

INDOOR PHOTOGRAFOMETRY USING UAVS WITH PROTECTIVE STRUCTURES: ISSUES AND PRECISION TESTS

G. Caroti¹, A. Piemonte¹, I. Martínez-Espejo Zaragoza¹, G.Brambilla²

¹ Dipartimento di Ingegneria Civile e Industriale, Università di Pisa, 56123 Pisa, Italy - gabriella.caroti@ing.unipi.it

² ALPHA Robotix Srl, Via Cisanello 145, 56124, Pisa, Italy - giovanni.brambilla@alpharobotix.com

Commission ICWG III/IVa

KEY WORDS: UAV, Collision-Tolerant Drone, Photogrammetry, 3D Model, safety survey, BVLOS control

ABSTRACT:

Management of disaster scenarios requires applying emergency procedures ensuring maximum safety and protection for field operators. Actual conditions of disaster sites are labelled as “Triple-D: Dull, Dusty, Dangerous” areas. It is well known that in this kind of areas and situations remote surveying systems are at their very best effective, and among these UAVs currently are an effective and performing field tool. Indoor spaces are a particularly complex scenario for this kind of surveys. In this case, technological advances currently offer micro-UAV systems, featuring 360° protective cages, which are able to collect video streams while flying in very tight spaces. Such cases require manual control of the vehicle, with the operator piloting the aircraft without prior knowledge of the status quo of the survey object and therefore without prior planning of flight paths. A possible benefit in terms of knowledge of the survey object could lay in the creation of a 3D model based on images extracted by video streams; to date, widely tested methods and techniques are available for processing UAV-borne video streams to obtain such models. Anyway, the protective cage and the need to use, in these operating conditions, wide-angle lenses present some issues linked to ever-changing image framing, due to the presence of the cage wires on the field of view. The present work focused on this issue. Using this type of UAVs, video streams have been collected in different environments, both indoors and outdoors, testing several procedures for photogrammetric processing in order to assess the ability to create 3D models. These have been tested for reliability based on data collection conditions, also assessing the level of automation and speed attainable in post-processing. The present paper describes the different tests carried out and the related results.

1. INTRODUCTION

Use of UAV-borne photogrammetry is to date a widespread methodology in surveying (Pajares, 2015, Martínez-Espejo Zaragoza, et al., 2017). In order to produce orthophotographs of the territory, these aircrafts can operate by means of autopilot systems, which allow for an orderly and homogeneous image collection, with multiple possibilities as regards on-board camera types. In architectural surveys, UAVs are at times manually operated, e.g. for façades in urbanized areas (Eschmann et al., 2012) or infrastructures such as bridge intradoses (Ajayi et al., 2017). In case of surveys of small or indoor spaces, preventing user access due to safety issues, or in presence of obstacles, traditional UAVs are unable to operate, either for their very size or the lack of suitable anti-collision devices, and also because flight stabilization systems rely upon GPS signals, which are often unavailable indoors. For these reasons, the ability to withstand in-flight collisions is crucial.

The present study used a UAV designed for video inspection purposes for surveying in the latter conditions. Rather than performing metric surveys, the goal of this aircraft is to collect video streams flying in environments preventing Visual Line Of Sight (VLOS), with remote manual operation relying on subjective vision (BVLOS – Beyond Visual Line Of Sight). In this case, the aircraft has to withstand, thanks to a protective cage, any collision with the surrounding objects, while the arrangement of the camera must be stabilized and independent from the protective cage. The present paper describes some tests checking the ability to obtain metric data using frames extracted from the video streams collected by this kind of UAV.

2. MATERIALS

2.1 UAV Elios

Elios, produced by the Swiss firm Flyability, is a quadcopter UAV equipped with a revolving cage, providing full protection from collisions and drops (Figure 1). It has been designed “to crash and keep on going”, i.e. to retain its stability following any in-flight collision.



Figure 1 – Elios UAV

This feature is ensured by disjoining the protective cage from the inside frame of the UAV along three axes by means of gimbals. Unlike professional UAVs, designed to fly outdoors, Elios lacks any geolocalization systems linked to GNSS signals. The absence of such navigation systems rules out the planning of automated missions, with fixed speed and paths, so that flight is always manually operated. Its primary purpose is to perform video inspections in industrial environments and explore confined spaces, all the while ensuring operator safety. The ability to operate in extreme conditions makes this UAV a very versatile tool, both for routine operations, such as inspecting industrial and civil structures for maintenance planning, and in response to major disasters, such as earthquakes or landslides, when operator access is forbidden until securing of the area.

Flyability_Elios_Brochure-LW (2017) lists all technical specifications of the UAV. In Table 2, a list of the main technical specifications is reported.

Dimension	Fits in <400mm sphere
Video	FHD (1920 x 1080) at 30 fps, good low light performance, recorded on board and streamed to pilot and camera operator
Horizontal field of view	130 degrees
Vertical field of view	75 degrees
Total vertical field of view	215 degrees (considering payload up/down rotation)
Resolution	When flying in contact with a surface, Elios can gather close-up images with a sub-millimeter resolution of 0.2 mm/px.
Lighting system	Elios includes a LED array for artificial lighting of the scene

Table 2. Main technical specifications of Elios UAV

Notably, standard equipment for Elios provides an ultra-wide angle lens, consistently with the need to frame wide portions of the object even when flying in its close proximity.

2.2 Data Sets

The following data sets have been included in the test base:

- Data set IMG_1: video stream collected indoors (Great Hall, School of Engineering, Pisa University) with both natural light and on-board LED array, flying at 3m on average from the survey object (GSD approx. 3mm). 55 images have been extracted from the stream.
- Data set IMG_2: video stream collected outdoors (Section of Pisa urban walls – La Cittadella) with natural light only, flying at 10m on average from the survey object (GSD approx. 10mm). 23 images have been extracted from the stream.
- Data set IMG_3: video stream collected indoors (test panel bearing markers with known coordinates), with natural lighting only. The stream provided 13 images.
- Data set IMG_4: video stream collected indoors (rain water tank, Certosa di Pisa, Calci), with lighting from the on-board LED array only.
- Dataset IMG_5: video stream collected indoors, in a confined and inaccessible space (Filter chamber of a steam turbine), with lighting from the on-board LED array only.
- Data set HDS_1: TLS survey of the same scene as in IMG_1. Leica C10 ScanStation, resolution approx. 15pts/cm².

- Data set HDS_2: TLS survey of the same scene as in IMG_2. Leica C10 ScanStation, resolution approx. 25pts/cm².

3. METHODS AND RESULTS

All processing in this research has been carried out by Agisoft's PhotoScan software in bundle adjustment mode with no manual input of tie points.

3.1 Extracting images from video streams

This kind of UAV usually features uneven speed and irregular flight paths. As a consequence, there is no standard settings for the extraction of frames from video streams, which instead requires adapting to the different operating conditions.

In any case, frame extraction takes place by setting a sampling interval (0.5s, 1s, 2s, etc.). Direct experience has shown that, as a consequence of the sudden movements due to the almost continuous collisions, it is advisable to set the sampling interval low enough to avoid inadequate overlapping between consecutive images, if necessary oversample the stream.

3.2 Case 1 – Dataset IMG_1 processing

The first test took place on 55 images extracted from a video stream, collected in an indoor environment, of a brick wall section also including a stone band carrying an inscription. The flight scene was naturally lit but the on-board lighting system was turned on anyway. Due to the presence of the freely rotating protective cage, every image displays an ever-changing portion of the scene, resulting in an ever shifting, rather than static, frame. Although the cage is dark coloured, use of the on-board lighting system causes some reflections and consequently overexposure, with the cage wires appearing mostly white with some shaded areas (Figure 3).

Firstly, image quality parameter was calculated. Its values, ranging between 0.66 and 0.74, are better than 0.5, which experimental evidence has shown to be the lower limit for effective use. The procedure of image alignment, without manually adding any tie point, automatically detected 626 tie points, successfully aligning only 20% of the images. Qualitative analysis of reciprocal camera position and of the point cloud resulting from the automatic alignment shows a completely wrong calculation of external camera orientation parameters. The cause of this anomaly is quite obvious upon close examination of the tie points, which are mostly detected on the protective cage of the UAV (blue dots in Figure 3) rather than on the static survey scene.



Figure 3 – Sample image from dataset IMG_1 with superimposition of software-calculated tie points.

3.3 Case 2 – Dataset IMG_2 processing

Data set IMG_2 was planned to avoid the need for the additional lighting provided by the on-board LED array. In this case, in fact, image regions depicting cage elements were darker and more homogeneous in colour. The analysis of the quality parameter of the 23 images extracted from the video stream yielded values ranging between 0.65 and 0.72. The automatic alignment procedure detected 3770 tie points, which enabled to align all the images in the data set. Notably, most of the software-detected tie points in this data set lie outside the protective cage (Figure 4).



Figure 4 – Sample image from dataset IMG_2 with superimposition of software-calculated tie points.

Following a qualitative analysis of mutual camera position, the alignment appears to be correct; therefore, the model has been scaled and provided with a reference system, thanks to a set of 10 Ground Control Points (GCPs), whose coordinates have been extracted from the point cloud provided by data set HDS_2. This step has been performed via a simple 7-parameters Helmert transformation on double points. GCPs have therefore been excluded from the bundle adjustment process. The model obtained from data set IMG_2 has been then compared against that derived from HDS_2, which acted as a reference, on 10 additional Control Points (CPs), evenly spread across the survey object, resulting in a standard deviation = 0.042m. Obviously, when using the images to generate the texture for the model, this would also include portions of the protective cage (Figure 5).



Figure 5 – Presence of cage projection on the textured model.

A further check on the alignment step of these images provided the use of masks to delete the portions of the images depicting the protective cage.

3.4 Image masking

Although PhotoScan software provides some tools for image masking, it is advisable to pre-process the images in photo editing software, which include more advanced tools. Image masking was performed according to two methodologies, the

first providing user operation and the second relying on software procedures.

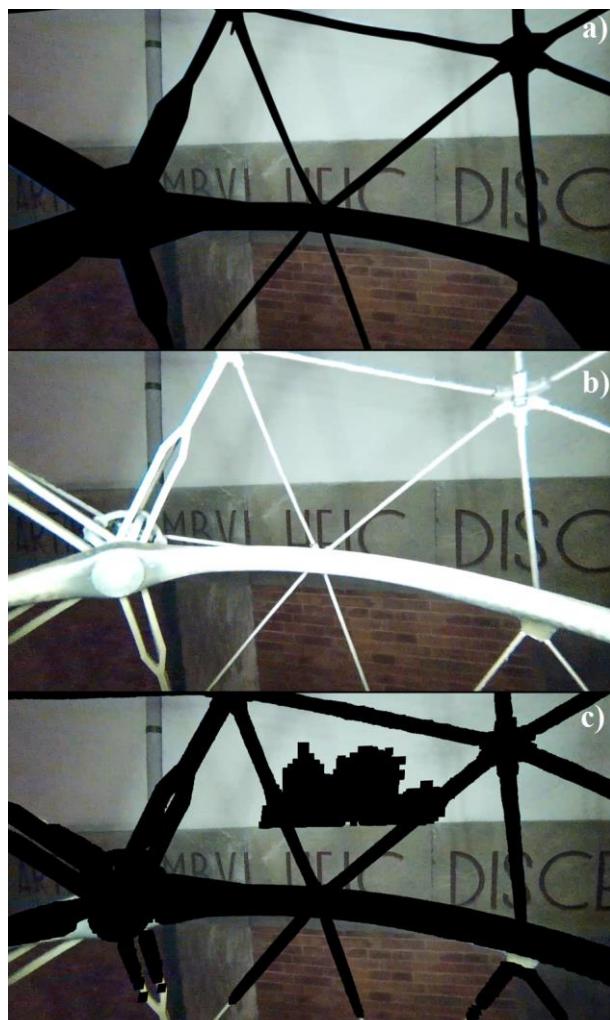


Figure 6 – a) User-defined image mask; b) Original image; c) Batch-automated image mask.

The first kind of mask achieves the best results in terms of image portions to exclude from further processing (Figure 6a) but has high requirements in relation to operator time (approx. 3 minutes per image). Automatic masking can be achieved via batch processing, which, by means of script instructions, applies firstly a colour-based mask with a certain sensitivity, and then expands it by about ten pixels, in order to include also the edges, whose colour is halfway between the cage elements and the background. Although this automated procedure cuts down time requirements by several magnitude orders, on the other hand it does not ensure complete masking of unwanted features and can sometimes filter out useful information (Figure 6c).

3.5 Case 3 – Dataset IMG_1 with mask processing

Processing of data set IMG_1 using user-defined masks achieved image alignment by means of 1209 tie points, automatically detected on the survey object only. Anyway, upon providing the model with a scale factor and a reference system by means of 10 GCPs, whose coordinates have been extracted from data set HDS_1, and cross-checking both models on 10 additional CPs, the resulting error is in the 400pxl range. This is also quite obvious from the check of the standard deviation on

the coordinates, ranging around 60cm. Taking into account the relevant radial distortion inherent to ultra-wide angle lenses, a further methodology test provided image pre-processing in order to minimize distortion, based on a pre-calibration of internal orientation parameters, and use of these “undistort” images as input of the alignment procedure (Teo, 2015, Balletti, et al., 2014, Hastedt, et al., 2016).

3.6 Camera pre-calibration

In order to obtain a set of camera pre-calibration parameters, data set IMG_3 provided a video stream of a flat panel carrying several targets, whose coordinates are defined with a 1mm precision (Figure 7).

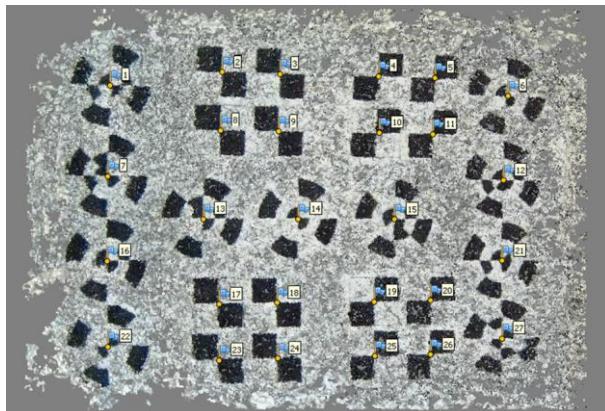


Figure 7 – Pre-calibration target panel.

Bundle adjustment of IMG_3 images, using all targets as tie points with known coordinates, yielded the results reported in Table 8.

f	1027.73 pixel
cx	-54.8143 pixel
cy	-72.7791 pixel
k1	-0.349736 pixel ⁻²
k2	-0.129576 pixel ⁻⁴
k3	-0.0212924 pixel ⁻⁶
p1	0.0930875 pixel ⁻²
p2	-0.00123303 pixel ⁻²

Table 8 – Camera pre-calibration parameters

Upon definition of these parameters, a custom Matlab script provided new data sets, containing images derived from the original ones corrected for distortion (Table 9).

distort images dataset	corresponding undistort images dataset
IMG_1 with mask	IMG_1_UNDIST
IMG_2 with mask	IMG_2_UNDIST
IMG_4 with mask	IMG_4_UNDIST
IMG_5 with mask	IMG_5_UNDIST

Table 9 – Undistort images dataset

Subsequent processing using PhotoScan software used these new data sets including “undistort” images and masks, without any calibration restraint, therefore entirely demanding modelling of residual distortion to the software.

3.7 Case 4 – Dataset IMG_1_UNDIST processing

Data set IMG_1_UNDIST was processed following the same procedures described for Case 3 (Figure 10). Images were aligned, with the software automatically detecting 448 Tie Points. Upon providing the model with a scale factor and a reference system, by means of the same GCPs used for Case 3, standard deviation on the same 10 CPs was 3cm (27pxl), i.e. at least one order of magnitude better than the previous one.



Figure 10 – Final 3D model of IMG_1_UNDIST.

3.8 Case 5 – Dataset IMG_4_UNDIST

Among the data sets used for the present work, IMG_4_UNDIST provides the most accurate simulation of an emergency survey, particularly for cases in which operators have no access to the survey area and must operate UAVs based on BVLOS view rather than on VLOS (Figure 11).



Figure 11 – UAV operator.

Besides, the survey object is a masonry tank (Figure 12) lacking any lighting system, where the relevant quantity of dust and small residue of building materials (mortar and crumbled bricks) and the rotor-induced turbulence often led to suspending solid particles in mid-air (Figure 13).



Figure 12 – Dark tank and UAV-borne lighting system.



Figure 13 – Turbulence and dust.

Applying the already-tested methodology, i.e. defining masks to filter the protective frame out of the images and using images corrected for distortion, it was possible to align these images and achieve a partial 3D model of the object (Figure 14).

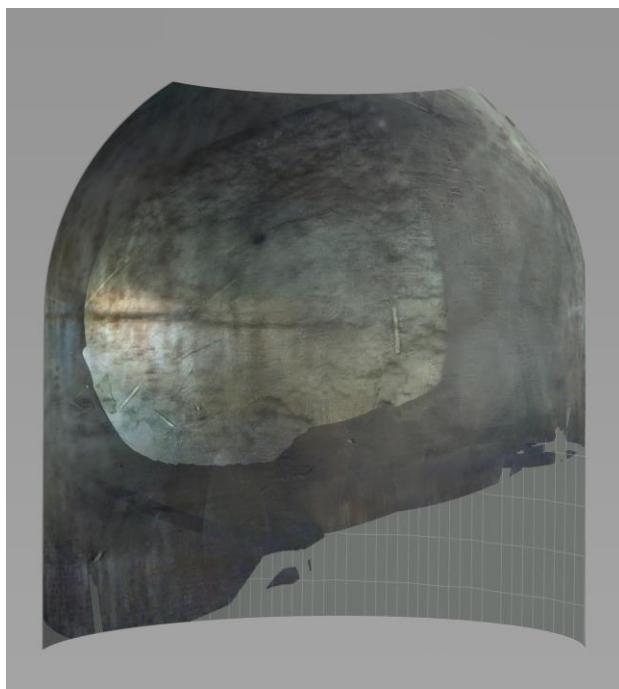


Figure 14 – Partial model of the tank.

The output highlights some issues of geometry definition, related to the aforementioned conditions (poor lighting, ever-shifting scene due to variable shadowing, floating dust, irregular flight paths etc.). In spite of these drawbacks, the model

provides geometry definition with a precision better than 10cm, assessed by comparing the model with the diameter of the cylindrical element which constitutes the lower portion of the tank.

3.9 Rectification of “undistort” images in the unfeasibility of modelling

In order to achieve data set IMG_5_UNDIST, the object of the last test, Elios explored a ventilation duct in the cooling system of a steam turbine. The UAV collected a video stream of a filter array (Figure 15a) in a chamber of the duct, aiming at the definition of geometry and mutual position. Starting from the initial hypothesis, which forewent any user-defined tie point, it was impossible to align the images of this data set even using “undistort” images and masks. Anyway, the availability of the “undistort” image set and the knowledge of some design dimension of each array element allowed to remove perspective effects (Cepolina, et al., 2015) and to achieve a metrically consistent graphic document (Figure 15b), albeit only for a single operator-defined plane (i.e. vertical face of the elements closer to the camera).

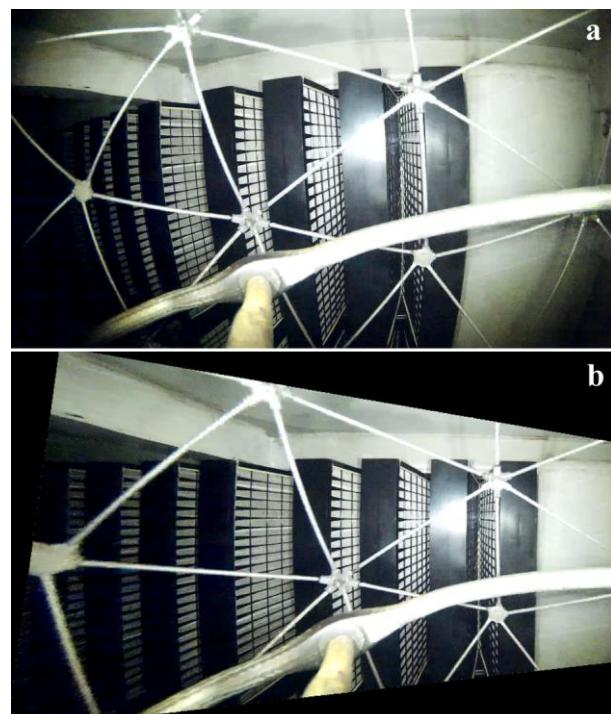


Figure 15 – a) Raw image b) Rectified “undistort” image

4. CONCLUSIONS

In the overall scenario of UAV systems aimed at image collection, the market recently offered a system able to fly in confined spaces, featuring different obstacles, designed “to crash and keep on going”. This is possible thanks to a protective frame providing full protection to both aircraft and video camera. These systems have been designed for video inspections in industrial environments, but their use is quickly expanding also to other scenarios, including emergencies and situations in which operator accessibility is precluded or dangerous.

The goal of the present work was to check system performance when using the collected video streams to achieve metrically

correct 3D models, by defining a workflow minimizing operator intervention. The first issue faced refers to the presence, on all images, of elements of the protective frame, which entails the inability to have static scenes. In most cases, the removal of the projection of such elements from the images enables their alignment thanks to correct automatic detection of the tie points on the survey object. Besides, due to the ultra-wide lens of the on-board camera of Elios, the distortion levels featured on the images entail a certain lack of consistency in image alignment, at the expense of the achievable geometric precision of the model. This issue can be addressed by using “undistort” image and mask sets, which enhances model correctness.

Overall, Elios proved able to perform large-scale surveys ensuring expeditious knowledge of object geometries, although some issues still need to be addressed in order to correctly perform stereo photogrammetry surveys. In order to improve overall usability, it would be advisable to change colour and/or texture of the protective frame of the UAV, so to minimize reflection of the light coming from the on-board LED array and to provide an easier target for the automatic definition of colour-based masks.

REFERENCES

- Ajayi, O. G., Salubi, A. A., Angbas, A. F. and Odigure, M. G., 2017. Generation of accurate digital elevation models from UAV acquired low percentage overlapping images, In: *Int. Journal of Remote Sens.*, 38:8-10, 3113-3134, DOI: 10.1080/01431161.2017.1285085
- Balletti, C., Guerra, F., Tsiontas, V. and Vernier, P., 2014. Calibration of Action Cameras for Photogrammetric Purposes. In: *Sensors 2014*, 14, 17471-17490.
- Cepolina, E.M., Caroti, G., Piemonte, A. and Rojas, P.G., 2015. Pedestrian modelling: Discomfort assessment. In: *17th International Conference on Harbor, Maritime and Multimodal Logistics Modelling and Simulation*, HMS 2015, pp. 51-56.
- Eschmann, C., Kuo, C., Kuo, C. and Boller, C., 2012. Unmanned aircraft systems for remote building inspection and monitoring. In: *Proceedings of the Sixth European Workshop on Structural Health Monitoring*, Dresden, Germany, 3–6 July 2012.
- Flyability_Elios_Brochure-LW, 2017.
https://www.flyability.com/wp-content/uploads/dlm_uploads/2017/11/Flyability_Elios_Brochure-LW.pdf (last visited 14/02/2018)
- Hastedt, H., Ekkel, T. and Luhmann, T., 2016. Evaluation of the quality of action cameras with wide-angle lenses in UAV photogrammetry. In *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B1, 851-859, <https://doi.org/10.5194/isprs-archives-XLI-B1-851-2016>.
- Martínez-Espejo Zaragoza, I., Caroti, G., Piemonte, A., Riedel, B., Tengen, D. and Niemeier, W., 2017. Structure from motion (SfM) processing of UAV images and combination with terrestrial laser scanning, applied for a 3D-documentation in a hazardous situation. *Geomatics, Natural Hazards and Risk*, 8 (2), pp. 1492-1504. DOI: 10.1080/19475705.2017.1345796
- Pajares, G., 2015. Overview and Current Status of Remote Sensing Applications Based on Unmanned Aerial Vehicles (UAVs). In: *Photogramm. Engineering & Remote Sens.*, Volume 81, Issue 4, April 2015, Pages 281-329, doi: 10.14358/PERS.81.4.281.
- Teo, T., 2015. Video-based point cloud generation using multiple action cameras. In *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-4/W5, 55-60, <https://doi.org/10.5194/isprsarchives-XL-4-W5-55-2015>.