Smartphone-based photogrammetry for the 3D modeling of a geomorphological structure

(Article begins on next page)

19 April 2024

Smartphone-Based Photogrammetry for the 3D Modeling of a Geomorphological Structure

Paolo Dabove[ID]**, Nives Grasso ***[ID]** and Marco Piras**[ID]

DIATI—Department of Environmental, Land and Infrastructure Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy; paolo.dabove@polito.it (P.D.); marco.piras@polito.it (M.P.)
* Correspondence: nives.grasso@polito.it; Tel.: +39-01-1090-7687

Abstract: The geomatic survey in the speleological field is one of the main activities that allows for the adding of both a scientific and popular value to cave exploration, and it is of fundamental importance for a detailed knowledge of the hypogean cavity. Today, the available instruments, such as laser scanners and metric cameras, allow us to quickly acquire data and obtain accurate three-dimensional models, but they are still expensive, require a careful planning phase of the survey, as well as some operator experience for their management. This work analyzes the performance of a smartphone device for a close-range photogrammetry approach for the extraction of accurate three-dimensional information of an underground cave. The image datasets that were acquired with a high-end smartphone were processed using the Structure from Motion (SfM)-based approach for dense point cloud generation: different image-matching algorithms implemented in a commercial and an open source software and in a smartphone application were tested. In order to assess the reachable accuracy of the proposed procedure, the achieved results were compared with a reference dense point cloud obtained with a professional camera or a terrestrial laser scanner. The approach has shown a good performance in terms of geometrical accuracies, computational time and applicability.

Keywords: low-cost sensors; close-range photogrammetry; structure from motion; rapid mapping; smartphone; underground caves

1. Introduction

Geomorphological and geological observation in underground environments such as caves could still represent a challenge. These environments are often prohibitive, given the difficulty of access and the conditions, even extreme, of humidity and temperature inside them. The continuous evolution of sensor-based 3D surveying and modeling techniques and the development of increasingly efficient systems for the visualization of digital data highlight the added value given by the use of these methods in the geological field. Indeed, the technological development of surveying instruments allowed passing from subjective descriptions to illustrate the environment [1], influenced by human perception, to full three-dimensional models. However, it is only in the last few years that lighter, cheaper and easier-to-introduce versions of surveying instruments have been sold on the market. The most recent electronic surveying equipment allows one to obtain a complete description of the environment and to represent all those morphological details that characterize the floors, walls, and roofs, which are fundamental for drawing conclusions on the evolution of the underground environments.

To date, beside instruments typical of classical topography, such as Total Stations (TS), both active sensors and passive sensors can be mainly used [2]. Active sensors are now widely used in underground caves to extract 3D information [3–11], and include all the scanning devices that emit a signal (in the form of a coded structural light pattern or through electromagnetic radiation in the form of a modulated laser light) and that record information from scene observation. Laser scanners are the most popular

solution, and today it is possible to find cheaper and smaller versions, such as profiler [12] and handheld laser scanner devices [13,14]. The active optical sensors [15] allow to directly obtain the spatial position of the detected points, sometimes coupled with the color information; the latter can be recorded by the sensor itself or by an external digital camera integrated into the instrument. These type of instruments have the main advantage of directly and quickly acquiring large amounts of data related to a complex geometry, with a very high accuracy. The raw output data of laser scanner sensors are unstructured point clouds, one for each acquisition point (in the case of a classical laser scanner), which need to be registered in a unique reference system to produce a single point cloud of the scene. Typically, in terrestrial laser scanners (TLS) a mirror rotates vertically and the whole instrument rotates horizontally, allowing a sequential scanning of the scene in both horizontal and vertical directions. There are also handled versions of laser scanners, but they do not always allow for the acquisition of the radiometric information of the object, or they are still very expensive and require some experience, especially in the post-processing of data. These instruments, however, can be easily transported and introduced in restricted access environments [16], but they have limited capabilities when it comes to describing details and edges, which become curved and approximate. Finally, 3D data delivered by active sensors are generally affected by the light and surface characteristics of the scanned materials, by the glass and by specular materials [2]. On the other hand, passive sensors attribute the radiometric information of the environment to a 2D matrix of pixels. The transition from the 2D data to the 3D model takes place thanks to different strategies, using a classical photogrammetric technique [17] or the more recent integration of Computer Vision (CV) algorithms [18], through stereo matching [19] and the Structure from Motion (SfM) [20] approaches. Image sensors such as digital cameras, action cams, and webcams, allow one to capture input data for a 3D structure reconstruction with these approaches. Specifically, the CV approach has been experimented since the years 2008–2009, and its integration in common software tools increased its potentialities [18]. The method permits the collection of very complex and complete information in a short amount of time, which is a great advantage for mapping and documentation purposes: moreover, this method can be applied for different case studies and to obtain information at different scales, both for small objects and for indoor or outdoor environments. Furthermore, the acquisition process is absolutely non-invasive, since it possible to collect data completely, even from remote areas, without touching the site at all. It also permits for the reaching of hardly accessible parts of the environment (such as the higher parts, which are difficult to record by means of other techniques, e.g., Terrestrial Laser Scanner, TLS). In fact, with TLS techniques it is sometimes not possible to extract the edges of an object, while this is possible with a photogrammetric restitution approach. Moreover, nowadays, thanks to the CV algorithms and to the new technologies, we are able to also use amateur cameras for photogrammetric purposes. The advantage of using passive sensors is that they do not require elaborate data acquisition planning, and data processing is also very easy for non-expert operators. Some recent studies described the application of 3D photogrammetric techniques on small areas of dark caves [6,21,22] or as an integration into the terrestrial laser scanning (TLS) approach [23].

A solution for the 3D mapping of natural caves can be found in terrestrial digital photogrammetry, using mass-market high-diffusion and low-cost technologies, such as smartphones or amateur commercial cameras for image acquisition. The limited size of image sensors represents an additional advantage, which allows one to install them in commonly used devices such as smartphones and tablets, giving anyone the ability to acquire data and extract 3D information. Thanks to the significant development of these devices and the embedded passive sensors, they represent a valid alternative to the classic survey technologies, which can be exploited for photogrammetric and 3D modeling purposes. Despite this, their performances are not yet comparable with high-end cameras, due to the quality of the sensors and the resolution of the images, which could affect the photogrammetric block alignment and, therefore, the 3D reconstruction; however, accurate results can be obtained, especially when performing a careful estimation of lens distortions [24]. In addition, one of the major advantages of using a smartphone is the ability to directly perform data processing with

specific applications installed on the mobile, now widely available in stores [25–29]. For example, today there are smartphones on the market, such as the Sony Xperia XZ1, sold with an integrated 3D modeling application [30]. In the literature, several studies are approaching 3D reconstruction methodologies involving the use of images acquired by smartphone devices or tablets for different analyses in multiple domains. In Aicardi et al. [24] and Fritsh and Syll [31], first attempts at exploiting the close-range photogrammetric technique using smartphone images were conducted, with the aim of reconstructing morphologically complex architectural elements, both for the documentation and the scientific dissemination of the cultural heritage and for the digital preservation of historical buildings and smaller objects of historical, architectural and artistic relevance. A further advantage derives from the easy availability of these tools, which makes them suitable for use in emergency situations when the goal is to obtain a complete description of the object in the shortest possible time. Since it is not possible to create a full object model with TLS or with photogrammetry, it is in both cases possible to interpolate the collected points in order to have a 3D model, removing the so-called shadows. This aim was pursued by Dabove et al. [32], who demonstrated the effectiveness of close-range photogrammetry for rapid mapping, using images acquired from tablets, in post-earthquake scenarios. Micheletti et al. [33] investigated the potentialities of the acquisition of high-resolution topographic and terrain data using hand-held smartphone technology for the generation of digital terrain models (DTM) of a river bank and an Alpine alluvial fan. These studies highlighted that an image-based approach that exploits mass-market devices allows for the reconstruction of geometrically correct models at different scales with a centimeter accuracy, or even higher. Moreover, the discussed methodology allows one to extract, in addition to the geometric properties of the object of study such as the shape and volume, its radiometric information [34], which characterizes each material; furthermore, this a fundamental aspect to take into account in order to define, for some applications, its properties and values.

The aim of this work is to investigate the potential of a smartphone device in the 3D reconstruction of objects of geomorphological interest and to propose a procedure that can be easily used, even by non-expert operators, based on the principles of photogrammetry. This refers in particular to researchers and experts interested in the geological and geomorphological evaluation of the different conformations present in an underground cave. Before applying this technique in an underground cave, the performances obtainable with a smartphone device are tested in a laboratory environment, in order to verify the precision and accuracy that are obtainable with these low-cost instruments. In addition, it is therefore important to define an operative methodology to be proposed in the field to non-expert users and which allows for a correct acquisition of images for the generation of 3D models; these are fundamental in order to have a good geological, geomorphological and tectonic knowledge of the subsoil, both for purely speleological or geological purposes and for the stability analysis of cavities.

The proposed methodology is designed to allow for acquisitions with easily transportable instruments (commercial cameras, tablets and smartphones) that are also available locally and that are used in limited spaces: often the elements of interest and the areas to be detected are located in areas that are difficult to reach, where it is possible to stop only for a short time, so that it is impossible to operate with LiDAR instruments. Moreover, if on the one hand the methods based on active sensors ensure rapidity in the survey and processing of dense clouds, on the other hand image-based procedures allow for the use of less expensive equipment, resulting in the same final product. The main novelty of this contribution is the use of the computer vision algorithms through the use of images acquired from a smartphone, exploiting both known software solutions and, above all, a smartphone application, which is quite a new approach for this type of application. Only the data from smartphone images were considered, without any integration into the model of other sorts of data (e.g., TLS point). Dense image matching techniques were then used to build a dense 3D point cloud, to be used as a support for geological and geomorphological investigations and any other activity, which could be supported by a dense and high-level-of-detail 3D model. With the aim of also involving non-expert users in the image-processing phase, low cost or open source tools were also tested: the acquired images

were processed using some of the best known software tools implementing CV algorithms, which are now also implemented in some smartphone applications (e.g., SCANN3D) [25], in order to evaluate their behavior in the processing and results. The possibility of reducing the costs of the survey with low-cost equipment, such as digital and amateur cameras, smartphones, tablets and action cams, is one of the main advantages of the photogrammetric technique, which also allows for a high repeatability of the survey. In this contribution, we propose a standardized procedure of 3D reconstruction of underground caves, through a photogrammetric survey with a smartphone, which allows us to obtain a high-quality result in terms of geometric accuracy. To define the methodology, we have analyzed the limits of this approach as the distance between the objects and as how image overlapping can influence the processing and the products. To this purpose, preliminary tests have been conducted in a controlled environment, using a rock sample as the object of study. Finally, the proposed approach was tested in a real underground cave.

2. Case Study

The present research was aimed at the three-dimensional reconstruction of a portion of the Cave of Bossea (CN), Italy. The Cave of Bossea, is considered among the most beautiful and geologically interesting caves in Europe, since it also has an underground karst laboratory managed by the Bossea Scientific Station of Italian Alpine Club (Club Alpino Italiano—CAI) of Cuneo and by the Department of Environmental, Land and Infrastructure Engineering of the Politecnico di Torino (Italy), with the collaboration of the Cuneo Department of Piedmont Regional Agency for the Protection of the Environment (Agenzia Regionale per la Protezione Ambientale—ARPA) and the Radiation Section of the ARPA Valle d'Aosta. This laboratory is of national importance for the study of water circulation in carbonate rocks, of the organization and evolution of karst aquifers, of genetic and lithogenic speleological processes, of atmospheric constituents, of the microclimate, and of the energy balance of the underground environment [35].

The preliminary test related to the proposed methodology was performed at the Geomatics Laboratory of DIATI (Department of Environmental Engineering, Land and Infrastructure) of the Politecnico di Torino, using a sample of calcareous material extracted from the Cave of Bossea (Figure 1).

Figure 1. The calcareous rock sample used for the test.

The sample measurements were about $60 \times 7 \times 3$ cm and aroused particular interest for its conformation in the study of water circulation in carbonate rocks. Moreover, it presents clear radiometric variations due to the characteristic components of the water that has been deposited over time and which represent the main study aspect of the object. The test was performed in optimal light and environmental conditions, in order to evaluate the potentialities of the procedures: this means that all images were collected in a controlled environment, both in terms of environmental (stable temperature of 25 °C and constant humidity of about 55%) and lighting conditions: the sample was settled on a non-reflective table and illuminated with professional lighting.

Subsequently, the work aimed to stress the methodology in a real underground cave environment, with variable levels of illumination and critical situations of humidity or temperature. To this purpose, three walls inside the cave were selected (Figure 2 and Table 1). The three chosen surfaces, characterized by the typical irregularity of these environments, are distinguished by the different lighting, punctual

in the first case, scarce in the second and diffused in the last one. Furthermore, due to a parapet that limits movement for safety reasons, it is not possible to get close enough to the various areas in order to inspect the various inlets with more detail. Moreover, most of the surfaces are moistened by a thin layer of water, which increases considerably in the most humid periods, and which makes them very reflective.

Wall 1　　　　　　　　　　　Wall 2　　　　　　　　　　　Wall 3

Figure 2. The selected areas of study inside the underground cave.

Table 1. Characteristics of the selected surfaces of the Cave of Bossea.

Study Area	Area Extension (LxH)	Avg. Acquisition Distance
Wall 1	7×3.50 m	4 m
Wall 2	13×4.40 m	Variable 3–6 m
Wall 3	4.50×3.30 m	Variable 1.50–5 m

3. D Model Generation: Methodology Assessment in a Controlled Environment

The photogrammetric technique allows us to determine the shape, size and position of an object, using as the main data the information contained in the images representing the object itself.

In the present work, a test is presented and performed with the purpose of analyzing the possibility of acquiring images with a common smartphone and using them for 3D modelling to obtain a detailed description of a geomorphological object and to evaluate the possibility of extending the procedure to a part of an underground cave.

3.1. Imaging Sensor Calibration

To solve the photogrammetric problem and extract metric information from the images, it is necessary to previously estimate the internal orientation parameters of the camera through a calibration procedure. The internal orientation consists on the estimation of the position of the projection center with respect to the fiducial system on the image plane [17], and it is defined by the coordinates of the principal point (x_0, y_0) and by the focal distance (f).

These parameters are referred to an ideal acquisition geometry, but in the case of using a non-metric camera, such as those integrated in smartphone devices, it is fundamental to take into account the distortion of the physical geometry of the lenses (radial and tangential distortions) and the real geometry of the sensor of the camera (skew). Therefore, in order to correctly estimate the internal orientation, it is necessary to know the magnitude of the distortions resulting from the optical system and the consequent deformations on the image.

Since the aim of this study is to verify the performances of popular mass-market devices in a survey of geomorphological structures of underground caves, a medium range smartphone has been considered: a OnePlus 5, whose characteristics are described in Table 2.

Table 2. Smartphone specifications.

	OnePlus 5
Sensor	*Rear 1:* CMOS 16MP 1/2.8″ Sony IMX 398 sensor and F1.7 lens *Rear 2:* CMOS 20MP 1/2.8″ Sony IMX 350 sensor and F2.6 lens *Front:* CMOS 16MP Sony IMX 371 sensor and F2.0 lens
Pixel pitch [µm]	*Rear 1:* 1.12 *Rear 2:* 1 *Front:* 1
Max image resolution [pixel]	3456×4608
Max video resolution [pixel]	3840×2160 (Rear camera 1: 4K 30 fps)
RAM	8 Gb
Dimension [mm]	$154.2 \times 74.1 \times 7.25$
Weight	153 g
Price	395–519 €

Generally, before using images for 3D reconstruction, it is important to analyze the intrinsic characteristics of the digital camera, which means knowing the parameters with which the images were taken, in order to study the acquisition geometry of the system.

The consequence of the diffusion of these devices has led, over the years, to the development of a series of algorithms and procedures for the reconstruction of the internal orientation parameters [17,36]. In this case study, we adopted the strategy of *the test field calibration* [37], which is based on the use of a special gridded panel with well-known coordinates or distances (Figure 3).

Figure 3. Wooden panel with checkerboards for the camera calibration. Each box measures 10 cm.

To perform the estimation, the free software Octave [38] was used; in particular, an adaptation of the open source Camera Calibration Toolbox for Matlab [39] allows us to obtain the calibration parameters from images of a checkerboard panel taken from the camera.

This panel has to be acquired from different positions, ensuring a good intersection of the rays and covering the entire sensor format. The images must be captured with the optical axis being perpendicular to the panel or convergent, and each image should have a relative rotation of 90° around the optical axis.

In our case, the checkerboard of 1×0.7 m was realized on a wooden panel to limit its deformations. The procedure then suggests carrying out a series of photos of the panel from different angles, which can then be inserted in the toolbox.

The calibration procedures must be conducted according to the sensor configurations (type of data acquired and resolution), which have been set during the survey, to calculate coherent intrinsic parameters. Therefore, the OnePlus 5 was calibrated using the frame extracted from the acquired videos. Indeed, the survey campaign was realized by employing the smartphone in the video mode in

order to ensure a complete, fast and stabilized data acquisition and in order to avoid the blur effect. Moreover, it would then be possible to exploit the same videos for further aims, such as the smartphone navigation of indoor environments [40].

20 frames selected from the recorded video were uploaded on the application, which made it possible to identify the grid corners by knowing their size (10 cm in our checkerboard panel) and to extract the camera parameters (Table 3).

Table 3. Estimated calibration parameters.

OnePlus 5		
Focal length	f [mm]	4.574
Principal point	x_0 [mm]	0.034
	y_0 [mm]	−0.017
Radial distortion	k_1	0.1764
	k_2	−0.1549
	k_3	−1.4567
Tangential distortion	p_1	0.0017
	p_2	−0.0131
Skew		−30.4633

3.2. Data Acquisition

The performed test addressed the assessment of the potentialities of a manual data acquisition for the updating of an object or a small area. The survey campaign was realized by employing the OnePlus 5 smartphone device in the video mode, in order to ensure a stabilized acquisition and to avoid the blur effect. At the same time, some images were captured with a NIKON D800E camera (Table 4), a high quality DSLR. Indeed, thanks to its performances from a technological point of view, it was employed for the generation of a reference dataset for the assessment of the geometric quality of the 3D point clouds produced with the smartphone.

Table 4. NIKON D800E specifications, according to the settings used during the data acquisitions.

NIKON D800E	
Sensor	35.9 × 24.0 mm CMOS (complementary metal-oxide semiconductor) sensor (Nikon FX format)
Focal length [mm]	24–50–60
Image resolution [pixel]	7360 × 4912
Shutter Type	Mechanical
Fastest–Slowest speed [s]	1/–30.0
Price	750–1200 €

The acquisition of the frames must be carefully designed and must take into account the distance between the different acquisition points, defined according to the scale of representation, the scale of the frame and the resolution of the image. In general, the acquisition geometry must be constructed in such a way that between one image and the next there is sufficient overlap both in the longitudinal and in the transversal directions, which for an approach with techniques deriving from Structure from Motion algorithms reaches values of 60–75% [41].

The smartphone device was held by the operator, who was moving around the object, trying to have a complete description: the auto-focus functionality of the camera sensors was disabled. In a few minutes it was possible to collect a dataset composed of more than 5000 frames, but only some of them were employed for the 3D reconstruction (Figure 4).
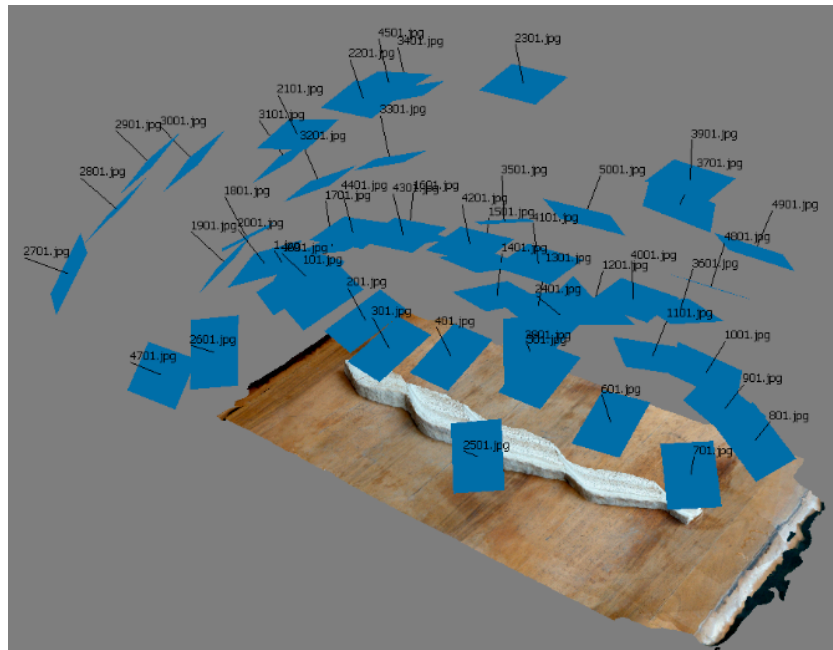
Figure 4. The camera's selected geometry configuration for data acquisition.

3.3. Data Processing

In order to evaluate the effectiveness of the smartphone-based methodology for the 3D reconstruction of geomorphological objects, the frames extracted from the video acquisition were processed and aligned through different software packages. To this purpose, two of the most popular pieces of software for image data processing were selected: the commercial solution Agisoft Metashape Professional (AMP) [42] and the open source VisualSfM (VSFM) [43]. Moreover, in order to evaluate the possibility of generating a point cloud in almost real-time and on the acquisition site, an application running on Android, SCANN3D, was tested. SCANN3D locally performs the 3D reconstruction, not on the cloud, in quite a rapid way, working entirely offline. The main benefit of the application is the real-time guidance during the image capturing phase, tracking multiple points in the viewpoints, which go from red to green when the camera has been moved sufficiently, ensuring a good image overlap. However, this application does not lock the focus and exposure after acquiring the first image, and this aspect can negatively affect the 3D reconstruction. In order to obtain a result comparable to that of the other software, it was decided to import and process the same image dataset.

These pieces of software implement different algorithms deriving from the Computer Vision (CV) for image processing and matching, allowing, almost automatically, for the easy elaboration of a large number of images. These image-based reconstruction methods are generally based on the Structure from Motion (SfM) [44] approach and the Dense Image Matching (DIM) or Multi View Stereo (MVS) algorithms [45] and are characterized by five main steps, which require, as their input data, only the acquired images:

- Feature extraction and identification of the matching points between images through different algorithms, such as the Scale Invariant Feature Transform (SIFT) algorithm [46], or its modified version [47];
- Outliers filtering, exploiting the Random Sample Consensus (RANSAC) algorithm [48], and a robust estimation of the acquisition geometry through linear models that impose geometric constraints;
- Estimate of the 3D coordinates of the object, the parameters of the internal and external orientation and the relative statistical information concerning the accuracy of the calculation, through the Bundle-Adjustment technique [49];

-     Dense surface reconstruction of the point cloud identified for the orientation, exploiting several algorithms, including the Exact, Smooth, Height-field and Fast methods, for the object surface description and its main discontinuities [45];
-     Dense point cloud interpolation and texturing in order to obtain a photorealistic visualization.

The used pieces of software differ in terms of the algorithms involved in the image processing and 3D reconstruction pipeline; however, they all lead to the generation of the dense point cloud and, except for VS, to the production of a textured surface model. The georeferencing of the model can be performed during the image processing phase by importing, together with the images, the EXIF files containing the cameras acquisition points; otherwise, it is possible to manually identify some common points between the images and to assign them their coordinates in the chosen reference system (e.g., topographically measured markers).

A first test was conducted using the advanced image-based software package AMP, which today is considered one of the best solutions for the generation of an accurate and complete dense point cloud in a very simple way. However, AMP can be defined as a "black box" since the implemented algorithms are not described in a deeper way.

Subsequently, we tested the performance of the open source software vs. the integrated algorithm CMVS/PMVS [50]. Naturally, all the processing steps of this tool for image elaboration are well known, thanks to the dissemination of the information provided by the developers and the scientific community. Finally, the same datasets processed by the previously mentioned software were elaborated with the SCANN3D application, in order to evaluate its performances.

Assuming that the conditions of particular environments require a certain speed in the acquisition of data, we wanted to assess what was the minimum number of images necessary to ensure a complete description of the object. To this end, three different datasets have been created with the frames extracted from the video, containing respectively 50, 25 and 12 images, in each of which an overlap between the images greater than 70% is guaranteed.

Naturally, in order to obtain comparable results in AMP and SCANN3D, the High and Ultra settings, respectively, were employed, while in VSFM it is not possible to select such a setting.

Data elaborations and analyses were conducted using a computer Hp EliteBook, mounting a processor Intel® Core™ i7-4600U CPU 2.1 GHz–2.7 GHz, with Ram of 16 GB and an architecture of 64 bit; to which all of the processing times refer.

Table 5 shows some information regarding the dense point clouds computation, which demonstrates how results can vary from one software to another. From an initial analysis of these data, it is clear that the AMP software is able to generate products that are much denser than for other software, although it requires a greater computational time. It is also possible to observe that, while AMP was able to reconstruct the photogrammetric model with both of the three datasets, on the other hand twelve images were not sufficient for the VSFM software to extract 3D information from the object. Furthermore, with the SCANN3D application, it was not possible to reconstruct the photogrammetric model with the dataset consisting of 50 images, probably because the processor implemented on the smartphone did not have a sufficient performance to face a similar computational cost.

Table 5. Information on the dense point clouds obtained from the smartphone images of the rock sample.

Software	Image Resolution	Images	Aligned Images	Points Dense Point Cloud	Computational Time
AMP	3840 × 2160 px	50	50	2,820,646	87 min
AMP	3840 × 2160 px	25	25	2,243,974	22 min
AMP	3840 × 2160 px	12	12	1,201,846	6 min
VSFM	3840 × 2160 px	50	50	1,108,386	25 min
VSFM	3840 × 2160 px	25	25	728,523	17 min
VSFM	3840 × 2160 px	12	2	-	-
SCANN3D	3840 × 2160 px	50	-	-	-
SCANN3D	3840 × 2160 px	25	Unknown	704,907	10 min
SCANN3D	3840 × 2160 px	12	Unknown	672,042	4 min

4. Reference Data

The validation of the resulting products was conducted through the comparison of each point cloud with the generated reference point cloud, starting from frames acquired by the NIKON D800E camera. In order to compare the models in the same local reference system, ten natural points on the rock sample were topographically measured with a total station and used for each dataset orientation through a bundle block adjustment. All measurements were adjusted with the MicroSurvey StarNet v.7.0 software in order to obtain the final coordinates, which were determined with a millimeter accuracy. Six points were used as control points for the point clouds georeferentiation, while four other points were employed as check points for the residuals evaluation.

The 33 images were processed with AMP, obtaining an estimate root mean square (RMSE) of less than 0.5 mm on the check points and a point cloud composed of more than two million points (Figure 5).

Figure 5. The colored 3D dense point cloud of the rock sample.

Table 6 summarizes the characteristics related to the reference dense point cloud.

Table 6. Characteristics of the reference dense point cloud obtained from the NIKON D800E images.

Software	Image Resolution	Images	Aligned Images	Points Dense Point Cloud	Computational Time
AMP	7360 × 4912 px	33	33	11,694,509	77 min

In order to further evaluate the accuracy of this reference model and, specifically, its scale, we used three points that were external to the rock sample, measured with the total station. The distances between these points have been computed and, the same distances, were measured on the reference 3D dense point cloud. The difference between these distances are less than 0.3 mm and confirm the accuracy of the model, considering that the average GSD corresponds to 0.032 mm/pixels.

5. Results and Comparisons

A first statistical analysis of the point clouds obtained with a low-cost tool was carried out by analyzing the estimated residuals on the control points and on the check points (Table 7). Indeed, for the

orientation of each dataset through a bundle block adjustment, the same six control points measured for the reference point cloud were employed, while the remaining four were used as check points. The obtained values are indicative of the overall geometric accuracy of the photogrammetric models generated through the SFM approach. However, this type of analysis can only be carried out on data processed with AMP and VSFM, as SCANN3D does not allow geo-referencing the model using control points, as the entire modeling process is automatic.

Table 7. Resulting errors obtained with the different software.

Software/Dataset	Control Points			Check Points		
	RMSE (X) [mm]	RMSE (Y) [mm]	RMSE (Z) [mm]	RMSE (X) [mm]	RMSE (Y) [mm]	RMSE (Z) [mm]
AMP/50 images	0.409	0.378	0.900	0.574	0.169	1.102
AMP/25 images	0.527	0.409	0.750	0.527	0.127	0.806
AMP/12 images	0.452	0.325	0.783	0.617	0.198	0.884
VSFM/50 images	Total error: 0.487 mm			-	-	-
VSFM/25 images	Total error: 1.505 mm			-	-	-

According to the obtained results (Table 7), for this type of object, increasing the number of images does not necessarily increase the quality of the final product. In fact, the constant radiometry could not allow for a correct alignment of the images; moreover, an excessive overlapping of the images, due to the limited dimensions of the object, could lead to a relatively inaccurate solution.

Some first observations of the photogrammetric models show that those generated through the AMP and VSFM software describe the object in its entirety, even if there is a decrease in the density of the points at the edge in the elevation of the rock sample, due to a non-optimal acquisition geometry.

In contrast, with the same image datasets, the SCANN3D application generates deformed models of the rock sample (Figure 6). In order to better assess the deviation of these products from the real object, the reference point cloud, realized as described in the previous paragraph, has been exploited to perform some statistical analyses on the models obtained from images acquired with the smartphone.
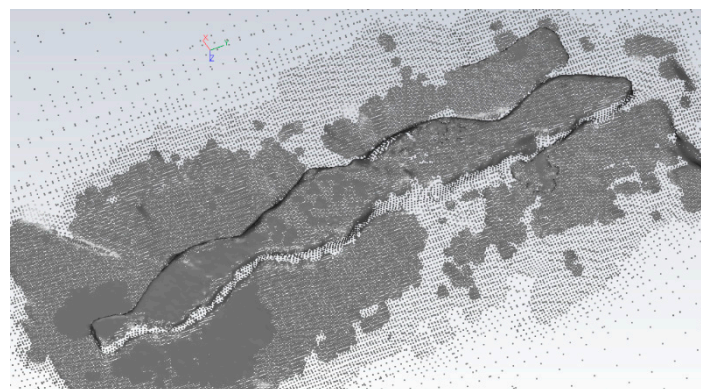
Figure 6. The point cloud model realized with the dataset of 25 images using the SCANN3D application.

5.1. C2C Distance Comparison

In order to assess, with smartphones, the quality and correspondence to reality of the SFM approach, the distance between the clouds was generated and the reference cloud was computed. To this end, the CloudCompare [51] open source software was used, allowing this estimate to be performed using the distance Cloud-to-Cloud (C2C) tool. The C2C tool exploits the Nearest Neighbor algorithm to compute the Euclidean distance between each point of the compared cloud and the nearest point of the reference cloud.

At this point, it was necessary to scale and register the models generated by SCANN3D in a reference system consistent with the other models. In this regard, CloudCompare offers several tools: a rough registration can be performed through pairs of equivalent points identified in the point cloud that is to be registered and in the reference model; subsequently, the automatic method based on the Iterative Closest Point (ICP) algorithm [52] allows the two models to be very finely recorded. This procedure was performed for both models deriving from SCANN3D, with a final RMSE of about 5 mm.

Since all of the point clouds were quite noisy in the area related to the supporting surface of the rock sample, the cloud portions that were not part of the study object were manually removed. The C2C tool was launched by setting the points that were more than 5 mm away from the reference model as outliers, since the errors that interest us are lower than this limit.

Figure 7 shows some significant results deriving from this procedure in histograms, statistical distributions and through the display in a color scale directly on the point clouds. The computation of the differences (Figure 7) confirmed the metrical accuracy of the models estimated by AMP and VSFM, and showed that the greater discrepancies correspond to the support surface and are greater than 5 mm. This error can be attributed to the noise affecting the point clouds, due to the reflective support surface combined with an external light source close to the test site, which probably caused an incorrect identification of the matching points.

Table 8 shows the statistical parameters derived from the analysis with the C2C tool, such as the maximal distance, average distance and standard deviation.

Table 8. The statistical results from the distance comparison between the smartphone point clouds and the reference point cloud generated with the NIKON D800E images.

Software/Dataset	Min [m]	Max [m]	Mean Distance [m]	Std Deviation [m]
AMP/50 images	0.000	0.005	0.0012	0.0011
AMP/25 images	0.000	0.005	0.0009	0.0010
AMP/12 images	0.000	0.005	0.0007	0.0008
VSFM/50 images	0.000	0.005	0.0009	0.0010
VSFM/25 images	0.000	0.005	0.0014	0.0009
SCANN3D/25 images	0.000	0.005	0.0027	0.0016
SCANN3D/12 images	0.000	0.005	0.0017	0.0016

From the analysis of these results, it emerges that the three-dimensional models obtained with AMP and VSFM have better statistical values, lower ones, or ones on the order of a millimeter. Furthermore, it must be taken into account that for the model obtained from SCANN3D by processing the dataset of 12 images, although the average distance from the reference model is less than 2 mm, the representation of the object is visibly incomplete, as shown in Figure 7. For a better understanding of these discrepancies, a further analysis on the error distributions is required, as reported in Table 9. The results that are obtained show that more than 84% of the point clouds realized with AMP and VSFM deviate from the reference data by less than 2 mm, while these results were not achieved by the SCANN3D application.

The comparison of the point clouds made with images acquired from the smartphone and the model generated from the NIKON data shows that the variations are minimal and that the models are almost the same. The only exception results from the model created through the SCANN3D app; in this case, the incompleteness of the model is not due to the performances of the smartphone camera but, rather, to the algorithms exploited by the application itself and to the computational performances of the smartphone.

Table 9. Point distances distribution.

Software/Dataset	Absolute Distance C2C [%]				
	D < 0.5 mm	0.5 < D < 0.8 mm	0.8 < D < 1 mm	1 < D < 2 mm	2 < D < 5 mm
AMP/50 images	~27.10	~21.81	~16.91	~18.05	~16.16
AMP/25 images	~44.02	~25.62	~7.70	~11.27	~11.40
AMP/12 images	~56.01	~18.18	~8.19	~11.72	~5.91
VSFM/50 images	~46.87	~22.87	~8.07	~11.31	~10.89
VSFM/25 images	~17.03	~14.09	~10.52	~37.86	~20.51
SCANN3D/25 images	~6.59	~6.52	~3.85	~22.04	~60.99
SCANN3D/12 images	~26.07	~14.36	~8.39	~19.99	~31.20
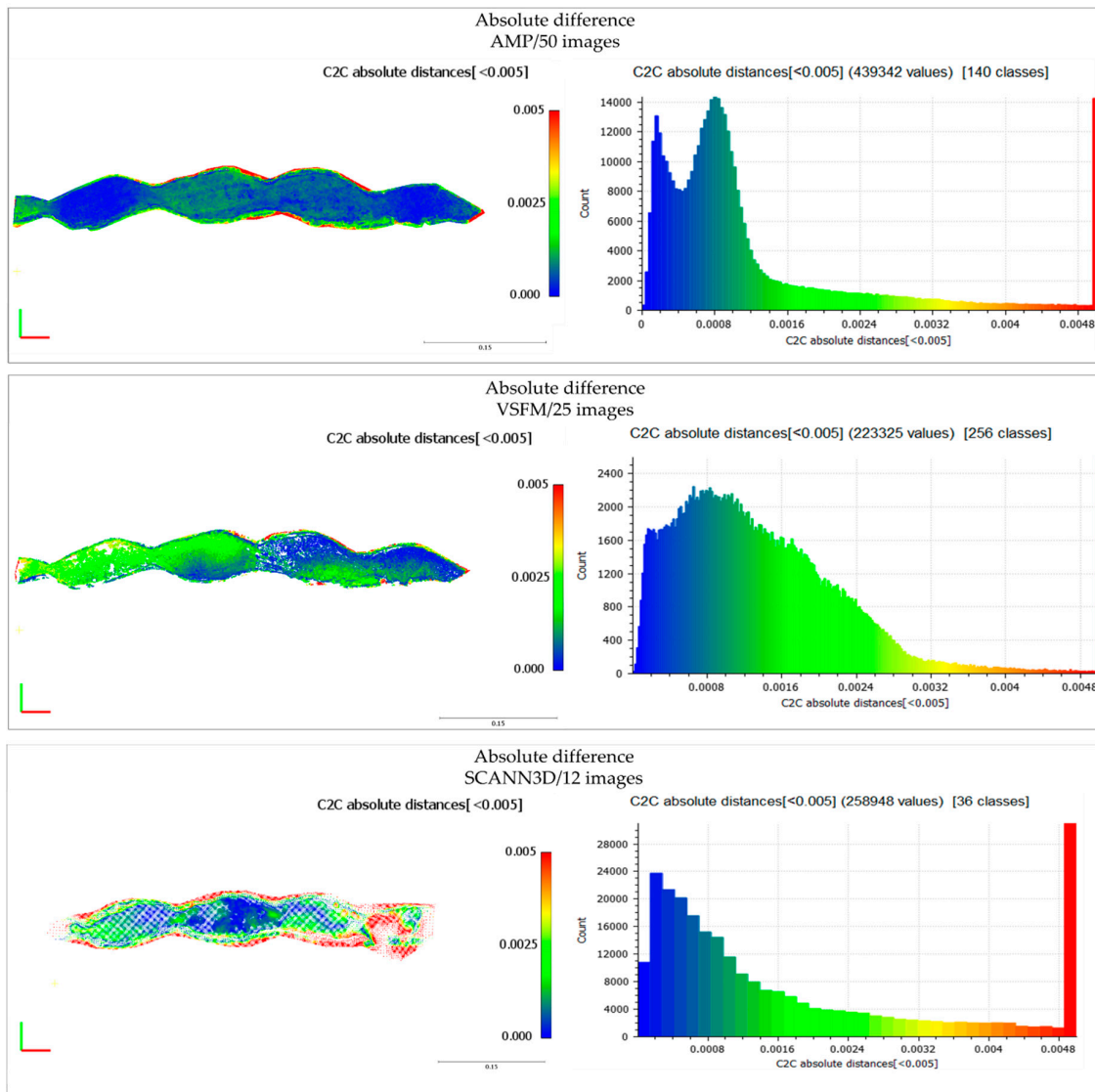
Figure 7. A comparison of the point clouds generated with different software products and the reference solution.

5.2. Estimate of the Size of the Pixel on the Rock Sample

As demonstrated, the procedure described so far is easily feasible, even by non-experienced operators, thanks to the enormous potential of modern software. However, depending on the objective of the survey, it is of primary importance to establish, as the first key step, the final definition of the product to be obtained, namely the representation scale. In the context of this research work,

it is fundamental to obtain 3D representations of objects at large or very large scales, in order to be able to conduct specific geomorphological analyses. These particular objects, as well as some areas of the underground caves, cannot always be inspected or, for different reasons, a prolonged on-site observation over time is not allowed. It is therefore essential to set the photogrammetric survey in such a way as to obtain the desired detail or, on the contrary, according to the operating conditions of the relevant site, in order to assess which resolution can be reached by the instruments that are available to us.

To this purpose, it is necessary to estimate which is the so called Ground Sample Distance (GSD) [17] achievable by the camera of our smartphone device, namely the distance between the center of two pixels expressed in the unit of measurement of the object, as a function of the distance acquisition. GSD was introduced in the field of aerial photogrammetry as fundamental information for flight planning [53]. However, by defining a target GSD for a specific project of terrestrial photogrammetry, it is possible to calculate the most suitable acquisition distance with respect to the object of study to be maintained.

Table 10 lists the characteristics of the processed data as shown in the APM report; these are average values calculated with respect to the images of the entire dataset.

Table 10. GSD estimated by the software AMP.

Number of Images	Avg Acquisition Distance [cm]	Avg GSD [mm/px]	Overlap [%]
50	37.7	0.0920	90
25	38.6	0.0942	85
312	43.3	0.1057	70

The GSD is a function of the acquisition distance, focal length and digital sensor size, as explained by the following formula:

$$GSD = \frac{d * p}{f} \tag{1}$$

where d is the image acquisition distance, p is the pixel pitch and f is the focal length.

The GSD values can be estimated by taking into account the characteristics of the OnePlus 5 smartphone and the focal length calculated in previous sections, in order to define the model resolution that can be reached with this specific device. Since the relationship between the GSD and the acquisition distance is directly proportional, it is easy to deduce that, if with an acquisition distance of 37.7 cm we have an average GSD of 0.0920 mm/px, at a distance of 100 cm the GSD will reach a size of 0.2440 mm/px, and at 200 cm it will increase to 0.4881 mm/px. However, it is important to mention that, for rather compact objects for which large-scale representations are desired, it is possible with such limited acquisition distances to have significant differences between the nearest parts of the object and the parts that are more distant from the camera. During the acquisition planning phase, it is therefore recommended that one specify both a target GSD and a minimum GSD to allow some flexibility because of object variation.

6. A Real Case Study: The Bossea Cave

As the procedure described above has given promising results, the same methodology has been exploited for the 3D modeling of three surfaces located inside a real cave: the Bossea Cave. Furthermore, in this case, the OnePlus 5 smartphone was used to capture several videos (about one minute for each area). All the frames, with a resolution equal to 3840 × 2160 pixel, were extracted, and many of them were selected to be further processed for the 3D model reconstruction. During the acquisition of the videos, no flash or other additional lights were used, other than those already present in the cave, since, although the lighting conditions were not optimal, we wanted to avoid increasing the reflectivity of the wet surfaces. As we were not able to cross the parapet (for safety reasons), we tried

to acquire as much information as possible by moving the smartphone to different heights and varying the acquisition angle.

Starting from the videos, a limited number of frames have been selected in order to guarantee sufficient overlap between them.

6.1. Data Processing

Subsequently, the image processing was conducted, exploiting the software AMP, as it combines a good accuracy of the result with the possibility of controlling the various process phases and a high density of the final 3D data. In order to guarantee a certain speed of data processing and 3D reconstruction, without renouncing the completeness of the description of the object, it was decided to use a *Medium* quality setting for the image alignment and a *High* quality setting for the point cloud densification.

Figure 8 and Table 11 show the characteristics of the resulting 3D dense point clouds reconstruction.
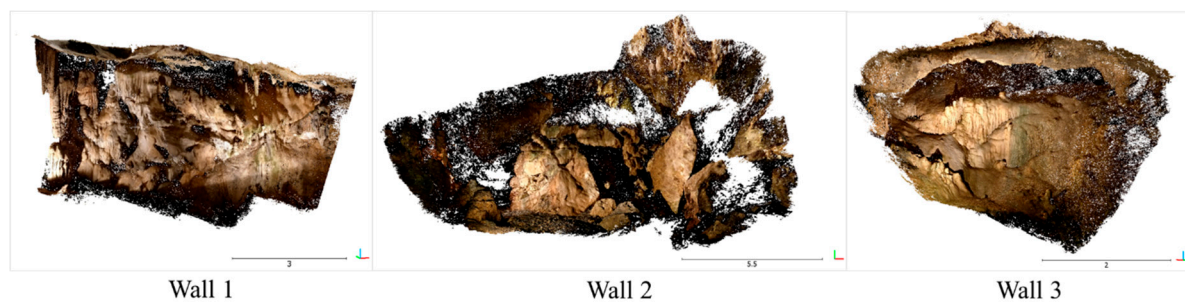

Wall 1   Wall 2   Wall 3

Figure 8. The resulting dense point clouds.

Table 11. The characteristics of the dense point clouds obtained with AMP from the smartphone images of the surfaces of the cave.

Study Area	Images	Aligned Images	Points Dense Point Cloud	Computational Time
Wall 1	138	138	7,948,534	47 min
Wall 2	106	106	8,872,074	32 min
Wall 3	90	90	5,688,908	23 min

The same datasets were subsequently processed by the SCANN3D application. However, likely due to the higher number of images and to the particularly poor light conditions, the application was able to reconstruct the models only partially, with a quality that was not excellent in terms of the number of points and completeness of the model, even when setting the highest quality level. Nevertheless, it should be pointed out that SCANN3D is not an application created for scientific purposes and requires specific environmental conditions and acquisition strategies in order to be able to correctly model an object through images.

6.2. Analysis of the Results

The geometric quality of the entire point cloud was estimated using as reference data a terrestrial laser scanner (TLS) survey of the entire cave, which was realized for a previous project.

The used instrument was the FARO Focus 3D. It is a well-known phase-shift laser scanner with a distance accuracy of up to ±0.002 m and a range from 0.6 m up to 120 m. For the registration of the LiDAR point clouds, some markers, spread over the area, were measured with topographic instruments. As usual, first a reference network was realized with Global Navigation Satellite System (GNSS) techniques to know high-precision vertices. The survey of the two vertices was made through a GNSS double frequency and multi-constellation receiver in static modality standing on each point for about 60 min. Afterwards, the total station was employed to realize a network of points and to measure

the position of the markers. All measurements were subsequently adjusted with the MicroSurvey StarNet v.7.0 software, in order to obtain the final coordinates, which were determined with a millimeter accuracy. The final model was obtained with the FARO Scene software, with a registration error of about 1 cm, compatible with the accuracy of the topographic network and the instrument itself.

The photogrammetric models were subsequently compared with the LiDAR point cloud using the software CloudCompare. To this purpose, it was necessary to scale the model and rototranslate it, so that it was placed in a reference system compatible with that of the reference model. Therefore, the ICP algorithm, implemented in the CloudCompare software, was used again. Subsequently, it was possible to estimate the Euclidean distances between the points of the two representations.

Figure 9 and Table 12 show some significant results deriving from this procedure in histograms and in statistical distributions. The computation of the differences confirmed the centimetric accuracy of the models generated by AMP, and showed that the surfaces have larger discrepancies where there are holes in the LiDAR point cloud. Other errors can be attributed to the noise affecting the point clouds, due to the low light and large shaded areas, as well as to the reflective surfaces resulting from high humidity. In order to improve the process of bundle block adjustment and facilitate the correct alignment of the images, some topographic measured markers should be used. Furthermore, some of the same markers could be employed as check points for a further evaluation of the accuracy of the dense reconstruction.
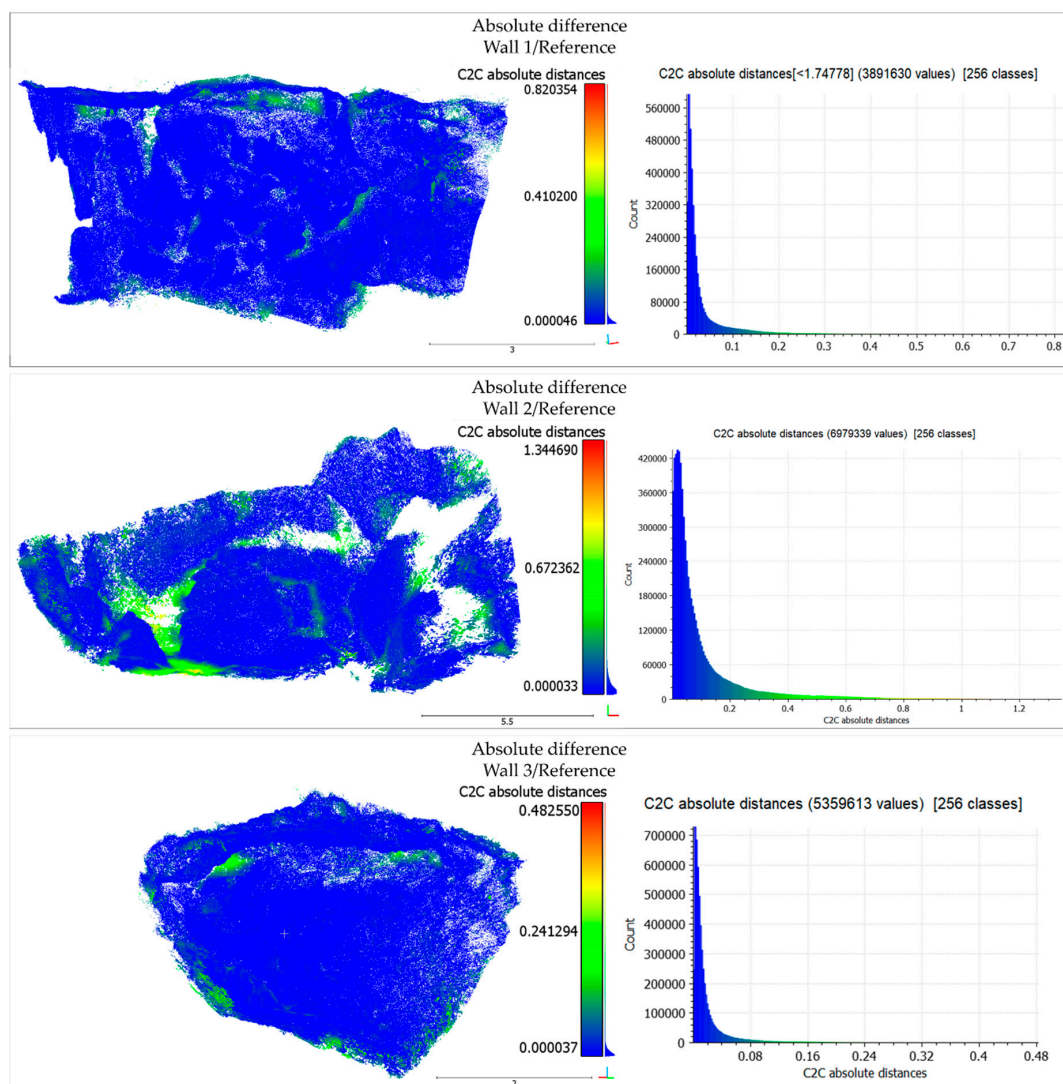

Figure 9. A comparison of the dense point clouds with the reference solution.

Table 12. The statistical results from the distance comparison between the smartphone point clouds and the reference point cloud generated with the terrestrial laser scanner.

Study Area	Min [m]	Max [m]	Mean Distance [m]	Std Deviation [m]
Wall 1	0.000	1.174	0.0303	0.0445
Wall 2	0.000	1.345	0.0958	0.1245
Wall 3	0.000	0.482	0.0173	0.0246

7. Conclusions

The aim of this work was to analyze and propose a low-cost technology and methodology in order to realize 3D representations of objects of geomorphological interest. The aspects examined concern the entire process, which, from the acquisition phase, leads to the analysis of the final result accuracy. The issue to be tackled was to find solutions able to guarantee a certain accuracy of the data but at the same time able to fulfill the requirements of portability, low-cost, speed in data acquisition and flexibility in order to adapt their use in complex environments. Nowadays, smartphone devices are accessible to everyone in the mass-market, and although they have not been created for metric purposes they solve some problems related to the acquisition of 3D data.

However, due to these devices being generally realized for amateur purposes, it has become crucial to study their characteristics, analyze their behavior and test performances in depth, before using them in surveying operations. The calibration problem was one of the addressed topics in this research, because in photogrammetry a good accuracy of the final model cannot ignore the correct calculation of the internal parameters of the camera.

Attention was dedicated to the planning of the acquisition procedures, in order to guarantee the capture of a limited number of images for the three-dimensional reconstruction, sufficiently overlapping between them. Moreover, it has been demonstrated how the acquisition geometry influences not only the accuracy but also the resolution of the 3D model, through the estimation of the GSD.

In particular, the proposed methodology was, first, assessed in a controlled environment, on a single small object (large scale) and, subsequently, tested on real cave walls (small scale). Many tests were performed in order to identify the limits and potentialities of the proposed solution in the generation of three-dimensional models of a rock sample, combining the use of different image datasets with the potentialities of the Structure from Motion approach, implemented in many pieces of software, both commercial (AMP) and open source (VSFM), and in a smartphone application (SCANN3D). The used software tools offer ease of use and automation in the data processing and are able to realize suitable products in quite a similar way, although there are some slight differences in the results, including in the free and open source software. The potentialities of the proposed approach were evaluated by analyzing the RMSEs with respect to topographically measured control points and check points and by comparing the dense point clouds with the reference model obtained from a professional camera. The obtained results demonstrated that the behavior of the different software tools in terms of performance is quite different. On the one hand, the data processed with software running on computer devices allow us to reach accuracies of about 1 mm on control points, which was also confirmed by the comparison with the reference point cloud; this makes them suitable models to be exploited for scientific analyses. On the other hand, it was proven that the algorithms exploited by the SCANN3D application must be optimized in order to be able to locally generate a 3D model on a smartphone platform. However, it is clear that "mobile" photogrammetry strongly depends on the computational power of the devices, and, above all, we still need hardware that is more powerful than that of the current generation of high-end smartphones, if the goal is to generate 3D models in real-time.

Given the good results of this first attempt to use smartphones for the generation of 3D models of objects of geomorphological interest, some preliminary tests were conducted in order to investigate how the proposed methodology behaves in different environmental conditions, such as in a real underground cave, in which the various operating conditions could affect the results:

- the lighting conditions: poor lighting conditions would require a higher ISO and a consequent loss of image quality; on the other hand, artificial lights and a high light variation modify the radiometry and create reflections both on wet surfaces and shadows, which negatively influence the reconstruction of the point cloud;
- the possibility of movement: the size of the paths, the distance from the walls and the possibility of reaching and seeing all the spaces influence the data acquisition operations;
- the humidity and temperature: these may affect the functionality of the sensors.

Furthermore, in this case, the obtained models showed that, by following a rapid acquisition process and using a limited number of inputs obtained from a mass-market sensor, it is possible to generate representations at scales large enough for the main geomorphological analyses. The procedure should be further stressed using different types of smartphones, which, in the most recent versions, integrate high-quality sensors and powerful processors.

In the future, the next step of this study will focus on deepening the possible application of smartphone data in underground caves for real-time navigation purposes, where GNSS receivers are not available and the environment could be quite critical or could involve limited accessibility. To this purpose, a possible solution could derive from visual odometry, which would be able to compute in real-time both the motion of the smartphone device and the 3D structure by exploiting the acquired 2D frames.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mattes, J. Underground fieldwork—A cultural and social history of cave cartography and surveying instruments in the 19th and at the beginning of the 20th century. *Int. J. Speleol.* **2015**, *44*, 251–266. [CrossRef]
2. Weinmann, M. *Reconstruction and Analysis of 3D Scenes: From Irregularly Distributed 3D Points to Object Classes*; Springer: Cham, Switzerland, 2016.
3. Brown, K.R.; Chalmers, A.; Saigol, T.; Green, C.; D'errico, F. An automated laser scan survey of the Upper Palaeolithic rock shelter of Cap Blanc. *J. Archaeol. Sci.* **2001**, *28*, 283–289. [CrossRef]
4. Westerman, A.R.; Pringle, J.K.; Hunter, G. Preliminary lidar survey results from Peak Cavern vestibule, Derbys, UK. *Cave Karst Sci.* **2004**, *30*, 129–130.
5. Canavese, E.P.; Forti, P.; Naseddu, A.; Ottelli, L.; Tedeschi, R. Laser scanning technology for the hypogean survey: The case of Santa Barbara karst system (Sardinia, Italy). *Acta Carsol.* **2011**, *40*, 65–77. [CrossRef]
6. Lerma, J.L.; Navarro, S.; Cabrelles, M.; Villaverde, V. Terrestrial laser scanning and close range photogrammetry for 3D archaeological documentation: The Upper Palaeolithic Cave of Parpalló as a case study. *J. Archaeol. Sci.* **2010**, *37*, 499–507. [CrossRef]
7. Grussenmeyer, P.; Landes, T.; Alby, E.; Carozza, L. High resolution 3D recording and modelling of the Bronze Age cave "Les Fraux" in Périgord (France). *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2010**, *38*, 262–267.
8. Puchol, O.G.; McClure, S.B.; Senabre, J.B.; Villa, F.C.; Porcelli, V. Increasing contextual information by merging existing archaeological data with state of the art laser scanning in the prehistoric funerary deposit of Pastora Cave, Eastern Spain. *J. Archaeol. Sci.* **2013**, *40*, 1593–1601. [CrossRef]
9. Hoblea, F.; Delannoy, J.J.; Jaillet, S.; Ployon, E.; Sadier, B. Digital tools for managing and promoting karst geosites in southeast France. *Geoheritage* **2014**, *6*, 113–127. [CrossRef]
10. Mohammed Oludare, I.; Pradhan, B. A decade of modern cave surveying with terrestrial laser scanning: A review of sensors, method and application development. *Int. J. Speleol.* **2016**, *45*, 71–88. [CrossRef]

11. Fabbri, S.; Sauro, F.; Santagata, T.; Rossi, G.; De Waele, J. High-resolution 3-D mapping using terrestrial laser scanning as a tool for geomorphological and speleogenetical studies in caves: An example from the Lessini mountains (North Italy). *Geomorphology* **2017**, *280*, 16–29. [CrossRef]

12. Bassani, M.; Grasso, N.; Piras, M. 3D GIS based evaluation of the available sight distance to assess safety of urban roads. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *40*, 137. [CrossRef]

13. Sirmacek, B.; Shen, Y.; Lindenbergh, R.C.; Diakite, A.A.; Zlatanova, S. Comparison of Zeb1 and Leica C10 indoor laser scanning point clouds. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *3*, 143–149. [CrossRef]

14. Chiabrando, F.; Sammartano, G.; Spanò, A. A comparison among different optimization levels in 3d multi-sensor models. A test case in emergency context: 2016 italian earthquake. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2017**, *42*, 155–162. [CrossRef]

15. Russo, M.; Remondino, F.; Guidi, G. Principali tecniche e strumenti per il rilievo tridimensionale in ambito archeologico. *Archeologia e Calcolatori* **2011**, *22*, 169–198.

16. Sammartano, G.; Spanò, A. Point clouds by SLAM-based mobile mapping systems: Accuracy and geometric content validation in multisensor survey and stand-alone acquisition. *Appl. Geomat.* **2018**, *10*, 317–339. [CrossRef]

17. Kraus, K. *Photogrammetry. V. 2, Advanced Methods and Applications*; Dummler: Bonn, Germany, 1997.

18. Szeliski, R. *Computer Vision: Algorithms and Applications*; Springer Science & Business Media: Berlin, Germany, 2010.

19. Kanade, T.; Okutomi, M. A stereo matching algorithm with an adaptive window: Theory and experiment. In Proceedings of the 1991 IEEE International Conference on Robotics and Automation, Sacramento, CA, USA, 9–11 April 1991; pp. 1088–1095.

20. Ullman, S. The interpretation of structure from motion. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* **1979**, *203*, 405–426.

21. Núñez, M.A.; Buill, F.; Edo, M. 3D model of the Can Sadurní cave. *J. Archaeol. Sci.* **2013**, *40*, 4420–4428. [CrossRef]

22. Citton, P.; Romano, M.; Salvador, I.; Avanzini, M. Reviewing the upper Pleistocene human footprints from the 'Sala dei Misteri'in the Grotta della Bàsura (Toirano, northern Italy) cave: An integrated morphometric and morpho-classificatory approach. *Quat. Sci. Rev.* **2017**, *169*, 50–64. [CrossRef]

23. De Waele, J.; Fabbri, S.; Santagata, T.; Chiarini, V.; Columbu, A.; Pisani, L. Geomorphological and speleogenetical observations using terrestrial laser scanning and 3D photogrammetry in a gypsum cave (Emilia Romagna, N. Italy). *Geomorphology* **2018**, *319*, 47–61. [CrossRef]

24. Aicardi, I.; Lingua, A.; Piras, M. Evaluation of mass market devices for the documentation of the cultural heritage. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2014**, *45*, 17–22. [CrossRef]

25. Google Play. Available online: https://play.google.com/store/apps/details?id=com.smartmobilevision.scann3d&hl=en (accessed on 11 April 2019).

26. App Store. Available online: https://apps.apple.com/us/app/3dsizeme/id964196048 (accessed on 29 July 2019).

27. Google Play. Available online: https://play.google.com/store/apps/details?id=com.sonymobile.scan3d&hl=en (accessed on 29 July 2019).

28. App Store. Available online: https://apps.apple.com/us/app/qlone-3d-scanner/id1229460906 (accessed on 29 July 2019).

29. App Store. Available online: https://apps.apple.com/us/app/canvas-by-occipital/id1169235377 (accessed on 29 July 2019).

30. Xperia XZ1. Available online: https://www.sonymobile.com/global-en/products/phones/xperia-xz1/#gref (accessed on 30 July 2019).

31. Fritsch, D.; Syll, M. Photogrammetric 3D reconstruction using mobile imaging. In *Mobile Devices and Multimedia: Enabling Technologies, Algorithms, and Applications*; International Society for Optics and Photonics: Bellingham, WA, USA, 2015; Volume 9411, p. 94110C.

32. Dabove, P.; Di Pietra, V.; Lingua, A.M. Close range photogrammetry with tablet technology in post-earthquake scenario: Sant'Agostino church in Amatrice. *GeoInformatica* **2018**, *22*, 463–477. [CrossRef]

33. Micheletti, N.; Chandler, J.H.; Lane, S.N. Investigating the geomorphological potential of freely available and accessible structure-from-motion photogrammetry using a smartphone. *Earth Surf. Processes Landf.* **2015**, *40*, 473–486. [CrossRef]

34. Aicardi, I.; Grasso, N.; Musci, M.A.; Piras, M. Geometrical 3D model generation of rock blocks using low-cost photogrammetry. *GEAM-GEoing. Ambien. Mineraria-GEAM-Geoeng. Environ. Min.* **2017**, *150*, 7–14.

35. Civita, M.; Vigna, B. Analysis of Bossea Cave Hydrogeological system (Maritime Alps, Italy). In *Proceeding of the Symposium 'Karst Water Resources'*; IHAS Publications: Wallingford, UK, 1985; Volume 161, pp. 101–113.

36. Remondino, F.; Fraser, C. Digital camera calibration methods: Considerations and comparisons. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2006**, *36*, 266–272.

37. Censi, A.; Scaramuzza, D. Calibration by correlation using metric embedding from nonmetric similarities. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 2357–2370. [CrossRef]

38. Octave. Available online: https://www.gnu.org/software/octave/ (accessed on 11 April 2019).

39. Camera Calibration Toolbox for Octave. Available online: http://www.sparetimelabs.com/cameracalib/cameracalib.php (accessed on 11 April 2019).

40. Dabove, P.; Di Pietra, V.; Lingua, A.M. Positioning Techniques with Smartphone Technology: Performances and Methodologies in Outdoor and Indoor Scenarios. *Smartphones Appl. Res. Perspect.* **2017**, *163*. [CrossRef]

41. Hartley, R.; Zisserman, A. *Multiple View Geometry in Computer Vision*; Cambridge University Press: New York, NY, USA, 2003.

42. Agisoft Metashape Professional. Available online: https://www.agisoft.com/ (accessed on 11 April 2019).

43. VisualSFM. Available online: http://ccwu.me/vsfm/ (accessed on 11 April 2019).

44. Snavely, N.; Seitz, S.M.; Szeliski, R. Photo tourism: Exploring photo collections in 3D. In *ACM Transactions on Graphics (TOG)*; ACM: New York, NY, USA, 2006; Volume 25, pp. 835–846.

45. Remondino, F.; Spera, M.G.; Nocerino, E.; Menna, F.; Nex, F. State of the art in high density image matching. *Photogramm. Rec.* **2014**, *29*, 144–166. [CrossRef]

46. Lowe, D.G. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* **2004**, *60*, 91–110. [CrossRef]

47. Farenzena, A.M.; Fusiello, A.; Gherardi, R. Structure-and-Motion Pipeline on a Hierarchical Cluster Tree. In Proceedings of the IEEE International Workshop on 3-D Digital Imaging and Modeling, Kyoto, Japan, 27 September–4 October 2009.

48. Fischler, M.A.; Bolles, R.C. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Commun. ACM* **1981**, *24*, 381–395. [CrossRef]

49. Triggs, B.; McLauchlan, P.F.; Hartley, R.I.; Fitzgibbon, A.W. Bundle adjustment—A modern synthesis. In *International Workshop on Vision Algorithms*; Springer: Berlin/Heidelberg, Germany, 1999; pp. 298–372.

50. Furukawa, Y.; Ponce, J. Accurate, dense, and robust multiview stereopsis. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 1362–1376. [CrossRef] [PubMed]

51. CloudCompare. Available online: https://www.cloudcompare.org/ (accessed on 11 April 2019).

52. Besl, P.J.; McKay, N.D. A Method for Registration of 3-D Shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* **1992**, *14*, 239–256. [CrossRef]

53. Neumann, K.J. Trends for digital aerial mapping cameras. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. (ISPRS)* **2008**, *28*, 551–554.