DOI: 10.1109/MRA.2023.3244423

New Era in Cultural Heritage Preservation

Cooperative Aerial Autonomy for Fast Digitalization of Difficult-to-Access Interiors of Historical Monuments

Pavel Petracek[*,×], Vit Kratky[*,×], Tomas Baca[*], Matej Petrlik[*], and Martin Saska[*]

***Abstract*—Digital documentation of large interiors of historical buildings is an exhausting task since most of the areas of interest are beyond typical human reach. We advocate the use of autonomous teams of multi-rotor Unmanned Aerial Vehicles (UAVs) to speed up the documentation process by several orders of magnitude while allowing for a repeatable, accurate, and condition-independent solution capable of precise collision-free operation at great heights. The proposed multi-robot approach allows for performing tasks requiring dynamic scene illumination in large-scale real-world scenarios, a process previously applicable only in small-scale laboratory-like conditions. Extensive experimental analyses range from single-UAV imaging to specialized lighting techniques requiring accurate coordination of multiple UAV. The system's robustness is demonstrated in more than two hundred autonomous flights in fifteen historical monuments requiring superior safety while lacking access to external localization. This unique experimental campaign, cooperated with restorers and conservators, brought numerous lessons transferable to other safety-critical robotic missions in documentation and inspection tasks.**

I. Autonomous Aerial Robotics for Heritage Digitalization

Digital documentation of large interiors of historical buildings is an exhausting task since most of the areas of interest are beyond typical human reach. We advocate the use of fully-autonomous teams of cooperating multi-rotor Unmanned Aerial Vehicles (UAVs) to speed up the documentation process by several orders of magnitude while allowing for a repeatable, accurate, and condition-independent solution capable of precise collision-free operation at great heights. In particular, we present a universal autonomy for UAVs cooperating aerially within a team while documenting the interiors of historical buildings for the purposes of restoration planning and documentation works, as well as for assessing the structural state of aging historical sites. We show that the proposed approach of active multi-robot cooperation enables performing documentation tasks requiring dynamic scene illumination in large-scale real-world scenarios, a process previously applicable only manually in areas easily accessible by humans.

The presented system was developed in cooperation with cultural heritage institutions as part of the Dronument project [1] and was deployed fully autonomously in numerous characteristically diverse historical monuments, as exhibited in Fig. 1 and Table III. The included experimental evaluation utilizes UAVs in multiple real-world documentation tasks, and discusses the quality of the obtained results used in subsequent restoration works, as well as suitability of particular techniques for UAVs. The analyses demonstrate the framework's robustness in single and multi-robot deployments in more than two hundred fully-autonomous flights in fifteen historical monuments. In these experiments, the aerial robots rely solely on onboard sensors without access to external localization such as global navigation satellite systems (GNSSs) or motion capture systems, which significantly increases deployability of the system. This unique, extensive, experimental campaign, which cooperated with restorers and conservators, brought numerous lessons learned that are transferable to other safety-critical robotic missions in documentation and inspection tasks. The system also serves as a large part of an official methodological study approved by the Czech National Heritage Institute for its high added value in heritage protection. The methodology (available at [1]) describes the proper usage of UAVs in historical structures for the first time and so prescribes the proposed system to be a standard in this application.

II. Background

Often serving educational, cultural, or social purpose, the preservation of cultural heritage as a valuable reminder of our history is in the greater interest of society. Cultural management and preservation of historical monuments became a relevant topic in the late 19^{th} and 20^{th} centuries when many valuable historical monuments were destroyed while establishing modern infrastructure. By introducing cultural heritage preservation into legislation, the monuments gained protection from human interference. However, being exposed to real-world conditions continually degrades historical buildings and artifacts within. This has initiated the endeavor to actively prevent the irreversible damage of cultural heritage by monitoring its condition and performing restoration and conservation works.

Conservation work on a historical artifact comprises four consecutive phases: the initial survey, the choice of restoration steps and costs evaluation, the actual restoration works, and continued monitoring of the restoration. Both the initial survey and monitoring phase require providing information about the artifact in digital form (usually camera imaging). Thus, these phases are considered a data collection task for which an aerial vehicle, capable of gathering data in a cost-effective and fast manner, can be of great help. This is especially true for areas of interest which are located beyond typical human reach, a situation often arising in tall historical buildings such as churches and cathedrals. Apart from planning restoration works, gathered digital materials can support the reconstruc-

[*] Authors are with the Department of Cybernetics, Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic.
× *Pavel Petracek and Vit Kratky are co-first authors.*
Corresponding author: pavel.petracek@fel.cvut.cz

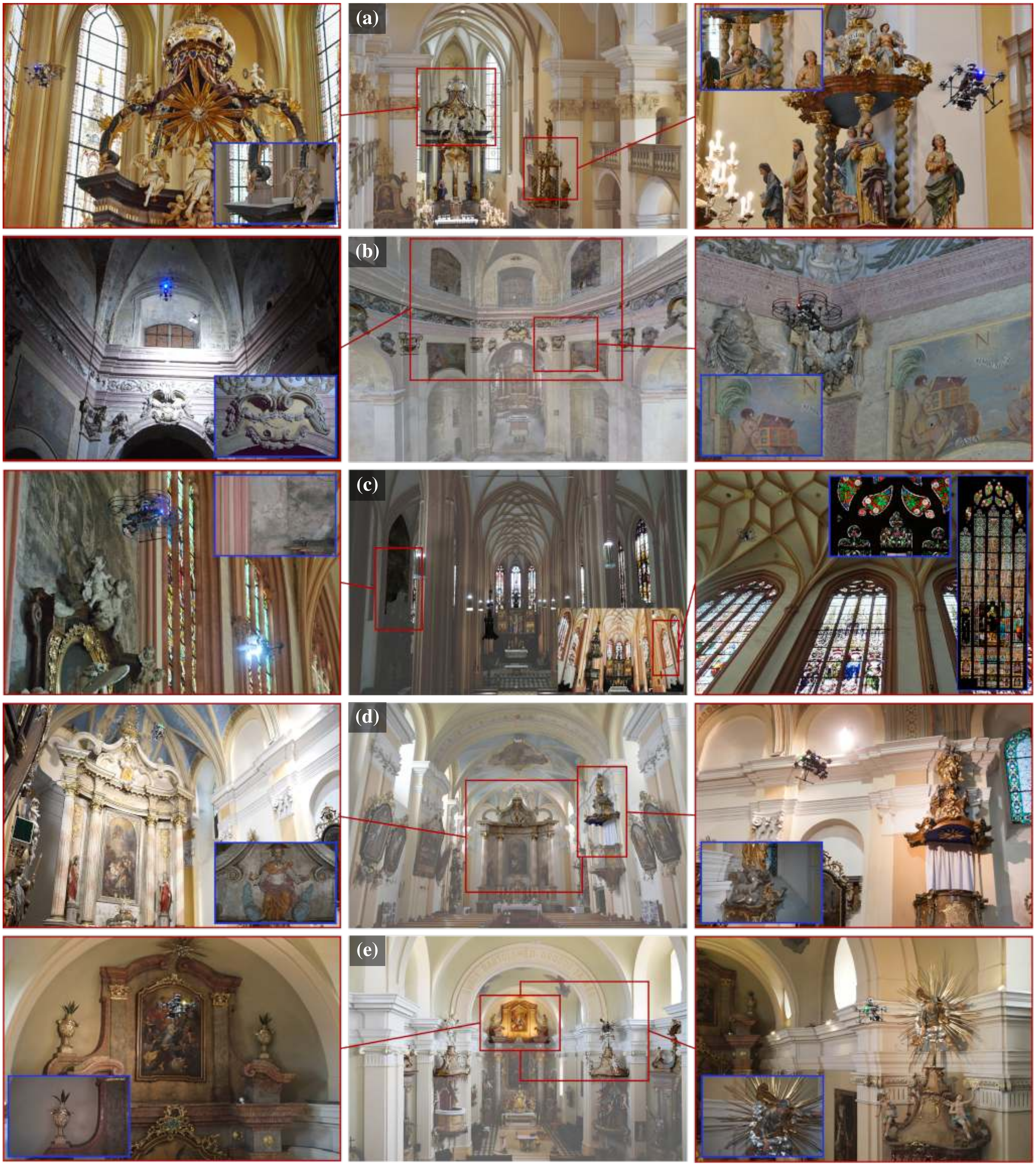


Fig. 1: Illustration of deployment of the presented methodology in selected historical buildings located in the Czech Republic — (a) Church of the Exaltation of the Holy Cross in Prostějov, (b) St. Anne and St. Jacob the Great Church in Stará Voda by Libavá, (c) Church of St. Maurice in Olomouc, (d) Church of the Nativity of the Virgin Mary in Nový Malín, and (e) Church of St. Bartholomew in Zábřeh. Center images show the interiors of the churches with highlighted objects of documentation interest. Side images show actual deployment of UAVs in the particular settings together with example images (highlighted in blue) captured by an onboard camera.

tion of a structure in the event of its sudden accidental destruction (e.g., the burning of the Notre-Dame Cathedral in 2019).

III. Robotics and Automation in Cultural Heritage Preservation

Documentation and digitalization of historical objects requires gathering various types of data, e.g., camera images in visible, infrared (IR) and ultraviolet (UV) spectra, and 3D models. The data gathering is demanding in both time and human resources, particularly in large buildings. This motivates the endeavor to automate data gathering by introducing mobile robotic solutions capable of fast autonomous documentation.The first level of mobile-robot automation can be achieved by applying Unmanned Ground Vehicles (UGVs) as carriers of the documentation sensors. A UGV equipped with a laser scanner and capable of autonomous navigation in constrained environments can sequentially visit several locations to collect a set of scans covering the entire operational space [2]. An advantage of this approach lies primarily in reducing necessary human participation in the scanning process, allowing for the collection of scans from potentially dangerous areas. Several systems applying such an approach were already developed and deployed for scanning historical monuments [3], [4].

Whereas the operational space of UGVs usually does not exceed typical human reach, multi-rotor UAVs capable of 3D navigation in confined environments can be applied for data collection tasks in difficult-to-access areas. In exteriors, UAV solutions abundantly utilize predefined GNSS poses for navigation [5]. In contrast to exteriors, the applicability of UAVs in interiors imposes additional challenges — lack of GNSS localization, navigation in a confined environment, and non-negligible aerodynamic effects. Because of that, UAV systems deployed for indoor data gathering are mainly limited to industrial inspections, with only a few works targeting UAV-based documentation of historical buildings. The specifics of such an application are targeted in this work.

For industrial inspections, the literature typically exploits the environment structure, such as known profiles of tunnels [6] or structured and well-lit warehouses [7]. More general solutions were introduced in the commercial sector introducing semi-autonomous UAV inspection systems[1]— DJI Mavic 3, Elios 3, or Skydio 2+™. In interiors, DJI provides image-based UAV stabilization, Elios allows for human-operated flight with LiDAR and camera-based stabilization and mapping with guarantees of environmental and mechanical protection, and Skydio™ offers automated camera-stabilized flight for interactive 3D reconstruction. Although all these solutions provide an assistive level of autonomy in UAV stabilization, the first two require human-in-the-loop navigation. None of the mentioned solutions offer full interior autonomy, repeatability, modularity, rotor nor sensory redundancy, imaging focusing on capturing high-quality details, and cooperative multi-robot deployment.

As mentioned, aerial data gathering inside historical buildings is rare. A specialized platform for assisting in cultural heritage monitoring called *HeritageBot* was introduced in [8]. However, no evidence of the deployment of this platform in historical monuments is presented. In [9], the authors propose an assistive system to manual control of the UAV during inspection tasks with the experimental deployment of the system inside and outside historical sites.

Among introduced solutions, the most advanced UAV-based systems with the high level of autonomy required for the interiors of historical buildings were introduced in our recent works [10]–[13]. In these publications, we introduced a preliminary application-tailored autonomous UAV system allowing for safe localization and navigation inside historical structures [10], the methodology and algorithms for the realization of advanced documentation techniques found in reflectance transformation imaging (RTI) [11] and raking light (RAK) [12], and an autonomous single-UAV system for realization of documentation missions [13]. All works provide a fully autonomous solution and the possibility of performing documentation techniques in difficult-to-access areas without using mobile lift platforms or scaffolding installation. Here, we progress beyond previous works by introducing a full 3D simultaneous localization and mapping (SLAM) methodology for indoor localization of robots; by advancing robustness to localization drifts and hard-to-detect obstacles with additional sensory redundancy; by improving path, trajectory, and mission planning; by using a UAV team to realize documentation techniques that could not be realized with only a single robot in principle; and by presenting the complete set of results achieved in the Dronument project that are summarized in numerous lessons learned during the unique experimental campaign within highly safety-critical missions.

IV. Documentation Techniques and Associated Constraints

The documentation techniques applied in the field of restoration and cultural heritage preservation aim to capture the current state of the object, survey a potential structural or artistic damage, and determine the age, author and possible dimensions of the elements by identifying the materials and techniques that have been used. For this purpose, diverse methods combining conventional photography in the visible spectrum, photography in invisible spectra making use of different reflective properties of materials, specialized lighting techniques applied for revealing structural details, and even invasive methods based on the collection of material samples are applied. In robotic context, all these methods are associated with varying requirements on sensory equipment, amount of cooperation, and external conditions (mainly illuminance). These relations are summarized in Table I, together with the studied documentation techniques.

The most common documentation technique providing initial information about the studied subject is standard visible spectrum photography (VIS). This technique is applicable to all types of studied objects, ranging from flat paintings and frescoes to 3D structures, including statues and altars. Since the documented areas of historical buildings are often dark, the obtained images suffer from insufficient lighting conditions. Hence, the VIS method often requires additional external lighting to locally increase illuminance, allowing for

[1]DJI Mavic 3: dji.com/cz/mavic-3, Elios 3: flyability.com/elios-3, Skydio 2+™: skydio.com/skydio-2-plus.

TABLE I: Recapitulative table of documentation tasks selected as realizable by aerial vehicles in interiors of historical structures. The squared check marks (☑) identify the realizable documentation methods which were experimentally applied in historical structures, as summarized in section VIII. The last column marks methods for which the ambient light is either required (✓), forbidden (✗), or arbitrary (unmarked).

			Realizable by		Required equipment and lighting conditions		
	Documentation technique		Single robot	Multiple robots	Onboard camera	Onboard light	Ambient light
Spectral analysis	visible spectrum:	photography (VIS)	☑		✓	✓	✓
		transmitography (VISTR)		✓	✓	✓	
		raking light (RAK)	☑	✓	✓	✓	
		three point lighting (TPL)		☑	✓	✓	
		reflectance transformation imaging (RTI)	☑	☑	✓	✓	✗
		light-induced luminescence (VIVL)	✓			✓	
	UV spectrum:	reflectography (UVR)	☑			✓	✗
		fluorescent photography (UVF)	☑	✓	✓	✓	
		false-color reflectography (UVRFC)	✓			✓	✗
	IR spectrum:	reflectography (IRR)	☑			✓	✗
		transmitography (IRRTR)	✓			✓	✗
		fluorescent photography (IRF)	☑	✓	✓	✓	
		false-color reflectography (IRRFC)	✓			✓	✗
	X-ray:	radiography		✓			
Others	3D reconstruction		☑	✓	✓		
	photogrammetry		✓	✓	✓		✓
	environmental monitoring		✓	✓			

the decreased exposure times required to avoid motion blur from instabilities of a multi-rotor vehicle.

Similar to aesthetic photography, light plays a significant role in restoration documentation. Documentation techniques capturing data in the visible spectrum make use of varying lighting intensity and illumination angles to enhance the quality and amount of information that can be derived from the gathered data. The main group of lighting techniques applicable during documentation tasks aims to highlight the 3D characteristics of captured objects, with three point lighting (TPL) being the most routine. TPL illuminates the object with several sources of luminance, each with different intensity and orientation with respect to the camera's optical axis, in order to provide an aesthetically pleasant and realistic view of the 3D object. Another widely used lighting technique is raking light (RAK), which focuses on revealing the surface details of flat objects. While TPL employs several light sources to avoid overshadowed areas, RAK applies a single light as parallel to the scene as possible. The illumination angle in RAK exploits the shadows to highlight the roughness of the surface.

A highly specialized documentation technique used in the field of restoration is the reflectance transformation imaging (RTI) — an image-based rendering method used for obtaining a representation of an image that enables displaying the image under an arbitrary direction of illumination. The necessary inputs of this method include a set of images of an object taken by a static camera, with each image being under illumination from a different but known direction. The captured images and the corresponding lighting vectors are then used for the computation of a polynomial texture map (PTM) representation of the image that enables an interactive illumination and view of the object. Another specialized documentation technique is visible spectrum transmitography (VISTR) which requires a light source to be positioned behind an object of interest (OoI) to transmit the light through this object. However, this method is mainly applied for canvas paintings and thus is rather impractical for realization by UAVs.

Multiple techniques exploit UV and IR lumination and its effects. While the methods based on the visible light focus on revealing structural characteristics and colors, the UV and IR methods aim primarily to identify the materials and hidden layers of artworks. The use of different spectra allows more precise dating of the paintings, as the glow of pigment combinations are unique to certain periods. The first group of methods applying UV and IR lights is based on capturing the fluorescent light in the visible spectrum emitted by an object after absorbing UV or IR radiation energy. These methods are called UV fluorescent photography (UVF) and IR fluorescent photography (IRF) and are used for, e.g., detecting zinc and titanium white (UVF) or cadmium red and Egyptian blue (IRF). The second group of methods applying UV and IR lights captures the reflected light in the corresponding spectra. These methods are called UV reflectography (UVR) and IR reflectography (IRR) and are applicable for, e.g., detecting restored areas, highlighting repairs and re-touchings, enhancing faded paintings (UVR) or reaching the underdrawing layer of paintings (IRR).

Except for VIS, all the above-mentioned methods require positioning the light at a certain angle with respect to the camera. Therefore, these methods are not fully realizable by a single UAV and require a multi-robot coordination. The particular methods can be realized in three different configurations dependent on the requirements of the task. The first configuration employs an autonomous multi-robot team consisting of a UAV carrying a documentation sensor and a set of supporting UAVs providing dynamic lighting of the documented scene. The second configuration applies the UAV as a carrier of the sensor whilst the light is provided by external sources. The third

TABLE II: Typical exposure times of the selected documentation techniques.

Technique	Spectrum	Exposure time (s)
visible spectrum photography	visible	≤ 0.2
raking light	visible	≤ 0.2
three point lighting	visible	≤ 0.2
reflectance transformation imaging	visible	≤ 0.2
UV fluorescent photography	UV	≤ 2.0
UV reflectography	UV	2.0
visible spectrum transmitography	visible	2.0
IR reflectography	IR	4.0
IR transmitography	IR	20.0
light-induced luminescence	visible	25.0
IR fluorescent photography	IR	30.0
radiography	X-ray	≥ 30.0

configuration uses the UAV for positioning the light whereas the data are captured by a static sensor from the ground.

The largest problem in realization of the techniques relying on a UAV carrying a camera is the exposure time required for sharp and detailed imaging. Table II summarizes that the exposure times for some of the methods reach tens of seconds. With constraints on image sharpness, such long times and natural *nonstaticity* of highly dynamical multi-rotor UAVs prevent the realization of these techniques in the camera-carrier mode with satisfactory results. Instead, imaging with a static camera and aerial lighting was investigated for some of these techniques.

The non-spectral tasks applied in the field of preservation mostly focus on the 3D reconstruction and environment monitoring through static sensors measuring physical quantities (e.g., temperature, humidity). The most common techniques applied in 3D reconstruction use visible spectrum images (photogrammetry) or scans produced by laser sensors. From the perspective of the proposed system, the data gathering process for 3D reconstruction does not differ from the realization of VIS and collection of raw data from onboard sensors used for localization and mapping. Monitoring the physical quantities in an environment requires attaching a sensor to the UAV frame and navigating it to the required area. If the measurement process requires permanent monitoring, the sensor must be attached at a specific position in the environment (e.g., adhered to a wall or placed on a mantel). This process is also realizable by UAVs but requires fine control, state estimation, and a mechanism for physical robot-to-environment interaction, as closely tackled in [12].

V. UAV-based Framework for Documentation of Cultural Heritage Interiors

The overall pipeline of the UAV-based framework for interior documentation in historical monuments is showcased in Fig. 2. The framework is composed of three main phases — the pre-deployment phase incorporating pre-flight data gathering and mission planning, the actual deployment of the system in interiors of historical buildings, and post-deployment phase, including processing and utilization of the collected data.

A. Pre-deployment Phase

The first step preceding the entire documentation process is obtaining a model of the environment used for safe navigation of the UAV, as well as for the specification of OoIs that should be scanned during documentation missions. For this purpose, a precise terrestrial 3D scanner Leica BLK360 is employed to obtain a set of scans that are later used for building a complete 3D representation of the target environment, both in form of a global point cloud and a 3D model with a colored texture. The colored 3D model serves for precise specification of the desired camera viewpoints and for presenting the documentation outputs to the public and the end users. The camera viewpoints specifications are made by experts of restoration or historical science who position a virtual camera within the 3D model of the environment using a viewpoint-selection tool shown in Fig. 3a. This tool shows a camera and its view and enables to save the camera viewpoint pose in the global coordinate frame. The optical properties of the camera can be parameterized with respect to the equipment available for real-world documentation, thus allowing for visualizing the desired photo to be captured from a given pose in the colored 3D model.

Given the point cloud representation of the environment and the set of to-be-captured images represented by the respective camera viewpoints in the global coordinate frame, the documentation mission plan is generated as follows. First, the problem of finding an optimal sequence σ^* of camera poses minimizing the overall traveled distance is defined and solved as the Traveling Salesman Problem (TSP). Considering the possible dimensionality of the problem, a solver using an efficient Lin-Kernighan heuristics [14] is employed for the solution of TSP to enable on-site plan generation. Constrained by available computational time, the mutual distance between particular pairs of poses within the solution of TSP are given either by the Euclidean distance or by the length of the collision-free path between the poses. Second, the consequent poses in σ^* are connected by the collision-free paths generated with the use of a grid-based planner [15]. This process creates a path connecting all the poses which can be generally unfeasible if limited flight time of a UAV is taken into account. Hence, the final set of plans $\mathbb{P} = \{\mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_n\}$ is obtained by splitting σ^* to a set of subsequences $\Sigma = \{\sigma_1, \sigma_2, \ldots, \sigma_n\}$, where $\sigma^* = \sigma_1 \cup \sigma_2 \cup \ldots \cup \sigma_n$, $\mathbf{P}_i \in \mathbb{P}$ is a collision free path connecting the initial pose with a sequence of poses in σ_i, and equation $\mathrm{t}(\mathbf{P}_i) < t_{max}$ holds, $\forall i \in \{1, \ldots, n\}$, for t_{max} being the maximum flight time of the UAV, and $\mathrm{t}(\mathbf{P}_i)$ being the time needed for following path $\mathbf{P}_i$.

To increase the mission safety, the final step of the pre-deployment phase verifies the paths planned for the documentation mission. First, each plan is verified by humans as collision-free by visualizing it in the 3D model of the environment. Second, the plan feasibility is verified by simulating the entire mission in the realistic Gazebo simulator using the virtual model of the environment with the same software and sensory plugins used during real-world missions. The goal of this two-stage process is to verify that all the generated paths are collision-free and do not traverse potentially risky parts of the environment. The mission specification and plan validation

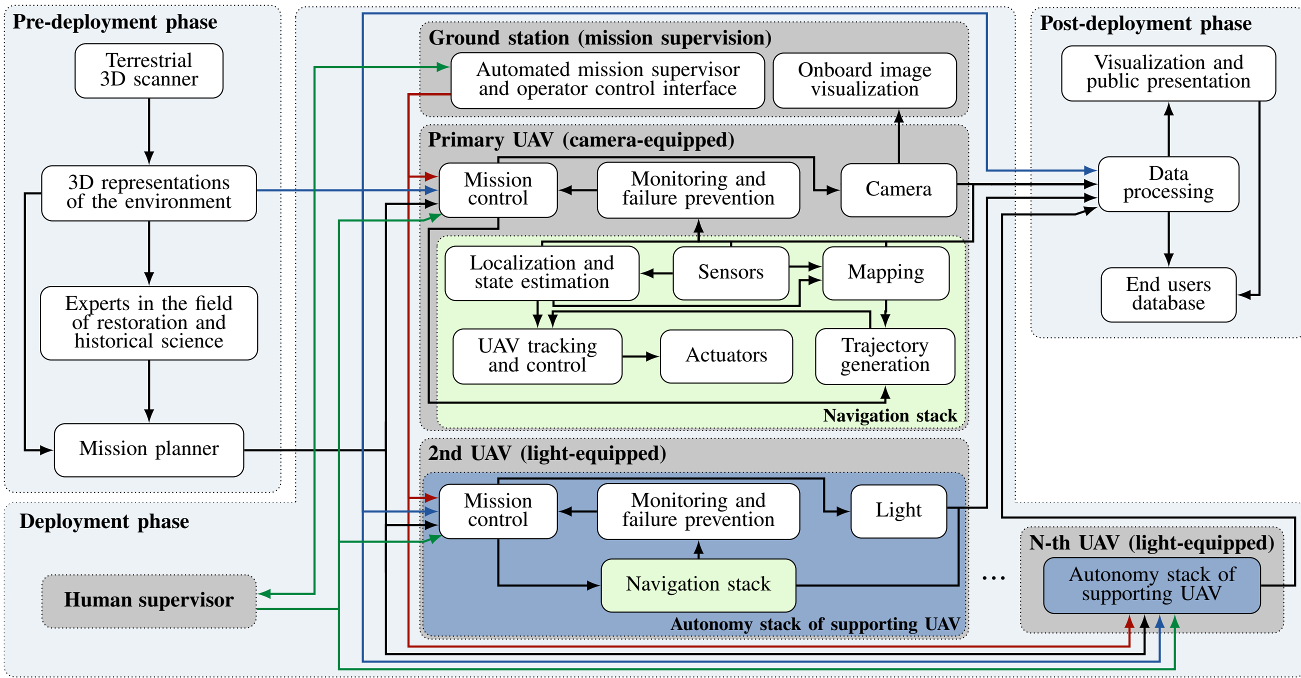


Fig. 2: High-level diagram of the three-phase architecture of the system designed for multi-UAV documentation of interiors of historical buildings. The 3D model of the environment and the mission plan are used as an a-priori generated input for the realization of the documentation mission itself. After the deployment phase, the data gathered during the mission are processed and provided to the end users.

is showcased in Fig. 3.

B. UAV Deployment

The actual system deployment is influenced by the application's specificity imposing strict safety-guarantee requirements. After the necessary hardware checks, all software components are initialized on the onboard computer of each UAV. After successful initialization, all the UAVs automatically align their reference frame with the common frame of coordination by matching their sensory data to the sparse interior map available to each UAV. The outputs of this phase are visually verified by the operator, who checks the correctness of the frames' alignment and validates the mission plan for the last time.

During the following autonomous mission, an automatic centralized supervisor (a ground station) checks the state of all the UAVs in real-time. This supervisor reacts to faults and allows for revealing many possible failures, even preventatively. Available safety actions include stopping all the airborne UAVs at the place at once, navigating them cooperatively to takeoff locations, and landing them at safe locations. Apart from the automatic supervisor, all these actions can be triggered by a human operator supervising the mission in parallel using the ground station. At last, a human operator serves as the final safety measure capable of landing the UAVs manually. The autonomy stack is described in section VI.

C. Post-deployment Phase

To increase the quality and range of the outputs, the data collected during autonomous flights in historical buildings are processed before being provided to the end users. This includes post-processing of onboard sensory data to increase accuracy of pose referencing associated with the captured data frames, stitching images into photomaps, or building a 3D model of the environment in areas occluded in ground-located scans. The generated data then serve for digitalization and archivation, pre- and post-restoration analyses, state assessment and monitoring, material analyses, photogrammetry, and for digital presentation to the public.

VI. Fully Autonomous, Cooperating UAVs

To benefit from extensively tested and field-verified methods, the proposed multi-UAV system is based on the open source MRS UAV system[2] developed within the authors' research group. In this section, let us summarize novel scientific results achieved within the presented project Dronument, whilst the MRS UAV system is described in detail in [16].

A. Reference Frame Alignment

The reference frames of the robots are aligned once during a pre-takeoff phase with each robot performing the alignment independently in four automated phases. This alignment process is mandatory for each robot as the supervising controller does not allow any robot to takeoff unless all robot frames are aligned with the global coordination frame (i.e., the map).

In the *data loading* phase, each robot loads the global map $\mathbf{M}$ and a single 3D LiDAR data-frame $\mathbf{D}$ to its memory, applies voxelization to both the objects for dimensionality reduction, and removes outliers in $\mathbf{D}$ using radius outlier filter. The z-axis of both the point clouds $\mathbf{M}$ and $\mathbf{D}$ is assumed to be

[2] github.com/ctu-mrs/mrs_uav_system

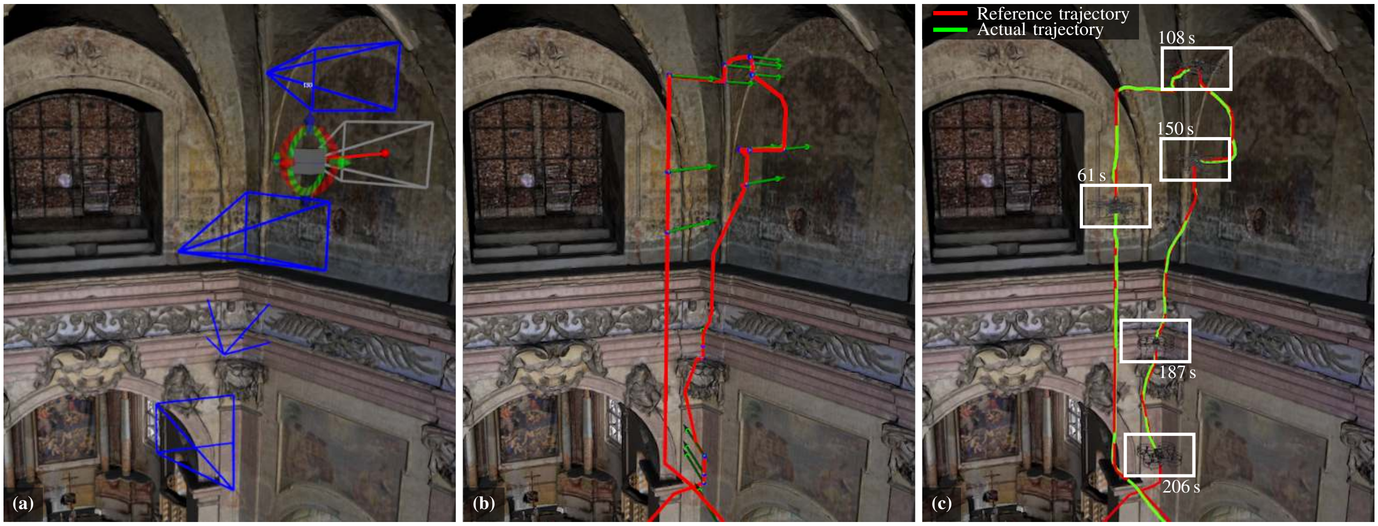


Fig. 3: Pre-deployment phase of the proposed framework — (a) specification of the documentation task by selecting a set of camera viewpoints within the 3D model of the environment, (b) planning trajectory of the robot (in red) which visits all the specified viewpoints (in green), and (c) verification of the mission plan in Gazebo simulator employing identical software that is used during real-world missions.

approximately parallel to the gravity vector. During *global correlation* phase, the origins and orientations of $\mathbf{M}$ and $\mathbf{D}$ are approximately matched. First, convex 3D-space hulls $\mathbf{H_M}$ and $\mathbf{H_D}$ are computed using Qhull [17] with a hull being represented as a set of undirected edges $\mathbf{H} = \{(\mathbf{v}_a, \mathbf{v}_b)_i\}$ (set of vertex pairs). Translation $\mathbf{t_D^M} \in \mathbb{R}^3$ of $\mathbf{D}$ to $\mathbf{M}$ is given as $\mathbf{t_D^M} = \mathbf{b_M} - \mathbf{b_D}$, where $\mathbf{b_X} \in \mathbb{R}^3$, $\mathbf{X} \in \{\mathbf{M}, \mathbf{D}\}$, represents a polyline barycenter of an edge set $\mathbf{X}$ as

$$\mathbf{b_X} = \frac{\sum_{(\mathbf{v}_a,\mathbf{v}_b)\in\mathbf{X}} \left[\mathbf{v}_a + (\mathbf{v}_b - \mathbf{v}_a)/2\right] ||\mathbf{v}_b - \mathbf{v}_a||_2}{\sum_{(\mathbf{v}_a,\mathbf{v}_b)\in\mathbf{X}} ||\mathbf{v}_b - \mathbf{v}_a||_2}. \tag{1}$$

The UAV is assumed to be taking off from ground locations, hence the grounds are coupled by setting z-axis translation to $\mathbf{t_D^M}(z) = \min_{\mathbf{p}\in\mathbf{M}} \mathbf{p}(z) - \min_{\mathbf{p}\in\mathbf{D}} \mathbf{p}(z)$, where $\mathbf{p}(z)$ denotes the z coordinate of point $\mathbf{p}$. Initial transformation to the consequent optimization phases is then given as

$$\mathbf{T_I} = \mathbf{T}(\mathbf{t_D^M})\mathbf{T}(\mathbf{t_D}, \mathbf{z}, \theta), \tag{2}$$

where $\mathbf{T} \in \mathbb{R}^{4\times4}$ is a general 3D transformation in the matrix form and $\mathbf{T}(\mathbf{t_D}, \mathbf{z}, \theta)$ is the matrix form of a z-axis rotation at a point $\mathbf{t_D} \in \mathbb{R}^3$ (the origin of $\mathbf{D}$) by angle θ. The rotation angle is given as $\theta = \theta_\mathbf{M} - \theta_\mathbf{D}$, where $\theta_\mathbf{X} = \arctan \xi_y^\mathbf{X} / \xi_x^\mathbf{X}$, $\xi^\mathbf{X} = (\xi_x^\mathbf{X}, \xi_y^\mathbf{X}, \xi_z^\mathbf{X}) = \arg\max_{\xi\in\Xi(\mathbf{X})} \sqrt{\xi_x^2 + \xi_y^2}$, and $\Xi(\mathbf{X})$ is the set of covariance matrix eigenvectors of the point cloud $\mathbf{X}$.

The following *global registration* phase copes with the lateral symmetry of the environments as typical of large historical structures. Several Iterative Closest Point (ICP) routines $\mathrm{ICP}(\mathbf{T})$ are performed in this phase, each with different initializations $\mathbf{T}$ and loosely set parameters for point association and convergence requirements. Given a number of desired initializations k, this phase selects $\theta^* = \arg\min_{\theta\in\Theta} \mathrm{ICP}(\mathbf{T_I}\mathbf{T}(\mathbf{t_D}, \mathbf{z}, \theta))$ where $\Theta = \{2\pi i/k \mid i \in \{0, 1, \ldots, k-1\}\}$. Final *fine-tuning optimization* phase estimates robot origin in the global coordinate frame $\mathbf{T_D^M}$ by running $\mathrm{ICP}(\mathbf{T_I}\mathbf{T}(\mathbf{t_D}, \mathbf{z}, \theta^*))$ optimization set with high-accuracy parameters and strict convergence criteria.

B. State Estimation, Localization, and Mapping

Estimating the 3D state of a UAV (i.e., pose and its derivatives) in real-time is crucial for the UAV mid-air control and 3D navigation. To keep the robot steady while airborne, follow reference trajectories, and avoid obstacles, the environment needs to be perceived with robot's onboard sensors (e.g., cameras, LiDARs). As state estimation, localization, and mapping are critical for collision-free flight, the utilized algorithms are based on well-tested works with implementation validated in differing real-world scenarios. To estimate the robot state, a bank of Kalman filters [16] extended with smoothing over a short past-measurements buffer fuses onboard inertial measurements with localization outputs, providing real-time feedback to the position control loop [16]. The localization and mapping systems utilize low-drift pose estimation LOAM [18]. An extensive evaluation in [15] showed that fusing LOAM efficiently with [16] provides sufficient accuracy and robustness even in safety-critical applications. The architecture of the control, state estimation, and localization pipelines is analogous to [15]. In contrast to [15], the mapping pipeline uses an a-priori map of the environment to derive a global frame for the robots' missions (see its calibration in subsection VI-A). The a-priori shared map enables multi-robot coordination and global mission planning, but also provides an additional safety level by allowing robust online analysis of localization drift and cross-checking of sensory measurements.

C. Navigation and Trajectory Tracking

The navigation of the UAVs during the mission follows a mission plan $\mathbf{P} \in \mathbb{P}$ generated in the pre-deployment phase, described in subsection V-A. This collision-free plan is represented by a sequence of triplets $\mathbf{P} = \left[(\mathbf{p}_{uav}, \mathbf{p}_{ooi}, \mathbb{I})_1, \ \ldots, \ (\mathbf{p}_{uav}, \mathbf{p}_{ooi}, \mathbb{I})_{|\mathbf{P}|}\right]$, where $\mathbb{I} \in \{0, 1\}$ is the acquisition flag. The triplets with $\mathbb{I} = 1$ specify the UAV poses $\mathbf{p}_{uav}$ in which capturing an image or illuminating the

OoI at pose $\mathbf{p}_{ooi}$ is required. The reference trajectory $\mathbf{R}$ is generated by uniform sampling of the collision-free path given as sequence of $\mathbf{p}_{uav} \in \mathbf{P}$ such that the sampling step respects the required velocity. The UAV is requested to stop at each pose $\mathbf{p}_{uav} \in \mathbf{P}$ where $\mathbb{I} = 1$ to improve the quality of data acquisition by minimizing deviation from the desired pose and reducing the motion blur that would occur in case of non-zero velocity during image capturing. The reference trajectory $\mathbf{R}$ then serves as an input to the trajectory tracking module using model predictive control (MPC). This module, described in our previous works [11], [19], produces a smooth collision-free trajectory while penalizing deviations from the original reference trajectory and respecting dynamic constraints of the UAV. The smooth-sampled reference trajectory is then passed into a feedback controller (implemented within the MRS UAV system [16]) handling tracking of the trajectory.

D. Multi-robot Coordination and Cooperation

Since the characteristics of the expected environment enable reliable use of standard communication channels, the cooperation algorithms rely on the information shared through a Wi-Fi interface among the UAVs and a ground station. Namely, the UAVs share their current poses, planned trajectories, and individual statuses based on the information from their onboard sensors. The same communication channel is utilized for commanding the UAVs from the ground station in case of emergency or a change in the mission plan, and for sharing specific messages among the UAVs during the realization of cooperative documentation techniques. The algorithms handling the autonomous flight are computed on board the UAVs.

During the cooperation, the reference trajectories of the UAVs are generated in a distributed manner on a short horizon corresponding to the optimization horizon used in the MPC-based trajectory tracking module [16]. By applying concepts of leader-follower architectures, the reference trajectories of supporting UAVs are generated with respect to the optimized trajectory of the primary UAV (leader), to the position and the desired distance of the UAV from the OoI, and to the desired lighting angle with respect to the optical axis of the documentation sensor on board the primary UAV. The coordination of the UAVs is part of trajectory optimization (see subsection VI-C) where both the current poses of the UAVs and their planned trajectories are considered to be part of constrained unfeasible space [19]. To prevent the downwash effect, this optimization is also constrained to not allow two nearby UAVs to fly above each other.

VII. Aerial Platforms

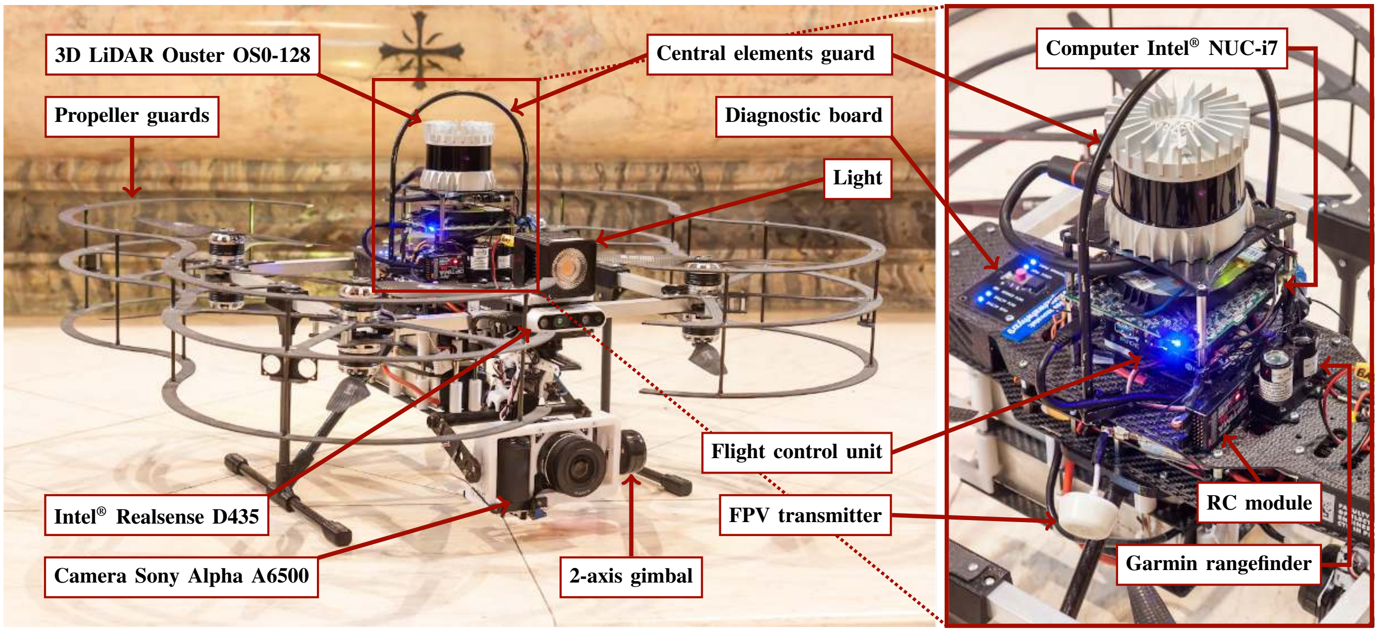
Two custom-made UAV platforms were designed specifically for the proposed application of deployment in interiors of buildings. Both the platforms, as shown in Fig. 4 and described in more detail in [20], support fully autonomous deployment within the tackled domain by carrying sensors for local environment perception together with a powerful computational unit handling the entire autonomous aerial mission. The primary platform is a heavy-weight (5.5 kg without payload) octo-rotor with dimensions of $78 \times 81 \times 40$ cm, capable of carrying up to 1.5 kg payload — enough for a mirrorless interchangeable-lens (MIL) camera with a suitable lens and 2-axis gimbal stabilization, as well as an onboard light source. This platform minimizes its dimensions while maximizing the payload capacities, is equipped with mechanical propeller guards, and carries sensory redundancy for active obstacle avoidance. The secondary platform is a lightweight (3 kg fully loaded) quad-rotor with dimensions of $68 \times 68 \times 30$ cm suited for assisting the primary UAV throughout a documentation process by providing the scene illumination, thus increasing the quality of the gathered digital materials. While cooperating, the supporting UAVs assist in performing tasks inexecutable by a single UAV in principle. As the primary payload, the secondary platform carries a set of high-power light sources. Both the platforms support flights in close proximity to obstacles and to other UAVs. However, relative distances are limited to a minimum of 2 m to limit the aerodynamic influence of downwash, ceiling, and ground effects on the UAV, and the contrary effect of the UAV on the environment (possible damage of not firmly attached objects and fragile plasters).

A. Sensors for Autonomy

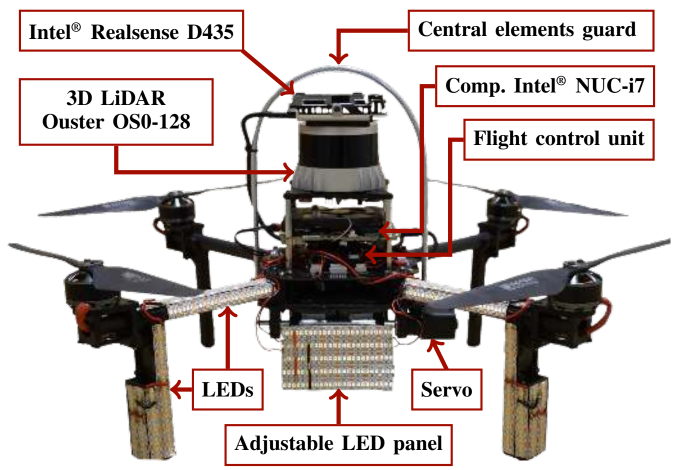
For autonomy in GNSS-denied environments, both platforms rely on onboard sensors only. The primary sensor is a 3D light detection and ranging (LiDAR) Ouster OS0-128 with 50 m detection range and 90° vertical FOV supported by a thermally-stabilized triple-redundancy inertial measurement unit, downward and upward looking point-distance sensors Garmin LiDAR Lite, and front-facing (primary UAV) or downward and upward-facing (secondary UAV) color-depth cameras Intel® Realsense D435 for sensory cross-checking in active obstacle avoidance. All the sensory data are processed by an Intel® NUC-i7 onboard computer which utilizes data in real-time algorithms handling the autonomous aerial mission. The low-level control (attitude stabilization) is handled by Pixhawk 2.1, an open-source autopilot used frequently by the robotic community. For safety reasons, the primary UAV carries a visible diagnostic RGB LED which indicates a possible failure to an operator who is authorized to override UAV autonomy for manual landing.

B. Payload

The payload equipment mountable on board the platforms is modular — cameras, lenses, and light sources can be easily interchanged for the purposes of a specific task. For general purposes, the primary UAV carries a 2-axis gimbal FlyDrotec capable of stabilizing up to 850 g payload. The stabilized axes are controllable, a feature useful mainly for controlling the pitch angle of a camera. Throughout our experiments, a MIL camera, the Sony Alpha A6500 with varying lenses, has been used for its integrated image-sensor stabilization, further minimizing the negative effect of mid-flight vibrations on the output image quality. Triggering image capture is automated via the onboard computer, whereas real-time imaging is transmitted to the ground for online visualization for the operator.



(a) Primary custom-made UAV application-tailored for documentation and inspection tasks in building interiors. The platform carries onboard sensors required for autonomous flight with equipment for acquiring high-quality documentation data, as well as a processing unit for handling autonomous flight, reasoning over the sensory data, obstacle avoidance, and the documentation mission.



(b) Secondary UAV tailored for supporting documentation tasks in building interiors. In contrast to the primary UAV (a), this platform is smaller and carries a high-power light instead of sensors for the documentation task.

(c) Comparison of the custom-made UAV platform (a) with lightweight commercial drone DJI Mavic Air 2, which carries a small camera sensor and does not support complex mission planning in building interiors.

Fig. 4: Aerial platforms used for documentation tasks in the Dronument project — primary UAV carrying documentation sensors (a), secondary UAV assisting in cooperative documentation (b), and commercial drone used for qualitative comparison (c). Both (a) and (b) carry environment-perception sensors and computational resources allowing fully autonomous deployment in interiors with poor lighting conditions.

VIII. Experiments and Results

The extreme requirements on safety imposed by the nature of the application requiring the deployment of UAVs in priceless historical buildings imply thorough validation of all the developed software and hardware solutions prior to their deployment in real-world missions. The software solutions ranging from the state estimation and control algorithms to high-level mission control were intensively tested with the use of Gazebo simulator and the MRS simulation package[3] providing realistic behavior of the UAVs. Running the same software with identical parametrization in simulation and on real hardware significantly simplifies the transfer of algorithms from the virtual environment to real-world applications. The 3D models built from the obtained 3D scans are directly used as the simulation environments for algorithms' testing. Together with simulated sensory noises and model inaccuracies, this makes the simulation as analogous to real-world conditions as possible. This methodology proves to be especially useful for discovering possible failures correlated with specific environments and validating the entire autonomous missions in an approximate copy of the real-world scenarios. Although the simulator is highly realistic, running the system in the real world introduces additional constraints. Therefore, even after thorough testing in virtual environments, the first deployments of the system

[3] github.com/ctu-mrs/simulation

TABLE III: Overview of three-phase UAV deployment in historical buildings within the Dronument project. The first phase focused on specifying the use cases, developing the methodology, designing the system, and performing preliminary experiments, including manually controlled flights. The second phase investigated autonomous multi-robot coordination in cooperative documentation and experimented with imaging outside the visible spectrum and with the physical interaction of UAVs with the environment. The third phase deployed the system in a full-operation mode for gathering data valuable to end users and for validating the methodology and overall performance of the autonomy.

Object (approximate floor area over which the system operated)	flights	images	flight time (h:mm:ss)	flight dist. (m)	maximum height (m)	min. obst. dist. (m)	multi-robot	applied methods	
Archbishop's Chateau in Kroměříž (UNESCO, 420 m^2)	10	202	0:24:30	320	9.0	2.0	✗	VIS	
Vranov nad Dyjí State Chateau (410 m^2)	18	1049	1:26:44	1020	12.0	3.0	✗	VIS	
Klein Family Mausoleum in Sobotín (30 m^2)	7	274	0:17:37	120	3.8	1.6	✗	VIS	**Phase 1 2017–2019**
Rondel at State Chateau and Castle Jindřichův Hradec (140 m^2)	8	660	0:51:32	940	7.8	2.1	✗	VIS	
Chapel of All Saints at Chateau Telč (UNESCO, 84 m^2)	6	190	0:14:04	145	4.8	1.7	✗	VIS	
Church of St. Mary Magdalene in Chlumín (224 m^2)	8	86	0:23:40	146	4.8	1.5	✓	RTI	
Church of the Holy Trinity in Běhařovice (252 m^2)	6	56	0:07:10	30	5.2	2.2	✗	IRF, UVF, IRR	**Phase 2 2019–2021**
Church of St. Maurice in Olomouc (1160 m^2)	27	971	1:31:38	1340	16.8	1.5	✓	VIS, TPL, RAK	
Church of St. Anne and St. Jacob the Great in Stará Voda (505 m^2)	95	6022	4:54:10	4540	19.5	1.7	✓	VIS, TPL, RTI RAK, UVR, IRR	
Church of the Exaltation of the Holy Cross in Prostějov (570 m^2)	7	548	0:19:49	308	15.2	1.7	✗	VIS	
Church of Our Lady of the Snows in Olomouc (918 m^2)	3	255	0:10:06	185	17.1	2.4	✗	VIS	**Phase 3 2021–2022**
Church of the Assumption of the Virgin Mary in Cholina (409 m^2)	3	82	0:08:51	132	7.5	1.4	✗	VIS	
Church of the Nativity of the Virgin Mary in Nový Malín (282 m^2)	2	129	0:06:06	68	8.8	1.8	✗	VIS	
Church of the Holy Trinity in Kopřivná (367 m^2)	4	211	0:17:50	247	11.4	2.0	✗	VIS	
Church of St. Bartholomew in Zábřeh (616 m^2)	4	263	0:18:23	258	12.4	1.4	✗	VIS	
Total (6387 m^2)	208	10998	11:32:10	9799	19.5	1.4	✓	VIS, TPL, RTI RAK, UVR, IRR UVF, IRF	

were preceded by test flights in mock-up scenarios and testing interiors in order to reveal potential problems related to transfer of the system from simulation to real hardware.

The final version of the system, as presented in this manuscript, builds on preliminary versions and architectures of both software and hardware stacks and integrates experience from over a year and a half period of experimental deployments. During the experimental campaigns, remaining sources of potential failures were identified and the UAV system upgraded to reach the desired performance and reliability while increasing the number of realizable documentation techniques. The entire system was, to this day, deployed in real-world missions in fifteen historical buildings of various characteristics (summarized in Table III), including one of the largest Baroque halls in the Czech Republic at State Chateau Vranov nad Dyjí and the UNESCO World Heritage Sites, Archbishop's Chateau in Kroměříž and Chateau Telč. Almost twelve airborne hours in more than two hundred flights have been performed for purposes of documentation missions in the given structures. Such an extensive experimental campaign provides an exhaustive validation of the system in real-world conditions and supports its applicability in GNSS-denied environments by identifying and overcoming challenges imposed by specific scenarios. The following sections describe the documentation techniques realized by the system in these structures. The OoIs of the presented documentation missions are showcased in Fig. 1.

A. Visible Spectrum Photography

Imaging in the visible spectrum is the most frequently applied technique as it includes methods providing the widest range of practical information while being relatively easy to perform. Within the fifteen historical structures, OoIs of various characteristics have been imaged by autonomous UAVs. These OoIs range from artistic elements, such as paintings, stained-glass windows, mosaics, stuccoes, and murals located in the most upper parts of the main naves, to complex 3D structures, such as window frames and altars up to 20 m high.

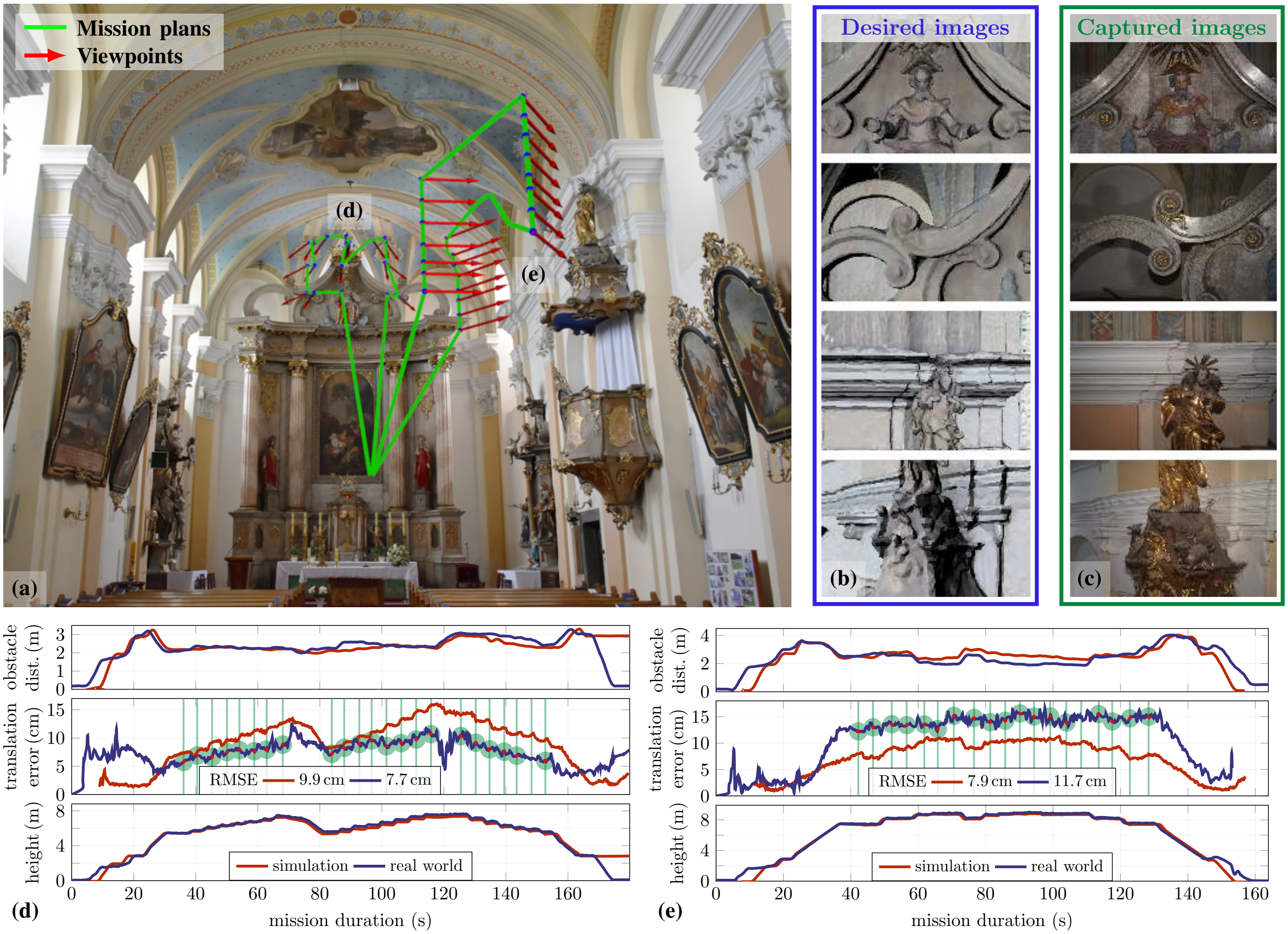


Fig. 5: An example documentation mission in the Church of the Nativity of the Virgin Mary in Nový Malín. The documentation mission was divided into two separate flights (a) focused on documenting the upper part of the altar (d) and baldachin of the pulpit (e). When going from the top, the rows in (d) and (e) show the minimal distance from the UAV frame to an obstacle, the 3D position error with image acquisition times (green vertical lines and red dots), and the height above ground. The desired imagery specified in the 3D model and the images captured on board are compared in (b) and (c). The simulation data are averaged from 5 runs each.

Additionally, objects may include structural damage, such as crevices, cracks, or fractures.

An example of fully-autonomous documentation of a single interior is provided in Fig. 5 depicting the documentation of a baroque church. The specified viewpoints were focused on documentation of two OoIs — the upper part of the altar reaching a height of 10 m and the baldachin of the pulpit. The automated process of viewpoints' specification and autonomous navigation has enabled fast realization of the documentation process in just two single-UAV flights lasting only 366 s in total. With mission specification being part of the pre-deployment phase, the overall time required for in-site deployment reached only 80 min, including equipment unpacking, flight test, mission validation and execution, and packing. Such a high level of autonomy in the process demonstrates superiority via fast, safe, effective, and repeatable data capturing when compared to the slow, imprecise, and dangerous manual control of the UAV in obstacle-filled environments by even a highly trained human operator. Even with assistive systems (stabilization and collision prevention) guiding the human in navigation, manual operation is unsafe in losses of line of sight in the presence of obstacles and inefficient in time and accuracy required to reach the desired viewpoints. Apart from higher efficiency and safety of autonomy in contrast to human-controlled flying, a fully autonomous system allows flight in close proximity to obstacles, enlarging the operational space of the UAV. This is advantageous particularly when documenting elevated OoIs where the inaccuracy in estimating the UAV's distance to the ceiling is proportional to the distance from the human eye, thus making manual navigation in these areas unsafe.

The VIS method can be performed with commercially available products (e.g., DJI Mavic) offering semi-autonomous solutions in small and lightweight packages. However, the limited level of autonomy and sensory modularity makes the realization of the missions in large interiors prolonged (the proposed system is on average ten times faster in the same task), non-repeatable, or even impossible in conditions unfavorable to onboard perception or the desired documentation technique. In Fig. 6, the images obtained by the proposed

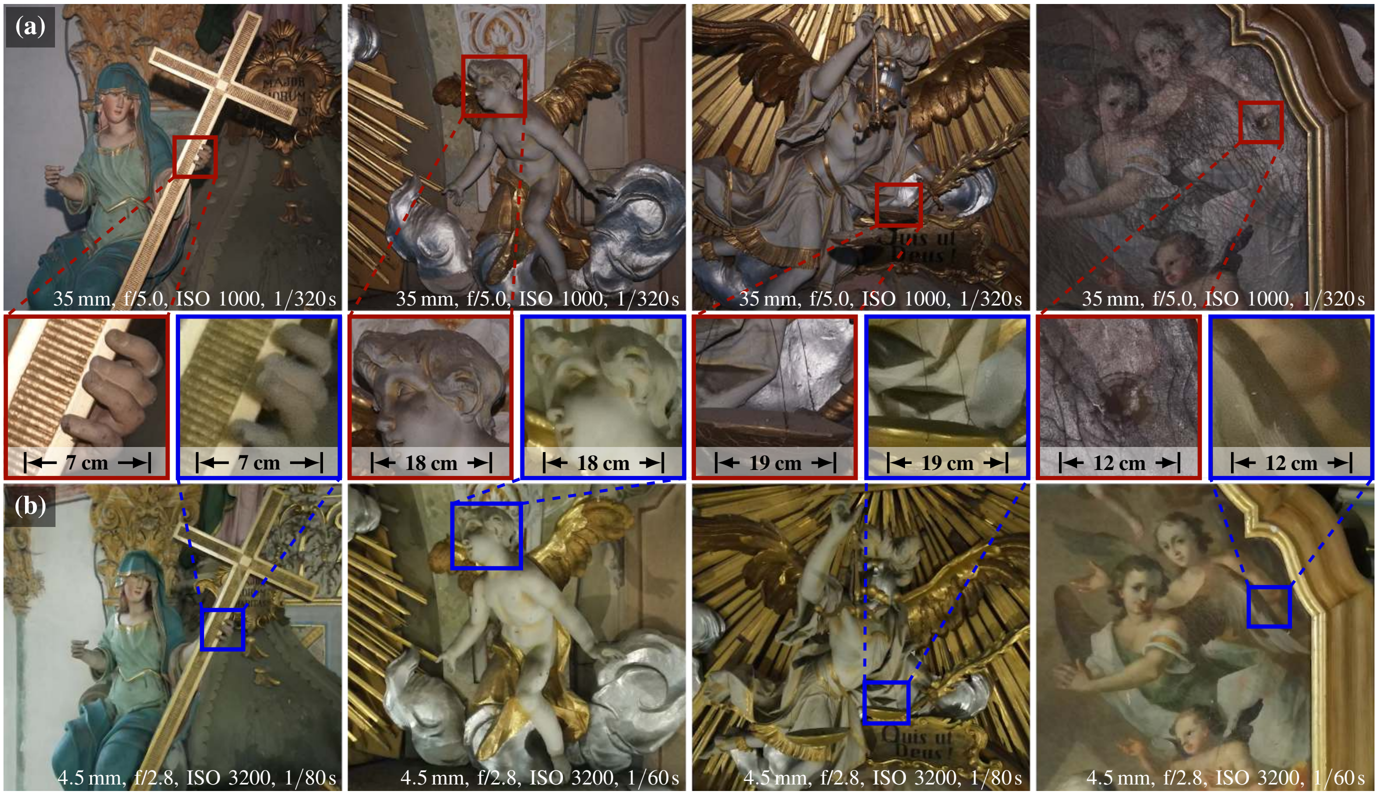


Fig. 6: Image outputs of VIS methodology as taken by the onboard MIL camera Sony Alpha A6500 (a) and commercial solution DJI Mavic Air 2 (b). Direct comparison of details of the images in the middle row shows that the proposed solution is superior in capturing high-quality details. This highlights the last column in which a hole in the painting is visible in top and absent in bottom image. Although the commercial solution is small and lightweight, its small sensor size of 6.4×4.8 mm hinders usability in interior documentation.

system are qualitatively compared to the ones obtained with a commercial product DJI Mavic Air 2. The figure highlights the superior performance of MIL camera imaging allowing for capturing high-resolution details of the OoIs while maintaining a safer distance from the obstacles.

Although VIS realized by a single UAV is a powerful technique, a multi-robot approach is often unavoidable if the lighting conditions are insufficient or documentation of an OoI requires non-direct lighting. An example OoI requiring additional lighting is the mural of St. Christopher in the late Gothic Church of St. Maurice in Olomouc, the documentation of which is shown in Fig. 7. Insufficient external lighting on the mural did not allow capturing bright, high-quality images without the motion blur effect arising from deviations in the reference pose over a long exposure time. Thus, to improve the quality of the images, a secondary UAV provides side lighting (approximately 45° with respect to the camera optical axis), lowering exposure times and highlighting details on the mural, such as small crevices invisible to the human eye from the ground. In the same church, 23 stained-glass windows (each about 8–34 m^2 large) were able to be documented with a single UAV as the windows were well illuminated by the outdoor light and could be captured with short exposure times without additional lighting. The individual images of the mural and the stained-glass windows were rectified and stitched together to compose singular high-resolution orthophotos of each object. The orthophotos were used to assess the state of the OoIs for subsequent restoration works and for enhancing the texture of the 3D model of the church[4]. As compared well in [13], the aerial-based orthophotos outperform the ground-based orthophotos in terms of quality of detail, quality of rectification due to perpendicular optical angles, and absence of occlusions.

B. *Reflectance Transformation Imaging*

RTI method requires a static camera and a dynamic light with a known history of poses. To validate whether the proposed system is feasible for RTI, it was applied to document a vault located 11 m above ground in St. Anne and St. Jacob the Great Church in Stará Voda (see Fig. 1b). This OoI was specifically selected as it can be photographed from a balcony on the opposite side of the central nave, thus allowing for the realization of the RTI technique in two comparable configurations: 1) with the camera (with telephoto lens) mounted on a static tripod with a clear, but misaligned view on the vault and 2) with the camera mounted on board the primary UAV. In both configurations, the light was carried on board the secondary UAV, with the directions of illumination being derived from the poses of this UAV, as estimated on board during the flight.

The comparison of results obtained in each configuration is presented in Fig. 8. The image representation produced from images captured by the tripod-mounted camera yields

[4]Selected OoIs and mapping and 3D reconstruction examples of documented historical structures can be found at mrs.felk.cvut.cz/3d-model-viewer.

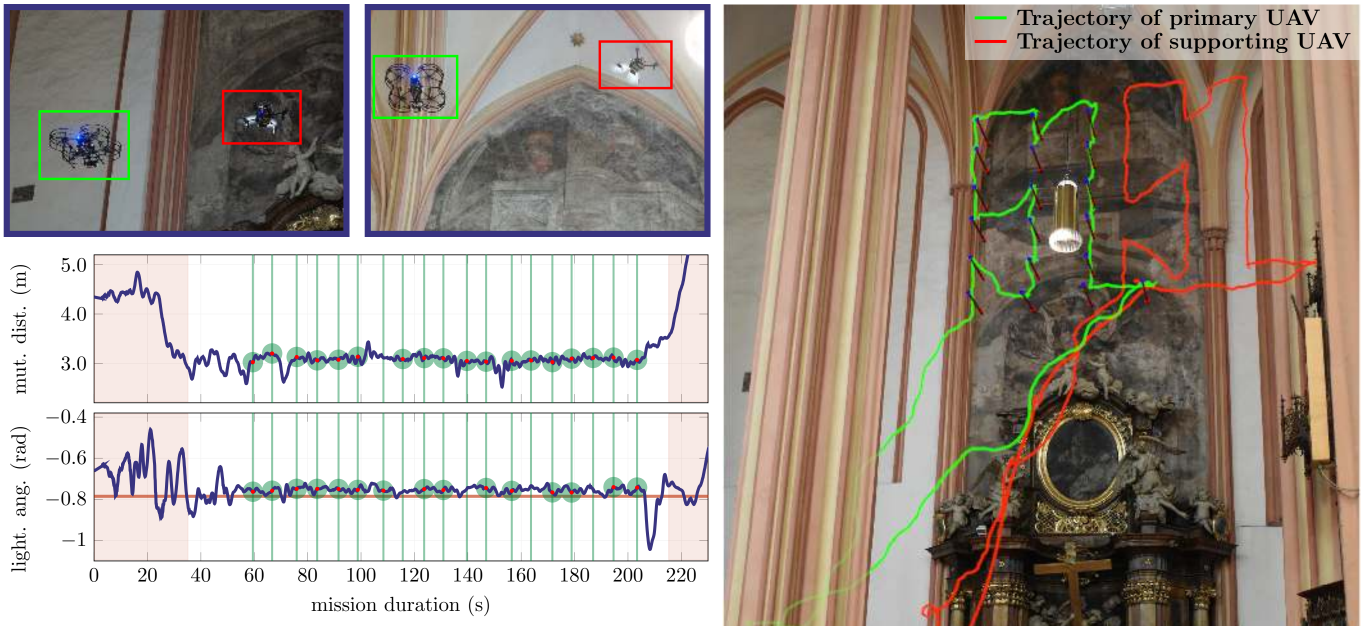


Fig. 7: Deployment of a multi-robot formation for detailed documentation of the late Gothic mural of St. Christopher using additional lighting for enhancing the quality of gathered data. The graphs show mutual distance of the UAVs during the cooperative flight and the angle between the camera optical axis and light, together with the time occasions of image capturing (green lines and circles). The red horizontal line denotes the required angle of lighting. The red areas mark parts of the mission in which the UAVs are not required to maintain the formation.

higher quality, as the choice of the OoI and usage of a telephoto lens fully compensate for the main disadvantages of the methodology in this particular case. These disadvantages are primarily smaller operational space, lower detail resolution of the resulting image caused by the large distance of the camera from the OoI, and often unavoidable occlusions. Although the fully UAV-based approach yields lower image quality since the camera's pose is not static over time, it has wider operational space and enables imaging from appropriate angles, as was verified for other OoIs in the church that could not be reasonably captured by a static camera at all. The non-staticity of the camera's reference pose misaligns the images; thus, their sub-pixel post-alignment is required to avoid blur in the resulting PTM. The experiment shows that the fully UAV-based approach yields comparable results to the single-UAV approach, which is favorable when the OoI can be photographed from the ground — an impossible scenario for most OoIs in difficult-to-reach areas of historical buildings.

C. Raking Light and Environmental Monitoring

A common feature of raking light documentation and monitoring of environmental conditions with UAVs stands in the need for robot-environment interaction. In the former, a light is attached to the wall illuminating a planar OoI from a direction perpendicular to the optical axis of the camera. This method is known to highlight even the smallest crevices and cracks in the planar surface. For the latter, a wireless sensor (e.g., for measuring humidity or temperature) is attached to the wall to measure the environmental conditions over longer periods of time. For the purpose of physical environment-UAV interaction itself, we researched a UAV equipped with a system for admittance-based control allowing for stabilization while being attached to a planar surface (and possibly interacting with it) [12]. Before using this technology, the involved risks must be compared to the payoff, particularly inside historical buildings. To minimize the risks, it is more convenient to interact with structural (not artistic) parts of the buildings. The system was successfully tested in real-world mock-up scenarios (see Fig. 9c) with walls of sufficiently good condition.

D. IR and UV Photography

Realization of UVF and IRF (fluorescent photography) is methodically similar to VIS with the equipment being a standard MIL camera and a source of light at appropriate frequency. In contrast to VIS, the light emitted by the object illuminated by an IR or UV light source in the visible spectrum is lower. Thus, these methods require higher exposure times, as specified in Table II. The higher exposure times put stricter requirements on image stabilization in the presence of onboard vibrations, inaccuracies, and disturbances that cause UAVs to deviate from their reference pose.

Realization of the UV and IR reflectography requires a camera without UV and IR filters and exposure times of tens of seconds. This makes the use of UAVs for imaging in UV and IR reflectography unfeasible. However, supporting ground-based imaging with aerial lighting is applicable. The UAVs can carry (relatively close to the OoI) high-power LEDs radiating in the desired spectrum. The IR and UV-based methods were tested in St. Anne and St. Jacob the Great Church, Stará Voda (see Fig. 9) and in Church of the Holy Trinity, Běhařovice. The experiments showed that the proposed system can be used in realization of the UV and IR-based methods in historical structures, even in limited lighting conditions.

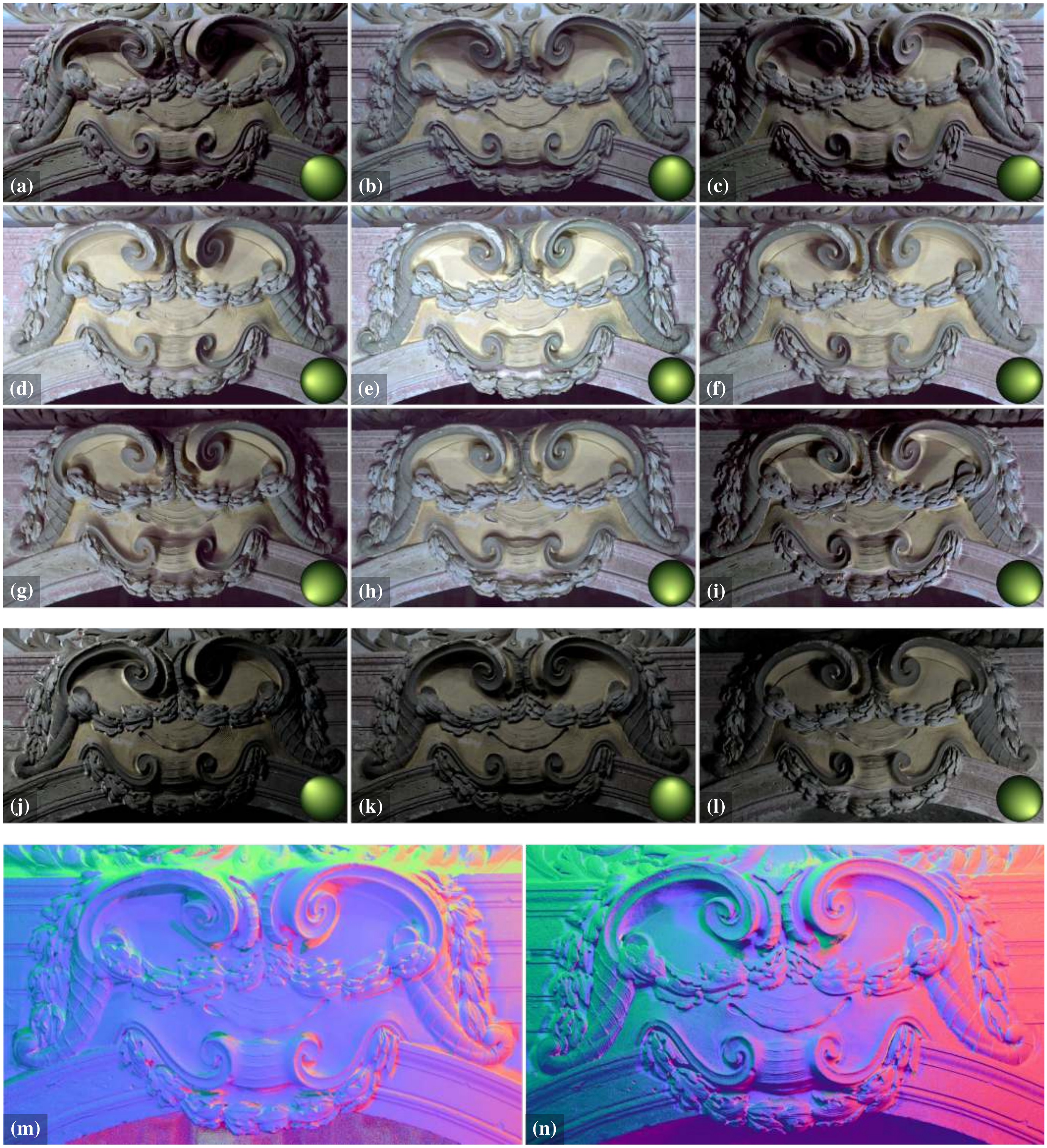


Fig. 8: Comparison of polynomial texture maps (PTM) obtained with a fully UAV-based RTI approach with camera carried by a UAV (a)–(i) and PTM obtained from images taken by a camera mounted on a static tripod (j)–(l). In both cases, the dynamic positioning of light is provided by the secondary UAV. The bottom row shows the normal maps encoded in RGB for fully UAV-based approach (m) and a single UAV approach (n).

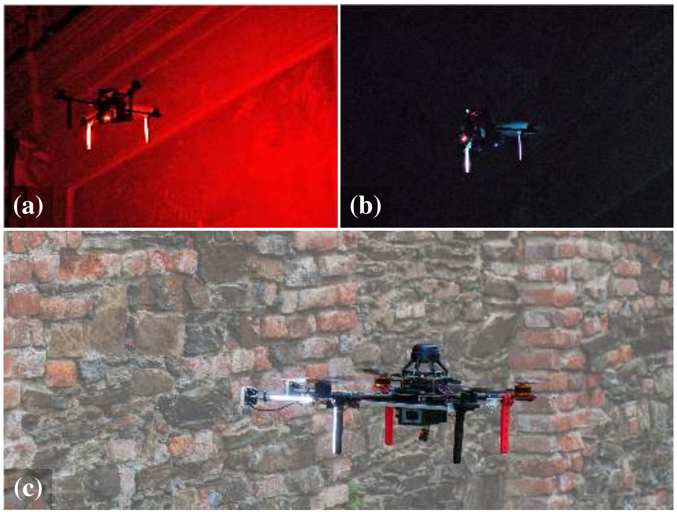


Fig. 9: Deployment of UAVs carrying IR (a) and UV (b) source of light, and a frame-extension mechanism for physical attachment and interaction with static planar surfaces (c).

E. Mapping and 3D Reconstruction

The capacities of UAVs allow capturing the interior under difficult-to-reach angles, not only for imaging purposes, but also for spatial mapping of the structures. Although terrestrial laser scanners yield the most accurate maps, these devices cannot, in principle, document occluded spaces, whereas the larger operational space of UAVs allows for minimizing these occlusions. This advantage is showcased in Fig. 10 where above-ledge areas could not be reconstructed from scans captured at ground. The potential for accurate 3D mapping using UAVs is immense; however, is not the main purpose of the proposed system which outputs dense 3D maps only as a byproduct to the photo-documentation task. The onboard-UAV-built maps contain larger amounts of noise as the mobile laser-scanning technology is less accurate (lightweight, low-power, and moving while scanning) than static scanners, making it harder to align the captured scans, even in post-processing. To achieve the best results for 3D reconstruction, we recommend leveraging the advantages of both methodologies simultaneously.

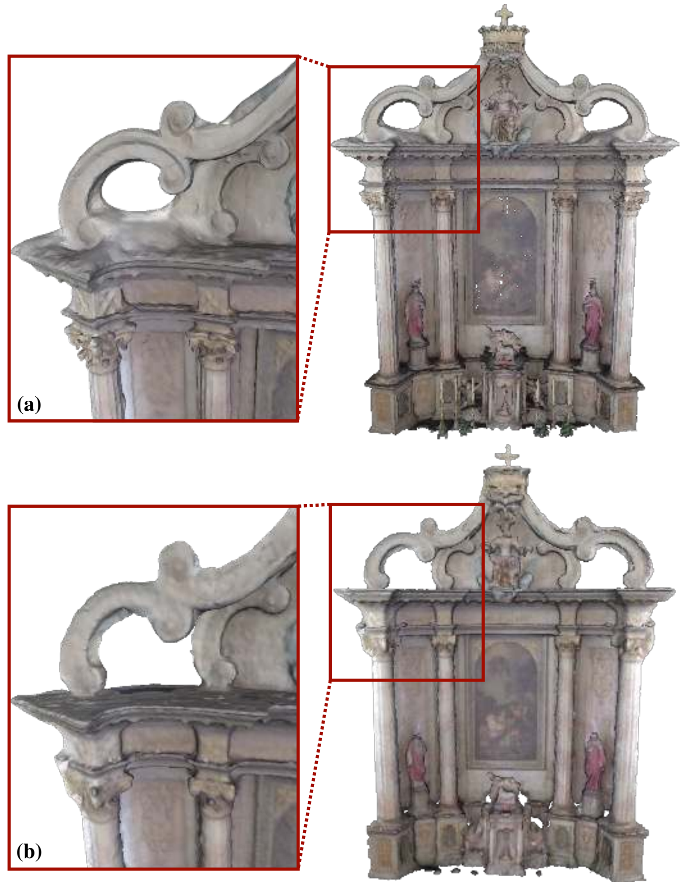


Fig. 10: 3D reconstruction of the altar at the Church of the Nativity of the Virgin Mary in Nový Malín, Czech Republic. The altar reconstructions were done using scans obtained by (a) terrestrial laser scanner Leica BLK360 and (b) Ouster OS0-128 mounted on board an autonomous UAV during the deployment shown in Fig. 5. The meshes were created with the Poisson surface reconstruction and colored using the panoramic RGB images captured by the terrestrial scanner.

IX. Discussion

The proposed UAV-based system for documenting historical monuments of differing structures, dimensions, and complexity has demonstrated its wide applicability in real-world documentation tasks, ranging from RGB photography and 3D mapping to multi-robot RTI in areas high above the ground. The high level of autonomy, the ability to fly beyond the visual line of sight between the UAV and a human operator, and the deployability in low lighting conditions (using a worldwide unique method of dynamic illumination by a cooperating UAV team) enable to gather crucial data for heritage protection and documentation that was not possible before. This universally novel system has been used in the very first fully-autonomous multi-robot real-world deployments in such complex and safety-demanding interior structures.

However, deploying mobile robots inherently poses risks to the environment, humans, and equipment therein. This requires careful justification of the UAVs' use that, in our experience, tends to be needlessly overused — conventional technology provides a safer and better quality solution in many documentation tasks. A common example is imaging the interior ceiling or low-height OoIs, where using a static camera with a long-focus lens was identified to be a more appropriate solution. Manual-control UAV solutions are also sufficient if the task is small-scale, the lighting conditions are feasible, repeatability is not required, and the OoIs are few. The need for multi-UAV teams in tasks achievable with sufficient quality by a single UAV, such as the selected example of single-UAV RTI presented in Fig. 8, should also be considered prior a full-scale deployment.

X. Conclusion

This work has presented a universally novel study on an autonomous multi-robot UAV-based system for realization of advanced documentation techniques in culturally valuable environments. The system showcases the immense potential of mobile robots for fast, accurate, and mobile digitalization of difficult-to-access interiors. The hardware and software architectures of the self-contained autonomous-UAV-based system were introduced and experimentally validated through almost twelve hours of flight time in more than two hundred real-world flights of single-UAVs and multi-UAV teams in fifteen historical monuments of varying structures. The system design has emerged from close cooperation with a team of restorers, and the data collected during the autonomous missions has been used by the end users in successive restoration works.

The study also assists in identifying the current challenges and future directions of research in aerial documentation and inspection. Based on the high added value for heritage protection, the system has been approved by the Czech National Heritage Institute for indoor usage and is accompanied by an official methodology (available at [1]) describing the proper usage of UAVs in historical structures. It is the first methodology of this authority for using UAVs in historical buildings and so prescribes the system to be a standard in this application.

XI. Acknowledgment

This work was supported by the Ministry of Culture of the Czech Republic through project no. DG18P02OVV069 in program NAKI II, by the Ministry of Education of the Czech Republic through OP VVV funded project CZ.02.1.01/0.0/0.0/16 019/0000765 "Research Center for Informatics," by the European Union's Horizon 2020 research and innovation program AERIAL-CORE under grant agreement no. 871479, by CTU grant no. SGS20/174/OHK3/3T/13, and by the Czech Science Foundation (GAČR) under research project No. 20-10280S.

The authors would like to thank Nicolas Staub for an initial survey of the projects' applicability in documentation and restoration, Jan Bednar for his assistance in creating 3D models for public presentation, Pavel Stoudek for preparing hardware platforms used during the experimental analyses, and Vojtech Krajicek for technical consultation of the restoration techniques. At last, we would like to thank to representatives of the Czech National Heritage Institute, namely Milan Skobrtal and Michaela Cadilova, for our fruitful cooperation during the project and for the opportunity to deploy the proposed system in real-world structures.

References

[1] Multi-Robot Systems group, CTU FEE. (2022) Dronument (Dron & Dronument). [Online]. Available: http://mrs.felk.cvut.cz/dronument

[2] S. Prieto, B. Quintana, A. Adan, and A. Vazquez, "As-is building-structure reconstruction from a probabilistic next best scan approach," *Robotics and Autonomous Systems*, vol. 94, pp. 186–207, 2017.

[3] P. S. Blaer and P. K. Allen, "Data acquisition and view planning for 3-D modeling tasks," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2007, pp. 417–422.

[4] D. Borrmann, R. Heß, H. R. Houshiar, D. Eck, K. Schilling, and A. Nüchter, "Robotic mapping of cultural heritage sites," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XL-5/W4, pp. 9–16, 2015.

[5] T. Bakirman, B. Bayram, B. Akpinar, M. F. Karabulut, O. C. Bayrak, A. Yigitoglu, and D. Z. Seker, "Implementation of ultra-light uav systems for cultural heritage documentation," *Journal of Cultural Heritage*, vol. 44, pp. 174–184, 2020.

[6] T. Özaslan, G. Loianno, J. Keller, C. J. Taylor, V. Kumar, J. M. Wozencraft, and T. Hood, "Autonomous navigation and mapping for inspection of penstocks and tunnels with mavs," *IEEE Robotics and Automation Letters*, vol. 2, no. 3, pp. 1740–1747, 2017.

[7] M. Beul, D. Droeschel, M. Nieuwenhuisen, J. Quenzel, S. Houben, and S. Behnke, "Fast autonomous flight in warehouses for inventory applications," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3121–3128, 2018.

[8] M. Ceccarelli, D. Cafolla, G. Carbone, M. Russo, M. Cigola, L. J. Senatore, A. Gallozzi, R. Di Maccio, F. Ferrante, F. Bolici, S. Supino, N. Colella, M. Bianchi, C. Intrisano, G. Recinto, A. Micheli, D. Vistocco, M. R. Nuccio, and M. Porcelli, "HeritageBot Service Robot assisting in Cultural Heritage," in *IEEE International Conference on Robotic Computing*, 2017, pp. 440–445.

[9] N. Hallermann, G. Morgenthal, and V. Rodehorst, "Vision-based monitoring of heritage monuments — Unmanned Aerial Systems (UAS) for detailed inspection and high-accurate survey of structures," *WIT Transactions on The Built Environment*, vol. 153, pp. 621–632, 2015.

[10] P. Petracek, V. Kratky, and M. Saska, "Dronument: System for Reliable Deployment of Micro Aerial Vehicles in Dark Areas of Large Historical Monuments," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2078–2085, 2020.

[11] V. Kratky, P. Petracek, V. Spurny, and M. Saska, "Autonomous Reflectance Transformation Imaging by a Team of Unmanned Aerial Vehicles," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2302–2309, 2020.

[12] D. Smrcka, T. Baca, T. Nascimento, and M. Saska, "Admittance Force-Based UAV-Wall Stabilization and Press Exertion for Documentation and Inspection of Historical Buildings," in *International Conference on Unmanned Aircraft Systems*, 2021, pp. 552–559.

[13] V. Kratky, P. Petracek, T. Nascimento, M. Cadilova, M. Skobrtal, P. Stoudek, and M. Saska, "Safe Documentation of Historical Monuments by an Autonomous Unmanned Aerial Vehicle," *ISPRS International Journal of Geo-Information*, vol. 10, no. 11, pp. 738/1–16, 2021.

[14] S. Lin and B. W. Kernighan, "An effective heuristic algorithm for the traveling-salesman problem," *Operations Research*, vol. 21, no. 2, pp. 498–516, 1973.

[15] V. Kratky, P. Petracek, T. Baca, and M. Saska, "An autonomous unmanned aerial vehicle system for fast exploration of large complex indoor environments," *Journal of Field Robotics*, vol. 38, no. 8, pp. 1036–1058, 2021.

[16] T. Baca, M. Petrlik, M. Vrba, V. Spurny, R. Penicka, D. Hert, and M. Saska, "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*, vol. 102, no. 26, pp. 1–28, 2021.

[17] C. B. Barber, D. P. Dobkin, and H. Huhdanpaa, "The Quickhull Algorithm for Convex Hulls," *ACM Trans. Math. Softw.*, vol. 22, no. 4, pp. 469–483, 1996.

[18] J. Zhang and S. Singh, "Low-drift and real-time lidar odometry and mapping," *Autonomous Robots*, vol. 41, no. 2, pp. 401–416, 2017.

[19] M. Saska, V. Kratky, V. Spurny, and T. Baca, "Documentation of dark areas of large historical buildings by a formation of unmanned aerial vehicles using model predictive control," in *IEEE International Conference on Emerging Technologies and Factory Automation*, 2017, pp. 1–8.

[20] D. Hert, T. Baca, P. Petracek, V. Kratky, V. Spurny, M. Petrlik, M. Vrba, D. Zaitlik, P. Stoudek, V. Walter, P. Stepan, J. Horyna, V. Pritzl, G. Silano, D. Bonilla Licea, P. Stibinger, R. Penicka, T. Nascimento, and M. Saska, "MRS Modular UAV Hardware Platforms for Supporting Research in Real-World Outdoor and Indoor Environments," in *IEEE International Conference on Unmanned Aircraft Systems*, 2022, pp. 1264–1273.