A flexible and swift approach for 3D image–based survey in a cave

L. Alessandri, V. Baiocchi, S. Del Pizzo, F. Di Ciaccio, M. Onori, M. F. Rolfo & S. Troisi

Applied Geomatics

ISSN 1866-9298

Appl Geomat
DOI 10.1007/s12518-020-00309-4

Springer

Applied Geomatics
https://doi.org/10.1007/s12518-020-00309-4

ORIGINAL PAPER

A flexible and swift approach for 3D image–based survey in a cave

L. Alessandri[1] · V. Baiocchi[2] · S. Del Pizzo[3] · F. Di Ciaccio[3] · M. Onori[2] · M. F. Rolfo[4] · S. Troisi[3]

Received: 20 January 2020 / Accepted: 1 April 2020

Abstract
In the geomatics field, modelling and georeferencing complex speleological structures are some of the most challenging issues. The use of conventional survey methods (for example, those employing total stations or terrestrial laser scanner) becomes more difficult, especially because of the space constraints and the often critical light conditions. In this work, a flexible and swift methodology to survey an in-progress excavation is presented, through image-based modelling techniques. The proposed approach allows obtaining a reliable and georeferenced three-dimensional model of the underground environments, preserving the integrity of the scene. The 3D model is scaled and georeferenced through three ground control points located just outside the cave, using data acquired by a double-frequency GNSS receiver in static session mode. Further targets were employed to check the deformation of the model inside the cave. The surveys were conducted on two archaeological sites: La Sassa cave and Guattari cave, both located in southern Latium.

Keywords Photogrammetry · Structure from motion · Cultural heritage · 3D modelling · Georeferencing · La Sassa cave · Guattari cave

Introduction

The archaeological excavation is a deliberate and scientific process of destroying (pre)historical evidence to obtain observations (Lucas 2001). Thus, archaeologists need to have this evidence documented in the best possible way, often in a short time. In cave environments, this process can be very challenging: caves are generally characterized by a lack of light, flowing water, high humidity, limited space and complex morphology. Their mapping is a challenging task that cannot be always performed by traditional topographic surveying methods (Gallay et al. 2015; Radicioni et al. 2019). Furthermore, some areas can only be reached by expert speleologists, because some passages could be very narrow and, in some cases, very dangerous if not tackled with adequate equipment and training. These conditions do not always allow using traditional topographic instruments or to employ the terrestrial laser scanning (TLS). Specifically, several limitations apply to TLS when used in cave environments, such as the difficulty of transportation and positioning (Angelini et al. 2017; Doneus and Neubauer 2005a, b; Stocchi et al. 2017). On the other hand, traditional topographic instrumentation,

✉ V. Baiocchi
valerio.baiocchi@uniroma1.it

L. Alessandri
l.alessandri@rug.nl

S. Del Pizzo
silvio.delpizzo@uniparthenope.it

F. Di Ciaccio
fabiana.diciaccio@uniparthenope.it

M. Onori
matteo.onori@uniroma1.it

M. F. Rolfo
rolfo@uniroma2.it

S. Troisi
salvatore.troisi@uniparthenope.it

[1] GIA, University of Groningen, Poststraat 6, 9712ER Groningen, The Netherlands

[2] DICEA, Sapienza University of Rome, Via Eudossiana 18, 00184 Rome, Italy

[3] Centro Direzionale Isola C4, Parthenope University of Naples, Naples, Italy

[4] Department of History, Culture and Society, University of Tor Vergata, Via Columbia 1, 00133 Rome, Italy

Published online: 12 May 2020

due to lateral refraction effects, should be integrated by a gyroscope (Velasco-Gómez et al. 2016) which, however, increases costs and makes the survey more complex (Baiocchi et al. 2016). In such harsh environments, only SLAM (simultaneous localization and mapping) image-based technique provides acceptable results in acquisition time and automation. Such solutions are generally based on the optimization of the classical SfM algorithms (Mouragnon et al. 2009). The scientific literature about the performances of these methodologies under different severe environments is increasing (Eyre et al. 2016; Mandelli et al. 2017; Sammartano and Spanò 2018). Although the elaboration time plays a fundamental role in this application, it is not so important as long as the instrument versatility. The SLAM approach assures real-time results, but it needs a great power computing system that cannot be provided by the device or by a remote server, due to a huge challenge to establish a communication in a cave environment. On the other hand, a prompt mapping of the site is essential to decide how to continue the excavation as well as to obtain quality observations about archaeologic deposits. Cave mapping is often accomplished with the aid of a fibreglass measuring tape or a laser rangefinder, a compass and an inclinometer (Gunn 2003). In the wider portions of the caves, a total station might be used. This approach is very versatile, but it is very time consuming especially when acquiring details about the passage walls, profiles, and cross-sections. Moreover, it relies on the operator's experience at interpreting and hand drawing the trend of the cave walls between survey stations. Therefore, it is necessary to adopt an approach which is able to achieve a high degree of details of hypogeal scenes and to provide an automatic solution.

In this paper, a methodology based on digital photogrammetry with a low-cost camera is described. The proposed approach allows achieving an automatic and reliable solution through a video recording of the scene that, at the end of the process, provides a three-dimensional model. This procedure is generally faster than the sketching procedure and can be partially performed in a laboratory rather than directly in the cave. Cross-sections, orthophoto, 3D measurement about length, areas or volumes can be extracted from the final 3D model. The morphology of the cave is also very important to reconstruct its ancient appearance and to gain new data on the cave formation. The great advantage is that such information is obtained in a non-subjective way, which is unbiased and scientifically repeatable (Jordan 2017). Digital photogrammetry has proven remarkable robustness and versatility in many fields (Bösemann 2005). A skilled person, who has the appropriate knowledge of all the involved steps, can determine a 3D model starting from a dataset of images. Indeed, several software programs allow the user to perform a correct workflow, thanks to the development of Structure from Motion (SfM) algorithms that are becoming increasingly reliable (Micheletti et al. 2015). As said above, a typical feature of archaeological caves is the presence of narrow spaces. In such conditions, the use of a wide-angle camera with a high field of view, such as a fisheye lens (Troisi et al. 2017; Perfetti et al. 2017) is particularly recommended to increase the overlap from two images and to decrease the number of images needed to obtain a reliable 3D model. This solution ensures a larger overlap between two successive frames, even if it creates a larger distortion than the standard lens.

In this paper, a setup workflow used for the survey of two archaeological caves is described. The second section contains a deep description of the setup survey that is divided into two main steps: GNSS and photogrammetric setup, the former being essential to georeference the final 3D model. The third section illustrates the features of the two case studies and the strategies adopted to find a reliable solution.

Motivations and methodology

Cave environments are exceptional systems which can provide strong evidence about sedimentary or tectonic structures. They constitute an important source of data from an archaeological point of view since they are generally more protected than the surface sites.

In these cases, 3D models frequently support and complete the research: the use of high-resolution 3D models of caves can improve the quality and quantity of data needed for the study (e.g. Abu Zeid et al. 2019; Rivero et al. 2019; Ruiz López et al. 2019).

Among the main methods generally used to acquire 3D datasets, those based on tachometric surveying instruments are the most complex to be deployed in these contexts, due to their intensive acquisition process and complex data processing. On the other hand, terrestrial laser scanning (TLS) systems give high-resolution and high-precision results but are generally very costly and often impossible to use due to space limitations.

In such environments, digital photogrammetry is becoming more and more used, since it may be able to provide new contributions to the traditional interpretations. Even if still challenging, this technology is based on digital camera images and constitutes a cheaper and more flexible method in various environments. These cameras are lightweight and can be embedded in mobile and remote shooting systems, and the results can be elaborated with commercial software packages, for example, Agisoft Metashape (Triantafyllou et al. 2019). In some specific and more critical cases, as through very narrow spaces where even the most capable operator may face difficulties (see Fig. 1), a method based on the Structure from Motion (SfM) algorithms allows a reduction of the survey time and complexity, still assuring good results (Troisi et al. 2017). The proposed methodology was tested in specific cave sites where the archaeological excavations were underway.

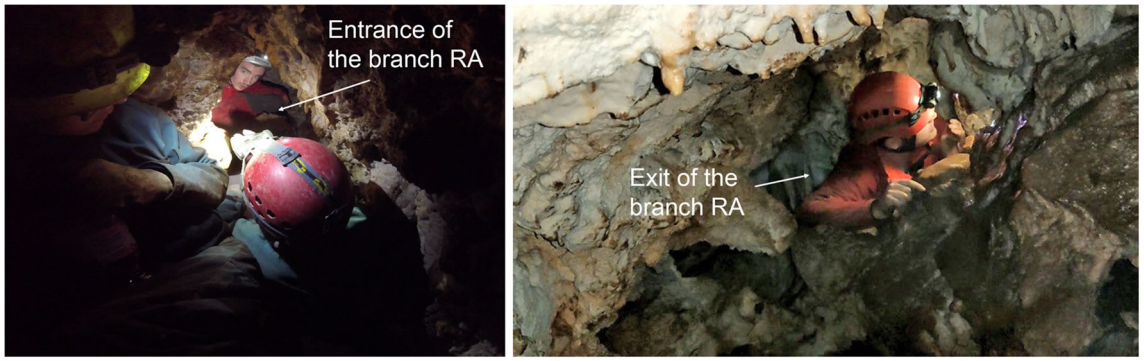

Fig. 1 The entrance (left) and exit (right) of the area called RA. At the end of the 70-cm circa wide tunnel, the Copper Age baby burial was found. Note the headlamps of the archaeologists, used for the photogrammetric survey

These scenarios helped the survey setup since some areas were already properly lighted to allow a safe excavation. Some other lights were placed in the darker areas. A light source was located behind the camera, and during the acquisition phase, the operator made it follow its movements to overcome the critical light conditions characteristic of narrow spaces. For example, during specific passages, the headlamp of the operator was used, as shown in Fig. 1. Such an approach was already experienced in Perfetti et al. (2018). This efficient and robust technique is widely used since it automatizes the photogrammetric procedures and lowers the costs, especially if compared with other methods.

Georeferentiation: GNSS setup

Every survey in an indoor or hypogea environment should be expressed in a local reference system (LSR) linked to a global reference system, such as the geocentric coordinate system ETRS89, Frame ETRF2000. Since the relation between the two systems is of fundamental importance for the model georeferencing, the employment of rigorous topographic methods, based on the use of a total station and GNSS receivers, is necessary. The achievable accuracy and precision are related to the adopted techniques (i.e. open and/or close traverses) and influenced by several parameters, for example, the morphology of the site and its position and extension (Baiocchi et al. 2017).

The model georeferencing has been achieved through a 3D Helmert transformation. A minimum of three ground control points (GCPs) for each cloud must be determined with high accuracy in the global reference system to get some degrees of redundancy. In this minimal configuration, a redundancy number of 2 is obtained. A higher number of GCPs are generally preferred, but that would have impacted on the survey execution speed, which was supposed to be swift. As a standard procedure, three tripods were stationed in the outside of the caves with a Topcon Legacy-E double-frequency GNSS receiver, ensuring the maximum visibility for the antenna to allow easy collimation during the elaboration of the images (Fig. 2).

The GNSS receiver acquired data for 40 min on each tripod. The data were then elaborated in post-processing mode using the nearest permanent station data managed by the Lazio Region, framed in the ETRF2000-RDN2008 system (EPSG: 7792) (Fig. 3).

The post-processing was performed using Topcon Pinnacle 1.0 (www.topcon.com), which is able to handle the GNSS data in the original format. The baseline estimation has also been achieved in the open-source software RTKLIB v. 2.4.2 (www.rtklib.com) with basically the same results. After the estimate of the coordinates of the points, the ellipsoidal heights of the antenna were reduced to the geoid according to the ITALGEO2005 model. In Italy, the latter provides better results than the global models (Barzaghi et al. 2014).

Photogrammetric survey

The photogrammetric workflow employed to obtain the final 3D model generally consists of three phases: acquisition, orientation and final plotting.

The acquisition

Modern software solutions allow reconstructing the 3D model starting from the image matching techniques. The images are elaborated using well-established algorithms to extract the geometry of the physical model through the detection of interest points on 2D images. Nevertheless, a loss of information may occur while acquiring photos; thus, a point on one image may not exist on the following one. It is also possible to have more than one correspondence or not even one, because of repetitive texture or acquisition noise.

The acquisition aims at assuring a minimum overlapping percentage of 65–85% between two successive images, meaning that all the points belonging to the area of interest are recorded at least three times (Monna et al. 2018). For this reason, while setting up the survey, the operators placed

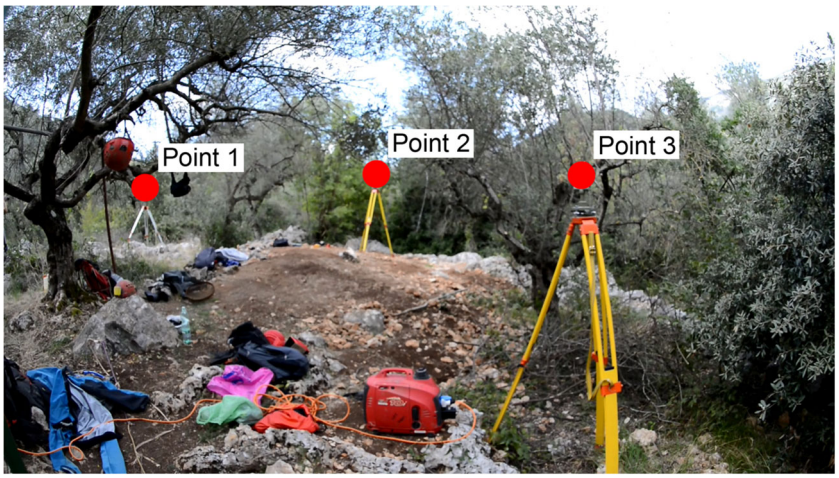

Fig. 2 The photogrammetric targets placed on tripods, where GNSS receivers were previously stationed

several targets inside and outside the caves to strengthen the camera network and to constrain the deformations of the models using the distances between the targets. Specifically, the internal targets were used to control the scale of the model, measuring the relative distances between them.

The surveys were performed with a full-frame DSLR Nikon D800E camera equipped with a Nikkor 16-mm fisheye lens (Fig. 4). The camera was in video mode, and the focus was set to infinity, at the highest aperture, to obtain a redundant number of images and to enlarge the pixel size on the sensor (useful in dark environment). The selected lens provides a diagonal FOV (field of view) of almost 180° which is particularly recommended both in narrow spaces (Perfetti et al. 2017) and when the site has a more prominent extension in length than in the other two dimensions, as in the case of tunnels (Troisi et al. 2017).

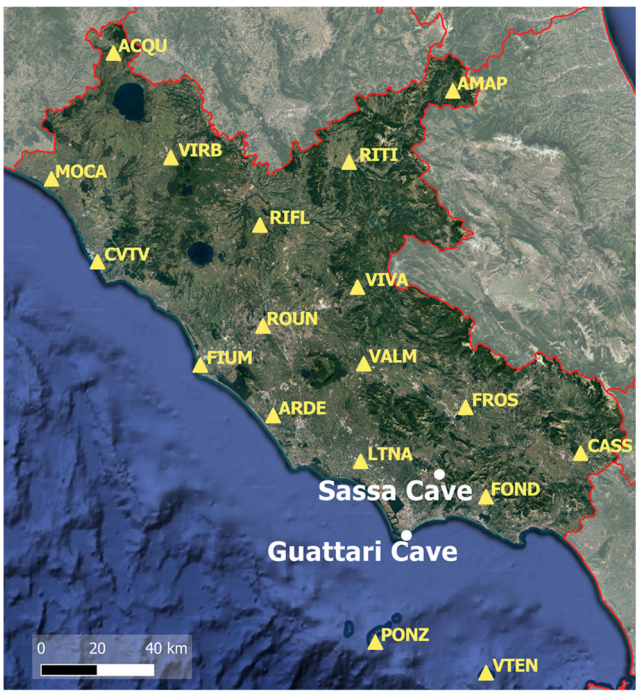

Fig. 3 The GNSS network of the Lazio region

The acquisition phase began outside, to allow the identification of the three GCP targets needed to georeference the model, then moved into the cave. The transition between the very bright environment and the dark location had been gradually performed, to avoid unexpected light variations (even if the ISO sensitivity parameter helps the automatic adjustment of the brightness gradient). Care should also be taken to avoid capturing moving features, such as people or shadows, which can affect the overall acquisition. The videos were then elaborated to collect the images. Providing specific guidelines on the minimum number of pictures required is difficult, so the sampling time may vary. A great number of images will produce a large number of matching features and more redundancy, but it will increase the computing time of the modelling step too.

The quality of the images is strongly affected by the locations of the light sources:

- A source placed behind the moving operator will cause different shadows on subsequent images of the same scene (Fig. 5).
- A light directed towards the camera will produce saturated images (Fig. 6).
- Lens flare patterns created by a light source spread widely throughout the scene and change position with camera movements (Fig. 6).

The elaboration

The images obtained from the videos were processed using Agisoft Metashape software which employs a user-friendly interface for the entire reconstruction workflow. The elaboration can be subdivided into three main phases:

Fig. 4 The Nikon D800E and Nikkor 16-mm fisheye lens used are shown on the top. Below, a generic frame recorded during the survey

1. Camera calibration and images orientation
2. Model scaling and georeferencing
3. Dense point cloud generation

The final accuracy and reliability of the model are strictly related to the correct orientation of the images, which depends on the camera calibration. Starting from the mathematical model chosen for the couple camera-lens, the calibration aims at determining the camera interior orientation parameters. The employed classical Brown model estimates these parameters through a combination of principal distance, principal point coordinates and pixel size and distortion effects modelling parameters. Especially in close-range photogrammetry, the interior orientation is obtained from several images within a self-calibration bundle adjustment. This is well-known as "on-the-job" calibration since it is directly performed during the 3D structure survey. Nevertheless, some specific geometric configurations of the camera are required to obtain coherent results. The test-field points should cover the sensor area to the greatest extent possible and the acquisition should be made with a 90-degree rotation around the view direction with all the photos pointing towards the test-field centre. Finally, no zoom should be applied to keep the interior orientation fixed. These technical objectives assure reliable results in terms of principal distance precision, internal parameter stability, decorrelation between distortion coefficients and accuracy of their estimation. Even if this procedure can be exploited before the survey, the network camera geometry (Fig. 7) could not be suitable for a self-calibration. This is the case of environments characterized by restricted spaces, where physical parameter (e.g. temperature and humidity) variations can affect the camera stability, reducing the calibration accuracy and the reliability of the overall survey.

A joint processing approach, useful in difficult and harsh contexts like this, is then proposed. It combines the "on-the-job" procedure with the classic self-calibration: the pre-calibration is performed in a laboratory, and then the parameters are refined during the images processing stage. For the first step, a calibration test field consisting of a 1-m cube built with aluminium profiles was used. Classic circular targets and

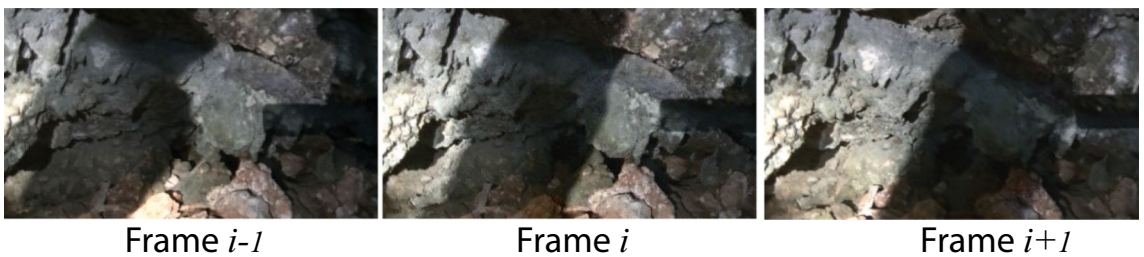

Fig. 5 The frame sequences that show how the shadow of the operator influences the light on the scene

Fig. 6 A portion of the frame is saturated due to direct light turned towards the lens

codified ones were applied both in the internal and external parts of the cube to create a 3D test field. The cube (shown in Fig. 8) has to be captured from different positions to ensure a good intersection of the camera rays. Some of the images were taken, with a 90-degree relative rotation around the optical axis. The videos were recorded in the same resolutions of those recorded for the surveys (1920 × 1080 and 1280 × 720).

The thus obtained pre-calibration parameters are used as the initial calibration; their refinement is then made during the elaboration, following the "on-the-job" approach. These parameters are in fact affected by unavoidable change factors, mainly related to different environmental conditions: this method allows obtaining reliable bundle adjustment solutions, helping the algorithm convergence.

Fig. 7 Typical camera network geometry

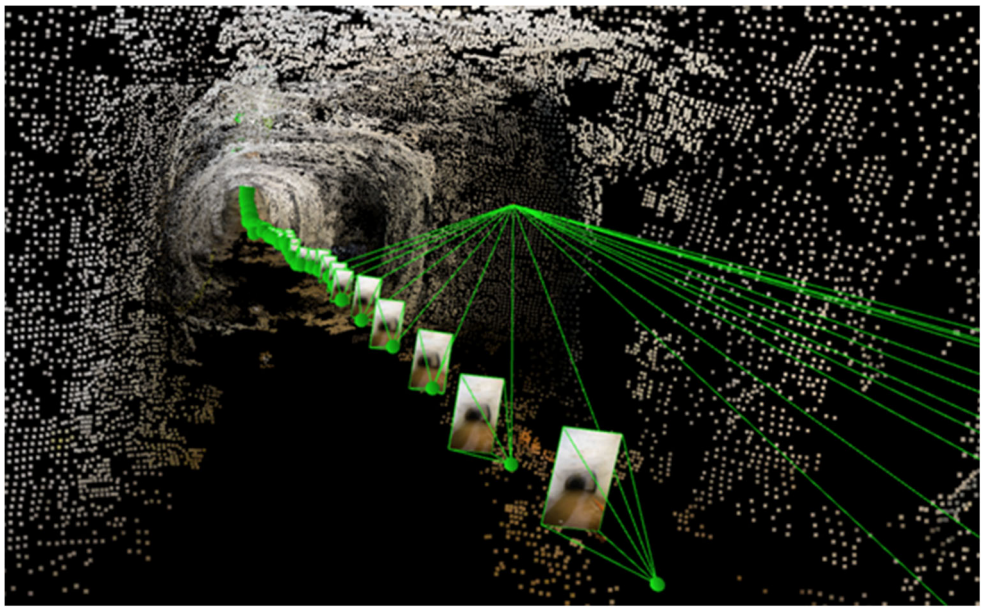

Fig. 8 Test-field calibration employed for the camera calibration

The followed procedure, fully automated in Metashape, is based on an automatic tie-point extractor algorithm, known as blob detector, which provides highly distinctive features invariant to image scaling and rotation, and partial invariant to brightness (Bay et al. 2006; Lowe 1999). This algorithm is particularly useful in the studied sites since it allows preserving the tie-points multiplicity; moreover, it provides a reliable solution for managing thousands of images in a reasonable time. The positions and orientations of each camera station are estimated through a bundle adjustment in the free network and, together with the 3D tie-points cloud of each image, result in a first photogrammetric model (the sparse point cloud), which is not georeferenced nor scaled. The model is oriented using the internal camera parameters obtained through the camera calibration and the reprojection total error is computed.

The second step aims at adjusting the model coherently with the scale constraints measured during the survey setup and the set of GCPs determined in the GNSS setup, for which the total error can be estimated. Other targets located in the scene help to estimate the deformation of the model, rerunning the bundle block adjustment algorithm in minimal constrain mode to adjust the camera parameters according to the "on-the-job" self-calibration. In this optimization phase, either the external and/or the internal orientation parameters are refined to best fit GCPs and scale constraints. The latter are mainly used in the inner part of the cave, far from the GCPs.

Finally, a cloud densification process is carried out. The obtained dense point cloud represents the final 3D model that can be analysed to evaluate the morphology and topography of the cave surface.

Afterwards, a triangular mesh model can be generated. The original colour is obtained from the pictures combined into a texture map which is wrapped on to the object.

Case studies

The ancient entrance and the deep branch of the La Sassa cave

La Sassa cave was discovered in 2015, during a systematic survey of the natural caves in the Ausoni Mountains run by two speleological groups: the Gruppo Grotte Castelli Romani and the Speleo Club Roma (Alessandri et al. 2019a; Alessandri and Rolfo 2015).

In the cave, a multi-layered deposit has been investigated, ranging from late Pleistocene to the II World War, when the cave was used as a shelter by the locals. A rich Bronze Age deposit and a Copper Age necropolis with hundreds of scattered human bones have also been found. The current entrance is a 2.5-m wide hole in the ceiling of the main room (room 1). However, some stratigraphic evidence suggested that, at least during the Copper Age and the Bronze Age, the entrance might have been elsewhere and specifically near a large niche which was completely filled with soil and rock debris at the beginning of the excavation. The 3D model was used to check the hypothesis and to observe the surface just above the niche (Fig. 9).

Moreover, a Copper Age baby burial was found in the deepest branch of the cave (area RA, around 12 m deep from the surface) which was particularly difficult to map with traditional methods due to the space constraints (Fig. 10). The 3D model allowed the precise localization of the area and gave new evidence about the deposit formation (and preservation).

La Sassa cave: results

Before starting the acquisition, three targets were positioned outside the cave (in the same positions of the GNSS antennas,

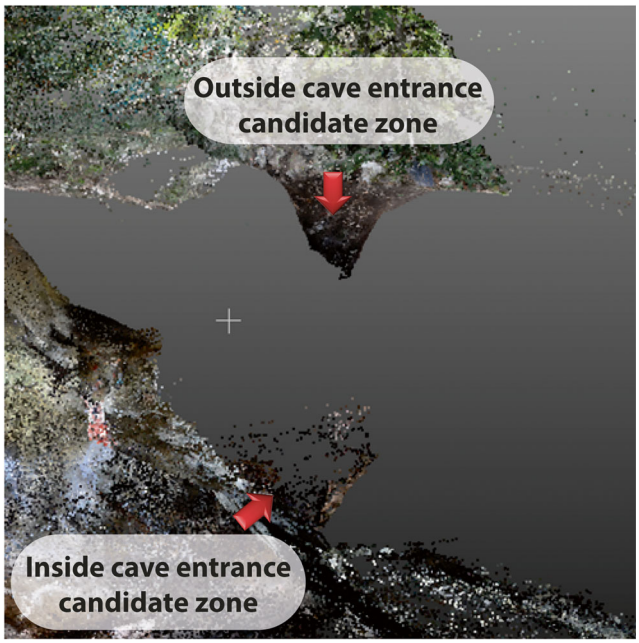

Fig. 9 Portion of the 3D model used to confirm the hypothesis of the Copper Age entrance

Fig. 11). Other targets were placed inside, to strengthen the camera network and to check the final results. The Nikon camera was set in video mode and recorded in standard HD 1080p (each frame is set to a resolution of 1920 × 1080 and the recording was performed in a progressive way) at 25 fps (frames per second), to obtain an 18.7-μm pixel size on the sensor (useful in dark environment) and to obtain a redundant number of images. The operator recorded a series of short-length videos because the cave terrain was highly irregular and there were no defined paths. During the pauses, two operators strategically placed some light sources to uniform the lighting of the scenario, carefully avoiding locating them in front of the camera or within the scene and to film people or moving objects (even if in some cases the shadow of the operator can still be seen). The acquisition phase lasted about 1 h: 30 min was spent for the 15 short videos recording, few minutes to measure the relative distances between the internal targets using steel tape, and the time left to settle the light sources and to evaluate the best path to be followed for a safe survey. It has to be noted that in the same cave a terrestrial laser scanner was used to make a survey: here, almost 6 h was employed to perform 12 laser stations to obtain the 3D model of the central, bigger portion of the cave. The inner part (the so-called sector RA), which has now been reconstructed using the proposed approach, could not be acquired due to the narrow spaces.

The 15 videos were sampled one frame every ten, collecting a total of 3899 images. These were oriented in a unique project in Agisoft Metashape, which provided a sparse points cloud of 524,168 points. The blob detector has been very important to preserve the multiplicity of the tie-points, allowing processing the images in a reasonable time. The bundle adjustment process provided an estimation of the camera poses and tie-points 3D positions. The three-dimensional model was scaled and georeferenced according to the GCPs

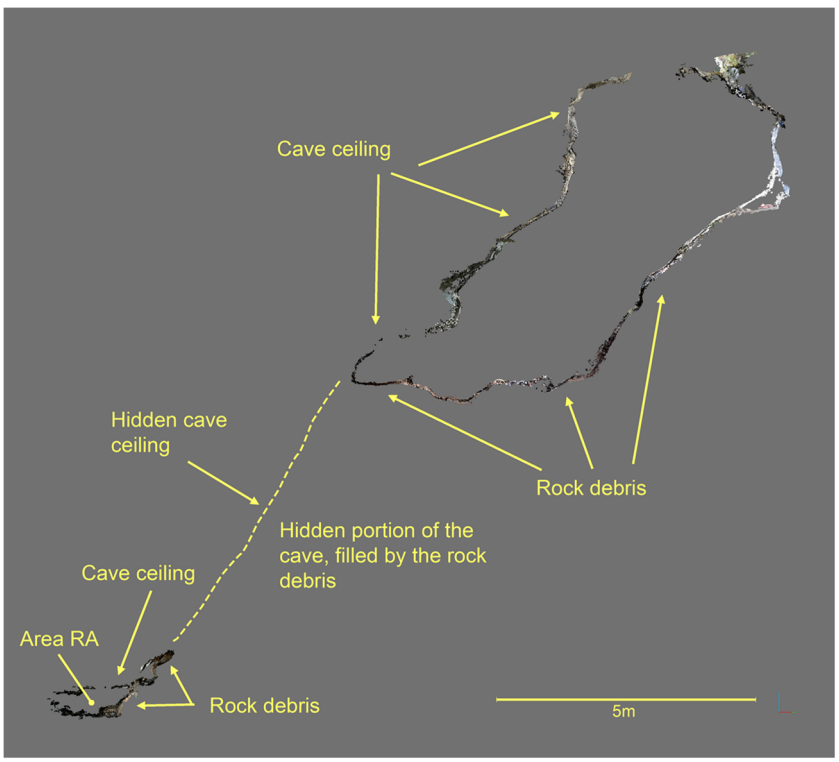

Fig. 10 Slice of the La Sassa cave. The area RA is "trapped" between the original walls of the cave and the post-Copper Age rock debris

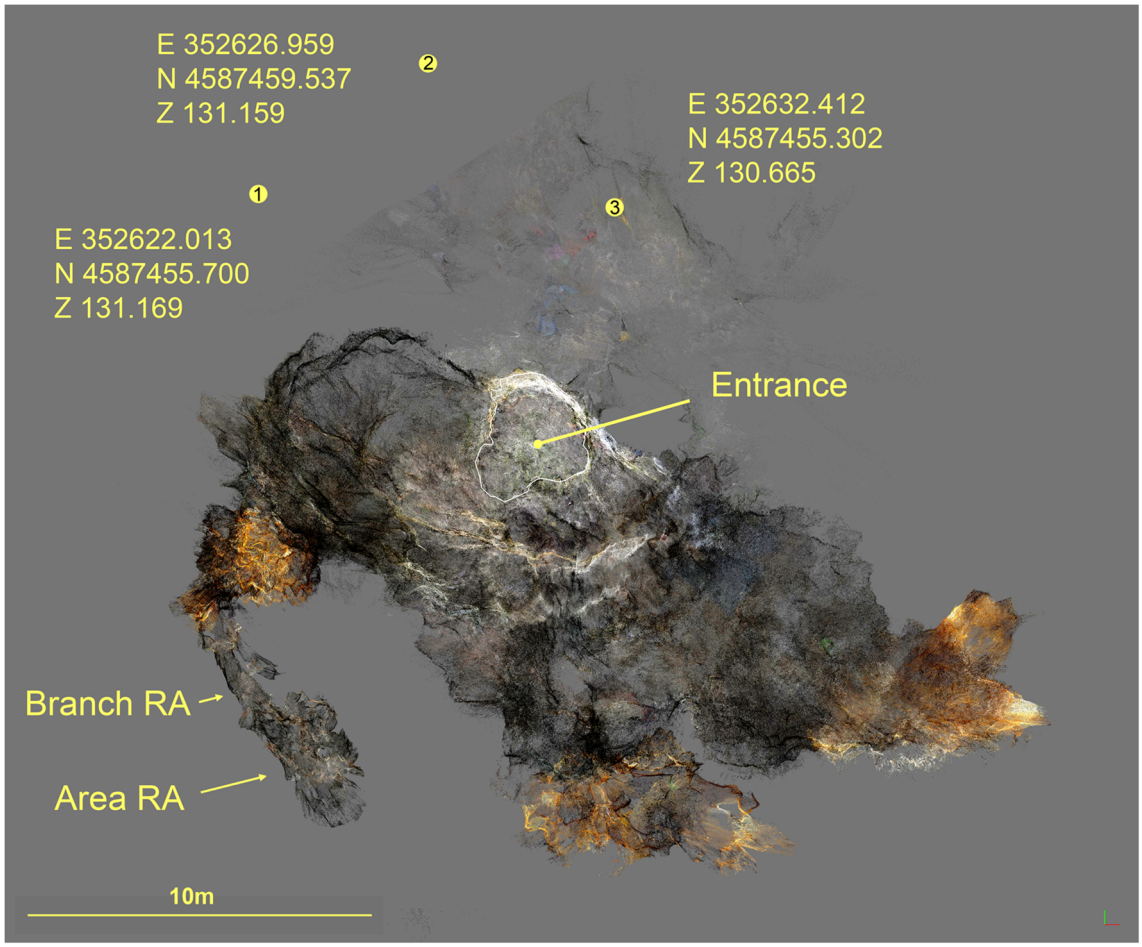

Fig. 11 Planimetric position of the external GCPs

and to the distances measured by a steel tape between internal targets, then the bundle adjustment algorithm was rerun to estimate the camera parameters according to the "on-the-job" self-calibration, providing a reprojection total error of 1.6 pixels. The total error on the GCPs was about 0.1 m, more than acceptable for this survey.

Finally, the image matching algorithm for the dense point cloud worked with an image downscaled to 25%, generating 6,910,198 points (Fig. 12).

To evaluate the reliability of the 3D model, an error analysis was performed using three targets as checkpoints. During the video recording, their coordinates were measured both on the way and back and differently labelled, allowing the software to distinguish them. The comparison between the two sets of 3D coordinates provided an estimation of both the cloud point precision and the model distortion, as reported in Table 1.

A first morphological approach allowed inspecting the surface of the cave and consequently to determine possible passages. The ancient entrance hypothesis was confirmed by the 3D model: as can be seen in Fig. 9, the outside excavation is compatible with the inside candidate entrance.

The paleo-beaches of the Guattari cave

The entrance of the Guattari cave is located in the Circeo promontory, in southern Latium, on the eastern slopes of the Mount Morrone. It was discovered by chance on 24 February 1939, during some quarry activities. One day later, the Italian archaeologist A. C. Blanc, who was warned of the discovery by the landowner, entered the cave and identified in a lateral branch one of the best-preserved Neanderthal skulls in Italy (Ascenzi 1991; Blanc 1939; Sergi 1974). The cave was then investigated by prof. L. Cardini and prof. A.G. Segre, until 1950. Two Neanderthal mandibles were also found at a later stage (Arnaud et al. 2015; Sergi 1954; Sergi and Ascenzi 1955). In all the cave, the surface of the paleosoil was characterised by the presence of several faunal remains. At the time of discovery, the skull was probably lying inside a circle of small stones, with the foramen magnum clearly widened and visible on top. These data, together with the isolated location, led some authors to hypothesise an episode of ritual cannibalism (Blanc 1961). However, the damage patterns, which is consistent with carnivore activity, the vertebrate faunal assemblage (40% cervids, 40% bovids, 10% hyena, 7%

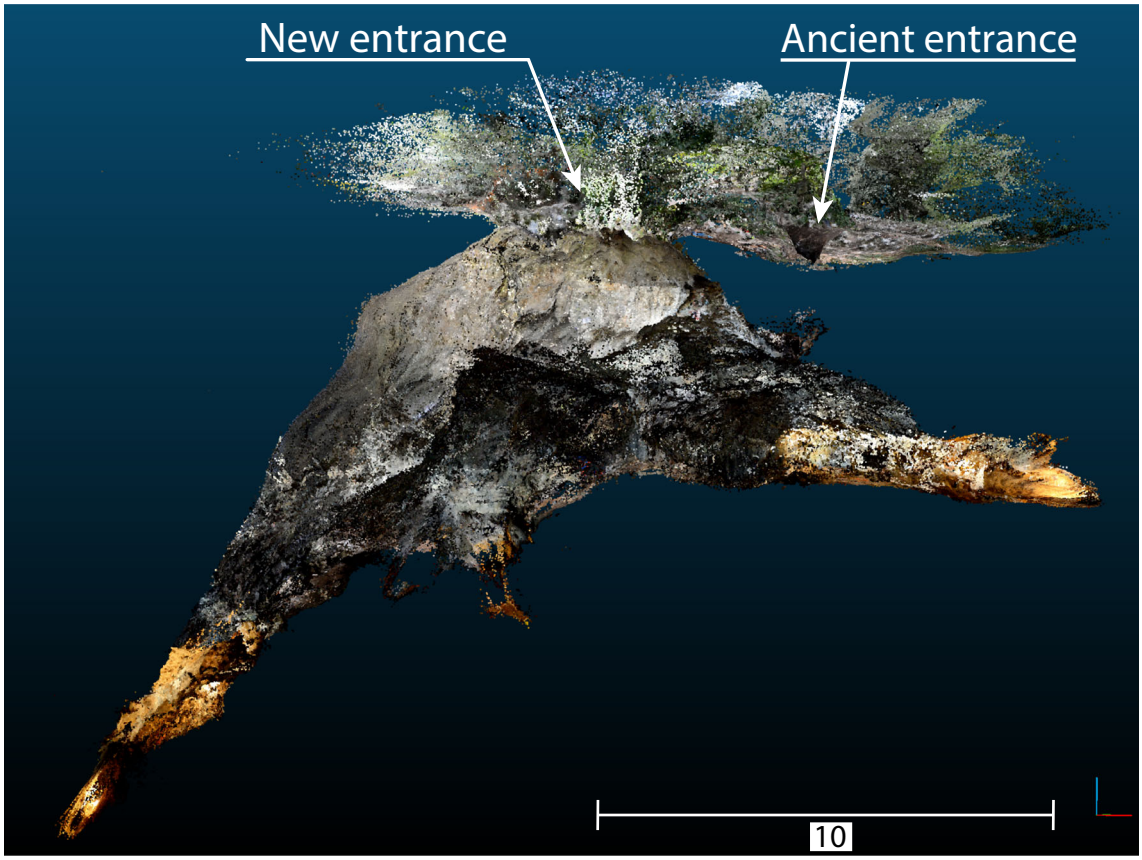

Fig. 12 The dense point cloud obtained after the image matching processing

horse, and 3% other mammals) and the particular pattern of fragmentation, along with the presence of hyena coprolites, all point to a hyena den (White et al. 1991). At the moment, the Guattari finds constitute a point of reference for the paleoenvironment and paleoclimate reconstructions of the Pleistocene Italian Tyrrhenian coast.

In 2019, the relevant Archaeological Superintendence promoted a series of initiatives both to assess the previous knowledge and to plan future researches. The chronology of the deposit and its origin are now once more under investigation. At the moment, some radiometric analyses are available to indirectly date the skull. A 51 ± 3 ka BP (kilo-annum Before Present) was obtained from calcite samples which were adherents to the adjacent stones and bones (U-series). Average date of 57 ± 6 ka BP for the surface and the layer 1 has been calculated from *Bos primigenius* teeth using the electron spin resonance method (Schwarcz et al. 1991a, b). This data suggests that the skull might have been deposited between 57 ka BP, the formation of the last deposit, and 51 ka BP, the start of the calcite formation. The development of new absolute dating techniques (uranium-thorium) and the identification of residual layers of a paleo-sea level dated at MIS (marine isotope stages) 5.1 (77 ± 9 ka BP) in the cave suggest a critical review of the layer chronologies and sequences (Marra et al. 2019). A 3D model of the cave and its deposit (Alessandri et al. 2019b) becomes then necessary to adequately link the absolute heights of the layer sequence with the reconstructed paleo-sea level (Fig. 13).

Table 1 The 3D targets' coordinate differences, in metres

Target Id	ΔX	ΔY	ΔZ	Total
6 (66)	− 0.012	− 0.041	0.001	0.042
7 (77)	− 0.031	− 0.021	− 0.003	0.037
8 (88)	0.018	0.000	0.019	0.026
Mean	− 0.008	− 0.020	0.005	0.035

Guattari cave: results

In this case, the operators decided to combine a camera and a smartphone to test the reliability of low-cost technologies in 3D model reconstructions (Dabove et al. 2019). For this reason, they assembled a handcrafted system consisting of a 25-cm bar on which a Nikon D800E camera with a 16-mm fisheye lens and a Huawei P9 mobile phone were installed. The videos were recorded with 1280 × 720-px resolution at 30 fps by the first device and in standard HD 1920 × 1080-px resolution at 60 fps by the second one.

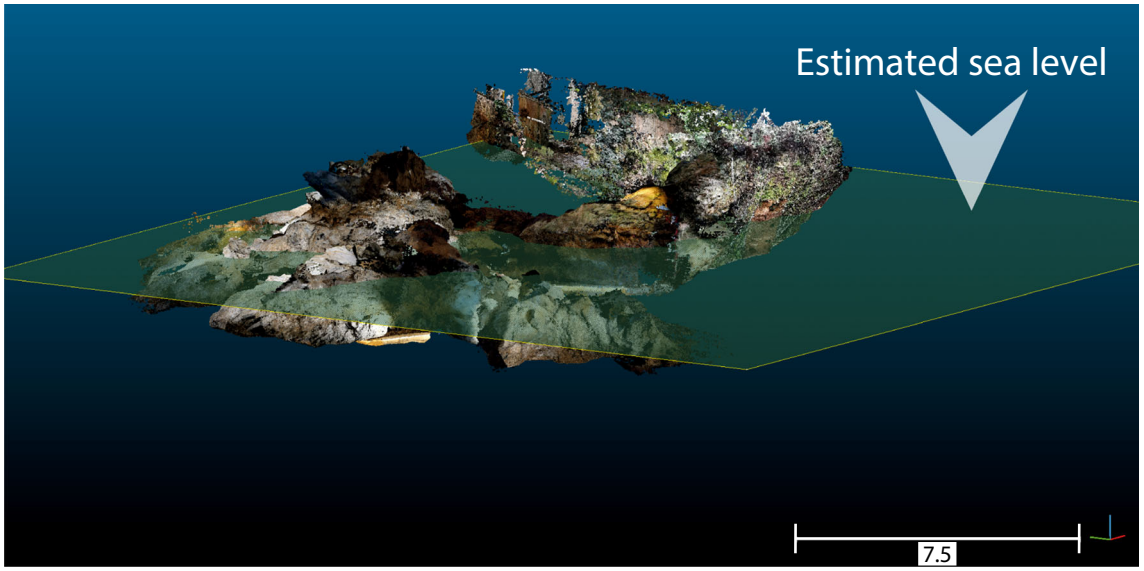

Fig. 13 Estimated paleo-sea level overlaid to georeferred 3D model of Guattari cave

The distances between 3 pairs of targets, placed at the entrance and inside the cave, were accurately measured to correctly scale the model and to constrain the deformations. During the survey, the operator recorded the 3 GCPs outside the cave and all the other targets, carefully avoiding fast movements and consequent unexpected light variations.

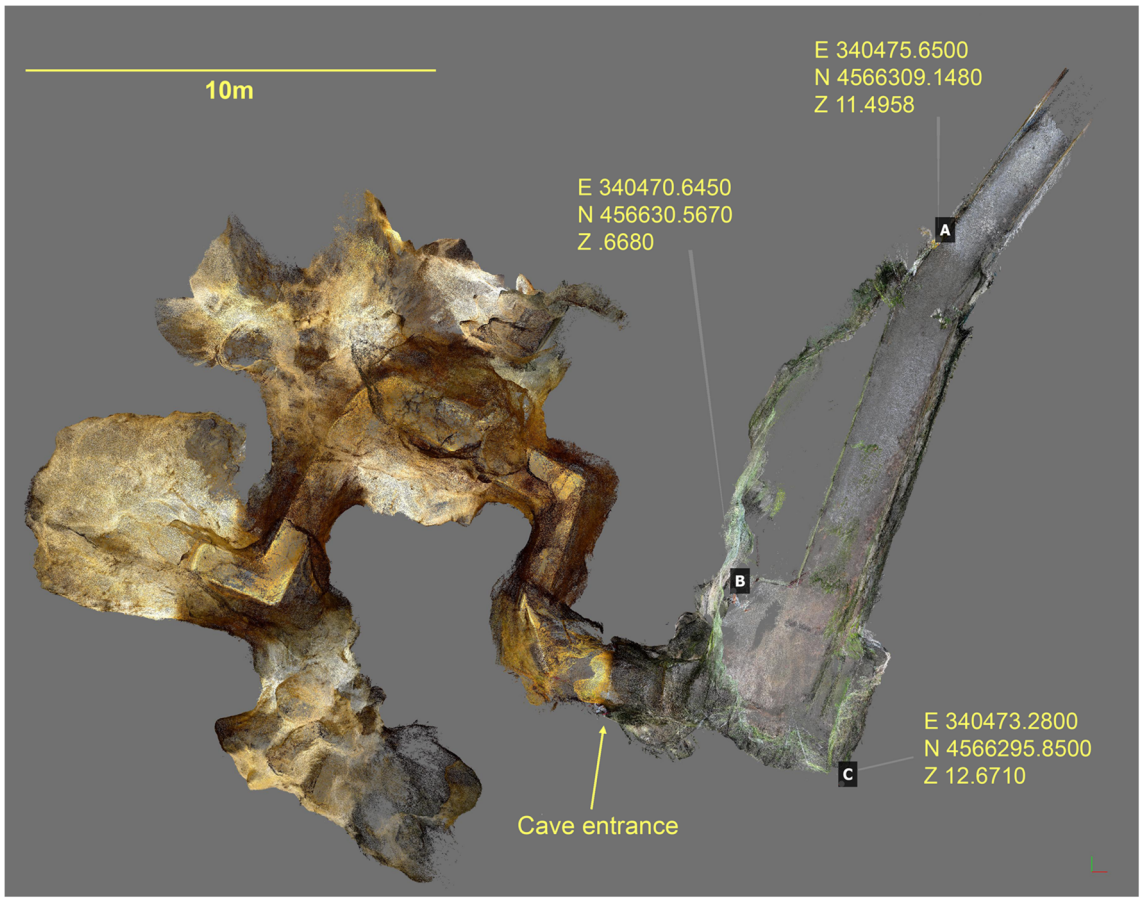

Fig. 14 The final dense point cloud of Guattari cave with the GCP positions

Table 2 Error comparison on scale constraints

Targets pairs	Measured distances	Nikon error	Huawei error
28–48	1.8	0.240	0.049
52–72	1.495	0.067	0.024
27–30	2.545	0.084	0.081
28–29	1.29	0.175	0.034

The fixed light sources, though, projected shadows on the walls which fortunately only resulted in a colour variation in the affected areas.

The total 10-min acquisition was sampled to obtain a frame every 1/3 s. The resulting image sets (more than 3000 for each camera) were oriented in Agisoft Metashape 1.5.0 after the cameras "on-the-job" self-calibration. The 3D models were then georeferenced and scaled using the 3 GCPs (Fig. 14).

The total errors on the GCPs represent the first difference between the sparse point clouds: a mean of 14 cm and a maximum of 18 cm values for the Nikon, a mean of 2 cm and a maximum of less than 3 cm values for the Huawei. A similar trend is visible in the scale constraint errors: Table 2 reports (in meters) the distances measured between some target pairs and the relative errors for both the cameras.

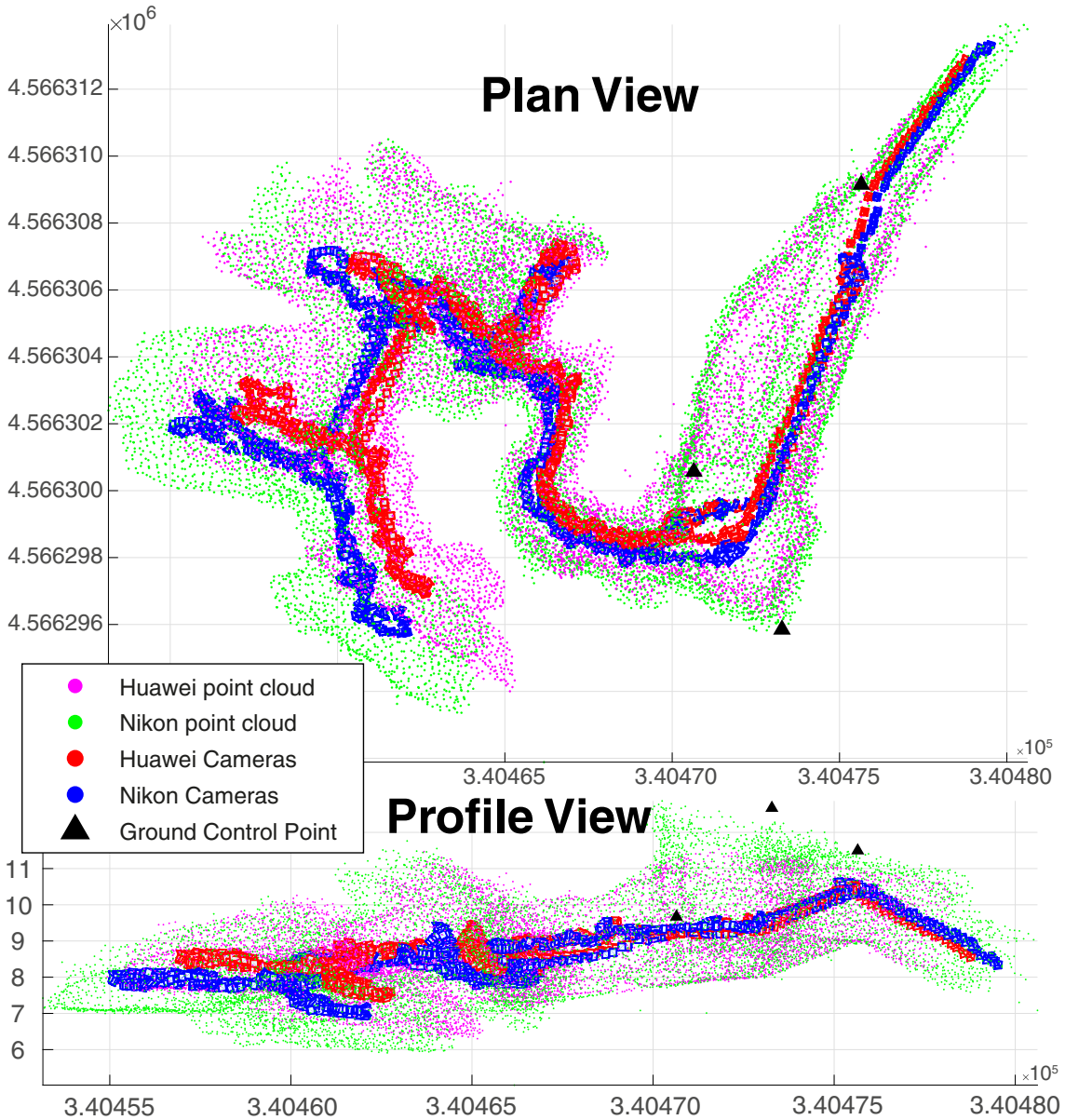

Fig. 15 The comparison between the two point clouds obtained by the Huawei and Nikon cameras

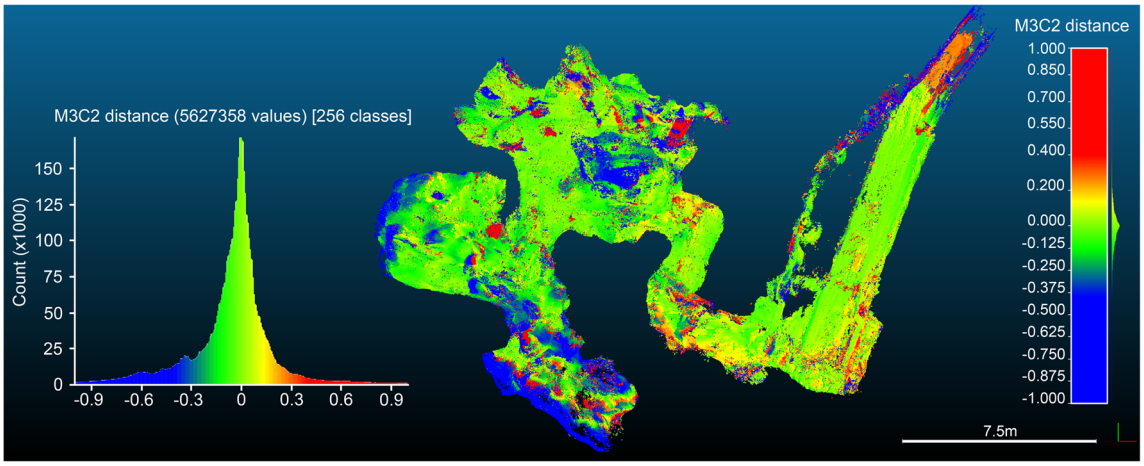

Fig. 16 The coloured distance histogram and the error map between the two registered point clouds

The external orientation parameters of the two cameras were estimated through the bundle block adjustment, and then the dense point clouds were obtained. A second difference showed up in this step. The software detected a different amount of points, 6 million for the Nikon and 14 million for the Huawei, even if the models were elaborated with the same parameters (Fig. 14): this is due to the different resolution of the images extracted from the videos. At this point, a more detailed comparison of the two models was necessary. Firstly, the point clouds were plotted together with the camera positions in MATLAB. Figure 15 shows an initial correspondence in the right part of the plot which decreases moving to the left, through the inner part of the cave, where the distance from the GCPs increases.

To verify if this issue was only related to a scale problem, the models were registered using the iterative closest point (ICP) algorithm in CloudCompare v. 2.6.3. This procedure provides the best models overlapping by computing the Helmert transformation parameters which minimise the cloud point distances. One of the two clouds needs to be taken as reference, so the Huawei one was chosen because of the lower error values it gave. The Nikon cloud was then transformed to get the best fit. The clouds were then compared again, calculating the signed distances along the normal direction of the reference model using the M3C2 plugin of CloudCompare (CloudCompare, 2.10.2 Zephyrus 2019). Figure 16 shows the resulting distances, summarised with a mean of − 0.059 m and a standard deviation of 0.031 m. Moreover, 12.3% of the distances came out to be greater than 0.5 m, while 2.4% of them were over 1 m.

The statistical analysis of the results evidenced an anisotropic distortion in the Nikon model, confirming that a simple Helmert transformation cannot minimise the differences between the two models.

Conclusions

In this work, a flexible and swift approach to survey hypogeum environments is described.

The aim is to propose and test a methodology to obtain 3D models by different image sets, each with a specific resolution. The proposed solution allows extremely fast survey operations, consisting of a simple video acquisition of the scene. Two different methods are then proposed to control the solutions: through checkpoints or through check scale bars, which both need to be previously placed in the scene.

This procedure was tested in two southern Latium caves, where archaeological researches are currently ongoing: La Sassa and Guattari caves.

The elaboration was performed in the Agisoft Metashape environment, which automatizes the entire workflow. More than 3000 images were oriented to obtain the primitive point clouds, which were then scaled and georeferenced using the 3 GCPs strategically placed outside the caves. The densification process finally provided the 3D model of the caves.

This procedure allows consistently reducing the survey time while ensuring a reasonable accuracy for archaeological scopes. The 3D model results prove to be a reliable way to register the caves' characteristics, allowing to easily test previously made archaeological hypothesis.

References

Abu Zeid N, Bignardi S, Russo P, Peresani M (2019) Deep in a Paleolithic archive: integrated geophysical investigations and laser-scanner reconstruction at Fumane Cave, Italy. J Archaeol Sci Reports 27:101976. https://doi.org/10.1016/j.jasrep.2019.101976

Alessandri L, Rolfo MF (2015) L'utilizzo delle cavità naturali nella media età del Bronzo: nuovi dati dal Lazio meridionale. Boll Unione Stor e Arte 10:109–126

Alessandri L, Baiocchi V, Del Pizzo S, Rolfo MF, Troisi S (2019a) Photogrammetric survey with fisheye lens for the characterization of the La Sassa Cave. Int Arch Photogramm Remote Sens Spat Inf Sci XLII-2(W9):25–32. https://doi.org/10.5194/isprs-archives-XLII-2-W9-25-2019

Alessandri L, Baiocchi V, Del Pizzo S, Di Ciaccio F, Onori M, Rolfo MF, Troisi S (2019b) Three-dimensional survey of Guattari cave with traditional and mobile phone cameras. Int Arch Photogramm Remote Sens Spat Inf Sci XLII-2(W11):37–41. https://doi.org/10.5194/isprs-archives-XLII-2-W11-37-2019

Angelini MG, Baiocchi V, Costantino D, Garzia F (2017) Scan to BIM for 3D reconstruction of the papal basilica of Saint Francis in Assisi in Italy. Int Arch Photogramm Remote Sens Spat Inf Sci XLII-5(W1):47–54. https://doi.org/10.5194/isprs-archives-XLII-5-W1-47-2017

Arnaud J, Peretto C, Grimaud-Hervé D (2015) The Grotta Guattari mandibular remains in the Italian human evolutionary context: a morphological and morphometrical overlook of the Neanderthal jaw. Quat Int 388:206–217. https://doi.org/10.1016/j.quaint.2015.08.030

Ascenzi A (1991) A short account of the discovery of the Mount Circeo Neandertal cranium. Quaternaria Nova 1:69–80

Baiocchi V, Barbarella M, D'Alessio MT, Lelo K, Troisi S (2016) The sundial of Augustus and its survey: unresolved issues and possible solutions. Acta Geod Geophys 51:527–540. https://doi.org/10.1007/s40328-015-0142-4

Baiocchi V, Barbarella M, Del Pizzo S, Giannone F, Troisi S, Piccaro C, Marcantonio D (2017) Augusto's sundial: image-based modeling for reverse engeneering purposes. Int Arch Photogramm Remote Sens Spat Inf Sci XLII-2(W3):63–69. https://doi.org/10.5194/isprs-archives-XLII-2-W3-63-2017

Barzaghi R, Betti B, Carrion D, Gentile G, Maseroli R, Sacerdote F (2014) Orthometric correction and normal heights for Italian levelling network: a case study. Appl Geomatics 6(1):17–25. https://doi.org/10.1007/s12518-013-0121-9

Bay H, Tuytelaars T, Van Gool L (2006) Surf: speeded up robust features. In: Leonardis A, Bischof H, Pinz A (eds) Computer vision – ECCV 2006. ECCV 2006. Lecture notes in computer science, vol 3951: 404-417 Springer, Berlin, Heidelberg. https://doi.org/10.1007/11744023_32

Blanc AC (1939) L'uomo fossile del Monte Circeo. Un cranio neandertaliano nella Grotta Guattari a San Felice Circeo. Rivista Antropologica 32:1–18

Blanc AC (1961) Some evidence for the ideologies of early man. In: Washburn SL (ed) Social life of early man. Routledge, London, pp 119–136

Bösemann W (2005) Advances in photogrammetric measurement solutions. Comput Ind 56(8–9):886–893

Dabove, P,., Grasso, N,., & Piras, M,. (2019). Smartphone-based photogrammetry for the 3D modeling of a geomorphological structure. Appl Sci , 9(18). https://doi.org/10.3390/app9183884), 3884

Doneus M, Neubauer W. (2005a) 3D laser scanners on archaeological excavations. In: Proceedings of the XX International Symposium CIPA, Torino, pp. 226–231

Doneus M, Neubauer W (2005b) Laser scanners for 3D documentation of stratigraphic excavations. In: Baltsavias E, Gruen A, Van Gool L, Pateraki M (eds) Recording, modeling and visualization of cultural heritage. Taylor & Francis Group, London, pp 193–203

Eyre M, Wetherelt A, Coggan J (2016) Evaluation of automated underground mapping solutions for mining and civil engineering applications. J Appl Remote Sens 10(4):46011. https://doi.org/10.1117/1.JRS.10.046011

Gallay M, Kaňuk J, Hochmuth Z, Meneely JD, Hofierka J, Sedlák V (2015) Large-scale and high-resolution 3-D cave mapping by terrestrial laser scanning: a case study of the Domica Cave, Slovakia. Int J Speleol 44(3):277–291. https://doi.org/10.5038/1827-806X.44.3.6

Gunn J (ed) (2003) Encyclopedia of caves and karst science. Fitzroy Dearborn, London

Jordan JH (2017) Modeling Ozark Caves with structure-from-motion photogrammetry: an assessment of stand-alone photogrammetry for 3-dimensional cave survey. Thesis, University of Arkansas. http://scholarworks.uark.edu/etd/2406

Lowe DG (1999) Object recognition from local scale-invariant features. In: Computer vision. The proceedings of the seventh IEEE international conference on computer vision, Vol. 2, pp 1150-1157. https://doi.org/10.1109/ICCV.1999.790410

Lucas G (2001) Destruction and the rhetoric of excavation. Nor Archaeol Rev 34:35–46. https://doi.org/10.1080/00293650119347

Mandelli A, Fassi F, Perfetti L, Polari C (2017) Testing different survey techniques to model architectonic narrow spaces. Int Arch Photogramm Remote Sens Spat Inf Sci XLII-2(W5):505–511. https://doi.org/10.5194/isprs-archives-XLII-2-W5-505-2017

Marra F, Bahain J-J, Jicha BR, Nomade S, Palladino DM, Pereira A, Tolomei C, Voinchet P, Anzidei M, Aureli D, Ceruleo P, Falgueres C, Florindo F, Gatta M, Ghaleb B, La Rosa M, Peretto C, Petronio C, Rocca R, Rolfo MF, Salari L, Smedile A, Tombret O (2019) Reconstruction of the MIS 5.5, 5.3 and 5.1 coastal terraces in Latium (central Italy): a re-evaluation of the sea-level history in the Mediterranean Sea during the last interglacial. Quat Int 525:54–77. https://doi.org/10.1016/j.quaint.2019.09.001

Micheletti N, Chandler JH, Lane SN (2015) Structure from motion (SFM) photogrammetry. In: Clarke LE, Nield JM (eds) Geomorphological techniques, chap. 2, sec. 2.2. British Society for Geomorphology, London

Monna F, Esin Y, Magail J, Granjon L, Navarro N, Wilczek J, Chateau C (2018) Documenting carved stones by 3D modelling–example of Mongolian deer stones. J Cult Herit 34:116–128. https://doi.org/10.1016/j.culher.2018.04.021

Mouragnon E, Lhuillier M, Dhome M, Dekeyser F, Sayd P (2009) Generic and real-time structure from motion using local bundle adjustment. Image Vis Comput 27(8):1178–1193. https://doi.org/10.1016/j.imavis.2008.11.006

Perfetti L, Polari C, Fassi F (2017) Fisheye photogrammetry: tests and methodologies for the survey of narrow spaces. Int Arch Photogramm Remote Sens Spat Inf Sci XLII-2(W3):573–580. https://doi.org/10.5194/isprs-archives-XLII-2-W3-573-2017

Perfetti L, Polari C, Fassi F, et al. (2018) Fisheye Photogrammetry to Survey Narrow Spaces in Architecture and a Hypogea Environment. In Latest Developments in Reality-Based 3D Surveying and Modelling; Remondino, F., Georgopoulos, A., González-Aguilera, D., Agrafiotis, P., Eds.; MDPI: Basel, Switzerland pp 3–28

Radicioni F, Rossi G, Tosi G, Marsili R (2019) Non contact shape and dimension measurements by LIDAR techniques of one of the biggest Italian caverns. J Phys Conf Ser 1249:012019. https://doi.org/10.1088/1742-6596/1249/1/012019

Rivero O, Ruiz-López JF, Intxaurbe I, Salazar S, Garate D (2019) On the limits of 3D capture: a new method to approach the photogrammetric recording of palaeolithic thin incised engravings in Atxurra Cave (northern Spain). Digit Appl Archaeol Cult Herit 14:e00106. https://doi.org/10.1016/j.daach.2019.e00106

Ruiz López JF, Hoyer CT, Rebentisch A, Roesch AM, Herkert K, Huber N, Floss H (2019) Tool mark analyses for the identification of palaeolithic art and modern graffiti. The case of Grottes d'Agneux in Rully (Saône-et-Loire, France). Digit Appl Archaeol Cult Herit 14:e00107. https://doi.org/10.1016/j.daach.2019.e00107

Sammartano G, Spanò A (2018) Point clouds by SLAM-based mobile mapping systems: accuracy and geometric content validation in

multisensory survey and stand-alone acquisition. Appl Geomatics 10:317–339. https://doi.org/10.1007/s12518-018-0221-7

Schwarcz HP, Bietti A, Buhay WM, Stiner MC, Grun R, Segre A (1991a) On the reexamination of Grotta Guattari: uranium-series and electron-spin-resonance dates. Curr Anthropol 32:313–316

Schwarcz HP, Buhay W, Grun R, Stiner MC, Kuhn S, Miller GH (1991b) Absolute dating of sites in coastal Lazio. Quaternaria Nova 1:51–67

Sergi S (1954) La mandibola neandertaliana Circeo II. Rivista di Antropologia 41:305–344

Sergi S (1974) Il cranio Neandertaliano del Monte Circeo. Academia Nazionale del Lincei, Roma

Sergi S, Ascenzi A (1955) La mandibola neandertaliana Circeo III (Mandibola B). Rivista di Antropologia 42:337–403

Stocchi P, Antonioli F, Montagna P, Pepe F, Lo Presti V, Caruso A, Corradino M, Dardanelli G, Renda P, Frank N, Douville E, Thil F, de Boer B, Ruggieri R, Sciortino R, Pierre C (2017) A stalactite record of four relative sea-level highstands during the Middle Pleistocene Transition. Quat Sci Rev 173:92–100. https://doi.org/10.1016/j.quascirev.2017.08.008

Triantafyllou A, Watlet A, Le Mouélic S, Camelbeeck T, Civet F, Kaufmann O, Quinif Y, Vandycke S (2019) 3-D digital outcrop model for analysis of brittle deformation and lithological mapping (Lorette cave, Belgium). J Struct Geol 120:55–66. https://doi.org/10.1016/j.jsg.2019.01.001

Troisi S, Baiocchi V, Del Pizzo S, Giannone F (2017) A prompt methodology to georeference complex hypogea environments. Int Arch Photogramm Remote Sens Spat Inf Sci XLII-2(W3): 639–644. https://doi.org/10.5194/isprs-archives-XLII-2-W3-639-2017

Velasco-Gómez J, Prieto JF, Molina I, Herrero T, Fábrega J, PérezMartín E (2016) Use of the gyrotheodolite in underground networks of long high-speed railway tunnels. Surv Rev 48(350):329–337. https://doi.org/10.1179/1752270615Y.0000000043

White TD, Toth N, Chase PG, Clark GA, Conrad NJ, Cook J, d'Errico F, Donahue RE, Gargett RH, Giacobini G, Pike-Tay A, Turner A (1991) The question of ritual cannibalism at Grotta Guattari [and comments and replies]. Curr Anthropol 32(2):118–138