Autonomous Cave Exploration using Aerial Robots

Mihir Dharmadhikari[1], Huan Nguyen[2], Frank Mascarich[1], Nikhil Khedekar[1], and Kostas Alexis[2]

***Abstract*— In this paper we present the complete system design for an aerial robot capable of autonomous exploration inside natural cave environments. Cave networks involve diverse and complicated topologies, complex geometries and degraded conditions rendering the process of robotic mapping a particularly daunting adventure. In response to these challenges, we outline the core algorithmic modules relating to localization and mapping, exploration path planning and control, alongside the developed perception and computing solutions onboard an aerial robot tailored to undertake such complex tasks given no prior information for the cave environments in which it is deployed. A set of extensive results is presented including both simulation studies in multi-branching and maze-like cave environments, as well as field experiments inside the Moaning Caverns natural cave environment in California, US.**

I. INTRODUCTION

Autonomous exploration of cave environments has recently attracted the interest of the scientific community. Pioneering studies test and field robotic systems in cave settings thus pushing the frontier of autonomy and resilience across the domains of mapping, path planning and broadly autonomy [1–3]. Accelerated by the kick-off of the DARPA Subterranean Challenge, research efforts around the world investigate the particular challenges for long-term underground autonomy. Caves, in particular, correspond to an ultimate test for robotic systems as they present diverse, complex, often extremely large-scale and at the same time, at instances particularly narrow geometries, while also imposing conditions of visual degradation. However, the extreme potential of a) scientific exploration both in Earth caves but also in off-worlds such as in Martian and Lunar lava tubes, alongside b) the benefits of possible robotic subterranean search and rescue call for the necessary contributions in the domain.

However, autonomous exploration of cave rooms and networks is a daunting adventure. In such missions, robots can be asked to enter complex underground environments without any prior information, localize in these degraded and sensor-deprived subterranean settings, search and map them in detail, detect objects and broadly entities of interest, while communications can be nearly impossible to establish and thus resilient autonomy is required. Responding to these challenges, this work presents a complete system for autonomous aerial robotic exploration of caverns and cave networks including a) the robotic platform depicted in Figure 1, b) its onboard perception and computing solution, c) the localization and mapping algorithm that facilitates autonomy, and d) the path planning module that allows for efficient exploration given no prior map of the environment. The system relies on the multi-modal fusion of diverse sensor data including LiDAR, visible-light and thermal camera vision, alongside inertial sensing. Furthermore, its autonomy-enabling planning module reflects key observations with respect to the topologies encountered in cave networks and ensures both efficient local behaviors and ability to negotiate large-scale maps.

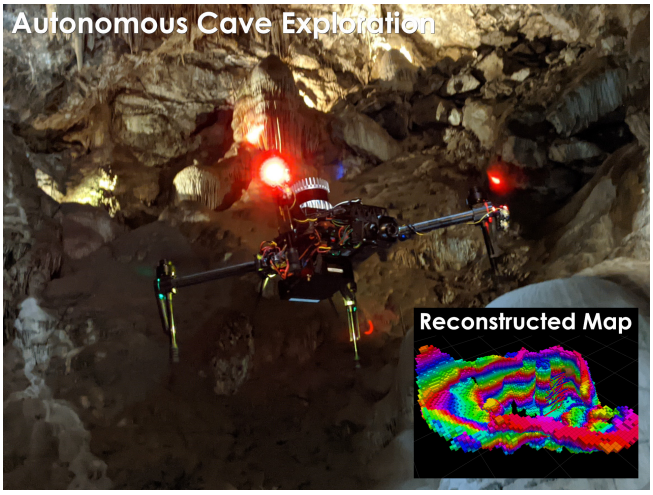

Fig. 1. Instance of autonomous exploration inside the Moaning Caverns. The presented aerial robotic scout is deployed to autonomously explore this large cavern room given no prior map. It localizes and maps the environment and builds a volumetric representation which in turn is exploited for the purposes of path autonomous path planning. A video recording from this result can be found at `https://youtu.be/0U0-PxMuGdk`

To evaluate the overall system solution, a set of studies were conducted both in simulation and experimentally. In terms of simulations, the path planning module is tested in particular with a model of the aerial robotic system tested in two sets of complex cave multi-branching and maze-like networks and over multiple times to derive statistical insights. Subsequently, a field experiment study is presented involving the autonomous exploration of a large room of the Moaning Caverns subterranean cave network.

The remainder of this paper is organized as follows. Section II outlines related work, followed by the system description in Section III. The autonomy-enabling modules for localization and mapping, as well as path planning are described in Section IV. Evaluation studies are detailed in Section V, followed by conclusions drawn in Section VI.

This material is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) under Agreement No. HR00111820045. The presented content and ideas are solely those of the authors.

[1] The authors are with the University of Nevada, Reno, 1664 N. Virginia, 89557, Reno, NV, USA, [2] The authors are with the Norwegian University of Science and Technology, O. S. Bragstads Plass 2D, 7034, Trondheim, Norway `mdharmadhikari@nevada.unr.edu`

II. RELATED WORK

A set of research works has considered the problem of autonomous robotic exploration of cave environments [1, 2, 4] with a significant subset of those focusing on extra-terrestrial cave configurations such as Lunar and Martian lava tubes [3, 5, 6]. In relation to the latter, NASA and other organizations have a set of activities to approach the potential of off-world robotic lava tube exploration [7, 8]. From a broader perspective, the research community has recently focused on the problem of autonomous subterranean exploration. With the kick-off of the DARPA Subterranean Challenge in September 2018, the domain of underground robotic autonomy has seen rapid growth. The contributions in [9, 10] present strategies for robust multi-sensor subterranean localization and mapping. The works in [11–14] present novel legged systems - including quadrupeds and hexapods - deployed in subterranean settings. Considering the potential of multi-robot collaboration, the authors in [15–18] present strategies for multi-agent exploration, map sharing and networked communications in the underground domain. These recent contributions build on years of research in the niche space of subterranean autonomy with a set of publications addressing challenges in exploration planning [19–22], localization [23], specialized robotic system development [24–27], navigation systems [28, 29] and more.

III. SYSTEM OVERVIEW

Autonomous exploration of cave environments presents unique features involving areas with both very wide geometry and also settings with narrow passages, while the underlying terrain and perceptual conditions can be extremely challenging. Aerial robots offer unique advantages to negotiate and autonomously explore such settings due to their elevated perspective and the fact that they are not bound to any terrain limitations. Despite their limited endurance, aerial robots can use their speed and flexible navigation to cover more area in short times. Furthermore, for certain large cave rooms or complex vertical geometries, they may correspond to the only option.

System Design: The developed robotic explorer for cave environments build on the DJI Matrice 100 quadrotor airframe. The system integrates an INTEL NUC Core-i7 Single Board Computer (NUC7i7BNH) onboard to enable the computing power towards localization and mapping, and path planning. The robot specifically integrates an OUSTER OS-1 LiDAR, a FLIR Blackfly S visual camera, a FLIR Tau2 thermal camera, and a VectorNav VN-100 Inertial Measurement Unit (IMU). The combined fusion of these sensing modalities enables the robust localization and mapping in the challenging, dark cave environments and facilitates autonomous operation. To better support the need for high quality visual observations in low-light environments, the system integrates four high-power Cree LEDs, which are synchronized with the visible light camera's shutter producing a power-efficient flashing behavior. Finally, the robot communicates with its base station primarily (but not only) through its onboard 5.8GHz WiFi radio. As reliable high-bandwidth communications are often not possible to be established underground, the system operates in fully autonomous mode, attempting to transmit its data to the base station in real time but not relying on it for any human input. If communications are denied during sections of the mission, the derived data are transmitted once a communication link is re-established (e.g., when the robot returns to the take-off location). The developed aerial robotic scout for cave exploration, called "Charlie", is depicted in Figure 2.

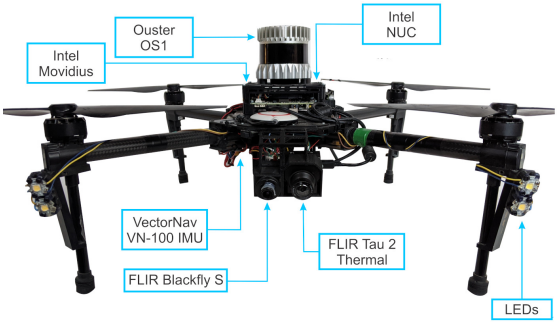

Fig. 2. The developed Charlie Aerial Scout.

Software Architecture: The software architecture and key functionalities onboard the Charlie aerial robot are presented in Figure 3. The system uses the exploration path planning algorithm briefly summarized in Section IV, alongside a multi-modal localization and mapping solution for GPS-denied and perceptually-degraded operation. The planned paths are tracked given the estimates of onboard odometry through a linear model predictive controller [30] and the calculated low-level commands are tracked by the autopilot electronics of the system.

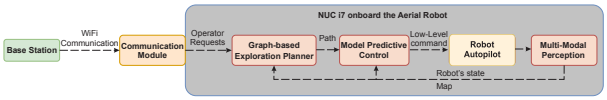

Fig. 3. Schematic system overview of the core functional modules of the Charlie aerial robot. All the components inside the gray shaded area are executed onboard the robot. Localization and mapping, planning and position control run on its INTEL NUC board.

Additional Sensing Payload: The presented robotic system fuses a set of sensor streams to facilitate its localization and mapping capabilities. As described above, this includes the integration of an OUSTER OS1 3D LiDAR with 64 Channels, a FLIR Blackfly visual camera, a FLIR Tau2 thermal camera and its IMU. At the same time, the robot integrates an additional sensing payload that may reflect further mission needs. For example, the robot integrates a Sensirion SCD30 CO_2 and temperature sensor and has a remaining capacity of 100g for a carefully selected additional sensing payload (e.g., a multispectral camera to characterize rock formations).

Subterranean Communications: To facilitate wireless communications underground - to the extent that this is functionally possible - the system integrates two types of antennas. First, it integrates WiFi at 5.8GHz which is communicating to the ground station either to share maps or receive high-level guidance commands. For that purpose, the ground station integrates a $4.9 - 5.8$GHz 19dBi Dual Polarized Flat Panel Antenna high-gain panel. In addition, the system integrates LoRa communications based on the Semtech SX1262 radio module. The latter is used as an alternative channel for telemetry feedback and crucial commands (e.g., return-to-home or immediate landing). The two channels operate in a complementary fashion and generally it is not observed that the LoRa module has persistently longer range than the utilized WiFi solution. However, there are environment configurations that one performs better than the other thus their complementary integration is adopted.

IV. EXPLORATION AUTONOMY

This section details the core path planning and perception solutions enabling autonomous exploration of cave environments onboard the presented aerial robotic scout.

A. Exploration Path Planning

Cave environments often involve both very wide rooms and networks of narrow passages with multiple branching points. In fact, the geometrical diversity of cave environments around the world is vast. For example, the Hang Son Doong cave - located in Central Vietnam - has a total volume of $3.84 \times 10^7 \text{m}^3$ and is considered the largest cave in the world (in terms of volume). The Mammoth Cave in Kentucky, US involves extremely wide cave rooms (e.g., the Rotunda Room) combined with a very wide and complex network of cave passages located under Flint Ridge. The Mammoth-Flint Ridge Cave System involves the longest known cave system in the world with more than 640km of surveyed passageways. These are just indicative examples as Earth involves a massive amount of different cave environments that pose unique complexities and diverse challenges.

Reflecting the diversity of cave environments, the work in [31] studied the topologies and morphologies of Karst systems. Karst is a topography formed from the dissolution of soluble rocks (e.g., limestone dolomite, and gypsum) and is characterized by underground drainage systems with sinkholes and caves. As a geologic environment, a Karst contains interconnected conduits which tend to develop so as to allow fluid flow. At the same time, the permeability structure evolved as a consequence of carbonate dissolution. Studying cave patterns and configurations based on geometric (vertical index, turtuosity, curvature, width-height ratio) and topological (ramification index, meshedness, network density) the authors of [31] propose a classification based on four types of patterns, namely a) branchwork cave, b) elongated branchwork, c) anastomatic maze, and d) angular maze. Furthermore they propose the morphological organization based on four categories, namely a) vadose branchwork (VB), b) water-table cave (WTC), c) looping cave (LC), and d) angular maze (AM). The proposed patterns are presented in Figure 4.

As shown in this analysis, cave environments can present a diverse level of complexity. Since both tree-like and more interconnected graph topologies are possible, we are motivated to utilize graphs as the data structure to facilitate efficient and resourceful path planning inside cave configurations. This is best understood by observing Figure 4 which presents 2D cave patterns as defined by the authors in [31] and via the associated Figure 5 presenting the application of vertex coloring in two such 2D cave patterns illustrating that well-established graph operations can help to assess the complexity of such topological configurations.

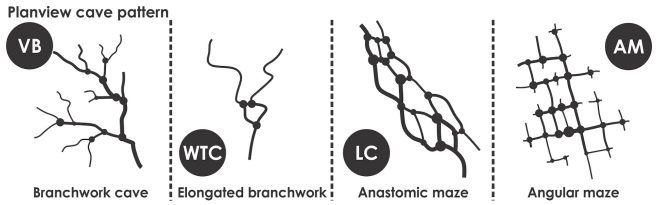

Fig. 4. 2D drawings of idealized cross-sections for cave karst systems presenting the planview of the associated patterns. Caves can present diverse topologies as discussed in [31]. It is highlighted that these visualizations do not correspond to any actual cave data but only indicative drawings graph topologies inspired by data presented in [31].

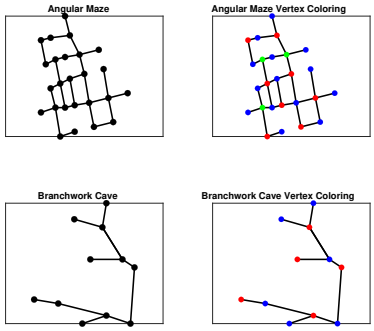

Fig. 5. Application of vertex graph coloring in simplified 2D planview cave patterns. The chromatic number of a graph or graph subset allows to evaluate the complexity of the multi-branching involved. More complex horizontal or vertical multi-way intersections increase the chromatic number. Maze-like caves involve components imposing the need to raise the chromatic number (3) whereas the tree-like structure of a branchwork cave keeps the chromatic number low (2). More complex and 3D cave topologies may present a chromatic number equal to 4. Note that these visualizations do not correspond to any actual cave data but only indicative of graph topologies.

In response to these facts, the developed Graph-Based exploration path Planner (GBPlanner) - presented and detailed in [14, 32] - corresponds to a methodology tailored to subterranean settings and especially caves which allows for efficient volumetric exploration despite the large-scale and geometrically-diverse character of caves, while at the same time offering a) safe return-to-home functionality, and b) solution resourcefulness when the exploration process reaches a dead-end (e.g. a mine heading). The exploration planning problem is defined as follows:

Problem 1 (Exploration Problem): Given an occupancy map $\mathbb{M}$, find a collision-free path $\sigma^* = \{\xi_i\}$ ($\xi_i = [x_i, y_i, z_i, \psi_i]^T$ being the flat rigid body configuration) to guide the robot towards unmapped areas and ensure the exploration of the perceivable volume V_E within the total and initially unknown volume V. Under the assumption of a depth sensor $\mathbb{S}$ with maximum effective range $d_{\max}$ and perception that stops at surfaces, the perceivable volume is defined as $V_E = V \setminus V_{res}$, where V_{res} is the residual volume for which no robot configuration exists to map it.

The implemented method, GBPlanner, is based on the architecture outlined in Figure 6. As cave environments can be particularly large and at the same time involve - at sections - narrow geometries, the method is organized based on a bifurcation between a local and a global stage. The first is responsible for efficient exploration in the region where the robot currently is, while the second stage handles the tasks of auto-homing and re-positioning towards frontiers of the exploration space when the local exploration is exhausted.

The cave environments involve various vertical passages and incline slopes. This requires dedicated perception sensors and modifications in the planning strategy. In order to tackle this, we utilise a vertical exploration mode within the local planner that biases the sampler to sample more densely along the vertical axis. This assumes that a depth sensor unveiling the space above and below the robot is available. Different sensor configurations involving a combination of depth cameras and LiDARs were tested in simulation. Usage of a 3D LiDAR with large vertical Field of View (FOV) (e.g., 90deg) was found to greatly support vertical passages exploration, while more widely used LiDARs with more narrow vertial FOV (e.g., 30deg) perform well in long corridors but may lead to slower vertical mapping.

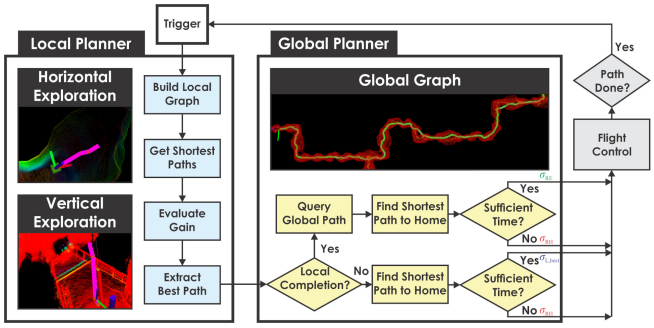

Fig. 6. Key steps of the proposed planner. The local planner operates within a window around the current robot location, samples a graph and identifies the path that maximizes the exploration gain. The global planner offers two functionalities, namely a) re-position to the frontier of the exploration space when a "dead-end" is reached, and b) return-to-home. For both tasks, the global planner utilizes a graph incrementally built during the robot operation thus saving the time that it would take to sample a new graph from scratch.

B. Local Planner

The local planning stage of GBPlanner runs within a local map $\mathbb{M}_L \subset \mathbb{M}$ that is centered around the current robot configuration ξ_0. The selection of $\mathbb{M}_L$ is such that it is both sufficiently large to allow to progress in the exploration task and also small enough to ensure computational efficiency. This local method first samples an undirected random graph $\mathbb{G}_L$ starting from ξ_0. Using Dijkstra's algorithm, it then identifies the shortest paths from the current robot location to all the sampled vertices $\{\sigma_i\}$. Each of these paths σ_i, $i = 1...n$ is then evaluated with respect to the exploration gain $\mathcal{G}$ it accrues based on the score:

$$\mathcal{G}(\sigma_i) = e^{-\gamma_{\mathcal{S}}\mathcal{S}(\sigma_i,\sigma_{exp})} \sum_{j=1}^{m_i} \mathbf{VolumeGain}(\nu_j^i) e^{-\gamma_{\mathcal{D}}\mathcal{D}(\nu_1^i,\nu_j^i)} \quad (1)$$

where $\mathbf{VolumeGain}$ represents the new volume to be explored at every vertex ν_j^i along this path, m_i is the number of vertices in σ_i, $\mathcal{S}(\sigma_i, \sigma_{exp})$, $\mathcal{D}(\nu_1^i, \nu_j^i)$ are weight functions with tunable factors $\gamma_{\mathcal{S}}, \gamma_{\mathcal{D}} > 0$. $\mathcal{D}(\nu_1^i, \nu_j^i)$ is the cumulative Euclidean distance from a vertex ν_j^i to the root ν_1^i along σ_i. Furthermore, $\mathcal{S}(\sigma_i, \sigma_{exp})$ is the similarity distance metric between σ_i and a pseudo-straight path σ_{exp} of identical length along the current exploring direction, that penalizes the paths having larger deviation from the current exploring direction that appear locally beneficial but may hinder the long-term exploration rate. This is essential when the system has to negotiate branchwork- and maze-caves with multiple branching points. Given the dominant exploring direction ϕ_{exp}, the pseudo-straight path σ_{exp} is generated along ϕ_{exp} having length equal to that of σ_i. Subsequently, using the Dynamic Time Warping method [33] we derive the similarity distance metric. Finally, the path $\sigma_{L,best}$ maximizing the exploration gain is selected, improved for safety by modifying vertices to be further away from obstacles and provided as a reference to the robot. The whole process is repeated in an iterative fashion. As the local planner finds solutions only within the local map subset $\mathbb{M}_L$ it is possible to be unable to derive a solution with sufficient gain, while such a solution can in principle be available in the overall map. Thus, when the local planner reports that all its solutions have a gain below a threshold $g_\epsilon > 0$, then the global planner is engaged.

C. Global Planner

When the robot enters a new cave room of significant size, it will be the local planning stage that runs several iterations until most of it is explored and mapped. However, when this is achieved or if the robot battery approaches its limits, the the global planning stage undertakes two important responsibilities. The first is to deliver autonomous homing, and the second is the detection of frontiers of the exploration space and guide the robot towards them when the battery is sufficient and the local exploration potential is exhausted. To deliver this functionality, the method incrementally builds a global graph $\mathbb{G}_G$. At every iteration, this graph is used to derive a return-to-home path σ_{RH} after a re-sampling step has taken place to locally enrich the graph. Furthermore, the same graph is used to detect vertices of it that provide maximum anticipated exploration gain and thus can stand as frontiers of the explored space towards directions that are not exploited so far by the local planner. Given these two capabilities, the method either uses $\mathbb{G}_G$ to derive σ_{RH} as per the

battery limitations, or to derive a path σ_{RE} that re-positions the robot towards the best possible frontier of the explored space such that the anticipation for future exploration is maximized and the system maintains battery to also return to its home location. Details on the frontier selection process can be found in [14]. When the system reaches the selected frontier, the local planner is re-triggered. The global planner is critical in caves due to their multi-branched geometry. In a Branchwork cave, the robot will explore using the local planner till it reaches a dead end. The global planner will then guide the robot near the closest branching point and the local planner will be re-engaged until another dead end is reached. This will be repeated resulting in an efficient exploration of the entire cave. Maze-like caves are more complicated. The current frontier selection methodology may lead to less efficient behavior when multiple cycles are present in the graph. Hence, in future work we consider the addition of complexity metrics, in relation to the number of branching directions and other properties, possibly through vertex/edge coloring. This is highlighted in the presented results.

D. Multi-Modal Localization and Mapping

To facilitate autonomous operation inside the challenging underground cave environments, a resilient localization and mapping solution is required. Responding to this need, our team has developed a Complementary multi-modal Simultaneous Localization And Mapping solution as presented in [10]. The method - called CompSLAM - fuses data from LiDAR, visual and thermal cameras, alongside IMU cues as presented in Figure 7. As depicted, camera frames from both (or either) visible-light and thermal vision sensors are combined with IMU data and fused to derive a first high-rate (e.g., 20Hz) estimate of the robot's motion. The deployed multi-camera/IMU estimation solution is based on ROVIO, the Robust Visual-Inertial Odometry framework detailed in [34]. ROVIO was extended to allow for the utilization of the full radiometric information of LongWave InfraRed (LWIR) thermal cameras as detailed in [35, 36]. A key element of its functionality relates to the fact that the data association front-end over camera data considers the full 16bit data format of radiometric information coming from the thermal sensor, as opposed to the majority of literature on thermal vision-based odometry estimation where the radiometric data are rescaled to conventional 8bit images. In subterranean environments and especially in caves, thermal differences are often minimal thus exploitation of the full sensor information is essential.

This Visual-Thermal-Inertial Odometry system provides its high update rate estimate for the robot trajectory and tracked landmarks, alongside the associated covariance. As the covariance matrix represents the level of confidence to the mean estimate, we associate the health of this multi-camera/inertial system via the D-optimality metric [37] applied over the pose covariance. If this estimate is evaluated as healthy then it is subsequently provided as a prior to the LiDAR Odometry and LiDAR Mapping modules of the multi-modal CompSLAM architecture. The LiDAR

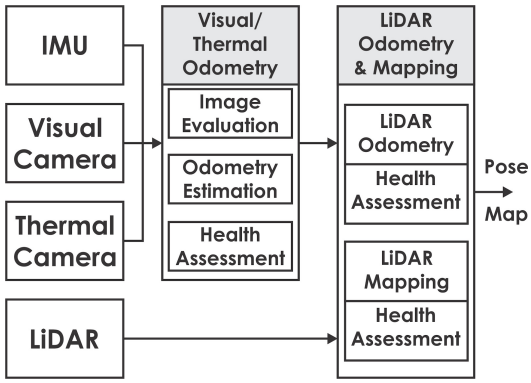

Fig. 7. Architectural overview of the complementary multi-modal localization and mapping solution running onboard the Charlie aerial robot.

Odometry module is used to identify the odometry estimate using the LiDAR data via scan-to-scan alignment. This is performed by minimizing point-to-line and point-to-plane distances between extracted line and plane features in a manner similar to the work in [38]. Importantly, we can use the eigenvalues of the underlying optimization matrix to detect possible ill-conditioning across certain dimensions. Ill-conditioning of the LiDAR scan-to-scan alignment is possible due to geometric self-similarities in the environment. When this condition is detected, the respective estimates are replaced by those derived by the visual/thermal-inertial estimation system as long as its health metric has indicated that those estimates are associated with low D-optimality values. The LiDAR Odometry step takes place at 10Hz. Beyond this step, CompSLAM also performs LiDAR Mapping - at a 2Hz rate - with the goal to reduce errors accumulated due to odometry estimation drift. In the LiDAR Mapping step, a scan-to-submap alignment takes place which is likely to be less prone to fitting errors. The health of this step is evaluated again via the eigenvalues of the involved optimization matrix. The overall architecture of CompSLAM provides resilient pose estimates and mapping in GPS-denied and degraded conditions, as long as either the visible/thermal-inertial or LiDAR-based modules remain informative at any given time. As the conditions of degradation for visual cameras, thermal vision and LiDAR are generally different to each other, the system maintains its ability for robust performance in a very large variety of environments. In terms of practical experience, in our field tests in cave environments so far, we have identified that LiDAR and visual camera data (given sufficient onboard illumination) are the two most commonly reliable exteroceptive data onboard the aerial robot. Thermal vision is often challenged inside natural caves as the environment thermal gradients are often not significant. However in certain cases, for example when water is involved or when openings to the ground are available, thermal vision provides particularly valuable information that is not affected by other conditions of degredation (e.g., dust). At the same time, as the sensitivity of low-cost thermal cameras increases so will do their utility and usefulness in cave environments.

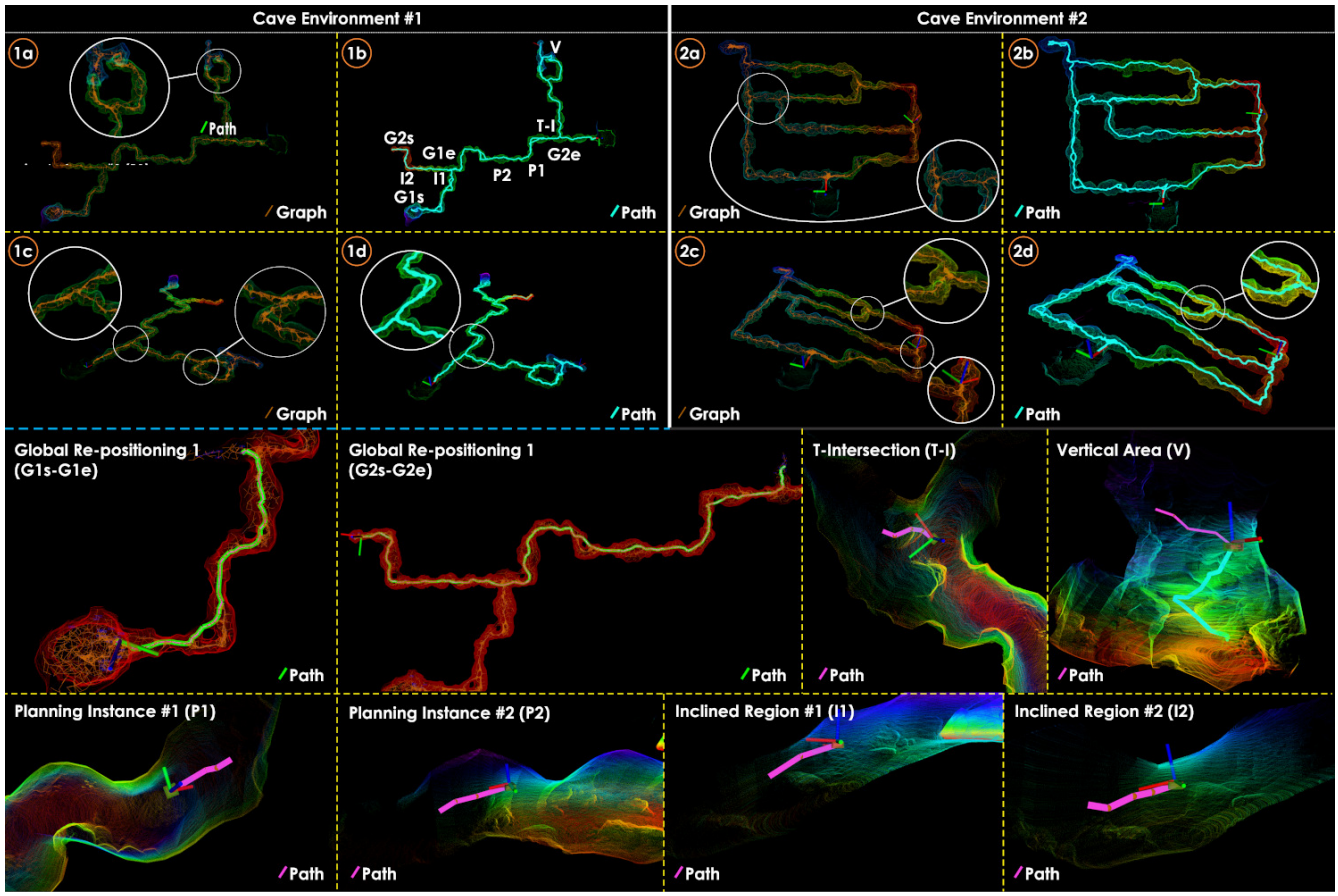

Fig. 8. Simulation missions inside two cave environments. Cave Environment #1 presents a branchwork cave with incline slopes towards the right and vertical passage towards the top as shown in 1b. Cave Environment #2 is a maze like cave with several incline slopes. Sub-figures 1a, 1c and 2a, 2c show the explored environment and the global graph whereas sub-figures 1b, 1d and 2b, 2d show the robot's path in Cave Environment #1 and #2 respectively. The bottom half of the figure shows few local and global planning instances from one of the missions in Cave Environment #1. The presented instances demonstrate the planner's capability to plan in vertical passages, inclined slopes, intersections.

V. EVALUATION STUDIES

A set of evaluation studies, involving both simulations and field experiments, were performed to assess the capacity of the proposed robotic cave exploration.

A. Simulation Studies

Simulation studies were conducted in two cave environments to test the Exploration Path Planner. The simulated environments were hundreds of meters in scale and involved large open spaces, vertical passages, inclined slopes and complex networks. The RotorS simulator [39] was used for the simulations with a quadrotor MAV model integrating a 3D LiDAR with FOV $[F_H, F_V] = [360, 90]^\circ$ (analogous to an OUSTER OS0 sensor) and maximum range $d_{\max} = 50$m. The large vertical FOV F_V (now available in off-the-shelf sensors) is particularly important for efficient mapping and path planning inside the cave environments having varying height conditions and vertical passages. The bounds of the local map $\mathbb{M}_L$ for the local planner were set to *length*×*width*×*height* = $30 \times 30 \times 8$m. Five fixed-time missions were conducted in each cave environment. The robot was started from the same location in every mission. The average speed of the robot for was set to 2m/s. Initially, the local planner was triggered and when local exploration was exhausted the robot was re-positioned to an exploration frontier based on the global planning functionality. The first cave environment is a Branchwork cave whereas the second is a maze. The second environment due to its topology requires longer exploration time for the same amount of volume and can result in sub-optimal global behavior. Figure 8 shows the explored map in both the caves with the path followed by the robot and the global graph spanned by the Global Planner. The bottom half of the figure shows indicative instances of local and global planning steps in the first cave environment. The performance of the planner in terms of explored volume over time can be seen in Figure 9. The constant volume parts in the plot correspond to the global re-positioning instances.

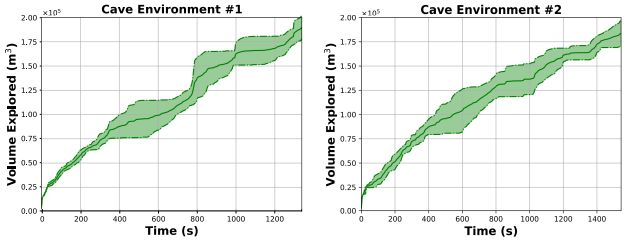

Fig. 9. Exploration rate (mean value and bounds over 5 executions) for each of the two cave simulation environments.

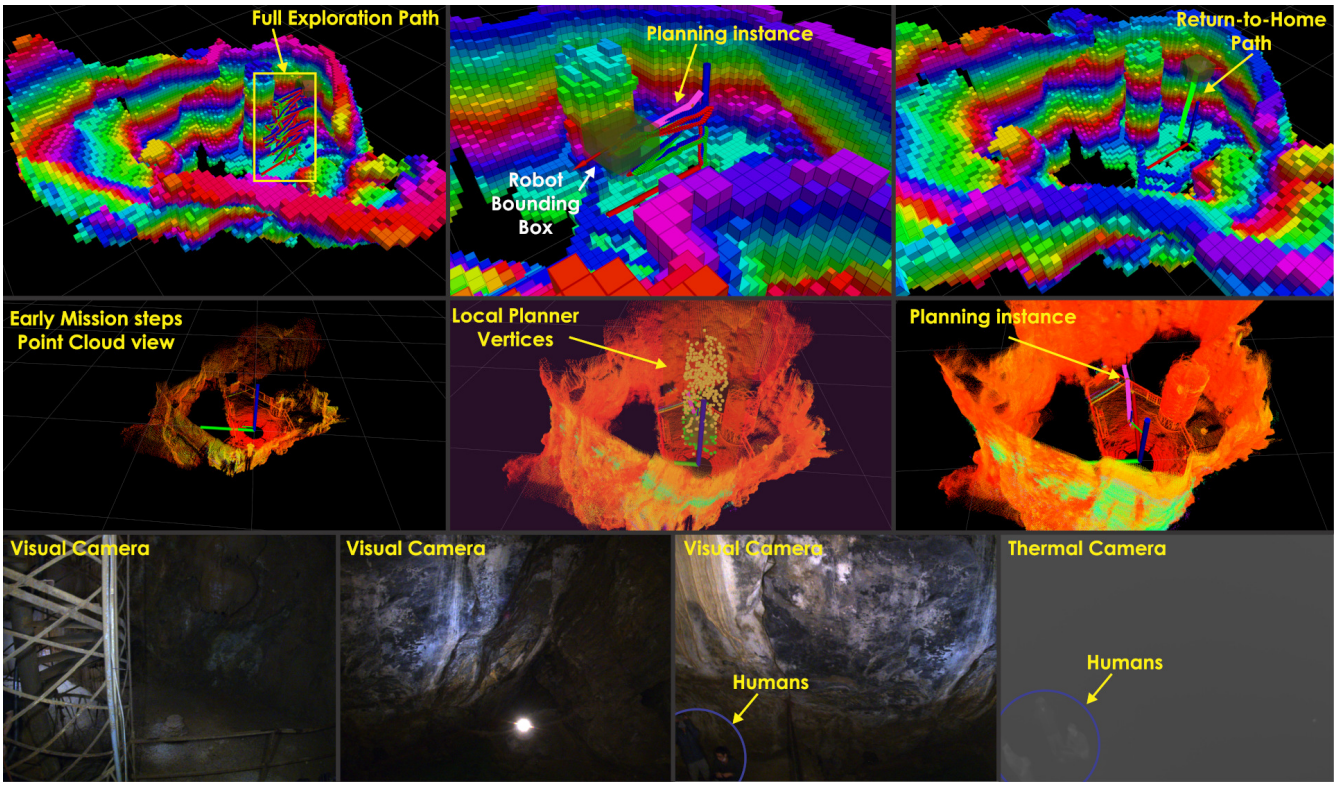

Fig. 10. Autonomous exploration mission inside a room of the Moaning Caverns national park complex subterranean cave settings. The first row presents volumetric data as used in planning, alongside the full robot path, an instance of planning (pink path) and the return-to-home path (green). The second row presents point cloud data as the map gets built incrementally, alongside an instance of the local planner vertices (green for admissible vertices and yellow for vertices outside of the mapped area), as well as an instance of planning path. Finally, the last row presents visual and thermal camera data from the two vision sensors onboard the robot. A video recording of the experiment is available at `https://youtu.be/0U0-PxMuGdk`

B. Field Experiment in the Moaning Caverns

The presented aerial robotic cave exploration system was fielded in the Moaning Caverns in California, US. In particular, the robot was tested inside a large room of the underground voids of the Moaning Caverns with the goal to reconstruct a complete 3D map and simultaneously acquire visual and thermal camera observations. The Charlie Aerial Scout as shown in Figure 2 was used in this mission. The bounds of the local map $\mathbb{M}_L$ for local planner were set to $length \times width \times height = 7 \times 4 \times 6$m. Figure 10 presents detailed instances from this field deployment and highlights key steps of the path planning and mapping process. During most of the mission, the system employs paths that allow it to incrementally explore this tall cave room based on its OUSTER OS1 LiDAR sensor onboard. As shown, the system is able to autonomously explore the cave environment, reconstruct its map and collect associated informative visual and thermal camera observations. Throughout this whole mission (after take-off) there was no human teleoperation or other interaction with the robot. This highlights the autonomy capabilities of the presented robotic system.

VI. CONCLUSIONS

In this work, the complete design of an aerial robot capable of autonomous exploration of cave networks is presented. The system relies on two core functionalities for multi-modal localization and mapping, as well as path planning for enabling resilient autonomy in such diverse and challenging environments. In particular, the system fuses LiDAR, visible-light and thermal cameras and inertial sensing for robust and resourceful perception in the GPS-denied, dark, often obscurants-filled and geoemetrically complex cave settings. Building on top of its perception capabilities, the system implements a graph-based exploration path planning algorithm tailored to the topological patterns observed in caves. To evaluate the proposed solution, a set of simulation studies and field experiments were conducted and presented in detail.

REFERENCES

[1] W. Tabib, K. Goel, J. Yao, C. Boirum, and N. Michael, "Autonomous cave surveying with an aerial robot," *arXiv preprint arXiv:2003.13883*, 2020.

[2] G. Zhang, B. Shang, Y. Chen, and H. Moyes, "Smartcavedrone: 3d cave mapping using uavs as robotic co-archaeologists," in *2017 International Conference on Unmanned Aircraft Systems (ICUAS)*. IEEE, 2017, pp. 1052–1057.

[3] A. Husain, H. Jones, B. Kannan, U. Wong, T. Pimentel, S. Tang, S. Daftry, S. Huber, and W. L. Whittaker, "Mapping planetary caves with an autonomous, heterogeneous robot team," in *2013 IEEE Aerospace Conference*. IEEE, 2013, pp. 1–13.

[4] K. Otsu, S. Tepsuporn, R. Thakker, T. S. Vaquero, J. A. Edlund, W. Walsh, G. Miles, T. Heywood, M. T. Wolf, and A.-A. Agha-Mohammadi, "Supervised autonomy for communication-degraded subterranean exploration by a robot team," in *2020 IEEE Aerospace Conference*. IEEE, 2020, pp. 1–9.

[5] T. Vaquero, M. Troesch, and S. Chien, "An approach for autonomous multi-rover collaboration for mars cave exploration: Preliminary results," in *International Symposium on Artificial Intelligence, Robotics, and Automation in Space (i-SAIRAS 2018). Also appears at the ICAPS PlanRob*, 2018.

[6] F. Klaesson, P. Nilsson, T. Vaquero, S. Tepsuporn, A. D. Ames, and R. M. Murray, "Planning and optimization for multi-robot planetary cave exploration under intermittent connectivity constraints," 2020.

[7] T. N. Titus, C. Phillips-Lander, P. Boston, J. Wynne, and L. Kerber, "Planetary cave exploration progresses," *Eos, Earth and Space Science News*, vol. 101, 2020.

[8] S.-K. Kim, A. Bouman, G. Salhotra, D. D. Fan, K. Otsu, J. Burdick, and A.-a. Agha-mohammadi, "Plgrim: Hierarchical value learning for large-scale exploration in unknown environments," *arXiv preprint arXiv:2102.05633*, 2021.

[9] M. Palieri, B. Morrell, A. Thakur, K. Ebadi, J. Nash, A. Chatterjee, C. Kanellakis, L. Carlone, C. Guaragnella, and A.-a. Agha-mohammadi, "Locus: A multi-sensor lidar-centric solution for high-precision odometry and 3d mapping in real-time," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 421–428, 2020.

[10] S. Khattak, H. Nguyen, F. Mascarich, T. Dang, and K. Alexis, "Complementary multi–modal sensor fusion for resilient robot pose estimation in subterranean environments," in *2020 International Conference on Unmanned Aircraft Systems (ICUAS)*, 2020, pp. 1024–1029.

[11] A. Bouman, M. F. Ginting, N. Alatur, M. Palieri, D. D. Fan, T. Touma, T. Pailevanian, S.-K. Kim, K. Otsu, J. Burdick, and A.-a. Agha-mohammadi, "Autonomous Spot: Long-Range Autonomous Exploration of Extreme Environments with Legged Locomotion," in *IEEE International Conference on Intelligent Robots and Systems*, 2020.

[12] I. D. Miller, F. Cladera, A. Cowley, S. S. Shivakumar, E. S. Lee, L. Jarin-Lipschitz, A. Bhat, N. Rodrigues, A. Zhou, A. Cohen *et al.*, "Mine tunnel exploration using multiple quadrupedal robots," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2840–2847, 2020.

[13] R. Steindl, T. Molnar, F. Talbot, N. Hudson, B. Tam, S. Murrell, and N. Kottege, "Bruce–design and development of a dynamic hexapod robot," *arXiv preprint arXiv:2011.00523*, 2020.

[14] T. Dang, M. Tranzatto, S. Khattak, F. Mascarich, K. Alexis, and M. Hutter, "Graph-based subterranean exploration path planning using aerial and legged robots," *Journal of Field Robotics*, 2020.

[15] T. Rouček, M. Pecka, P. Čížek, T. Petříček, J. Bayer, V. Šalanskỳ, D. Heřt, M. Petrlík, T. Báča, V. Spurnỳ *et al.*, "Darpa subterranean challenge: Multi-robotic exploration of underground environments," in *International Conference on Modelling and Simulation for Autonomous Systesm*. Springer, 2019, pp. 274–290.

[16] M. Zoula, M. Prágr, and J. Faigl, "On building communication maps in subterranean environments."

[17] K. Goel, W. Tabib, and N. Michael, "Rapid and high-fidelity subsurface exploration with multiple aerial robots," 2020.

[18] M. Corah, C. O'Meadhra, K. Goel, and N. Michael, "Communication-efficient planning and mapping for multi-robot exploration in large environments," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1715–1721, 2019.

[19] D. Silver, D. Ferguson, A. Morris, and S. Thayer, "Topological exploration of subterranean environments," *Journal of Field Robotics*, vol. 23, no. 6-7, pp. 395–415, 2006.

[20] C. Baker, A. Morris, D. Ferguson, S. Thayer, C. Whittaker, Z. Omohundro, C. Reverte, W. Whittaker, D. Hahnel, and S. Thrun, "A campaign in autonomous mine mapping," in *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004*, vol. 2. IEEE, 2004.

[21] J. G. R. III, R. E. Sherrill, A. Schang, S. L. Meadows, E. P. Cox, B. Byrne, D. G. Baran, J. W. C. III, and K. M. Brink, "Distributed subterranean exploration and mapping with teams of UAVs," in *Ground/Air Multisensor Interoperability, Integration, and Networking for Persistent ISR VIII*, T. Pham and M. A. Kolodny, Eds., vol. 10190, International Society for Optics and Photonics. SPIE, 2017, pp. 285 – 301. [Online]. Available: https://doi.org/10.1117/12.2266316

[22] S. S. Mansouri, M. Castaño, C. Kanellakis, and G. Nikolakopoulos, "Autonomous mav navigation in underground mines using darkness contours detection," in *International Conference on Computer Vision Systems*. Springer, 2019, pp. 164–174.

[23] A. Jacobson, F. Zeng, D. Smith, N. Boswell, T. Peynot, and M. Milford, "What localizes beneath: A metric multisensor localization and mapping system for autonomous underground mining vehicles," *Journal of Field Robotics*, 2020.

[24] A. Morris, D. Ferguson, Z. Omohundro, D. Bradley, D. Silver, C. Baker, S. Thayer, C. Whittaker, and W. Whittaker, "Recent developments in subterranean robotics," *Journal of Field Robotics*, vol. 23, no. 1, pp. 35–57, 2006.

[25] P. Novák, J. Babjak, T. Kot, P. Olivka, and W. Moczulski, "Exploration mobile robot for coal mines," in *International Workshop on Modelling and Simulation for Autonomous Systems*. Springer, 2015, pp. 209–215.

[26] A. Maity, S. Majumder, and D. N. Ray, "Amphibian subterranean robot for mine exploration," in *2013 International Conference on Robotics, Biomimetics, Intelligent Computational Systems*, 2013, pp. 242–246.

[27] D. Tardioli, L. Riazuelo, D. Sicignano, C. Rizzo, F. Lera, J. L. Villarroel, and L. Montano, "Ground robotics in tunnels: Keys and lessons learned after 10 years of research and experiments," *Journal of Field Robotics*, vol. 36, no. 6, pp. 1074–1101, 2019.

[28] J. N. Bakambu and V. Polotski, "Autonomous system for navigation and surveying in underground mines," *Journal of Field Robotics*, vol. 24, no. 10, pp. 829–847, 2007.

[29] R. Lösch, S. Grehl, M. Donner, C. Buhl, and B. Jung, "Design of an autonomous robot for mapping, navigation, and manipulation in underground mines," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 1407–1412.

[30] M. Kamel, T. Stastny, K. Alexis, and R. Siegwart, "Model predictive control for trajectory tracking of unmanned aerial vehicles using robot operating system," in *Robot operating system (ROS)*. Springer, 2017, pp. 3–39.

[31] J. Jouves, S. Viseur, B. Arfib, C. Baudement, H. Camus, P. Collon, and Y. Guglielmi, "Speleogenesis, geometry, and topology of caves: A quantitative study of 3d karst conduits," *Geomorphology*, vol. 298, pp. 86–106, 2017.

[32] T. Dang, F. Mascarich, S. Khattak, C. Papachristos, and K. Alexis, "Graph-based path planning for autonomous robotic exploration in subterranean environments," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 3105–3112.

[33] A. G. Bachrach, "Trajectory bundle estimation for perception-driven planning," Ph.D. dissertation, Massachusetts Institute of Technology, 2013.

[34] M. Bloesch, M. Burri, S. Omari, M. Hutter, and R. Siegwart, "Iterated extended kalman filter based visual-inertial odometry using direct photometric feedback," *The International Journal of Robotics Research*, vol. 36, no. 10, pp. 1053–1072, 2017.

[35] S. Khattak, F. Mascarich, T. Dang, C. Papachristos, and K. Alexis, "Robust thermal-inertial localization for aerial robots: A case for direct methods," in *2019 International Conference on Unmanned Aircraft Systems (ICUAS)*, 2019, pp. 1061–1068.

[36] S. Khattak, C. Papachristos, and K. Alexis, "Keyframe-based thermal–inertial odometry," *Journal of Field Robotics*, vol. 37, no. 4, pp. 552–579, 2020.

[37] H. Carrillo, I. Reid, and J. A. Castellanos, "On the comparison of uncertainty criteria for active slam," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*. IEEE, 2012, pp. 2080–2087.

[38] J. Zhang and S. Singh, "LOAM: Lidar odometry and mapping in real-time," in *Robotics: Science and Systems X*. Robotics: Science and Systems Foundation, jul 2014. [Online]. Available: https://doi.org/10.15607%2Frss.2014.x.007

[39] F. Furrer, M. Burri, M. Achtelik, and R. Siegwart, "Rotors-a modular gazebo mav simulator framework," in *Robot Operating System (ROS)*. Springer, 2016, pp. 595–625.