360° IMAGE ORIENTATION AND RECONSTRUCTION WITH CAMERA POSITIONS CONSTRAINED BY GNSS MEASUREMENTS

M. Previtali [1] *, L. Barazzetti [1], F. Roncoroni [2], Y. Cao [1], M. Scaioni [1]

[1] Department of Architecture, Built Environment and Construction Engineering, Politecnico di Milano, via Ponzio 31, 20133, Milan, Italy - (mattia.previtali, luigi.barazzetti, yuwei.cao, marco.scaioni)@polimi.it
[2] Polo Territoriale di Lecco, via Previati 1/c, Lecco, 23900, Italy - fabio.roncoroni@polimi.it

KEY WORDS: Photogrammetry, 360° images, GNSS, Direct Georeferencing, Image alignment

ABSTRACT:

Photogrammetric applications using 360°images are becoming more and more popular in different fields, such as cultural heritage documentation of narrow spaces; civil, architectural, and environmental projects like tunnel surveying; mapping of urban city centres, etc. The popularity of 360° photogrammetry relates to the high productivity of the acquisition phase, giving the opportunity to capture the entire scene around the user in a relatively short time. On the other hand, the photogrammetric workflow needs ground control points (GCPs), well distributed over the survey area, to georeference the produced 3D data. Placing, measuring on-field, and identifying GCP on images is time-consuming and sometimes even not feasible due to environmental conditions. While effective solutions exist for UAV-based projects, direct georeferencing and GNSS assisted photogrammetry is still not fully exploited for ground-based acquisitions. This paper aims at presenting a solution coupling 360°images and high-precision GNSS systems for direct georeferencing of outdoor projects without the need for manually measuring GCPs. Three different acquisition modes for 360°images and GNSS data are presented, and orientation results are compared with manually measured Check Points.

1. INTRODUCTION

The availability of georeferenced 3D models is becoming a major demand in different application fields such as civil engineering, architecture, and environmental sciences. Among the different techniques for generating 3D point clouds and models, photogrammetry is gaining a lot of attention due to the possibility of using low-cost instruments. In particular, 360° cameras are receiving significant attention from different researchers (Humpe, 2020; Barazzetti et al., 2022; Murtiyoso et al., 2022). Indeed, in recent years, rapid technological advances in 360° cameras and the use of such sensors for amateurish applications have reduced hardware costs and enhanced image/video quality. A 360° camera consists of multiple cameras capturing in different directions. The acquired images are then stitched together to generate a complete 360° panorama. The acquisition of 360° images speeds up the acquisition phase and can represent an interesting alternative to traditional frame cameras in some application fields. The images generated by 360° cameras have a resolution of around 16-24 megapixels, while videos can be captured at a 5k resolution of 5120 x 2880 pixels, although some camera models offer even higher resolutions. Different authors (Gottardi and Guerra, 2018; Barazzetti et al., 2019; Teppati Losè et al., 2021; Janiszewski et al., 2022) have discussed the results achievable using images acquired with low-cost 360° cameras, showing different case studies, and testing metric accuracy and model completeness with laser scanning or traditional photogrammetry. The main applications of 360° cameras are currently devoted to cultural heritage applications, documentation of narrow spaces, mapping of city centres, and mapping of forestry areas. Even if 360° image based photogrammetry allows for the automatic reconstruction of 3D models from images, ground control points (GCPs) are required for scaling and georeferencing the model. However, placing GCPs can be difficult (some areas can be difficult to access or

identifying stable points can be questionable) and time-consuming, especially if GCPs are distributed on a wide area. Measuring GCPs is a time-consuming task when processing the image block since each single GCP must be manually measured in a set of images with possible mistakes by the operator. Problems connected with GCPs have led researchers to explore alternative methods, such as using Global Navigation Satellite System (GNSS) modules, in the case of outdoor areas. This approach proved suitable in the case of unmanned aerial vehicles (UAVs). Real-time kinematic (RTK) and post-processed kinematic (PPK) technologies have been applied to reach centimetre-level accuracies. This approach, called GNSS-assisted photogrammetry, uses antenna phase centre (APC) coordinates to scale and georeference the 3D model and add constraints to the bundle block adjustment. While most literature on this topic focuses on UAV-based applications for GNSS-assisted photogrammetry, limited attention is paid to ground-based applications, notwithstanding that terrestrial photogrammetry is a valid alternative where UAV flights are not possible (Morelli et al., 2022).

In this paper, we propose an alternative to GCPs for terrestrial photogrammetry using a geodetic pole, a 360° camera, and a GNSS antenna, coupled together to create a versatile setup that can be used to survey areas that cannot be reached by UAV solutions. In particular, we are proposing three operative survey pipelines: *"static"*, *"stop and go"* and *"kinematic"*, which use geotagged photos and frames from 5.7k videos to significantly decrease the acquisition time. The aim is to minimise survey cost and time to obtain a scaled and georeferenced photogrammetric model without manually measured GCPs.

The remaining of the paper is structured as follows: section 2 presents a review of the state of the art for 360° image photogrammetric processing and direct georeferencing; the presented acquisition system is described in section 3; section 4 describes the three operative workflows developed; test and

---

* Corresponding author

results are discussed in section 5 and sections 6 presents conclusions and future works.

2. RELATED WORKS

Over the past decade, there has been an exponential increase in market demand for lower-cost and consumer grade cameras, which has given rise to a wide range of new sensors available today. The potential to use spherical and cylindrical images captured with a nonmetric camera for photogrammetric applications is one of the latest and most important research topics among researchers working on photogrammetry. The market for spherical images and videos, also known as panoramas, has experienced a rapid growth over the last few years. These images are not only being used as documentation, visualisation and sharing tools but also for 360-degree videos, Augmented and Virtual Reality, and other applications. From a photogrammetric perspective, they provide some interesting advantages with respect to traditional frame cameras, like the largest field of view, which can be exploited for several applications. In Fangi, 2007 the concept of photogrammetric reconstructions based upon spherical cameras was first introduced and the mathematical model of spherical cameras was set out. From a practical point of view, the generation of spherical images was carried out through the stitching of different images acquired around a nodal point from the same camera, following the Computer Vision approach developed by Szeliski and Shum, 1997. Starting from those works, over the past few years, thanks to the development of consumer grade 360° cameras, the topic of spherical cameras has gained a lot of attention. Nowadays, spherical cameras are composed of a pair or more synchronised cameras that shoot in different directions all around the device. The individual photographs are then combined in a unique 360° image, which is generally represented as an equirectangular projection (Fangi & Nardinocchi, 2013). Consumer grade cameras are not specifically designed for metric purposes. For this reason, researchers are exploring the possibility of using them as a source for 3D measurement, working in two main directions: improving the image stitching (Lee et al., 2020) and developing photogrammetric procedures to cope with specific issues related to 360° image processing (Fangi et al., 2018; Janiszewski et al. 2022). Indeed, some practical aspects can pose significant problems when processing 360° images, even if the mathematical framework for spherical bundle adjustment is well defined. In the last few years, spherical cameras have been used for the survey of narrow spaces such as tunnels and caves and for the documentation of some peculiar architectural structures like the indoor of belltowers (Teppati Losè et al., 2021) and for the documentation of historical city centres (Barazzetti et al., 2022). However, GCPs are always highlighted as fundamental to obtaining reliable results (Barazzetti et al., 2022). On the other hand, integration between UAV data and GNSS sensors demonstrated really promising results for aided GNSS photogrammetry and direct georeferencing. The processing of data acquired by GNSS receivers on board of UAV platforms with the RTK/PPK (real-time kinematic, post-processed kinematic) approach proved accuracies comparable with those obtained by using GCPs (Tomaštík et al., 2019; Ekaso et al., 2020). Indeed, the idea of using GNSS or an inertial platform and GNSS to support bundle block adjustment is not new. Since the early 2000s, the combination of GNSS data, inertial platform, and aerial images was tested (Foralni and Pinto, 2002). Such combination was mainly designed for aerial photogrammetry to reduce the number of ground control points. As previously mentioned, this concept was then reused for UAV platforms. Ground applications were tested too (Forlani et

al., 2014). However, due to the high cost of the instruments and the difficulties of using them due to the bulkiness of some solutions, they were never fully exploited. In the last few years, the development of low-cost GNSS receivers and the creation of more compact cameras have revitalized interest in ground acquisitions and GNSS assisted photogrammetry. In Morelli et al., 2022 a low-cost multi-frequency antenna is used in combination with an action cam (GoPro HERO9) for the survey of vertical surfaces like building facades; in Tomaštík and Everett, 2023 smartphone images are georeferenced under tree canopy using low-cost (u-Blox) GNSS receivers. The present work is in the same research direction and tries to test the applicability of GNSS aided photogrammetry to 360° images.

3. ACQUISITION SYSTEM

The acquisition system consists of a geodetic pole with an Emlid Reach RS2 antenna on the top end and an Insta ONE X2 attached to the pole with a short arm (Figure 1). This configuration ensures that the camera and antenna maintain their relative orientation throughout the survey. Additionally, the device can function as a traditional GNSS antenna for surveying Ground Control Points (GCPs) since the antenna is aligned with the pole screw. Both the GNSS antenna and 360° camera can be easily controlled with an APP on a smartphone. So that users can manage both the Insta ONE X2 parameters such as shutter speed, colour balance, and frame rate and the GNSS acquisition parameters.

Figure 1: The acquisition system.

4. DATA ACQUISITION

For the surveying stage, three different approaches (Figure 2 - 4) were developed and tested for data acquisition.

The first approach, defined as *"static approach"* (Figure 2), requires the measurement of a small number of well-distributed images, e.g., one image every 20-30 m, 360° photos of the survey object along with their position recorded in RTK mode (*"geotagged images"*), and a set of well-distributed GCPs that can be measured in RTK as well. Once those images are recorded, a photogrammetric survey can be carried out with a 360° video acquisition at 5.7k resolution. Then the video sequence is processed in a combined bundle adjustment with 360° geotagged images and (if necessary) GCPs. The geotagged

images and GCPs are used to constrain the bundle block adjustment.
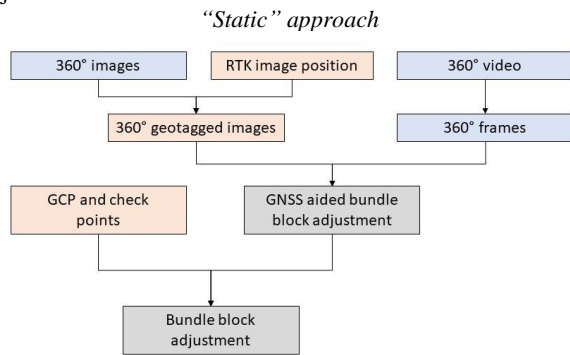
"Static" approach

Figure 2: Workflow for the *"static approach"*.

The second approach, defined as *"stop and go"* (Figure 3), involves a continuous recording of photos (derived as frames from a video sequence) and GNSS observations. Some stops of a few seconds are carried out during the acquisition, e.g., one stop every 5-10 m. In the processing phase, those stops are identified both in the GNSSS measurements and in the video sequence (*"geotagged frames"*). The position of the geotagged frames is used inside the bundle block adjustment of images extracted from the video with a high weight (precision set to 2.0 cm) to constrain the adjustment, while a lower weight (precision set to 2 m) is assigned to the position of the other frames. The lower weight assigned to general frames extracted from the video is due to the fact that no precise time synchronisation is considered for this acquisition approach.
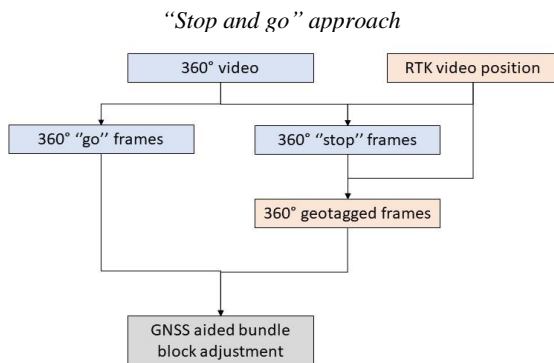
"Stop and go" approach

Figure 3: Workflow for the *"stop and go approach"*.

The third approach, named *"kinematic approach"* (Figure 4), requires continuous observation of GNSS RTK data and photos derived as a set of frames from a video sequence. The images can be synchronised using the GPS timestamp from the Insta ONE X2 internal receiver, providing more data to constrain the bundle block adjustment.
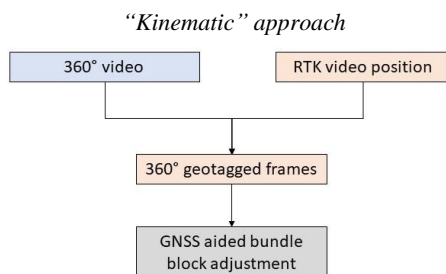
"Kinematic" approach

Figure 4: Workflow for the *"kinematic approach"*.

5. TESTS AND RESULTS

The presented approaches were tested in a courtyard of the Politecnico di Milano (Lecco Campus). The courtyard is surrounded by buildings with two or three floors. The data acquisition was carried out along a closed loop around the courtyard, a few metres from the buildings' façades. The GNSS receiver had good visibility of the sky to the north, while on the south-east side, there were other buildings that may have modest impact on satellite visibility. The three presented approaches were tested on the same path (Figure 5).
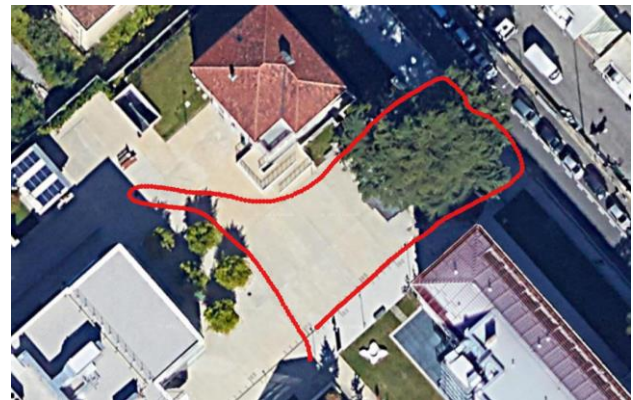

Figure 5: The path used for testing the three presented approaches (red line).

The length of the path is approximately 120m. A set of four points (in correspondence of manholes) were measured in RTK mode and used as Check Points (CPs) to evaluate the metric accuracy of the bundle adjustment. In addition, a further test was carried out, considering only 360° geotagged images. This test was setup to provide a reference and a comparison set for the three approaches presented in this paper.

5.1 Test of the three proposed approaches

As previously mentioned, the first test was carried out using only geotagged images. A set of 92 images and their positions were acquired. The distance between the different images is approximately a little bit more than 1.0 m and for each 360° image its position measured in RTK was recorded. The acquisition time in this operating mode is approximately 18 minutes, making it a solution not feasible for real-world applications. However, this acquisition serves as a refence for the evaluation of the accuracy of the three operative approaches presented in this paper.

The second test was carried out considering the *"static approach"*. A first video (resolution 5.7k) was recorded along the defined path (time of video acquisition approximately 6 minutes). And then, in a few positions (12), a set of static 360° images were acquired along with their position measured in RTK mode. A total of 10 images were acquired. The influence of the number of "static" 360° on georeferencing accuracy is tested, and results are discussed in Section 5.2 using different configurations of the geotagged images. The video sequence is then sampled at 1 HZ (i.e., 1 frame per second extracted from the video), obtaining 362 equirectangular frames.

The same path was followed using the *"stop and go approach"*. Two different solutions were tested: the first one with ten stops along the trail and the second one with 5 stops during the acquisition. Each stop lasts about 10 seconds. The video was sampled at 1 HZ, and to identify "stop" frames from the video, the Structural Similarity Index Measure (SSIM), see also Wang

et al., 2004, is computed between consecutive frames. A peak whose length is approximately 10 frames of the SSIM is evaluates as a "stop" position. Only one frame is extracted among the frames considered as "stop" and the corresponding GPS position is associated.

In the third approach, i.e., the *"kinematic"*, the GPS timestamp of the camera and the one recorded by the GNSS antenna are used. It is worth to mention that while the timestamp recorded by the Insta 360° cam refers to UTC time, the one recorded by the Emlid receiver is the GPS one, and a shift of 18 seconds has to be taken into consideration. The sampling frequency of the GNSS antenna is 0.2 s, while that of the Insta cam is 0.1 s (even if some inhomogeneities were observed).

Image orientation was carried out using Agisoft Metashape v 1.7.3.

5.2 Orientation results

To evaluate the accuracy of the orientation and of the georeferencing, both residuals on image positions and residuals on four check points (Figure 6) were considered. Check Points position was measured in RTK mode using the same receivers adopted for the presented solution, i.e., Emlid Reach RS2. The antenna used as a master is on the roof of the main building of the university campus (the baseline is approximately 100 m).

Figure 6: Position of the check points (blue diamonds) used for the evaluation of the accuracy of the image orientation.

The results of those comparisons are summarised in Figure 7 and Table 1 for residuals on image coordinates. Instead, Table 2 presents discrepancies on check points.

Approach	Geotag images	X error (cm)	Y error (cm)	Z error (cm)	XY error (cm)	Total error (cm)
Geotagged only	92	2.0	1.6	1.5	2.7	2.9
Static	12	2.0	1.3	1.2	2.4	2.7
Static	6	1.6	1.6	0.7	2.3	2.4
Stop and go	12	2.3	1.5	1.7	2.8	3.2
Stop and go	6	2.2	1.7	1.1	2.4	2.7
Kinematic	357	17.5	12.5	2.2	21.5	21.7

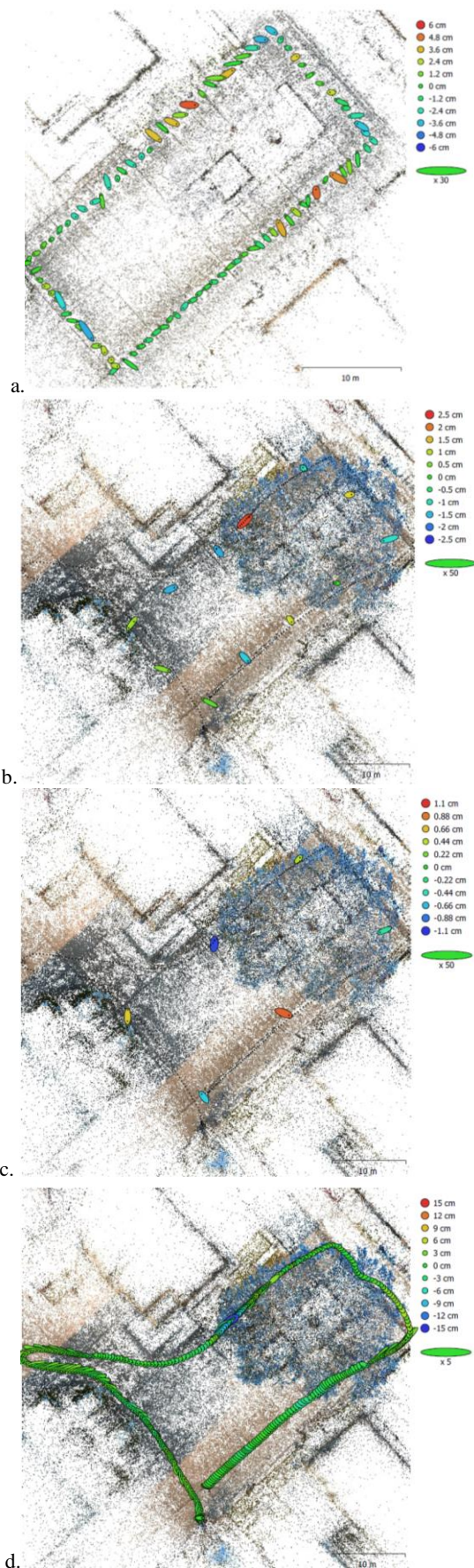Table 1. Average residuals on image position for the presented tests.

Figure 7: Residuals on camera position for: geotagged only images (a); static approach with 12 geotagged positions (b) and 6 geotagged positions (c); and kinematic approach (d).

Approach	Check points	X error (cm)	Y error (cm)	Z error (cm)	Total error (cm)
Geotagged only (92)	4	1.7	2.2	1.0	2.9
Static (12)	4	0.8	1.9	1.2	2.5
Static (6)	4	1.1	1.9	1.2	2.8
Stop and go (12)	4	0.8	2.8	1.9	3.5
Stop and go (6)	4	1.4	2.8	1.9	3.3
Kinematic (357)	4	12.3	5.1	1.3	17.1

Table 2. Average discrepancies on four check points for the presented tests.

Accuracy evaluation of the adjustment on average residuals on image positions and CPs shows that for the "geotagged only image" and the "static" approach, results comparable with RTK precision: in the order of ± 2.0 cm. The number of geotagged images does not seem to significantly influence the accuracy of the static approach. In the case of the "stop and go" accuracy seems a little bit lower ± 3.0 - 4.0 cm. This is probably due to the fact that a single position, the average of all the positions identified as "stop", is assigned to the 360° image to create geotagged data. Small movements of the pole during the "stop" may determine a lower accuracy in the definition of the position. For the "kinematic" mode instead, the horizontal accuracy is about 5 times worse ± 15.0 – 20.0 cm, and it is probably due to some problems in the synchronisation. Indeed, a few tenths of a second of offset in the synchronisation between the two data may lead to some centimetres of errors, considering a normal walking speed of 1 m/s. In addition, it can be observed that the mismatch between GPS position and camera position estimated from the bundle adjustment tends to increase along the path.

5.3 Point cloud comparison

As a final check about the metric accuracy the reconstructed point cloud derived from the "static approach" project with 6 geotagged images was compared with a set of reference scans acquired with a Faro Focus X130. In particular, three scans were acquired of the courtyard of the Politecnico campus and were registered together using an ICP approach. The same approach was used to co-register laser scans and photogrammetric cloud. Comparison between the two point clouds is performed on common areas by using the software CloudCompare (https://www.danielgm.net/cc/). Results in terms of unsigned cloud to cloud distance are reported in Figure 8. A false colour representation is used to evidence on the reference point cloud (the laser scanning one) discrepancies with respect to the photogrammetric one. The comparison between point clouds shows an average discrepancy of 2.7 cm for the vertical surfaces (i.e., building facades) and a little bit higher discrepancy (3.5 cm) for the horizontal paving of the courtyard. Overall accuracies are in good agreement with results obtained on GCPs and CPs. The distribution of discrepancies does not show significant bias in the reconstructed dense point cloud. Some boundary effects are visible in correspondence of areas at the boundary of the image block characterised by lower image coverage.
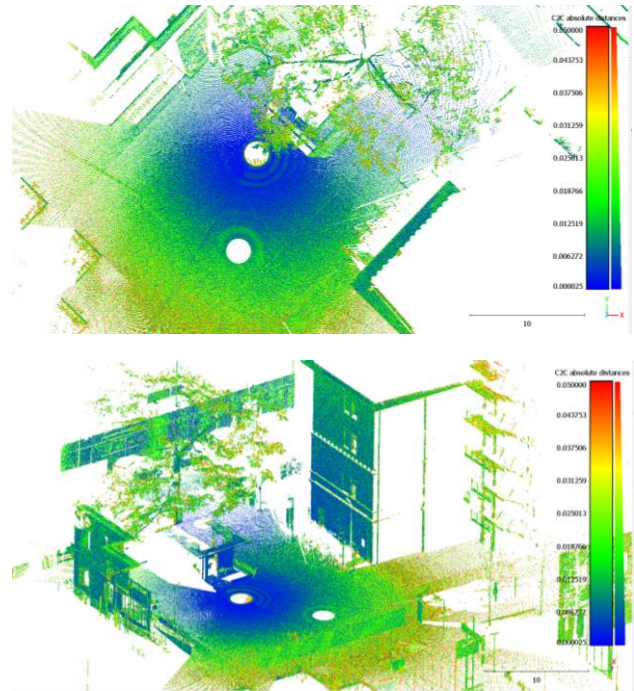

Figure 8: Comparison between laser scanning and 360° images ("static approach") of the Politecnico courtyard. Results are colorised according to discrepancies in a top view (top) and a side view (bottom).

6. CONCLUSIONS AND FUTURE WORKS

Photogrammetric applications relying on 360° images are becoming more and more poplar. The possibility to acquire with 5.7k resolution 360° video allows to acquire in a relatively reduced amount of time large areas. However, the large amount of acquired data poses some problems both for their processing (orientation and dense cloud generation for hundreds of thousands of images can be time consuming) and, using standard workflows, the need of acquiring a set of GCPs to obtain accurate and reliable results. In both situation the availability of positioning information for the acquired images can be beneficial. Indeed, as shown in Barazzetti et al., 2022 the availability of approximate positions (acquired with a smartphone) can speed up both matching and orientation stages with the initial creation of approximated Exterior Orientation parameters. However, the low quality of EO parameters acquired with a smartphone does not allow for direct utilization of the results and measurement of ground control points is still needed. In this paper we tested the possibility to increase accuracy of initial EO parameters by coupling a 360° camera with a topographic GNSS antenna and the possibility of using such parameters for GNSS assisted bundle block adjustment without the need of further GCPs. We have proposed three operational procedure for data acquisition named as *"static"*, *"stop and go"* and *"kinematic"*. While the "static" approach requires after the acquisition of the 5.7k resolution 360° video a second acquisition step of a set of 360° geotagged images. Both the "stop and go" and "kinematic" requires a single step of acquisition reducing this way the operational time of the survey. Accuracy evaluation of the adjustment on CPs for the "static" and the "stop and go" acquisition schemes show residuals on CPs similar to RTK precision: in the order of ± 2.0 cm for the "static" and a little bit lower accuracy ± 4.0 cm for the "stop and go". For the "kinematic" mode instead, the horizontal accuracy

is about 5 times worse ± 15.0 – 20.0 cm, and it is probably due to some problems in the synchronization. In future works we will better investigate the synchronization of the GPS with the 360° camera to improve the results of the "kinematic" approach and we will consider the effect of different path organization to the final orientation accuracy. In addition the integration with other sensors like inertial navigation system will be tested.

REFERENCES

Barazzetti, L., Previtali, M., Roncoroni, F., Valente, R., 2019. Connecting inside and outside through 360 imagery for close-range photogrammetry. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 42, 87-92.

Barazzetti, L., Previtali, M., Roncoroni, F., 2022. 3D MODELING WITH 5K 360° VIDEOS. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 46(2), 65-71.

Ekaso, D., Nex, F., & Kerle, N., 2020. Accuracy assessment of real-time kinematics (RTK) measurements on unmanned aerial vehicles (UAV) for direct geo-referencing. Geo-spatial information science, 23(2), 165-181.

Fangi, G. 2007. The Multi-image spherical Panoramas as a tool for Architectural Survey-XXI International CIPA Symposium, 1-6 October 2007, Atene, ISPRS International Archive–vol XXXVI-5/C53–ISSN 1682-1750–CIPA Archives vol. XXI-2007 ISSN, 0256-1840.

Fangi, G., & Nardinocchi, C. 2013. Photogrammetric processing of spherical panoramas. The photogrammetric record, 28(143), 293-311

Fangi, G., Pierdicca, R., Sturari, M., & Malinverni, E. S. (2018). IMPROVING SPHERICAL PHOTOGRAMMETRY USING 360° OMNI-CAMERAS: USE CASES AND NEW APPLICATIONS. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, 42(2).

Forlani G, Pinto L, 2002. Integrated INS/DGPS systems: Calibration and combined block adjustment, Proceedings, OEEPE Workshop "Integrated Sensor Orientation", Hannover, Sept. 17–18, 2001, OEEPE Official Publication N°. 43, 2002, pp. 85–96.

Forlani, G., Pinto, L., Roncella, R., & Pagliari, D., 2014. Terrestrial photogrammetry without ground control points. Earth Science Informatics, 7, 71-81.

Gottardi, C., Guerra, F., 2018. SPHERICAL IMAGES FOR CULTURAL HERITAGE: SURVEY AND DOCUMENTATION WITH THE NIKON KM360. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, 42(2).

Humpe, A., 2020. Bridge inspection with an off-the-shelf 360° camera drone. Drones, 4(4), 67.

Janiszewski, M., Torkan, M., Uotinen, L., Rinne, M., 2022. Rapid Photogrammetry with a 360-Degree Camera for Tunnel Mapping. Remote Sensing, 14(21), 5494.

Morelli, L., Menna, F., Vitti, A., Remondino, F., 2022. Action Cams and Low-Cost Multi-Frequency Antennas for Gnss Assisted Photogrammetric Applications Without Ground Control Points. ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 48, 171-176.

Murtiyoso, A., Hristova, H., Rehush, N., Griess, V., 2022. Low-cost mapping of forest under-storey vegetation using spherical photogrammetry. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 48, 185-190.

Szeliski, R., & Shum, H. Y. 1997. Creating full view panoramic image mosaics and environment maps. In Proceedings of the 24th annual conference on Computer graphics and interactive techniques (pp. 251-258).

Teppati Losè, L., Chiabrando, F., & Giulio Tonolo, F., 2021. Documentation of complex environments using 360 cameras. The Santa Marta Belltower in Montanaro. Remote Sensing, 13(18), 3633.

Tomaštík, J., Mokroš, M., Surový, P., Grznárová, A., & Merganič, J., 2019. UAV RTK/PPK method—an optimal solution for mapping inaccessible forested areas?. Remote sensing, 11(6), 721.

Tomaštík, J., & Everett, T., 2023. Static Positioning under Tree Canopy Using Low-Cost GNSS Receivers and Adapted RTKLIB Software. Sensors, 23(6), 3136.

Wang, Z., Bovik, A. C., Sheikh, H. R., Simoncelli, E. P., 2004. Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing, 13(4), 600-612.