IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

Autonomous Exploration with Terrestrial-Aerial Bimodal Vehicles

Yuman Gao[1,*], Ruibin Zhang[1,*], Tiancheng Lai[1,*], Yanjun Cao[1], Chao Xu[1], and Fei Gao[1]

***Abstract*—Terrestrial-aerial bimodal vehicles, which integrate the high mobility of aerial robots with the long endurance of ground robots, offer significant potential for autonomous exploration. Given the inherent energy and time constraints in practical exploration tasks, we present a hierarchical framework for the bimodal vehicle to utilize its flexible locomotion modalities for exploration. Beginning with extracting environmental information to identify informative regions, we generate a set of potential bimodal viewpoints. To adaptively manage energy and time constraints, we introduce an extended Monte Carlo Tree Search approach that strategically optimizes both modality selection and viewpoint sequencing. Combined with an improved bimodal vehicle motion planner, we present a complete bimodal energy- and time-aware exploration system. Extensive simulations and deployment on a customized real-world platform demonstrate the effectiveness of our system.**

***Index Terms*—Aerial Systems: Applications, Task and Motion Planning, Aerial Systems: Perception and Autonomy**

I. Introduction

Autonomous exploration has gained increasing attention in both academia and industry, with applications in search and rescue, engineering surveying, and tunnel inspection. In recent years, researchers have proposed considerable exploration strategies and deployed them on uncrewed aerial vehicles (UAVs) and uncrewed ground vehicles (UGVs). However, exploration performance is constrained by the kinodynamic characteristics of mobile robots. Although aerial robots offer high mobility and a broad field of view (FoV), their endurance is significantly shorter than that of ground robots, limiting their ability to support large-scale and long-term exploration. Ground robots, especially wheeled vehicles, face challenges when navigating complex and rugged terrains, restricting exploration to wide, flat areas. To break the above hardware constraints, collaborative aerial-ground exploration systems consisting of UAV-UGV robots are proposed [1]–[5]. However, adopting such systems introduces issues of multi-robot SLAM, planning, and coordination, which greatly increase the complexity of the problem.

Manuscript received: February, 18, 2025; Revised May, 25, 2025; Accepted July, 28, 2025.

This paper was recommended for publication by Editor Giuseppe Loianno upon evaluation of the Associate Editor and Reviewers' comments. This work was supported by the National Key R&D Program of China under grant No. 2023YFB4706600, the Zhejiang Provincial Science and Technology Plan Project under Grant No. 2024C01170 and the National Natural Science Foundation of China under Grant No. 62322314.

*Indicates equal contribution.

[1]All authors are from the Institute of Cyber-Systems and Control, College of Control Science and Engineering, Zhejiang University, Hangzhou 310027, China, and also from Huzhou Institute, Zhejiang University, Huzhou 313000, China. (Corresponding Author: Fei Gao)

E-mail: {ymgao, fgaoaa}@zju.edu.cn

Digital Object Identifier (DOI): see top of this page.

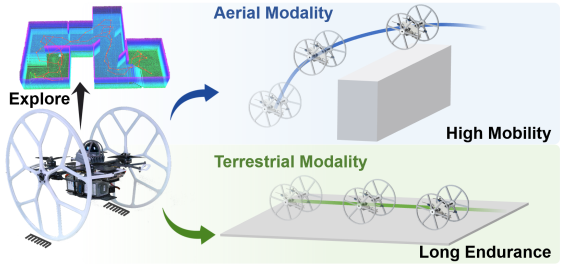

Fig. 1: The TABV integrates two modalities into a single platform, offering significant potential for autonomous exploration.

To address the above issues, we propose a hierarchical exploration framework making use of a terrestrial-aerial bimodal vehicle (TABV) [6]. With this type of vehicle, the long endurance of UGVs and high mobility and broad FoV of UAVs can be integrated into a single robot system as shown in Fig. 1, showcasing great potential in exploration. Moreover, in exploration tasks, especially in search and rescue scenarios, energy and time constraints should be considered, as robots operate with finite battery capacity and are typically expected to complete missions within a reasonable time frame. However, these constraints are often overlooked in previous studies despite their significance. With bimodal capability, TABV exhibits enhanced flexibility for these constraints. To leverage the unique characteristics of the TABV for exploration, we start by generating bimodal viewpoints according to the frontier of the known environment. Subsequently, we introduce an adaptive exploration planner that enables the TABV to select a suitable modality to complete the exploration under given energy and time constraints. We propose the Bimodal Monte Carlo Tree Search (BM-MCTS) method to determine the traverse sequence of the generated viewpoints. Then, we adopt and improve the bimodal motion planner from our previous work [6] for trajectory generation. To demonstrate and validate the proposed method, we conduct extensive exploration tests in various scenes in the simulation. Furthermore, we deploy our system into a customized TABV platform to conduct real-world experiments.

Contributions of this paper are summarized as follows:

1) A hierarchical exploration framework for TABV, featuring a bimodal viewpoint generation module based on two alternative coverage strategies, and an energy- and time-aware decision-making mechanism that fully exploits the robot's bimodal locomotion capability.
2) An adaptive BM-MCTS approach for information-driven exploration, enabling flexible modality and viewpoint selection under energy and time constraints.
3) Integrating the exploration planner with an enhanced bimodal motion planner, featuring terrain perception and modality-aware planning, forming a complete au-

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

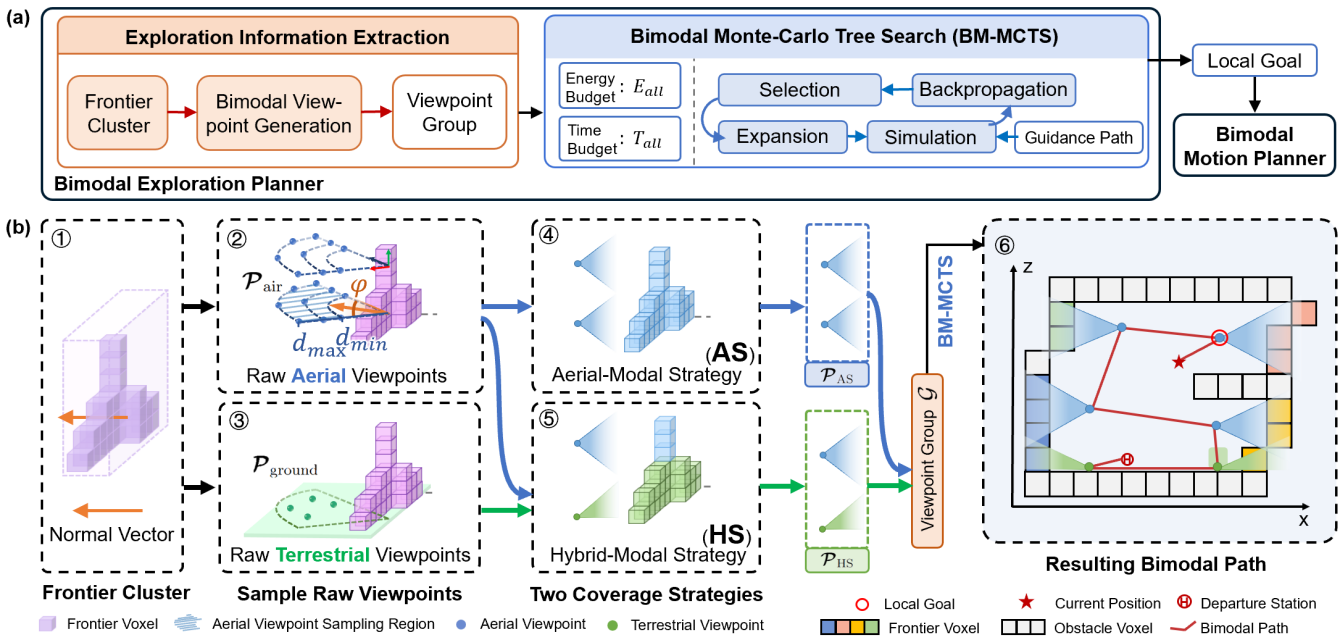

Fig. 2: (a): An overview of the proposed TABV exploration framework. (b): Module details: ①: A frontier cluster C whose surface normal is computed via Principal Component Analysis (PCA) and oriented toward the known space. ②-③: Raw aerial viewpoints $\mathcal{P}_{\text{aerial}}$ and raw terrestrail viewpoints $\mathcal{P}_{\text{ground}}$ generation. $\mathcal{P}_{\text{aerial}}$ are generated via cylindrical coordinate sampling within $[d_{min}, d_{max}]$ range and φ angle around the cluster's normal at multiple heights. ④-⑤: Two alternative strategies to select bimodal viewpoints to fully cover C, resulting $\mathcal{P}_{\text{AS}}$ and $\mathcal{P}_{\text{HS}}$. ⑥: An example of the resulting bimodal path and the next local goal.

tonomous TABV system deployed on the real platform.

II. Related Work

Autonomous exploration problem has been tackled using various strategies on multiple robot platforms. Among the various exploration methods, frontier-based methods make unknown environment information implied in frontiers to guide exploration [7]–[11]. Taking a set of frontiers or sampled viewpoints as goals, a global path traversing the goals is found through different approaches, such as the shortest distance criterion [7], minimum velocity change criterion [8], the balanced reward between information gain and path cost [12], or the solution of Traveling Salesman Problem (TSP) [9]–[11]. Meanwhile, some works utilize MCTS to find non-myopic solutions for more complex such as decentralized, long-horizon, and multi-agent exploration tasks [13]–[15].

To combine the advantages of aerial and ground mobile robots, researchers also focus on developing collaborative UAV-UGV exploration systems. Butzke et al. [1] mount a drone on a ground robot as a backup. When the ground robot encounters high, invisible areas, the drone takes off to cover them. Wang et al. [2] use a centralized approach to plan UAV-UGV exploration trajectories. Ground robots are preferable for open areas, while aerial robots are preferable for cluttered environments. Ropero et al. [3] propose a path-planning algorithm for cooperative UGV–UAV exploration. Their strategy employs the ground robot as a mobile charging station to address the aerial robot's energy constraints, while the aerial robot reaches target points to overcome the ground robot's functionality limits. However, these multi-robot systems introduce increased system-level complexity, making deployment more challenging.

Not only the joint planning problem and communication problem between two platforms need to be solved, but the multi-robot SLAM problem also has to be concerned for an integrated system. To achieve scalability to multi-robot SLAM loop closures, Team CSIRO [5] directly models the frontiers on point clouds to avoid dense volumetric map representations. Qin et al. [4] implement a two-layered exploration strategy, where the ground robot generates a coarse environment model, and the aerial robot produces the 3D fine mapping according to the coarse map.

Combining the advantages of ground and aerial robots without introducing the complexity of multiple platforms, TABVs have become a research hotspot [16]–[19]. However, works using TABVs for exploration are rare. Rollocopter, the TABV of Team CoSTAR has been deployed for DARPA Subterranean Challenge. However, as just one among dozens of platforms, they utilize a general exploration planner without specifically addressing the schedule of the bimodal vehicle [20].

In this work, we extend the MCTS method to address the bimodal, energy- and time-constrained exploration problem. By integrating it with the improved bimodal motion planner, we form a complete TABV exploration framework.

III. Problem Statement and System Overview

The goal of the proposed method is to explore an initially unknown but bounded 3D space using a TABV under a given energy budget E_{all} and time budget T_{all}. Importantly, the objective is *not* full coverage of the environment, but to collect as much informative data as possible and ensure that the robot can return to the departure station with the data, which is a more reliable approach in practical application such as communication-denied post-disaster environments.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

To achieve this, we formulate the exploration task as selecting a sequence of viewpoints and their associated modalities that maximizes the perception of unknown space, while ensuring execution within the available energy and time budgets. Formally, the problem is defined as:

$$\boldsymbol{\mathcal{P}}^* = \text{argmax}_{\boldsymbol{\mathcal{P}}}\, IG(\boldsymbol{\mathcal{P}}) \tag{1a}$$

$$\text{s.t.} \quad E_r(\boldsymbol{\mathcal{P}}) \geq 0, \tag{1b}$$

$$T_r(\boldsymbol{\mathcal{P}}) \geq 0, \tag{1c}$$

where $\boldsymbol{\mathcal{P}} = \{\mathcal{P}_i \mid i = 0, \ldots, n\}$ is a sequence of selected viewpoints. $IG(\boldsymbol{\mathcal{P}})$ measures the information gain along the trajectory, and $E_r(\boldsymbol{\mathcal{P}})$, $T_r(\boldsymbol{\mathcal{P}})$ denote the remaining energy and time after visiting all selected viewpoints and returning to the departure station.

To cope with the inequality constraints and environmental uncertainty, we convert (1) into an unconstrained optimization problem using a penalty function approach:

$$\boldsymbol{\mathcal{P}}^* = \text{argmin}_{\boldsymbol{\mathcal{P}}} \left(-IG(\boldsymbol{\mathcal{P}}) + \kappa_{E_r}(E_r(\boldsymbol{\mathcal{P}})) + \kappa_{T_r}(T_r(\boldsymbol{\mathcal{P}}))\right), \tag{2}$$

where $\kappa_{E_r}(\cdot)$ and $\kappa_{T_r}(\cdot)$ are exponential penalty terms based on remaining energy and time, which will be detailed in Sec. V-A. Considering the practical scenario, exceeding the energy limit may prevent return and thus risk mission failure, while moderate time overruns are more tolerable. Therefore, we adopt a steeper penalty curve for energy, while the time-related penalty is relatively moderate and designed to take effect only when the remaining energy is already sufficient. This formulation encourages safe and efficient planning by maintaining a margin to cope with environmental uncertainty.

An overview of our TABV autonomous exploration system is presented in Fig. 2. We first extract the exploration information from environment and generate potential viewpoints (Sec. IV). Then we conduct BM-MCTS to determine the traverse sequence of the viewpoints (Sec. V) under the given energy and time budget. Finally, we use the bimodal motion planner to generate bimodal trajectories (Sec. VI).

IV. Exploration Information Extraction

A. Bimodal Viewpoints Generation

Similar to classic frontier-based exploration, frontiers are defined as free voxels adjacent to unknown space. As the robot moves, new frontiers appear. The robot iteratively plans viewpoints to cover frontiers and replans as the map updates.

To efficiently generate viewpoints covering frontiers, we first group the frontier voxels into clusters. For each frontier cluster C, we consider two strategies for generating viewpoints to achieve full coverage: i) **Aerial-Modal Strategy (AS)**: The cluster is covered exclusively using aerial viewpoints. ii) **Hybrid-Modal Strategy (HS)**: The cluster is first covered using terrestrial viewpoints; if full coverage cannot be achieved, additional aerial viewpoints are selected to complete the coverage. The AS and HS sets provide diverse modality-specific candidate viewpoints, which serve as inputs for the subsequent decision-making in BM-MCTS.

Algorithm 1: BM-MCTS

```
Main Function Search(v_0):
    while number of iterations is less than threshold do
        v_s = Select(v_0)
        [succ, v_e] = Expand(v_s, Reward)
        if succ then
            Simulate(v_e)
            BackPropagate(v_s, v_e)
    return BestChild(v_0)
```

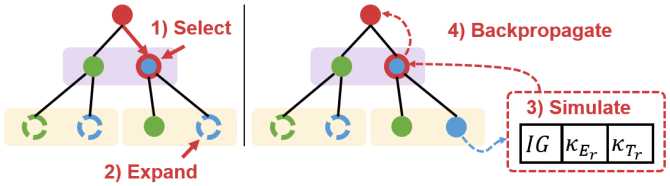

Fig. 3: An illustration of the process of BM-MCTS.

Specifically, as shown in Fig. 2(b), for each C, we first sample raw candidate terrestrial and aerial viewpoints. Raw aerial viewpoints $\boldsymbol{\mathcal{P}}_{\text{aerial}}$ are generated using cylindrical coordinate sampling around C. Raw terrestrial viewpoints $\boldsymbol{\mathcal{P}}_{\text{ground}}$ are sampled from nearby traversable ground voxels extracted from the grid map, with a fixed vertical offset. Given these raw viewpoints, the problem of selecting bimodal viewpoints to cover C exhibits submodularity [21]. We apply a greedy method to efficiently solve it and obtain AS and HS viewpoint sets. Details and algorithm are provided in the Sec. 1 of the supplementary material [22].

Finally, we obtain two sets of candidate viewpoints representing alternative coverage strategies: $\boldsymbol{\mathcal{P}}_{\text{AS}}$ for pure aerial-modal coverage, and $\boldsymbol{\mathcal{P}}_{\text{HS}}$ for hybrid-modal coverage. Each viewpoint is defined as $\mathcal{P}_i = (\mathbf{p}, \phi)$, where $\mathbf{p}$ is the position and ϕ is the yaw angle. The IG of $\mathcal{P}_i$ is defined as the number of visible frontier voxels within $\mathcal{P}_i$'s FoV. All viewpoints belonging to the same cluster compose a viewpoint group $\mathcal{G} = \{\boldsymbol{\mathcal{P}}_{\text{AS}}, \boldsymbol{\mathcal{P}}_{\text{HS}}\}$.

B. Energy and Time Cost between Viewpoints

As we choose E_r and T_r as criteria to achieve energy- and time-aware planning, the energy and time cost should be modeled for the estimation process. Given two viewpoints $\mathcal{P}_i$ and $\mathcal{P}_j$, the time cost between them is defined as:

$$T(\mathcal{P}_i, \mathcal{P}_j, M) = \max\left\{\frac{\textit{length}(\mathcal{P}_i, \mathcal{P}_j)}{v_{M,\max}}, \frac{\textit{dyaw}(\mathcal{P}_i, \mathcal{P}_j)}{\omega_{\text{M},\max}}\right\}, \tag{3}$$

where $\textit{length}(\mathcal{P}_i, \mathcal{P}_j)$ is the length of the searched path between $\mathcal{P}_i$ and $\mathcal{P}_j$, $\textit{dyaw}(\mathcal{P}_i, \mathcal{P}_j)$ is the minimum yaw angle difference between the two viewpoints, $v_{M,\max}$ and $\omega_{\text{M},\max}$ are maximum speed and maximum yaw angular speed in modality M, respectively. For the energy cost, we use a simple constant power model, and the energy cost between them is defined as:

$$E(\mathcal{P}_i, \mathcal{P}_j, M) = P_M \cdot T(\mathcal{P}_i, \mathcal{P}_j, M), \tag{4}$$

where P_M is the average power under modality M. M denotes the modality: $M \in \{T, A\}$. T and A indicate ter-

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

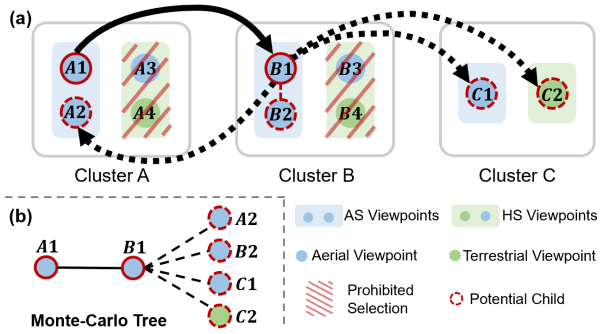

Fig. 4: An example for determining the potential children. (a): The potential children of the viewpoint $B1$ of the frontier cluster B. If the viewpoint belongs to a cluster that has already been expanded, then the viewpoints with the other modality are prohibited from selection. (b): The corresponding Monte Carlo tree. Each branch represents a viewpoint traversal sequence.

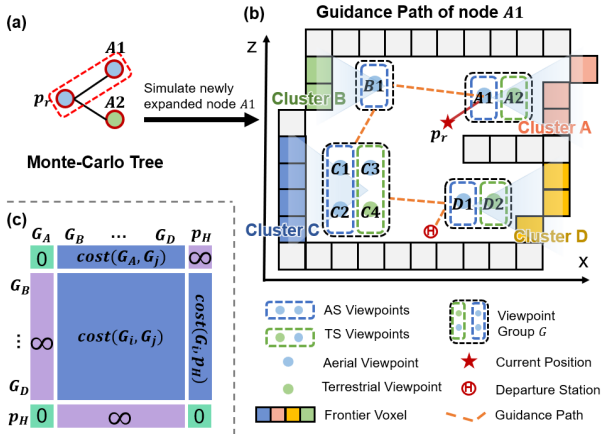

Fig. 5: An example of the guidance path generation for a newly expanded node. (a): The expanding Monte Carlo tree, where node $A1$ is newly expanded and requires simulation. (b): The guidance path generation for node $A1$. In this case, $A1$ is chosen in $\mathcal{G}_A$ to cover cluster A, and the path from robot's position p_r to $A1$ is determined. Then we solve a grouped TSP to get the whole guidance path that traverses all clusters and returns home. (c): The cost matrix of grouped TSP. The purple region denotes the infinity connection, and the green region denotes the zero connection.

restrial modality and aerial modality, respectively. Specially, we denote $M = (T + A)/2$ as the average modality which means $v_{(T+A)/2,\max} = (v_{T,\max} + v_{A,\max})/2$ and $P_{(T+A)/2} = (P_T + P_A)/2$.

And if the time and energy cost between $\mathcal{G}$s need to be estimated, taking $E(\mathcal{G}_i, \mathcal{G}_j, M)$ for example, we use an average viewpoint with the mean position and yaw angle of all viewpoints in $\mathcal{G}$ for calculation.

According to existing bimodal vehicle design papers [18, 23, 24], the average power in aerial modality is about $5 \sim 8\times$ higher ($7.2\times$ for ours) than terrestrial modality. And the maximum speed in terrestrial modality is about $1m/s$. However, the maximum speed in aerial modality can not be achieved $5 \sim 8\times$ faster in exploration due to safety and perception accuracy. So the terrestrial modality has an advantage in energy cost, while the aerial modality results in less time cost.

V. Bimodal Monte Carlo Tree Search

The Monte Carlo Tree Search [25, 26] is a planning method for finding the optimal decision within a given horizon under limited computational resources. Based on that, we propose an extended MCTS method, called BM-MCTS, to select the optimal viewpoint sequence with two potential modalities. The BM-MCTS method consists of four key steps: selection, expansion, simulation, and backpropagation, as shown in Fig. 3. After the required iterations, the best child node of the root is selected as the next local goal. And the best branch corresponds to the resulting bimodal path. The whole algorithm is shown as Alg. 1.

A. Tree Structure and Reward

First, we define the structure of the tree. Each tree node v corresponds to a viewpoint $\mathcal{P}$, and each branch determines a viewpoint traversal sequence, as shown in Fig. 4(b).

Each node v is associated with several attributes. Specifically, $E_R(v)$ and $T_R(v)$ denote the estimated remaining energy/time when the robot reaches the viewpoint $\mathcal{P}$ corresponding to node v, by following the sequence of viewpoints along the current search branch. $E_r(v)$ and $T_r(v)$ represent the estimated remaining energy/time *after* the robot reaches $\mathcal{P}$, completes visiting the remaining selected viewpoints, and returns to the departure station.

As for connections between nodes, it is necessary to specify the potential children of each node. In the case of a viewpoint $\mathcal{P}_i$, its potential child $\mathcal{P}_j$ should satisfy the following conditions: 1). $\mathcal{P}_j$ is not already included in the branch of $\mathcal{P}_i$. 2). The distance between $\mathcal{P}_i$ and $\mathcal{P}_j$ is smaller than the threshold. 3). If the viewpoint group to which $\mathcal{P}_j$ belongs has been expanded in the branch of $\mathcal{P}_i$, $\mathcal{P}_j$ should maintain the same sampling mode, as shown in Fig. 4. If no candidate $\mathcal{P}_j$ satisfies all the conditions, we relax the second condition to allow feasible selection.

Second, we define the reward of the nodes for tree search. As formulated in (2), the reward function consists of *process gain*, represented by IG, and *terminal cost*, represented by E_r and T_r. The process gain is related to all nodes in the subtree of this node. While the terminal cost is only related to the leaf nodes in the subtree of this node. Each node v has reward $\mathbf{R}(v) = [R_p(v), R_t(v)]$, where R_p is the process gain and R_t is the terminal cost:

$$R_p(v) = IG(v)/n_{\text{IG}}(v), \tag{5a}$$

$$R_t(v) = \kappa_{E_r}(E_r(v)) + \kappa_{T_r}(T_r(v)), \tag{5b}$$

where $n_{\text{IG}}(v)$ denotes the discounted visitation count of node v, which is incremented by a fixed weight $\gamma_{\text{IG}} = 0.8$ during each backpropagation step (Alg. 2). Correspondingly, $IG(v)$ accumulates the discounted IG from its child nodes during backpropagation. $R_p(v) = IG(v)/n_{\text{IG}}(v)$ therefore reflects the discounted average IG across the subtree rooted at v. This formulation mitigates the bias caused by uneven subtree sizes and emphasizes information from nearby nodes. $\kappa_{E_r}(E_r(v))$ and $\kappa_{T_r}(T_r(v))$ denote the average terminal energy and time cost computed over all *leaf* nodes in the subtree rooted at v. The costs are constructed using the exponential functions defined as $\kappa_{E_r}(x) = exp(-a_1 \cdot x/E_{all} + b_1)$, $\kappa_{T_r}(x) = exp(-a_2 \cdot x/T_{all} + b_2)$, where a_i, b_i are preset hyperparameters.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

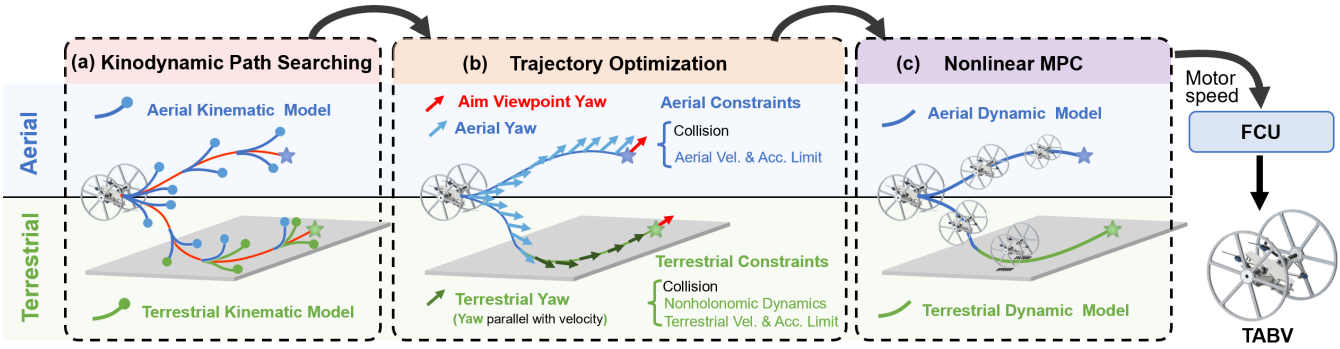

Fig. 6: The hierarchical bimodal motion planning framework. (a): The kinodynamic path search front end. (b): The flatness-differential-based spatial-temporal trajectory optimization back end. (c): The NMPC module [6] to calculate the desired motor speed.

B. Selection and Expansion Process

In the selection process, we recursively choose the optimal node from the root until reaching a node with unexpanded potential children. The selection policy follows the upper confidence bound (UCB) rule [26], balancing exploration and exploitation. In the function *BestChild*(v), we compute the UCB score $\boldsymbol{U}(v_k)$ for each child node v_k of v:

$$\boldsymbol{U}(v_k) = G(v_k) - \sqrt{2\ln n_s / n(v_k)}, \tag{6a}$$

$$G(v_k) = -N(R_p(v_k)) + R_t(v_k), \tag{6b}$$

where $n(v_k)$ is the number of times v_k has been selected so far. $n_s = \sum_{k=1}^{K} n(v_k)$, where K is the number of v's child nodes. $N(x)$ linearly maps x to $[\epsilon, 1]$, where $\epsilon = 0.05$ represents the smallest normalized gain. Among all child nodes, the one with the minimum $\boldsymbol{U}(\cdot)$ is selected.

Given the selected node v_i to be expanded corresponding to viewpoint $\mathcal{P}_i$, its potential children are determined as Sec. V-A. A child v_{i+1} is randomly selected from v_i's unexpanded potential children. Then we update the left energy E_R and left time T_R when robot reaches v_{i+1}:

$$E_R(v_{i+1}) = E_R(v_i) - E(\mathcal{P}_i, \mathcal{P}_{i+1}, M(\mathcal{P}_{i+1})), \tag{7}$$

where $E(\mathcal{P}_i, \mathcal{P}_{i+1}, M(\mathcal{P}_{i+1}))$ is the energy consumption from $\mathcal{P}_i$ to $\mathcal{P}_{i+1}$, and the modality is determined by the latter one. As E_r and T_r are handled similarly, we only present the energy-related equations here and in the following sections.

C. Simulation Process with Guidance Path

The simulation process updates the reward of the newly expanded node, which requires evaluating its IG, E_r, and T_r. For IG, $IG(v_i)$ is initialized as the number of visible frontier voxels within $\mathcal{P}_i$'s FoV, serving as an estimate of newly gathered information. To estimate the energy and time required for traversing through all clusters and returning home, we solve an extended grouped Traveling Salesman Problem to generate a guiding path, as shown in Fig. 5.

We design the cost matrix, as shown in Fig. 5(c), to set the departure station p_H as the final destination. The cost between viewpoint groups is defined as the travel time between them, given by $T(\mathcal{G}_i, \mathcal{G}_j, (T+A)/2)$. Since the exact modality of nodes is unknown during evaluation, we set $M = (T+A)/2$.

The main computational bottleneck in the simulation process is the cost matrix calculation. To efficiently estimate the feasible path length, we maintain a global topo-graph that records the visited positions at intervals and connects nodes within a distance threshold, as shown in Fig. 7(c). By performing an $A^\star$ search on the topo-graph, we obtain a fast conservative path length estimation. Moreover, paths between viewpoint groups are incrementally updated during viewpoint generation, further improving efficiency.

Then the simulation step is executed to update E_r, T_r:

$$\begin{aligned} E_r(v_i) = & E_R(v_i) - \sum_{q=i}^{k-1} E(\mathcal{G}_q, \mathcal{G}_{q+1}, (T+A)/2) \\ & - E(\mathcal{G}_k, \mathbf{p}_H), \end{aligned} \tag{8}$$

where v_i corresponds to $\mathcal{P}_i \in \mathcal{G}_i$, and the last two terms of the equation represent the estimated energy consumption for the robot traveling from the current viewpoint group to the last one and returning home according to the guidance path.

D. Backpropagation Process

The obtained reward should be backpropagated to each visited node to update the reward of those nodes, preparing for the next iteration of selection. As the process gain R_p is related to all the nodes on the same branch, the cumulative backpropagation is adopted. While the terminal cost R_t is only related to the leaf node on the branch, we adopt the average backpropagation for it. The backpropagation process is detailed in Alg. 2.

E. Prune Condition

To reduce the search space and improve efficiency, a child node is pruned and no longer expanded if the remaining energy

Algorithm 2: Backward Process

Function *BackPropagate*(v, v_k)**:**
- $n(v) = n(v) + 1$
- $n_{\text{IG}}(v) = n_{\text{IG}}(v) + \gamma_{\text{IG}}$
- $IG(v) = IG(v) + \gamma_{\text{IG}} IG(v_k)$
- $\kappa_{E_r}(v) = \sum_{v_c \in v.kid} \kappa_{E_r}(v_c)$ / Num($v.kid$)
- $\kappa_{T_r}(v) = \sum_{v_c \in v.kid} \kappa_{T_r}(v_c)$ / Num($v.kid$)
- *BackPropagate*($v.parent$, v)

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

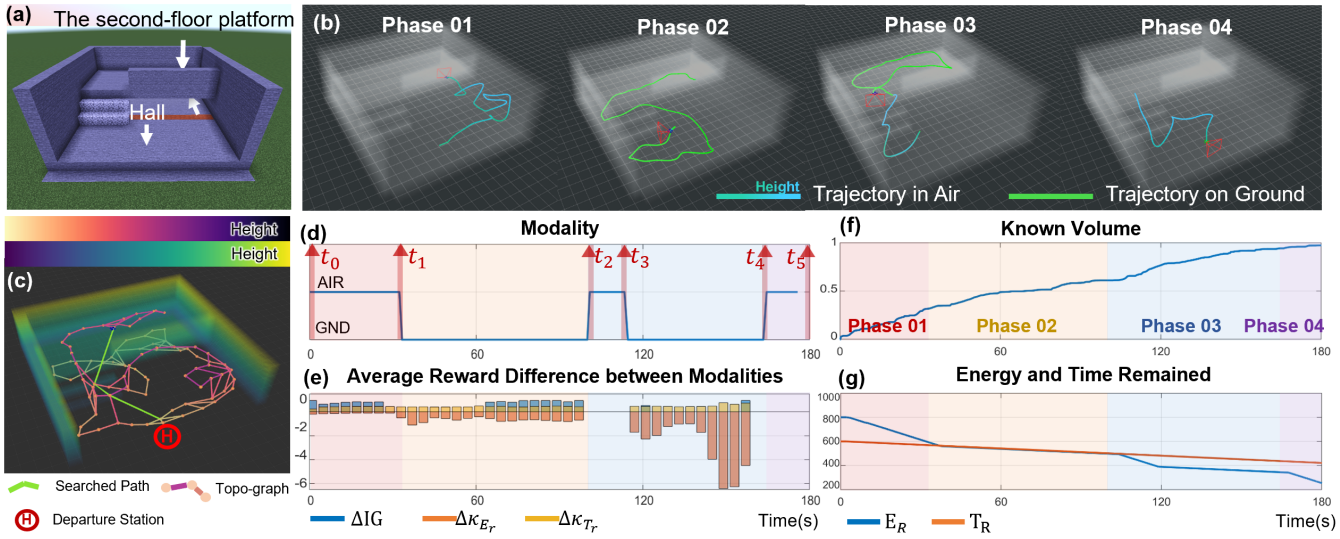

Fig. 7: (a) A two-story house scene for TABV exploration. (b) Four exploration phases. *Phase 01*: Fly to explore most of the hall. *Phase 02*: Cover the first-floor platform, then fly upstairs. *Phase 03*: Roll to cover the second-floor platform. *Phase 04*: Cover the remaining hall and return home. (c) Topo-graph illustration. (d) TABV modalities. (e) Average reward difference between aerial and terrestrial child nodes at the second depth in the Monte Carlo tree. (f) Coverage ratio over time. (g) Remaining energy and time.

after reaching node v_i is less than the energy required to return home from v_i:

$$E_R(v_i) - E(\mathcal{G}_i, \mathbf{p}_H) < 0, \tag{9}$$

where v_i corresponds to $\mathcal{P}_i \in \mathcal{G}_i$.

Further details of algorithm analysis are presented in Sec. 5 of the supplementary material [22].

VI. Bimodal Motion Planning

Given the next goal from BM-MCTS, we use a hierarchical bimodal motion planning method to generate terrestrial-aerial hybrid trajectories and control the TABV to execute. Building upon our previous work [6], we briefly summarize the overall pipeline here and highlight the key enhancements.

As shown in Fig. 6, the planner follows a standard hierarchical architecture with a kinodynamic path search front end and a spatial-temporal trajectory optimization back end. To better support our exploration framework, we introduce several key enhancements.

First, instead of assuming a fixed ground plane, we perform online terrain perception using incremental ground segmentation. This allows the system to dynamically identify traversable surfaces in complex, multi-level environments without relying on predefined structural assumptions.

Second, we introduce modality-aware planning to replace the prior approach that consistently favored the terrestrial modality. In the front end, motion primitives are selected according to the target modality: aerial primitives for aerial targets, and bimodal primitives with penalties on aerial motion for terrestrial targets. In the back end, modality-specific constraints are applied—such as nonholonomic dynamics for terrestrial segments. Additionally, we also integrate bimodal yaw planning into the back end.

Lastly, to enhance safety, we compute the Euclidean Signed Distance Field (ESDF) for each ground segment to query the distance to edges, ensuring the TABV flies safely near edges and prevents falls.

Finally, by incorporating this bimodal motion planner, a complete autonomous terrestrial-aerial exploration framework is established. More details are provided in the Sec. 2 of the supplementary material [22].

VII. Result

A. Simulation

To validate our TABV exploration system, we conduct simulations and phased analyses in multiple multilayered buildings. Moreover, we analyze the adaptability of the BM-MCTS under different budgets, as well as the relationship between solution quality and iteration times. Based on our TABV platform, the power of the aerial modality is 7 times that of the terrestrial modality. So we set $P_T = 1$, and $P_A = 7$, meaning that one second of movement in terrestrial mode consumes 1 unit of energy, while one second in aerial mode consumes 7 units. To avoid redundancy, we omit the units for time and energy in the remainder of this section. Moreover, the maximum velocity is set as $v_{T,\max} = 0.5$ m/s, $v_{A,\max} = 1.0$ m/s.

*1) **Bimodal Exploration for A Two-story House** :* We construct a 15 m×15 m×6 m two-story house in MineCraft as an exploration scene (Fig. 7(a)). The first floor of the house is $2m$ high, connected by a staircase to a second-floor platform. The energy budget is set to 1000, and the time budget is set to 600. The hyper-parameters κ_{E_r} and κ_{T_r} are set identically to those described in Sec. 9 of the supplementary material [22]. In Fig. 7(e), the average reward difference is defined as $\Delta\bar{\mathbf{R}} = \bar{\mathbf{R}}(v_A) - \bar{\mathbf{R}}(v_G)$, where $\bar{\mathbf{R}}(v_A)$ (or $\bar{\mathbf{R}}(v_G)$) is the average reward of the root node's aerial (or terrestrial) modality child nodes.

We divided the whole exploration process into four phases for detailed analysis (Fig. 7). In Phase 01, both of the remaining energy and time are sufficient, so the IG reward

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

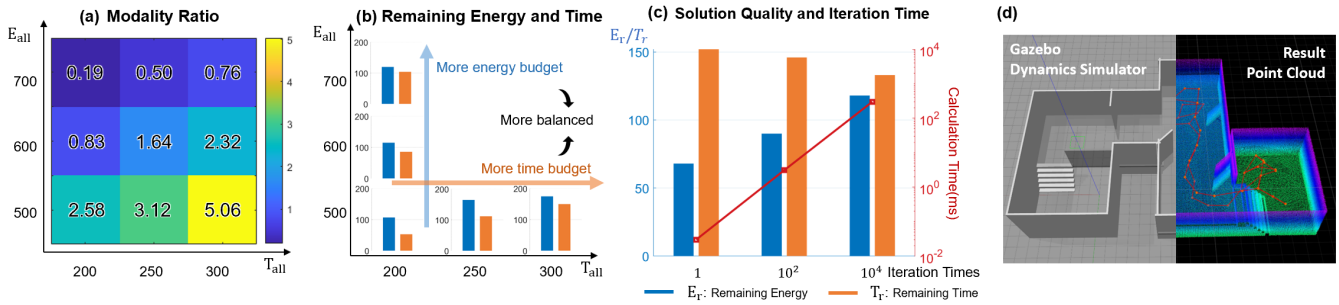

Fig. 8: (a)-(b): Performance under different energy/time budgets. (a): The ratio of modality under different budgets, where the value represents the ratio of time on the ground to time in the air. (b): The remaining energy/time when exploration is completed. (c): The algorithm calculation time and the remaining energy/time when exploration is completed under different iteration times. (d): More simulation scenario and exploration result provided in the supplementary material [22].

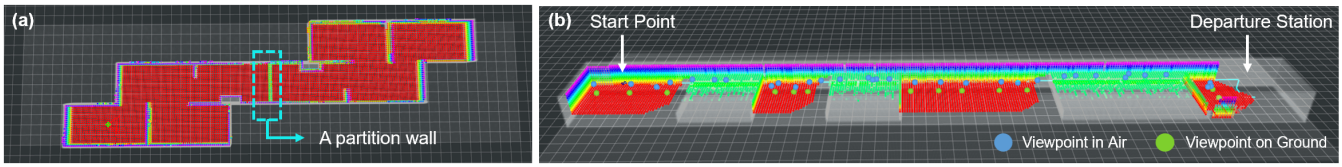

Fig. 9: Simulation scenes. (a): A multi-room scene with a 0.5m height partition wall. (b): A scene with a series of viewpoints.

takes advantage. The TABV switches to aerial modality for more information gain along with a shorter time. At t_1, the TABV gets into the first-floor platform. As the platform height is only 2 m, the IG reward is almost the same under two modalities. Since a lot of energy is consumed in Phase 01, energy counts more than time t_1, leading the TABV to choose the terrestrial modality for the entire first-floor platform. At t_2, after covering all terrestrial viewpoints, the TABV takes off and lands on the second floor at t_3, then rolls to cover the second-floor platform to save energy. In Phase 04, with only aerial viewpoints remaining in the hall, the TABV flies to cover them and returns home.

*2) **Adaptability to Different Budgets**:* We show the performance of our method under different energy and time budgets. The simulation scene is a $40\ \mathrm{m} \times 16\ \mathrm{m} \times 2\ \mathrm{m}$ office-like area, which is completely separated by a $1\ \mathrm{m}$ high partition wall (as shown in Fig. 9(a)). We define the modality ratio as the time on the ground divided by the time in the air. Experiments are conducted under varying energy and time budgets, with five trials for each setting. The results are shown in Fig. 8. As the time budget increases and the energy budget decreases, TABV tends to favor the terrestrial modality. We further analyze the remaining energy and time at the end of exploration, as shown in Fig. 8(b). With the same energy budget, a larger time budget results in more remaining energy; with the same time budget, a larger energy budget results in more remaining time. This demonstrates that the increasing one provides greater flexibility for more efficient use of the other, highlighting the flexibility of the BM-MCTS method in balancing energy and time, and enabling the TABV to adapt to different budgets.

*3) **Solution Quality and Iteration Times**:* Since the BM-MCTS method can produce solutions at any time when reaching the iteration threshold, we analyze the relationship between the iteration times and the quality of the solution. We built a one-way scene as Fig. 9(b) is shown to ensure consistent exploration direction in the test. The TABV first flies along a preset trajectory from the departure station to the start point to generate a series of bimodal viewpoints. Then the exploration starts with energy and time budget set to 400 and 200, respectively. For each preset iteration time, we simulate for ten times. As shown in Fig. 8(c), as the maximum iteration time increases, the remaining energy and time become more balanced. This is because the energy and time consumption estimates become more accurate as the search tree expands, at the cost of increased calculation time.

Further simulation tests and comprehensive comparisons are presented in Sec. 6-8 of the supplementary material [22].

B. Real-World Experiment

To demonstrate and verify the proposed approaches in real-world environments, we use a customized TABV platform, as shown in Fig. 10. The TABV weighs 1.85 kg, equipped with a Livox 360 and a Jetson Xavier NX. We use FAST-LIO2 [27] for localization. All the algorithms are running onboard. The TABV consumes 648 W in the aerial modality and 90 W in the terrestrial modality, making the aerial power consumption 7.2 times higher. We set a 20 Wh serve energy budget and a 900 s abundant time budget for the TABV to explore an underground parking garage. The TABV presents the adaptation to fly across uneven stairs and switch to terrestrial modality to explore with low energy cost. Finally, the TABV returns home safely when the energy budget is used up.

VIII. Conclusion

In this paper, we develop a hierarchical scheme to drive the TABV to explore under given energy and time budget. With this scheme, the TABV can flexibly respond to different environments and energy/time constraints by changing the modality. A detailed analysis of the system's limitations is provided in Sec. 11 of the supplementary material [22]. For future work, we will add environment prediction for more accurate energy/time consumption estimation.

IEEE Robotics and Automation Letters (RA-L) paper, presented at ICRA 2026, Vienna, Austria. Cite as RA-L paper.

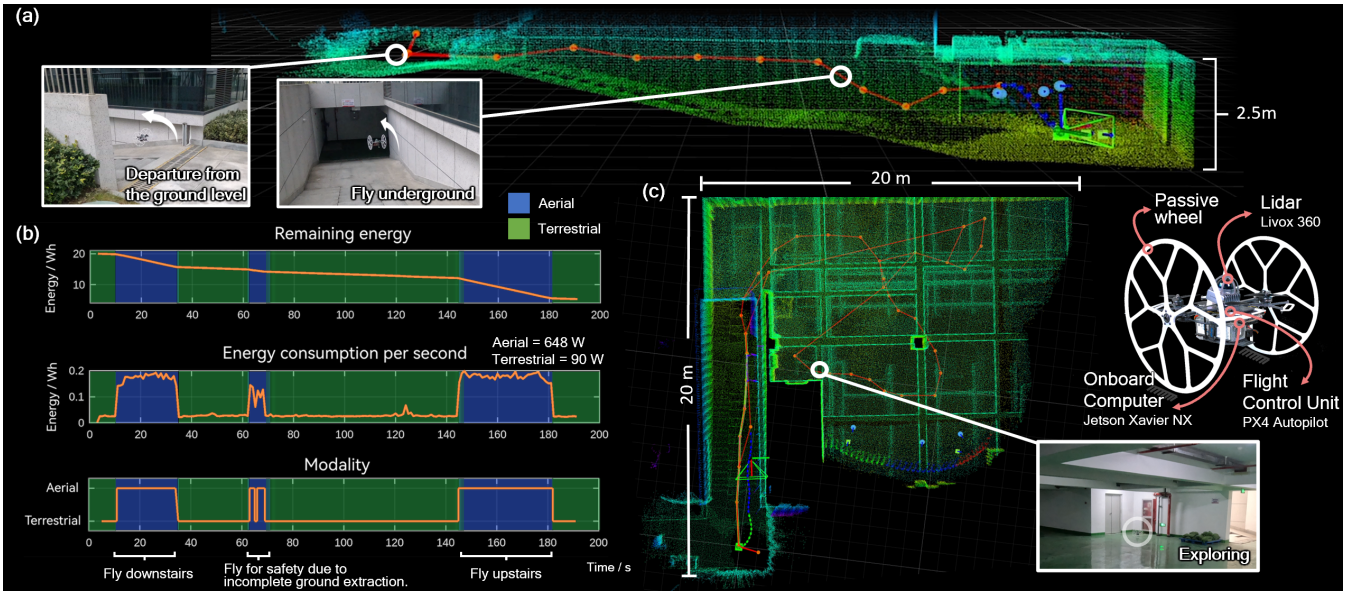

Fig. 10: Real-world experiment. (a): Side view of the entrance to the underground parking garage. (b): The energy and modality curves. (c): Top view of the explored map.

REFERENCES

[1] J. Butzke, A. Dornbush, and M. Likhachev, "3-d exploration with an air-ground robotic system," in *Proc. of the IEEE/RSJ Intl. Conf. on Intell. Robots and Syst.*, 2015, pp. 3241–3248.

[2] L. Wang, F. Gao, F. Cai, and S. Shen, "Crash: A collaborative aerial-ground exploration system using hybrid-frontier method," in *Proc. of the IEEE Intl. Conf. on Robot. and Bio. (ROBIO)*, 2018, p. 2259.

[3] F. Ropero, P. Muñoz, and M. D. R-Moreno, "Terra: A path planning algorithm for cooperative ugv–uav exploration," *Engineering Applications of Artificial Intelligence*, vol. 78, pp. 260–272, 2019.

[4] H. Qin, Z. Meng, W. Meng, X. Chen, H. Sun, F. Lin, and M. H. Ang, "Autonomous exploration and mapping system using heterogeneous uavs and ugvs in gps-denied environments," *IEEE Trans. on Veh. Tech.*, vol. 68, no. 2, pp. 1339–1350, 2019.

[5] J. Williams, S. Jiang, M. O'Brien, G. Wagner, E. Hernandez, M. Cox, A. Pitt, R. Arkin, and N. Hudson, "Online 3d frontier-based ugv and uav exploration using direct point cloud visibility," in *2020 IEEE Intl. Conf. on Multisensor Fusion and Integration for Intell. Sys. (MFI)*. IEEE, 2020, pp. 263–270.

[6] R. Zhang, J. Lin, Y. Wu, Y. Gao, C. Wang, C. Xu, Y. Cao, and F. Gao, "Model-based planning and control for terrestrial-aerial bimodal vehicles with passive wheels," in *Proc. of the IEEE/RSJ Intl. Conf. on Intell. Robots and Syst.* IEEE, 2023, pp. 1070–1077.

[7] B. Yamauchi, "A frontier-based approach for autonomous exploration," in *Proc. of the IEEE Intl. Symp. on Comput. Intell. in Robot. and Autom.* IEEE, 1997, pp. 146–151.

[8] T. Cieslewski, E. Kaufmann, and D. Scaramuzza, "Rapid exploration with multi-rotors: A frontier selection method for high speed flight," in *Proc. of the IEEE/RSJ Intl. Conf. on Intell. Robots and Syst.* IEEE, 2017, pp. 2135–2142.

[9] C. Cao, H. Zhu, H. Choset, and J. Zhang, "Tare: A hierarchical framework for efficiently exploring complex 3d environments," in *Proc. of Robot.: Sci. and Syst. (RSS)*, 2021.

[10] B. Zhou, Y. Zhang, X. Chen, and S. Shen, "Fuel: Fast uav exploration using incremental frontier structure and hierarchical planning," *IEEE Robot. Autom. Lett.*, vol. 6, no. 2, pp. 779–786, 2021.

[11] Y. Gao, Y. Wang, X. Zhong, T. Yang, M. Wang, Z. Xu, Y. Wang, Y. Lin, C. Xu, and F. Gao, "Meeting-merging-mission: A multi-robot coordinate framework for large-scale communication-limited exploration," in *Proc. of the IEEE/RSJ Intl. Conf. on Intell. Robots and Syst.* IEEE, 2022, pp. 13 700–13 707.

[12] A. Bircher, M. Kamel, K. Alexis, H. Oleynikova, and R. Siegwart, "Receding horizon" next-best-view" planner for 3d exploration," in *Proc. of the IEEE Intl. Conf. on Robot. and Autom. (ICRA)*. IEEE, 2016, pp. 1462–1468.

[13] G. Best, O. M. Cliff, T. Patten, R. R. Mettu, and R. Fitch, "Dec-mcts: Decentralized planning for multi-robot active perception," *Intl. J. Robot. Research (IJRR)*, vol. 38, no. 2-3, pp. 316–337, 2019.

[14] K. M. Seiler, F. H. Kong, and R. Fitch, "Multi-horizon multi-agent planning using decentralised monte carlo tree search," *IEEE Robot. Autom. Lett.*, 2024.

[15] S. Bone, L. Bartolomei, F. Kennel-Maushart, and M. Chli, "Decentralised multi-robot exploration using monte carlo tree search," in *Proc. of the IEEE/RSJ Intl. Conf. on Intell. Robots and Syst.* IEEE, 2023, pp. 7354–7361.

[16] Z. Zheng, Q. Cai, J. Wang, X. Xu, M. Cao, H. Yu, J. Li, J. Meng, and G. Lu, "Capsulebot: A novel hybrid aerial-ground bi-copter robot with two actuated-wheel-rotors," *IEEE Robot. Autom. Lett.*, vol. 10, no. 1, pp. 120–127, 2025.

[17] M. Cao, X. Xu, S. Yuan, K. Cao, K. Liu, and L. Xie, "Doublebee: A hybrid aerial-ground robot with two active wheels," in *Proc. of the IEEE/RSJ Intl. Conf. on Intell. Robots and Syst.*, 2023, pp. 6962–6969.

[18] J. Lin, R. Zhang, N. Pan, C. Xu, and F. Gao, "Skater: A novel bi-modal bi-copter robot for adaptive locomotion in air and diverse terrain," *IEEE Robot. Autom. Lett.*, vol. 9, no. 7, pp. 6392–6399, 2024.

[19] J. Tang, R. Zhang, K. Beyduz, Y. Jiang, C. Wiebe, H. Zhang, O. Asoro, and M. W. Mueller, "Duawlfin: A drone with unified actuation for wheeled locomotion and flight operation," *arXiv preprint arXiv:2505.13836*, 2025.

[20] B. Morrell, R. Thakker, À. Santamaria Navarro, A. Bouman, X. Lei, J. Edlund, T. Pailevanian, T. S. Vaquero, Y. L. Chang, T. Touma *et al.*, "Nebula: Team costar's robotic autonomy solution that won phase ii of darpa subterranean challenge," *Field robotics*, vol. 2, pp. 1432–1506, 2022.

[21] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, "An analysis of approximations for maximizing submodular set functions—i," *Mathematical programming*, vol. 14, pp. 265–294, 1978.

[22] Y. Gao, *Supplementary Materials for Autonomous Exploration with Terrestrial-Aerial Bimodal Vehicles*, May 2025. [Online]. Available: https://doi.org/10.5281/zenodo.16545839

[23] D. D. Fan, R. Thakker, T. Bartlett, M. B. Miled, L. Kim, E. Theodorou, and A.-a. Agha-mohammadi, "Autonomous hybrid ground/aerial mobility in unknown environments," in *Proc. of the IEEE/RSJ Intl. Conf. on Intell. Robots and Syst.* IEEE, 2019, pp. 3070–3077.

[24] A. Kalantari, T. Touma, L. Kim, R. Jitosho, K. Strickland, B. T. Lopez, and A.-A. Agha-Mohammadi, "Drivocopter: A concept hybrid aerial/ground vehicle for long-endurance mobility," in *IEEE Aero. Conf.* IEEE, 2020, pp. 1–10.

[25] R. Coulom, "Efficient selectivity and backup operators in monte-carlo tree search," in *International conference on computers and games*. Springer, 2006, pp. 72–83.

[26] W. Chen and L. Liu, "Pareto monte carlo tree search for multi-objective informative planning," in *Proc. of Robot.: Sci. and Syst. (RSS)*, 2019.

[27] W. Xu, Y. Cai, D. He, J. Lin, and F. Zhang, "Fast-lio2: Fast direct lidar-inertial odometry," *IEEE Trans. Robot. (TRO)*, vol. 38, no. 4, pp. 2053–2073, 2022.