IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

Monitoring Ancient Buildings: Real Deployment of an IoT System enhanced by UAVs and Virtual Reality

MANLIO BACCO[1], PAOLO BARSOCCHI[1], PIETRO CASSARÁ[1], DANILA GERMANESE[1], ALBERTO GOTTA[1], GIUSEPPE RICCARDO LEONE[1], DAVIDE MORONI[1], MARIA ANTONIETTA PASCALI[1] and MARCO TAMPUCCI[1]

[1]Institute of Information Science and Technologies (ISTI), National Research Council (CNR), Pisa, Italy

corresponding author: Pietro Cassará, e-mail: pietro.cassara@isti.cnr.it

ABSTRACT The historical buildings of a nation are the tangible signs of its history and culture. Their preservation deserves considerable attention, being of primary importance from a historical, cultural, and economic point of view. Having a scalable and reliable monitoring system plays an important role in the Structural Health Monitoring (SHM): therefore, this paper proposes an Internet of Things (IoT) architecture for a remote monitoring system that is able to integrate, through the Virtual Reality (VR) paradigm, the environmental and mechanical data acquired by a wireless sensor network set on three ancient buildings with the images and context information acquired by an Unmanned Aerial Vehicle (UAV). Moreover, the information provided by the UAV allows to promptly inspect the critical structural damage, such as the patterns of cracks in the structural components of the building being monitored. Our approach opens new scenarios to support SHM activities, because an operator can interact with real-time data retrieved from a Wireless Sensor Network (WSN) by means of the VR environment.

INDEX TERMS Internet of Things, Structural Health Monitoring, Virtual Reality, Unmanned Aerial Vehicles, Condition Monitoring

I. INTRODUCTION

Cultural Heritage (CH) offers value and attractiveness to the cities and places of a nation. Therefore, the exploitation and preservation of it is of utmost interest. SHM aims at providing asset managers with actionable information [1] to make informed decisions. It can provide damage detection and characterization, risk assessment, or can be used to hypothesize future performance. Recently, the use of digital sensors has gradually replaced analog and/or mechanical devices in SHM systems; moreover, the advent of the IoT paradigm has brought interworking capabilities, opening to the use of sensor networks for the remote monitoring of monuments and/or structures of great impact [2]. Finally, information systems allow for processing the data produced by these sensors, whether they are images, vibration, acceleration data, environmental parameters, thus opening to the techniques able to provide complex aggregate information, alarms, or actuation without direct human intervention. Nowadays, fixed installations only are concerned, temporary or definitive [2], [3], which allow for a level of monitoring and implementation that is limited to what assumed at the design stage. The advent of UAVs is reversing this logic and lays the foundation for a dynamic concept of remote monitoring of the infrastructure. As a matter of fact, it is now conceivable to perform a survey even on parts of structures where no fixed detection systems were provided, or to integrate an existing monitoring system with other information captured by an UAV flying on the structure of interest. The use of UAVs for photogrammetry activities is already widely adopted today [4], [5] and has led to a significant reduction in costs and to an enhancement of services offered in the application field, ranging from video shooting to goods delivery [6], [7]. In the field of SHM systems, a joint use of fixed sensor networks, mobile sensors (e.g. cameras and Light Detection and Ranging (LIDAR) radars) on board UAVs, and VR techniques provides an additional powerful system to support operators in the field.

In this work, an UAV was used to obtain a 3D reconstruction of a structure with the exact position of the sensors deployed on itself, in order to allow the operator to dynamically

interact with the sensors in a VR environment: monitoring data retrieved by the sensor network can be visualized in real time, in order to infer the structure health. Such a system opens new scenarios for SHM and for new applications of Artificial Intelligence (AI) techniques to automatically detect anomalies and generate alarms or simulate a forecast in the VR environment. As already mentioned, the UAV payload was also used for sensing: the UAV camera is used to inspect the crack patterns of a monitored structure with high measurement accuracy. The camera calibration is obtained by detecting the markers posed aside the crack and by estimating the 3D position.

This article analyses and describes the architecture of a remote monitoring system, herein referred to as the *MOSCARDO system*, which integrates a network of fixed sensors and an UAV. The proposed system allows to interact on a 3D graphic model of the monitored structure with the relative sensing information, showing the location of sensors on the structure and displaying instantaneous and/or historical records in a VR environment. Section II provides a survey of the literature, highlighting how the proposed system is in the vanguard of the SHM domain. Section III describes the use cases and the real testbed that was performed. Section IV describes the IoT architecture of the proposed system and the real case studies, while Section IV-B presents the advanced monitoring capabilities based on the use of an UAV. Then, the methods used for image processing and VR, as well as their validation, are presented in Section V. Section VI highlights the current and future role of the proposed monitoring system for decision makers, beyond pure numerical results. Eventually, the conclusions can be read in Section VII.

II. STATE OF THE ART

Nowadays, SHM is being increasingly used in practical applications as, for instance, the monitoring of CH described in this work, or to perform predictive maintenance activities. SHM can be seen as an enabler of the objectives of the 11th Sustainable Development Goal[1], which advocates to *strengthen efforts to protect and safeguard the world's cultural and natural heritage*. More generally, SHM is one of the bricks to build the so-called *smart cities*, where an increasing degree of intelligence and automation will bring *urban intelligence* into play. For instance, such an idea is in [10], which discusses smart cities and the necessity to preserve and revitalize the past, as for instance in the case of CH. Along with this, the authors argue that livability, revitalization, and sustainability are needed, because of the increasing number of people leaving rural areas in favor of cities. In order to aim towards a smart city targeting those values, the authors propose an IoT architecture composed of different layers for sensing, interconnecting, handling data, and providing services at the very top. The case study under consideration sees the city of Trento, Italy, benefiting of a context-aware recommendation system for tourism and CH, namely Treesight. Another core activity in this context is environmental monitoring [16], [17], because the preservation of CH may exploit environmental data, in order to enrich data coming from typical SHM deployments.

Focusing on systems built to meet the necessities of SHM explicitly, WSNs stand out as a well suited solution to deploy sensor nodes in an economical manner for this purpose [18]. In general, the idea is relying on simple and small sensor nodes to collect raw data to be delivered to a control center for analysis purposes. From a networking viewpoint, single and multi-hop networks can be used towards this aim, as well as gateways or other related network devices used for the purposes of synchronization, computing and storing, and other coordination activities. According to the type of data to be collected, small or large amount of them can be expected, with a direct repercussion on the expected lifetime of sensor nodes [19], especially if battery powered. Different approaches for setting up low-cost monitoring systems are discussed in [20]–[23]. Reference [20] describes an SHM system exploiting accelerometers in smartphones to monitor structural vibrations. In other words, the authors propose to rely on widely diffused devices as smartphones to detect and measure sinusoidal vibrations, so to build a citizens' network for crowdsourcing data for analytics purposes. The presented results, similarly to those in [21], can be considered as relevant in this field, when taking into account the low cost of smartphones with respect to high-end sensor nodes typically used in such monitoring scenarios. Another low-cost technique is explored in [22], based on the use of moiré patterns: two fully overlapped circular fringes are mounted on both sides of a crack in a wall. Then, pictures taken with a smartphones can be collected for processing by means of a mobile app developed for such a purpose: in fact, any displacement causes the two fringes to misalign, which is a phenomenon that can be captured with a phone camera. The authors compare the results achieved with this technique with those achievable with digital micrometers, eventually estimating the measurement accuracy to be $\pm 5 \times 10^{-2}$mm.

WSNs used for SHM should satisfy requirements such as: network scalability, high synchronization, optimal sensor placement, energy efficiency, and clustering. According to the authors in [18], several research issues are still open, as for instance: damage prediction, energy harvesting, the use of mobile phones as sensor nodes, and large amount of data that can put network infrastructures under pressure. The focus on using WSNs, more specifically on IoT as a paradigm for data collecting and data analysis for SHM, is discussed in [24]. The authors survey Low-Power Wide Area Network (LPWAN) in the context of SHM, considering, for instance, the use of Long Range (LoRa) protocol to deliver small amount of data in an energy efficient manner, as proposed in this work. The monitoring of buildings relies on the use of lightweight, reliable, and open network stacks, enabling interoperability and lowering integration costs. On this topic, the use of Message Queuing Telemetry Transport (MQTT) or

[1]The Sustainable Development Goals can be read here: sustainabledevelopment.un.org

SHM system	year	cloud/edge computing	wireless sensor networks			unmanned vehicles	energy harvesting	network scalability	high synchronization	data analysis		VR/AR
			environ. data	*mechanical*	*images*					*Big Data/ML*	*event-driven*	
[8]	2008				✓						✓	
[3]	2011	✓	✓	✓					✓		✓	
[9]	2013		✓	✓					✓		✓	
[10]	2015	✓	✓					✓	✓	✓		
[11]	2015									✓		
[2]	2017	✓	✓	✓				✓	✓		✓	
[12]	2017				✓					✓		
[13]	2017				✓	✓					✓	
[14]	2018				✓	✓				✓		
[15]	2019	✓			✓	✓		✓	✓	✓		
MOSCARDO	2019	✓	✓	✓	✓	✓	✓	✓	✓		✓	VR

TABLE 1: Comparison among implemented functionalities in relevant monitoring systems for ancient buildings in the literature and the one proposed in this work, namely MOSCARDO.

Constrained Application Protocol (CoAP) protocols is a de facto standard [25] because they can provide lightweight solutions for IoT data exchanges in an energy efficient manner. The Publish / Subscribe (PUB/SUB) paradigm, connecting data publishers to data consumers through a rendezvous entity, allows flexible and customisable implementations.

Looking to closely related works, the ones in [2], [3] can be referenced. In [3], the authors describe the design and deployment of a SHM system based on the use of Memsic Imote2 nodes to collect acceleration data, temperature, humidity, and light level. Sixteen sensor nodes have been deployed in a basilica, aiming at monitoring the building that suffered severe damage in the 2009's earthquake in central Italy. Similarly to this work, the authors exploit a WSN to collect data at a local sink, then delivered to a remote center via a 3G connection. In a similar fashion, reference [2] describes an IoT-based system to monitor a bell tower. The authors analyze the impact of large moving crowds on the structure, during a large festival hosted in the city of Lucca, Italy; thanks to this and to the oscillations due to the bell, it was possible to provide a characterization of the structure.

Two other cases of interest are here analysed: the use of UAVs in this context, and the increasing role played by Machine Learning (ML)-based techniques.

ML-based techniques have been used to provide preliminary assessments of potential upcoming dangerous situations that may be recognised by jointly analysing heterogeneous quantities and exploiting historical data. For instance, in [11] the authors discuss the use of data clustering and Support Vector Machines (SVMs) to classify the data streams coming from a wired SHM system. The aim is in providing a hazard management method related to earthquake disaster. The authors modify the data classifiers in order to support online training. Data are collected from accelerometers, labelled in an unsupervised way, then classified, clustered, and used to train the SHM process. The output is binary: a structure has been damaged (or not) by an earthquake. In [12], Convolutional Neural Networks (CNNs) are used to assess the damage to bridges in China. Real-time data are fed to a CNN, aiming at classifying possible bridge cracks, in order to reduce the so-called *risk coefficients*. The authors are convinced that the method proposed in [12] could effectively solve the problems of low fracture diagnosis efficiency and high risk factor, thus proposing intelligent systems for bridge safety. Bridges are monitored also in [9], which anyway focuses on the performance of the wireless system rather than results related to SHM. In [14], the authors show that the integration of CNN and other classifiers can improve detection accuracy for images of masonry structures. The training phase is performed on a dataset containing images of cracks from masonry structures (created using a digital camera and an UAV), in order to extract significant features. Other classifiers, such as SVMs, are used to enhance the classification ability, reaching a detection accuracy of approximately 86% in the validation stages and 74% in the testing stage.

The role played by aerial vehicles is rapidly increasing in this domain, seeing the use of Micro Air Vehicles (MAVs) and UAVs [26] to overcome the drawbacks of manual visual inspection. In addition, the use of on-board digital image processing can specifically address the characterization of crack patterns. In [13], the authors describe a processing technique to measure cracks thicker than 0.1 mm with the maximum length estimation error of 7.3% on concrete surface; this approach is also complemented with a CNN, to improve its feasibility with respect to real world conditions (e.g. [27], [28]). In [15], a system for infrastructure monitoring based on the use of multiple UAVs is proposed: the paper is focused, mostly, on the real-time path planning and on the introduction of a classical yet effective image segmentation method based on iterative thresholding for crack detection and quantification. Moving from the domain of civil buildings and infrastructures to the domain of CH, specific features and constraints must be considered: for instance, the physical and chemical properties of the material inspected can significantly vary, even in the same structure; the condition in which data are collected may be challenging (with area difficult to access, or variable environmental conditions); the methods used for acquiring data and monitoring relevant parameters should be not invasive. In the CH domain, most of the methods using MAVs or UAVs for the data acquisition, and aiming at the detection and characterization of the crack pattern, are based on 2D image processing and digital photogrammetry. In [8], close range photogrammetry is applied to the analysis of structural damage detected in a masonry structure (Basilica da Ascension, north-west of

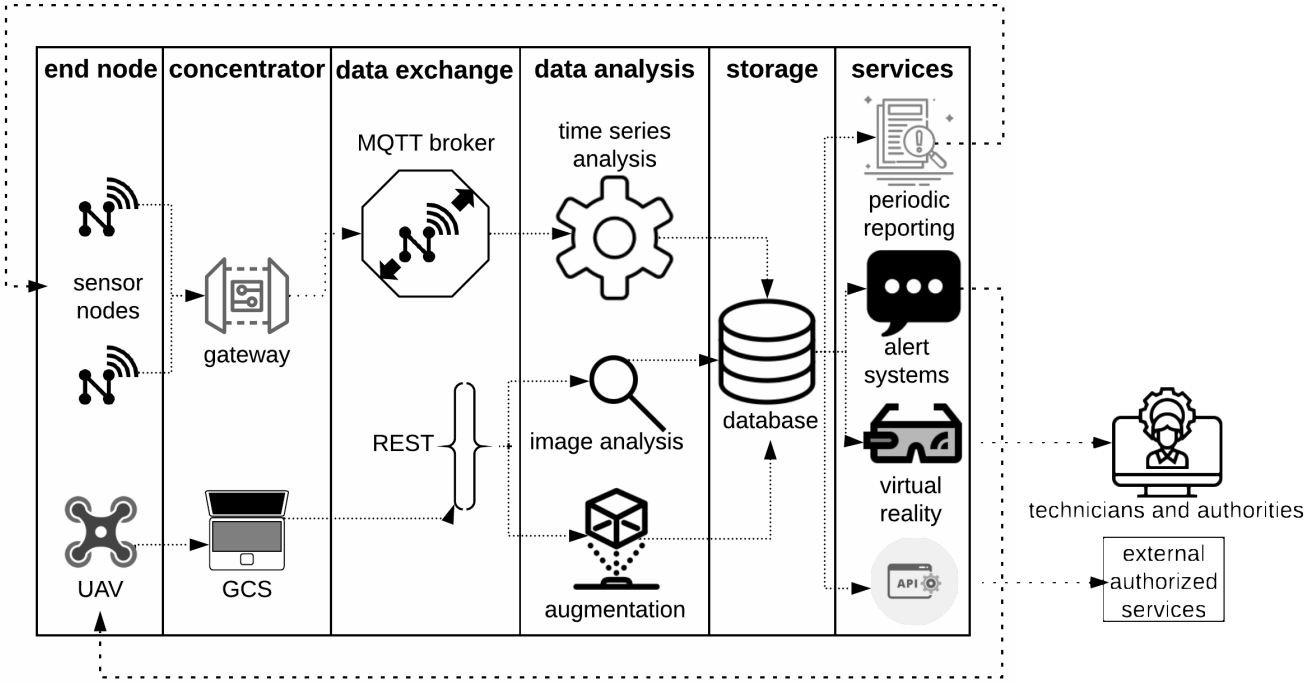


FIGURE 1: The proposed IoT system, namely MOSCARDO, for monitoring ancient buildings via a modular IoT-based system: sources of raw data on the left, data exchange and data analysis in the middle, available services on the right.

Spain). 3D point clouds of the monitored damage (a crack) are gathered at different epochs; shape parameters of each point cloud are extracted and used to assess the damage and track its evolution. In [29], the authors perform a stereo-photogrammetric scanning of the Bridge of the Towers, Italy, by means of an UAV. A large amount of high-resolution photos were analysed and processed in order to reconstruct the 3D model of the bridge. Starting from the 3D model, the crack pattern is monitored, by mapping the damage onto two orthophotos, which represent the two sides of the bridge. Other remote techniques have been used in this context, further than UAVs. For instance, in [30], the authors survey the use of Global Navigation Satellite System (GNSS)-based technology, like multi-sensor and multi-constellation data acquisition techniques, to support monitoring activities for towers, buildings, and bridges. Six GNSS systems are considered in the paper to be used jointly with Real Time Kinematic (RTK) techniques to improve the achievable accuracy in monitoring activities on the ground. Anyway, it should be noted that GNSS technology is intrinsically limited by multipath errors and low sample rates, among others. Thus, GNSS systems coupled with accelerometer units close to the monitored buildings provide an increased accuracy, as in the case of Robotic Total Station (RTS) integration with GNSS. Pseudolites are considered in [30] as well, to enhance satellite-based systems, which are anyway able to identify modal frequencies up to 10 Hz, enough to accurately determine cable forces of bridges from GNSS data.

As shown in Table 1, the system proposed in this work covers almost all the selected features, thus pushing further both the level of complexity and integration that a monitoring system provides. In particular, our implementation allows interacting with the monitored buildings and the deployed sensor nodes by means of a VR environment. That allows for easily visualising both current and historical readings, further than knowing the exact location of the monitoring points all over the buildings under consideration. Thanks to that, remote technicians can also access information in real-time, especially during or immediately after events of interest.

III. THE PROPOSED MONITORING SYSTEM AND THE THREE REAL CASE STUDIES

The whole system is configured as shown in Figure 1: data collected by sensor nodes and by the UAV is sent to the remote server for data processing and storing. On the one hand, data is pushed from sensor nodes towards a local gateway, hosting simple services: that allows to ensure the correct functioning of the local network at each site and to take care of the actual data delivery towards the remote server. Therefore it avoids that intermittent connections force sensor nodes to stay awake longer than necessary. On the other hand, the data collected by the UAV is sent to a close Ground Control Station (GCS), acting similarly to the aforementioned local gateway. The latter relies on the use of the MQTT protocol, to reliably deliver collected data in a PUB/SUB fashion: the local gateway hosts, among others, a MQTT broker configured as a MQTT bridge, connected to the remote gateway, which hosts a MQTT broker as well, feeding Fast Fourier

Transform (FFT) of data to the remote database. Thus, from the local to the remote gateway, PUB/SUB data exchanges occur in a reliable way, delivering data in one direction (from sensor nodes to the remote server) and commands in the other one (from the remote server to the sensor nodes). The other logical data link, from the UAV to the remote server, is based on the use of a Representational State Transfer (REST)-ful Application Programming Interface (API), transferring data to a remote image analysis service. Both results are stored into the remote database for providing historical and real-time data querying services, aiming at feeding analysis and alert systems and at comparing fresh data with historical ones. The main source of information for building periodic reports is the WSN, while the UAV can be promptly used, in case of alerts, as visual moving system able to provide real-time and high-quality video feeds. The VR features can help technicians to visualise both instantaneous and historical data in a human-friendly manner, by immediately recognising the sensor nodes and their locations. Eventually, the API allows external systems, once authorized, to connect to our network, to receive collected data in a real-time manner.

In the following, Sections IV-B and IV-C focus both on sensor nodes, used in the WSN, and on the UAV role. Figures 3 and 5 show the position of the sensor nodes on the three buildings used as case studies and the UAV tasks, respectively. More details about the sensors nodes are reported in Tables 3 and 4, as well as the related statistics in Table 2. Section IV presents the concentrator module, that means, the local gateway and the GCS that collects data from close sensor nodes and from the UAV, respectively. Data exchange occurs as described in Section IV-A and according to the infrastructure depicted in Figure 4. Sections IV-D and V provide a detailed description about the data analysis module, that means, the used algorithms for time series analysis and the methods used for image processing and VR. The results are then reported in Tables 6 and 7. Section IV-D also provides a description of the data storage module. Both data storage module and on-top services run at the Monitoring Control Center (MCC), which is described in Section IV-D. In addition, Figures 9 and 10 show an example of 3D reconstruction and the VR environment. The 3D reconstruction and the VR environment can be seen at http://moscardo.isti.cnr.it.

The MOSCARDO architecture has been designed to provide technicians and authorities, in the field of cultural heritage preservation, with key insights into the condition of the monitored structure. Therefore, we can affirm that it can be considered as a support system for decision makers, especially during or immediately after events of interest. The whole system has been installed and tested in three different locations, visible in Figure 2, having different requirements, installation and data collection challenges. The first installation (see Figure 2c) is in large tower, namely *Mastio di Matilde*, pertaining to the Leghorn's Old Fortress, next to the sea and to the city's port, mixing indoor and outdoor sensor nodes. This large old tower is close to the port, thus being exposed to strong winds and salt air, further than the vibrations due to the engines of large ships departing closely. Sensor nodes have been installed indoor and a weather station outdoor. The second installation (see Figure 2b) is in a waterway below a large square, namely *Voltone*, with sensor nodes very close to the water below and to passing boats. The streets all around the square above the waterway are part of the Leghorn's arterial roads, thus exposing the waterway to varying loads [31]. The installation in the waterway has required the presence of specialized personnel, able to climb the inside walls for installing small devices in precise locations. The third installation (see Figure 2a) is in a large tower in the medieval city of San Gimignano, Siena. It is the tallest tower in the city, measuring approximately 54 meters, and daily visited by a large number of tourists. Sensor nodes have been installed indoor and a weather station outdoor. Several preliminary campaigns have been conducted to identify the most suitable placement locations for the sensor nodes in the monitored buildings, in order to: (*i*) install the minimum number of devices necessary to support SHM activities; (*ii*) limiting obtrusive installations; (*iii*) respecting enforced regulations on any placement on ancient surfaces (like walls). Installing the minimum number of sensor nodes, which anyway suffice to monitoring activities, has the advantage of reducing the amount of data to be delivered, thus reducing energy consumption as well. The main criterion, used for choosing the most suitable placement locations, aims at reducing any possible correlation among the data collected by the sensor nodes. In Figure 3, the sensors placements for the three cases study are shown. The sensors placement has been performed so to achieve minimal visual impact, while contemporary providing the desired detection of the structure modes.

IV. MONITORING SYSTEM

The system designed to support SHM purposes relies on PUB/SUB data exchanges, thus leveraging IoT communication protocols. In particular, the data collected from the sensor nodes is delivered to the remote server by means of MQTT, an IoT protocol for reliable data exchanges that decouples data producers from data consumers by means of a rendezvous node, referred to as broker. MQTT offers intrinsically reliable data exchanges because it is TCP-based, contrarily to the plain version of CoAP [25], [32], an UDP-based solution for RESTful data exchanges.

Looking at Figure 4, the blue blocks on the left can be identified, connecting the WSNs at each monitored site to the remote server via a 3G/4G data link. The local gateway, based on a Raspberry Pi and equipped with both WiFi and LoRa connectivity, collects data from close sensor nodes. The LoRa concentrator is based on the *iC880A* radio module transmitting and receiving at a frequency of 868MHz. LoRa-based sensor nodes, such as environmental sensors, generate tiny amount of raw data in a periodical manner (up to 12B/s for each node), exploiting an energy-efficient radio transceiver. WiFi-based sensor nodes generate really moderate amounts

(a) *Torre Grossa* (San Gimignano) (b) *Voltone*, Republic Square (Leghorn) (c) *Mastio di Matilde*, Old Fortress (Leghorn)

FIGURE 2: Three case studies under analysis in the MOSCARDO project; all sites are located in Tuscany, Italy. The *Mastio di Matilde* is the large round tower (in the top center) pertaining to the Old Fortress in Leghorn; the *Voltone* is the waterway below a large town square; the *Torre Grossa* is the tallest of the historical towers in the city of San Gimignano, close to Siena.

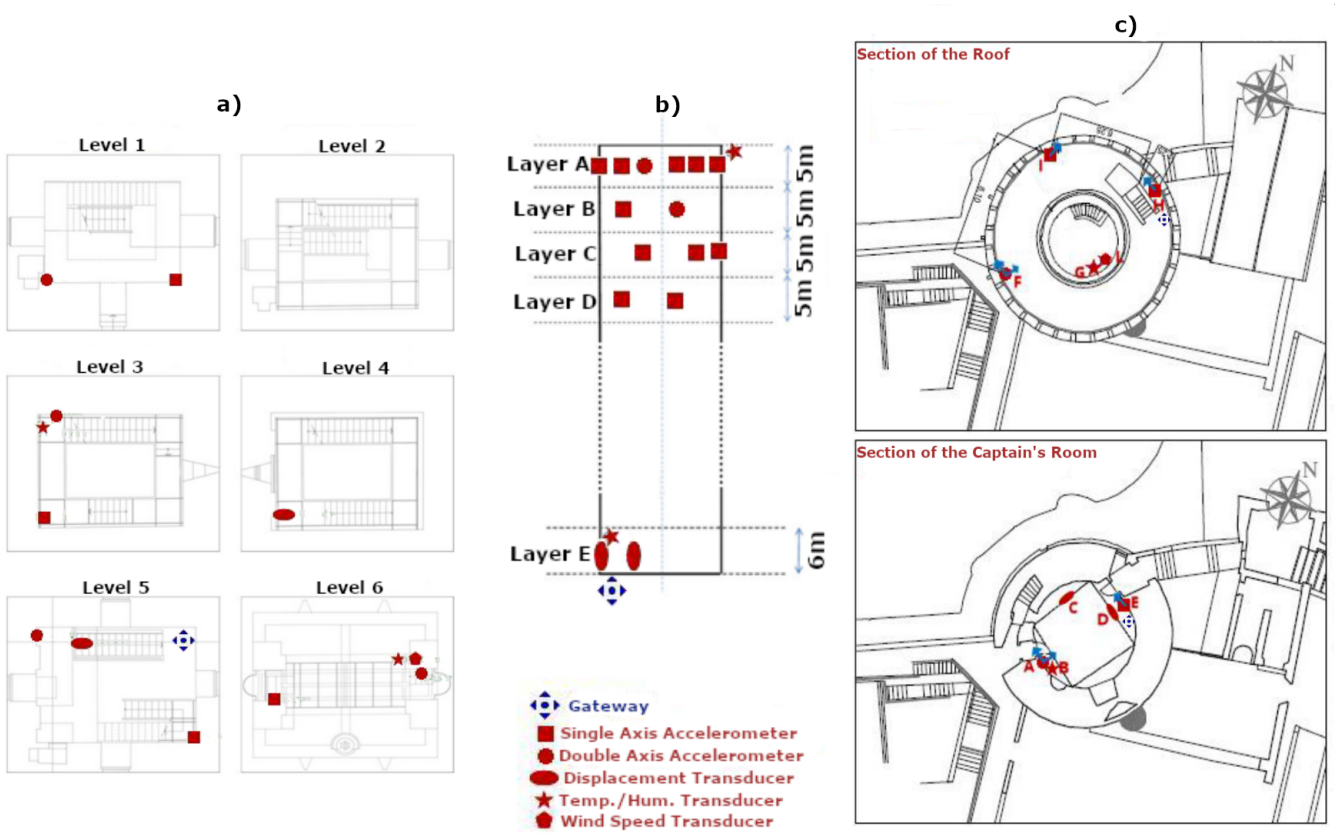


FIGURE 3: Position of the sensor nodes in the three buildings used as cases studies: a) Torre Grossa, b) Voltone, and c) Mastio di Matilde. The legend can be read in the middle bottom.

of raw data in a periodical manner (up to 300B/s for each node). A lightweight implementation of an MQTT client, running on sensor nodes themselves, publishes data on the local broker. Note that all WiFi sensor nodes collect data for T_W seconds in a temporally synchronized manner. In the first building (see Figure 2c), seven WiFi-based and one LoRa-based sensor nodes were installed; in the second building (see Figure 2b), fourteen WiFi and two LoRa, while, in the third building (see Figure 2a), ten WiFi and two Lora sensor nodes. Totally, 31 WiFi and 5 LoRa sensor nodes have been installed. Table 2 provides information on those sensor nodes, and on the average net and gross bitrate transferred by the proposed system.

The local gateway hosts an MQTT broker, namely Mosquitto, to receive data from both WiFi and LoRa nodes. The latter ones exploits an open-source LoRaWAN network-server, namely LoRaServer[2], running as a daemon on the local gateway to encapsulate received LoRa payloads into

[2]Documentation and source code available at: loraserver.io.

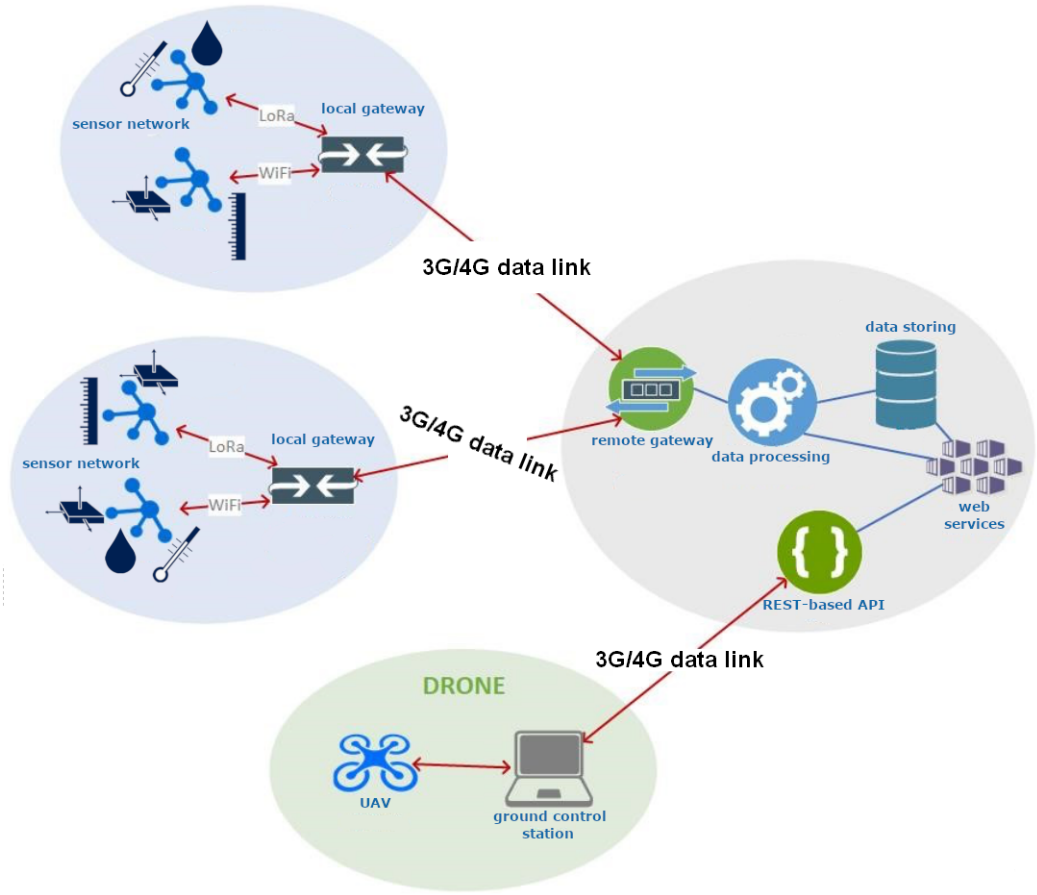


FIGURE 4: An overview of the communication infrastructure in use.

outgoing MQTT messages. In the end, the local broker receives data from both WiFi and LoRa sensor nodes; being configured as an MQTT bridge connected to the remote MQTT broker at the MCC (described in Section IV-D), the received messages are immediately transferred towards the latter, if Internet connectivity is available. In fact, during our large testbed (still running), 3G/4G connectivity has been intermittent for several reasons: such issue was taken into account during the design phase, motivating the choice of a reliable protocol solution and the aforementioned transparent bridge configuration. Such a setup, in our experience, relieves sensor nodes from the burden of dealing with intermittent Internet connectivity, because the gateway is always on and reachable (local network) and can persist data until a successful data transfer to the remote server, so, in the end, allowing the development of lightweight and energy-efficient sensor nodes.

In what follows, PUB/SUB data exchanges and MQTT topics are detailed in Section IV-A, and the UAV-based solution is described in Section IV-B. Then, the manuscript focuses on the deployed sensor nodes in Section IV-C. The functionalities provided by the MCC are described in Section IV-D, and eventually Section V deepens the techniques for image analysis and related results.

A. PUB/SUB DATA EXCHANGES

Our system implements state-of-the-art IoT solutions for data transmission and collection, thus ensuring interoperability as a core feature. Reliable data transfers have been taken into account thanks to the use of the MQTT protocol: the PUB/SUB architecture rests then on the gateways installed in the buildings, hosting rendezvous functionalities for both WiFi and LoRa sensor nodes. MQTT topics in use adhere to the following format and exchanging data payloads in JavaScript Object Notation (JSON) format. More specifically, in the case of sensor nodes, the following topic structure has been used. Each topic is prefixed by *moscardo/[building_id]*, where [building_id] is the ID of the site under monitoring. Then, the next element in the topic structure can be *node* or *gateway*, the former in the case of data exchanges with sensor nodes and the latter in the case of data exchanges with gateways. Unless explicitly stated below, the payload of each message is ignored (flag value) if the message does not need to carry any actual data. In the case of sensor nodes, taking into account that multiple sensors can be connected to a single node, the topic structure is as follows:

- sensor readings: *[node_id]/(measure_type)/ [sensor_id]/data* where (measure_type) can assume values:

	Mastio di Matilde	**Voltone**	**Torre Grossa**	**avg bitrate**
WiFi sensor nodes	7	14	10	(net) ≈2.4 Kb/s
LoRa sensor nodes	1	2	2	(net) ≈96 b/s
aggregated	10 (incl. 2 gtw)	17 (incl. 1 gtw)	13 (incl. 1 gtw)	(net) ≈56.8 Kb/s (gross) ≈68 Kb/s

TABLE 2: Sensor nodes and gateways (gtw) installed on the three monitored sites (see Fig. 2) and average bitrate: in the first two rows, the net bitrate is shown per sensor node in the last column; in the last row, aggregated values are instead shown, both net and gross (i.e., including control traffic and protocol stack overhead for the latter).

1) *accel* for accelerometers;
2) *th* for temperature and humidity;
3) *strain* for strain;
4) *disp* for displacement;
5) *bar* for atmospheric pressure;
6) *wind* for wind velocity and direction.

- sensor commands: *[node_id]/cmd/[command]* where [command] can assume values:
  1) *status* for status reports;
  2) *reboot* to reboot the sensor node;
  3) *update* to trigger a firmware update, if a new version is available;
  4) *set* [JSON payload] to set the configuration parameters;
  5) *get* to read the configuration parameters.
- errors: *[node_id]/error* in the case of errors.

In the case of gateways, the topic structure is equivalent:

- gateway commands: *[gw_id]/cmd/[command]* where [command] can assume values:
  1) *status* to collect detailed info about the gateway status;
  2) *ntp_restart* to restart the NTP service;
  3) *mqtt_restart* to restart the MQTT broker;
  4) *vpn_start* / *vpn_stop* to enable or disable the VPN client connection from the gateway to a remote VPN server for maintenance;
  5) *set* [JSON payload] to set the configuration parameters.

B. UNMANNED AERIAL VEHICLE (UAV)

The monitoring system makes use of an UAV, which can be considered as a mobile node of the WSN, to be deployed on-demand or in a periodical manner. In our scenario, it has been used to monitor the crack patterns of the structure in a regular fashion; also, it may be used for specific surveys close to the sensors should there be anomalies in the data stream and in the case of alerts, whenever other kinds of intervention might not be viable. There are several valuable advantages in using an UAV to monitor the structure: (*i*) improved accessibility to difficult areas; (*ii*) improved safety for operators that are not required to reach potentially hazardous areas (*iii*) time-effective surveys; (*iv*) repeatability of measurements, useful for long-term monitoring activities; (*v*) possibility to perform promptly structural assessment in case of alerts.

As shown in Figure 5, the task of the UAV is twofold: in the initial phase, it is used to acquire high resolution video footage of the structure to generate a 3D model for the VR environment; in the monitoring phase, it is able to follow a predefined path to take high quality pictures of all the cracks and spots under observation. To allow the former tasks, the sensing payload of the UAV is composed by two vision units: a Canon EOS M, capable of very high-quality 18MP still pictures, and an embedded smart camera that features the NVIDIA Jetson TX1 as computing module. The first unit is used mainly to take images of the cracks and it is remotely operated by the UAV pilot, while the smart camera compresses and sends a real-time video stream to the GCS; this helps the pilot during the navigation and it enables the Simultaneous Localization and Mapping (SLAM) software module, needed to position of the UAV in the virtual environment in real time. The UAV delivers an H.264 compressed video stream over an UDP-based protocol stack. 802.11g connectivity has been used to transfer data in Visual Line of Sight (VLoS) trials [33], while 4G-based data transmissions have been used in Beyond Visual Line of Sight (BVLoS) scenarios [34].

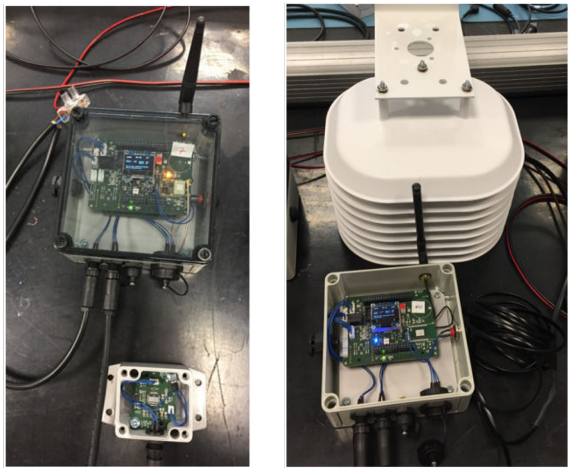
C. DEPLOYED SENSOR NODES

The sensor nodes used in our monitoring application are based on commercial components, carefully chosen to respect the requirements of SHM applications, which aims at performing Operational Modal Analysis (OMA). A careful analysis of the literature [1], [16], [18], in the field of SHM, has shown typical requirements to be met by the sensor nodes, meaning that the hardware should be able to reveal at least what follows: accelerations vary in the range $\pm 2 \div 2.5g$, where g is the standard gravity; the bandwidth (BW) of the modes falls in the range 0.5÷50 [Hz], and the admissible noise level is ω_N, in the range 1.4-14$\times 10^{-3}g$. The noise level ω_N determines the minimum measurement resolution, labeled with (RM), because it is directly proportional to the ratio $RM/\sqrt{BW}$. In our case, the minimum resolution is 0.01÷0.1 $10^{-3}g$, for an useful signal bandwidth of about 50 [Hz]. The sensitivity of the transducer is related to the analog-to-digital converter (ADC). In fact, measuring a minimum signal value RM of $0.01 \times 10^{-3}g$, by using an ADC with voltage levels between 0÷ 3.3 [V], with a minimum number of bits equal to 20 bits, the minimum useful value to vary the least significant bit is $3.3/2^{20} = 3.14$ [μV]. Hence, the transducer sensitivity must be greater than $3.14 \times 10^{-6}/0.01 \times 10^{-3} = 314$ [mV/g]. A final consideration about the ADC is related to its sampling frequency, which must be at least twice the maximum band of detectable frequencies, which in our

FIGURE 5: The two main different tasks of the UAV in the proposed IoT system: on the left, image acquisition for 3D reconstruction; on the right, periodical image acquisition for marker-based structural defects monitoring.

case is 100 [Hz]. Strain-gauge transducers should measure strains that generate displacements in the range of a few millimeters, with a resolution of a tenth of a millimeter. Instead, transducers have to measure displacements in the range of a few tens of millimetres. In both cases, the sensitivity of the transducer must guarantee a voltage level enough to vary the least significant bit of the ADC. Temperature and humidity transducers must be able to measure temperatures between - 10 and + 50 $^{\circ}C$, and relative humidity between 30% and 90%. An accuracy of $\pm$ 0.5 $^{\circ}C$ and $\pm$ 0.5% is considered sufficient for the temperature and the relative humidity, respectively.

The sensor node was designed as a set of logical blocks, where the latter can be interconnected through commercial hardware connectors: such a choice allowed a rather general design phase whose output can be adapted to the specific characteristics of the site and/or to the quantity to be monitored. The blocks, composing the system, are the following ones: the core block, the signal transduction block, the communication block, and the power block. Figures 6a (accelerometer) and 6b (temperature and humidity) show two sensor nodes configured for the mechanical and environmental quantities acquisition, respectively. The core block takes care of controlling all the other sub-blocks, i.e., managing the communication logic, conditioning the transduced signals, formatting (timestamp included) and storing of the acquired data. The core block is based on a 32-bit ARM Cortex-M



(a) Accelerometer node (b) Temperature/Humidity node

FIGURE 6: The casing in use: on the left (6a), the transducer is visible in the lower part of the figure in the light gray encasing, while the other hardware blocks are in the transparent IP67 casing above (a WiFi antenna is visible). On the right (6b), the environmental sensor nodes are in the accordion-like casing above, and the other hardware blocks in the open casing below (a LoRa antenna is visible).

by ST-Microelectronics, namely the STM32F411RE device. The conversion and conditioning of the signals is performed by an active low-pass filter, by a conversion circuit of the signal from differential to single-end, and by an ADC with variable gain amplification (PGA). The low-pass filter allows for a first detrending of the acquired signal, and it is based on a low-noise operational amplifier suitably configured to guarantee the lowest degradation of the acquired signal. The commercial device used for signal filtering is the AD8572. Such a choice has been dictated by the need of ultra-low offset, drift, and bias current. The AD8572 is able to eliminate the inter-modulation effects from interaction of the chopping function with the signal frequency in AC applications by using a spread-spectrum, auto-zero technique. With an offset voltage of 1 μV and a drift of 0.005 $\mu V/^{\circ}C$, the AD8572 is perfectly suited for applications, where error sources must be reduced as much as possible, such as in the case of strain gauges and accelerometers. The ADC adopted in our system is the ADS124S06, it operates with a precision of 24 bits and it is configured to ensure the least degradation of the converted signal. This is possible thanks to its PGA that allows an effective conversion of very weak signals, using an amplification algorithm optimized on the physical characteristics of the acquired signal. The transduction block involves different type of devices: accelerometers and strain gauges for the detection of mechanical quantities; humidity, temperature, and wind speed for the detection of environmental quantities. Transducers of mechanical quantities have a rather high sensitivity, as shown in Table 3, because they must enable the modal analysis of the structure on the basis of environmental forces. Furthermore, in order to limit the effects of common mode noise, the acquisition of the transduced signal is performed in differential mode. The communication block deals with formatting the digital signals from the ADC in JSON format, establishing communication with the gateway through the MQTT communication protocol, ensuring time synchronization through the Network Time Protocol (NTP), and managing the transceiver interface. A reliable and precise time reference is critical to allow for time series analyses at the MCC, by taking into account the spatial-time correlation of the acquired quantities. The available transceiver interfaces for the sensor node are shown in Table 4. Finally, the power supply block has been designed to use a solar-powered battery system enabling the energy harvesting paradigm. Table 4 shows some of the most important features of the components installed in the sensor nodes.

All the installed sensor nodes perform a $T_W = 900$ seconds long acquisition every hour. In the case of mechanical data, the sampling frequency is set to 100Hz, considering 3D coordinates; in the case of environmental data, the sampling frequency is set to 1Hz.

D. MONITORING CONTROL CENTER (MCC)

All the collected data are reliably delivered via MQTT to a remote cloud platform, which can be referred to as MCC. Looking at Figure 1, data analysis, storage, and on-top services run at the MCC. Raw data incoming from the WSNs deployed at the three different locations, earlier described, are analysed: the data coming from the sensor nodes pass preliminary validity checks to filter out incorrect measurements, and then significant modal frequencies are extracted and correlated with environmental data. Data coming from strain gauge and displacement sensors help in monitoring existing fractures on the walls. As in Figure 1, data coming from the WSN are collected and then forwarded by the MQTT broker at the MCC; data coming from the UAV arrive at the MCC via a REST-based API. Once the data elaboration has been performed, raw and elaborated data are stored in the database, which is used to feed the developed services on top of the lower layers of the system.

Three types of services have been developed, plus an API: all those services are used by authorized personnel and services, like local institutions in charge of the maintenance and care of the three aforementioned sites. The first service delivers periodic reporting, also allowing to browse historical raw and elaborated data: in such a way, the state of the building can be analysed taking into account natural decay and the maintenance actions to prevent and reduce it. It is one of most importance sources of information, because the understanding of long-term effects due to undertaken actions can be corroborated from numerical data provided by the remote monitoring system. Further than that, an alert service is in place: it is designed to immediately report any readings out of expected ranges, triggering alert procedures. On this, one of the procedures designed is based on the use of an UAV: it can be quickly deployed in an alert case to visually confirm, if possible, any sudden deterioration that triggered the alert. The use of a UAV allows to perform from the above such an inspection, in a safe manner, avoiding that technicians and other personnel enter the building, potentially exposing them to dangerous situations. If the UAV cannot detect anything significant, or in the rare case of a false alert, an additional assessment by authorized personnel can be performed on site. As aforementioned, the UAV is used in a periodical manner to complement data coming from the WSN with a visual inspection, allowing a technician to virtually navigate the building from the outside and seeing instantaneous and historical readings associated to the sensor nodes. Eventually, data stored in the database can be accessed by authorized external subjects via a REST-based API.

V. IMAGE ANALYSIS

In this section is provided the description of the image processing services and of the interface of virtual reality developed. A marker-based image processing method is specifically designed and developed to monitor crack patterns in ancient structures, while an interface of virtual reality provides real time information about the structure and the sensors on it.

	Device	Producer	Model	Main Features
Mechanical Transducers	Analog Output MEMS Single-Axis Accelerometer	Colibris Safran	VS1002	range $\pm$ 2 g, bandwidth 0-700Hz, noise in band 7 $\mu g/\sqrt{Hz}$ sensitivity 1.35 V/g
	Analog Output Strain-Gauge	Tokyo Sokki Kenkyujo	PI-2-300	range $\pm$ 2 mm, rated output 2 mV/V, gauge length 300 mm sensitivity 2$\cdot 10^{-3}$ $strain/mm$
	Analog Rectilinear Displacement Transducer	Gefran	PZ67-A	range $\pm$ 10 mm resolution infinite, linearity 0.5%
Environmental Transducers	14 bit-Digital Temperature/Humidity Transducer	Davis Instruments	Davis 7346	temp. range -40 $\div$ +124 $^\circ C$ temp. accuracy $\pm$0.5 $^\circ C$ temp. resolution 0.01$^\circ C$ hum. range 0 $\div$ 100 $\%RH$ hum. accuracy $\pm$ 3 $\%RH$ hum. resolution 0.05 $\%RH$
	Analog Output Wind Transducer	Davis Instruments	Davis 6410	wind speed range 1$\div$322 Km/h wind speed accuracy $\pm$3 Km/h wind direction accuracy $\pm$ 3 $^\circ$

TABLE 3: The transducers in use in our testbed.

	Device	Producer	Model	Main Features
Signal Elaboration	ADC	Texas Instrument	ADS124S06	Sigma-Delta 24-bits, 6 channels single-ended or 3 differential channels Programmable Gain Amplifier
Communications	WiFi Transceiver LoRa Transceiver	ST-Microelectronic MultiTech	SPWF04 MultiConnect xDot	WiFi 802.11 b/g/n LoRa/LoRaWan 1.01

TABLE 4: The hardware components for signal conditioning and communications in use in our testbed.

A. IMAGE PROCESSING

The main goal of this task was to implement a robust and computationally inexpensive algorithm able to automatically inspect the crack patterns of an ancient building, with no detriment in measurement accuracy.

As regards crack monitoring, our main case study was the Old Fortress, which shows sides quite far and definitely not co-planar. This makes very difficult to achieve absolute and accurate measures of the crack opening with standard methods. Also, the walls are partially surrounded by the sea, hence difficult to be monitored. For such structures, which are irregular, subject to environmental agents and to seasonal changes (e.g. for weeds growing on it), photogrammetric reconstruction quality may be not enough to monitor over time the crack pattern variations. Hence, a set of coded markers was used, to provide a complete and stable 3D information about specific fiducial points along the crack, to be tracked over time. Under the assumption that the fiducial markers were placed along the crack in the most critical points (i.e., most stressed and subject to variations), our image-processing method aimed at a reliable long-term crack monitoring.

Our solution is based on the SLAM approach and optimized for the creation of 3D map of the markers visible in the images acquired. It was inspired by the work in [35], in which the authors: *(i)* define a dictionary of coded planar square markers; and *(ii)* address the mapping problem as a variant of the sparse bundle adjustment problem, by solving the corresponding graph-pose problem. A sequence of frames of the same scene, as well as camera calibration parameters, has been acquired. At each frame, the markers are detected (using the ArUco library[3]), and their 3D pose is estimated. The optimization is done thanks to the simultaneous sub-pixel detection of the corners of all the markers placed in the scene, by minimizing their re-projection errors in all the observed frames. The output of the algorithm is a complete 3D information about each marker, thus enabling to track the variations of the crack pattern over time. In particular, such data includes: *(i)* the set of the 3D coordinates of the corners of each marker; *(ii)* the set of the Euclidean distances between the barycenters of each pair of markers; and *(iii)* the angle variations between the reference frame associated to each marker. More in detail, the algorithm works as follows: (1) a sequence of frames of the same scene (at least six) is acquired; (2) at each frame, the graph-pose is estimated by minimizing the re-projection error in the detection of the marker corners; (3) for each marker, the coordinates of its corners are assessed in the 3D space, as well as the coordinates of the barycenter, and the associated reference frame; (4) the Euclidean distances between the markers' barycenters are computed; (5) the change in markers' poses is assessed

[3]ArUco is a library for Augmented Reality (AR) applications based on OpenCV, available at: https://www.uco.es/investiga/grupos/ava/node/26.

by computing the Euler angles {yaw, pitch, roll} associated to the rotation matrix, which sends the reference frame of one marker into the reference frame of the other. An example about how the image processing algorithm computes the Euclidean distance between markers is shown in Figure 7. Since there is not any absolute reference frame, in the present study, the pose variation of the markers is tracked over time, by computing the rotation matrix between the frame associated to each pair of marker at each acquisition. All the pairs in the marker configuration are considered in the algorithm, in order to account for any possible pose variation.

1) Controlled experiments

In order to assess the robustness of the overall marker-based SLAM algorithm, the accuracy of ArUco marker detection and the repeatability of the measurements, some tests in a controlled setting were preliminary conducted.

In such experiments, a crack opening (of around 1 mm up to 5 mm) was simulated and a set of images was acquired at about 150 cm from the target. Six planar markers, with side length of 5.5 cm, were fixed on two identical boxes (three markers on left, three on right). The left box was allowed sliding integral with a mobile axis of a coordinatographer, while the right one was fixed to the table. The certified accuracy of the coordinatographer is of 0.1 mm. For simplicity, a planar motion was simulated and only some distances between the pair of markers were considered: 5 steps by 5 mm, and other 5 steps by 1 mm, by moving the left box far from the right one along one axis, as shown in Table 5; nonetheless, the approach can be easily applied, as it is, to a 3D variation of the marker configuration.

Furthermore, the marker-based SLAM method was compared with a more traditional one, which uses the 3D translation vectors directly estimated from the camera pose parameters, in the computation of the 3D position of the markers. Such tests showed that the marker-based SLAM method is quite a stable method, able to provide an accuracy of less than 1 mm with only six images of the scene. Further details about the methods, as well as the results, are described in [36], [37].

2) Real testbed

Experiments were carried out in real settings, as described in [38], which reports the initial results of the crack monitoring of Old Fortress walls in Leghorn (Figure 2c), located in the area called *Bastione della Capitana*. This is the only case study, where large cracks need to be monitored. Six pairs of markers, four small pairs (0.1 m side length) and two larger ones (0.2 m side length), have been glued to the wall. A sequence of 6 images were acquired 3 m, 6.5 m and 9 m far from the wall. All markers were correctly detected, as shown in Figure 8. Further acquisition campaigns, performed over a period of four months, allowed for extending such preliminary results, as shown in Table 7. As expected, the computed distances are very close to the ground truth, in most cases. In one case (pair of markers *#27-#20*), the proposed method reached a sub-millimiter accuracy in estimating the distance.

In Table 6, details are provided on the data acquired during the whole experimentation, especially on the two marker pairs #38-#23 and #3-#4: the Euler angles (theta, phi, and psi [degrees]) are shown, the norm of the vector made of Euler angles [degrees], and the distance [meters] between barycenters. The standard deviation is provided as well. This experimentation leads to the two following key points: the feasibility of the proposed method for the crack monitoring through UAV, and the robustness of such measurement method outdoor in an uncontrolled environment. Actually, since the markers are not co-planar, the distance between the barycenters are measured with a flexible meter: this cannot be considered an accurate measure, but allow us to understand whether the distances computed with our method are plausible. In addition, the low standard deviation computed over the 4-month period through regular acquisitions (with no critical events over such a period) shows the robustness of our measures. A final remark about how to use these data in the long-term monitoring: after acquiring data over about a year (without critical events), and estimating the crack pattern variation as the linear and pose variations associated to the configuration of markers, those measurements can be used, along with their standard deviations, to define alert thresholds or pejorative trends, so triggering further investigations or recovery actions.

3) Considerations

The results discussed so far demonstrate that irregular ancient structures, subject to environmental agents and to seasonal changes, a marker-based SLAM algorithm provides a complete and robust 3D information about specific critical points along the crack. Therefore, the proposed image processing method allows for both an accurate quantitative analysis of the crack and a comparison of the crack measurements over time. Furthermore, compared to other minimally-invasive, state-of-the-art methods for crack monitoring, as in Table 8, the solution herein proposed can be considered as a reasonable trade-off between the minimal invasiveness required to work with cultural heritage, the cost-effectiveness of the technological resources, and the measurement accuracy required for structural defects monitoring.

B. VIRTUAL REALITY

As shown in Figures 9 and 10, the 3D environment and VR solutions have been adopted in order to enrich the remote inspection providing a high-quality reconstruction of the monitored sites together with the location of installed sensors and their readings. Moreover, VR offers valid support during periodical or extraordinary inspections, giving the UAV operator a quick look of the explored site and of the installed sensors with the vehicle correctly placed in the virtual scene in real-time.

In order to provide a useful tool to the UAV operator and to the users who are responsible for monitoring the structures,

Pair of markers	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}
$d(A,B)$	4.33	6.14	5.32	4.53	4.26	2	1.10	0.70	0.90	1.5
$d(C,D)$	4.60	5.41	5.63	4.6	4.38	1.79	0.71	1.10	0.78	1.63
$d(E,F)$	4.70	5.58	6.2	4.31	4.02	1.96	1.26	0.64	0.8	1.8
actual Δ	5	5	5	5	5	1	1	1	1	1

TABLE 5: Simulation of the crack opening: three pairs of markers (A and B, C and D, E and F) moving away from each other, in five steps by 5 mm ($T_1, \ldots, T_5$) and 5 steps by 1 mm ($T_6, \ldots, T_{10}$). All the distance values are expressed in mm.

Pair of markers	**Pose parameters**	**I** *acq.*	**II** *acq.*	**III** *acq.*	**IV** *acq.*	**V** *acq.*	**VI** *acq.*	**VII** *acq.*	**St.Dev.**
#38-#23	theta	392.72	359.79	359.75	359.76	359.73	359.90	360.13	**0.15**
	phi	-0.05	-0.30	0.11	0.00	0.98	0.18	-1.44	**0.72**
	psi	360.16	360.65	359.62	360.60	359.22	358.91	359.52	**0.68**
	norm	26.83	26.84	26.82	26.84	26.83	26.81	26.80	**0.01**
	baryc. distance	0.3897	0.3914	0.3942	0.3943	0.3951	0.3942	0.3931	**0.0020**
#3-#4	theta	179.32	179.27	179.20	179.25	179.22	179.22	179.04	**0.09**
	phi	1.86	1.11	2.92	1.30	2.16	-0.61	2.21	**1.13**
	psi	359.55	357.16	358.39	358.71	356.52	363.97	359.99	**2.44**
	norm	23.25	23.18	23.25	23.22	23.19	23.29	23.26	**0.04**
	baryc. distance	0.5584	0.5581	0.5603	0.5601	0.5615	0.5606	0.5599	**0.0012**

TABLE 6: The table focuses on two marker pairs (#38-#23 and #3-#4); data was acquired at about 6 mt during the whole experimentation (acquisitions I to VII). The pose variation is appreciated estimating the Euler angles (first three rows, in degree) associated to the rotation matrix sending one marker's frame into the other marker's frame; also the norm of the (theta, phi, psi) vector is computed (fourth row, in degrees). The last row, for each pair, is the Euclidean distance between barycenters is expressed in meter.

Pair of markers	**I** *acq.*	**II** *acq.*	**III** *acq.*	**IV** *acq.*	**V** *acq.*	**VI** *acq.*	**VII** *acq.*	**Avg**	**St.Dev.**	*Ground Truth (GT)*	**Avg - GT**
#27-#20	0.329	0.332	0.330	0.330	0.330	0.331	0.332	**0.331**	0.001	**0.331**	**0.000**
#38-#23	0.386	0.387	0.387	0.387	0.388	0.388	0.388	**0.387**	0.000	**0.385**	**0.002**
#3-#4	0.555	0.552	0.553	0.555	0.555	0.555	0.555	**0.554**	0.001	**0.551**	**0.003**
#31-#26	0.339	0.334	0.334	0.334	0.335	0.335	0.335	**0.334**	0.000	**0.327**	**0.007**
#25-#18	0.428	0.428	0.428	0.429	0.428	0.428	0.428	**0.428**	0.000	**0.424**	**0.004**
#5-#10	0.522	0.540	0.524	0.530	0.525	0.520	0.542	**0.530**	0.008	**0.518**	**0.011**

TABLE 7: Computed distances between the barycenters of each pair of markers. The data refers to image acquisitions performed 3m far from the crack. Last two columns show the ground truth and the absolute value of the difference between the ground truth and the average value of the evaluated distance. All the distances are expressed in meters.

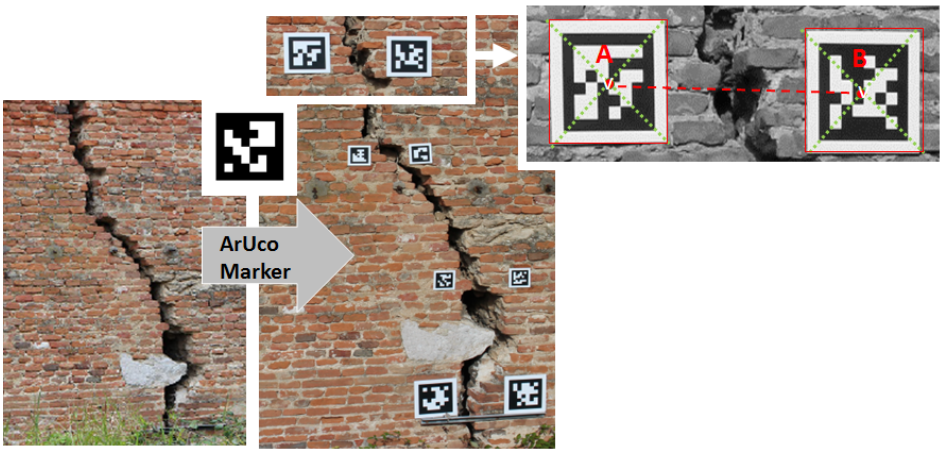


FIGURE 7: How the marker-based SLAM algorithm works from image acquisition to the barycenter detection for each marker in the scene.

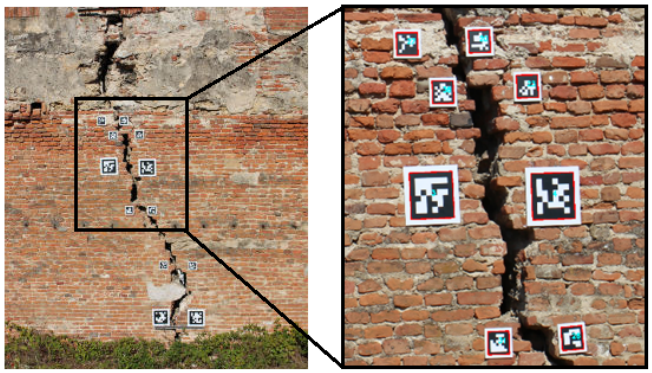


FIGURE 8: The ArUco markers glued to the wall of the Old Fortress (Leghorn, Italy) were correctly detected by the marker-based SLAM algorithm.

the 3D model reconstruction of the ancient structure has to be as close as possible to the reality (see Figure 9). In this way, the operator will easily plan UAV flights and rely on that to estimate the vehicle position. In this scenario, photogrammetric techniques come in handy to virtual reconstruct the monitored site. Aiming at ensuring reconstruction quality, a reconstruction chain has been established that exploits Agisoft Metashape[4] (a commercial offline software) and Meshlab[5] for mesh downsampling and fine correction [48]. Reconstructions have been obtained employing Metashape with the images acquired during the first UAV survey.

Generated reconstructions are used to create virtual 3D scenes that are populated by adding information about installed sensors and markers and linking to the available records in the database (see Figure 10). In detail, each scene is built through Unity Game Engine[6]; thanks to this, all the linear transformations required to perform movement inside the scene, object placement, lighting and collision management are entrusted to the game engine. Thanks to a set of dedicated developed scripts, the game engine is able to connect with the database in the MCC to retrieve the list of sensors installed in the site, populate the scene properly with the sensors and, for each sensor, retrieve acquired data and their elaboration. Inside the scene, sensors are represented as an icon with color depending on sensor status. Moreover, for each sensor, a panel, containing its name and a preview of retrieved values, is shown. As shown in Figure 10, from this panel, users can access all retrieved data that will be displayed in a dedicated portion of the screen. The developed 3D virtual environment[7] offers the possibility to show, in real-time, the UAV position inside the monitored site. This functionality is useful, as mentioned before, for UAV operators who can have another point of view of the vehicle and its surroundings and the relative sensor positions. Further, in the case of an alert that inhibits on-site inspection, the UAV and the operator can still operate quickly and safe. The virtual environment is directly connected with the SLAM

[4]Agisoft Metashape at: https://www.agisoft.com/
[5]Meshlab at: http://www.meshlab.net/
[6]Unity Game Engine at: https://unity.com/

[7]The MOSCARDO 3D virtual environment is available at: http://moscardo.isti.cnr.it

Method	*Pros*	*Cons*
UAV + Stereo-photogrammetric scanning [29]	• High resolution photos • Reconstruction of a very accurate 3D model of the inspected structure • Non invasive	• Highly demanding with respect to hardware/software resources • A large amount of photos is needed • Prone to errors in the case of irregular structures/surfaces
Installation of sensors that evaluate the response of the concrete to an electrical/acoustic stimulus [39], [40]	• The intrinsic physical properties of the material under investigation (concrete) are exploited • A quantitative analysis is performed	• Invasive (an electrical/acoustic stimulus has to be applied to the concrete) • Highly demanding with respect to hardware resources • Suitable for concrete walls only
MAV+ High resolution camera + edge detection algorithms [4], [41]	• The identification of structural defects is automatically performed • Non invasive	• Not applicable in the case of light-coloured walls/surfaces • Suitable for homogeneous materials
Reflective targets + camera (manual acquisition) + photogrammetry [42], [43]	• Just a few photos are sufficient • A quantitative analysis of the structural defect can be performed	• Light conditions-dependent • Minimally invasive • Not applicable in the case of area difficult to access, or variable/dangerous environmental conditions
Active/Passive thermography + edge detection algorithms [44], [45], [46]	• Light condition- independent • Allows for detection of water seepage and/or other structural damages due to humidity	• In the case of active termography, a heat stimulus is needed • Highly demanding with respect to hardware resources • Climatic conditions may affect the emissivity of the material and, hence, the acquisition
LiDAR-equipped UAV [47]	• Highly accurate • Non invasive • Able to detect the depth of the defects	• Highly demanding with respect to hardware resources • High expensive solution
Our method: UAV + ArUco Coded Markers + close range photogrammetry	• The markers may serve as waypoints for the UAV flight plan • Light conditions independent • The markers are coded, thus, they can be easily and rapidly detected • Both planar and non-planar displacement can be monitored • High accuracy when performing a quantitative analysis of the crack	• Minimally invasive • At least six photos are needed

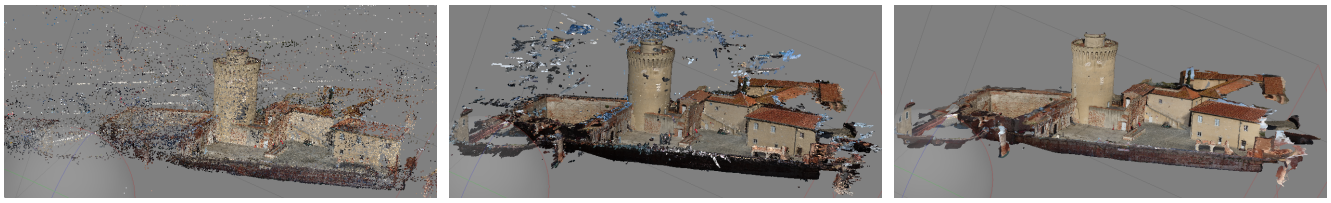
TABLE 8: Comparison between the state-of-the-art, minimally-invasive methods for crack monitoring and the proposed one.

software module, operating on the GCS, to retrieve the live UAV position. SLAM software module estimates UAV motion, exploiting ORB-SLAM[8] [49]; UAV movements are expressed as absolute shifts from the starting point. In this scenario, the absolute shift is a better choice with respect to incremental shifts, since it avoids issues due to possible packet loss. Finally, each UAV inspection and its movements are stored in the database and can be accessed and run in playback mode directly from the virtual environment.

VI. BEYOND NUMERICAL RESULTS

From a technical point of view, the use of a WSN and of an UAV can be quite effective in acquiring a rich set of data to provide both experts and local authorities with a comprehensive description over time of the health of any ancient structure. Furthermore, the design of such a support system for decision making met the constraints of modularity, flexibility, and minimally invasiveness; hence, its features allow to customize the installation and the monitoring process, thus respecting the peculiarities of the inspected architectural asset. Beyond numerical results, the real deployment of MOSCARDO system clearly pointed out the positive acceptance of such a system by local authorities, whom asked for the prolongation of the monitoring in two cases: Torre Grossa in San Gimignano, and Mastio di Matilde in Leghorn. Indeed, the local authorities demonstrated a strong interest in collecting data on long temporal scales (months or years) to be analysed as described in this work. Long-

[8]The software is available at: http://webdiis.unizar.es/~raulmur/orbslam



(a) *Sparse point cloud* (b) *Dense point cloud* (c) *Mesh*

FIGURE 9: The three phases of 3D reconstruction of the *Mastio di Matilde* in the Old Fortress (Leghorn)

FIGURE 10: The VR environment: users can explore the virtual reconstruction of a monitored structure having, in the meantime, direct access to values measured by the sensors deployed on it.

term data collecting allows also the fine tuning the thresholds currently in use in the system to trigger any alerts. Research scientists, including the authors, have been and are still collaborating with the local decision makers on the planning of future actions, in order to exploit and further improve the MOSCARDO system.

VII. CONCLUSIONS

This paper describes the design and the implementation of a monitoring system that integrates the environmental and mechanical data acquired through an IoT network with the images and the context information collected by means of an UAV. The images collected from above are used to build a 3D reconstruction of the ancient structures, showing the exact position of the deployed sensor nodes, thus allowing an operator to dynamically interact with the real-time readings collected by the IoT network, exploiting a PUB/SUB paradigm. Such a 3D virtual tool offers the possibility to show the virtual point of view of the UAV. The VR tool can be accessed online at this URL: http://moscardo.isti.cnr.it.

The UAV was also used for sensing purposes. In fact, the on-board camera has been used to automatically inspect and measure the crack patterns of the monitored structures with high accuracy, as shown by the numerical results presented in this work. The proposed monitoring system has been deployed in three different case studies, so proving its versatility and scalability in different operating conditions. Furthermore, the system has proved to be reliable and robust to rather severe operating conditions, in fact being tested for more than twelve months. In fact, despite harsh environmental conditions, i.e., strong sun in summer and sustained rainfalls in winter, close seacoast, the system has proved to be able to automatically restore all deployed services with minimal human intervention. Our study has some limitations, anyway: regarding the image-based method for crack monitoring, it is essential for the UAV to be in stable hovering for properly acquiring high resolution images without blur effects, further than be close to the target (from 1.5 to 2 meters). Furthermore, the configuration of the markers installed along a crack may greatly affect the accuracy of the analysis, which is based on the evaluation of the distance variations between those, thus requiring careful attention. Anyway, we highlight again the low cost of such a monitoring solution.

As future tests, we plan to increase the number of samples for each acquisition related to the image processing method (six images per session have been acquired in the presented

tests), which would result in a lowering of the measurement error. Finally, in order to improve the accuracy of the markers pose estimation, the ChArUco diamond [50] markers[9] may be used instead of the single ArUco markers. The detection of a diamond marker takes advantage of the known relative position of the markers in it, and it will increase robustness and accuracy of the marker pose estimation.

ACKNOWLEDGMENT

This work was carried out in the framework of the MOSCARDO project, co-funded by the Tuscany Region (Italy) under the Regional Implementation Program for Underutilized Areas Fund (PAR FAS 2007-2013) and the Research Facilitation Fund (FAR) of the Ministry of Education, University and Research (MIUR). The authors also wish to thank all the other stakeholders of the MOSCARDO consortium: Capitaneria di Porto, Autorità Portuale di Livorno, and the CNR UAV pilot, Andrea Berton.

REFERENCES

[1] J. P. Lynch, C. R. Farrar, and J. E. Michaels, "Structural Health Monitoring: Technological Advances to Practical Implementations [scanning the issue]," Proceedings of the IEEE, vol. 104, no. 8, pp. 1508–1512, 2016.

[2] P. Barsocchi, P. Cassara, F. Mavilia, and D. Pellegrini, "Sensing a City's State of Health: Structural Monitoring System by Internet-of-Things Wireless Sensing Devices," IEEE Consumer Electronics Magazine, vol. 7, no. 2, pp. 22–31, 2018.

[3] V. Gattulli, F. Graziosi, F. Federici, F. Potenza, A. Colarieti, and M. Lepidi, "Structural Health Monitoring of the Basilica S. Maria di Collemaggio," in Proceedings of The Fifth International Conference on Structural Engineering, Mechanics and Computation (SEMC 2013), 2013.

[4] C. Eschmann, C.-M. Kuo, C.-H. Kuo, and C. Boller, "Unmanned Aircraft Systems for Remote Building Inspection and Monitoring," in Proceedings of the 6th European Workshop on Structural Health Monitoring, Dresden, Germany, vol. 36, 2012.

[5] G. Morgenthal and N. Hallermann, "Quality Assessment of Unmanned Aerial Vehicle (UAV)-based Visual Inspection of Structures," Advances in Structural Engineering, vol. 17, no. 3, pp. 289–302, 2014.

[6] E. T. Bokeno, T. M. Bort, S. S. Burns, M. Rucidlo, W. Wei, D. L. Wires et al., "Package Delivery by Means of an Automated Multi-copter UAS/UAV Dispatched from a Conventional Delivery Vehicle," Jul. 14 2016, uS Patent App. 14/989,870.

[7] R. Haarbrink and E. Koers, "Helicopter UAV for Photogrammetry and Rapid Response," in 2nd International Workshop on The Future of Remote Sensing, ISPRS Inter-Commission Working Group I/V Autonomous Navigation, vol. 1. Citeseer, 2006.

[8] J. Armesto, P. Arias, J. Roca, and H. Lorenzo, "Monitoring and Assessing Structural Damage in Historical Buildings," The Photogrammetric Record, vol. 21, pp. 269–291, 2006.

[9] Z. Liu, Y. Yu, G. Liu, J. Wang, and X. Mao, "Design of a Wireless Measurement System based on WSNs for Large Bridges," Measurement, vol. 50, pp. 324–330, 2014.

[10] Y. Sun, H. Song, A. J. Jara, and R. Bie, "Internet of Things and Big Data Analytics for Smart and Connected Communities," IEEE access, vol. 4, pp. 766–773, 2016.

[11] X. Li, W. Yu, and S. Villegas, "Structural Health Monitoring of Building Structures with Online Data Mining Methods," IEEE Systems Journal, vol. 10, no. 3, pp. 1291–1300, 2016.

[12] L. Zhang, G. Zhou, Y. Han, H. Lin, and Y. Wu, "Application of Internet of Things Technology and Convolutional Neural Network Model in Bridge Crack Detection," IEEE Access, vol. 6, pp. 39 442–39 451, 2018.

[13] K. Hyunjun, L. Junhwa, A. Eunjong, C. Soojin, S. Myoungsu, and S. Sung-Han, "Concrete Crack Identification Using a UAV Incorporating Hybrid Image Processing," Sensors, vol. 17, no. 2052, pp. 1–14, 2017.

[14] C. Krisada, S. Mayank, A. Luqman, K. Wasif, and P. Nakhon, "Crack Detection in Historical Structures based on Convolutional Neural Network," International Journal of GEOMATE, vol. 15, pp. 240–251, 2018.

[15] M. D. Phung, T. H. Dinh, Q. P. Ha et al., "System Architecture for Real-Time Surface Inspection Using Multiple UAVs," IEEE Systems Journal, 2019.

[16] Q. Lin, F. Zhang, W. Jiang, and H. Wu, "Environmental Monitoring of Ancient Buildings Based on a Wireless Sensor Network," Sensors, vol. 18, no. 12, p. 4234, 2018.

[17] M. Bacco, F. Delmastro, E. Ferro, and A. Gotta, "Environmental Monitoring for Smart Cities," IEEE Sensors Journal, vol. 17, no. 23, pp. 7767–7774, 2017.

[18] A. B. Noel, A. Abdaoui, T. Elfouly, M. H. Ahmed, A. Badawy, and M. S. Shehata, "Structural Health Monitoring using Wireless Sensor Networks: a Comprehensive Survey," IEEE Communications Surveys & Tutorials, vol. 19, no. 3, pp. 1403–1423, 2017.

[19] M. Z. A. Bhuiyan, G. Wang, J. Cao, and J. Wu, "Sensor Placement with Multiple Objectives for Structural Health Monitoring," ACM Transactions on Sensor Networks (TOSN), vol. 10, no. 4, p. 68, 2014.

[20] M. Feng, Y. Fukuda, M. Mizuta, and E. Ozer, "Citizen Sensors for SHM: Use of Accelerometer Data from Smartphones," Sensors, vol. 15, no. 2, pp. 2980–2998, 2015.

[21] Y. Yu, R. Han, X. Zhao, X. Mao, W. Hu, D. Jiao, M. Li, and J. Ou, "Initial Validation of Mobile-Structural Health Monitoring Method using Smartphones," International Journal of Distributed Sensor Networks, vol. 11, no. 2, p. 274391, 2015.

[22] M. M. Ratnam, B. Y. Ooi, and K. S. Yen, "Novel moiré-based Crack Monitoring System with Smartphone Interface and Cloud Processing," Structural Control and Health Monitoring, vol. 26, no. 10, p. e2420, 2019.

[23] K. Máthé and L. Buşoniu, "Vision and Control for UAVs: A Survey of General Methods and of Inexpensive Platforms for Infrastructure Inspection," Sensors, vol. 15, no. 7, pp. 14 887–14 916, 2015.

[24] C. A. Tokognon, B. Gao, G. Y. Tian, and Y. Yan, "Structural Health Monitoring Framework based on Internet of Things: a Survey," IEEE Internet of Things Journal, vol. 4, no. 3, pp. 619–635, 2017.

[25] M. Bacco, L. Boero, P. Cassará, M. Colucci, A. Gotta, M. Marchese, and F. Patrone, "IoT Applications and Services in Space Information Networks," IEEE Wireless Communications Magazine, vol. PP, pp. 1–7, 2019.

[26] C. Kanellakis and G. Nikolakopoulos, "Survey on Computer Vision for UAVs: Current Developments and Trends," Journal of Intelligent and Robotic Systems, vol. 87, pp. 141–168, 2017.

[27] B. Kim and S. Cho, "Automated Vision-Based Detection of Cracks on Concrete Surfaces Using a Deep Learning Technique," MDPI Sensors, vol. 18, no. 10, p. 3452, 2018.

[28] F. Kucuksubasia and A. Sorguc, "Transfer Learning-Based Crack Detection by Autonomous UAVs," in 35th International Symposium on Automation and Robotics in Construction, no. 135. ISARC, 2018.

[29] M. Marialuisa, d. C. Gerardo, R. Ivan, M. Marialaura, N. Andrea, and d. F. Gianmarco, "3D Photogrammetric Reconstruction by Drone Scanning for FE Analysis and Crack Pattern Mapping of the Bridge of the Towers, Spoleto," Key Engineering Materials, vol. 747, pp. 423–430, 2017.

[30] J. Yu, X. Meng, B. Yan, B. Xu, Q. Fan, and Y. Xie, "Global Navigation Satellite System-based Positioning Technology for Structural Health Monitoring: a Review," Structural Control and Health Monitoring, 2019.

[31] M. Girardi, C. Padovani, and D. Pellegrini, "The NOSA-ITACA code for the safety assessment of ancient constructions: A case study in Livorno," Advances in Engineering Software, vol. 89, pp. 64–76, 2015.

[32] M. Bacco, P. Cassarà, M. Colucci, and A. Gotta, "Modeling Reliable M2M/IoT Traffic over Random Access Satellite Links in Non-saturated Conditions," IEEE Journal on Selected Areas in Communications, vol. 36, no. 4, 2018.

[33] M. Bacco, S. Chessa, M. Di Benedetto, D. Fabbri, M. Girolami, A. Gotta, D. Moroni, M. A. Pascali, and V. Pellegrini, "UAVs and UAV swarms for Civilian Applications: Communications and Image Processing in the SCIADRO Project," in International Conference on Wireless and Satellite Systems. Springer, 2017, pp. 115–124.

[34] M. Bacco, P. Cassará, A. Gotta, and V. Pellegrini, "Real-Time Multipath Multimedia Traffic in Cellular Networks for Command and Control Applications," in Vehicular Technology Conference (fall), Honolulu, Hawaii, USA. IEEE, September 2019, pp. 1–5.

[35] R. Munoz-Salinas, M. J. Marin-Jimenez, E. Yeguas-Bolivar, and R. Medina-Carnicer, "Mapping and Localization from Planar Markers," Pattern Recognition, vol. 73, pp. 158–171, 2018.

[9] A diamond marker is a chessboard composed by 3x3 squares and 4 ArUco markers inside the white squares.

[36] D. Germanese, G. Leone, D. Moroni, M. Pascali, and M. Tampucci, "Long-Term Monitoring of Crack Patterns in Historic Structures Using UAVs and Planar Markers: A Preliminary Study," Journal of Imaging, vol. 4, no. 8, p. 99, 2018.
[37] D. Germanese, G. R. Leone, D. Moroni, M. A. Pascali, and M. Tampucci, "Towards Structural Monitoring and 3D Documentation of Architectural Heritage Using UAV," in Multimedia and Network Information Systems, Proceedings of the 11th International Conference MISSI 2018. Springer, 2019, pp. 1–11.
[38] D. Germanese, M. A. Pascali, A. Berton, G. R. Leone, D. Moroni, B. Jalil, and M. Tampucci, "Architectural Heritage: 3D Documentation and Structural Monitoring using UAV," in Proceeding of 1st Int. Workshop on Visual Pattern Extraction and Recognition for Cultural Heritage Understanding (VIPERC2019), Jan 30, 2019, Pisa, Italy. CEUR-WS, 2019, pp. 1–12.
[39] N. Rajic, "In situ Monitoring of Crack Growth in Mild Steel under Closure Conditions Using a Piezotransducer Array," Journal of Intelligent Material Systems and Structures, vol. 11, pp. 696–702, 2000.
[40] T. Hou, , and J. Lynch, "Electrical Impedance Tomographic Methods for Sensing Strain Fields and Crack Damage in Cementitious Structures," Journal of Intelligent Material Systems and Structures, vol. 20, pp. 1363–1379, 2008.
[41] A. Ellenberg, A. Kontsos, I. Bartoli, and A. Pradhan, "Masonry Crack Detection Application of an Unmanned Aerial Vehicle," in Proceedings of Computing in Civil and Building Engineering, ASCE2014, 2014.
[42] S. Nishiyama, N. Mianakata, T. Kikuchi, and T. Yano, "Improved Digital Photogrammetry Technique for Crack Monitoring," Advanced Engineering Informatics, vol. 29, pp. 851–858, 2015.
[43] J. Valenca, D. Dias-da Costa, E. Julio, H. Araujo, and H. Costa, "Automatic Crack Monitoring Using Photogrammetry and Image Processing," Measurement, vol. 46, pp. 433–441, 2013.
[44] M. Clark, D. McCann, and M. Forde, "Application of Infrared Thermography to the Non-Destructive Testing of Concrete and Masonry Bridges," NDT & E International, vol. 36, pp. 265–275, 2003.
[45] C. Maierhofer and M. Roellig, "Active Thermography for the Characterization of Surfaces and Interfaces of Historic Masonry Structures," in Proceedings of NDTCE, Non-destructive Testing in Civil Engineering, 2009.
[46] F. Sham, N. Chen, and L. Long, "Surface Crack Detection by Flash Thermography on Concrete Surface," Insight, vol. 50, pp. 240–243, 2008.
[47] N. Bolourian, M. Soltani, A. AH, and A. Hammad, "High Level Framework for Bridge Inspection Using LiDAR-equipped UAV," in 34th International Symposium on Automation and Robotics in Construction (ISARC 2017), vol. 34, 2017, pp. 1–6.
[48] M. Magrini, D. Moroni, M. A. Pascali, M. Reggiannini, O. Salvetti, and M. Tampucci, "Marine Virtual Environment to Access Underwater Cultural Heritage," in Proceeding of 2016 International Workshop on Computational Intelligence for Multimedia Understanding (IWCIM), October 27-28, 2016, Reggio Calabria, Italy. IEEE, 2016, pp. 1–5.
[49] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardos, "ORB-SLAM: A Versatile and Accurate Monocular SLAM System." in IEEE Transactions on Robotics, vol. 31. IEEE, 2015, pp. 1147–1163.
[50] OpenCV, "Detection of Diamond Markers," 2019, https://docs.opencv.org/master/d5/d07/tutorial_charuco_diamond_detection.html.