Crowdsensing Close-Range Photogrammetry for Accurately Reconstructing a Digital Twin of a Cultural Heritage Building Using a Smartphone and a Compact Camera

Cristina Monterisi, Alessandra Capolupo(✉), and Eufemia Tarantino

Department of Civil, Environmental, Land, Construction and Chemistry (DICATECh), Politecnico di Bari, Via Orabona 4, 70125 Bari, Italy
alessandra.capolupo@poliba.it

Abstract. The development of user-friendly and cost-effective survey technologies is critical in the protection and preservation of cultural heritage structures. Although they are commonly reconstructed using digital photogrammetry techniques, and with the integration of Terrestrial Laser scanners and Remotely Piloted Aircraft Systems, they are now increasingly being modeled by crowdsensed systems, which are easily accessible even to non-expert users.

As such, the goal of this research is to evaluate the performance of a smartphone and a commercial compact camera in reconstructing a detailed and accurate digital twin of a cultural heritage object. In both cases, the close-range photogrammetric technique, based on the combination of Structure for Motion and Computer Vision approaches, was used. Those methods were tested on the Ognissanti Church, located in Valenzano, Italy. MicMac Graphic User Interface and CloudCompare, two open-source software and user-friendly interfaces, were used throughout the process. Thus, once camera calibration, preprocessing, and processing phases were completed, both collected databases were compared on the base of tie points image matching, average residuals in block-bundle adjustment resolution, and processing times. Despite variations in acquisition resolution and instrumental stability, and the fact that the point cloud from smartphone camera is 3 times less dense than that from compact camera, picture matching is equivalent after 15 to 20 orientation repetitions. The resulting two clouds were almost overlapping, with an average distance of 0.01 m. These findings also matched previous outcomes in the literature for small-volume reconstruction, substantiating the notion of performance independence from the size of the modified item.

Keywords: Digital twin · Smartphone-based photogrammetry · Cultural Heritage · 3D metric reconstruction · Computer Vision · Structure for Motion

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
O. Gervasi et al. (Eds.): ICCSA 2023 Workshops, LNCS 14107, pp. 234–250, 2023.
https://doi.org/10.1007/978-3-031-37114-1_16

1 Introduction

Digitization assumes a key role among the strategic axes of the National Recovery and Resilience Plan (PNRR) under the "Next Generation EU" program. The European Union has recently moved toward producing digital twins integrating data collection, aggregation, artificial intelligence, and analytic and virtual representations to preserve and make accessible cultural and historical heritage. In achieving this, it is essential to plan for continuous monitoring that ensures accurate estimates of surveyed areas and volumes [1].

Digital photogrammetry, both aerial and terrestrial, enables the reconstruction of metrically accurate and efficient digital models. It involves Computer Vision (CV) algorithms, for isolated objects and monuments with regular geometries or complex shapes modeling [2], Structure for Motion (SfM) [3–5], and Dense Multi-View 3D Reconstruction (DMVR) [6] techniques used in photographic campaigns with images overlay.

For years digital photogrammetry has been used for 3D reconstruction in Cultural Heritage (CH) applications [7, 8], often coupled with surveys using Terrestrial Laser Scanner (TLS) [9–16] and Remotely Piloted Aircraft Systems (RPAS) platforms [10, 14, 16–19]. Classical surveys (e.g. total station and Global Navigation Satellite System (GNSS)) combine effectively with TLS, which provides dense geometry information and very fine textured dense point clouds despite being a time-consuming process, and with RPAS platforms, which are characterized by high efficiency in covering large areas in a limited time. [10]. However, photogrammetry has usually been performed with commercial cameras and technologies, which are high-cost and not available to all users.

Therefore, with the increasingly widespread use of smartphones in recent years, Close-Range Photogrammetry (CRP) and crowdsensing have shown to be cheaper and user-friendly technologies, available even for all non-expert photogrammetry users, but still reliable for reconstructing 3D models [17, 20].

Nowadays, smartphones are equipped with cameras, even multiple ones, with increasing resolution, often higher than compact cameras. In addition, images can be shared on the web via Wi-Fi, providing near real-time measurements [21]. Moreover, apps such as Polycam, available on both IOS and Android devices, have recently been developed for in-app cloud-based 3D model reconstruction, with results comparable to SfM approaches [22].

During the last few years, several studies investigate the photogrammetric accuracy of 3D reconstructions obtained from smartphones, particularly for modeling geomorphological, geotechnical, and geological structures [23–25], historic buildings [26], including by stereovision [27], for industrial [28, 29] and medical [30, 31] applications. Although crowdsensing applications from smartphones are increasingly widespread, one of the limitations associated with this technology is the accuracy of model reconstruction. To achieve this, some studies include and analyze the calibration [32], pre-processing [33], and processing phase using proprietary software such as Pix4D [34] (https://www.pix4d.com/), Agisoft Metashape [35] (https://www.agisoft.com/) or open software such as Apero and MicMac (https://micmac.ensg.eu/index.php/) [36].

Only a few studies have compared the performance of smartphones and Digital Single-Lens Reflex (DSLR) cameras, both of which affordable and widely available ways of surveying. However, all studies refer to small-object modeling. In particular, Saif et al. [21] verified the reliability of the self-calibration and modeling approach with both cameras in the metric survey of areas and small geometric volumes, even though SmartPhone cameras (SPCs) had more lens distortion than DSLRs..

To respond to the lack of application of studies on large object modeling and comparison in the calibration phase, this paper aims to evaluate the potential of smartphones as on-site collection close-range photogrammetry tools for geometric modeling of CH buildings, comparing their reconstruction performance in terms of accuracy and point clouds, with digital compact cameras. In accordance with the European INfrastructure for SPatial InfoRmation in Europe (INSPIRE) directive [37] and Free-and-Open-Source Softwares (FOSS), in this study the open source MicMac software for the calibration and 3D reconstruction phase was tested, specifically using of the newly introduced user-friendly MicMac Graphic User Interface (GUI) interface [38].

The paper is organized in three main sections: i) "Material and methods" (Sect. 2), ii) "Results and discussion" (Sect. 3), and, lastly, iii) "Conclusion" (Sect. 4). In Sect. 2 background, methodology, dataset definition, pre-processing, processing, and post-processing phases needed for an accurate 3D modeling of a CH site, are described. Section 3 goes over the study findings and discussion on crowdsensing potentialities and reconstruction performances using smartphone and DSLR cameras. The final part highlights the process's merits and limitations as well as potential research advances.

2 Material and Methods

2.1 Study Site and Workflow Definition

The Romanic Church of Ognissanti, Valenzano, Italy (Fig. 1), is a historic rock building from the 12th century. Built of limestone tuff ashlars, it is located in the Apulian countryside, among olive groves. Abandoned for almost a century, it has remained almost intact in its Romanic style. Twentieth-century restorations were focused on repairing the cracks found on the roof slab, closing the entrance vestibule, maintaining the interior area, and removing the bell gable [39].

Restored in the 20th century, particularly regarding structural and roofing securing and internal maintenance efforts, it is now inaccessible and at risk of collapse. Therefore, this heritage needs careful and continuous monitoring of its static condition to plan suitable ordinary and extraordinary maintenance.

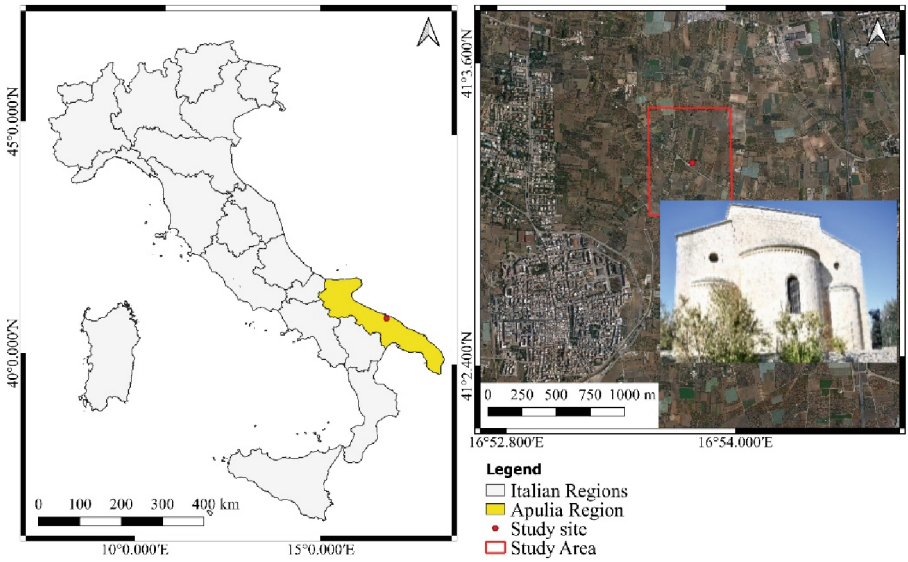


Fig. 1. Study site position in the Italian context on the left and overview of the study area on the right. A detailed of Ognissanti's Church is reported on the bottom-right.

After careful planning, the surveys carried out undershooting conditions to ensure the comparability of the surveyed data. The pre-processing and processing steps were conducted using open source MicMac (through the new user-friendly MicMac GUI interface) and CloudCompare software, focusing on the influence of the matching and calibration phases on the reconstruction accuracy, size, and noise of the generated dense point clouds. In this regard, statistical investigation methods were employed to compare reconstruction performances and estimate the accuracy of the 3D models acquired from both smartphone and DSLR cameras. Finally, the reliability of such results was investigated and discussed before statistically evaluating the correlation among the results of the previous processing phases. The operational workflow adopted is summarized in Fig. 2.

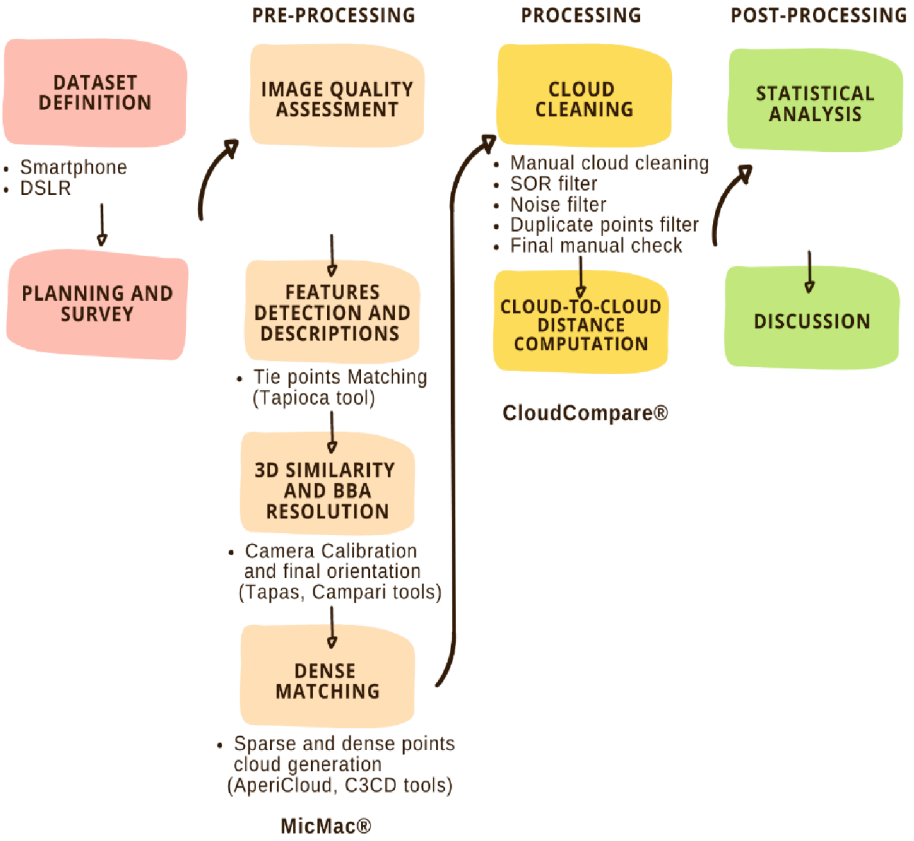


Fig. 2. Workflow of the survey from smartphone and Digital Single-Lens Reflex (DSLR) cameras. BBA: Bundle Block Adjustment; C3CD: Culture 3D Cloud; SOR filter: Statistical Outlier Removal filter.

2.2 Field Data Campaign and Dataset Description

In this research, photogrammetric data about Ognissanti's Church was captured using an ordinary DSLR camera (Nikon D-3300) and a smartphone (Xiaomi Redmi 10 C)..

The camera's specifications are listed in Table 1.

The survey was designed to be completed in a single day with uniform weather conditions (negligible cloud cover and wind) and monitoring the sun's position on the horizon to minimize direct reflections. By setting the ground pixel size as constant, depending on the focal length and the camera pixel size, the capture distance was calculated for both sensors, as shown in Table 2.

To overcome the difficulties caused by the survey's relevant noise conditions, the horizontal capture distance between the two sensors was kept constant. The presence of trees and dry-stone walls around the structure to be examined affected image matching.

Table 1. Utilized cameras Specifics.

Specifications	DSLR camera (Nikon D-3300)	Smartphone camera (Xiaomi Redmi 10C)
Camera model	D-3300	Samsung SSKKN1 Isocell
Sensor type	CMOS	CMOS
Sensor size	APS-C (23.5 mm x 15.6 mm)	~ 1/5”
Pixel size	3.84 μm	~ 0.72 μm
Camera Resolution	24 MP	Dual 50 MP + 2 MP
Focal Length (35 mm equivalent)	18- 55 mm	26.4 mm
Native ISO	10–12,800	100 – 6,400
Aperture	f/3.5-f/5.6	f/1.8-f/2.4
Image Format	RAW (NEF), JPEG	JPG

Table 2. Conditions and parameters of the acquisition survey.

Acquisition conditions and parameters	DSLR camera (Nikon D-3300)	Smartphone camera (Xiaomi Redmi 10C)
Acquisition data Acquisition hours	2022/11/17 h. 9.30 a.m. – 11.30 a.m	2022/11/17 h. 9.30 a.m. – 11.30 a.m
Acquisition distance	10.8 m	15.0 m
Captured photos number	148	152
Horizontal Photos distance	0.96 m	0.96 m
Photos Overlap	$\geq$ 80%	$\geq$ 80%

2.3 Data Processing

The collected images were transferred to a PC which specifications are given in Table 3.

Table 3. Configuration of the PC used for images processing.

PC model	HP Omen 45L
Operating System	Windows 64 bit
RAM	64 GB
CPU	Intel(R) Core(TM) i9–12900 K 3.20 GHz; 12th gen
GPUs	NVIDIA GeForce RTX 3090

Following a preliminary quality assessment phase that measures images contrast, blurred images are eliminated (with estimated quality less than 0.6 in a range from 0 to 1).

The program setting allowed editing the *exif* metadata (Focal Length, Focal Length equivalent to 35mm, Camera Model Name) for both acquisition sensors. The Tapioca tool, an operational interface of the bigger Pastis, was then used to perform feature detection and description, as well as the identification of matching tie points across pictures. At this step, four chunks (image blocks), one for each side of the building, were created and, for each of them, the same parameters were set to ensure comparability of results. It was possible to automatically compute the tie points by first considering pairs of images at a lower resolution and then detail matching at full resolution by using the Scale Invariant Feature Transform++ (SIFT++) algorithm [40, 41], whose advantage is to integrate the tie point generation in a way compatible with the global pipeline, with Multiple Scale (MulScale) mode.

Moreover, the ResiduesTapas tool was used to compute the camera orientation with a free-network approach. At the Internal Orientation (IO) step, the Radial Extended radial distortion model with 10 degrees of freedom (focal, principal point coordinates, distortion point coordinates, and 5 radial distortion coefficients) [42] was used. Micmac used Campari tool capabilities to address Block Bundle Adjustment (BBA) incorporating IO parameters using the Levenberg-Marquard optimization method [43, 44], which is suitable for handling nonlinear least squares problems. In the final orientation step, the four previously created chunks were merged. Lastly, using the Apericloud and Culture 3D Cloud (C3CD) tools, it was possible to automatically reconstruct the new sparse and dense points by setting the maximum resolution using the BigMag algorithm [45, 46].

After manually cleaning the point clouds to remove outliers from the object to be modeled (e.g., dry stone walls and trees), the Statistical Outlier Removal (SOR) filter and noise filter were applied [47] and duplicate points were removed.

Filtering parameters were selected through an iterative process including manual interpretation in order to offer the best cleaning of the dense point cloud while preserving relevant relief characteristics such as scattered points portraying the building's top. A last inspection allowed for the removal of any remaining isolated points or identified things that were discordant with the study site. After manual alignment and a subsample of the two clouds, the Cloud-to-Cloud (C2C) tool, in conjunction with the Multiscale Model to Model Cloud Comparison (M3C2) plugin [48], made it possible to calculate the distance between points acquired by DSLR and smartphone along the normal surface to the reference direction, while keeping track of 3D variations in the orientation of the surface, after manual alignment and a subsample of the two clouds.

It was then possible to estimate a confidence interval dependent on points cloud roughness and registration error for each measured distance.

2.4 Performance Reconstruction and Accuracy Assessment

The performance reconstruction and accuracy assessment step was aimed at testing: a) datasets quality; b) performance reconstruction in terms of processing time, quality, and cleanness of the generated point clouds; and c) accuracy of the photogrammetric measurements made with smartphones and DSLRs.

As highlighted by Triggs et al. [49], the photogrammetric products reconstruction accuracy depends on processed images quality, acquisition distance, and camera type. Therefore, the study considered these factors all. While conducting BBA using SfM (see Subsection "Data Processing"), the first factor was assessed (with the goal of minimizing its impacts) by utilizing the picture quality tool and selecting a SIFT++ technique.

Attaching an error matrix to the percentage of tie points retained in the BBA stage in matching the images and analyzing the residuals and their variation by the number of iterations in the calibration stage allows to compare calibration and modeling performance between the two cameras.

Pearson's correlation [50] was used to evaluate the alignment matrix between point clouds. Lastly, the errors that affected the final measurements of the oriented image blocks were evaluated using the Root Mean Square Error (RMSE).

3 Results and Discussion

In this study, a close-range photogrammetric crowdsensing approach was used to reconstruct a digital twin of a heritage building by comparing smartphones and compact cameras performances.

Survey planning allowed the acquisition of 152 and 146 images from smartphone and DSLR, respectively. Following the image quality assessment, 3 and 1 images were discarded from the two datasets, respectively. Inside MicMac software, all images with unsuitable lighting and environmental conditions were discarded, tie point matching and local calibration, and final alignment of the models were performed. Since MicMac is an open and programmable software, it allows to manage all preprocessing parameters. In this regard, it permits to choose the calibration model, as not all software, even commercial ones like Agisoft Metashape, does.

After the pre-processing step, the percentages of tie points kept in the image matching during the BBA resolution phase were calculated for the 112 images collected during both surveys (Table 4).

Table 4. Tie points maintained in images matching during the BBA phase (in percentage).

		DSLR (ref)				
		>99 (%)	95 - 99	80–95	50–80	<50
Smartphone	>99 (%)	89	1	1	0	0
	95–99 (%)	1	1	0	0	1
	80–95 (%)	6	3	2	1	0
	50–80 (%)	3	1	0	0	0
	<50 (%)	0	0	0	1	1

The results reveal precise image matching. Despite the significant number of trees present, only two shots from each camera had a proportion of tie points retained in the

final orientation of less than 50%. A value more than 80% is represented by 84% and 94% of images shot by smartphones and DSLRs, respectively. In both smartphone and DSLR polls, 89 images had a value more than 99%. Table 5 depicts elaboration timing of whole pre-processing step.

Table 5. Elaboration time during pre-processing phases in MicMac

Pre-processing phase	Smartphone camera Elaboration time (min)	DSLR camera Elaboration time (min)
Image Quality Assessment	0.5	1.0
Tie point matching	240.0	150.0
Camera calibration and final orientation	180.0	90.0
Dense matching	8.0	12.0

Even though crowdsensing systems such as smartphones and DSLRs reduce survey time and allow for user-friendly surveying and retrieval, the preprocessing phase requires substantial processing time. When Micmac is used as a preprocessing tool in conjunction with the PC configuration outlined in Table 3, the processing times are much longer than those measured in trials using commercial software. [21]. Pre-processing times for data from mobile devices are roughly twice as lengthy as those for tiny cameras.

However, it greatly increases processing times when compared to standard photogrammetry from TLS and RPAS [11, 16–18, 51, 53], especially when utilizing MicMac, rather than black-box software like Agisoft Metashape [21, 54].

An analysis of average residual variation in relation to the number of iterations of the final orientation step (BBA resolution/ image matching) was conducted on a sample of 15 images acquired from both cameras (for a total of 30 paired images). Samples were selected by identifying 3 major clusters: 4 images with higherthan 80% BBA coverage, 4 images with about 50% BBA coverage, and 4 images with less than 20% BBA coverage. Moreover, the first and last survey images, as well as a random acquisition, were taken into account. Indicative results of the 3 identified groups are displayed in Fig. 3.

The further synthetic analysis highlights that, as environmental conditions become worse, more iterations are required before the residual achieves acceptable values (smaller than 1 px). The gap between the residual from the smartphone and the residual from the DSLR drops substantially after a series of 15 and 25 repetitions in pictures with increased noise produced by tree cover. It should be noted that the residuals from DSLR acquisitions are more stable, yielding satisfactory results even with a small number of repeats (less than 5). However, although the value of the resulting residual is always lower in images acquired from DSLR (with an average of 0.39 px), the residual from smartphone data also reaches sub-pixel values (with an average of 0,67 px).

The topographic conformation of the site and the survey distance precluded a total reconstruction of the building. The noise caused by the significant presence of trees and the dry-stone walls surrounding the building affected the modeling result.

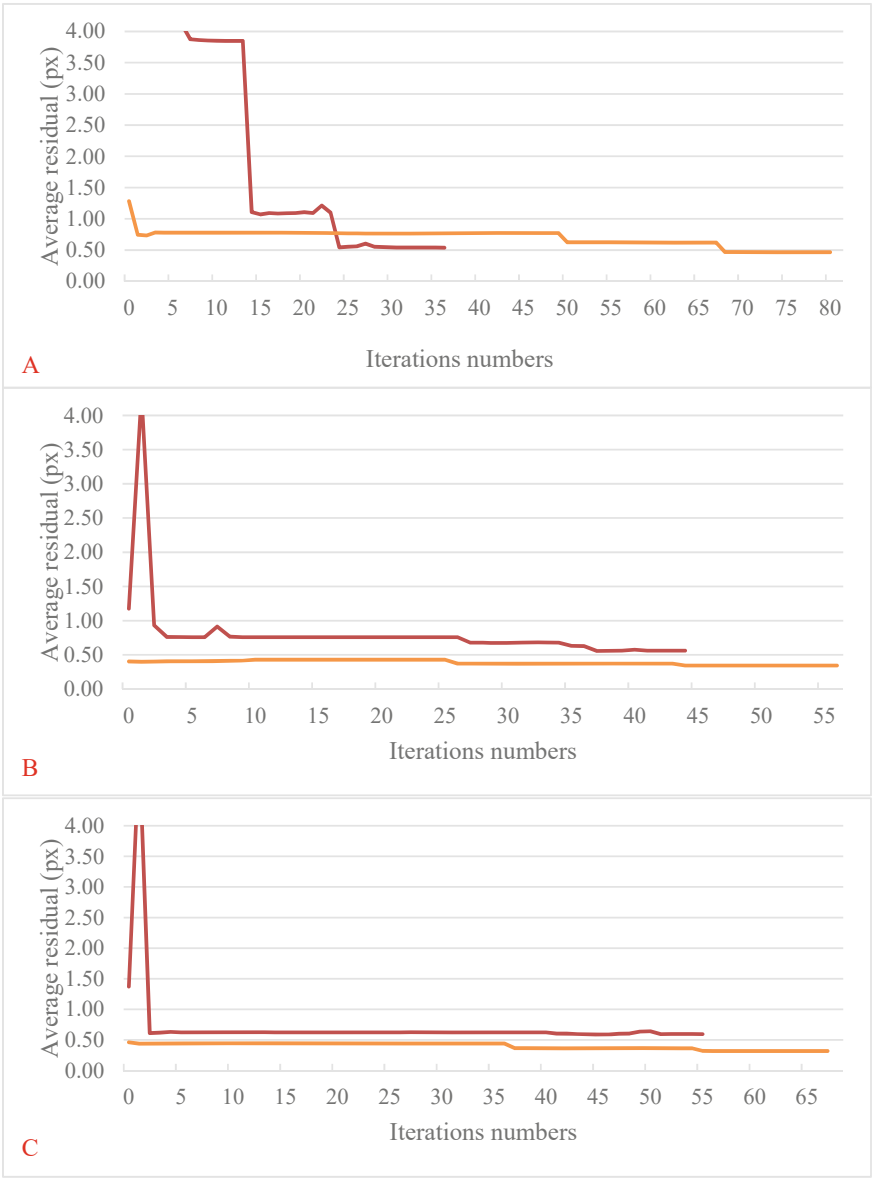


Fig. 3. Average Residual in function on iterations number in images matching during the Block Bundle Adjustment (BBA) phase. A) Reference image with more than 80% tree cover; B) Reference image with about 50% tree cover; c) Reference image with less than 20% tree cover.

The analysis of the reconstructed dense clouds is summarized in Table 6. The *DSLR camera* and *Smartphone camera* columns represent the size of each dense point cloud. The $\Delta\,(DPC_i\,points - DPC_{i\text{-}1})\,(\%)$ and $\Delta\,(DPC_i\,points - DPC_{i\text{-}1})\,(\%)$ columns represent the percentage of point decrease after applying the pre-processing filters. The point

Table 6. DPC points number on processing phases. DPC: Dense Point Cloud; SPC: Sparse Points Cloud

Points Cloud specifications	DSLR camera	Δ (DPC_i points – DPC_{i-1}) (%)	Smartphone camera	Δ (DPC_i points – DPC_{i-1}) (%)
SPC	5,750,814		957,338	
DPC	16,195,314		6,693,286	
DPC after SOR filter	9,967,825	-38.45	3,155,806	-52.85
DPC after noise filter	6,298,317	-36.81	2,329,822	-26.17
DPC after duplicate points filter	6,298,317	0.00	2,329,822	0.00
DPC after manual check	6,296,217	-0.03	2,328,969	-0.04
Δ (DPC_i points – DPC_0) (%)		-61.12		-65.20

cloud obtained by DSLR, after manually removing the elements from the survey target (dry-stone walls and trees), is approximately three times denser than that generated by a smartphone. It presents better geometric quality and a significant outlier values reduction. The data also reveal that both point clouds contain no duplicate points after applying SOR and noise filtering. The research provides evidence on cloud point number independence from the smartphone or digital camera utilized, which confirms the results reported by Saif et al. [21] on the modeling of tiny geometric objects. It also implies that the same characteristic is unaffected by the size of the discovered item. The gap indicates that the dense cloud cover provided by DSLR gives a more accurate depiction of the identified object information. Consequently, despite having a camera resolution twice that of the tested DSLR, the smartphone camera performs worse in terms of reconstructed model quality. After all, images for the compact camera are captured in RAW format, with no auto-adjustment or image compression. Instead, the JPG format compresses images to reduce their size.

In order to metrically compare the two generated models accuracy, alignment was required. The models were then uploaded to CloudCompare for the processing stage and aligned with point pairs-picking for registration. The DSLT model was scaled on a 1.68 factor. The RMSE was used to analyze the error impacting the final alignment from oriented image blocks. It yielded a reasonably low value (0.42 m), indicating that the two developed models fit well. DSLR dense point cloud, as seen in Fig. 4, provides a higher level of details and definition but does not include the reconstruction of the building roof. This is especially due to the lower acquisition distance between the camera and the object, which was selected to maintain the same detected pixel size between both cameras.

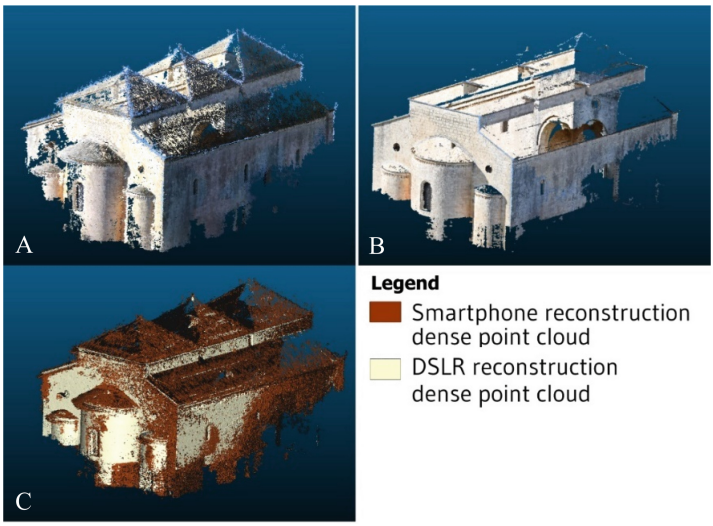


Fig. 4. Dense Points cloud alignment. A. Smartphone reconstruction; B. Reflex reconstruction; 3. Reconstruction alignment overlay.

Lastly, the distance between the modeled dense point clouds was computed using C2C tool (Fig. 5). This suggests that most of the points are less than 0.3 m away from the corresponding ones in the direction normal to the plane of the modeled surface. Building facades differ less across the two models, while roof parts, which have a lesser concentration of points, differ the most. This is explained in part by the lack of top modeling points in the DSLR data.

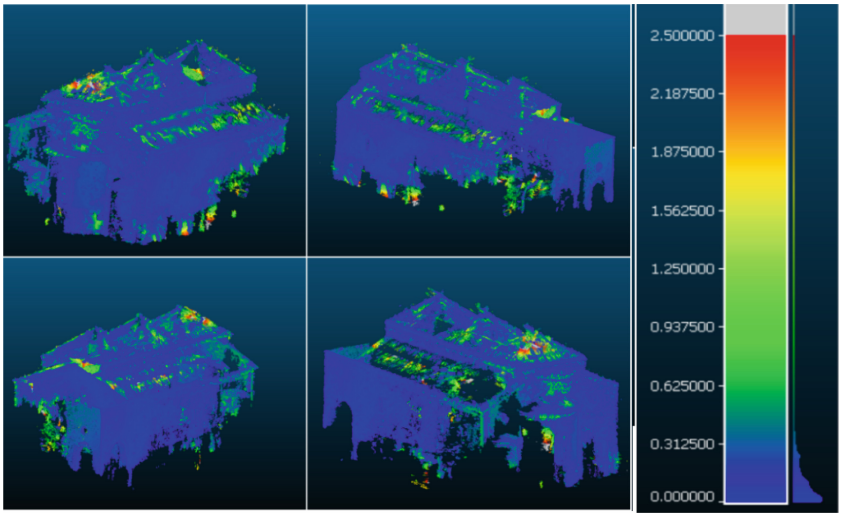


Fig. 5. Cloud-to-cloud distance computation (DSLR reference data)

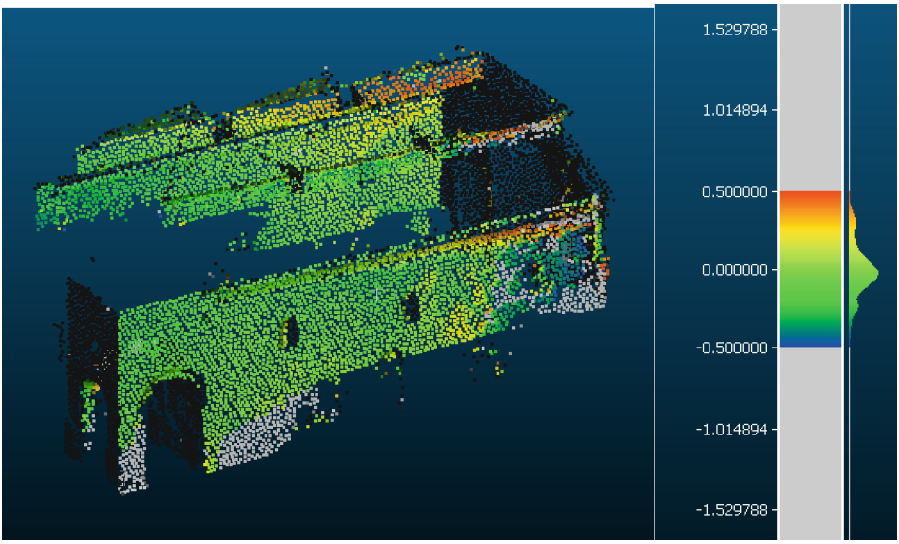


Fig. 6. M3C2 tool Dense points cloud distance

M3C2 tool results in determining the distances between the two point clouds are shown in Fig. 6. Distances distribution can be assimilated to a normal curve with values between -0.5m and 0.5m, a mean of -0.01 m, and a standard deviation of 0.29.

Although model's accuracy may be affected by factors, including image quality, image acquisition strategy, camera calibration, and SfM algorithm [55], the error obtained from a combination of TLS and drone-mounted cameras, was compared. While the average reconstruction error generated by crowdsensing is equivalent to that produced by camera reconstruction on the drone and TLS [10–12, 16], the standard deviation is approximately ten times larger than that produced by the combination of TLS and RPAS camera modeling [13]. However, the acquisition tactics employed a more approachable, user-friendly, cost-effective, and direct approach.

Moreover, this paper presents comparable results in large object monitoring to the literature on small object modeling [21], in terms of dense point clouds size ratio, gaussian distribution curve of point cloud distances between cameras, mean absolute distances, and standard deviation of 3D reconstruction. The ability to reconstruct accurate 3D metric models using user-friendly survey procedures is critical for major cultural heritage improvement and monitoring.

4 Conclusion

In literature, a few researchers have investigated the metric accuracy of reconstruction of 3-dimensional models by crowdsensing close-range photogrammetry. In particular, some works compared the metric reconstruction performance of user-friendly and readily available technologies such as smartphones and compact cameras. Among them, all have focused on the survey of small objects.

The performance of reconstruction combining smartphone and DSLR data capture to the modeling of large objects, namely a Cultural Heritage building located in Valenzano (Italy), was investigated in this work.

In both the calibration and alignment phases, the study shows encouraging results for 3D model reconstruction. Although the average preprocessing residues, processing time, and dense point cloud size are lower than that of a DSLR camera, the alignment phase reveals a consistent overlap between the two point clouds, with cloud-to-cloud differences ranging from 0.5 m to 0.5 m and, on average, settling around 0.01 m. Those findings suggests that, like compact cameras, smartphones may be used to collect on-site photogrammetric data immediately. Despite increased processing time and early residual instability in the first iterations of the BBA resolution phase, it produces an accurate reconstruction of the 3D object. Furthermore, by providing data comparable to tiny object analysis, this study solidifies the results' independence from the size of the identified items and the noise around the survey zone. Despite their speed in surveying, the processing periods of these cameras are frequently longer than those of consolidated photogrammetric acquisition methods.

This study's conclusions are confined to the testing of only two cameras and one calibration technique. Future studies might quantitatively examine the association between camera characteristics (sensor size, resolution, focal length) and model quality. Furthermore, in order to make accurate 3D heritage modeling accessible to non-expert users, further research might look at reconstruction quality in connection to calibration methodologies and pre-processing technologies, such as in-app processing.

References

1. Remondino, F.: Heritage recording and 3D modeling with photogrammetry and 3D scanning. Remote Sens. **3**, 1104–1138 (2011)
2. Remondino, F., El-Hakim, S.: Image-based 3D modeling: a review. Photogramm. Rec. **21**, 269–291 (2006)
3. Robertson, D., Cipolla, R. Structure from motion. Practical Image Processing and Computer Vision. John Wiley: Hoboken, NJ, USA (2009)
4. Capolupo, A., Saponaro, M., Borgogno Mondino, E., Tarantino, E.: Combining interior orientation variables to predict the accuracy of Rpas-Sfm 3D models. Remote Sens. **12**(17), 2674 (2020)
5. Caradonna, G., Tarantino, E., Scaioni, M., Figorito, B.: Multi-image 3D reconstruction: a photogrammetric and structure from motion comparative analysis. In: Gervasi, O., et al. ()eds. Computational Science and Its Applications–ICCSA 2018: 18th International Conference, Melbourne, VIC, Australia, July 2–5, 2018, Proceedings, Part V, vol. 18, pp. 305–316. Springer, Heidelberg (2018). https://doi.org/10.1007/978-3-319-95174-4_25
6. Furukawa, Y., Ponce, J.: Accurate, dense, and robust multiview stereopsis. IEEE Trans. Pattern Anal. Mach. Intell. **32**, 1362–1376 (2010)
7. Guarnieri, A., Pirotti, F., Vettore, A.: Cultural heritage interactive 3D models on the web: an approach using open source and free software. J. Cult. Herit. **11**(3), 350–353 (2010)
8. Balletti, C., Bertellini, B., Gottardi, C., Guerra, F.: Geomatics techniques for the enhancement and preservation of cultural heritage. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., **XLII-2/W11**, 133–140 (2019)

9. Cianci, M.G., Colaceci, S.: Laser scanner and UAV for the 2D and 3D reconstructions of cultural heritage. Sci. Res. Inf. Technol. Ricerca Scientifica e Tecnologie dell'Informazione **12**(2), 43–54 (2022)
10. Xu, Z., Wu, L., Shen, Y., Li, F., Wang, Q., Wang, R.: Tridimensional reconstruction applied to cultural heritage with the use of camera-equipped UAV and terrestrial laser scanner. Remote Sens. **6**, 10413–10434 (2014)
11. Tiwari, P.S., Pande, H., Gupta, S., Chandan, G., Samwal, E., Agarwal, S. Damage detection and virtual reconstruction of built heritage: an approach using high-recolution range and intensity data. J. Ind. Soc. Remote Sens. (2023)
12. Fassi, F., Fregonese, L., Ackermann, S., De Troia, V. Comparison between laser scanning and automated 3D modelling techniquesd to reconstruct complex and extensive cultural heritage areas. Int. Arch. Photogram. Remote Sens. Spat. Inf. Sci. **XL-5/W1** (2013)
13. Tysiac, P., Sieńska, A., Tarnowska, M., Kedziorski, P., Jagoda M. Combination of terrestrial laser scanning and UAV photogrammetry for 3D modelling and degradation assessment of heritage building based on a lighting analysis: case study—St. Adalbert Church in Gdansk, Poland. Herit. Sci. **11**, 53 (2023)
14. Saponaro, M., Capolupo, A., Turso, A., Tarantino, E.: Cloud-to-cloud assessment of UAV and TLS 3D reconstructions of cultural heritage monuments: the case of Torre Zozzoli. In: Eighth International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2020), vol. 11524, pp. 60–71. SPIE (2020)
15. Saponaro, M., Tarantino, E., Fratino, U.: Geometric accuracy evaluation of geospatial data using low-cost sensors on small UAVs. In: Gervasi, O., et al. (eds.) ICCSA 2018. LNCS, vol. 10964, pp. 364–374. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-95174-4_29
16. Capolupo, A.: Accuracy assessment of cultural heritage models extracting 3D point cloud geometric features with RPAS SfM-MVS and TLS techniques. Drones **5**(4), 145 (2021)
17. Capolupo. A.; Maltese, A.; Saponaro M.; Costantino, D. Integration of terrestrial laser scanning and UAV-SFM technique to generate a detailed 3D textured model of a heritage building. Earth Resources and Environmental Remote Sensing/GIS Applications (SPIE) XI Conference (2020)
18. Pepe, M.; Alfio, V.S.; Costantino, D. UAV platforms and the SfM-MVS approach in the 3D surveys and modelling: a review in the cultural heritage field. Appl. Sci. **12** (2022)
19. Cutugno, M., Robustelli, U., Pugliano, G.: Structure-from-motion 3D reconstruction of the historical overpass Ponte Della Cerra: a comparison between MicMac® open source software and Metashape®. Drones **6**, 242 (2022)
20. Boccia, L., Capolupo, A., Esposito, G., Mansueto, G., Tarantino, E.: A crowd-sensing system for geomatics applications. In: Misra, S., et al. (eds.) ICCSA 2019. LNCS, vol. 11622, pp. 297–312. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-24305-0_23
21. Saif, W., Alshibani, A.: Smartphone-based photogrammetry assessment in comparison with a compact camera for construction management applications. Appl. Sci. **12**, 1053 (2022)
22. Eker, R., Elvanoglu, N., Ucar, Z., Bilici, 3, Aydın, A.: 3D Modelling of a historic windmill: PPK-aided terrestrial photogrammetry vs smartphone app. ISPRS—Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. **XLIII-B2-2022 XXIV** (2022)
23. Dabove, P., Grasso, N., Piras, M.: Smartphone-based photogrammetry for the 3D modeling of a geomorphological structure. Appl. Sci. **9**, 3884 (2019)
24. Tavani, S., Granado, P., Riccardi, U., Seers, T., Corradetti, A.: Terrestrial SfM-MVS photogrammetry from smartphone sensors. Geomorphology **367**, 107318 (2020)
25. Alessandri, L., et al.: The fusion of external and internal 3d photogrammetric models as a tool to investigate the ancient human/cave interaction: the La Sassa case study. ISPRS—ISPRS—Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. **XLIII-B2–2020**, 1443–1450 (2020)

26. Bolognesi, C. M., Fiorillo, F. Virtual representations of cultural heritage: sharable and implementable case study to be enjoyed and maintained by the community. Buildings **13**, 410 (2023)
27. Masiero, A., Fissore, F., Piragnolo, M., Guarnieri, A., Pirotti, F., Vettore, A.: Initial evaluation of 3D reconstruction of close objects with Smartphone Stereo Vision. ISPRS—Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. **XLII-1**, 289–293 (2018)
28. Yu, L., Lubineau, G. A smartphone camera and built-in gyroscope based application for non-contact yet accurate off-axis structural displacement measurements. Measurement **167**, 108449 (2021)
29. Najathulla, B.C., Deshpande, A.S., Khandelwal, M.: Smartphone camera-based micron-scale displacement measurement: development and application in soft actuators. Instrum. Sci. Technol. **50**, 616–625 (2022)
30. Stark, E., Haffner, O., Kuˇcera, E.: Low-cost method for 3D body measurement based on photogrammetry using smartphone. Electronics **11**, 1048 (2022)
31. Matuzeviˇcius, D., Serackis, A.: Three-dimensional human head reconstruction using smartphone-based close-range video photogrammetry. Appl. Sci. **12**, 229 (2021)
32. Remondino, F., Fraser, C.S.: Digital camera calibration methods: considerations and comparisons. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. **36**, 266–272 (2006)
33. Jasinska, A., Pyka, K., Pastucha, E., Midtiby, H.S. A simple way to reduce 3D model deformation in smartphone photogrammetry. Sensors **23**, 728 (2023)
34. PIX4D Homepage. https://www.pix4d.com/. Accessed 27 Feb 2023
35. Agisoft Metashape Homepage. https://www.agisoft.com/. Accessed 27 Feb 2023
36. MicMac Homepage. https://micmac.ensg.eu/index.php/. 27 Feb 2023
37. Wheeler, D.A. Why Open Source Software/Free Software (OSS/FS). http://www.dwheeler.com/oss_fs_why.html. 27 Feb 2023
38. MicMac Graphic User Interface (GUI), http://clintons3d.com/plugins/other/micmacgui/index.html. Accessed 02 Mar 2023
39. De Cadilhac, R.L.:'arte della costruzione in pietra: chiese di Puglia con cupole in asse dal secolo XI al XVI. Gangemi Editore, Roma, Italy (2011)
40. Lowe, D.G.: Object recognition from local scale-invariant features. In: Proceedings of the Seventh IEEE International Conference on Computer Vision, Kerkyra, Greece, vol. 2, pp. 1150–1157 (1999)
41. GitHub SIFT++ Homepage. https://github.com/davidstutz/vedaldi2006-siftpp. Accessed 27 Feb 2023
42. Zhang, Z.: A flexible new technique for camera calibration. IEEE Trans. Pattern Anal. Mach. Intell. **22**(11), 1330–1334 (2000)
43. Levenberg, K.: A method for the solution of certain nonlinear problems in least squares. Quart. Appl. Math. **2**, 164–166 (1944)
44. Marquardt, D.W.: An algorithm for least-squares estimation of nonlinear inequalities. SIAM J. Appl. Math. **11**, 431–441 (1963)
45. MicMac documentation, MicMac, Apero, Pastis, and Other Beverages in a Nutshell! (2022)
46. Rupnik, E., Daakir, M., Pierrot Deseilligny, M.: MicMac – a free, open-source solution for photogrammetry. Open Geospat. Data Softw. Stand. **2**(1), 1–9 (2017). https://doi.org/10.1186/s40965-017-0027-2
47. Rusu, R.B., Cousins, S.: Point cloud library (pcl). In: IEEE International Conference on Robotics and Automation, pp. 1–4 (2011)
48. Lague, D., Brodu, N., Leroux, J.: Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (N-Z). ISPRS J. Photogramm. Remote Sens. **82**, 10–26 (2013)

49. Triggs, B., Mclauchlan, P.F., Hartley, R., FitzGibbon, A.W.: Bundle Adjustment -A Modern Synthesis. Lecture Notes in Computer Science. Springer Science and Business Media LLC, vol. 1883, pp. 298–372 (2000)
50. Pearson, K.: On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling, the London, Edinburgh, and Dublin. Philos. Mag. J. Sci. **50**(302), 157–175 (1990)
51. Saponaro, M., Capolupo, A., Caporusso, G., Tarantino, E.: Influence of co-alignment procedures on the co-registration accuracy of multi-epoch SFM points clouds. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. **43**, 231–238 (2021)
52. Saponaro, M., Capolupo, A., Caporusso, G., Borgogno Mondino, E., Tarantino, E.. Predicting the accuracy of photogrammetric 3D reconstruction from camera calibration parameters through a multivariate statistical approach. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLIII-B2–2020 **43**, 479–486 (2020)
53. Saponaro, M., et al.: Data fusion of terrestrial laser scanner and remotely piloted aircraft systems points clouds for monitoring the coastal protection systems. Aquat. Ecosyst. Health Manag. **23**(4), 389–397 (2020)
54. Saponaro, M., Turso, A., Tarantino, E.: Parallel Development of Comparable Photogrammetric Workflows Based on UAV Data Inside SW Platforms. In: Gervasi, O., et al. (eds.) ICCSA 2020. LNCS, vol. 12252, pp. 693–708. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-58811-3_50
55. Eker, R., Alkan, E., Aydın, A.: A comparative analysis of UAV-RTK and UAV-PPK methods in mapping different surface types. Eur. J. Forest Eng. **7**(1), 12–25 (2021)