded into the planner.

V. MAPPING

The measurement output from YOLO is fused onto the GP map in two steps: FoV estimation and data fusion. The estimated FoV at the measurement location of the UAV is used to select the region in the YOLO output that is to be fused onto the map. The field is then updated using recursive filtering based on Kalman Filter (KF). Given a uniform initial mean \( \mu^- \) and spatial correlations \( P^- \) given in Equation 5, the map \( p(\zeta|X) \sim GP(\mu^-, P^-) \) is used as a prior onto which new sensor measurements are fused. Let \( z = [z_1, ..., z_m]^T \) denote the \( m \) independent measurements for the human occupancy given by YOLO at all the points \( [x_1, ..., x_m]^T \subset X \) in the FoV. The maximum a posteriori estimator is used to fuse the measurements \( z \) with the prior map \( p(\zeta|X) \), as formulated as:

\[
\arg \max_\zeta p(\zeta|z, X)
\]  

The posterior density \( p(\zeta|z, X) \propto p(z|\zeta, X) \times p(\zeta|X) \sim GP(\mu^+, P^+) \) is computed directly by applying the KF update equations [16]:

\[
\mu^+ = \mu^- + K v
\]

\[
P^+ = P^- - K H P^-,
\]

where \( K = P^- H^T S^{-1} \) is the Kalman gain and \( v = z - H \mu^- \) and \( S = H P^- H^T + R \) are the measurement and covariance innovations. \( H \) is an \( m \times n \) matrix representing a linear sensor model that intrinsically selects the FoV \( \{\zeta_1, ..., \zeta_m\} \) observed through \( z \), and \( R \) is a diagonal \( m \times m \) matrix of the height dependent variances \( \sigma^2_{z,i} \) associated with each measurement \( z_i \). Therefore, \( H \) controls the measurement, and \( R \) controls the influence of these measurement on map updation. The height-dependent variance model \( \sigma^2_{z,i} \) given in [2] is used for map updation. If the image is captured at low altitudes, the associated uncertainty in measurement is lower, making a higher influence on the map updation. An image captured at high altitudes has a lower influence on the map updation due to the high uncertainty associated with the measurement.

VI. PLANNING

This section describes the main contribution of the paper by proposing an adaptive planning algorithm for autonomous data acquisition.

A. Trajectory parametrization

A polynomial trajectory \( \psi \) is parameterized by a sequence of \( N \) control waypoints to be visited by the UAV, defined as \( C = [c_1, ..., c_N] \), where the first waypoint \( c_1 \) represents the current UAV location. The polynomial trajectory connects these control points using \( N - 1 \) \( k \)-order spline segments for minimum-snap dynamics as given in [17]. The trajectory \( \psi \) represents the fixed-horizon path for the UAV for data gathering.

B. Uncertainty reduction as the objective

Many IPP approaches use uncertainty reduction as the objective for the IPP problem without using the target detections for planning [2, 5]. The information gain can be represented as

\[
O_{in fo}(T) = Tr(P^-) - Tr(P^+),
\]

where \( Tr(.) \) denotes the trace of the matrix, \( P^- \) is the prior covariance, and \( P^+ \) is the posterior covariance given by Equation 8, evaluated after all the measurements along the trajectory \( T \). However, the planner is explorative and does not take target detection into account while planning.

C. Layered optimization for adaptive planning

In this section we propose a layered optimization approach that facilitates target re-observation, rendering the planner with exploration-exploitation trade-off capability, which is crucial for robustness against wrong detections. We define an information-theoretic objective function based on the acquisition function that is used in the BO framework. We use the Upper Confidence Bound (UCB) [18] defined by Equation 10 as the acquisition function.

\[
UCB(x) = \mu(x) + \kappa \sigma(x),
\]

Here \( \kappa \) is the exploration-exploitation tuning parameter, and \( \mu \) and \( \sigma^2 \) are the mean and variance of the GP. We define Acquisition View (AV) as the sum total of acquisition function values within the FoV, given by

\[
AV = \sum_{x \in FoV} UCB(x).
\]

The sensor performance curve given in Figure 1 is modelled as a normal distribution \( N(h_{opt}, \sigma_1) \), and is used as a reward function for the objective function. The information gain is then defined as the combination of sensor performance and AV as

\[
O_{inf o} = \begin{cases} 
AV \times \left[ \frac{1}{\sigma_1 \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{h-h_{opt}}{\sigma_1} \right)^2} \right] & h < h_{sat}, \\
0 & \text{otherwise}.
\end{cases}
\]

The AV increases with altitude, while the sensor performance follows the pattern in Figure 1. Therefore, this layer of initial optimization based on BO given by Equation 12 encodes the trade-off between sensor performance and FoV by facilitating target re-observation through a balanced exploration-exploitation strategy that is embedded within the UCB. This objective contributes to the next optimization layer to evaluate the optimal 3D path.

D. Optimization

The trajectory optimization is performed in two steps: evaluation of initial coarse myopic solution followed by the standard optimization routine. Based on the current UAV pose \( R_0 \), the set of \( N - 1 \) control waypoints for the trajectory are estimated, which acts as an initial myopic
Algorithm 1 Function to perform the greedy search.

1: function COARSEGREEDYSEARCH($R_0$)
2: Initialize empty queue $C_{init}$.
3: Insert $R_0$ into $C_{init}$.
4: Initialize $GP_{sim}$ to $GP$.
5: for $i = 1 : N - 1$ do
6: \hspace{1em} $nbvp = NextBestViewPoint(C_{init}[i], GP_{sim})$
7: \hspace{1em} Insert $nbvp$ into $C_{init}$.
8: \hspace{1em} Update $GP_{sim}$ map at $nbvp$.
9: return $C_{init}$

Solution $C_{init}$ for the optimizer. The control waypoints are the next best viewpoints which are estimated by randomly sampling points from the environment at different altitudes and greedily maximizing the IPP objective given in Equation 1, assuming a measurement update at these locations. Next best points are sampled such that the line joining it and the previous waypoint in $C_{init}$ is collision free. The greedy search algorithm is given in Algorithm 1, while the target search algorithm using IPP is given in Algorithm 2.

Algorithm 2 Algorithm to solve the IPP problem for target search using a UAV.

1: Create ESDF map
2: Initialize the $GP$
3: if $t_{flight} < Budget$ then \hspace{1em} Replan
4: \hspace{1em} $C_{initial} = CoarseGreedySearch(R_0)$.
5: \hspace{1em} $C = CMAES(C_{init})$
6: \hspace{1em} while not isCollisionFreePath($C$, ESDF) do
7: \hspace{2em} Fly along the path $C$
8: \hspace{2em} Take measurements
9: \hspace{2em} Update $GP$ map at measurement locations
10: \hspace{2em} Update current UAV pose $R_0$
11: else
12: \hspace{1em} Land the UAV.

VII. EXPERIMENTAL RESULTS

This section deals with performing a number of simulated flight experiments and benchmarking our planner. Experiments on wide range of environments generalize our planner with respect to the environmental complexity.

A. Benchmarking planner

The environment ($30m \times 30m \times 26m$) shown in Figure 2 with an obstacle of size ($4m \times 10m \times 26m$) is used for benchmarking our planner. The field is discretized such that each cell in the GP map is of dimensions $0.75m \times 0.75m$. The flight parameters include a fixed flight time budget of 150s, reference speed of $5m/s$ and reference acceleration of $3m/s^2$ [17]. The camera FoV is set to $(45^\circ, 60^\circ)$.

We use the Root Square Error (RSE) as the evaluation metric for the experiments. It is defined by

$$RSE = \sum_{\forall \epsilon \in M} \left[ M(\epsilon) - GT(\epsilon) \right]^2,$$

where $\epsilon$ is a grid cell in the map $M$, and $GT$ is the ground truth map of human occupancy. RSE equally weighs the false positive and false negative target detection in the map. Therefore, the best planner should ideally decrease the RSE at the fastest rate.

Results of 25 flight simulations, one using variance reduction as the information objective $O_{info}$ given by Equation 9, while the other using the $O_{info}$ given by Equation 12, are shown in Figure 3. Both the planners can be observed to decrease the RSE at the same rate at the beginning during which the UAV is performing coverage of the field (exploration). Afterwards, the acquisition based planner concentrates on exploitation and flies above the lower half of the field, while the variance based planner purely performs exploration and circles above the whole field. This results in a superior performance of our acquisition based planner through re-observation of the target. Exploitation renders the planner with robustness against wrong target detection, which makes the search feasible even with a sensor with
relatively poor performance. Upon wrong detection, the re-
observation strategy helps in correcting the map.

Our planner is then benchmarked against coverage planner,
which is one of the state-of-the-art planner [19], and random
sampling planner. Coverage planner demonstrated the best
performance at an altitude of 10m - in adherence to the $h_{opt}$
in Figure 1 - and was used for benchmarking. It can be
observed from Figure 4 that our planner outperforms other
planners in reducing the map error at the fastest rate.

![Figure 4: Benchmarking the planner.](image)

**B. Environment complexity**

This section aims at generalizing the OA-IPP planner
with respect to environment complexity. Two types of envi-
ronments are considered: one representing cities of historic
significance containing low height buildings like in Zurich,
while the other representing modern cities with tall skyscrap-
ers like in New York City. A UAV is capable of flying over
buildings in Zurich with good sensor performance. However,
this is not possible in cities with skyscrapers due to the low
$h_{sat}$ of the sensor. We consider both these scenarios and
vary the obstacle density in the environment by randomly
changing the location of the obstacles at the beginning of
each test flight.

![Figure 5: Environment complexity - small buildings](image)

Results of 100 flight simulations with randomly placed
cuboidal obstacles of dimension $4m \times 4m \times 13m$ in the
environment at different obstacle density are shown in Figure
5. It can be observed that the planner is capable of decreasing
the variance of the field, despite the increasing obstacle
density. The same experiment was repeated with an obstacle
size of $4m \times 4m \times 26m$ and the result is shown in Figure
6. It can be observed that the planner is still capable of
reducing the variance of the field. However, the increasing
obstacle density occludes the ground field more for the UAV,
resulting in an increased difficulty to map. Therefore, it can
be concluded that the planner is employable for a wide range
of environments and obstacle complexities.

![Figure 6: Environment complexity - tall buildings](image)

**C. Realistic simulation**

This section deals with providing a realistic simulation of
our algorithm in a search and rescue scenario in an urban
environment shown in Figure 8, with a ground truth for the
human occupancy as shown in Figure 7a. The ground truth
can be compared to the humans in top view of Figure 8. The
simulation was performed with the same flight parameters
from the previous simulations, and the result is plotted in
Figure 7b. It can be observed that the algorithm is successful
in finding all the 7 victims lying on the ground after flying
for 150s, demonstrating the success of our algorithm despite
multiple false detections. The final map is crucial and should
be accurate because it is directly used by the search and
rescue team.

![Figure 7: Realistic simulation results.](image)
VIII. CONCLUSIONS AND FUTURE WORK

The paper introduced a novel obstacle-aware IPP algorithm that is applicable for general target search problems using a UAV. The planner simultaneously trades off between coverage, obstacle avoidance, target re-observation, altitude dependent sensor performance, flight time and FoV to generate the optimal, finite-horizon 3D polynomial path in an obstacle filled environment. The proposed layered optimization approach facilitates a balanced exploration-exploitation strategy which makes it robust against false detections. Extensive simulations show that our planner outperforms purely exploratory IPP planner, coverage planner and random sampling planner in terms of search efficiency. The algorithm was generalized for a wide range of environment complexities and obstacle densities. It successfully found all the humans on the ground in a realistic search and rescue simulation, despite multiple false human detections by the sensor, demonstrating the robustness provided by the layered optimization approach.

The main drawback with the planner is the underlying assumption of a known and static environment. Moreover, a non-temporal field was assumed for target occupancy. Future research will investigate planning with dynamic obstacles and fields. Multi-UAV collaboration for large area search is also a promising direction for future work.

REFERENCES


