POLITECNICO DI TORINO
Repository ISTITUZIONALE

3D modeling of industrial heritage building using COTSs system: Test, limits and performances

Original
3D modeling of industrial heritage building using COTSs system: Test, limits and performances / Piras, M.; Di Pietra, V.; Visintini, D.. - In: INTERNATIONAL ARCHIVES OF THE PHOTOGRAMMETRY, REMOTE SENSING AND SPATIAL INFORMATION SCIENCES. - ISSN 1682-1750. - ELETTRONICO. - 42:(2017), pp. 281-288. (Intervento presentato al convegno 4th ISPRS International Conference on Unmanned Aerial Vehicles in Geomatics, UAV-g 2017 tenutosi a Bonn nel 2017) [10.5194/isprs-archives-XLII-2-W6-281-2017].

Availability:
This version is available at: 11583/2691847 since: 2017-11-15T08:57:36Z

Publisher:
International Society for Photogrammetry and Remote Sensing

Published
DOI:10.5194/isprs-archives-XLII-2-W6-281-2017

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

12 October 2024

3D MODELING OF INDUSTRIAL HERITAGE BUILDING USING COTSs SYSTEM: TEST, LIMITS AND PERFORMANCES

M. Piras [a], V. Di Pietra [a]*, D. Visintini [b]

[a] Department of Environmental, Land and Infrastructure Engineering (DIATI), Polytechnic of Torino
corso Duca degli Abruzzi 24, I-10129 Torino, Italy – marco.piras@polito.it, vincenzo.dipietra@polito.it
[b] Polytechnic Department of Engineering and Architecture (DPIA), University of Udine
via delle Scienze 206, I-33100 Udine, Italy – domenico.visintini@uniud.it

Commission VI, WG VI/4

KEY WORDS: Benchmark session, UAV photogrammetry, camera "on-the-job" calibration, oblique images, COTS

ABSTRACT:

The role of UAV systems in applied geomatics is continuously increasing in several applications as inspection, surveying and geospatial data. This evolution is mainly due to two factors: new technologies and new algorithms for data processing. About technologies, from some years ago there is a very wide use of commercial UAV even COTSs (Commercial On-The-Shelf) systems. Moreover, these UAVs allow to easily acquire oblique images, giving the possibility to overcome the limitations of the nadir approach related to the field of view and occlusions. In order to test potential and issue of COTSs systems, the Italian Society of Photogrammetry and Topography (SIFET) has organised the SBM2017, which is a benchmark where all people can participate in a shared experience. This benchmark, called "Photogrammetry with oblique images from UAV: potentialities and challenges", permits to collect considerations from the users, highlight the potential of these systems, define the critical aspects and the technological challenges and compare distinct approaches and software. The case study is the "Fornace Penna" in Scicli (Ragusa, Italy), an inaccessible monument of industrial architecture from the early 1900s. The datasets (images and video) have been acquired from three different UAVs system: Parrot Bebop 2, DJI Phantom 4 and Flytop Flynovex. The aim of this benchmark is to generate the 3D model of the "Fornace Penna", making an analysis considering different software, imaging geometry and processing strategies. This paper describes the surveying strategies, the methodologies and five different photogrammetric obtained results (sensor calibration, external orientation, dense point cloud and two orthophotos), using separately - the single images and the frames extracted from the video - acquired with the DJI system.

1. INTRODUCTION

Nowadays, the role of UAV system in geomatics is daily increasing, in particular there are very different application fields as inspection, surveying and monitoring, archaeology, cultural heritage, environmental data acquisition, etc. (Du, Liu and Du, 2010; Hague, Kung and Suter, 2012; Wallace et al., 2012; Baiocchi, Dominici and Mormile, 2013; Boccardo et al., 2015; Masiero, Fissore and Vettore, 2017). This particular phenomenon is due to main factors: new available technologies (devices, sensors, systems) and new algorithms for data processing.

Starting from technologies point of view, from some years ago there is a very wide use of commercial UAV even COTSs (Commercial On-The-Shelf) system (Austin, 2011). These last systems are UAVs which are usually employed by hobbyist or for fun, but there are some solution, not even very expensive (< 2.000 €), which can be also used for some geomatics application. Moreover, the quality of the COTS sensors installed on the system as digital camera, GNSS receiver, inertial plaftorm are very interesting from the performance point of view.

On the other hand, the Computer Vision algorithms as *Structure from Motion* (SfM) and *Dense Image Matching* (DIM), included in the classical photogrammetric procedures and the integration of sensors and data, have provide comprehensive tools for manage all the aspect of the spatial information science. 3D reconstruction and visualization, spatial analysis, scene interpretation, environmental monitoring and autonomous flight are examples of the widely range of applications.

Since few years ago, the strategy is partially changed due to the use of oblique images. New methods in photogrammetric procedure are based on the use of images acquired from different point of view in a data fusion approach. These UAVs allow to easily acquire such kind of images, giving the possibility to overcome the limitations of the nadir approach related to the field of view and occlusions. The critical aspects, in these cases, are related to scale factor and the quality of the products released. Recent photogrammetric SfM software manages these different data in various ways, and also the semi-automatic procedures of each user can hugely change the results, in terms of camera calibration parameters, point cloud reconstruction and so on.

In order to test potential and issue of COTSs systems, the Italian Society of Photogrammetry and Topography (SIFET) has organised the SIFET Benchmark 2017, inviting national and European university, research institute, professional and private companies to participate in a shared experience for the evaluation of UAV surveying and data processing.

Scientific tests by distributing data to participants and by evaluating their obtained results is a well-known and consolidate activity, also in the photogrammetric field, as well as in the particular subfield of oblique images, reminding here the important recent tests "*ISPRS/EuroSDR Benchmark on High Density Image Matching for DSM Computation"* (Cavegn et al., 2014) and "*ISPRS Benchmark for multi-platform photogrammetry*" (Nex et al., 2015). Also SIFET has proposed in 2016 a first Italian test on UAV images with the Benchmark *"On the use of UAV images for 3D reconstruction: a joint experience among users"* (Mancini et al., 2016). The SIFET benchmark of 2017 (from now on simply "SBM2017") is called *"Photogrammetry with oblique images from UAV: potentialities and challenges"* and it was aimed to analyse the potential of these systems, defining the critical aspects and the technological challenges of UAV oblique images, comparing distinct approaches and software. The case study is the "Fornace Penna" in Scicli (Italy), an inaccessible monument of industrial

architecture from the early 1900s, whose inside space can be wholly surveyed only by aerial close range photogrammetry. The datasets have been acquired from three different UAVs system, with different payload, camera sensors and configuration of flight/acquisition. In particular, Parrot Bebop 2 quadcopter, DJI Phantom 4 quadcopter and Flytop Flynovex hexacopter have been used to acquire aerial images and videos in nadir, oblique and horizontal assets. An innovative aspect is the comparison between the 3D model generated using only the images and the one generated using only the frames extracted by the video. On both case, the same UAV has been used to collect the imaging data, changing the acquisition modality during the flights.
For the completeness of the furnace 3D model, composed by both external and internal wall sides of the building, oblique images and video become fundamental since, by using instead only nadir images, is very hard to reconstruct the various walls composing, in turn, just the "fully" model of the monument.

In the following chapters, the details about the used COTSs UAV systems (section 2), the test site (section 3), and the campaign surveying strategies (section 4) are described. After that, the image processing (section 5) and the obtained results (section 6) are presented: these our elaborations have been carried out onto only images and video acquired by the DJI Phantom 4 (Sabatini et al., 2013) and have been conducted following the same recommendations given to the SBM2017 participants.

2. COTS UAV SYSTEMS DESCRIPTION

As already told, also in the UAV market, from some years ago there is a very wide use of COTSs systems. As remarked in the acronym definition or in the equivalent term "prêt-à-porter", these UAV, from one side, can be immediately used "as is" from (also) a non-expert user but, on the other side, fixed hardware and software commercial configuration does not allow to improve the surveying performances. Anyway, it is interesting to analyse which limits and potentialities characterize such kind of UAVs considering, above all, the consequent image processing by SfM software, where various orientation and modeling strategies can be pursued; off course, this is one of the goals of the SBM2017.
DJI Phantom 4 and Bepop Parrot 2 quadcopters surely are COTSs systems, which main flight features are listed in Table 1:

UAV	PARROT BEBOP 2	DJI PHANTOM 4
Weight	500 g	1.380 g
Size	38,2 x 32,8 x 8,9 cm	48 x 48 x 19 cm
Max ascent speed	21 km/h	22 km/h
Max descent speed	60 km/h	14 km/h
Max speed	60 km/h	57,6 km/h
Max altitude	150 m	120 m
GNSS mode	GPS/GLONASS	GPS/GLONASS
Gimbal controllable pitch range	from –90° to 0°	from –90° to +30°
Max flight time	25 min	28 min
Radio control	Tablet and smartphone	Tablet and smartphone
Max transmission distance	2 km (Wi-Fi: 300 m)	3,5 km CE 5 km FCC

Table 1: Main features of COTs UAV Bepop 2 and Phantom 4.

Regarding instead image and video sensors, their principal characteristics are reported in Table 2, where one can notice how these COTSs systems allows nowadays the acquisition of high-resolution images, 14 and 12 Mpixel respectively, on the same 1/2,33" sensor. Concerning the 16:9 achievable videos, Phantom 4 has definitely better performances, shooting up to UHD (4K) movies (8 Mpixel), with respect to the only FHD of Parrot 2.

UAV	PARROT BEBOP 2	DJI PHANTOM 4
Imaging sensor	1/2,33"	1/2,33"
Objective	Sunny Fisheye	94° FoV
Focal length (optical)	1,83 mm ?	3,61 mm
Maximum format	3.320x4.096 pixel	4.000x3.000 pixel
Image acquisition	Single, Timelapse	Single, Timelapse, HDR
Video acquisition	FHD 1.920x1.080 p.	from HD 1.280x720 p. to UHD 4.096x2.160 p.
SD card	8 GB (internal)	6 GB (micro SD)
File format	JPG, RAW, MP4	JPG, RAW, MP4, MPEG

Table 2: Main imaging features of Bepop 2 and Phantom 4.

Since this paper is focused on DJI Phantom 4 imaging data, single images of Fornace Penna here processed have 4.000x3.000 pixel, while 30 Hz video sequences have full HD 1.920x1.080 pixel resolution. Yet from these values and considering that acquisition flights had similar geometries, a very different pixel *Ground Sample Distance* (GSD) arises from images and from video, as will be later better explained. For sake of simplicity, from now on when we consider any data of the single images, it will be simply called "images" while, referring to frames anyway extracted from video sequences, the term "videos" will be used.

3. THE TEST SITE

The data collection for the SIFET Benchmark 2017 (SBM2017) was made on the Fornace Penna (Figure 3), an important industrial heritage building located in Scicli (Ragusa, Sicily, Italy). This vast industrial plant (86,8 m long, 25,2 m wide, 14,8 m maximum height) was built between 1909 and 1912 according to the wish of Baron Guglielmo Penna on the design of Eng. Ignazio Emmolo, who also construct and direct the plant. The furnace, used for the production of bricks and tiles, was severely damaged by a fire in 1924; the roof, the slabs and all the wooden parts were destroyed and this caused the end of the activity.

Figure 3. Fornace Penna in Scicli: view from an UAV image acquired for the test (left) and inaccessible inner space (right).

The surviving structure today are not many, the fire and the weariness of time have left only the exposed masonry in ashlar blocks. Furthermore, the degradation due to lack of maintenance has aggravated the situation of the building, which today is at risk of structural collapse. Despite the state of degradation of the structure, the municipality shows a high interest in the recovery of this important industrial heritage. The Authority for Cultural Heritage puts the monumental bond in 2008 and in 2016 the Court of Ragusa puts the property under seizure.
In order to carry out any kind of consolidation/restoration activity on the furnace and a future reuse of spaces, accurate knowledge of the actual state of the building is necessary. *Terrestrial Laser Scanning* (TLS) or photogrammetric surveying techniques, compared in Gonizzi Barsanti, Remondino and Visintini (2012), generally satisfies this need. In our case, the risk of structural collapse and the inaccessibility of the site due to seizure, require instead different solutions to acquire spatial information.
In this context, aerial close-range photogrammetry represents a useful tool for acquiring geometric measurements of the object without risk for the personal safety and without overcome the authority access limitations. In fact, the absence of the furnace's

wooden roof permits to flight over the building and to acquire the images of the internal spaces needed to accomplish the photogrammetric procedure. In this way, *Digital Surface Model* (DSM) and digital orthophotos could be provided to extract information on the structural walls and terrain elevation profile. Moreover, thanks to the capability of the new algorithms of SfM to process oblique images and the directional cameras mounted on the UAVs, is possible to collect also information on the interior walls and occluded parts that otherwise would not be visible from the classic nadir acquisition.
This aspect of the surveying has generated interest from SIFET scientific committee, who saw in the furnace an excellent case study to assess the potential of aerial photogrammetry exploiting oblique images. Gathering the availability of three UAV systems, the "SBM2017 working group" has been constituted with experts from different Italian universities and professionals. In February 2017, this team hence performed the UAV photogrammetric surveying of Fornace Penna, over an area of about 20.000 m^2.

4. DATA ACQUISITION

4.1 Terrestrial data acquisition (and processing)

In the furnace area, SBM2017 working group carried out a huge terrestrial measurement campaign, with Leica GS08 GPS receiver, Leica TS02 e TCRP1201 total stations, and Faro X330 TLS. In particular, a topographic network of 20 vertexes around the furnace was performed: these ground points, materialized by 50x50 cm plastic chessboards (targets) were clearly visible in UAV images. From such vertexes, by means of double/triple intersections, were surveyed the position of 23 20x20 cm paper chessboards glued onto the walls and of 37 natural points in the block edges. The overall scheme is reported in Figure 4, were also targets 18 movable target for TLS registration are drawn.

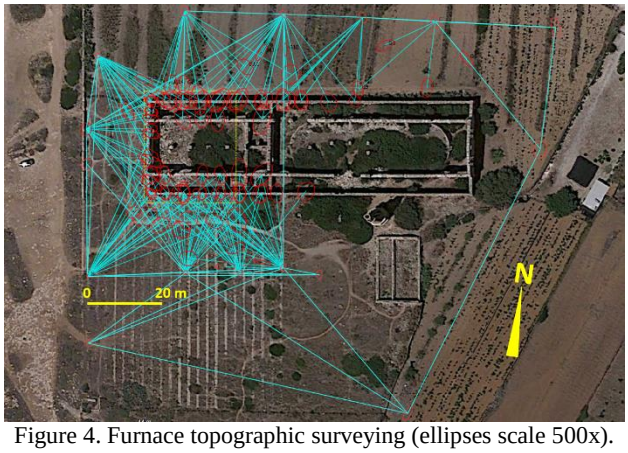

Figure 4. Furnace topographic surveying (ellipses scale 500x).

A XYZ local reference system was suitably adopted (Figure 5 at left) and the test area has been bounded in 25x30x16 m, in the western part of the furnace. The 12 clouds acquired by TLS, once registered (Figure 5 at right), have constituted a unique cloud of of 104 Mpoints as "truth" to evaluate the clouds from UAVs.

Figure 5. Test area for SBM2017: bounding box and XYZ local reference system (at left), TLS global cloud (at right).

XYZ coordinate of 20 ground and 23 wall target points, suitably documented, were made available to the SBM2017 participants for the orientation step, while those of the natural control points, displaced in higher part of the walls, remained unknown to them.

4.2 UAV imaging data acquisition

Relating to the test area only, the nine different flight, acquired by manually piloting the DJI Phantom 4, are here described:
Nadir/oblique/horizontal image acquisition (totally 266 images):

- block 1 (79 images): five longitudinal (along the furnace length X) nadir strips from 25 m height (e.g. Figure 6 on left),
- block 2 (60 images): six longitudinal 45° oblique strips from 25 m height (e.g. Figure 6 on right),
- block 3 (83 images): eight transversal (Y direction) 45° oblique strips from 25 m height (e.g. Figure 7 on left),
- block 4 (15 images): three horizontal strips of the North façade only, acquired from "near" distance,
- block 5 (14 images): three horizontal strips of the West façade only (e.g. Figure 7 on right), acquired from "near" distance,
- block 6 (15 images): two horizontal strips of the South façade only, acquired from "near" distance.

Figure 6. Image examples: nadir (left) and 45° oblique (right).

Figure 7. Image examples: 45° oblique transversal (left) and horizontal (right), this last from a strip on a single façade.

45° oblique video acquisition (totally 3'41"):

- video 1 (2'59"): six longitudinal strips from 25 m, potentially 5.370 different frames (e.g. Figure 8 on left),
- video 2 (25"): two "near" strips of North façade, 750 frames,
- video 3 (17"): one "far" strip of South façade, 510 frames (e.g. Figure 8 on right).

Figure 8. Frames from video examples: 45° oblique longitudinal (left) and from a strip on a single façade (right).

The pixel GSD is the fundamental data in nowadays SfM photogrammetric processing, having the same role of the early concept of "scale of the image" (or scale factor) S_i in analogical photogrammetry. Despite of the image format, any significant change of GSD/S_i implies a variation in the accuracy of XYZ 3D coordinate computed from xy 2D images, although SfM multi-image approach anyway improves the achievable results.

In this sense, the nadir and horizontal acquired images have computable GSD values and potential surveying accuracy, while this is quite impossible for 45° oblique ones, as can be noticed already from previous figures. For nadir or horizontal images, the GSD values are simply given from the ratio between the relative height H or the distance D [m] and the focal length c [pixel]. Off course, the larger is the focal length and/or the lower is the height (distance), the smaller will be the GSD. Having c fixed by the UAV sensor system, the height have to be planned and fulfilled.

SBM2017 idea was to warrant a nadir GSD of around 1 cm for all three UAV systems but, unfortunately, this was not possible for Parrot Bepop 2 images (dataset 1) having a very short focal length, a part their enormous radial deformation for which can be classified as "spherical images". Available information about focal length are pretty incoherent: image EXIF data report values of 1,1 mm (real) and 6 mm (equivalent to 35 mm), while the manual reports 1,83 mm, a value that looks like more realistic since, transformed in pixel, agree with the mean focal length, 1.405 pixel, estimated just thanks to SBM2017. Anyway, with this little focal length (very wide angle), the fight height assuring 1 cm GSD should be around 14 m, then impossible, since lower the furnace height (14,8 m). An elevation of 25 m was adopted.

From technical specifications of DIJ Phantom 4 images (dataset 2), the declared equivalent focal length is 20 mm, while in image EXIF data is reported the optical value of 3,61 mm onto a 1/2,33" CCD sensor 6,08x4,56 mm: considering that such 4:3 sensor has a diagonal of 7,6 m, the equivalent value should be 3,51 mm. This time, starting from such values, the corresponding pixel focal length are similar, 2.250 or 2.314: it means that, to assure a 1 cm GSD, the flight height should be of 22-23 m. In truth, it was fixed to 25 m (the same of Bepop 2), to flight adequately far from the furnace walls, having a maximum elevation of nearly 13 m in the test area, where the western façade is 1,8 m lower than eastern one, due to wall breakdown, well visible in Figure 6 and 7 on right. Just these higher parts of the monument are closer to the camera and here the GSD value is practically 5 mm only.

Referring to Flynovex system, it mounts a Sony Alpha a6000 professional camera with a focal length c (real) of 16 mm onto a APS-C CMOS sensor 23,5x15,6 mm, equivalent to 4.000 pixel. In this case, from the adopted height of 50 m, the GSD value is 1,25 cm; only single images were acquired (dataset 3).

The situation dramatically changes if we consider oblique images, where the orthogonal distances from the sensor are extremely variable, ranging from less than 15 m even to 50 m (see e.g. Figure 6 right, Figure 7 left and Figure 8 left).
It is important to remember that GSD of frames from video at least doubles since resolution is halved respect to single shots.
Concerning horizontal images, namely "the most possible oblique ones", we can state that UAV systems make possible the *photogrammetrist' dream* to acquire images always parallel to the object of interest. As very well-known, by taking terrestrial images the walls appears tilted or a great part of ground is depicted: now instead, by easily elevating the point of shot, the ideal situation of normal images become reality (e.g. Figure 7 right) also for high buildings to surveying. For such kind of images acquired in front of a single façade furnace (block 4, 5, 6 and video 2) the reported term "near" means that distance D is less than 10 m and was chosen "on-the-fly", as well as the "stereoscopic base" for the images, simply to warrant a triple image overlap. Corresponding GSD are in the order of 5 mm only for external (closest) walls for images, while is obviously larger for video. To this regard, 45° oblique video 3, unfortunately acquired only from "far", is the worse imaging data source.

5. SFM IMAGE PROCESSING

As stated before, all the imaging data acquired by three different UAV systems were distributed to the SBM2017 participants (universities, professionals and companies) in order to be processed from them with whatever photogrammetric software and to test their obtained "package results". In particular, the organizing committee has required to send back this results:
1. Estimated parameters of image calibration,
2. Coordinates from images of targets and natural control points,
3. *Dense Points Cloud* (DPC) from images,
4. 1:100 scale plan, in form of digital orthophoto,
5. 1:50 scale South façade elevation, in form of digital orthophoto.

Furthermore, a report were requested, with the description of the pre-processing steps, the used images, software and hardware, the orientation strategies and the time of processing.
This section presents only the data processing made by the authors with the same DJI Phantom 4 dataset of the SBM2017 and, as suggested, not mixing images and video frames.

The 266 4.000x3.000 pixel images acquired by the DJI Phantom 4 UAV have been processed by means of the well-known commercial software Agisoft PhotoScan (version 1.2.6.2834), following the standard workflow proposed:
1. Feature matching: namely, in order, detection of points stable under viewpoint and lighting variations, generation of a descriptor based on local neighbourhood, use of the descriptors to detect and define corresponding across the images (similar to the well-known SIFT approach).
2. Bundle-adjustment: solving for camera internal and external orientation, by starting from approximate camera locations and refines them later using a bundle-adjustment algorithm.
3. Dense surface reconstruction: this step is particularly important and, first, makes use of suitable DIM algorithms to produce the DPC and, later, exploits pair-wise depth map computation or multi-view approach to build the DSM.
4. Texture mapping: by parametrizing the obtained DSM, possibly cutting it in smaller pieces, and then blends source photos to form a texture atlas of the DSM.

For each step shown above, Agisoft PhotoScan permits to set some parameters to downscale image resolution or limits the number of faces generated in the triangulation procedure. In particular, the alignment (relative orientation) step and the DPC reconstruction were executed with "Medium accuracy", whose consequence is an image downscaling by factor of 4. Also the quality of the triangulated DSM was fixed to "Medium": in this case, the maximum number of polygons is limited to 1/15 of the number of points of the DPC.

As described in paragraph 4.2, the acquired dataset is composed by images taken with different viewing direction with respect to the building. Some images are taken with a nadir image configuration, other in an oblique camera direction and other horizontals, namely normals to the furnace's façades. This complicate image configuration could cause problems with some algorithms of Structure from Motion.
In fact, following the standard procedure, not all the 266 images have been oriented during the alignment steps. In particular, some images that shots the South and West portion of the furnace, do not have sufficient overlapping with the adjacent images and also have different acquisition geometry respect to the other images. To overcome this problem, the first step is to align different "chunks" of images subdivided according to the direction of the camera. At the end of the alignment, all the chunks are relative oriented in its own local reference system. Then, each chunk is aligned with the others thanks to a point based algorithm implemented in PhotoScan. Once that all the

chunks are in a common local reference systems, the estimated external orientation parameter are extracted and used as input for a new process with all the images computed in the same chunk. This is similar to using the GPS information stored in the EXIF file of the images. Thanks to this approach, is possible to align all the images and so to effectively test the software capability to process oblique images. After the bundle-adjustment, where the XYZ target coordinates have been exploited, the images are calibrated (result 1) and target and natural point (re)computed (result 2). Later, the DPC is computed and exported in ply format (result 3), the plan orthophoto in jpg format and its reference frame information (result 4) and, finally, the South façade orthoimage (result 5). Data processing was realized using a PC with Windows10 Pro ×64, Intel Core i7 3.60 GHz, 128 GB of RAM. The total data process required 5 hours and 54 minutes.

The same procedure has been applied to the dataset of videos acquired by the DJI Phantom 4. In this case, a simple Matlab routine was used to extract frames from the various mpeg videos. In total, "only" 214 frames (on 6.300 potentially) were extracted and processed with the same parameters adopted in the images elaboration. At the end, once solved the bundle-adjustment (result 1 an 2), the DPC of the furnace (result 3) and the textured 3D model are again obtained; from these, orthoimages of the plant (result 4) and the façade (result 5) are created in jpg format. All the process onto such frames required 3 hours and 47 minutes.

Summarizing, the same five results required to the SBM2017 participants have produced by us, either for images or for videos.

6. RESULTS

Only our results are here explained and commented, anyway taking into account mean results or the best ones of the SBM2017. To this regard, the comparison can be done onto ten different "results package" for images and six for videos: moreover, among such sixteen results, five have been obtained by using other SfM software as Pix4D Mapper (two results), 3DF Zephyr Aerial (two results) and PixySFM (one result).
The results will be presented firstly for images, later for videos.

6.1 Evaluation of image calibration

As known, in photogrammetry the availability of the internal (intrinsic) parameters of a camera is a fundamental requirement, sharing the cameras world in *metric* and *non-metric* if this is fulfilled or not. However, nowadays such binary categorization is outdated since, by means of SfM software, whatever camera can be easily calibrated, even with an "on-the-job" approach.
The availability of "well estimated parameters" gives anyway the reference on which evaluate any calibration process: for our DJI image and video sensor, unfortunately, we have not such data.
Estimated calibration parameters are here reported:

- centre coordinate: C_x = 1.967,8 pixel, C_y = 1.496,6 pixel;
- focal length: $F_x = F_y$ = 2.311,3 pixel;
- Brown' curve distortion coefficients: K_1 = 0,005286, K_2 = –0,016469, K_3 = 0,007936; P_1 = 0,000328, P_2 = –0,000059;
- skew factor = 0,429218.

Values of centre coordinate are not distant to half part of sensor size (2.000 and 1.500 pixel) and also the estimated focal length is in agreement with attended mean value of 2.300 reported in paragraph 4.2. These four values are moreover fully equivalent with the means from those estimated in SBM2017, that are:

- C_x = 1.967,5 ± 12,0 pixel, C_y = 1.496,7 ± 3,7 pixel;
- F_x = 2.329,2 ± 1,5 pixel, F_y = 2.329,5 ± 1,4 pixel.

Regarding the coefficient of radial distortion, an interesting evaluation is given not on the K_i values, but on the corresponding curve, reported in Figure 9. Nine comparable curves represent the estimated distortion of DJI Phantom 4 images, not to high since maximum values are 4-5 pixel for the maximum radii of 2.500 pixel. Particular name of curves arises from SBM2017 aim to assure the anonymity results: anyway, "PS" stands for "PhotoScan", "PD" for "Pix4D", "Ze" for "Zephyr", "Px" for "PixySFM" for which the calibration has been estimated out with two different process (1 and 2). Bold grey curve 2_I_PS is related to our K_i estimated values and it appears coherent with the other.

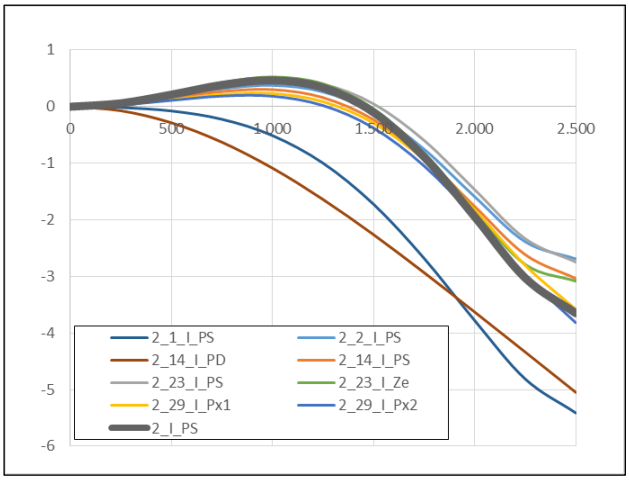

Figure 9. Estimated curves of radial distortion for image sensor.

The corresponding results estimated for video sensor are instead:

- centre coordinate: C_x = 951,2 pixel, C_y = 546,8 pixel;
- focal length: F_x =1.153,8 pixel, F_y = 1.146,3 pixel;
- Brown' curve distortion coefficients: K_1 = –0,004951, K_2 = –0,048203, K_3 = 0,029155; P_1 = 0,000133, P_2 = –0,000251.

Expected centre coordinate are 960 and 540 pixel, while no information about focal length is given; anyway also now our estimations agree with means values coming out from SBM2017:

- C_x = 947,2 ± 17,0 pixel, C_y = 548,7 ± 18,3 pixel;
- F_x = 1.165,5 ± 21,0 pixel, F_y = 1.169,1 ± 17,1 pixel.

It must be stressed the higher variance among estimations of video internal parameters. The same consideration can be done for the distortion curve coefficients that now is also characterized by not negligible deformation, in order of 20 pixel for the maximum radii. As visible in Figure 10, our estimated curve 2_V_PS, represented by a bold green line, lies among other SBM2017 curves, which variability in now higher, also 10 pixel.

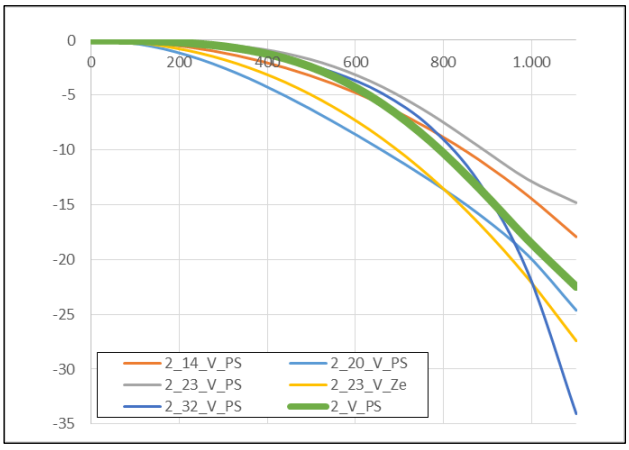

Figure 10. Estimated curves of radial distortion for video sensor.

6.2 Evaluation of target and control point coordinate

The analyses of XYZ coordinates of object points estimated from images is a classical method to evaluate the estimated images

exterior (external) orientation. Discrepancies among photogrammetric and topographic coordinates on target points used for orientation is a first accuracy evaluation: mean values, reported in Table 11, give an order idea of such discrepancies that are acceptably restricted in 4 cm (on walls) or 6 cm (on ground).

		ΔX [m]	ΔY [m]	ΔZ [m]	ΔTOT [m]
13 ground targets	mean	–0,006	0,004	0,029	0,062
	st.dev.	0,039	0,031	0,043	0,034
15 walls targets	mean	–0,007	0,001	–0,011	0,036
	st.dev.	0,029	0,024	0,008	0,016

Table 11. Errors on ground/wall targets: images vs topography.

In spite of this, much more interesting is the analysis onto 33 natural control points, not anyway involved in the orientation process and whose topographic coordinates were unknown in SBM2017; the resultant discrepancies are reassumed in Table 12.

	ΔX [m]	ΔY [m]	ΔZ [m]	ΔTOT [m]
Max	0,057	0,068	0,064	
Min	–0,046	–0,097	–0,084	
mean	–0,002	–0,004	–0,014	0,047
st.dev.	0,027	0,034	0,027	0,023

Table 12. Errors on natural points: images vs topography.

From mean and standard deviation values of Table 12 we can say that the 3D-error magnitude on unknown natural points is 4,7 ± 2,3 cm: this result is satisfactory, taking into account that a global orientation of all 266 has been pursued. Following instead the strategy to separately orient the single blocks acquired for each façade, e.g. blocks 4, 5 and 6, surely allows to reduce these errors, being such images constrained onto "few" points; working in this way, some SBM2017 participants obtained 2-4 cm 3D-errors.

Considering now the same values coming out from processing of video frames, Tables 13 and 14 summarize the errors.

		ΔX [m]	ΔY [m]	ΔZ [m]	ΔTOT [m]
13 ground targets	mean	*–0,007*	*0,015*	*0,035*	*0,052*
	st.dev.	*0,043*	*0,052*	*0,054*	*0,064*
15 walls targets	mean	*0,015*	*–0,009*	*0,012*	*0,044*
	st.dev.	*0,049*	*0,054*	*0,061*	*0,060*

Table 13. Errors on ground/wall targets: video vs topography.

	ΔX [m]	ΔY [m]	ΔZ [m]	ΔTOT [m]
Max	*0,072*	*0,058*	*0,094*	
Min	*-0,069*	*-0,107*	*-0,091*	
mean	*0,022*	*-0,014*	*-0,034*	*0,047*
st.dev.	*0,107*	*0,134*	*0,107*	*0,123*

Table 14. Errors on natural points: video vs topography.

As can be straight away seen, errors dramatically increase: these was forecastable simply remembering the quite halving of resolution or, conversely, the doubling of pixel GSD. Moreover, the presence of a strong radial distortion requires care in its estimation and any not modelled deformation introduces errors.

6.3 Evaluation of dense point cloud

The comparison between the DPC obtained by images versus the DPC coming from TLS surveying, computed by means of the well-known CloudCompare software, gives a comprehensive indication of the correctness of the performed UAV surveying. Figure 15 shows such 3D distances, computed only in the interval 0÷50 cm, coloured from the minimum value (blue) to maximum one (red). In the evaluation of these "cloud to cloud" distances, it must be taken into account that some parts of the internal walls are not fully measurable from TLS, since it is not allow to enter in the furnace, while these are measurable from above by UAV.

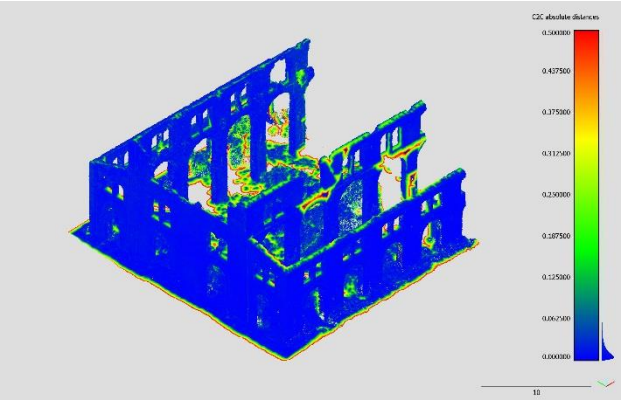

Figure 15. Distances between DPC from images vs DPC by TLS.

Also the distances along X, Y, Z directions have been computed, where X and Y values are the more significant: standard deviations are 0,074 m on X, 0,073 m on Y, 0,056 m on Z, and 0,100 m as 3D cloud-to-cloud difference. Figure 16 displays the X, Y, Z, 3D distances of eleven groups, in order: nine SBM2017 participants, our values, and the SBM2017 mean; a part some cases, similar X and Y values around 7 cm have been obtained.

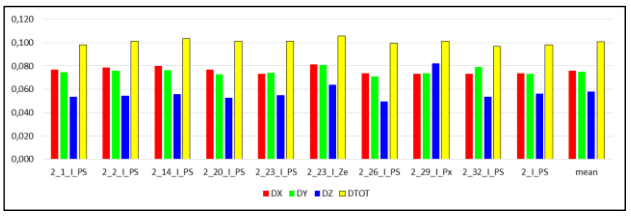

Figure 16. X, Y, Z, 3D distances of various DPC from images.

For DPC coming out from video, the cloud comparison evidence bigger differences, around 8-9 cm for X,Y directions.

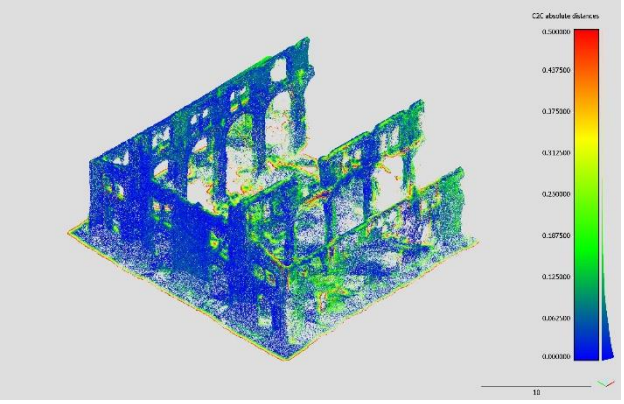

Figure 17. Distances between DPC from video vs DPC by TLS.

Considering five SBM2017 results, a higher variability arises, anyway stating an increasing in the "UAV vs TLS" distances.

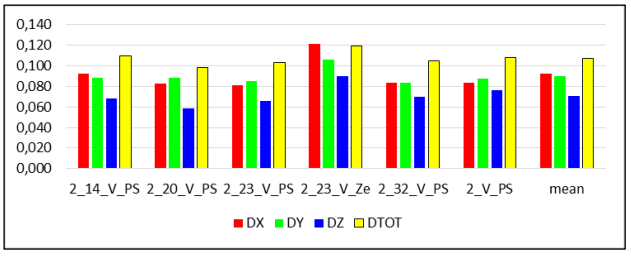

Figure 18. X, Y, Z, 3D distances of various DPC from video.

6.4 Evaluation of 1:100 scale orthophoto plan (25x30 m)

The quality of an orthophoto is determined from a lot of aspects, sharable in those depending from the image internal parameters (result 1) or external ones (result 2), and those from the DSM. It must be underline how DSM strictly depends, in turn, from the dense points cloud (result 3): coming out from the same images, accuracy/error of result 1 and 2 are so doubly crucial. On the other side, it is also true that, from the same DPC, very different DSMS can be produced by changing the parameters adopted in the modeling. Further influence on the final radiometric orthophoto quality is given from the efficiency of the blending algorithm in the multi-image texturing step.
For all these reasons, the orthophoto quality evaluation is not a trivial task, becoming quite impossible to carry out in numerical form. The idea to compare the orthophoto to test versus a "perfect orthofoto" by considering differences in the pixel colors, is not realistic in our case, being an external open area with grass and vegetation, wall irregularity, changing shadows and so on. Concluding, the evaluation has been expressed by means of an ordinal decreasing scale of judgements *A*, *B*, *C*, and so on.
We can state that orthophoto is the last product of a (semi-)black box flow-chart began with the UAV acquisition of the images: it is then essential to anyhow evaluate it, as a sort of "final verdict". In the follows, our orthophoto will be compared with other three from SBM2017, choosing among those having various votes.

Regarding specifically the 1:100 scale plan of 25x30 m test area, the required *"Resampled Sample Distance"* (RSD) has been fixed (again) to 1 cm, so to create an image of 2.500x3.000 pixel that can be printed by 300 dpi at 1:100 scale with a good quality. The quality of the obtained plans from images in Figure 19 looks very good, although here represented at a very low scale (1:595!).

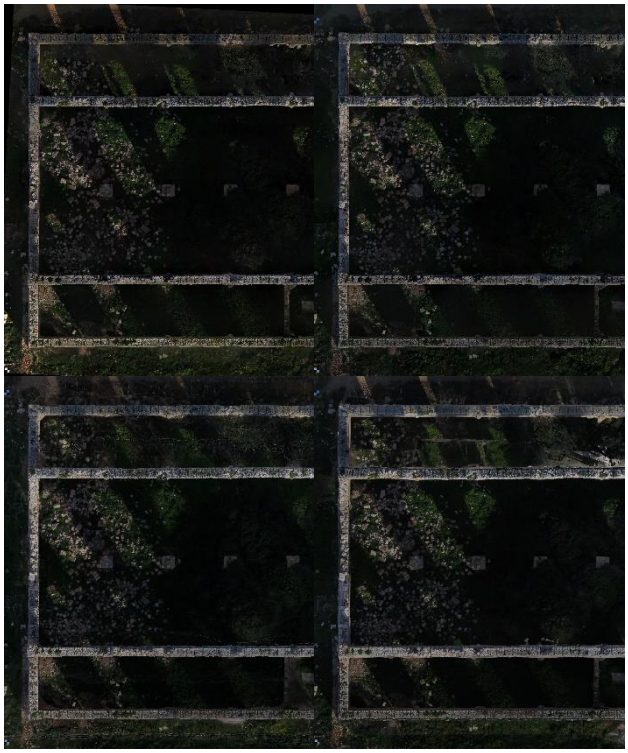

Figure 19. Comparison among different plans from images: from us (a), from SBM2017 participants (b, c, d).

Observing with more care Figure 19, in the third (c) and forth (d) image, is visible the effect of the texturing errors in the area of the northern "nave" of furnace, worse for plan d respect to plan c. Plan a and b votes are *A*, while *A*-- for plan c, and *B* for plan d.

Evaluating now the plans obtained from videos, grouped in Figure 20, the quality significantly degrades, as a logical consequence of what already seen in paragraph 6.1, 6.2 and 6.3. Besides, having acquired only oblique video, with frames like in Figure 8 on left, the resampling quality cannot be excellent.

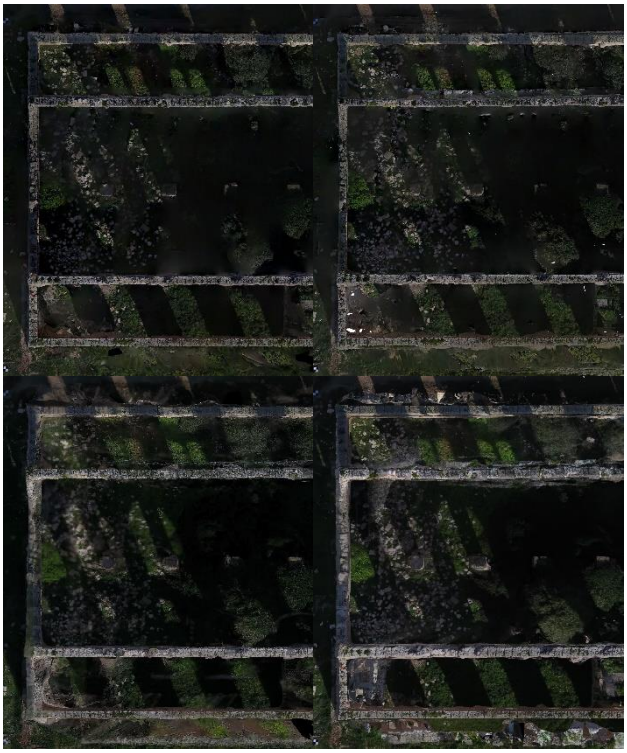

Figure 20. Comparison among different plans from video: from us (a), from SBM2017 participants (b, c, d).

In our orthophoto (a) some central parts have a sort of fog and the vertical wall wrongly appears in lower left corner (vote *B*). Such errors are common with plan b (vote *B*), where the same corner has some empty pixel: this furnace part is definitely not well depicted in the oblique videos. Third (c) and forth (d) plan have walls and pilasters bubbled and other errors (vote *C*).

6.5 Evaluation of the 1:50 scale orthophoto elevation (23,5x13 m)

The 1:50 scale representation of the South façade involves a 23,5x13 m vertical portion of the test area, with RSD fixed to 5 mm, so to produce now an image of 4.700x2.600 pixel that can be printed by 300 dpi at 1:50 scale with a good quality.
Once again, the quality of the obtained elevations from images in Figure 21 looks more than good, although represented at a scale (1:560) eleven times lower! The DSM is now more complex since it must have holes, in correspondence of building windows or gates, not so easy to be created, making texturing problems.

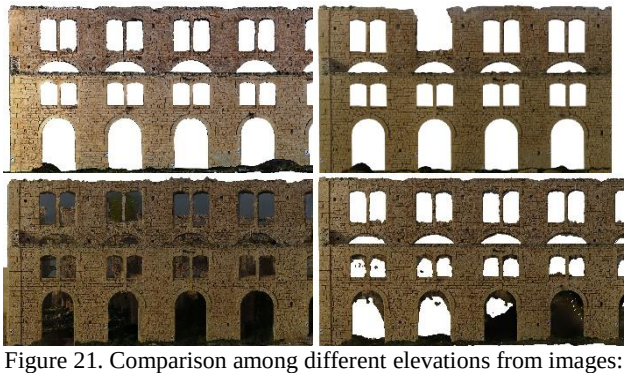

Figure 21. Comparison among different elevations from images: from us (a), from SBM2017 participants (b, c, d).

Four elevations reported in Figure 21 have not great difference in the masonry walls: the second (b) is a little bit better (vote *A*) than the others (vote *A*-). Windows and gates areas are instead correctly/incorrectly represented, but this a minor problem, anyway solving by photo-editing software. A nice situation is given from the lacuna in correspondence of the second upper mullioned, that is correctly lacking in the ortophoto b, while is filled in the others with the blocks of the corresponding part in the North façade, although from internal and not external side.

The same South façade orthophoto, but starting from video, has not been created but with inadequate quality and it is not here presented. Our result is similar to those obtained from two SBM2017 participants (on ten having these dataset), reported in Figure 22, and evaluated in the order with *B* and with *C*.

Figure 22. Comparison among different elevations from video.

The motivation of this flop is not due to software limits or to user inabilities, but to the acquired video! Seeing again Figure 8 on right, a 1.920x1.080 frame producing these bad orthophotos, the GSD value a posteriori computed results there about 4 cm. The statement of "worse imaging data source" (of SBM2017) written at the end of section 4, can be now fully understood.

Concluding these two paragraph regarding the orthophoto production with a final remark, the evaluation process here presented is simply macroscopic: further investigations have to be conducted analysing in detail at great scale, at least at the true 1:100 and 1:50 scale, the various plans and elevations.

7. CONCLUSIONS

This paper presents limits and performances of the 3D modeling of an industrial heritage building by UAV COTS system used in a benchmark carried out on a Sicilian damaged 1900s furnace.
In particular, the SfM processing of nadir, oblique and horizontal images/video acquired by the low-cost DJI Phantom 4 system is here presented and discussed. Imaging sensors have good performances (12 Mpixel) for 4:3 format images, that could be kept enough similar for UHD video (8 Mpixel), despite the loss for adapt to 16:9 format; unfortunately our final video had a only HD (2 Mpixel) resolution and this great decay gave worse results. In other words, the fundamental importance of the pixel GSD is confirmed: no matter if COTS digital sensor are not calibrated, since this can be adequately done "on-the-job" by SfM tools.
Other final consideration regards the geometry of most part of our images, (unconventionally) acquired with a "non nadir" direction: if the main surveying goal are the building walls and not its roof, as in this case, such oblique images are essentials, though arising GSD variations and possible occlusions.
The obtained final surveying results, dense point cloud and orthophotos, have a good quality if coming out from images, but not the same from (low-resolution) video. Anyway, new analysis on other our results, as well as of relating to the whole SIFET Benchmark 2017, have to be suitably developed and refined.

ACKNOWLEDGEMENTS

The authors thanks the Italian Society of Photogrammetry and Topography (SIFET) to have given them the opportunity to work on this very interesting test on surveying by UAV systems.
An acknowledge to the participants of SIFET Benchmark 2017 working on this dataset, whose some results are here presented in anonymous form: Marco Scaioni - Polytechnic of Milano, Marco Severin - SoluTOP Pavia di Udine, Yuri Taddia - University of Ferrara, Francesco Mancini - University of Modena and Reggio Emilia, Francesca Fissore - University of Padua, Donatella Ebolese and Marcello La Guardia - University of Palermo, Salvatore Troisi - Parthenope University of Napoli.
A special thanks for the cooperation goes to the other components of the SBM2017 working group (the "Dream Team") Martina Ballarin, Giuseppe Furfaro, Aldo Guastella and Andrea Scianna.

REFERENCES

Austin, R., 2011. Unmanned aircraft systems: UAVS design, development and deployment (Vol. 54). John Wiley & Sons.

Baiocchi, V., Dominici, D., Mormile, M., 2013. UAV application in post-seismic environment. *The International Archives Photogrammetry. Remote Sensing and Spatial Information Sciences*, XL-1 W, 2, pp. 21-25.

Boccardo, P., Chiabrando, F., Dutto, F., Tonolo, F.G., Lingua, A., 2015. UAV deployment exercise for mapping purposes: Evaluation of emergency response applications. *Sensors*, 15(7), pp. 15717-15737.

Cavegn, S., Haala, N., Nebiker, S., Rothermel, M., Tutzauer, P., 2014. Benchmarking high density image matching for oblique airborne imagery. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-3, pp. 45-52.

Du, D., Liu, L., Du, X., 2010. A low-cost attitude estimation system for UAV application. In: *Control and Decision Conference (CCDC), Chinese, IEEE*, pp. 4489-4492.

Gonizzi Barsanti S., Remondino F., Visintini, D., 2012. Photogrammetry and laser scanning for archaeological site 3D modeling - some critical issues. In: *Proc. of the 2nd Workshop on "The New Technologies for Aquileia"*, 948, pp. B1-B10.

Hague, D., Kung, H. T., Suter, B. 2006. Field experimentation of cots-based UAV networking. In *Military Communications Conference, 2006. MILCOM 2006. IEEE*, pp. 1-7.

Mancini, F., Castagnetti, C., Rossi, P., Guerra, F., Ballarin, M., Visintini, D., Garraffa, A., Stecchi, F., Aicardi, I., Grasso, N., Noardo, F., De Luca, V., Ratto Cavagnaro, M., Caroti, G., Piemonte, A., Pinto, L., Santise, M., Albano, V., Curuni, M., Cerratti, M., Abate, G., 2016. Risultati ottenuti nell'ambito della sessione "benchmark" - uso di immagini UAV per la ricostruzione 3D: esperienze condivise tra utenti (in Italian), *Bollettino della SIFET*, 3, pp. 1-9.

Masiero, A, Fissore, F., Vettore, A., 2017. A Low Cost UWB Based Solution for Direct Georeferencing UAV Photogrammetry. *Remote Sensing*, 9, 414, 21 pages.

Nex, F., Gerke, M., Remondino, F., Przybilla, H.J., Bäumker, M., Zurhorst, A., 2015. ISPRS benchmark for multi-platform photogrammetry. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, II-3/W4, pp. 135-142.

Sabatini, R., Ramasamy, S., Gardi, A., Salazar, L.R., 2013. Low-cost sensors data fusion for small size unmanned aerial vehicles navigation and guidance. *International Journal of Unmanned Systems Engineering*, 1(3), 16.

Wallace, L., Lucieer, A., Watson, C., Turner, D., 2012. Development of a UAV-LiDAR system with application to forest inventory. *Remote Sensing*, 4(6), pp. 1519-1543.