3D LOW-COST MULTISPECTRAL UAV SYSTEMS: SURVEY AND ANALYSIS OF TORRESINO DA POLVERE OF SAN MARCO IN BERGAMO

Alessio Cardaci1, Antonella Versaci2, Pietro Azzola1

1School of Engineering, University of Bergamo, Italy
2Faculty of Engineering and Architecture, University of Enna ‘Kore’, Italy
alessio.cardaci@unibg.it – antonella.versaci@unikore.it – pietro.azzola@unibg.it

Commission II

KEY WORDS: 3D laser scanning, digital photogrammetry, multispectral images, cultural heritage, conservation.

ABSTRACT:
The Torresini da Polvere are special constructions built by the Venetian Republic between the 16th and 18th centuries. They are unique architectural structures, characterized by a pyramidal roof, for the preservation of the gunpowder. The torresino of San Marco in Bergamo is one of the best-preserved in the world and was then the subject of detailed research. The powder magazine was measured by professional, ground-based and airborne instruments. The paper shows a comparison between the existing structure and the photogrammetric model obtained by low-cost UAV systems to formulate an accurate estimate, including the pros and cons of different survey systems. It also aims to deepen the aspects associated with 3D reconstruction using multispectral imagery, from the investigation to data processing, to create models with NDVI mapping for the study of the building’s decay caused by biological agents.

1. INTRODUCTION
The practice of survey is increasingly enriching with new contents capable of acquiring information on the thermo-biogeochemical characteristics, the material degradation and both structural and energetic criticalities of the built heritage. It also provides new opportunities to describe and understand the landscape through its natural components, soil health and pollution index. All this data is expressed in various forms, from traditional 2D drawings, to shortened (both conical and parallel) 3D views, 4D video renderings up to the new ‘interactive and immersive’ frontiers of the 5D.
The survey is therefore evolving, proposing not only spatial geometries but also information characterizing matter. The traditional RGB models are now flanked by thermographic and multispectral models able to analyse the existing and providing additional information on temperature, humidity, spectral signatures and, in general, with the physical-chemical characterization of surfaces. Models created on the basis of direct observation, integrated in the measurement with active sensors (3D laser scanning) and passive (digital photogrammetry), ground and unmanned aerial vehicles.

Data processing has progressed steadily since the traditional redesign of simple thematic maps into 3D restitution and image analysis. All this to be able to quickly monitor changes in the built environment and green spaces, to obtain quantitative data on vegetation and tree species and to investigate spontaneous mutations and biodiversity. The territorial and built surveys are confronted today, on the one hand, with the profound change brought about by innovation and the use of complex digital documentation systems and on the other hand, with the need for a multidisciplinary dialogue able to promote in-depth studies in specialized sectors.
The combined use of traditional surveying techniques and remote-piloted aeronautical systems has opened up new scenarios (Rakha & Gorodetsky 2019). Drones enable the capture of images from the zenith and ‘bird’s eye’ views from locations very close to the building: images very difficult to catch by common helicopter for aerial photography and impossible to obtain by satellite remote sensing.

2. LOW COST UAV SYSTEMS: RGB & MULTISPECTRAL IMAGES
Optical sensors generate a digital image through a range of radiation measurements that are returned by colour pixels: a digital image is therefore the display of measured values in the electromagnetic spectrum (ES), a point representing the energy emitted by the bodies. The electromagnetic spectrum is continuous and infinite, but a purely conventional subdivision identifies as an ‘optical spectrum’ (from 1mm to 10nm) only the infrared, visible and ultraviolet fields. The adjective ‘optical’ is related to the ability of the radiation, in these frequencies, to impress properly processed photographic plates. Ultraviolet (UV) to infrared (IR) radiations, although not always visible to the human eye, are therefore able to return images of the thermal state and the degradation of the organic materials of the photographed body. Visible light (betwe-
en 400 nm and 750 nm), so called because it includes all wavelengths perceived by the human eye as a sequence of colours ranging from red to purple, is the most used because it gives a plausible image of reality.

Each digital image is distinguished by three parameters: spatial resolution (the size of the surface represented by a single pixel), spectral resolution (the width of the spectral band that the sensor can acquire) and radiometric resolution (the encoding in number of bits with which each pixel is stored). In addition, they can be divided into panchromatic, when they continuously use the entire band of the spectrum (both optical and visible), or multispectral when they have discontinuous information and related only to some specifications.

The market now offers low-cost sensors for both visible and IR spectrums; however, there is a lack of cheap and easy-to-use instruments for the UV field.

The aerial survey made by Unmanned Aerial Vehicle (UAV) equipped with high-definition cameras - both in the visible and infrared spectrum - allow for a high level of spatial detail combined with the huge speed of data acquisition and analysis. Thanks to the new systems with conventional and multi-spectral sensors, it is possible to carry out more concentrated temporal analyses in the most delicate phenomenological phases. For the study of the natural and/or built landscape, they allow to evaluate the ‘state of health’ of the buildings and the ‘physiological condition’ of the vegetation with non-destructive and touchless measurements. The building’s decay mapping, linked to meteoric and atmospheric actions (biological patina, efflorescence, accumulation of moisture, etc.), can therefore become automatic. These phenomena are indeed conditioned by clearly identifiable spectral signatures associated with some frequencies, as is the case for vegetation systems (Shafi et al. 2019).

Selective analysis not covering the entire visible spectrum, but based on images recorded with several bands thus allows an automatic process of recognition of the decrease. Each band is very small in frequency extension and has a different ability to reflect incident radiation. Spectrum bands considered to be important for monitoring are green (555-580 nm), red (665-700 nm) and near infrared (NIR, 740-900 nm). The evaluation of the data is based on vegetative indices defined by various mathematical combinations of flora reflectance within the different bands of the spectrum.

The two most important indices used are the NDVI (Normalized Difference Vegetation Index) and the Red/NIR (the division between red and near infrared) which combine reflection in the red and infrared bands. The NDVI is positively correlated with the amount of plant biomass per unit area and consequently with the vigour of the crop. The index assumes values between -1 and +1; in particular, from 0.1 to 0.3 indicates a bare or slightly grassed surface (therefore free of efflorescence and biological patina), whereas for plant biomass the index assumes values greater than 0.5. The NDVI index is related to the high Red/NIR values (greater than 1) because they can indicate the presence of patina and surface impurities.

The DJI® Phantom 4 Multispectral Drone is a low-cost drone specifically designed for multi-spectral aerial imagery, particularly suitable for assessing building and vegetation conditions. Thanks to six 2.08 MegaPixel by CMOS 1/2.9-inch sensor (five monochrome for multispectral acquisition and one tri-band for acquisition in the visible spectrum), it is able to provide information in the Red-Edge (RE: 730 ± 16 nm), Near Infrared (NIR: 840 ± 26 nm), Green (G: 560 ± 16 nm), Red (R: 650 ± 16 nm), Blue (B: 450 ± 16 nm) and in space color perceptible to the human eye (RGB: 380 - 780 nm). The quality of RGB information compared to the DJI® Phantom 4 PRO with a 20 MegaPixel by the CMOS 1-inch sensor is therefore much lower in quality and resolution (fig. 1).

The combined use of several drones and sensors is therefore a recommended practice to achieve reliable metric results. Specifically, the studio will compare the differences between the models obtained from shots with both a DJI® Phantom 4 PRO and a DJI® Phantom 4 Multispectral.

![Figure 1. Low-Cost UAV systems: DJI® Phantom 4 Multispectral and DJI® Phantom 4 Pro V2.0](image-url)
3. THE CASE STUDY OF THE TORRESINO DA POLVERE OF SAN MARCO IN BERGAMO

Surveys were performed at different levels, from territory to building. For this reason, a particular case study was selected: the Torresino da Polvere of San Marco in Bergamo, a structure, particularly suitable both for conducting reflections at the architectural scale and for experimenting within a countryside area characterised by a lot of vegetation (fig. 2).

The city of Bergamo, since 1428 the westernmost outpost of the Republic of Venice on the border with the Duchy of Milan, was the strategic crossroads for trade towards the markets of central Europe. A fortress designed to reflect the evolution of war technologies and siege strategies resulting from the discovery of ‘black powder’ and the introduction of firearms. The invention of gunpowder and the use of early artillery mark this crucial moment in the history of the transition from the Middle Ages to Modernity. The Venetian Republic quickly understood their potential. The Arsenale di Venezia was among the oldest factories to produce the powerful blend. It was directly governed by the Council of Ten, the supreme authority of the Serenissima. The deposit took place inside small unusual and curious buildings, the so-called Torri da Polvere. These are unique architectures whose elegant regularity and aesthetic beauty is the result of a cultured design, based on careful and competent geometrical research (Cardaci & Versaci 2022). Although the architect who created this singular typology is unknown, we know that the ‘model’ was presented, in June 1565, to the Council of Ten which decreed that it be replicated “in the dimensions indicated by the experts” (Panciera 2002) in all the domains of both the Stato de Mar and the Stato de Terra.

The construction in Bergamo of a “toresella per meter la motione della polvere a monte dell’aloziamenti alla porta di Sant’Alessandro” (Cappellini 1987) is proven by a contract of enchantment dated December 1580. Between 1580 and 1581 the construction of two towers began (the upper and lower one called San Marco): the first lying along the bastion of Castagneta in the upper part of the ridge, the second in the lower valley near the gate of San Lorenzo. The upper building was completed in 1582 and storage of the first barrels began in 1595. The lower tower, on the other hand, was finished in 1598, but it stayed for a long time without the lead roof. This soon led to the onset of some critical issues: the rain, filtered by the coating too thin and badly folded, forced to remove the dust because it was wet, and partly reduced as mud.

Historical sources document the solution to the roofing problem and attest that in 1612, all barrels for the defence of the city were secure in both depots. The toresini, despite minor drawbacks, have been used without interruption for about two centuries. In 1759 restoration work started, including the removal of the lead roofs, but the work was never completed. They ceased their function following the transformations of the fortress in the 19th century at the behest of the Austrians and they will be used as simple warehouses.

The toresini of Bergamo are now in an almost unchanged condition, as if time had not grieved them with great wounds, incorporations, or radical transformations. Today they look like small buildings without buttresses, both inserted into a bucolic context.

They are made of a thick bag masonry: the outer layer of square stone and the interior of brick contain a stone core fixed with a good quality lime. The roof consists of the same conglomerate - moulded with a high semi-octahedral pyramid - and is covered with sandstone. Above, lead plates were initially placed to render it impervious to rainwater. In the past, they were surrounded by a high wall that isolated and protected them. The interior has a single covered room with a brick pavilion vault. The walls were covered with thick wooden planks to protect them against moisture. The floor, also made of wood, was elevated to create a small, aerated cavity. These architectures, both for their limited size and because they are inserted into a natural environment, therefore allow to combine the documentation of a precious heritage with an interesting experiment with Low-Cost survey systems.
4. METHODOLOGICAL PRACTICE

The survey was preceded by a careful planning of the acquisition phases, performed with the use of active and passive sensors. The building was digitized using 3D laser scanning instruments to create a precise reference model for comparing different photogrammetric artefacts. They were obtained through photos made with low-cost cameras, reconstructing multiple prototypes with different complexity and precision (Bolognesi et al. 2015; Brilakis et al. 2011).

The activities were carried out in just a few hours only thanks to the preparation of a dense network of markers for geo-referencing all the scans and the photo captures into a single reference system. The markers, different in shape and layout so that they could be automatically recognized by the different parameter of the software Agisoft@ Metashape PRO, have made up the GCP (Ground Control Point) and the GCC (Ground Control Constraint) of the system. In particular, the GPCs, four universal collimated manually markers, have allowed both the correct georeferencing and the verification of the error; the GCCs have instead been used as ‘constraints’ of reference for the models within the various photogrammetric software. It was decided to compare the models based on common points with known coordinates and to delegate the data processing to ICP algorithms for shape registration. The spatial coordinates of the makers were measured with active sensor technology (with a Topcon@ GT 1200 robotic total station). The acquisition and treatment phases were undertaken as follows.

The survey began with the acquisition with a TLS phase difference instrument for reconstruction an efficient and accurate model for the compare with the photogrammetric models. Specifically, 22 scans were performed by a Faro© Focus 150 HDR, positioned at a distance of a few metres from the object. The scans were carried out radially at different altitudes to maximize the coverage of the interested surface. The resolution of the instrument, set in a value 1/4 (a quarter of the maximum precision), ensured an average distance between the points about 2 mm; the ‘x-control’ value was set to have a strong noise reduction and high angular precision (7.7mm/10m). The point cloud registration was performed after recording the individual scans operated by using the proprietary software (fig. 3). A first scanning approach was performed using a pre-registration based on the targets to which the cloud cleaning followed. All points outside the building’s geometry (land sections, trees, etc.) were eliminated; the overlap of areas common to several scans employing ICP algorithms was improved. This gave a cloud of over 38,553,982 points with an average overlap error of 5.82 mm, a maximum value of 14.24 mm and over 48.22% included within the instrumental uncertainty of 4 mm.

The subsequent phases have deepened the photogrammetric study to verify the peculiarities of the models acquired with different sensors; in particular, artefacts were created from three image sequences:

- multispectral: photos captured with the DJI® Phantom 4 Multispectral thanks to five 2.08 MegaPixel monochrome cameras with CMOS 1/2.9-inch sensor in the fields of Red-Edge (RE: 730 ± 16 nm), Near Infrared (NIR: 840 ± 26 nm), Green (G: 560 ± 16 nm).
• RGB: photos captured with the DJI® Phantom 4 PRO V2.0 thanks to a 20 MegaPixel camera with 1-inch CMOS sensor in the entire visible spectrum (400 nm and 750 nm).
• multispectral-RGB: photos captured with the DJI® Phantom 4 Multispectral thanks to the small 2.08 MegaPixel camera with CMOS 1/2.9-inch sensor in the entire visible spectrum (400 nm and 750 nm).

The flight plan was the same for all shots taken with both UAVs and designed with DJI® Fly and DJI® Virtual Flight simulators, reliable and easy-to-use software useful for planning aerial missions to predict ‘what the drone will see’ before its use. The ope-
rations were conducted at the same time on the same day to ensure consistent and uniform lighting. The shooting made it possible to obtain about 380 high quality frames for each camera.

Data processing took place with Agisoft® Metashape PRO software by setting the same workflow for all three sequences. The image alignment was carried out very carefully without reducing and/or sampling the photographs used at their maximum resolution (with a high increase in calculation times). To achieve greater accuracy, the correction of the ‘internal orientation’ parameters of the camera was performed based on the bundle adjustment calculations. The realignment was then performed based on the corrected values. This process slightly improved the result, in the order of 1%; the correction proved to be more effective on the set of RGB images captured with the DJI® Phantom 4 PRO V2.0 and almost influential on the sets of Multispectral and RBG images captured with the DJI® Phantom 4 Multispectral (table 1). An unexpected thing because this technique is instead very effective in terrestrial photogrammetry with full frame cameras. Among the possible reasons is to consider the low quality of the optics (in low-cost sensors with a reduced number of lenses and, often, poor sharpness and brightness). Most importantly, in the case of the multispectral sensor, the existence of six non-axial cameras with very different characteristics was poorly managed by the software which returned unlikely values of the radial distortion and tangential distortion coefficients. These coefficients, fortunately with very low numerical values close to zero, did not affect the final calculation: they still allowed us to evaluate the poor adaptability of the algorithm to monochrome sensors.

The sets, both calibrated and non-calibrated ones, were processed to obtain the Dense Point Cloud, the Depth Maps, the Mesh Model and the Texture Model (fig. 4). Data processing was done with just one software, but, for each application, in different ways. An innovative methodological procedure was used - in addition to the traditional process flow - related to a “filtering” of dense clouds based on algorithms of evaluation of the quality of the points. An integrated treatment based on the classification of points and linked to a semi-automatic machine learning approach of semantic segmentation.

Figure 4. 3D photogrammetric models (by DJI® Phantom 4 PRO V2.0 & DJI® Phantom 4 Multispectral)

Figure 5. Georeferencing error of the models reconstructed from the GNSS positioning data of the two different UAVs

This contribution has been peer-reviewed.
Figure 6. Comparison between the TLS model and photogrammetric models in the same reference system
The study of the results on the characteristics made it possible to ascertain, as was presumed, a better 3D image-based reconstruction from the images of the DJI® Phantom 4 PRO V2.0 compared to those of the DJI® Phantom 4 Multispectral. The model obtained by the RGB sensor has in fact a density of points and a number of meshes ten times higher than those reconstructed by the multispectral sensor. The same RMS (root mean square) values of the re-protection error are about double from each other. The Mean Key Point Size (the value is a mean tie point scale averaged across all projections) is, oddly, the same for all sets (tab. 2).

Another interesting point concerns the geo-referencing of models. A UAV system is equipped with a GNSS antenna that allows us to have metric indications for the orientation, vertical computation and the scale of the models. DJI® Phantom drones do not (theoretically) need topographic support for determining the coordinates of significant markers. The experimentation has instead shown how the global positioning system, in addition to not being accurate, influences the reconstructions (fig. 5).

It is interesting to highlight how this value is higher (around 3%) for the DJI® Phantom 4 PRO V2.0 model compared to the DJI® Phantom 4 Multispectral. This highlights a higher reliability GNSS system of one medium than the other (Table 3).

The quantitative comparison was carried out after scaling the various models on a network of topographically measured targets, integrated with the 3D laser scanner model described above (Mader et al. 2016; Furukawa et al. 2021). It allowed observing how the overlap uncertainty is in the order of one centimetre between the TLS model and the one reconstructed from the RGB images but much higher (with maximum values of almost 10 cm.) with the reconstructions of the images acquired by the multispectral sensor (fig. 6).

5. CONCLUSION

The trial aims to highlight the fundamental role of acquisition and processing to achieve not only qualitative but also quantitative virtual reconstructions. It wants to demonstrate that it is possible to obtain accurate and reliable 3D models even with acquisition with Low-Cost sensors if correctly processed. The effective and correct processing of data is therefore an essential element for a serious intervention of knowledge, documentation and preservation which cannot be ignored (fig. 7).

REFERENCE


Figure 7. The metric model with NDVI mapping reconstructed from multispectral data processing