UAV multi-image matching approach for architectural survey in complex environments

UAV multi-image matching approach for architectural survey in complex environments / Calantropio, Alessio; Chiabrando, Filiberto; Einaudi, Davide. - ELETTRONICO. - (2020), pp. 98-105. ((Intervento presentato al convegno D-SITE, Drones - Systems of Information on culTural hEritage. For a spatial and social investigation tenutosi a Online.

ALESSIO CALANTROPIO, FILIBERTO CHIABRANDO, DAVIDE EINAUDI

Polytechnic of Turin, Department of Architecture and Design (DAD), Laboratory of Geomatics for Cultural Heritage (Lab G4CH),
Turin, Italy

alessio.calantropio@polito.it, filiberto.chiabrando@polito.it,
davide.einaudi@polito.it

ABSTRACT

Photogrammetric survey using UAV shows nowadays some aspects, especially related to the optimization of the operative workflows and best practices, that still need to be investigated. This work concerns the use of a small UAV for the documentation of an historical architectural complex, in which space constraints arises. The adoption of a rapid mapping workflow using frames extracted from videos is discussed, together with the exploitation of an automatic procedure for the acquisition of 360° shots, used for ensuring the minimum required overlap for a reliable and accurate image orientation.

Keywords:
UAV photogrammetry, 360° acquisition, Cultural Heritage, Image Orientation.

UAV MULTI-IMAGE MATCHING APPROACH FOR ARCHITECTURAL SURVEY IN COMPLEX ENVIRONMENTS

1. INTRODUCTION

The use of UAVs (Unmanned Aerial Vehicles) is nowadays a widely adopted survey methodology, due to the wide diffusion of new and performant platforms and because of the easiness of handling the gathered raw data (images). Unfortunately, the enormous spread of image-based survey methods and techniques has not been supported by an equivalent diffusion of the validation strategies of the generated product, especially related to the quality control of the metric precision and accuracy of the georeferenced spatial data. The risk nowadays is to produce and disseminate 3D and 2D metric products with a not verified (and therefore probably not correct) geometric and radiometric information.

A special attention must be focused on the proper use of non-metric cameras, as these devices needs to be tested and compared, in terms of reliability and quality of the generated products, with more consolidated instruments and survey techniques.The tests carried out in this research have been performed in order to evaluate a consistent workflow for image-based rapid-survey of architectural objects, with complex shapes and in narrow space conditions. This work concerns the use of a lightweight UAV for the documentation of an historical architectural complex, in which space and time constraints arises. A rapid mapping workflow, consisting in using frames extracted from videos has been tested in the framework of this research. In view of speeding up the survey phases, without sacrificing the quality of the final 3D products, it has been experimented an automatic procedure for the acquisition of 360° shots, used for ensuring the minimum required overlap for a reliable image orientation. Photogrammetric 3D reconstructions using terrestrial 360° acquisition have been already conducted (Perfetti, Polari, Fassi 2018; Barazzetti, Previtali, Roncoroni 2018), and there are also very few examples of the use of spherical camera on light or very light UAV (Calantropio, Chiabrando, Einaudi, Teppati Losè 2019). In this research only normal frame images acquired with a 360° approach have been considered (without stitching the captured images in a single spherical panorama).

2. CASE STUDY

The *Borgo Medievale* is a reconstruction of a medieval hamlet, which was conceived for the 1884 Italian Exhibition. The Borgo Medievale were designed by Alfredo D'Andrade to symbolize the medieval Italian styles (Pagella 2011). Surrounded by walls and protected by a tower with a drawbridge, the Borgo is composed of a Church (the object of the presented tests), some houses, ecc. All these structures are related to real medieval architectures of two Italian regions, Piedmont and Valle d'Aosta. The buildings of the Borgo are located directly on a 7 meters wide royal road; its morphology fits very well the aims of this research. At first, an area of the Borgo that can be considered difficult to survey using a traditional photogrammetric workflow has been selected, in order to gain feedbacks concerning the potentialities of the proposed approach.

The selected area is the little square in front of the Church, a narrow space faced by 4 buildings. The Church is, moreover, higher than 17 meters and is in front of an 11 meters tall building, at a relative distance of only 7 meters. It is easy to understand that this morphology introduces different issues in planning and executing a terrestrial survey, since it requires an elevated number of control points, apart from other significative issues related with obtaining metric data related to the higher portions of the taller buildings.

Nevertheless, performing UAV flights in this area is particularly difficult, due to the low reliability of the on-board GPS in a narrow space surrounded by buildings, and issues in connecting different parts of the photogrammetric block related to facades of different objects. The main problem is related to the reconstruction of the relation between the different buildings, given the difficulty in guaranteeing the necessary image overlaps. The possibility of using drones embedded with a 360° camera sensor is to be avoided in this condition, because of lack of a reliable GNSS signal and the necessity to disable the sense-and-avoid system to attach the external camera. Moreover, the use of an UAV capable of carrying an additional device, would have required (due to the weight of the drone and the related operation typology) the closure of the site, as provided for in the Italian Civil Aviation Authority amendment "Remotely Piloted Aerial Vehicles Regulation" (ENAC 2019).

Figure 1. Ortophoto of the Borgo Medievale in Torino; highlighted in red the little square, object of the survey carried out in this research.

3. MATERIAL AND METHODS

Due to issues related to space constraints of the case study, the platform DJI Spark has been selected for carrying out this survey. DJI SPARK is a lightweight UAV with a take-off mass below 300g, considered "unoffensive" by ENAC (Ente Nazionale Aviazione Civile) thanks to the use of a 3D printed lightening kit. The use of this UAV has already been tested for the generation of 3D model with a photogrammetric approach (Adami, Fregonese, Gallo, Helder, Pepe, Treccani 2019), using frames extracted from video (Calantropio, Chiabrando, Rinaudo, Teppati Losè 2018; Carnevali, Ippoliti, Lanfranchi, Menconero, Russo, Russo 2018) and in this research using the automatic 360° shots function for ensuring that the necessary minimum overlap is respected. The employed UAV, especially for its unoffensive capability, allows a fast, low-profile and easy to authorize operation, ensuring a safe and stable flight without altering the VPS (Visual Positioning System). Before discussing in depth the followed workflow and the subsequent results, it is necessary to define that the main aim of the test presented in this article is to reduce the time required for carrying out the data acquisition and processing, as this procedure could be easily transferred to emergency scenarios (Giordan et al. 2017; Murphy et al. 2008). Because of this, all the acquisition have been performed in a semi-automatic way, using frames extracted from videos acquired following parallel photogrammetric stripes, with the optical axis of the camera perpendicular in respect to the façade of the buildings (Einaudi 2019). This granted the reduction of the acquisition time, as well as allowed the pilot to focus only on the flight operation, as an intelligent extraction of the frames at an ideal frequency will satisfy the optimal overlap.

The accuracy of the reconstruction has been evaluated using a set of 10 GCPs (Ground Control Points) and of 4 CPs (Check Points) measured using a traditional topographic approach with total station (LEICA Viva TS16 – angular accuracy Hz and V 0.1 mgon – distance range 1.5 m to 3500 m). Moreover, a C2C (Cloud to Cloud) comparison has been performed between the dense cloud obtained following the photogrammetric process, and a single TLS (Terrestrial Laser Scanning) cloud, used as a ground truth.

As reported in the introduction, the automatic 360° acquisition of frame images needs some preliminary considerations; because this semi-automatic procedure is functional to the generation of spherical panoramas, it is necessary to acquire multiple shots in order to ensure the optimal overlap between the images, that has to be acquired using (almost) the same camera position. This procedure can be executed by an UAV using an integrated function that autonomously acquire the necessary number of images for ensuring a good result of the final product.

There are different mobile applications (apps) that allows to plan the automatic 360° acquisition of frame images; for the purpose of this research two apps have been testes: LITCHI and DJI GO 4; according to the camera of the DJI SPARK the optimal number of frame images to be acquired is 46.

With the aim of defining a fast-photogrammetric acquisition workflow, the followed approach has been used to acquire videos and subsequently extract the necessary number of frames. Considering that during the fights the speed of the UAV was (more or less) constant, and that the video has been recorded at 1080p with a framerate of 30 fps, frames have been extracted each 2 second (1 frame each 60 frames). It is important to note that the frame extraction procedure lead to a loss of the EXIF information (aperture, focal length, etc.). This is very important to consider especially when important information (typology of the employed sensor and its pixel size) are used during the camera calibration and image orientation steps.

For the generation of the sparse point clouds, the software Agisoft Metashape has been used, setting the quality at medium (downscaling the original images by factor of 4 – i.e. 2 times by each side).

4. RESULTS

The objective of this test is to demonstrate the potentialities of the automatic 360° acquisition for enhancing the orientation of frames extracted from videos.

Figure 2. Different view of the position of the three automatic 360° acquisitions. Highlighted is the one used in the dataset B.

In order to evaluate the potentialities offered by this approach, three different datasets have been used; one dataset contained only the frames extracted from the video, without the use of the automatically acquired 360° frames (Dataset A – 305 images). The second dataset (Dataset B – 351 images) contained all the frames extracted from the video plus one of the acquired 360° frames (+46 images). A third dataset (Dataset C – 443 images) contained again all the frames extracted, together with all the 3 acquired 360° frames (+138 images). Concerning the dataset A, about half of the images haven't been aligned during the orientation phase. As expected, this is because there is not enough overlap between the two sub-datasets employed (West building and East building), this led to the generation of only one façade, arbitrary chosen from the software. One strategy to avoid this problem is to process the two sub-datasets in a separate way, or to introduce some images to cover lack of overlap.

Following this last consideration, a 360° acquired set of frame images has been added (dataset B), and this allowed to align all the frame; therefore, the whole scene has been successfully reconstructed. In order to evaluate if only one 360° acquisition (46 images) is enough to enhance the alignment, a third dataset (C) containing a total of three 360° acquisitions (138 images) has been processed. As it is possible to observe in the following table 2, the residuals on GCPs and CPs are lower (almost the half) when more than one 360° acquisition is used. We can say that the results obtained in the third dataset C reports an accuracy compatible with the restitution of products at a scale of 1:100 (precision of 2 cm). For obtaining these results, no pre calibration of the sensor has been done, but only a self-calibration during the processing phase. In order to validate the results, a C2C comparison has been done between the photogrammetric dense point cloud (Dataset C) and a TLS acquisition of the same architectural space, used in this test as a ground truth. The TLS acquisition has been performed using a Faro Focus 330 (CAM2) with a ranging error of ± 2 mm.

The average density (number of points per m2) of the point clouds at 1 and 4 meters starting from the ground level is respectively 48k and 38k points for the TLS point cloud and 21k and 20k points for the

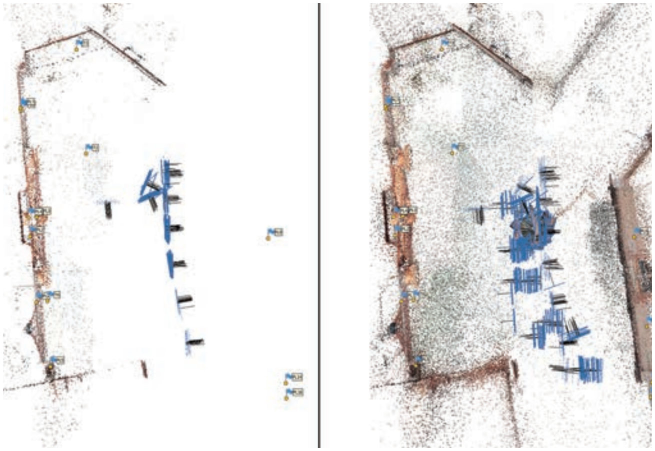

Figure 3. On the left, image aligned in the dataset A (only the Church is successfully reconstructed); on the right, all the images are successfully aligned (dataset B and C).

DATASET	A (Frame only)	B (Frame + 1 360°)	C (Frame + 3 360°)
IMAGES [aligned/total]	160/305	351/351	443/443
N. of tie points	38,091	372,094	505,401
Matching time	2 h 25 m	20 h 00 m	28 h 27 m
Alignment time	1 h 08 m	5 h 15 m	6 h 13 m

Table 1. Information regarding the three processed datasets. The processing has been carried out using the same settings in Agisoft

DATASET	Number of GCPs	X error [cm]	Y error [cm]	Z error [cm]	XY error [cm]	XYZ error [cm]
B	10	2.5	1.4	1.5	2.9	3.3
C	10	1.0	0.6	0.5	1.2	1.3
DATASET	Number of CPs	X error [cm]	Y error [cm]	Z error [cm]	XY error [cm]	XYZ error [cm]
B	4	3.2	2.7	1.2	4.2	4.4
C	4	1.2	1.8	0.4	2.1	2.2

Table 2. Table showing the residuals of GCPs (Ground Control Points) and CPs (Check Points) for the two datasets B and C.

photogrammetric point cloud. In the validation phase, the registration of the TLS and the photogrammetric cloud reported a mean error of 20 mm and a standard deviation of 32 mm. Analysing the results of the C2C comparison for the Church façade, the points with a residual error of ± 1 cm are the 35% of the results, and the ones with the residual error of ± 2 cm are the 85%. Analysing the results of the C2C comparison for the building façade, the points with a residual error of ± 1 cm are the 69% of the results, and the ones with the residual error of ± 2 cm are the 67%. We can thus say that the generated model is validated for a survey at a scale of 1:100.

5. DISCUSSION AND CONCLUSIONS

Thanks to the integration of the adopted techniques, which are the extraction of frames from video on one hand, and the generation of linking images for ensuring the minimum overlap using an automatic function on the other hand, it is possible to generate products useful for the architectural survey, like high-quality orthophotos of the façade and elevations, that allows the correct description of the surveyed architectural heritage. Talking in deep about the façade of the Church, as it is possible to observe from the already discussed results, the level of reached accuracy allows to produce drawings at the scale of 1:100, with the possibility of generating accurate 3D models. Concerning the second façade (the facing building), the same considerations can be done. In order to generate good quality orthomosaic, the followed workflow began with a cleaning operation of point cloud, i.e. to segment it correctly in order to delete wrong reconstructed or not necessary geometries. After the point cloud cleaning and segmentation, the next step is the generation of a 3D mesh, on which the images (after the un-distortion process) are projected for the creation of a texture. For a better result, it has been created a mesh with a high number of triangles, and to furtherly improve the obtained results, the geometries have been refined using an external software of 3D modelling. Then the improved mesh has been imported again in the software Agisoft Metashape, in order to allow a better generation of the applied texture. Following this, an ortoprojection plane has been defined; this has been done using several (at least 3) points surveyed on the façade that

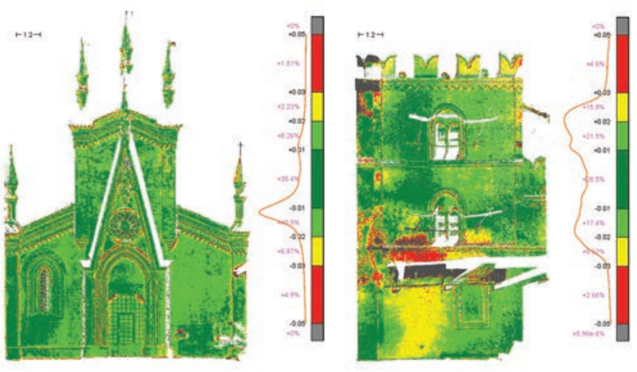

Figure 4. Different C2C analysis concerning the Church (on the left) and the facing building (on the right). The colour shows, for each point of the clouds, the relative distance between the photogrammetric and the TLS point clouds.

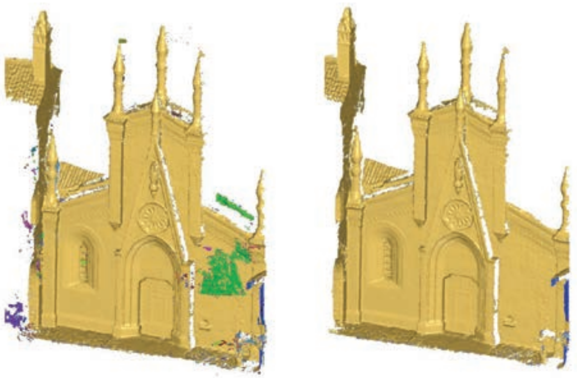

Figure 5. Triangulated mesh before the segmentation and cleaning process (left); final version of the 3D model (right).

allowed the definition of a mean plane with a least square approximation method. After the generation of a correct ortoprojection plane (parallel to the mean plane of the façade) an orthophoto has been generated. The ortophoto, that is the last step of the followed procedure, could be used to produce traditional 2D drawings, useful for a complete documentation of the surveyed object.

According to the achieved tests and obtained results, is possible to state that nowadays the flexibility of the UAV platforms, able to acquire high resolution data in different ways (video and or images) in conjunction with the SfM (Structure from Motion) software that can process different kind of data in a common environment, allow to explore non-conventional surveying strategies that are difficult to imagine some years ago.

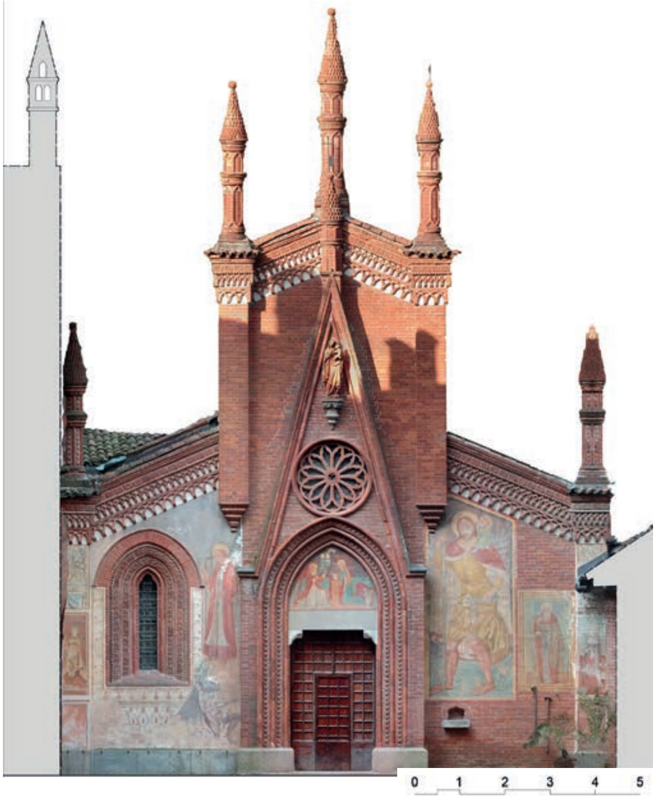

Figure 6. On the left, Orthophoto of the Church, generated following the photogrammetric approach.

One the one hand it is clearly proven that the availability of a large number of images allows to improve the rigidity of the photogrammetric block, on the other hand it is important to underline that, especially in complex scenario, the use of GCPs and CPs allows to define in the correct way the accuracy of the final results.

This is a crucial outcome for an architectural representation, where the scale of the drawing is strictly related to the precision of the photogrammetric process.

Acknowledgements and COI

The authors would like to acknowledge the direction of the Borgo Medievale of the Comune di Torino for their availability during the survey phase of this research. Many thanks to the colleagues Lorenzo Teppati Losè and Giacomo Patrucco for their help in acquiring the data. The authors disclose that a MOU is in force between the authors' institution and the manufacturer of the UAV used. However, the authors declare that the manufacturer had no role in this study, and no conflict of interest exists.

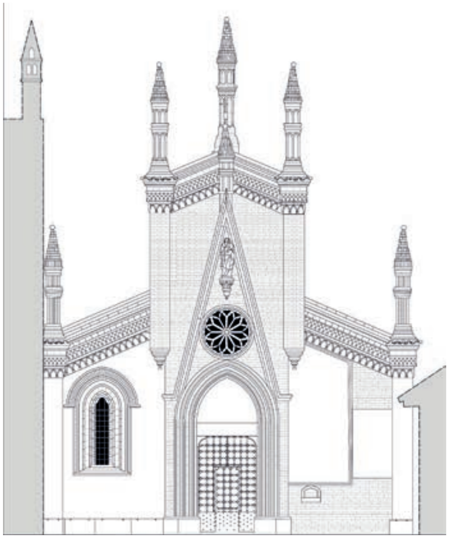

Figure 7. On the right, Traditional 2D drawing (elevation of the Church), produced starting from the previously generated orthophoto.

BIBLIOGRAPHY

Adami A., Fregonese L., Gallo M., Helder J., Pepe M., Treccani D. (2019). *Ultra Light Uav Systems For The Metrical Documentation Of Cultural Heritage: Applications For Architecture And Archaeology. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives [online]*. 29 November 2019. Vol. XLII-2/W17, 2019, pp. 15–21.

Barazzetti L., Previtali M., Roncoroni F. (2018). *Can We Use Low-Cost 360 Degree Cameras To Create Accurate 3d Models? International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives [online]*. 4 June 2018, Vol. XLII-2, pp. 69–75.

Calantropio A., Chiabrando F., Einaudi D., Teppati Losè L. (2019). *360° images for uav multisensor data fusion: First tests and results. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives [online]*. 4 June 2019, Vol. XLII-2/W13, pp. 227–234.

Calantropio A., Chiabrando F., Rinaudo F., Teppati Losè L. (2018). *Use and evaluation of a short range small quadcopter and a portable imaging laser for built heritage 3D documentation. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives [online]*. 10 October 2018, Vol XLII-1, pp. 71-78.

Carnevali L., Ippoliti E., Lanfranchi F., Menconero S., Russo M., Russo V. (2018). *Close-range MINI-UAVS photogrammetry for architecture survey. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives [online]*. 30 May 2018, Vol. XLII-2, pp. 217–224.

Einaudi D. (2019). *Utilizzo di sistemi UAV multi-sensore per la modellazione 3d e la realtà virtuale. Tesi di Laurea Magistrale, Politecnico di Torino*. 28 February 2019.

ENAC, 2019. R*egolamento - Mezzi Aerei a Pilotaggio Remoto [online]. 2019. [Accessed 6 December 2019]*. Available from: https://www.enac.gov.it/la-normativa/normativa-enac/regolamenti/regolamenti-ad-hoc/regolamento-mezzi-aerei-pilotaggio-remoto.

Giordan D., Manconi A., Remondino F., Nex F. (2017). *Use of unmanned aerial vehicles in monitoring application and management of natural hazards. Geomatics, Natural Hazards and Risk*. Volume 8, Issue 1, Taylor and Francis Ltd, 2017, pp. 1-41.

Murphy R., Steimle E., Griffin C., Cullins C., Hall M., Pratt K. (2008). *Cooperative use of unmanned sea surface and micro aerial vehicles at Hurricane Wilma. Journal of Field Robotics [online]*. March 2008, Vol. 25, no. 3, pp. 164–180.

Pagella E. (2011). *Il Borgo Medievale. Nuovi Studi*. Torino: Edizioni Fondazione Torino Musei. ISBN 978-88-88103-83-9.

Perfetti L., Polari C., Fassi F., (2018). *Fisheye multi-camera system calibration for surveying narrow and complex architectures. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives [online]*. 30 May 2018, Vol. XLII-2, pp. 877-883.