

Extinguishing Real Fires by Fully Autonomous Multirotor UAVs in the MBZIRC 2020 Competition

Viktor Walter¹, Vojtěch Spurný¹, Matěj Petrлік¹, Tomáš Báča¹, David Žaitlík¹, Lyubomyr Demkiv² and Martin Saska¹

¹Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague, <http://mrs.felk.cvut.cz>

²Robotics Group of SoftServe Inc. and with the Institute of Computer Science and Information Technologies, Lviv Polytechnic National University, Ukraine

Abstract: In this paper, we describe a system for combating real fires with sprayed liquid extinguishing agent using a team of multirotor UAVs. The system design relies on onboard sensors and operates without the need for human intervention throughout its complex mission, from its takeoff to landing. The core UAV platform can estimate its state, self-localize, navigate and locate and extinguishing fires. Thermal and RGB cameras are used, each with a specialized computer vision subsystem and are combined with planar LIDAR for 3D localization of fires on multistory building facades. The system conducts aerial firefighting with a software stack that addresses flight dynamics and sensor limitations and a liquid-spraying subsystem incorporating a rigidly-attached water nozzle. The approach presented in this paper was motivated by the Mohamed Bin Zayed International Robotics Challenge (MBZIRC 2020) firefighting scenario, which focused on coordinated multi-UAV teams that can autonomously combat high-rise building fires. The MBZIRC series places particular emphasis on fast and reliable deployment of robots in realistic conditions. These contests promote development of real-world applications that are greatly needed by society, but which still exceed State-of-the-Art in the robotics community. To our knowledge, our system was the only MBZIRC 2020 contender to extinguish a facade fire successfully in autonomous mode without using an RTK-GNSS system. Our approach contributed to victory in the overall competition and we have now adapted it into an industrial prototype for a firefighting UAV. A video attachment to this paper is available at <http://mrs.felk.cvut.cz/fr2020firechallenge-facedefires>.

Keywords: emergency response, aerial robotics

1. Introduction

There is currently an increasing demand for Unmanned Aerial Vehicles (UAVs) in the disaster-relief and emergency-response sectors. In 2017 alone, worldwide fire casualties reached 120,632 people

Received: 1 October 2020; revised: 19 February 2021; accepted: 7 May 2021; published: 29 March 2022.

Correspondence: Viktor Walter, Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague, <http://mrs.felk.cvut.cz>, Email: viktor.walter@fel.cvut.cz

This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © 2022 Walter, Spurný, Petrлік, Báča, Žaitlík, Demkiv and Saska

(Ritchie, 2018) and death from fire or heat was the 4th leading cause of death by accidental injury in 2019, with over 110,000 victims (WHO, 2020). Additionally, there were 1,318,500 fires in the United States during 2018, causing 3,655 deaths and 15,200 injuries. The total losses from fires in 2018 were equal to \$25.6 billion US dollars (USFA, 2018; NFPA, 2019).

Robots present one possible solution to helping firefighters and decreasing the amount of fire-related casualties. The robots can aid firefighters in detecting and locating fires within buildings and extinguishing them, both indoors and outdoors (Safetymanagement, 2018). Such systems can also provide firefighters with the information needed to protect the lives of firefighters during incidents. For example, among the robots used in disaster response is

- THOR (Tactical Hazardous Operations Robot), a humanoid robot designed by Virginia Tech University and employed by the U.S. Navy's Shipboard Autonomous Firefighting Robot (SAFFiR) program. Its goal is to assist sailors in firefighting onboard ships. Unfortunately, this robot is slow and may be vulnerable to water damage (VT, 2018; Hopkins et al., 2015).
- Thermite Robot is a system created by Howe and Howe Technologies for the U.S. Army. It is a remote-controlled mobile robot equipped with a hose capable of delivering 1900 liters of water per minute (Firehouse, 2018).
- TAF 20 and TAF 35 (Turbine Aided Firefighting Machine) by Emicontrols (a subsidiary of the TechnoAlpin Group) operates well in small spaces, such as tunnels. The robot can move obstacles with its bulldozer blade and clear the environment of smoke and gasses by binding them within its turbine. It is also able to discharge 3,500 liters of water per minute (Emicontrols, 2018a; Emicontrols, 2018b).
- TC800-FF by Tecdrone is a remotely operated robot capable of working indoors and outdoors. It is equipped with multiple sensors to retrieve data from disaster-stricken areas and to then send it back to remote operators (Robotpompiers, 2018).
- Fire OX (Fire Product Search, 2018) by Lockheed Martin is capable of carrying a water tank, standing out among other disaster robots due to its capability of semi-autonomous operation.

These are merely samples of existing solutions. Other firefighting robotic systems are also in use (Lattimer, 2020; ScienceBusiness, 2020; Peskoe-Yang, 2019).

While these robots share a common need for remote operation, they are also unsuitable for combatting fire on tall buildings. Though they can ascend stairs, when speed is essential these systems are outperformed by UAVs. The robots used during the fire at Notre Dame Cathedral (ScienceBusiness, 2020; Peskoe-Yang, 2019) could not help firefighters abate the fire, but rather helped them fight with assessing the consequences and to save priceless artefacts. However, even though UAVs were used as well, they were only applied to monitor the scene and to select locations for deploying hoses by humans. Thermal cameras, which would have been highly practical for the scenario, were not used (The Verge, 2019).

Moreover, operating within environments with low visibility makes it harder for a remote operator to control the robot. Such limitations increase the importance of autonomous robots in navigating smoky or hazardous environments, since they are able to navigate using data from sensors of diverse modalities, such as rangefinders, thermal and visual cameras, etc.

Fires on high-rise building facades pose a particularly relevant challenge and opportunity for robots. An emergency evacuation in such a building can be difficult and may thus lead to injuries or the loss of human life and since such fires are out of reach for conventional, ground-based firefighting methods, an aerial solution is called for. The timing aspect of this problem is the decisive factor for the amount of damage the fire will cause. According to recent analyses, the incipient stages of fires on the cladding materials of modern high-rise buildings take several minutes, during which the size of the fire is relatively manageable. After this phase, the fire rapidly spreads beyond the control of firefighting efforts and considerable damages are then unavoidable (Chen et al., 2019b; Guillaume et al., 2018; McKenna et al., 2018). This observation is consistent with the timeline of the fire in the Grenfell Tower in June 2017 (London Fire Brigade, 2018). One strategic option would

maintain a team of UAVs with firefighting capabilities on-site or ready for delivery by the local first responders deployed by local fire departments. Such systems could be deployed significantly sooner than a manned helicopter mission, which might follow later.

In order to fully exploit the potential of UAVs in firefighting, especially in high-rise buildings, and to achieve reliable operation, the UAVs themselves need to be autonomous with capabilities going beyond merely following a static, predefined trajectory. The precise location of a fire source, especially in hard-to-reach locations such as high-rise buildings, is difficult to specify accurately from a takeoff location. The fire source may not even be directly visible from the operator's vantage point, apart from smoke plumes indicating an approximate region of interest.

On the other hand, the human factor in UAV-based firefighting can be indispensable. For example, with vision-based fire localization onboard a UAV, finding fire sources without at least an approximate prior estimate of the region of interest can be time-consuming. Human oversight can specify, for example, the range of floors or the building face where the fire is most likely located, based on external observation, sensors inside the building, and other information sources available to the human operators. This input can then be used by the UAV system to navigate quickly and autonomously to a location where onboard sensors can identify and locate the fire sources. While merely flying close to a point outside of a building in an otherwise open space is simple to execute, it is time-consuming if the distance is great. For this reason, manual control of the UAV in the approach phase would waste human time and attention that could better be applied to other critical aspects of the entire firefighting mission.

Autonomy is additionally indispensable in the procedure of extinguishing the fires themselves. Since direct mechanical contact with the environment by UAVs is hazardous, the fires may have to be doused with a liquid stream or fire-extinguishing projectiles. For such scenarios a fire-fighting UAV must be able to aim at the localized fire sources. The ideal position at which the extinguishing agent should be deployed at any given moment will change with time and with targets being affected by extinguishing. Controlling this operation manually is unfeasible in conditions typical of the exteriors of higher floors of high-rise buildings. In such locations, the UAV and the fire source can be too far away from ground-based operators for them to effectively take aim, even if the UAV system features a low-latency onboard video streaming system. Wind often present in such conditions would also make manually correcting aim highly challenging, as had been observed during the Mohamed Bin Zayed International Robotics Challenge (MBZIRC 2020) Challenge 3 where some teams employed teleoperated solutions without success.

Saving endangered people and property is an often-cited motivation for the development of new robotic systems. However, this factor is rarely as immediately applicable as it was in Challenge 3 of the MBZIRC 2020 competition which was directly inspired by the need to extinguish fires in difficult to access locations, such as the high-rise buildings discussed above. This challenge required the use of a team of firefighting robots to locate and extinguish fires, both within and without a building. Solving the challenge required multiple robotic tasks to be addressed beyond just hardware design, including navigation, searching for objects, computer vision with 3D localization, sensor fusion, and implementation of active extinguishing procedures.

This paper describes our approach to one specific subtask of the challenge: extinguishing real fires on the facade of a building with water deployed by a multirotor UAV. The approach we employed was based on using thermal cameras combined with a specialized thermal vision and estimation subsystem, as well as custom-designed searching and in-flight extinguishing maneuvers. To the best of our knowledge, we were the only team in this subtask to not only successfully use full autonomy during the flight, but to also rely solely on local sensors for localization of the targets and control of the UAVs.

The main contributions of this paper include the overall system design, as a model for realistic UAV-based firefighting, and the observations made during iterative development and deployment stages. Novel aspects of our system include the onboard, hybrid, vision-based localization of building-facade fires and fire-extinguishing capability, using dynamic body-pose-based aiming and water-jet control. Among the lessons learned in deployment, we emphasize the need for robust computer

vision, specifically robust with respect to the highly dynamic appearance of fires in RGB and thermal images. Other lessons include the possibility of exploiting the transparency of flames to blue light in order to observe surfaces behind the flames. There is also a need for robust aiming and localization suitable even for a maneuvering multirotor UAV. More specific technical aspects are also included, such as addressing the severe interference of USB 3.x devices with Global Navigation Satellite System (GNSS) equipment.

2. Related Work

Using UAVs in firefighting has been previously investigated in several works. One situation where UAVs can help is in monitoring of outdoor fires. As reported in (Merino et al., 2012), a system of multiple UAVs can be used for automatic forest fire monitoring using RGB and infrared cameras. Real experiments with forest fire monitoring in a national park have already been conducted by the Hungarian fire department (Restas, 2006). The authors of (Viguria et al., 2010) describe a task allocation strategy for the distributed cooperation of ground and aerial robot teams in fire detection and extinguishing. In (Chen et al., 2019a), a UAV system is designed to extinguish a fire by dropping a fire-extinguishing capsule on it. Similarly, in (Saikin et al., 2020) the authors employ a UAV system for delivering a liquid fire-extinguishing agent by a dynamic dropping maneuver, maximizing release velocity to reduce dispersion.

However, none of the systems above address certain aspects, such as environmental interaction, precision control for extinguishing, or flight close to buildings (much less flight inside them). In order to solve these tasks autonomously, the following challenges need to be addressed: Precise localization near obstacles with minimal reliance on GNSS, autonomous real-time detection of fires using onboard sensors, reliable stabilization of UAV during extinguishing, motion planning, and reactive obstacle avoidance.

Multirotor UAVs, also known as drones, are widely used in multiple areas of industry and the public sector, as is evident in surveys (Shakhatreh et al., 2019). For instance, the CB Insight report (CBInsight, 2020) lists 38 technology sectors where UAVs are highly applicable were listed, among which were education, space, defense, disaster relief, etc. An extensive survey of the application of drones for bridge inspection is performed in (Hubbard and Hubbard, 2020). While the methodology presented in the paper establishes a basic framework intended for bridge inspection, the approach can be tailored for inspections in other domains, such as the firefighting mission under consideration. For the inspection, a mission definition system must be set up. Such multirotor UAV-oriented architecture is presented in (Portas et al., 2018). Another mission guided framework was presented in (Hrabia et al., 2019). The investigation of automated inspection applications is conducted in (Tudevtagva et al., 2017) for high voltage transmission lines, and in (Ashour et al., 2016) for inspecting and regulating construction sites. As mentioned in the Deloitte analysis report (Schatsky and Ream, 2016), UAVs can increase the speed of wind turbine inspection up to 10 times in comparison to traditional methods. In the report published by (Measure, 2016), benefits and use cases of UAVs in power generation facilities are also analyzed. However, these systems are primarily intended only to assist in gathering information for human operators using manual control and telemetry.

During a firefighting mission, one of the objectives is to detect specific objects, typically flames, and to determine their coordinates. This can be done in multiple ways. One of the methods that is becoming increasingly common in practical use is computer vision. Images provide a rich information source on a relatively large section of the environments, but they require specialized processing to be utilized in autonomous systems. A survey regarding the computer vision methods and corresponding data-sets for UAV flying at different altitudes (e.g., eye-level, low-medium, aerial) is discussed in (Cazzato et al., 2020). Among the most popular ways is the utilization of the machine learning (ML) approach. For instance in (Nousi et al., 2019), a Robot Operating System (ROS)-based environment was used to investigate the applicability of various approaches (executed on Nvidia Jetson TX2) for real-time object detection and tracking. The results obtained in this paper testify

to the possibility of running Neural Network (NN)-based detection on an embedded device. While such ML methods are capable of exceptional object recognition capabilities in certain situations, they present numerous challenges. Besides the considerable onboard computational power necessary for such a hazardous application, the highly varied appearance of fires and the unpredictable backgrounds and lighting in real-world situations can make it unfeasible to obtain sufficiently representative training datasets.

In general, for projects where the goal is to detect fire, a vision based approach using classical color imaging alone may not be enough as the fire itself may be obscured from cameras by various solid surfaces or smoke. For this purpose, in (Pecho et al., 2019), a thermal camera for indoor building inspection was used. The thermal imaging proved to be a good source of data on the interior fires, but the UAV was only piloted manually, with the thermal images being transmitted for appraisal by human operators instead of being processed automatically. Thermal cameras are also useful for the detection of other objects beside fires or when, for example, detection is challenging due to light conditions in the visible spectrum (Andraši et al., 2017). Simultaneous utilisation of both thermal and RGB cameras is also a common form of sensor fusion. For instance in (Vidas et al., 2013), such a sensor combination is used for the construction of a dense 3D model of an investigated object. It is used for crowd monitoring in (Schulte et al., 2017) and for small UAV detection in (Goecks et al., 2020).

The robotic extinguishing of fire with a water stream has been previously explored, but primarily with ground robots (Liljeback et al., 2006; Rangan et al., 2013) that are mechanically stabilized by contact with the ground. The kinematics of a water stream deployed from a UAV were explored in (Lee et al., 2020) and aiming with UAV body position was explored in (Vrba et al., 2019). However, these works were not focused on correctly aiming for extended periods of time in the context of firefighting. Firefighting using UAVs has already also been a topic of other robotics competitions. (Qin et al., 2016) describes the design and implementation of a firefighting UAV for outdoor applications designed specifically for the IMAV 2015 competition.

3. MBZIRC 2020 Challenge 3

This paper is mainly motivated by the MBZIRC 2020 competition. The competition was divided into three challenges plus a Grand Challenge and final Exhibition, all of which our team participated in, each motivated by a different set of real-world problems. Challenge 3 was focused on robotic firefighting in various conditions and this paper discusses our approach to its first subtask, specifically extinguishing real fires that were emitted by a gas pipe on the facade of a building with a water stream.

3.1. Challenge subtasks

Challenge 3 of the MBZIRC 2020 comprised four distinct parts: extinguishing real fires on the exterior of a building by UAV with water, extinguishing fire analogues in the interior of a building by UAV with water, extinguishing fire analogues in the interior of a building by Unmanned Ground Vehicle (UGV) with water, and extinguishing free-standing fire analogues by UAV with fire blankets.

Each of the three challenges progressed as follows: first, each participating team was given three rehearsal runs on the dates of the 20th, 21st, and 22nd of February, 2020. These rehearsals took 15 min each for Challenge 3. Afterwards, two scored attempts were performed as part of the competition proper on the 23rd and 24th of February, 2020, once again for 15 min each. Each challenge of MBZIRC 2020 had a separate ranking based on points achieved with manual control. The teams scoring no points on autonomous control were always ranked below the lowest scoring team with at least some degree of autonomy, reflecting the significant difference in relevance to the field of robotics and the difficulty of execution. The final result of a team was selected as the better result of two attempts. Lastly, the Grand Challenge phase of the competition took place on the

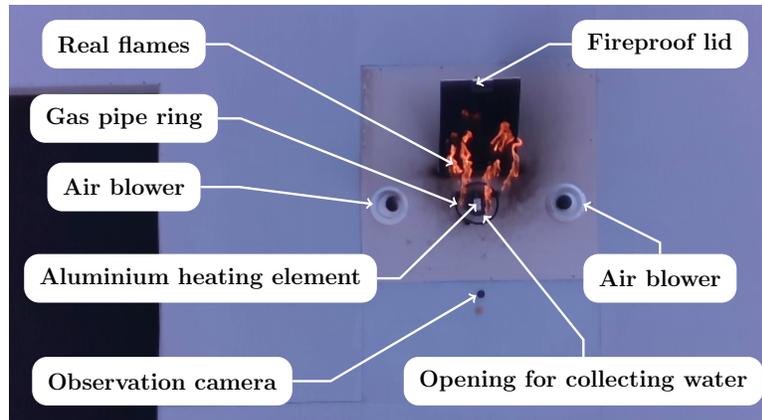


Figure 1. Description of the fire objects used in the competition.

25th of February, where parts of two other challenges (aerial grasping and wall building) were to be performed simultaneously with Challenge 3.

This model of competition enabled a fair comparison of diverse, competing solutions. Each team was afforded an equal number of public attempts and limited rehearsal time. This makes the results more trustworthy than if they were potentially cherry-picked from laboratory testing. Nevertheless, such a small number of attempts makes it impossible to obtain statistically significant test results on the performance of the systems. To address this observation, we complement these results with statistical evaluation of the precision of our aiming procedure.

3.2. Facade firefighting

The Challenge 3 subtask of extinguishing facade fires took place around the facade of a purposefully-built structure resembling a three story building of realistic proportions (referred to hereby as “building”). The real fires (henceforth collectively referred to as “fire objects”) were housed in structures located on the facade of the building comprised of multiple parts (see Figure 1):

- A circular ring of perforated metal piping producing ignited gasses — approximately 25 cm diameter
- A circular opening in the facade with a dark interior, containing a system for measurement for the volume of collected water - approximately 15 cm diameter
- An anodized aluminium heating element in the center of the opening — rectangle of 6×3.5 cm
- Air blowers on two sides of the ring active on the ground floor of the building.

These objects were spread out on 3 of the 4 sides of the building, but only one fire object was activated per floor during each run of the subtask. The rest were inactive, meaning that the fire was out and the entire object was covered by an opaque lid of the same color as the building walls. The lid was also closed on the active objects once 1 L of water had been collected through the opening. One of the objects had air blowers activated on the sides, complicating stabilization of the UAV in front of it.

While the specification of the competition task corresponds to a simplified situation compared to a real firefighting mission, it provides a well-modelled representation for such real-life missions. This is because a successful solution in the competition must contain, in some form, all the elements that would otherwise be involved in a real, autonomous, aerial, firefighting system, namely self-localization, navigation, fire source localization, and extinguishing capabilities.

For the liquid-based fire extinguishing subtasks of Challenge 3, the scoring scheme was based on the weighted sum of the amount of water deposited into each specific target. Up to 1 L of water

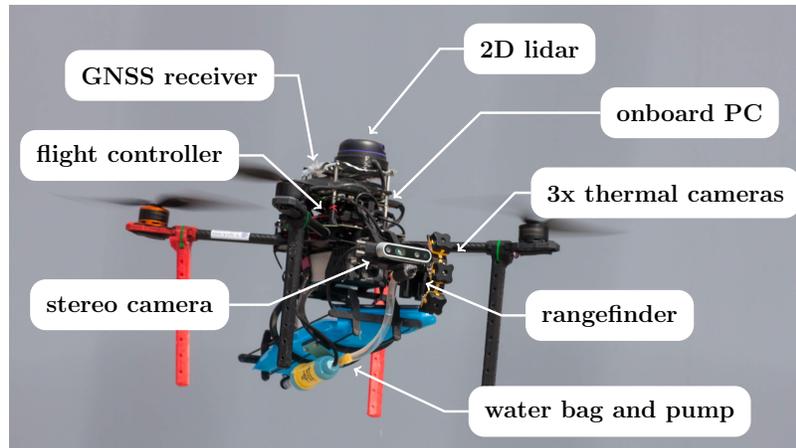


Figure 2. The description of components of the deployed UAV platform for liquid-based fire extinguishing.

could be deposited into each of the six targets. Relevant to the subtask described in this section, the fire objects on the upper floors of the building facade counted as 8 points each multiplied by the amount of deposited water in liters. The total maximum possible score in Challenge 3 of all subtasks was 100 points.

4. Platform Description

Our team participated in all challenges of the MBZIRC 2020 competition. To allow for re-usability of the system and its spare parts, our intention was to select a base UAV platform that could be used in all challenges with the possibility to modify the sensors and actuators. The UAV platform with complete sensory equipment for the task solved in this paper can be seen in Figure 2. The selected base platform is created mostly from commercially available off-the-shelf components and 3D printed parts. The platform is built from the *Tarot T650* quadrotor frame, the *PixHawk 4* flight controller¹, and an onboard computer. This frame satisfies the size limitations for the competition - diagonal dimension without propellers is 650 mm, with the 15 inch propellers attached expanding it to maximum diagonal span of 995 mm, or aligned with the front-back axis making the highest dimension 800 mm. The height of the fully equipped UAV is 410 mm. The payload capacity is sufficient for carrying additional sensors and fire extinguishing equipment. The entire unit weighs 4.1 kg, with the full water bag mentioned below increasing this weight by up to additional 1 kg. The onboard computer is Intel NUC8i7BEH² that contains Intel i7-8559U CPU and 8 GB of RAM, and runs the Ubuntu 18.04 LTS operating system and ROS(Quigley et al., 2009) Melodic middleware. For general localization, the UAV carries GNSS receiver based on the *Ublox Neo-M8N* module. Furthermore, the UAV is equipped with the *RPLIDAR A3*³, which is a 360° 2D LIDAR. This sensor provides 1,600 samples per second and can detect obstacles up to a 25 m radius depending on the setting of the sensor. We also use the *Realsense D435* camera⁴ with field of view (FoV) ($H \times V \times D$) $87^\circ \pm 3^\circ \times 58^\circ \pm 1^\circ \times 95^\circ \pm 3^\circ$ and a range of up to 10 m, but for the facade fire detection, it

¹ https://github.com/PX4/px4_user_guide/raw/master/assets/flight_controller/pixhawk4/pixhawk4_technical_data_sheet.pdf - Accessed: 03-30-2022

² https://www.intel.com/content/dam/support/us/en/documents/mini-pcs/NUC8i3BE_NUC8i5BE_NUC8i7BE_TechProdSpec.pdf - Accessed: 03-30-2022

³ https://www.generationrobots.com/media/LD310_SLAMTEC_rplidar_datasheet_A3M1_v1.0_en.pdf - Accessed: 03-30-2022

⁴ <https://www.intelrealsense.com/wp-content/uploads/2020/06/Intel-RealSense-D400-Series-Datasheet-June-2020.pdf> - Accessed: 03-30-2022

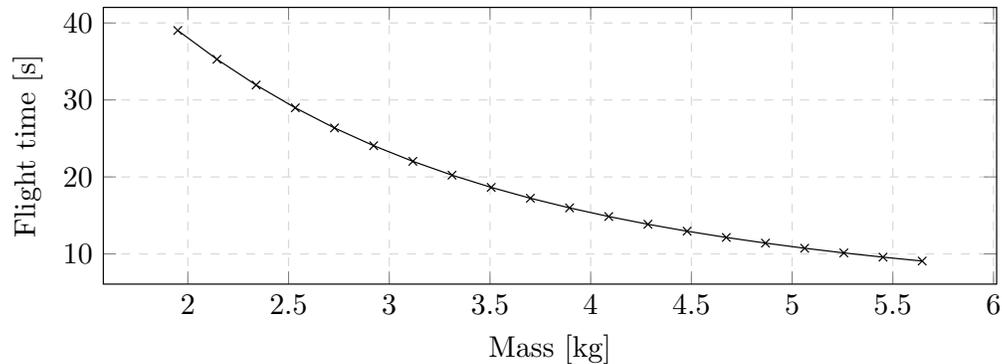


Figure 3. Empirical evaluation of the flight time of our UAV platform in relation to its total mass including all equipment.⁵

is only used as an RGB camera for precise target localization from close distances. Detection of the fires from greater distance is done using a set of three thermal cameras – *TeraRanger Evo Thermal 33*⁶. This thermal camera is cheap, small, and lightweight (only 12 g) which is very important for a limited payload, as is the case here. The cameras are arranged vertically with one pointing forwards and the two others above and below it with an orientation at 30° upwards and downwards respectively, from the first thermal camera - see Figure 2. The UAV is further equipped with down facing *Garmin LIDAR-Lite v3*⁷ laser rangefinder used as a precise altimeter.

To extinguish fires, the UAV is equipped with a water bag and a pump. The capacity of the bag was limited to 1 L of water to maintain high maneuverability of the system. This maneuverability is vital for flight in an environment with possible strong air currents close to hazards. The pump drives the water through a nozzle with a diameter of 4 mm and can fully deplete the bag in 25 s. The nozzle is rigidly attached to the UAV frame and is oriented towards the front with the spraying tip located 2 cm below and 2 cm in front of the *Realsense* camera. Aiming is done using the motion of the UAV itself.

Early in development, we had considered the option of having the nozzle independently actuated to allow for aiming uncoupled from the motion of the UAV. However for the design of the small-scale UAV for the purposes of the competition, we avoided this approach due to various development considerations and limitations. Among other things, these included the increased likelihood of unpredictable behavior in untested states brought about by added mechanical and subsequent software complexity in the system, added weight to UAV with limited carrying capacity, and precision of the available actuators. Additionally, our experience afforded us the expectation that the repair work of UAVs damaged in testing and the competition runs would be significantly more difficult and costly.

It should be noted that while water was used, because of the specifications of the competition, the platform could easily have been adapted to other liquid fire extinguishing agents, as would be demanded by real-world conditions of use.

The use of an extinguishing agent with a better weight to effect ratio would be especially beneficial when the flight time of UAV is considered. The batteries used in the competition on our platforms had the capacity of 177.6 W h, allowing for 10 min of flight with the water bag full and up to 14 min with the bag empty (see Figure 3). However, the hardware solution described here is a prototype and a larger UAV platform could allow for greater payloads and longer flight times more suitable for real missions.

⁵ https://ctu-mrs.github.io/docs/hardware/motor_tests.html

⁶ https://terabee.b-cdn.net/wp-content/uploads/2020/05/evo-thermal_specsheet.pdf - Accessed: 03-30-2022

⁷ http://static.garmin.com/pumac/LIDAR_Lite_v3_Operation_Manual_and_Technical_Specifications.pdf - Accessed: 03-30-2022

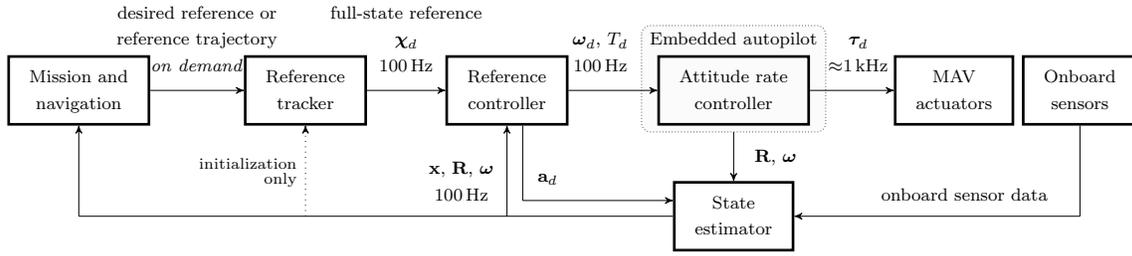


Figure 4. A diagram of the control system architecture. A *Mission and navigation* software supplies a 3D position and heading reference (\mathbf{r}_d, η_d) or a time-parametrized reference trajectory $\{(\mathbf{r}_d, \eta_d)_1, (\mathbf{r}_d, \eta_d)_2, \dots, (\mathbf{r}_d, \eta_d)_k\}$ to a reference tracker. The *Reference tracker* creates a smooth and feasible reference χ_d for a reference feedback controller. The feedback *Reference controller* produces desired thrust and angular velocities (T_d, ω_d) for the Pixhawk embedded flight controller. A *State estimator* fuses data from onboard sensors to create an estimate of the UAV translation and rotation $(\mathbf{x}, \mathbf{R}, \omega)$.

5. Software System Structure

This section describes the software components of the proposed system. All of the detailed components are executed on the onboard *Intel NUC-i7* PC.

5.1. Control and estimation of the UAV state

The UAV is controlled by a multi-layer control pipeline as depicted in Figure 4. Since the focus of this paper is for outdoor firefighting, only the basic structure of the control architecture will be described here. For a more detailed description of the whole control software platform, we refer the reader to (Baca et al., 2021). The desired trajectory supplied by a *Mission and Navigation* module (in our case, by a high-level motion planning and fire extinguishing procedure) is first processed by the *Reference tracker* (Baca et al., 2018) based on model predictive control to obtain a smooth and feasible reference for the *Reference controller*. The *Reference controller* uses this reference to provide the SE(3) geometric state feedback control (Lee et al., 2010) of the translational dynamics and orientation of the UAV. The attitude rate and thrust commands generated by the *Reference controller* are sent to the embedded *Attitude rate controller* in the flight control unit of the UAV. The feedback loop of the *Reference controller* is closed by the *State estimator*, which fuses data from onboard sensors with the UAV attitude to obtain a precise and reliable state estimate.

The state estimation process uses Kalman filtering to estimate the 3-D position of the UAV and its heading angle, along with their respective first and second derivatives. The UAV state is divided into lateral, altitude, and heading parts. Such decoupling facilitates tuning of the filter and smaller system matrices save computation resources. The lateral filter uses position corrections from GNSS and heading filter corrections from magnetometers. The altitude estimation fuses data from the built-in barometer with measurements from the laser rangefinder.

5.2. Motion planning and exploration

The precise positions of the fire objects are unknown ahead of the mission, therefore the facade must be surveyed for their localization. The strategy used here was to allow each UAVs to circumnavigate the building at altitudes corresponding to individual floors at a distance of 4.5 m while aiming their thermal cameras towards the walls of the building. This gives us thermal coverage (see section 6.1 for the thermal vision) of a vertical strip of the building facade of 9.5×2.7 m due to the extended vertical FoV of the three thermal cameras. The detection from the RGB camera (see section 6.2) is also active, but in this phase it is more likely that the fire objects are located using thermal vision due to the greater robustness of thermal vision in this subtask.

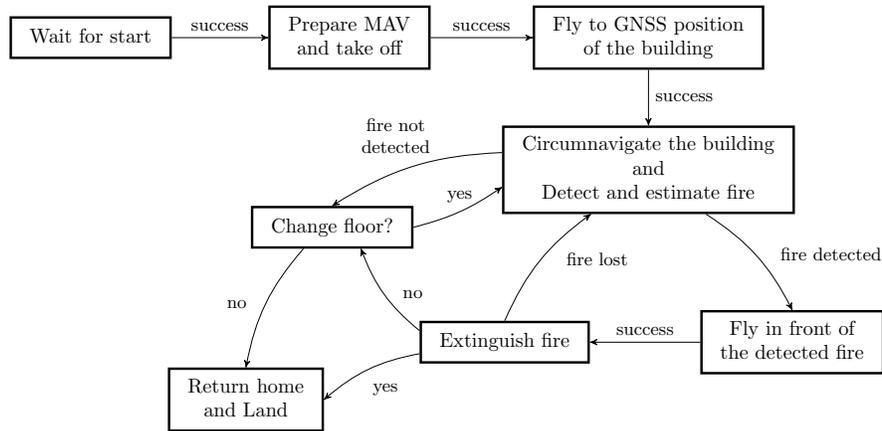


Figure 5. A high-level diagram of the main state machine. This figure shows the behavior and decision-making implemented on the UAV for the facade firefighting from takeoff to return.

Once the fire object is located, the UAV stops circumnavigating and flies in front of the object at the distance of 1.5 m from its estimated center along the normal of the corresponding wall. After reaching this position, control is handed over to the extinguishing sub-system (see section 6.4). As long as the fire object is not lost, the UAV depletes all the water being carried during the extinguishing maneuver. If the target is lost, the UAV continues to search for fire objects. After depletion of the water, the UAV flies back to its starting position and lands. In the case of circumnavigating the building without detecting a fire object, the UAV can change the flying altitude and start the search for the fires again or it can return to its starting position to land (depending on user defined settings). The latter option was not used during the competition. The complete strategy of the task is shown in the diagram in Figure 5.

In the competition, multiple UAVs are deployed in the same environment and may possibly collide with each other. To avoid this, the UAVs share their future trajectories to detect possible collisions and react to them. In the case of a threat of collision, the UAVs have different set priorities and the one with lower priority will replan its trajectory by ascending to a higher flying altitude. This collision avoidance technique is explained in detail in (Baca et al., 2018).

6. Detection, localization, and extinguishing of fire objects

The most easily detected feature of the fire objects are the real flames produced from their gas pipe ring. These are shown in the image of the thermal cameras as large blobs of more than 200 °C and if directly observed, they can be consistently detected from a significant distance. Depending on various factors such as air currents and water being deployed, their appearance in the thermal images can morph significantly (see Figure 6). The flames are observed by the thermal cameras as opaque and significantly hotter than the gas pipe rings that produce them. They are in fact so much hotter than the central aluminium heating element, that the element itself is practically undetectable (although it does occasionally visibly reflect the heat of the flames). The elements are heated to 120 °C, but our cameras observe them as having approximately 70 °C, due to the material having the emissivity value of 0.55 (Minkina and Dudzik, 2009), as opposed to the value of 0.95 that our cameras⁸ presume.

Given how the opening that collects the deployed water for evaluation is smaller than the typical area of the observed flames and is obscured by them, it is nearly impossible to aim accurately into the opening based on the thermal images alone (see Figure 9 - A1, A2 for example of unsuitable thermal

⁸ https://terabee.b-cdn.net/wp-content/uploads/2020/05/evo-thermal_specsheet.pdf - Accessed: 03-30-2022

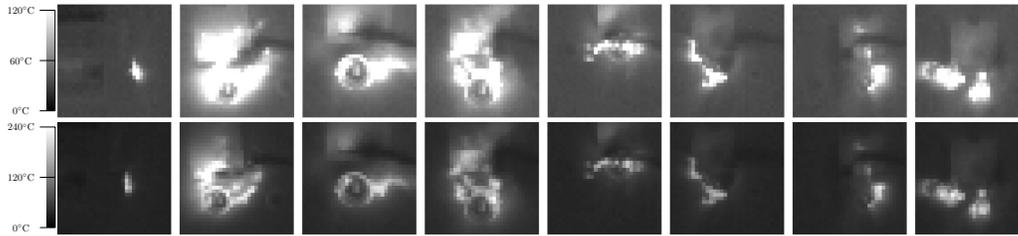


Figure 6. Thermal images of the fire objects taken with our thermal camera shown in two different thermal ranges. With the exception of the first image, these were taken within 30 s and the flames themselves are always visible. Note how different the fire object appears due to the effects of random air flow, pressure from the presence of a UAV, the water stream (visible in some images as a darker band starting from the right where the nozzle was), and other factors.

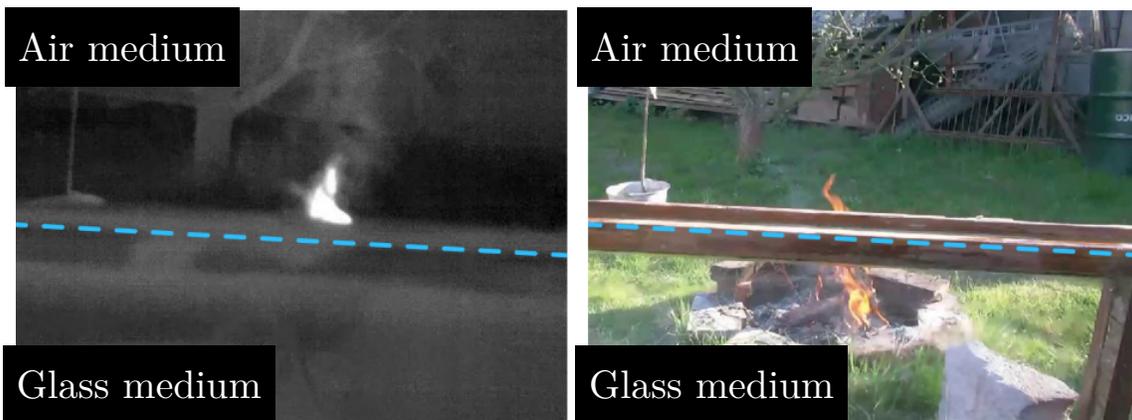


Figure 7. An example of a significant limitation of thermal vision that RGB vision can address: thermal cameras do not see heat sources through glass, posing a hurdle when searching for fires behind windows. This motivates the use of fusion of RGB and thermal vision in robotic firefighting even for the purpose of real deployment. While such conditions were beyond the scope of the competition where the target was only a small, specific region behind flames, the RGB vision can be adapted for detection of fires behind surfaces opaque to infrared radiation as a fallback.

image), despite it being possible to temporarily extinguish the flames. While in theory an applicable strategy could be to first extinguish the flames and then to aim at the aluminium heating element in the window when flames are not present, this would waste a significant portion of the limited amount of carried water and time and would additionally raise the requirement for dynamically changing detection parameters. Instead, we have decided to combine two modes of vision - thermal and RGB - to address their relative drawbacks and benefit from the strengths of both. We have based the additional vision system on an onboard RGB camera to enable aiming at the center of the opening for collecting water when the UAV has sufficiently closed in on the fire object. Using such hybrid vision in firefighting is not merely a reaction to the conditions of the competition, it is additionally motivated by the real-world problem that thermal cameras can not see through glass that blocks infrared (IR) radiation (see Figure 7) while RGB cameras can compensate by detecting fires through windows.

6.1. Thermal vision

The primary goal in the design of the thermal vision subsystem was to localize the center of the opening for water collection in the images from thermal cameras. We have observed in our tests

that the flame seen in the thermal image often covers the center we are attempting to target and that the central aluminium heating element is often too dim by itself to be consistently detected. This was made more challenging by the fact that the small resolution of the lightweight thermal cameras does not lend itself well to advanced computer vision algorithms. These observations led us to develop a simple technique for localizing a “best guess” image position for targeting. This technique was designed after surveying previously recorded datasets of thermal images of the real fire objects. From these images, we selected samples from such frames where simultaneously:

- The flame was not split into multiple individual flames fully separated by colder regions
- The colder interior of the gas pipe ring was fully enclosed by the higher temperature of the piping and flames,
- The aluminium heating element is visible and not connected to the surrounding flames and piping.

For these frames, we tuned a double temperature threshold, such that the binarized image \mathbf{B}_{temp} of pixels from the thermal image $\mathbf{I}_{\mathbf{T}}$ within these two values makes the pipe and flames a distinct contour from the aluminium heating element. In order to emphasize small objects such as the aluminium heating element, we additionally AND-combine the binarized image with thresholded matrix \mathbf{B}_{diff} of the thermal image Laplacian $\mathbf{I}_{\mathbf{L}}$ (representing a required local differential). In our case, these thresholds were as follows:

- Temperature must be above 60 °C - this selected only hot objects in the image
- Temperature must be below 250 °C - this is meant to generate additional negative contours in the hottest areas of the largest flames
- The local differential must be above 10 °C - this emphasized thin objects, such as the gas pipe ring and the heating element to increase robustness to ambient temperature shifts.

We then selected the centroid $[x, y]$ of the innermost contour \mathcal{D}_{sel} as the image target position, regardless of whether it is a positive or negative value in the binarized image, as long as the contour is not the background itself. By the innermost contour, we mean to say the one which is enclosed within the most other contours. If multiple contours $\mathcal{D}_{\mathbf{L}}$ enclosed by this number of other contours exist, we select the one with the largest area. This thermal image processing approach is summarized in the Algorithm 1. The algorithm leads to one of three cases of target selection, as follows: In the rare case that the image satisfies the requirements we had for the sample images, the output image position will correspond to the heating element in the center of the opening where we would ideally deploy water (Figure 8 - A1,A2). Otherwise, when the interior of the pipe ring is fully enclosed in an area of higher temperature, but the aluminium heating element is either not visible (Figure 8 - B1) or seen as connected to the surrounding higher temperature, the innermost contour will be the centroid of the enclosed darker section. Although it is not in the center, this is still inside the opening (Figure 8 - B2) which would usually lead to the successful deployment of water. Lastly, when the colder center is not visible or not fully enclosed by a hotter temperature, the output will be the centroid of the hotter visible area (Figure 8 - C1,C2) and will lead to water being deployed into the flames themselves to temporarily extinguish them. This in turn occasionally leads to the observed situation being closer to the sample images, due to the flames restarting slowly. As this occurs, the ring tends to be more clearly visible as enclosing the colder central region. This third case is also the closest to real firefighting, where the goal is not to precisely target a small, difficult to see area inside of a flame, but rather to actually extinguish the easier to detect flame itself. Aiming for the centroid of the heat silhouette is a good strategy for extinguishing the flame by cooling down its core. Additionally, from a distance the fire object almost always appears as a simple continuous contour (Figure 8 - C1). This detection represents the direction towards the flame and is useful for the phases of flight where the UAV is searching for new fire objects.

Algorithm 1. Retrieval of the estimated center of fire object from thermal images

```

1: function GET_THERMAL_TARGET_IN_IMAGE
2:    $I_l \leftarrow \text{GET\_LATEST\_THERMAL\_IMAGE}$  ▷  $32 \times 32$  matrix of Celsius temperatures
3:    $K_L \leftarrow \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$  ▷ Convolution kernel for approximating spatial second derivative
4:    $I_L \leftarrow \text{CONVOLVE}(I_l, K_L)$  ▷  $32 \times 32$  Laplacian image matrix
5:    $I_L \leftarrow I_L \cdot 0.25$  ▷ Per-element scaling to account for the integer values of  $K_L$ 
6:    $B_{\text{temp}} \leftarrow (250.0^\circ\text{C} > I_l > 60.0^\circ\text{C})$  ▷ Logical binary  $32 \times 32$  matrix -
    $\forall i, j : B_{\text{temp}_{ij}} = (250.0^\circ\text{C} > I_{ij} > 60.0^\circ\text{C})$ 
7:    $B_{\text{diff}} \leftarrow I_L > 10.0^\circ\text{C}$  ▷ Logical binary  $32 \times 32$  matrix -  $\forall i, j : B_{\text{diff}_{ij}} = (I_{Lij} > 10.0^\circ\text{C})$ 
8:    $B_{\text{comb}} \leftarrow B_{\text{temp}} \wedge B_{\text{diff}}$  ▷ Per matrix element logical AND
9:    $C \leftarrow \text{GET\_CONTOURS}(B_{\text{comb}})$  ▷ Selects elements  $B_{\text{comb}_{ij}}$  with neighbors of
   different value and groups their indices into sets of 8-adjacent true elements and 4-adjacent false elements:
    $C = \{C_1, C_2, \dots\}, C_N = \{[i_{N1}, j_{N1}], [i_{N2}, j_{N2}], \dots\}$ 
10:   $h \leftarrow 0_{|C|}$  ▷ Initialize hierarchy vector for count of enclosing contours per contour:  $h = [h_1, h_2, \dots, h_N]$ 
11:   $m_h \leftarrow 0$  ▷ An integer to track the current maximum enclosing count
12:  for all  $C_N \in C$  do
13:    for all  $C_M \in C$  do
14:      if  $M \neq N$  then
15:        if IS_ENCLOSING( $C_M, C_N$ ) then ▷ Check if  $C_M$  totally encloses  $C_N$ 
16:           $h_N \leftarrow h_N + 1$  ▷ Increment the appropriate element of hierarchy vector
17:          if  $h_N > m_h$  then
18:             $m_h \leftarrow m_h + 1$ 
19:          end if
20:        end if
21:      end if
22:    end for
23:  end for
24:   $D \leftarrow \{C_K \in C : |C_K| = m_h\}$  ▷ Get the innermost contours
25:   $D_{\text{sel}} \leftarrow \underset{D_L \in D}{\text{argmax}} |D_L|$  ▷ From the innermost contours, select the largest
26:   $[x, y] \leftarrow \frac{\sum_{D_{\text{sel}}} D_{\text{sel}}}{|D_{\text{sel}}|}$  ▷ Approximate contour centroid as the average x and y coordinate of its elements
   return  $[x, y]$ 
27: end function ▷ Note:  $|D|$  denotes the number of elements in set  $D$ 

```

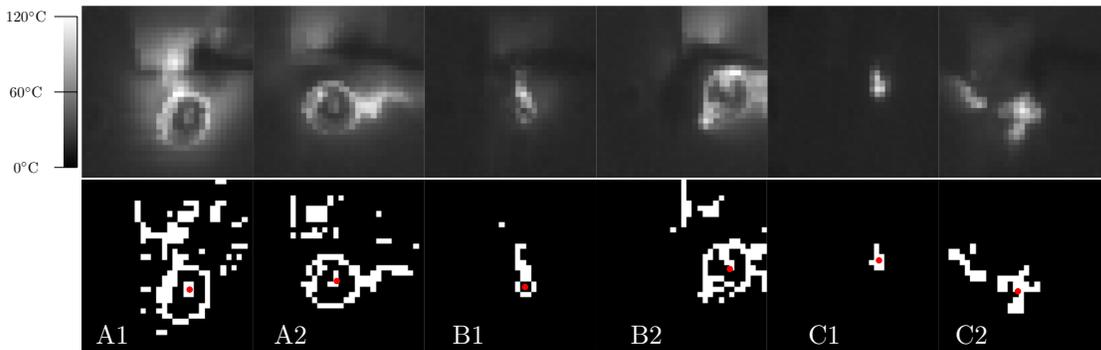


Figure 8. Thermal views (top row) and contour processing (bottom row) of the fire object for the three mentioned cases. The red dot denotes the selected contour centroid on output. A1 and A2 show cases where both the colder enclosed center of the ring and the aluminium heating element are visible. B1 and B2 are cases when only the colder center can be retrieved. C1 and C2 show the most common case where we may only directly see continuous individual flames and thus select the largest of them.

6.2. RGB vision

Since the goal of the competition was to deposit the maximum amount of water into the openings of the fire objects as opposed to extinguishing the flames, the fire itself poses a challenge in this task. Its shape and size is constantly changing under the influence of various factors such as wind, air currents from the UAV, local fuel distribution, and the extinguishing procedure. Additionally, as mentioned above, it obscures static features that we are specifically interested in localizing. Specifically, in order to locate the center of the opening for collecting water, we had to detect and localize the opening itself, the aluminium heating element, or the gas pipe ring that surrounded it. The challenge for RGB vision was, therefore, primarily in how to consistently detect the static features despite the presence of flames. We had decided to focus our detection on the gas pipe rings. They were the largest feature to surround the opening and possessed a shape that would allow us to easily retrieve the center of the opening itself. In order to see “through” the flame when it obscured the gas pipe ring and center of the opening, we exploited our observation that the flames were mostly transparent to blue light. Therefore, we have used only the blue channel image matrix \mathbf{I}_B to detect the contour of the gas pipe ring around the water collection opening.

The information from the other color channels was used for retrieval of the contours of the flames (see Figure 9 - C for extracted flame contours). Even though they were not sufficient for precise aiming, their detection was trivial and allowed the system to crop out the axis-aligned rectangle

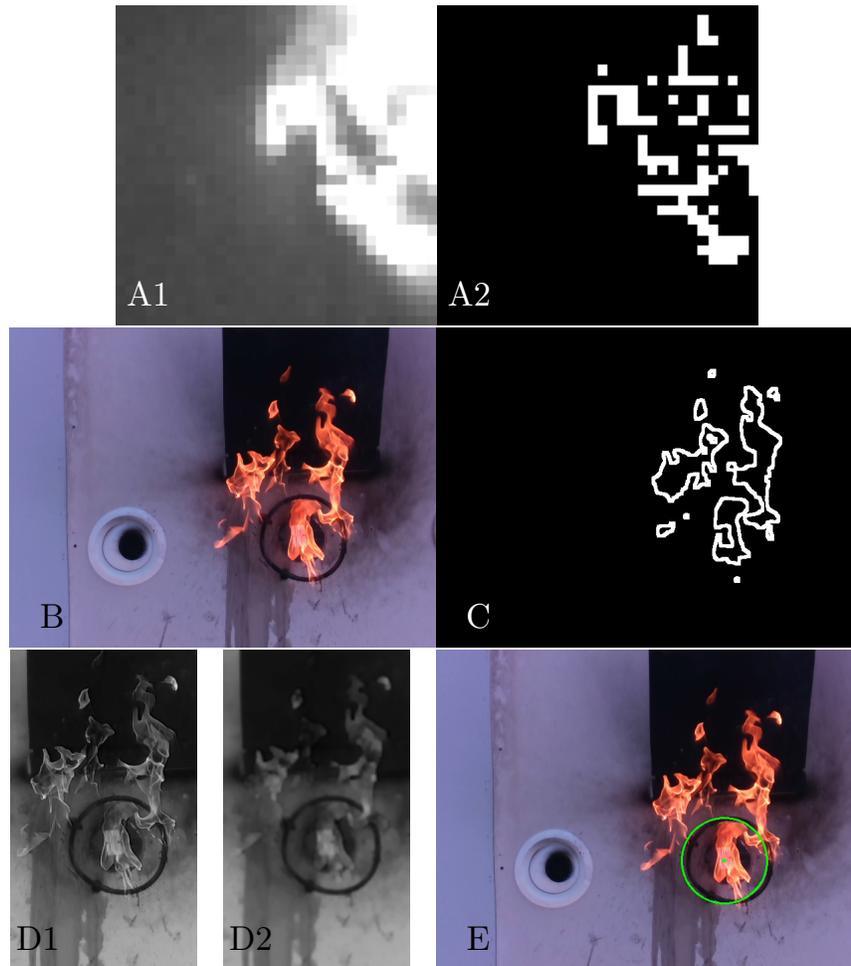


Figure 9. Steps of target detection with RGB vision in the case that the thermal vision is unsuitable for aiming.

$[s_x, s_y, w_x, w_y]$ within, which was where the gas pipe ring and the opening were most likely located (see Figure 9 - D1). This rectangle was not a tight bounding box only enclosing the flames, but rather a rectangle which contained the flames and took into account the possible influence of external disturbances. The bounding box was used to select a sub-matrix $\mathbf{I}_{\mathbf{B}_{\text{sub}}}$ from the blue channel image $\mathbf{I}_{\mathbf{B}}$ for further processing, thus saving processing power and reducing the chance of false detections. The image $\mathbf{I}_{\mathbf{B}_{\text{sub}}}$ was blurred slightly (see Figure 9 - D2) to suppress traces of the flames still visible in the form of high-image frequency features. The Canny Edge Detector was used to subsequently produce a binary image $\mathbf{B}_{\mathbf{B}_{\text{sub}}}$. Next, contour detection was applied on the image $\mathbf{B}_{\mathbf{B}_{\text{sub}}}$. For each positive contour, \mathcal{D}_M its smallest enclosing circle with a center at $[s_{cx}, s_{cy}]$ and radius s_{cr} was calculated. Comparison of these circles with their corresponding contours has been done using area comparison and the IOU (intersection-over-union) approach. Additional filtering was applied, taking into account the approximate expected image area of the circle t_a based on the physical dimensions of the gas pipe ring and the distance to the wall d that was retrieved from LIDAR readings. If the contour \mathcal{D}_M passes through the filtering described above, the center of the enclosing circle $[s_{cx}, s_{cy}]$ is used as the image position of the extinguishing target. For cases when more than one compliant contour existed in the image (meaning some object in addition to the opening was still in consideration), an additional strategy had to be applied. The following were considered:

- Accept the circle with the center closer to the center of the flame bounding box. This may be challenging when the wind level is high and, therefore, the flames are moving further away from the opening
- Discard the detection with a number of enclosing circles above one. Depending on lightning conditions, this may lead to a significant period without detection of the opening, even when the UAV is in front of it
- Track the previous detections of the opening and select the enclosing circle with a center that is closer to previous detections. This case presumes that the UAV is not moving with respect to the fire object. Alternatively, such motion must be taken into account and may be challenging due to multiple factors, including wind influence, GNSS inaccuracy, or computational complexity. Such factors make it impossible to include this data within the real-time factor. Additionally, if a number of previous readings were wrong, this may lead to further error accumulation.

The third approach was chosen for the competition. When the number of enclosing circles was equal to one, all previous readings were discarded allowing us to overcome the issue of error accumulation. The steps of the visual detection approach outlined here are expanded upon in the Algorithm 2. After the enclosing circle was selected, its center was taken as the center of the opening for water collection (see Figure 9 - E) and the fire extinguishing was directed accordingly.

6.3. Fire localization in 3D

We assumed that the thermal cameras conform to the pinhole camera model and derived their focal distance from the pixel resolution w per side and their FoV ϵ per side. Upon detection of the lowest level contour in the thermal image, we calculate the average image coordinates of the pixels of the contour as x and y . These coordinates are converted to direction vectors using the assumed camera model as

$$\mathbf{v}_{\mathbf{t}} = \begin{bmatrix} v_{t_x} \\ v_{t_y} \\ v_{t_z} \end{bmatrix} = \begin{bmatrix} 1/f & 0 & -((w-1)/2)/f \\ 0 & 1/f & -((w-1)/2)/f \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \text{ where } f = \frac{(w/2)}{\tan(\epsilon/2)}. \quad (1)$$

Subsequently, we normalize $\mathbf{v}_{\mathbf{t}}$ and transform it into a coordinate frame centered in the optical center of the camera, with the x -axis pointing forwards, the y -axis to the left, and the z -axis upwards. We call this coordinate frame the thermal base frame. The direction vector after this transformation is denoted as $\hat{\mathbf{v}}_{\mathbf{f}}$.

Algorithm 2. Retrieval of the estimated center of a fire object from RGB images

```

1: function GET_RGB_TARGET_IN_IMAGE
2:    $\mathcal{O} \leftarrow \emptyset$  ▷ initialize output set
3:    $\mathbf{I}_1 \leftarrow$  GET_LATEST_RGB_IMAGE ▷ 3-channel color image matrix from camera
4:    $[\mathbf{I}_R, \mathbf{I}_G, \mathbf{I}_B] \leftarrow$  SPLIT_RGB_CHANNELS( $\mathbf{I}_1$ ) ▷  $\forall i, j, X \in R, G, B : I_{Xij} \in \{0, \dots, 255\}$ 
5:    $[\mathbf{I}_H, \mathbf{I}_S, \mathbf{I}_V] \leftarrow$  CONVERT_TO_HSV( $\mathbf{I}_1$ ) ▷ Express the input in terms of hue, saturation and value images:
    $\forall i, j : I_{Hij} \in (0, 2\pi) ; I_{Sij} \in (0, 1) ; I_{Vij} \in (0, 1)$ 
6:    $\mathbf{B}_{\text{flame}} \leftarrow (0 < \mathbf{I}_H < 1.22) \wedge (0.52 < \mathbf{I}_S < 1) \wedge (0.40 < \mathbf{I}_V < 1)$  ▷ Binary matrix with true at elements
   where the HSV values fall within an empirically obtained ranges for flames
7:    $\mathbf{B}_{\text{flame}} \leftarrow$  MORPHOLOGY_OPEN( $\mathbf{B}_{\text{flame}}, \mathbf{J}_3$ ) ▷ Morphological opening with structuring element of  $3 \times 3$ 
   matrix of ones to suppress small random blobs
8:    $\mathcal{C} \leftarrow$  GET_POSITIVE_CONTOURS( $\mathbf{B}_{\text{flame}}$ ) ▷ Selects elements where  $B_{\text{flame},j} = \text{true}$  with neighbors of false
   value and groups their indices into 8-adjacent sets:  $\mathcal{C} = \{\mathcal{C}_1, \mathcal{C}_2, \dots\}$ ,  $\mathcal{C}_N = \{[i_{N1}, j_{N1}], [i_{N2}, j_{N2}], \dots\}$ 
9:    $\mathcal{C}_{\text{sel}} \leftarrow \underset{\mathcal{C}_M \in \mathcal{C}}{\text{argmax}} |\mathcal{C}_M|$  ▷ Select the largest contour
10:   $[s_x, s_y, w_x, w_y] \leftarrow$  GET_AABB( $\mathcal{C}_{\text{sel}}$ ) ▷ Smallest axis-aligned bounding box of  $\mathcal{C}_{\text{sel}}$  with center  $[s_x, s_y]$  and
   dimensions  $[w_x, w_y]$ 
11:   $w_x \leftarrow 1.2 \cdot w_x$  ▷ Expand the horizontal range to account for flame scatter by wind
12:   $w_y \leftarrow$  GET_VERTICAL_DIMENSION( $\mathbf{I}_1$ ) ▷ Expand the vertical range to account for flame buoyancy
13:   $s_y \leftarrow 0.5 \cdot$  GET_VERTICAL_DIMENSION( $\mathbf{I}_1$ ) ▷ Shift center to cover the entire image height
14:   $\mathbf{I}_{\text{Bsub}} \leftarrow$  SUBMATRIX( $\mathbf{I}_B, [s_x, s_y, w_x, w_y]$ ) ▷ Section of  $\mathbf{I}_B$  within the bounding box
15:   $\mathbf{I}_{\text{Bsub}} \leftarrow$  CONVOLVE( $\mathbf{I}_{\text{Bsub}}, \text{GAUSSN\_KERNEL}(5)$ ) ▷ Apply Gaussian blur with kernel size of 5
16:   $\mathbf{B}_{\text{Bsub}} \leftarrow$  CANNY( $\mathbf{I}_{\text{Bsub}}, 35, 93$ ) ▷ Apply the Canny Edge Detector with the empirically discovered
   hysteresis thresholds
17:   $\mathbf{B}_{\text{Bsub}} \leftarrow$  MORPHOLOGY_DILATE( $\mathbf{B}_{\text{Bsub}}, \mathbf{J}_3$ ) ▷ Morphological dilation with structuring element of  $3 \times 3$ 
   matrix of ones
18:   $\mathcal{D} \leftarrow$  GET_POSITIVE_CONTOURS( $\mathbf{B}_{\text{Bsub}}$ ) ▷  $\mathcal{D} = \{\mathcal{D}_1, \mathcal{D}_2, \dots\}$ ,  $\mathcal{D}_M = \{[i_{M1}, j_{M1}], [i_{M2}, j_{M2}], \dots\}$ 
19:  for all  $\mathcal{D}_M \in \mathcal{D}$  do
20:     $[s_{cx}, s_{cy}, s_{cr}] \leftarrow$  MIN_ENCLOSING_CIRCLE( $\mathcal{D}_M$ ) ▷ Center and radius of smallest enclosing circle
21:     $a_M \leftarrow$  CONTOUR_INNER_AREA( $\mathcal{D}_M$ ) ▷ Number of elements enclosed by  $\mathcal{D}_M$ 
22:     $a_C \leftarrow \pi \cdot s_{cr}^2$  ▷ Area of the enclosing circle
23:     $f \leftarrow$  GET_FOCAL_LENGTH ▷ Camera focal length in pixels
24:     $d \leftarrow$  GET_LIDAR_WALL_DISTANCE ▷ Distance from the wall based on onboard LIDAR
25:     $t_a \leftarrow \pi \cdot (f \cdot (0.125^\circ C/d))^2$  ▷ Expected image area inside the gas pipe ring based on its radius and
   observing distance
26:     $v_{\text{IoU}} \leftarrow$  IOU(GET_AABB( $\mathcal{D}_M$ ),  $[s_{cx}, s_{cy}, 2 \cdot r, 2 \cdot r]$ ) ▷ Intersection over union of rectangles
27:     $v_{\text{adiff}} \leftarrow \frac{|a_M - a_C|}{a_C}$  ▷ Relative area difference
28:    if  $(a_M > 0.7 \cdot t_a) \wedge (v_{\text{IoU}} > 0.8) \wedge (v_{\text{adiff}} < 0.3)$  then
29:       $\mathcal{O} \leftarrow \mathcal{O} \cup \{s_{cx}, s_{cy}\}$  ▷ Add the current enclosing circle center
30:    end if
31:  end for
32:  if  $|\mathcal{O}| > 1$  then
33:     $\mathcal{O}_{\text{prev}} \leftarrow$  GET_PREVIOUS_DETECTIONS
34:     $\mathcal{O} \leftarrow \{\text{GET_CLOSEST_DETECTION}(\mathcal{O}, \mathcal{O}_{\text{prev}})\}$  ▷ Member of  $\mathcal{O}$  closest to any member of  $\mathcal{O}_{\text{prev}}$ 
35:  else if  $|\mathcal{O}| = 1$  then
36:    DISCARD_PREVIOUS_DETECTIONS ▷ To prevent long-term accumulation of error
37:  end if
38:  STORE_AMONG_PREVIOUS_DETECTIONS( $\mathcal{O}$ )
39: return  $\mathcal{O}$ 
40: end function ▷ Note:  $|\mathcal{D}|$  denotes the number of elements in set  $\mathcal{D}$ 

```

For the extinguishing action done by UAVs, we also needed a distance estimate. This was achieved by combining the direction vectors with a source of surface shape measurement. We used the onboard 2D LIDAR to estimate the outline of the objects in front of the cameras.

It should be noted that the LIDAR output data and the direction vector $\hat{\mathbf{v}}_f$ have to be expressed in a common coordinate frame. This was trivial due to our knowledge of the rigid UAV geometry, and we have chosen to transform all of these into the frame of the front-facing thermal camera. The

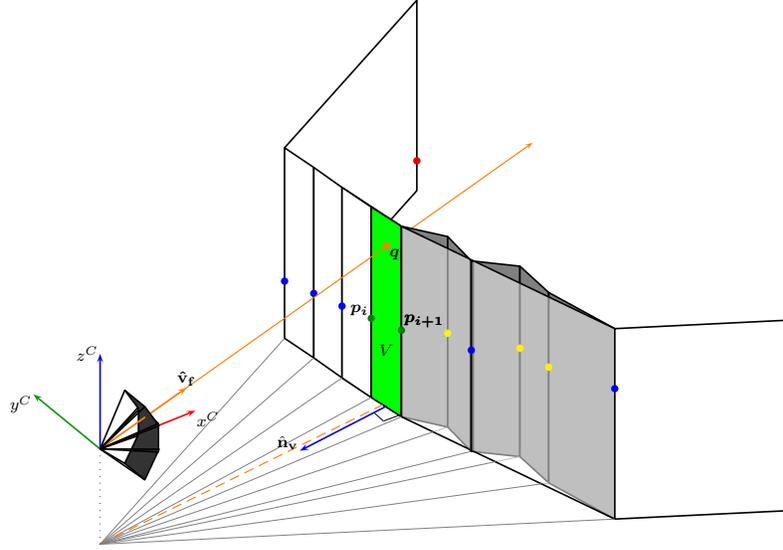


Figure 10. The 3D localization of fires on a wall based on vision and 2D LIDAR data.

time-delays between the different measurements were negligible compared to the UAV dynamics, with the output rate of the LIDAR being 20 Hz.

The output of the sensor was a set of coplanar 3D points \mathcal{P}_{lidar} ordered by their angle from the sensors (blue, green, and yellow points in Figure 10). Given that the targets were attached to a flat wall, we can presume that the outline points retrieved by the LIDAR represent the shape of the surface sufficiently well. However, the set of outline points must first be reduced to their convex hull in terms of their horizontal coordinates in order to remove measurements behind windows if they were captured by the LIDAR. This is necessary, since the fire objects can be located above or below windows, in which case if we tried to obtain their 3D positions, then they may be incorrectly evaluated as being on an interior wall of the building. The convex hull is obtained by a pseudo gift-wrapping algorithm as follows: we first remove from \mathcal{P}_{lidar} the points that are outside of the horizontal field of view of the thermal cameras. Next, we expand \mathcal{P}_{lidar} with two new points beyond the leftmost and rightmost of the remaining points in terms of their angle from the LIDAR. These points (red points in Figure 10) are generated by copying the previous leftmost and rightmost points and increasing their distance while preserving their bearing. We then select the three leftmost points (the first of which was just newly added) and check that they form a left-turning angle. If that is the case, we select the next triplet by sliding one point to the right and repeating the check on the next triplet. If a right turning point is found, the middle of the three points is discarded from \mathcal{P}_{lidar} . This process is repeated until no right turns remain and we are left with a convex hull where interior points (yellow points in Figure 10) were removed. We can now describe the surface in front of the cameras as an extrusion object composed of vertical plane segments defined by pairs of adjacent points remaining in \mathcal{P}_{lidar} (white and green surfaces in Figure 10). For each of the direction vectors retrieved from the thermal cameras, we select the vertical plane V (green plane in Figure 10) corresponding to the generating pair of 3D points \mathbf{p}_i and \mathbf{p}_{i+1} (green points in Figure 10) from \mathcal{P}_{lidar} that it passes in between of in terms of the yaw angle. We select this pair of points as

$$\{\mathbf{p}_i \in \mathcal{P}_{lidar} \mid \text{atan2}(p_{i_y}, p_{i_x}) > \text{atan2}(v_{f_y}, v_{f_x}) \wedge \text{atan2}(p_{i+1_y}, p_{i+1_x}) < \text{atan2}(v_{f_y}, v_{f_x})\}. \quad (2)$$

The 3D position of the estimated target is then obtained by calculating intersection point \mathbf{q} of the line (the orange line in Figure 10) defined by $\hat{\mathbf{v}}_f$ passing through the origin with the selected vertical plane V . The plane V is defined by a point and a normal vector, specifically in this case the

point \mathbf{p}_i and the normal vector \mathbf{n}_v , obtained as

$$\mathbf{n}_v = \begin{bmatrix} -(p_{i+1,y} - p_{i,y}) \\ p_{i+1,x} - p_{i,x} \\ 0 \end{bmatrix}. \quad (3)$$

The intersection \mathbf{q} is then calculated as

$$\mathbf{q} = \hat{\mathbf{v}}_f \cdot t, \quad (4)$$

where t is obtained using normalized vector of the surface normal $\hat{\mathbf{n}}_v$ as

$$t = \frac{\hat{\mathbf{n}}_v \cdot \mathbf{p}_i}{\hat{\mathbf{n}}_v \cdot \hat{\mathbf{v}}_f}. \quad (5)$$

The estimate of the surface normal is further used to steer the UAV to the ideal perpendicular position for extinguishing located $r_d = 1.5$ m from the fire object position estimate along the normal.

We do not consider single detection of the 3D position to be sufficient. Instead, we implemented a Kalman filter that stores multiple measurements as an array of trackers and refines each tracker state using new measurements.

The state vector of the Kalman filter used here is

$$\mathbf{q}_k = [q_{k_x}, q_{k_y}, q_{k_z}, q_{k_\eta}]^T, \quad (6)$$

where q_{k_x} , q_{k_y} , and q_{k_z} are the coordinates of the fire object in the world coordinate frame. The q_{k_η} represents the azimuth of the surface normal for that fire object. The filtering mechanism stores multiple such state vectors that are corresponding to multiple different detected fire objects. We update a specific state vector \mathbf{q}_k using a new estimate of the fire object position \mathbf{q} and normal \mathbf{n}_v , if \mathbf{q} is closer in the world frame than 1 m to $[q_{k_x}, q_{k_y}, q_{k_z}]^T$ and at the same time the horizontal component of \mathbf{n}_v is closer than 90° to q_{k_η} .

In that case, the state is refined in the standard Kalman filter correction step. Otherwise, we generate a new tracker from the measurement. To account for random errors, a state is only considered validated if at least 10 measurements have been associated with it. Additionally, the states have a decay time to discard old estimates after 10 s without updates.

Note, that the 3D pose estimation of the target position does not depend on the global self-localization of the UAV itself, as it directly uses the measurements from onboard sensors. However, we do use global localization of the UAV in order to convert the measured intersection \mathbf{q} and its normal \mathbf{n}_v into the world coordinate frame after calculating them. The main goal of this transformation is for the Kalman filter to better cope with the influence of the ego-motion of the observing UAV on the filter state vector \mathbf{q}_k . The global localization of the UAV is based on a sensor fusion from GNSS and auxiliary sensors, such as an altimeter and the inertial sensor built into the flight controller (Baca et al., 2021).

The Kalman filter brings three main benefits:

- The discarding of random errors
- The refinement of estimates as we approach the targets
- Preserving targets in memory (even if they are currently out of view).

The last point is especially significant as our thermal cameras only cover 33° horizontally at any moment. In this way, we can use their extended total vertical coverage angle to sweep the environment for fire objects.

6.4. Fire extinguishing

Upon obtaining the first validated fire object detection state in the Kalman filter array, the UAV is sent to a position 1.5 m in front of the given target along the estimated normal. The observation

angles with respect to the front wall from which the fire objects can be reliably seen are limited and the observation, as well as the extinguishing, has to be done from as close to a perpendicular position to the wall as possible. It is further useful to maintain perpendicular alignment for extinguishing as this maximizes robustness of the correct aiming with respect to drifting in an arbitrary direction.

As the UAV flies to the designated position, its estimate of the target position and surface normal improve thanks to obtaining new detections. Once it reaches the position, the control is given to the fire extinguishing sub-system. In this state (see Algorithm 3 for a pseudocode on the UAV behavior), the UAV is steered towards a specific position with respect to the detected fire object, defined the same way as the position to which the UAV was sent previously at the desired distance r_d of 1.5 m from the target along the estimated surface normal. However, continuously driving the UAV to this exact position, especially in the face of potential fire object estimation errors and disturbances such as air currents would lead to rapid tilting in attempts to correct the current position, especially close to the desired position. In our initial experiments, this led to the direction of the water stream being significantly unstable, since the water nozzle is rigidly attached to the body and thus follows the tilts. To address this issue, we have included a hysteresis to the steering. Specifically, we have defined two ranges of angle and distance offsets. These are illustrated in Figure 12. The angle ranges were defined as limits to angles formed by the surface normal and the line connecting the center of the fire object and the UAV - the inner range α_i was set to $\pm 5^\circ$ and the outer range α_o to $\pm 10^\circ$. The distance ranges are offsets from the desired extinguishing distance of 1.5 m, and these were set to ± 0.075 m for the inner range r_i and ± 0.15 m for the outer range r_o . Once the UAV has reached the inner ranges (red zone in Figure 12, it is forbidden to correct its horizontal (X-Y) coordinates despite any disturbances and it can only correct its altitude and heading as changing these does not generate tilt in the UAV. The UAV thus tends to drift or “float”. Continuously correcting the

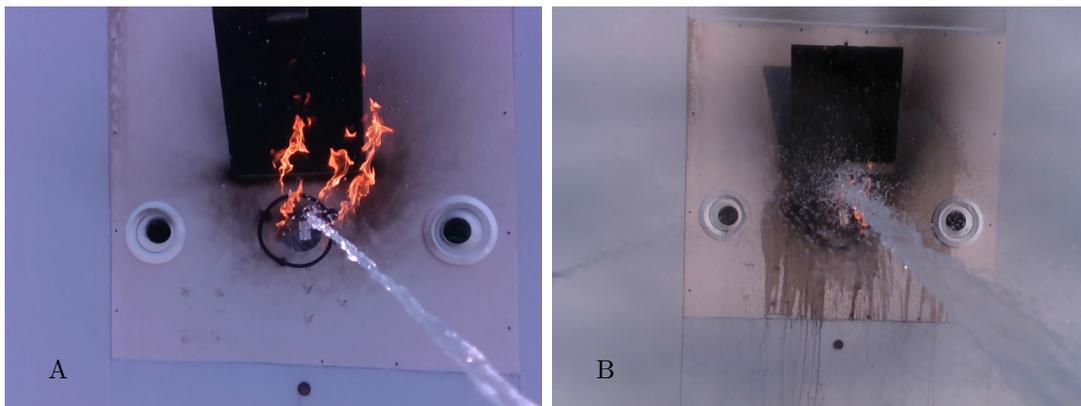


Figure 11. An onboard view of the fire extinguishing in the first run of the competition (A) and in the final exhibition (B).

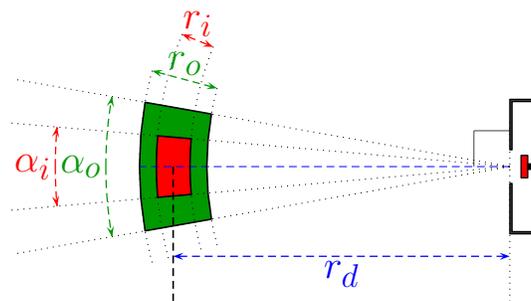


Figure 12. Steering hysteresis for fire extinguishing viewed from the top.

Algorithm 3. High-level control during fire extinguishing

```

1: function EXTINGUISH_FIRE
2:    $t_{qk} \leftarrow \text{GET\_CURRENT\_TIME}$  ▷ Initialize the pose estimate time stamp
3:    $W \leftarrow false; H \leftarrow false$  ▷ Initialize water pump state to off and the pose hysteresis to Controlling
4:    $w_t = w_{ti}$  ▷ Initialize the remaining water spraying time to measured value (25 s)
5:   while  $w_t > 0$  &  $\text{IS\_RECENT}(t_{qk})$  do ▷ Until the water is depleted or the target tracking is lost
6:      $q_k, t_{qk} \leftarrow \text{GET\_LATEST\_FIRE\_ESTIMATE}$  ▷  $[q_{k_x}, q_{k_y}, q_{k_z}, q_{k_\eta}]^T$ , time stamp
7:      $v_t, t_{vt} \leftarrow \text{GET\_LATEST\_THERMAL\_DIRECTION}$  ▷ Target direction form thermal vision, time stamp
8:      $v_v, t_{vv} \leftarrow \text{GET\_LATEST\_RGB\_DIRECTION}$  ▷ Target direction form RGB vision, time stamp
9:      $o_p \leftarrow \text{GET\_LATEST\_ODOMETRY}$  ▷ UAV pose in world frame  $[o_{p_x}, o_{p_y}, o_{p_z}, o_{p_\eta}]^T$ 
10:     $W_p \leftarrow W$  ▷ Store previous water pump state
11:     $\hat{n}_{vk} \leftarrow \text{ROTATE\_ABOUT\_Z}([1, 0, 0]^T, q_{k_h})$  ▷ Surface normal from heading in target pose estimate
12:     $s \leftarrow [q_{k_x}, q_{k_y}, q_{k_z}]^T + r_d \cdot \hat{n}_{vk}$  ▷ Calculate the desired extinguishing position
13:     $R_i, R_o \leftarrow \text{CHECK\_INNER\_OUTER\_RANGE}(o_p, s)$  ▷ Check compliance with the ranges in Figure 12.
14:    if  $H = false$  &  $R_i = true$  then
15:       $H \leftarrow true$  ▷ Set pose hysteresis state to Floating
16:    end if
17:    if  $H = true$  &  $R_o = false$  then
18:       $H \leftarrow false$  ▷ Set pose hysteresis state to Controlling
19:    end if
20:    if  $\text{IS\_RECENT}(t_{vv})$  then ▷ Fresh direction from RGB vision takes priority
21:       $d \leftarrow v_v$ 
22:       $e_z \leftarrow (v_{v_z} - z_z) \cdot r_d$ 
23:    else if  $\text{IS\_RECENT}(t_{vt})$  then ▷ Thermal vision is more consistent but less precise
24:       $d \leftarrow v_t$ 
25:       $e_z \leftarrow (v_{t_z} - z_z) \cdot r_d$ 
26:    else ▷ The pose estimate from Kalman filter is a fallback
27:       $d \leftarrow [q_{k_x}, q_{k_y}, q_{k_z}]^T - [o_{p_x}, o_{p_y}, o_{p_z}]^T$ 
28:       $e_z \leftarrow d_z$ 
29:    end if
30:     $e_\eta \leftarrow \text{GET\_HEADING\_DIFFERENCE}(z, d)$  ▷  $z$  is the direction of the water nozzle
31:    if  $H = true$  then
32:       $n = 0$  ▷ Send only one trajectory point
33:       $r_0 \leftarrow [o_{p_x}, o_{p_y}, o_{p_z} + e_z]^T, \eta_0 \leftarrow o_{p_\eta} + e_\eta$  ▷ Control only the heading and altitude
34:       $W \leftarrow true$  ▷ Deploy water
35:    else
36:       $c \leftarrow [s_x, s_y]^T - [o_{p_x}, o_{p_y}]^T$  ▷ Get horizontal offset from the desired extinguishing position
37:       $n \leftarrow \|c\| / (v_{max} \cdot t_s)$  ▷ We need  $n + 1$  trajectory points with time step of  $t_s$ 
38:       $c_i \leftarrow c \cdot (v_{max} \cdot t_s / \|c\|)$  ▷ We request returning speed of  $v_{max}$ 
39:      for  $i = 0$  to  $n$  do
40:         $r_i \leftarrow [o_{p_x} + i \cdot c_{ix}, o_{p_y} + i \cdot c_{iy}, o_{p_z} + e_z]^T, \eta_i \leftarrow o_{p_\eta} + e_\eta$ 
41:      end for
42:       $W \leftarrow false$  ▷ If pose hysteresis state is not Floating, turn the water pump off
43:    end if
44:     $\mathcal{P}_{ref} \leftarrow \{(r_0, \eta_0), \dots, (r_n, \eta_n)\}$ 
45:     $\text{SET\_TRAJECTORY\_REFERENCE}(\mathcal{P}_{ref})$  ▷ Send the trajectory points to the reference tracker
46:    if  $W \neq W_p$  then ▷ If the state has changed, turn the water pump on or off
47:       $\text{TOGGLE\_WATER\_PUMP}(W)$ 
48:    end if
49:    if  $W = true$  then
50:       $w_t \leftarrow w_t - dt$  ▷ Keep track of the remaining time of water deployment
51:    end if
52:     $\text{SLEEP}(dt)$  ▷ Enforce constant time period (0.1 s)
53:  end while
54:   $\text{TOGGLE\_WATER\_PUMP}(false)$ 
55:   $\text{YIELD\_CONTROL\_TO\_STATE\_MACHINE}$ 
56: end function

```

heading was necessary, otherwise the drifting would throw off the aim of the UAV. It is in this state that the UAV is allowed to spray water. The Z-coordinate and heading are controlled either to aim at the directly observed target (water collecting opening or significant flames, see section 6.1 and 6.2), or (if it is currently not visible *e.g.* due to being cooled down by the water) to aim at its estimated position from the Kalman filter. If the target is directly observed, the aim is more responsive to disturbances, while if the aim relies on the filter, the precision is lowered. Despite this, we have determined that for success of the mission it is better to spray water even when we do not have direct observation, since such observation was often lost and rapidly starting and stopping the water stream affected the precision more negatively than drifts in the filter. The UAV is only allowed to correct its X-Y coordinates again once it has been moved outside of the outer ranges (green zone in Figure 12), at which point spraying water is disabled. Tuning this behavior took a considerable amount of preparation time as its success was dependent on many variables not present in our simulation and testing was demanding in terms of time, personnel, and conditions. Apart from the actuated nozzle approach mentioned in section 4, a notable theoretical alternative to this approach - a paradigm that currently has the attention of the robotics and UAV communities (Rajappa et al., 2015; Convens et al., 2017; Tadokoro et al., 2017) - are fully-actuated platforms that allow for direct tilt control decoupled from horizontal translation. Such platforms are however highly experimental and we have thus decided for better-tested conventional platforms for this task.

7. Performance Results

7.1. Preparation

Before the competition officially began, the system was incrementally developed and rigorously tested. The preliminary testing and assembly of hardware components was done in Czech Republic. However, six weeks before the competition, the team relocated to a desert area in the United Arab Emirates. The reasons for this were twofold: the weather was unsuitable for flying in Czech Republic shortly before the competition due to it being winter. Our long-term experience says that the weather, lighting, and other local conditions have profound effects on the results of robotic deployment, especially in terms of sensor outputs, and properly preparing a robotic system to reliably operate in a certain region requires testing in the same region. For instance, the stronger sunlight and higher temperatures in UAE affected thermal imaging in ways that were not apparent in our central European country.

In UAE, we performed numerous tests for establishing the right approach to the task as described in the previous sections, as well as for tuning parameters. Initially, we performed tests with only the fire extinguishing subsystems on fire analogues (see Figure 15.). In these tests, we had shown that the aim was correct on an average of 93% of the time when performing the fire extinguishing actions (Section 6.4 (Spurny et al., 2021)), which shows the effectiveness of the 3D localization and extinguishing procedures in case of reliable target detection. In one of the latest stages of the development in the desert environment, we performed successful following of a wall, followed then by spraying water onto a fire analogue (this was captured on video⁹ and samples are in Figure 16). The fire analogue was identical to fire analogues used in other parts of the fire challenge of the MBZIRC 2020 competition, since we could not safely use open flames. These were detected only with thermal vision in a significantly less complex mode where only an aluminium heating element was detected as a small concentrated spot of elevated temperature. However, these tests allowed us to develop and tune the fire extinguishing procedure described in section 6.4. The plot of the extinguishing trajectory from such a test, shown in Figure 16, clearly demonstrates compliance of the UAV motion with the motion hysteresis from Figure 12.

These real-world tests were interspersed with frequent testing in Gazebo simulation (see Figure 13). This type of simulation testing is a staple of modern robotics and is extremely useful for

⁹ <http://mrs.felk.cvut.cz/fr2020firechallenge-facadefires>

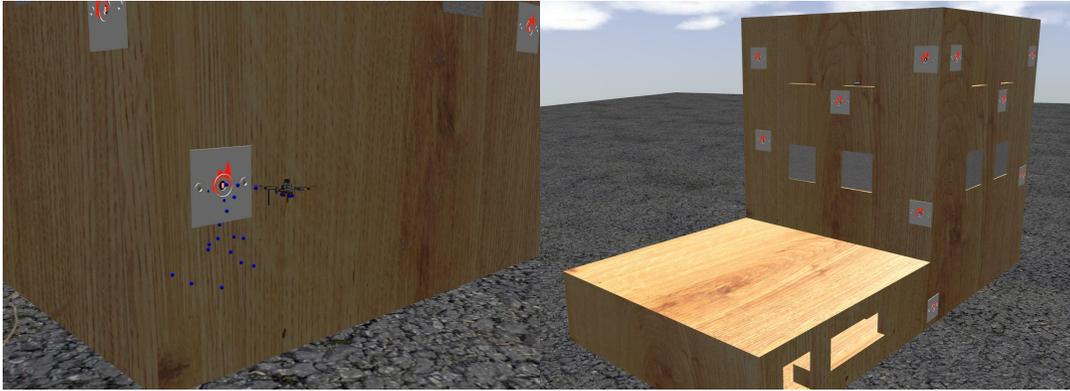


Figure 13. Extinguishing of facade fire objects and building overview in Gazebo simulator using our plugins.

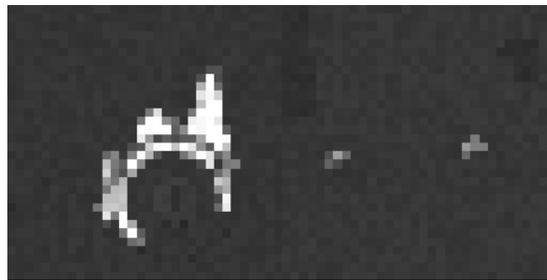


Figure 14. Outputs of the simulated version of our thermal cameras in Gazebo simulator.



Figure 15. Preliminary test in the desert with only fire extinguishing on a fire analogue.

saving resources and time when testing new changes to a system. In the simulator, in addition to emulating all the hardware onboard our *T650*-based UAVs, we have developed a plugin for emulating thermal cameras with outputs tuned to closely match the outputs of our real thermal cameras (see Figure 14), as well as the water spraying system. This software is open-source and is available online¹⁰. This allowed us to run the very same software, including the thermal vision, on the simulated UAVs as on the real UAVs, streamlining the development.

¹⁰ <https://github.com/ctu-mrs>

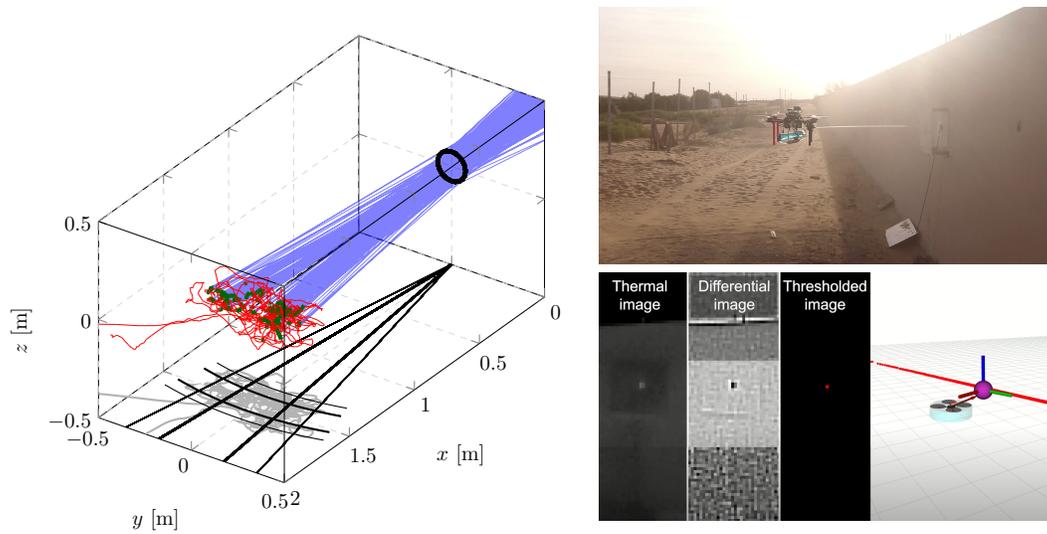


Figure 16. Extinguishing of fire analogue during our preparations in the desert. Green section of the trajectory on the left are points from which the water pump was activated, while blue lines represent the theoretical lines of water stream.



Figure 17. Bottom view of the UAV detecting a fire object during rehearsals of the competition.

7.2. Competition

Before the competition, we had been testing flights in the competition arena during rehearsal runs (Figure 17). This allowed us to check for potential negative effects of the environment, as well as to record datasets of the real fire objects. In the competition itself, we flew in two main trials of Challenge 3, as well as in the Grand Challenge of the competition and in the Final exhibition before the award ceremony. In the two rounds of trials, our team competed in the subtask of extinguishing the real fire objects described in this paper with two UAVs. We also competed in the subtasks of extinguishing ground fires with fire blankets with one UAV and extinguishing interior fires with a UGV.

Trial 1

In the first trial of the competition, one of our UAVs successfully sprayed water into one of the facade fire objects. This can be seen in Figures 18 and 11 - A. Unfortunately, this amount was very small and registered as 8 mL. This result was due to the conditions for spraying water being too strict in this phase. Specifically, the UAV was only permitted to spray water if both thermal and RGB vision provided fresh target localization to prevent wasting the limited capacity of carried water, as opposed to merely having active tracking from previous detection. This requirement was relaxed in subsequent runs to allow spraying in case of direct observation with either thermal vision or RGB vision, as long as the UAV was within the Floating phase of the pose hysteresis (see Algorithm 3).

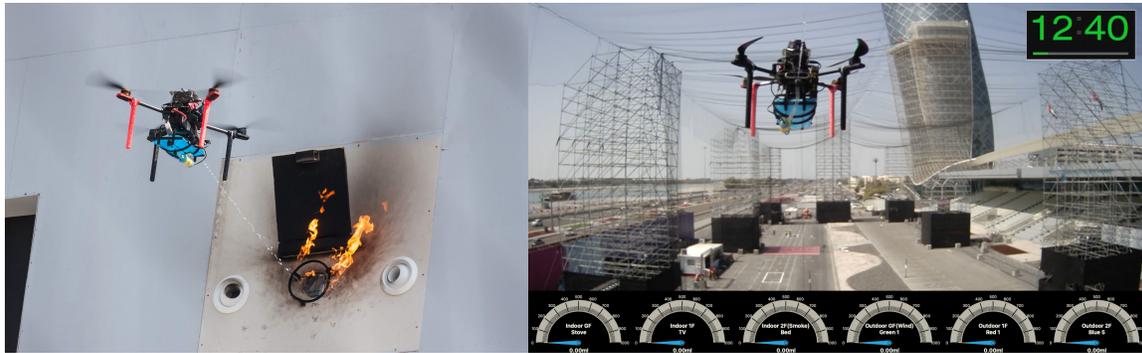


Figure 18. Active extinguishing of a facade fire object in the first trial of the competition.

Additionally, water spraying was allowed regardless of the directional offset between the water nozzle and the estimated target direction. This meant that in the subsequent runs, water would more often be sprayed off-target, but the target was more likely to be hit with significantly more water (see subsection Final Exhibition below). In this trial, we also lost a significant amount of time due to high interference with the GNSS signal. Because of this, one of the water spraying UAVs did not take off after an emergency landing and another for more than 5 min. The primary culprit was the influence of sensors connected using USB 3.x. One of these devices onboard our UAVs was a *Realsense D435* camera. The precision of most GNSS localization systems can be severely affected by the components transmitting data via the cable at frequencies close to those used by GNSS. See sheet¹¹ for a detailed description of USB 3 frequency interference. However, we had observed these effects before and made some changes to our UAVs to mitigate them. Notably, we added multiple layers of aluminium foil between the GNSS and the rest of the components, but we could not fully overcome the severity of this interference. Our team did, however, successfully deploy over 1 L of water into a fire analogue using a UGV.

Trial 2

In the second trial, we did not successfully deploy water into the fire objects due to two separate incidents. The GNSS of one of the UAVs recorded a large sudden jump of position estimation while close to the building. Even though we have implemented a virtual repulsion force based on the LIDAR data, the jump was too large for this safeguard to overcome and the UAV hit the wall of the building in an attempt to correct its perceived erroneous position. The second incident occurred when the other UAV correctly detected and approached a fire object above the roof of a lower part of the building. The onboard laser altimeter of the UAV suddenly detected the change in perceived altitude when crossing over the edge of the roof. The UAV did not immediately crash, but started to oscillate in altitude which was exacerbated by the flexible textile lining of the roof oscillating due to the air pressure from the propellers of this UAV. Additionally, the fire object that the UAV approached had its air blowers active. Upon approach, the two combined sources of disturbance fatally destabilized the UAV, making it crash on top of the roof. Despite these problems, our team did score seven points with our blanket placing UAV.

While the total maximum possible score in Challenge 3 was 100 points, the achieved score of the victorious team was only 12.2626. This was followed by two teams tied at 10 points each, and our team tied in fourth place with another team at 7 points. The time left when a team declares the end of their turn was used as a rank tie-breaker when possible.

¹¹ <https://www.intel.com/content/www/us/en/products/docs/io/universal-serial-bus/usb3-frequency-interference-paper.html> - Accessed: 03-30-2022

Grand Challenge

The Grand Challenge of the competition consisted of performing selected subtasks of each of the three challenges of the MBZIRC 2020 in one time-slot of 25 min. We had prepared two UAVs for tasks from Challenge 1, one UAV for the task from Challenge 2, and one UGV and one UAV for tasks from Challenge 3 since our team had participated in all three challenges with full robotic autonomy in each. The last UAV was equipped to perform the liquid-based fire extinguishing described in this paper. In order to qualify for the financial award, a team had to score at least some points autonomously from each of the three reduced challenges. Unfortunately, we had to suspend deployment of the fire extinguishing UAV in favor of repeated attempts at successful scoring points in the reduced Challenge 2. At this point of planning, we had already scored 72 points in the reduced Challenge 1, as well as 12.5 points in the reduced Challenge 3 thanks to the success of the UGV. Only points from Challenge 2 were absent due to various complications with the associated UAV. Having the water-spraying UAV in air could interfere in various ways with this UAV, and therefore we recalled it shortly after takeoff and grounded it for the rest of the Grand Challenge. Overall, we did not score points in the reduced Challenge 2. Regardless, we ended up with the best total ranking in the Grand Challenge, making us the victorious team.

Final exhibition

After the Grand Challenge and before the award ceremony where the rankings were to be announced, the teams with the best results so far were offered the chance to perform a final exhibition of their systems. As our team qualified, we prepared one UGV and one UAV for repetition of Challenge 2, and one UGV and two UAVs for repetition of Challenge 3. The last two UAVs were once again prepared to do the task described in this paper and did so successfully in this last run. The UAVs found one facade fire object each and proceeded to deploy water into them, even simultaneously for a short time. Despite one of them hitting the wall at the end of this attempt due to the same type of GNSS jump as in Trial 2, by that point they deposited 56 mL and 61 mL respectively into the openings. Lastly, after the other UAV had depleted its water onto the fire object, it safely returned to the landing area without damage. This UAV is shown extinguishing fire during the final exhibition in Figures 11 - B and 20 with its trajectory in this attempt shown in Figure 19.

8. Beyond the MBZIRC and Future Work

Based on the lessons learned in MBZIRC, we are currently working on various improvements to the system for realistic firefighting in order to address the main drawbacks of the solution optimized for the competition trials. During the competition, it was also notable that of the two deployed fire extinguishing UAVs, one had a 3D-printed water nozzle made with less precision. This resulted in the side effect of its water stream being spread noticeably wider (see Figure 11) which was detrimental in terms of the competition. This has even lead to more water spillage outside of the measuring receptacle during preliminary testing. However, in practical terms this appears to be the superior design as the more spread-out water stream tended to extinguish the actual flames more consistently.

Another observation made during the competition concerns the nozzle aiming procedure. The nozzle, which was rigidly attached to the UAV body, made the aiming complex due to the need to use the already constrained degrees of freedom of the UAV. The fire extinguishing was less flexible than with a gimbaled nozzle and was noticeably disturbed by drifting. We have previously addressed this issue by temporarily allowing the UAV to drift while only correcting the altitude and heading, but this solution solves the problem only partially in a real application. One potential solution to this problem would be to decouple the aiming from the UAV position control by using a gimbaled nozzle actuated by servos based on target localization. If that had been the case, hovering in front of the fire objects would have been significantly smoother and the extinguishing would have become more

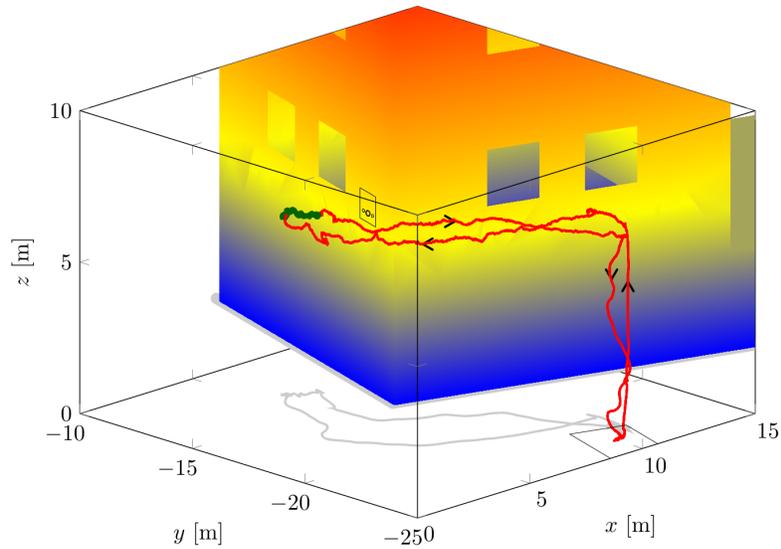


Figure 19. The trajectory of one of the extinguishing UAVs during the final exhibition, as recorded based on onboard GNSS receiver. The UAV took off at the designated area, circled around the building until it made a detection, and then localized a fire object. It then proceeded to approach it, spray water until the water was depleted, and then returned to the assumed position of the take-off area. The green section of the trajectory shows where the water pump was active. The notable drift during this time is difficult to analyze post-hoc and the most likely explanation for it is merely drift of the GNSS itself being dynamically compensated for by the sensor-based fire object position estimate. This also means that the shown position of the building is only approximate.

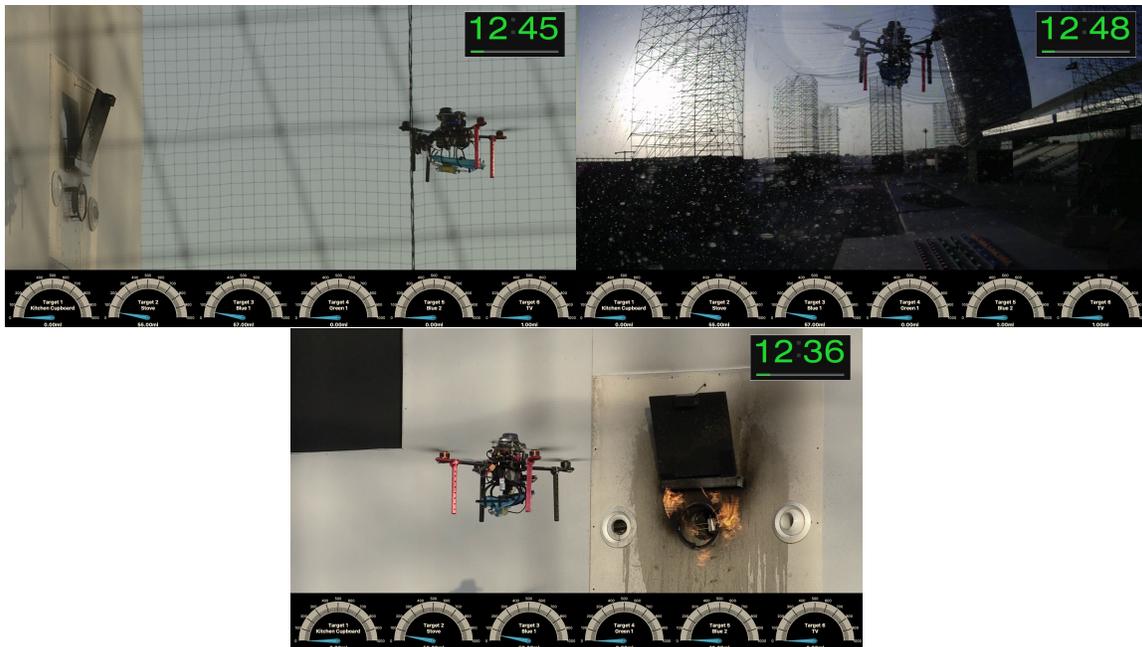


Figure 20. External views of the fire extinguishing in the final exhibition of MBZIRC 2020. The trajectory the UAV took is shown in Figure 19. The view from the onboard RGB camera for this flight is in Figure 11 - B and onboard thermal camera views from this flight are used in Figure 6.



Figure 21. Our working prototype of a firefighting UAV designed for using gas-based extinguishing capsule launcher - an evolution of the system developed for the MBZIRC 2020 competition.

consistent with fewer interruptions. We have not used this approach for the considerations discussed previously, but for a more stable UAV of larger scale equipped for a real firefighting deployment, this could also be a viable alternative to our approach, provided that the effects of the additional degrees of freedom in the system are properly studied.

However, we realized together with our industrial partners that spraying a liquid agent with a nozzle is altogether a sub-optimal approach to the task regardless, as it induces the need to involve a precisely-controlled UAV in situ for an extended period of time. Moreover, the firefighting agent needs to be dispersed throughout the entire inflamed room all at once to prevent re-ignition. To solve these requirements, we propose deploying capsules with a rigidly enclosed fire extinguishing agent into a localized fire using an onboard gas-based launcher. Such a solution makes it necessary to correctly aim using the thermal-based detection and servoing only for the instant before deployment of the capsule. The temporary disturbance that the launcher enacts on the UAV will not affect the firefighting effort itself. An overview of a prototype UAV with such a highly effective firefighting launching device can be seen in Figure 21.

9. Conclusion

A system designed for the fast extinguishing of fires on buildings by autonomous aerial robots equipped with onboard sensors and processing power was presented in this paper. The proposed solution enables scanning of a facade of a building while searching for hot regions indicating the presence of fire, localization of the fire source, and alignment of an onboard firefighting mechanism to precisely apply a liquid fire extinguishing agent. The novelty of the system lies in its onboard hybrid vision used for localizing facade fires, its capability of additionally observing surfaces behind flames, and the deployment of an extinguishing liquid based on body-pose-based dynamic aiming. In addition to the firefighting maneuvers, the system performs complex missions, takeoff to landing with full autonomy by exploiting onboard subsystems for UAV state estimation, localization, stabilization, and motion planning.

The fire localization and extinguishing systems of this solution are suitable for practical deployment, as was demonstrated during the MBZIRC 2020 competition that sought to emulate a real first-responders mission. According to the best of our knowledge, it was the only approach that achieved successful UAV-based firefighting of real fires in the MBZIRC 2020 without using RTK-GNSS or manual intervention - both of which were penalized in the competition as they do not follow the requirements of real-world applications. The presented solution was one of the key components of the system designed by our team - CTU-UPENN-NYU - that won the Grand challenge of the competition and was also successful in all three individual challenges. More importantly, the performance of the system in the MBZIRC 2020 competition motivated further research and development towards an industrial solution of this important task, as was the main goal of the MBZIRC 2020 organizers and its technical committee.



Figure 22. Photos from the MBZIRC 2020 competition in Abu Dhabi, United Arab Emirates.

Besides the system design itself, this paper also documents several practical observations made during the development and the competition itself on the performance of the equipment used. We believe that these can be of additional value for the future development of similar systems. We have documented the need for computer vision that is robust to the highly dynamic appearance of fires in RGB and thermal vision, while highlighting the individual drawbacks of each. For RGB vision, we have pointed out how the transparency of flames to blue light can be useful for observing surfaces of interest hidden behind flames. Lastly, we have shown the need for robust aiming and localization that is resistant to various types of platform drifting and have documented more specific adverse effects, such as the interference of USB 3.x connected devices with the onboard GNSS receivers.

Acknowledgments

The outstanding results within the framework of the MBZIRC 2020 project could not have been achieved without the full cooperation of each member of our team CTU-UPENN-NYU (Figure 22). This work has been supported by the Czech Science Foundation (GAČR) under research project No. 20-10280S, Research Center for Informatics project CZ.02.1.01/0.0/0.0/16_019/0000765, CTU grant No. SGS20/174/OHK3/3T/13, EU H2020 project AERIAL CORE No. 871479, and by Khalifa University via sponsorship in support of 15 selected teams for the MBZIRC 2020 competition.

ORCID

Viktor Walter  <https://orcid.org/0000-0001-8693-6261>
 Vojtěch Spurný  <https://orcid.org/0000-0002-9019-1634>
 Matěj Petrлік  <https://orcid.org/0000-0002-5337-9558>
 Tomáš Báča  <https://orcid.org/0000-0001-9649-8277>
 David Žaitlík  <https://orcid.org/0000-0001-9116-6265>
 Lyubomyr Demkiv  <https://orcid.org/0000-0002-2802-3461>
 Martin Saska  <https://orcid.org/0000-0001-7106-3816>

References

Andrašić, P., Radišić, T., Muštra, M., and Ivošević, J. (2017). Night-time detection of uavs using thermal infrared camera. *Transportation Research Procedia*, 28:183–190.

- Ashour, R., Taha, T., Mohamed, F., Hableel, E., Kheil, Y., Elsalamouny, M., Kadadha, M., Rangan, K., Dias, J., Seneviratne, L., and Cai, G. (2016). Site inspection drone: A solution for inspecting and regulating construction sites. pages 1–4.
- Baca, T., Hert, D., Loianno, G., Saska, M., and Kumar, V. (2018). Model Predictive Trajectory Tracking and Collision Avoidance for Reliable Outdoor Deployment of Unmanned Aerial Vehicles. In *IEEE IROS*, pages 1–8.
- Baca, T., Petrlik, M., Vrba, M., Spurny, V., Penicka, R., Hert, D., and Saska, M. (2021). The MRS UAV System: Pushing the Frontiers of Reproducible Research, Real-world Deployment, and Education with Autonomous Unmanned Aerial Vehicles. *Journal of Intelligent & Robotic Systems*, 102(26): 1–28.
- Cazzato, D., Cimarelli, C., Sanchez-Lopez, J., Voos, H., and Leo, M. (2020). A Survey of Computer Vision Methods for 2D Object Detection from Unmanned Aerial Vehicles. *Journal of Imaging*, 6:78.
- CBInsight (2020). 38 Ways Drones Will Impact Society: From Fighting War To Forecasting Weather, UAVs Change Everything. Available: <https://www.cbinsights.com/research/drone-impact-society-uav/> (Accessed: 01-09-2020).
- Chen, R., Cao, H., Cheng, H., and Xie, J. (2019a). Study on Urban Emergency Firefighting Flying Robots Based on UAV. In *IEEE IAEAC*, volume 1, pages 1890–1893.
- Chen, T., Yuen, A. C. Y., Yeoh, G., Yang, W., and Chan, Q. (2019b). Fire risk assessment of combustible exterior cladding using a collective numerical database. *Fire*, 2:11.
- Convens, B., Merckaert, K., Nicotra, M. M., Naldi, R., and Garone, E. (2017). Control of fully actuated unmanned aerial vehicles with actuator saturation. *IFAC-PapersOnLine*, 50(1):12715–12720.
- Emicontrols (2018a). 3 new taf20 firefighting robots for shanghai. Available: <https://www.emicontrols.com/en/stories/3-new-taf20-firefighting-robots-for-shanghai-16/> (Accessed: 06-17-2020).
- Emicontrols (2018b). Taf35 firefighting robot – for more safety when fighting fires. Available: <https://www.emicontrols.com/en/fire-fighting/mobile-fire-protection/firefighting-robot-taf35> (Accessed: 06-17-2020).
- Fire Product Search (2018). New fire ox can do what humans can't. Available: <https://www.fireproductsearch.com/new-fire-ox-can-humans-cant/> (Accessed: 06-17-2020).
- Firehouse (2018). Product of the day: Thermite firefighting robot. Available: <https://www.firehouse.com/operations-training/hoselines-water-appliances/product/21142939/howe-howe-technologies-inc-thermite-firefighting-robot> (Accessed: 06-17-2020).
- Goecks, V. G., Woods, G., and Valasek, J. (2020). Combining visible and infrared spectrum imagery using machine learning for small unmanned aerial system detection. 11394:113940Z.
- Guillaume, E., Fateh, T., Ukleja, S., Chiva, R., and Schillinger, R. (2018). Study of fire behaviour of facade mock-ups equipped with aluminium composite material-based claddings, using intermediate-scale test method. *Fire and Materials*.
- Hopkins, M. A., Griffin, R. J., Leonessa, A., Lattimer, B. Y., and Furukawa, T. (2015). Design of a compliant bipedal walking controller for the DARPA Robotics Challenge. In *IEEE-RAS International Conference on Humanoid Robots*, pages 831–837.
- Hrabia, C.-E., Hessler, A., Xu, Y., Seibert, J., Brehmer, J., and Albayrak, S. (2019). Effeu project: Towards mission-guided application of drones in safety and security environments. *Sensors*, 19:973.
- Hubbard, B. and Hubbard, S. (2020). Unmanned Aircraft Systems (UAS) for Bridge Inspection Safety. *Drones*, 4:40.
- Lattimer, B. Y. (2020). Robotics in firefighting. Available: https://www.sfpe.org/page/FPE_ET_Issue_100/Robotics-in-Firefighting.htm (Accessed: 06-17-2020).
- Lee, S. M., Chien, J. L., Tang, E., Lee, D., Liu, J., Lim, R., and Foong, S. (2020). Hybrid kinematics modelling for an aerial robot with visual controllable fluid ejection. In *2020 IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, pages 832–838.
- Lee, T., Leoky, M., and McClamroch, N. H. (2010). Geometric tracking control of a quadrotor UAV on SE(3). In *IEEE CDC*, pages 5420–5425.
- Liljeback, P., Stavdahl, O., and Beitnes, A. (2006). Snakefighter - development of a water hydraulic fire fighting snake robot. In *2006 9th International Conference on Control, Automation, Robotics and Vision*, pages 1–6.
- London Fire Brigade (2018). London fire brigade operational response to grenfell tower. Available: <https://www.grenfelltowerinquiry.org.uk/evidence/london-fire-brigade-operational-response> (Accessed: 01-09-2020).

- McKenna, S., Jones, N., Peck, G., Dickens, K., Pawelec, W., Oradei, S., Harris, S., Stec, A., and Hull, R. (2018). Fire behaviour of modern façade materials – understanding the grenfell tower fire. *Journal of Hazardous Materials*, 368.
- Measure (2016). Drones for power plants. Available: https://www.measure.com/hubfs/PDF%20Resources/Drones_PowerPlants.pdf (Accessed: 06-17-2020).
- Merino, L., Caballero, F., Martinez-de Dios, J. R., Maza, I., and Ollero, A. (2012). An Unmanned Aircraft System for Automatic Forest Fire Monitoring and Measurement. *Journal of Intelligent and Robotic Systems*, 65:533–548.
- Minkina, W. and Dudzik, S. (2009). *Infrared thermography: errors and uncertainties*. John Wiley & Sons.
- NFPA, N. F. P. A. (2019). U.s. fire statistics.
- Nousi, P., Mademlis, I., Karakostas, I., Tefas, A., and Pitas, I. (2019). Embedded UAV Real-Time Visual Object Detection and Tracking. In *IEEE RCAR*, pages 708–713.
- Pecho, P., Magdolenová, P., and Bugaj, M. (2019). Unmanned aerial vehicle technology in the process of early fire localization of buildings. *Transportation Research Procedia*, 40:461–468.
- Peskoe-Yang, L. (2019). Paris firefighters used this remote-controlled robot to extinguish the notre dame blaze. Available: <https://spectrum.ieee.org/automaton/robotics/industrial-robots/colossus-the-firefighting-robot-that-helped-save-notre-dame> (Accessed: 06-17-2020).
- Portas, J., Bergesio, L., Campaña, I., Vaquero-Melchor, D., López-Araquistain, J., Barbolla, A., and Casar, J. (2018). Drone mission definition and implementation for automated infrastructure inspection using airborne sensors. *Sensors*, 18:1170.
- Qin, H., Cui, J. Q., Li, J., Bi, Y., Lan, M., Shan, M., Liu, W., Wang, K., Lin, F., Zhang, Y. F., and Chen, B. M. (2016). Design and implementation of an unmanned aerial vehicle for autonomous firefighting missions. In *IEEE ICCA*, pages 62–67.
- Quigley, M., Conley, K., Gerkey, B., Faust, J., Foote, T., Leibs, J., Wheeler, R., and Ng, A. Y. (2009). ROS: an open-source Robot Operating System. In *IEEE ICRA workshop on open source software*, volume 3, page 5.
- Rajappa, S., Ryll, M., Bühlhoff, H. H., and Franchi, A. (2015). Modeling, control and design optimization for a fully-actuated hexarotor aerial vehicle with tilted propellers. In *IEEE ICRA*, pages 4006–4013.
- Rangan, M. K., Rakesh, S. M., Sandeep, G. S. P., and Suttur, C. S. (2013). A computer vision based approach for detection of fire and direction control for enhanced operation of fire fighting robot. In *2013 International Conference on Control, Automation, Robotics and Embedded Systems*, pages 1–6.
- Restas, A. (2006). Forest Fire Management Supporting by UAV Based Air Reconnaissance Results of Szendro Fire Department, Hungary. In *IEEE ISEIMA*, pages 73–77.
- Ritchie, H. (2018). Causes of death. *Our World in Data*. Available: <https://ourworldindata.org/causes-of-death> (Accessed: 06-17-2020).
- Robotpompier (2018). Te800-ff technical assistance and fire-fighting robot. Available: <https://www.robotpompier.com/en/> (Accessed: 06-17-2020).
- Safetymanagement (2018). The use of robotics in firefighting. Available: <https://safetymanagement.eku.edu/blog/the-use-of-robotics-in-firefighting/> (Accessed: 06-17-2020).
- Saikin, D. A., Baca, T., Gurtner, M., and Saska, M. (2020). Wildfire fighting by unmanned aerial system exploiting its time-varying mass. *IEEE Robotics and Automation Letters*, 5(2):2674–2681.
- Schatsky, D. and Ream, J. (2016). Drones mean business. Available: https://www2.deloitte.com/content/dam/Deloitte/ie/Documents/Technology/DUP_Signals-for-Strategists_Drones-mean-business.pdf (Accessed: 06-17-2020).
- Schulte, S., Hillen, F., and Prinz, T. (2017). Analysis of Combined Uav-Based Rgb and Thermal Remote Sensing Data: a New Approach To Crowd Monitoring. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 42:347.
- ScienceBusiness (2020). Robots flex their firefighting skills. Available: <https://sciencebusiness.net/robots-flex-their-firefighting-skills> (Accessed: 06-17-2020).
- Shakhatreh, H., Sawalmeh, A. H., Al-Fuqaha, A., Dou, Z., Almaita, E., Khalil, I., Othman, N. S., Khreishah, A., and Guizani, M. (2019). Unmanned aerial vehicles (UAVs): A survey on civil applications and key research challenges. *IEEE Access*, 7:48572–48634.
- Spurny, V., Pritzl, V., Walter, V., Petrlík, M., Baca, T., Stepan, P., Zaitlik, D., and Saska, M. (2021). Autonomous Firefighting Inside Buildings by an Unmanned Aerial Vehicle. *IEEE Access*, 9:15872–15890.
- Tadokoro, Y., Ibuki, T., and Sampei, M. (2017). Maneuverability analysis of a fully-actuated hexrotor uav considering tilt angles and arrangement of rotors. *IFAC-PapersOnLine*, 50(1):8981–8986.

- The Verge (2019). Dji drones helped track and stop the notre dame fire. Available: <https://www.theverge.com/2019/4/16/18410723/notre-dame-fire-dji-drones-tracking-stopped-thermal-cameras> (Accessed: 10-06-2021).
- Tudevdaya, U., Battseren, B., and Hardt, W. (2017). Unmanned aerial vehicle based automated inspection system for high voltage transmission lines. volume 1, pages 28–32.
- USFA, U. F. A. (2018). Fire loss in the united states during 2018.
- Vidas, S., Moghadam, P., and Bosse, M. (2013). 3D Thermal Mapping of Building Interiors using an RGB-D and Thermal Camera. *IEEE ICRA*, pages 2303–2310.
- Viguria, A., Maza, I., and Ollero, A. (2010). Distributed Service-Based Cooperation in Aerial/Ground Robot Teams Applied to Fire Detection and Extinguishing Missions. *Advanced Robotics*, 24(1-2):1–23.
- Vrba, M., Heřt, D., and Saska, M. (2019). Onboard marker-less detection and localization of non-cooperating drones for their safe interception by an autonomous aerial system. *IEEE Robotics and Automation Letters*, 4(4):3402–3409.
- VT (2018). Trec: Terrestrial robotics engineering and controls. Available: <http://www.me.vt.edu/research/laboratories/trec/> (Accessed: 06-17-2020).
- WHO (2020). Global health estimates: Leading causes of death. Available: <https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghle-leading-causes-of-death> (Accessed: 10-06-2021).

How to cite this article: Walter, V., Spurný, V., Petrlík, M., Báča, T., Žaitlík, D., Demkiv, L., & Saska, M. (2022). Extinguishing real fires by fully autonomous multirotor UAVs in the MBZIRC 2020 competition. *Field Robotics*, 2, 406–436.

Publisher's Note: Field Robotics does not accept any legal responsibility for errors, omissions or claims and does not provide any warranty, express or implied, with respect to information published in this article.