

Star-Searcher: A Complete and Efficient Aerial System for Autonomous Target Search in Complex Unknown Environments

Yiming Luo¹, Zixuan Zhuang¹, Neng Pan², Chen Feng³,
Shaojie Shen³, Fei Gao², Hui Cheng¹, Boyu Zhou^{1,†}

Abstract—This paper tackles the challenge of autonomous target search using unmanned aerial vehicles (UAVs) in complex unknown environments. To fill the gap in systematic approaches for this task, we introduce Star-Searcher, an aerial system featuring specialized sensor suites, mapping, and planning modules to optimize searching. Path planning challenges due to increased inspection requirements are addressed through a hierarchical planner with a visibility-based viewpoint clustering method. This simplifies planning by breaking it into global and local sub-problems, ensuring efficient global and local path coverage in real-time. Furthermore, our global path planning employs a history-aware mechanism to reduce motion inconsistency from frequent map changes, significantly enhancing search efficiency. We conduct comparisons with state-of-the-art methods in both simulation and the real world, demonstrating shorter flight paths, reduced time, and higher target search completeness. Our approach will be open-sourced for community benefit ¹.

Index Terms—Aerial Systems: Perception and Autonomy, Aerial Systems: Applications, Search and Rescue Robots

I. INTRODUCTION

Unmanned aerial vehicles (UAVs) are prized for their compact size and exceptional maneuverability, making them indispensable across various applications like disaster search and rescue, resource exploration, and environment monitoring. In these tasks, UAVs can effectively supplant humans in the exploration of entirely unknown and hazardous environments while simultaneously conducting target search. This paper focuses on the challenge of autonomous searching for targets in complex unknown environments using UAVs.

The challenge of autonomous target search is closely connected to the field of autonomous exploration, a fundamental domain in robotics that has garnered significant attention [1]–[5]. While both areas share certain similarities, they are fundamentally distinct. Autonomous exploration primarily focuses on mapping unknown regions as either occupied or free areas. In contrast, autonomous target search demands the UAV to perform two related but different tasks simultaneously, i.e., exploration and inspection. The former task only requires coarsely mapping the unknown space, while the latter demands meticulous visual inspections in the occupied spaces where potential targets may be located, with more rigorous constraints such as observation distance

[†] Corresponding Author.

¹ Sun Yat-Sen University, China.

² State Key Laboratory of Industrial Control Technology, Institute of Cyber-Systems and Control, Zhejiang University, Hangzhou, China.

³ Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong, China.

{yiming.luo2001@gmail.com, zhoubu23@mail.sysu.edu.cn}

¹<https://github.com/SYSU-STAR/STAR-Searcher>

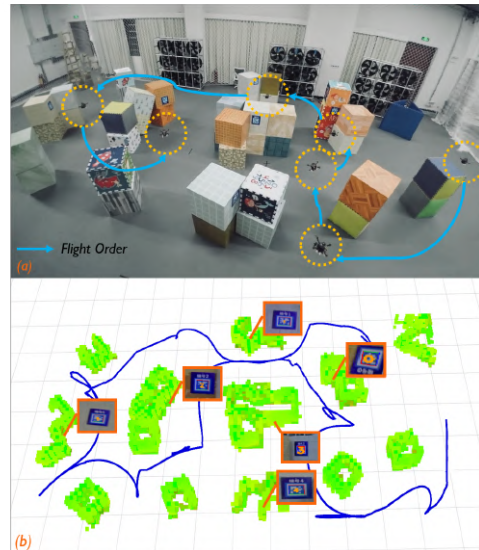


Fig. 1: (a) A test of autonomous target search conducted in a complex scene with six apriltags. (b) The apriltag search results and executed trajectory. Video of the experiments is available at: https://youtu.be/08ll_oo_DtU.

and viewing angles. Thus, an efficient system capable of handling the diverse perception requirements and generating motions that switch seamlessly between the two tasks is essential for fast target search. Currently, there is a gap in the existence of a systematic approach specifically tailored for autonomous target search—one that can ensure search completeness without compromising task efficiency.

Notably, autonomous target search presents a significant challenge in the context of path planning. The additional requirement of careful inspection introduces a substantial number of inspection viewpoints, leading to a considerably increased computational load for determining the shortest paths. Additionally, as the scene’s map and uninspected areas are incrementally constructed during the search process, the UAV must adapt its path when the map undergoes changes. Such alterations in the map can cause the newly planned path to noticeably deviate from the previous one, resulting in back-and-forth movements that detrimentally impact search efficiency. For a smooth and graceful flight, it is crucial to plan paths in real-time and ensure that consecutive paths remain consistent.

To tackle the challenges outlined above, we present **Star-Searcher: A Complete and Efficient Aerial System for Autonomous Target Search in Complex Unknown Environments**. Our aerial system incorporates specialized sensor

suites, mapping, and planning modules, all geared toward improving task efficiency and completeness. Our system takes advantage of various sensors and seamlessly switches between exploring unknown space and inspecting surface regions. We address the path planning challenges from two key angles. Firstly, we introduce a hierarchical planner supported by a visibility-based viewpoint clustering method. This approach decomposes the complex large-scale planning task into two more manageable sub-tasks: global path planning at the level of viewpoint clusters and local path planning at the level of individual viewpoints. Viewpoint clustering groups viewpoints separated by obstacles into distinct clusters and aggregates mutually visible ones into convex sets. This strategy provides regional guidance for generating reasonable global paths that sequentially visit different regions and ensures straightforward local paths for covering each convex set. Secondly, to mitigate motion inconsistency arising from frequent map changes, our global path planning incorporates a history-aware mechanism, taking into account both the historical movement tendency and visiting costs to all viewpoints. This prevents the occurrence of indecisive global paths across consecutive planning iterations, significantly enhancing search efficiency.

We compared our approach to state-of-the-art fast exploration and object-centric search methods in simulations. The results demonstrated superior search performance, with the shortest path length, flight time, and the highest completeness in all experiments. We further validated our system in complex real-world environments using entirely onboard devices. We plan to release our code as open source. In summary, our contributions are as follows:

- An aerial system with specialized sensor suites, mapping, and novel planning modules, which enables seamless exploration and inspection, achieving comprehensive and fast autonomous target search in complex unknown environments.
- A hierarchical planning method enhanced by visibility-based viewpoint clustering, allows the real-time generation of the global path at the level of viewpoint clusters and local path at the level of individual viewpoints with fewer detours.
- A history-aware mechanism for global path planning, which utilizes historical path information to prevent inconsistency in consecutive planning processes, significantly improving task efficiency.
- Extensive simulations and real-world experiments for validation. The source code will be made public.

II. RELATED WORKS

In autonomous target search, a critical step involves exploring unknown environments. Various methods for rapid autonomous exploration have been extensively researched, among which frontier-based approaches are among the most popular methods [6]–[11]. The concept of the frontier is initially introduced to delineate the boundaries between unknown and known areas [6]. They employ a greedy approach

that selects the nearest frontier at each step. Many subsequent methods have proposed more rational selection strategies to enhance exploration efficiency. A next-best-view selection strategy is introduced [12] and widely used. Each sampled viewpoint is evaluated using a utility function, which assesses the information gain achievable by visiting them and the required path length. The design of the utility function has also received more extensive consideration [13]–[16]. Recent methods formulate the problem of traversing all viewpoints as a traveling salesman problem [1, 2, 17], improving task efficiency.

Building upon exploration planning methods, several approaches have introduced object search techniques during the exploration process [3]–[5, 18]–[20]. Some of these methods involve conducting informative sampling to facilitate re-observations at higher resolutions when detecting objects [3]. Papatheodorou et al. [5] propose a utility function tailored to object-centric exploration, ensuring sufficient proximity for all background voxels. Kim et al. [4] apply a 2D space segmentation method and integrate search into the exploration. Meera et al. [21] employ a Gaussian Process-based model for target occupancy. Besides, motivated by the DARPA Subterranean Challenge, some systems have been developed to combine multiple robots to cooperate in searching objects [22]–[24]. A cooperative exploration strategy is proposed to enable the robots to coordinate their exploration while having the ability to explore individually [22]. Roucek et al. [23] develop a heterogeneous exploration robotic system and improve single-agent robustness. Other methods [24] combine multiple sensors in the task to improve the perception precision.

However, these methods trade-off between exploration and target search, lacking a more detailed consideration of autonomous target search and suffering from severe efficiency issues. In contrast, we explicitly define the problem of autonomous target search and design a suitable aerial platform. Furthermore, our history-aware hierarchical planner achieves fast search with a concise flight path.

III. PROBLEM FORMULATION

The problem of autonomous target search involves the search for an unknown number of targets detectable by visual sensors within a bounded 3D space $V \subset \mathbb{R}^3$. This space is represented as a set of cubic voxels. Continuous updates of the occupancy probability $P_o(v)$ for each voxel v incrementally map the initially unknown space $V_{unk} = V$ into two parts: $V_{free} \subset V$ (free space) and $V_{occ} \subset V$ (occupied space). Since the visible surfaces of the targets, denoted as V_{tar} , fall within the volume of V_{occ} , the UAV must conduct a thorough observation of all the voxels within V_{occ} to guarantee complete search. Voxels within V_{occ} that have not yet been observed with the desired precision by the visual sensor are labeled as V_{umi} . As most sensors' perception stops at surfaces, some hollow or corner spaces can not be mapped. These spaces are denoted by V_{res} . The task is considered fully completed when the entire area has

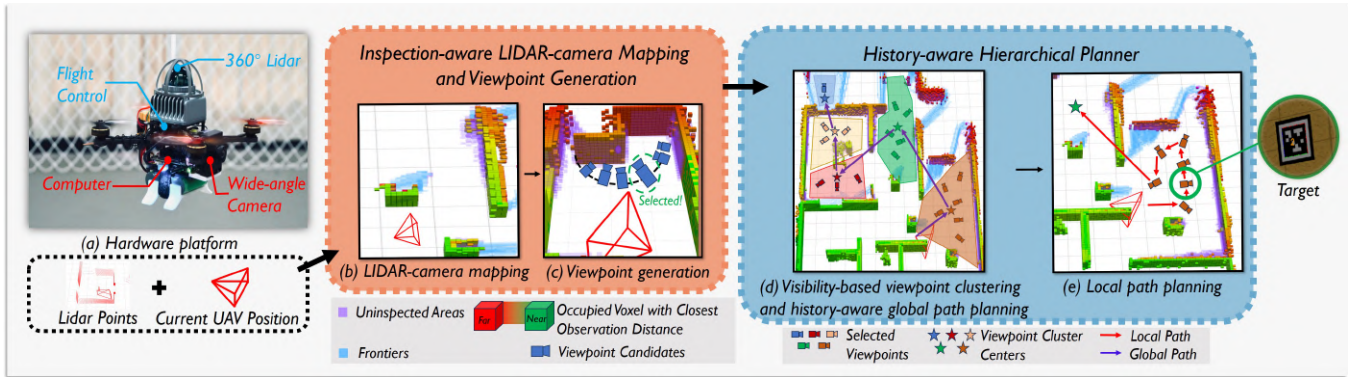


Fig. 2: An overview of our star-searcher system. (a) The hardware platform of our aerial system. (b) The UAV utilizes LIDAR point clouds and camera projection to update map information, cluster frontiers and uninspected areas. (c) For each cluster of frontiers and uninspected areas, a set of viewpoints is generated and scored based on information gain and viewing angle. The higher the score, the larger the size of the viewpoint. The one with the highest score is selected. (d) Visibility-based viewpoint clustering is performed, and history-aware global path planning is conducted based on this. (e) Local path planning.

been searched, i.e., when $V_{free} \cup V_{occ} = V \setminus V_{res}$, and $V_{uni} = \emptyset$.

IV. SYSTEM DESIGN

As explained in Sect. III, autonomous target search involves the simultaneous identification of occupied and free spaces while carefully observing uninspected areas V_{uni} . Improving task efficiency can be accomplished by employing various sensors to identify V_{occ} and update V_{uni} . We configure a UAV equipped with a 360-degree LIDAR and a wide-angle RGB camera, as shown in Fig. 2(a). Despite its conventional appearance, this platform consists of minimal sensor components carefully tailored to our specific problem. The wide sensing range of the 360-degree LIDAR facilitates the rapid acquisition of geometric information about the surrounding environment, enabling the UAV to quickly identify occupied areas for detailed inspection using the RGB camera. By utilizing the wide-angle camera, the UAV can cover more spaces and detect targets more rapidly.

The algorithm framework in Fig. 2(b)-(e) comprises mapping and planning modules. The mapping module represents the environment volumetrically, fusing LIDAR and camera data continuously for each voxel to update its occupancy and the closest observation distance from the camera (Sect. V-A). Frontiers and uninspected areas are then extracted and clustered, generating and selecting corresponding viewpoints with high information gain and suitable viewing angle (Sect. V-B). The history-aware hierarchical planner (Sect. VI) then utilizes the viewpoints and previous path planning results to plan global paths and local paths, achieving simultaneous identification of free and occupied spaces and thorough inspection of uninspected areas. A visibility-based viewpoint clustering method is employed during global path planning (Sect. VI-A) to create a more reasonable visiting order and reduce the computation burden.

V. INSPECTION-AWARE LIDAR-CAMERA MAPPING AND VIEWPOINT GENERATION

Traditional occupancy mapping used in exploration lacks information about the observation distances of voxels. This

limitation hinders the ability to plan inspection paths to ensure a thorough target search. Moreover, existing methods overlook the consideration of viewing angles towards uninspected surfaces, potentially leading to missed targets due to the detrimental effects of large viewing angles on target detection. Our mapping module addresses the first issue by integrating measurements from both LIDAR and the camera, providing the closest observation distance for each occupied voxel and enabling more precise path planning for thorough inspections. Additionally, we introduce a scoring mechanism for viewpoint selection, which prevents large viewing angles in uninspected areas.

A. LIDAR-camera Mapping with Inspection Information

Our volumetric environment representation is based on [25]. In addition to occupancy probability, we update the closest observation distance to the camera for each occupied voxel. When a new frame of LIDAR point clouds is acquired, all point clouds are employed to update the occupancy probability using the ray-casting method. Simultaneously, we can project the voxels already mapped as occupied into the camera coordinate frame, updating the closest observation distances for those within the camera's field of view (FOV) and not occluded by any occupied areas. Voxels with a closest observation distance greater than the maximum observation distance d_{max} are labeled as uninspected areas V_{uni} . They are subsequently removed from the uninspected areas when they are scanned within d_{max} . Continuous target detection is performed in RGB images, and when a target is detected, the corresponding voxels are mapped to V_{tar} through transformation from pixel coordinates to world coordinates. It's worth noting that additional information pertaining to observation accuracy can be integrated into the map using a similar approach. However, conducting a comprehensive study of these related factors falls beyond the scope of our research.

B. Viewpoint Generation

Given the information on occupancy and observation distance, both frontiers [6] and uninspected areas are extracted,

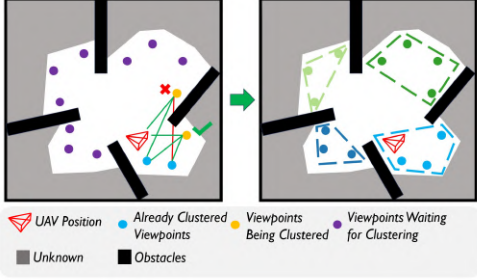


Fig. 3: Illustration of the visibility-based viewpoint clustering. If a viewpoint has collision-free rays to all viewpoints in the current cluster, it is assigned to that cluster.

as defined in Sect. III. Uninspected areas and frontiers imply the potential presence of targets or areas that can be expanded upon in the map. Hence, we perform viewpoint sampling in proximity to these areas. Similar to [1], we employ a PCA-based method to split excessively large clusters along axes. Several viewpoint positions and their corresponding yaw angles are sampled in the spherical space around the center of each cluster. Then we score all viewpoints of each cluster based on two criteria: information gain and viewing angle.

- **Information Gain:** The information obtained by the UAV at each viewpoint is determined by a weighted combination of the number of observable frontiers to the lidar $\mathcal{N}_{unknown}$ and the number of observable uninspected voxels to the camera $\mathcal{N}_{uninspected}$ within d_{max} , i.e.

$$S_{info} = \omega_{uni} \cdot \mathcal{N}_{uninspected} + \omega_{unk} \cdot \mathcal{N}_{unknown} \quad (1)$$

where ω_{uni} and ω_{unk} denote the weights for uninspected voxels and frontiers, respectively. We assign a larger value to ω_{uni} in experiments.

- **Viewing Angle:** To ensure accurate target detection and mitigate errors caused by extreme viewing angles, our method scores viewpoints based on the proximity between the vector from the center of the cluster to the viewpoint and the average normal vector of each cluster \mathbf{n}_{avg} , i.e.

$$S_{nor} = \frac{\mathbf{p}_{c,v} \cdot \mathbf{n}_{avg}}{\|\mathbf{p}_{c,v}\| \|\mathbf{n}_{avg}\|} \quad (2)$$

The average normal of each cluster is computed as the average normal of the center points of each voxel within the cluster by PCL package. The final score for each viewpoint is calculated as follows:

$$S_{VP} = S_{nor} \cdot S_{info} \quad (3)$$

Finally, we pick the highest-scoring viewpoint of each cluster as illustrated in Fig. 2(c).

VI. HISTORY-AWARE HIERARCHICAL PLANNER

In order to plan paths with fewer detours and revisits in real-time, simultaneously exploring unknown regions and covering uninspected areas, we adopt a hierarchical planning strategy. We first perform visibility-based viewpoint clustering (Sect. VI-A) and conduct global path planning (Sect. VI-B) to determine the visiting order of the viewpoint clusters.

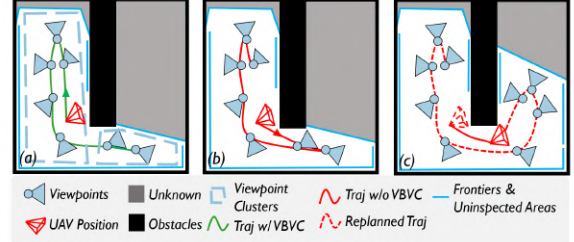


Fig. 4: A comparison between the planned trajectories with (fig. a) and without (fig. b) visibility-based viewpoint clustering (VBVC). Without viewpoint clustering, the UAV calculates the shortest path to visit all viewpoints and selects the viewpoint behind the obstacle as the next target. Once it reaches this viewpoint, it discovers a new area and generates additional viewpoints. As a result, it replans a trajectory as indicated by the red dashed line, leading to revisits later, as shown in fig. c.

Subsequently, local path planning is conducted within the first viewpoint cluster, efficiently covering the frontiers and uninspected areas (Sect. VI-C). Our global planner incorporates historical path information, maintaining the consistency between the replanned path and the previous ones (Sect. VI-B).

A. Visibility-based Viewpoint Clustering

By leveraging a broad sensing range, the UAV can identify occupied surfaces for thorough visual inspection rapidly. However, this results in a significant increase in the number of required viewpoints, also for those used to cover the frontier. Planning a single shortest path to visit each of these viewpoints would result in excessive computation time, making it impossible to respond promptly to environmental changes. To address this challenge, we introduce a viewpoint clustering method to intelligently divide the entire set of viewpoints into multiple subsets. We then generate global routes passing through these subsets and local paths that visit the viewpoints within the subset sequentially, creating an efficient path within a manageable computation time.

The details of the proposed visibility-based viewpoint clustering are illustrated in Fig. 3. We initiate the process by designating the current UAV position as the starting point for the first cluster. During each clustering iteration, viewpoints within a specified radius R_{vp} around the current cluster's center engage in ray-casting, prioritized based on their distances from the cluster's center. If a viewpoint's rays do not intersect with obstacles from any viewpoints in an existing cluster, that viewpoint is incorporated into the cluster, and the cluster's center is recalculated as the average position of all the viewpoints within that cluster. Once a cluster is finalized, we select the nearest unclustered viewpoint to its center to serve as the starting point for a new round of clustering. This process continues iteratively, gradually forming distinct clusters of viewpoints based on their mutual visibility and proximity until no viewpoints remain unclustered, guiding the path planning for efficient target search.

The proposed clustering method ensures mutual visibility of viewpoints within each cluster, offering several advantages. Firstly, visibility allows viewpoints in the same cluster

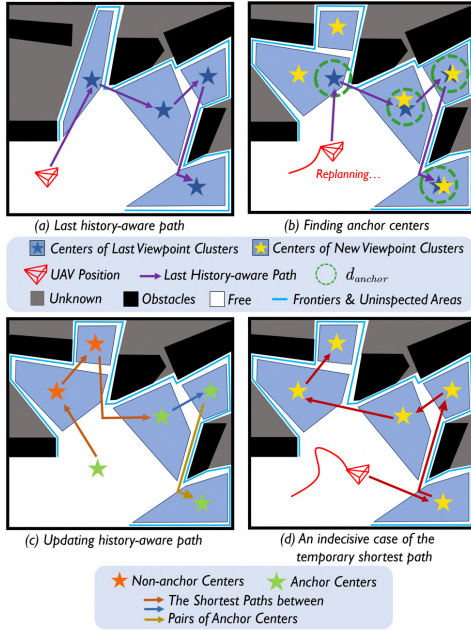


Fig. 5: Illustration of the history-aware global path planning. After finding the anchor centers, multiple collision-free shortest paths are calculated using the A* algorithm between pairs of anchor centers and concatenated to update the history-aware global path. Without the history-aware path, an indecisive trajectory occurs, as indicated by the red curve in (d).

to be encompassed by a collision-free convex set. This means the UAV can efficiently visit all viewpoints in the same cluster without the need to detour to avoid collisions. Secondly, viewpoints that are occluded by obstacles are naturally grouped into different clusters, providing implicit regional guidance for planning a sensible global path that subsequently visits separate regions. Additionally, this method effectively reduces severe revisits. For instance, Fig. 4(b)-(c) illustrates a scenario where the UAV prioritizes a viewpoint behind an obstacle, discovers another unexplored area, and diverts to it, leading to severe revisits later. In contrast, by thoroughly exploring one convex cluster before moving to the next, the UAV eliminates the need to circumvent obstacles and revisit previously inspected areas.

B. History-aware Global Path Planning

After the viewpoint clustering, our planner computes a short global path starting from the UAV’s current position, traversing the centers of all viewpoint clusters. This can be formulated as an Asymmetric Traveling Salesman Problem (ATSP) [1]. The planner replans every time the map is updated. However, a complete refresh of the global path after replanning can lead to drastic variations in the visiting order of different regions due to the similar path costs, causing indecisive flight as shown in Fig. 5(d). Actually, only a few regions are updated at a time, while other regions remain unchanged. For these unchanged regions, maintaining the visiting order from the previously planned path can ensure consistency in planning and flight. The newly updated regions should be reasonably integrated into the visitation planning for the unchanged regions. Our proposed method

leverages the relative visiting order from the previous global path to generate a new global path.

During each global path planning, we compute a temporary shortest path constructed by considering all viewpoint cluster centers and maintain a history-aware path. Although it requires additional computation, the update of the history-aware path incurs a minimal increase in computation time. The history-aware path is initialized using the temporary shortest path. Fig. 5 illustrates the updating process. To ensure the order of new viewpoint clusters aligns with the previous history-aware path, we first identify the nearest old viewpoint cluster for each new one and calculate the distances between their centers. Subsequently, the new viewpoint clusters are sorted based on the order of their nearest centers in the last history-aware path. In regions with minimal updates, where the new viewpoint clusters exhibit slight variations, centers with a distance change less than the threshold, d_{anchor} , are designated as anchor centers, and their associated viewpoint clusters are labeled as unchanged. The UAV’s position is also considered an anchor center. For the non-anchor viewpoint clusters, integration with anchor clusters is required to form a coherent new path. To achieve this, we select pairs of neighboring anchor clusters as start and end points, retrieve non-anchor clusters between them, and compute the shortest path. The associated problem is formulated as a TSP with fixed start and end nodes, which can be transformed into an ATSP [17]. Multiple such shortest paths between pairs of anchor clusters are concatenated to obtain the new history-aware path. Finally, the cost of the history-aware path is compared to the cost of the temporary shortest path. If the difference \mathbb{D}_{cost} between them is small, the history-aware path is adopted. Conversely, the temporary shortest path is adopted, and the history-aware path is reset to the temporary shortest path.

C. Local Path Planning

The global path provides a rational order for visiting different regions. We perform more detailed local path planning based on the guidance of the global path. Specifically, we select all viewpoints in the first viewpoint cluster to visit for local path planning. We construct an ATSP with the UAV’s current position as the starting point and the center of the second viewpoint cluster in the global path as the endpoint. To enhance the smoothness of UAV motion, we take into account the variations in the UAV’s velocity when calculating the visiting cost. We determine the cost associated with moving the UAV from its current position to each viewpoint by breaking down the UAV’s velocity into components aligned with and perpendicular to the line connecting the UAV to the viewpoint, modeling the motion in both directions as constant acceleration motion with an initial position at the UAV’s current position, final position at each viewpoint, initial velocity as v_{ali} and v_{per} , final displacement as the distance from the current position to the viewpoint l and 0, and calculate the corresponding cost by computing the time required for the motion in both directions. Consequently, the cost function from the UAV

position p to the viewpoint vp_i can be defined as follows:

$$t_{ali} = \begin{cases} \frac{\sqrt{v_{ali}^2 + 2a_{ali}l} - v_{ali}}{a_{ali}}, & \text{if } \frac{v_{max}^2 - v_{ali}^2}{2a_{ali}} < l \\ \frac{v_{max} - v_{ali}}{a_{ali}} + \frac{l - \frac{v_{max}^2 - v_{ali}^2}{2a_{ali}}}{v_{max}}, & \text{else} \end{cases} \quad (4)$$

$$\mathbb{C}(p, vp_i) = \max \left\{ t_{ali}, \frac{2v_{per}}{a_{per}}, \frac{|\phi - \phi_j|}{\omega_{max}} \right\} \quad (5)$$

ϕ and ω_{max} denote the yaw angle and the max angular velocity of the UAV respectively. The cost between two points vp_i and vp_j is defined as follows:

$$\mathbb{C}(vp_i, vp_j) = \max \left\{ \frac{L(vp_i, vp_j)}{v_{max}}, \frac{|\phi_i - \phi_j|}{\omega_{max}} \right\} \quad (6)$$

where $L(vp_i, vp_j)$ denotes the distance between two viewpoints. Once we obtain the local path, we follow the framework proposed in [26] to generate a smooth and safe trajectory, which is then sent to the controller for execution.

D. Computation Complexity Analysis

Due to the large number of generated viewpoints, attempting to use all of them for path planning would impose a significant computation burden. Within these extensive computation needs, the computation of the visiting cost between different viewpoints predominantly consumes a significant portion of resources. The farther apart two viewpoints are, the longer it takes to search for a path and calculate the visiting cost between them. However, the planned path often avoids consecutively visiting two points with a high visiting cost, making the computation unnecessary yet time-consuming. Our hierarchical planner only requires computing costs between the centers of viewpoint clusters during the global path planning and between the viewpoints within the first cluster to visit during the local path planning. Therefore, our method transforms the computation complexity from $O(N^2)$ to $O(n_1^2)$ for global path planning and $O(n_2^2)$ for local path planning. Here, N represents the total number of all viewpoints, n_1 is the number of viewpoint clusters, and n_2 is the number of viewpoints within the first cluster. Typically, $N \gg n_1, n_2$. Furthermore, the viewpoints in local path planning are typically in proximity, avoiding the need to calculate costs between distant viewpoints. Consequently, our planner, which can run at a frequency of 10 Hz in both simulation and real-world experiments, achieves good real-time capability.

VII. RESULTS

A. Implementation Details

We set $\omega_{uni} = 0.8$ and $\omega_{unk} = 0.2$ in Equ.1. In viewpoint clustering, we set R_{vp} to 3m. The ATSPs are solved using the Lin-Kernighan-Helsgaun heuristic solver [27]. For real-world experiments, we utilize the Mid360 LIDAR and an efficient LIDAR-inertial localization system [28]. A geometric controller [29] is employed to track the (x, y, z, ϕ) trajectory. All modules run on the Jetson Orin NX 16GB platform.

TABLE I: Simulation experiments.

Scene	Method	Flight time(s)		Path length(m)		Completeness(%)
		Avg	Std	Avg	Std	Avg
SubT	FUEL-3m [1]	195.19	11.66	265.75	16.12	100
	FUEL-4m [1]	177.09	4.17	245.63	6.92	95.0✗
	Semantic [5]	266.19	23.04	309.70	29.76	100
	Ours	153.36	3.46	191.65	5.31	100
Maze I	FUEL-3m [1]	256.48	5.62	344.14	4.60	100
	FUEL-4m [1]	236.99	7.58	328.37	16.33	77.5✗
	Semantic [5]	387.75	27.39	426.40	35.93	100
	Ours	216.94	6.04	281.69	14.12	100
Maze II	FUEL-3m [1]	394.58	31.30	559.96	36.26	100
	FUEL-4m [1]	308.01	40.36	468.71	35.89	68.8✗
	Semantic [5]	502.7	29.09	678.50	54.73	100
	Ours	303.15	21.78	458.83	26.62	100
Relic	FUEL-3m [1]	261.68	6.42	371.41	15.70	100
	FUEL-4m [1]	178.87	13.61	283.57	17.51	81.3✗
	Semantic [5]	368.77	19.66	563.66	30.18	100
	Ours	179.33	9.67	269.64	9.59	100

B. Benchmark Comparisons

We conduct simulation experiments in Gazebo, evaluating our method in four scenes: SubT [24] (68 m x 18 m x 2 m), maze I (33 m x 27 m x 2 m), maze II (60 m x 50 m x 2 m) and relic (10 m x 10 m x 7 m). We place 8 apriltags [30] in each scene, with an effective recognition distance set to 3.0 m. Our proposed method is compared to FUEL [1] (a fast exploration method) and Semantic [5] (a NBVP-based object search method). Since Semantic [5] has not been fully open-source, we use our implementation (excluding object mapping). The proposed method and Semantic [5] employ a UAV equipped with a LIDAR with a range of 8m and an RGB camera with a fov of [68, 51] degrees. The sensor used for FUEL [1] is a depth camera and we conduct two experiments with different sensing ranges for FUEL [1] to show the limitations of applying exploration methods directly to target search. In one experiment, the sensing range is equivalent to the effective recognition distance to ensure completeness while the other one is set to 4 m. Additionally, we configure FUEL [1] with the same sensor setting and observation distance constraint as the proposed method in the ablation study. The side length of each voxel is set to 0.1 m, and the maximum observation distance d_{max} is set to 3.0 m. We set the dynamics limits as $v_{max} = 2.0$ m/s, $a_{max} = 1.5$ m/s² and $\omega_{max} = 1.2$ rad/s. All methods are run with the same configuration for 5 times in each scene. We evaluate the efficiency (flight length and time) and search completeness (number of recognized apriltags) of each method. Table.I presents the results for each scene, while Fig. 6 displays the executed trajectories in different scenes.

The results show that many revisits and detours occur in the FUEL [1] and Semantic [5] due to the lack of effective planning methods for target search. In contrast, our method achieves high search completeness due to ensuring sufficient observation of each occupied voxel. Also, the proposed method demonstrates the shortest path length and flight time in this task. This is attributed to the viewpoint clustering method, which prevents detours and revisits. The history-aware global path planning also produces a more consistent path, enhancing efficiency.

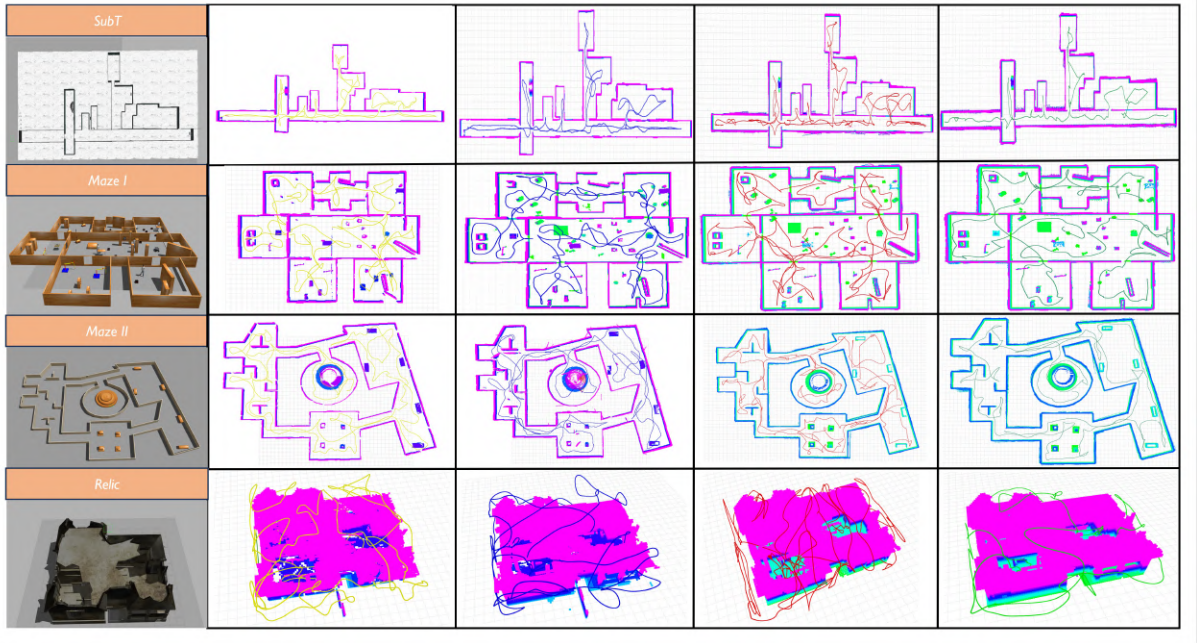


Fig. 6: Trajectories generated by the proposed method (green), FUEL-3m [1](yellow), FUEL-4m [1](blue) and Semantic [5](red) in simulation experiments.

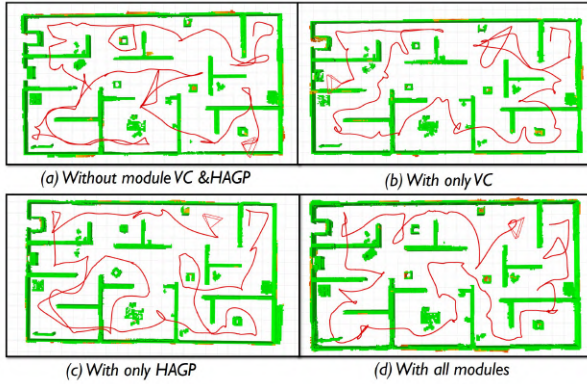


Fig. 7: Executed trajectories in the ablation study. We test our method without viewpoint clustering (module VC) or history-aware global planning (module HAGP).

C. Ablation Study

Table II and Fig. 7 illustrate our tests regarding viewpoint clustering (Module VC) and history-aware global planning (Module HAGP) in the small maze (24 m x 12 m x 2 m). When neither of the two modules is engaged, it is equivalent to FUEL [1] with the LIDAR-camera mapping module, solving the path by considering all viewpoints. Fig. 7(b) and the corresponding table data show that our viewpoint clustering module considers the visibility between viewpoints, reducing the detours across obstacles. Without the viewpoint clustering module, the computation time sharply increases as the number of viewpoints accumulates, causing lag in algorithm execution. The comparison between Exp 2 and Exp 4 demonstrates the effectiveness of history-aware global planning for improving planning consistency and efficiency. In conclusion, with all modules combined, our algorithm generates a concise and consistent path with fewer revisits, and it also exhibits good real-time capability.

TABLE II: Results in the ablation study.

Setting	Flight time(s)		Path length(m)	
	Avg	Std	Avg	Std
Exp 1 (Without module VC & HAGP)	126.25	9.89	146.15	4.46
Exp 2 (Only VC)	109.04	8.24	116.68	10.81
Exp 3 (Only HAGP)	129.47	16.65	143.17	9.21
Exp 4 (With all modules)	97.8	4.90	109.4	7.77

D. Real-world Experiments

We further validate our method in real-world experiments. Considering camera motion blur, we set the dynamics limits as $v_{\max} = 1.6$ m/s, $a_{\max} = 0.8$ m/s² and $\omega_{\max} = 0.9$ rad/s. The maximum observation distance d_{\max} is set to 2.0 m, within which the apriltag can be stably recognized. All of our modules run onboard without relying on external devices.

The first scene consists of a cluttered indoor environment (14 m x 7 m x 2 m) with 6 apriltags. The second scene is a maze (15 m x 8.5 m x 2 m) with 10 apriltags. Each apriltag has dimensions of 0.12 m x 0.12 m. The three-dimensional coordinates of each apriltag in the world frame are pre-measured and calibrated using laser ranging. Utilizing the UAV's pose obtained by the SLAM module [28] and the pose of the apriltags recognized in the camera frame, we could compute the apriltags' positions in the world frame. In each experiment, all apriltags are successfully recognized, with a coordinate error ranging from 0.2 m to 0.4 m. The time taken for the two scenes are 110 s and 180 s respectively, more details can be seen in the video. The online generated map and trajectory are shown in Fig. 1 and Fig. 8. These experiments validate the capability of our system in complex real-world scenarios.

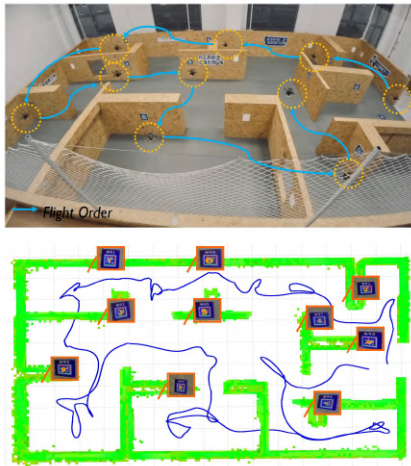


Fig. 8: Experiments conducted in a maze with 10 apriltags to be found.

VIII. CONCLUSION

In this paper, we propose a systematic solution designed to autonomous target search in complex unknown environments. An aerial system with specialized sensor suites, mapping, and planning modules is developed to enhance task efficiency and completeness. A hierarchical planner generates global and local paths with regional guidance provided by a visibility-based viewpoint clustering method in real-time. A history-aware mechanism is introduced to prevent motion inconsistency in consecutive global planning processes. Extensive simulation and real-world experiments validate the effectiveness of our proposed method.

We have identified several directions for future work: extending the current method to a multi-UAV swarm, considering conducting autonomous target search in complex dynamic environments, and for dynamic targets.

REFERENCES

- [1] B. Zhou, Y. Zhang, X. Chen, and S. Shen, "Fuel: Fast uav exploration using incremental frontier structure and hierarchical planning," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 779–786, 2021.
- [2] C. Cao, H. Zhu, H. Choset, and J. Zhang, "Tare: A hierarchical framework for efficiently exploring complex 3d environments," in *Robotics: Science and Systems*, vol. 5, 2021.
- [3] T. Dang, C. Papachristos, and K. Alexis, "Autonomous exploration and simultaneous object search using aerial robots," in *2018 IEEE Aerospace Conference*. IEEE, 2018, pp. 1–7.
- [4] H. Kim, H. Kim, S. Lee, and H. Lee, "Autonomous exploration in a cluttered environment for a mobile robot with 2d-map segmentation and object detection," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 6343–6350, 2022.
- [5] S. Papatheodorou, N. Funk, D. Tzoumanikas, C. Choi, B. Xu, and S. Leutenegger, "Finding things in the unknown: Semantic object-centric exploration with an MAV," in *IEEE International Conference on Robotics and Automation*, London, United Kingdom, May 2023.
- [6] B. Yamauchi, "A frontier-based approach for autonomous exploration," in *Proceedings 1997 IEEE International Symposium on Computational Intelligence in Robotics and Automation CIRA'97: Towards New Computational Principles for Robotics and Automation*. IEEE, 1997, pp. 146–151.
- [7] W. Gao, M. Booker, A. Adiwahono, M. Yuan, J. Wang, and Y. W. Yun, "An improved frontier-based approach for autonomous exploration," in *2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV)*. IEEE, 2018, pp. 292–297.
- [8] J. Faigl and M. Kulich, "On determination of goal candidates in frontier-based multi-robot exploration," in *2013 European Conference on Mobile Robots*. IEEE, 2013, pp. 210–215.

- [9] M. Kulich, J. Kubalík, and L. Přeučil, "An integrated approach to goal selection in mobile robot exploration," *Sensors*, vol. 19, no. 6, p. 1400, 2019.
- [10] C. Dornhege and A. Kleiner, "A frontier-void-based approach for autonomous exploration in 3d," *Advanced Robotics*, vol. 27, no. 6, pp. 459–468, 2013.
- [11] L. Heng, A. Gotovos, A. Krause, and M. Pollefeys, "Efficient visual exploration and coverage with a micro aerial vehicle in unknown environments," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2015, pp. 1071–1078.
- [12] A. Bircher, M. Kamel, K. Alexis, H. Oleynikova, and R. Siegwart, "Receding horizon" next-best-view" planner for 3d exploration," in *2016 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2016, pp. 1462–1468.
- [13] A. Akbari and S. Bernardini, "Informed autonomous exploration of subterranean environments," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 7957–7964, 2021.
- [14] L. Schmid, M. Pantic, R. Khanna, L. Ott, R. Siegwart, and J. Nieto, "An efficient sampling-based method for online informative path planning in unknown environments," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1500–1507, 2020.
- [15] Z. Xu, D. Deng, and K. Shimada, "Autonomous uav exploration of dynamic environments via incremental sampling and probabilistic roadmap," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 2729–2736, 2021.
- [16] H. H. González-Banos and J.-C. Latombe, "Navigation strategies for exploring indoor environments," *The International Journal of Robotics Research*, vol. 21, no. 10-11, pp. 829–848, 2002.
- [17] B. Zhou, H. Xu, and S. Shen, "Racer: Rapid collaborative exploration with a decentralized multi-uav system," *IEEE Transactions on Robotics*, 2023.
- [18] G. Best, J. Faigl, and R. Fitch, "Online planning for multi-robot active perception with self-organising maps," *Autonomous Robots*, vol. 42, pp. 715–738, 2018.
- [19] R. Ashour, T. Taha, J. M. M. Dias, L. Seneviratne, and N. Almoosa, "Exploration for object mapping guided by environmental semantics using uavs," *Remote Sensing*, vol. 12, no. 5, p. 891, 2020.
- [20] A. Asgharivaskasi and N. Atanasov, "Active bayesian multi-class mapping from range and semantic segmentation observations," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021, pp. 1–7.
- [21] A. A. Meera, M. Popović, A. Millane, and R. Siegwart, "Obstacle-aware adaptive informative path planning for uav-based target search," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 718–724.
- [22] M. Kulkarni, M. Dharmadhikari, M. Tranzatto, S. Zimmermann, V. Reijgwart, P. De Petris, H. Nguyen, N. Khedekar, C. Papachristos, L. Ott, *et al.*, "Autonomous teamed exploration of subterranean environments using legged and aerial robots," in *2022 International Conference on Robotics and Automation (ICRA)*. IEEE, 2022, pp. 3306–3313.
- [23] T. Roucek, M. Pecka, P. Cizek, T. Petricek, J. Bayer, V. Šalansky, T. Azayev, D. Hert, M. Petrlik, T. Báca, *et al.*, "System for multi-robotic exploration of underground environments ctu-cras-norlab in the darpa subterranean challenge," *arXiv preprint arXiv:2110.05911*, 2021.
- [24] G. Best, R. Garg, J. Keller, G. A. Hollinger, and S. Scherer, "Resilient multi-sensor exploration of multifarious environments with a team of aerial robots," in *Robotics: Science and Systems (RSS)*, 2022.
- [25] L. Han, F. Gao, B. Zhou, and S. Shen, "Fiesta: Fast incremental euclidean distance fields for online motion planning of aerial robots," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 4423–4430.
- [26] B. Zhou, F. Gao, L. Wang, C. Liu, and S. Shen, "Robust and efficient quadrotor trajectory generation for fast autonomous flight," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 3529–3536, 2019.
- [27] K. Helsgaun, "An effective implementation of the lin-kernighan traveling salesman heuristic," *European journal of operational research*, vol. 126, no. 1, pp. 106–130, 2000.
- [28] W. Xu and F. Zhang, "Fast-lio: A fast, robust lidar-inertial odometry package by tightly-coupled iterated kalman filter," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 3317–3324, 2021.
- [29] T. Lee, M. Leok, and N. H. McClamroch, "Geometric tracking control of a quadrotor uav on se (3)," in *49th IEEE conference on decision and control (CDC)*. IEEE, 2010, pp. 5420–5425.
- [30] J. Wang and E. Olson, "AprilTag 2: Efficient and robust fiducial detection," in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, oct 2016, pp. 4193–4198.