Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Full size cave 3D modelling using close range photogrammetry and comparison with laser scanning

Gael Cazes
gael.cazes@umontpellier.fr

University of Montpellier
Philippe Vernant
University of Montpellier
François Baleux
University of Toulouse-Jean Jaurès
Carla Giuliani
Aix-Marseille University
Jean-Philip Brugal
Aix-Marseille University
Johan Jouves

Research Article

Keywords: Photogrammetry, structure from motion, cave survey, 3D modelling, karstology

Posted Date: October 14th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-4849312/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Applied Geomatics on July 12th, 2025. See the published version at https://doi.org/10.1007/s12518-025-00643-5.

Abstract

Although underground topography methods have significantly evolved in the past two decades, it is still challenging and often costly to create comprehensive 3D models of caves or artificial cavities. Modern methods used by speleologists, involving the generation of a topographic 'skeleton' through laser pointer measurements, enable surveying extensive developments but suffer from a lack of resolution. Conversely, the use of Lidar technologies, while capable of obtaining sub-millimeter scans, implies the use of expensive equipment and precision electronics often unsuitable for this harsh underground environment. In this study, we propose to test the use of low-cost cameras (action-cam type) for the complete 3D modeling of a cave through photogrammetry aiming to compare the results with a Lidar survey based on a series of topographic markers precisely measured with a total station. Our findings indicate that photogrammetry is an approach significantly faster and more adaptable in the field, leading to a substantial reduction in artifacts and shadows compared to static Lidar usage. Although post-processing involving image correlation is more computationally intensive for photogrammetry, we explore various strategies to reduce the calculation times. Ultimately, we demonstrate that the residual positioning errors on the topographic markers are of similar centimetric magnitude to those of Lidar. Recent advancements in computing capabilities now make it feasible to consider the use of photogrammetry for extended underground developments, presenting a promising alternative to the more conventionally employed Lidar in such contexts.

1. Introduction and Research Questions

Thanks to the rapid evolution of laser scanning and photogrammetry technologies during the last decade, 3D modelling of natural features has become a standard approach especially in geomorphology (Telling et al., 2017; Westoby et al., 2012) and geoarchaeology (Magnani et al., 2020; Remondino & Campana, 2014). Despite this trend, conducting a detailed and precise survey of all type of cavities, natural (karst) or anthropic (*e.g.*, mine) (named here as general term: cave), remains challenging and often requires expensive hardware and time-consuming field-work.

In such constrained settings, caving explorations coupled with accurate surveying of the penetrable voids such as horizontal or vertical passages as well as chambers, provide the opportunity to characterize this specific environment which often show complex geometry. This knowledge has significant implications for applied and societal problems (water and mineral resources characterisation, sinkhole collapse hazards, design and construction in civil engineering…) along with various fundamental research topics related to geosciences (speleogenesis, climate research, biospeleology, hydrogeology…), archaeology and palaeontology. Cave surveying tools have undergone significant improvements with the widespread use of handheld laser rangefinders that have completely replaced more traditional distance measurement methods. Despite this technical evolution, the same fundamental approach has been used since A. Martel's precursor book "Les Abimes" (Martel, 1894), where the main objective is to produce a linear frame of the cave by linking precisely measured topographic stations around which the gallery walls are then sketched, both in plan and profile views. This method is now routinely performed with the use of electronic cave surveying systems (Heeb, 2008; Trimmis, 2018) allowing instant mapping through measurement for each laser shot between the base and the target of the distance, azimuth and tilt. This paperless approach is until now the most efficient way to get an accurate cave-survey, however it does not provide a final high-resolution 3D model. This is why 'Terrestrial Laser Scanning' (TLS) became recently a new approach for detailed 3D modelling in the broad field of karstology (Jaillet et al., 2019; Oludare & Pradhan, 2016). Despite the important evolution of laser technologies since the 2000' and their astonishing measurement density and accuracy, Lidar scanners are still relatively bulky, fragile and costly which make them unsuited in various underground conditions such as vertical shafts, sinks, sumps, lakes and narrow or muddy passages, as commonly found underground. At the opposite, numerous mentions of photogrammetric approaches (often referred as SfM, for 'Structure from Motion') in cave environments can be found in the literature. Nevertheless, this method is still mostly used to capture discrete objects or limited portions of galleries or chambers and rarely, to our knowledge, for high-resolution topography of developing cave passages. SfM is often seen to complement 3D laser scanning adding textural information and photorealism, which is particularly valuable in archaeological contexts (Beraldin et al., 2006; Burens-Carozza et al., 2013; Grussenmeyer et al., 2020; Lerma et al., 2010; Núñez et al., 2013). When the aim is to study karstic morphologies, photogrammetry has also been used in conjunction with TLS for small-scale, highly heterogeneous objects (Pukanská et al., 2020) or for sections where the nature of the underground progression prevents the transportation of TLS equipment (De Waele et al., 2018; Pukanská et al., 2023); here again the studies still primarily rely on LiDAR surveys. Other works have also used SfM techniques to obtain a 3D mesh of the cave's exterior environment, providing a geometric and geological context for the 3D scan, often using drones equipped with on-board cameras (Kruger et al., 2016). As explained in Fryer et al., (2005) and in Silvestre et al., (2015), the dominance of underground 3D studies based on laser scanning can be attributed to the challenges associated with photographing in dark and occasionally dusty/blurry contexts, as well as the computational complexity of image correlation which was the main technological barrier until recently. As a result, studies where SfM has been the main method to digitize natural caves are sparse and mostly applied to short-horizontal cave passages. Recent works have yet proven that visual methods could represent good alternatives to TLS for caves characterized by narrow passages (Alessandri et al., 2019), semi-submerged cave systems (Furlani et al., 2023; Lee, 2018; Nocerino et al., 2019) or vast underground chambers (Triantafyllou et al., 2019). However, the accuracy of the resulting 3D models was not assessed in detail in these studies and it remains unclear if SfM based 3D meshes can be accurate and free of significant drift for long cave passages. Some comparisons between SfM and control points (W/GCP for 'Wall/Ground Control Point') measured by the paperless/rangefinder method were however proposed (Jordan, 2017) where a ~ 80m passage of a natural cave was modelled using a GoPro type action camera and a floodlight. The results show a metrical-level error between the two methods. A more comprehensive methodological comparison was also recently proposed by Giordan et al. (2021) in both a 100m long artificial tunnel and a 105m long gallery in the Gazzano area (Italy). They concluded that SfM based data, although characterized by lower resolutions, were comparable with point clouds obtained with TLS at a centimeter to decimeter-level co-registering errors, with the exception of some portions of the natural cave where the ceiling was too high to be properly lit.

Given the limited number of works exclusively focused on photogrammetry-based reconstruction in constrained environments, it appears relevant to investigate the benefits of this approach, particularly in terms of error reduction (drift), by comparison to more traditional surveying techniques. With this objective in mind, our goal is to assess the performance of a low-cost sensor-based photogrammetry survey on the scale of a whole cave and to compare the resulting data with those obtained from static laser scanning. To achieve this, we propose a comprehensive 3D survey of a cave studied for its

archaeological and paleontological records which already provide precise topographic markers recorded during previous excavation campaigns and where comprehensive 3D data was needed.

2. Study site: the “Mas des Caves” cave

The ‘Mas des Caves’ is an unusual cave-complex developed in Miocene limestone (*Molasse*), located in Lunel-Viel (Hérault, south of France). The two main caves (LV I and LV IV) were in the past a single tunnel-shaped gallery but are now separated by a sinkhole. The sinkhole was at time the natural entrance to the cave and is now completely sealed which makes only LV I accessible. Early fieldwork was carried out in LVI, initially by Marcel de Serres (1828, 1839) and subsequently by E. Bonifay (EB) from 1962 to 1982 (Bonifay, 1968, 1976); new research projects have been ongoing since 2019 directed by J.Ph. Brugal (JPB) (Brugal et al., 2021, 2022). Originating from the sinkhole past dynamic, the infilling is ca.3 m thick and dated recently of late Middle Pleistocene, isotopic stage MIS7 (Falguères et al., 2024). It yields abundant and diversified paleontological remains (ca. 75 taxa) associated with lithic artefacts. The role of hominins in the cave stays to be better define (Le Grand, 1994) and part of the bone assemblage (especially herbivores) have been associated to hyena activities (Fosse, 1996). One of the primary scientific objectives in this cave is to enhance the understanding of the formation processes of this archaeo-paleontological deposit and to shed light on the role of hominins in this accumulation. To achieve this, spatial analyses are essential, combining bone and lithic data-sets within a stratigraphy and taphonomic informations, in addition to geological and topographical contexts.

Replacing the large spatialized dataset within a precise 3D model of the cave would greatly help to highlight dynamics of infill deposition and post-depositional events of the remains, allowing a better understanding of the general scattering of bones with all behavioral implications. This interdisciplinary approach can provide a better understanding of depositional dynamics within a revised stratigraphic framework, which will allow to better underline the human occupation in the cave. Correspondingly, the methodological cross comparison presented in this work is also aimed at delivering a final 3D model that would best fit these needs of spatialized statistical analyses of the remains within the cave topography.

3. Methodology of 3D cave mapping

3.1. Photogrammetry

3.1.1 Image acquisition

The method described here revolves around a high-density photographic capture aimed at generating a 3D digital model representing the cave's walls geometry using close range photogrammetry. With the objective of testing a lightweight and cost-effective 3D scanning method, particularly suited for challenging underground conditions, we opted to employ rugged consumer-grade action cameras (2x GoPro 8 Black, 12MP) for the image acquisition (Fig. 1). Due to its small 1/2.3-inch sensor, this device is less ideal for low-light situations and tends to produce lower-quality shots in cave settings. The sharpness of images tends to rapidly declines as the distance to the photographed wall or object increases, primarily due to the combined impact of insufficient captured light and camera movements. During the acquisition process, the scene was accordingly illuminated by a powerful portable LED projector (6000 lumens) on top of which were mounted the two cameras oriented along the axis of the light beam with overlapping fields of view. This setup enables the light source and cameras to be as close as possible and fixed on the same pivot to optimize overlap and minimize cast shadows. The assembly is handheld by the operator, allowing for unrestricted movement within the cave and the ability to point in different directions along the path. The various sequences of shots (at 1Hz) are carried out by moving slowly on the cave floor, with the primary goal of covering the entire visible wall surface while maximizing overlap between the shots (at least 60–80% in theory). In contrast to automated photographic acquisition methods, where final resolutions relative to overlap rates and object distances can be constrained (*e.g.*, drone flights and outdoor DEM generation), SfM scanning within a cave relies more on the operator's subjective judgment. Variables such as the pace of movement, the mode of camera pivot adjustment (preferring translation over rotation to limit motion blur), the ability to visualize the path to avoid obscured areas, the distance to the objects being photographed, and the modulation of the light source's intensity will all exert influence over the final outcomes, particularly regarding geometric accuracy and resolution. Therefore, despite the overarching goal of obtaining a high-quality mesh representation of the entire cave, greater attention and closer image captures were applied when navigating through the excavation area (closer to the natural entrance), which represents a smaller portion of the overall study site. Additionally, in order to subsequently georeference the data and compare it with the TLS point cloud, closer shots were also conducted around the topographic markers measured with the total station. In total, more than 18,000 images were taken corresponding to half a day of field work with an effective acquisition time of approximately 2 hours and 30 minutes (sequences of 1 second each with two cameras). This dataset amounts to a file size of about 40 GB (approximately 2.3 MB per image).

3.1.2 Image correlation and data processing

The entire photogrammetric calculation (mainly performed with Agisoft Metashape 1.8 Pro software) and post-processing workflow is schematically illustrated in Fig. 2. Due to the computational intensity of image correlation algorithms in terms of processing time and memory usage, it is generally not efficient/feasible to attempt aligning all images from the same project in a single pass. In order to prevent image correlation from being overloaded with unusable shots, a strategy for quantifying quality parameters was explored. This type of quality assesment is not straightforward in optic as it remains a relatively subjective concept (Sieberth et al., 2016) and can depend on various factors such as the presence of shadows, graininess, over/underexposure, optical blur, and motion blur. We used a Metashape’s build-in algorithm which involves quantifying the detected edge contrast gradients between different areas of the image in terms of tones or colours (edge detection spread (Narvekar & Karam, 2009; Ong et al., 2003). This sharpness proxy is calculated as the average of edge detection gradients within the same image. By ranking the dataset relative to this parameter, it allows exclusion of the least usable photos, which involved a reduction of around 20% of the number of images in the project. The weakness of this relatively simplistic approach lies in the fact that it does not take into account edge orientation and is therefore not well-suited for images subjected to significant motion blur, which despite their poor quality,

can generate sharp edge detections. Other methods involving frequency analyses (Rahtu et al., 2012), approaches based on the human perception of blur (Crete et al., 2007) or the use of more modern techniques involving AI and deep learning remain too intensive on such a large dataset.

Due to the exponential nature of computation times as a function of the number of images, another approach involving sorting the images into batches corresponding to distinct areas within the cave was tested before the alignment step. Each batch of images is sequentially aligned, with a certain number of common images between the chunk pairs to ensure alignment of all batches along the cave (Fig. 2). Such method reduces drastically the alignment step duration. Once performed, the image alignment was then optimized by reducing the amount of tie points (common points within the overlap area of adjacent images) used for the reconstruction based on 1) the weakest values of statistical error and variation in tie point placement and 2) the accuracy of point placement from local neighbours. At this step, the resulting optimized cloud and relative camera positions are only known in a relative space and needs to be aligned to a known referential for further analyses and cross comparison. In order to do so, thirteen reference topographic nails corresponding to a previously conducted work within the excavation area, still visible in the capture scene, were employed. These reference points had been measured using a Total Station and used daily by excavators during field-works and were thus considered reliable. The markers are still clearly visible and were predominantly placed on concrete support pillars (6 of the 13 points) as well as on the wall (7 of the 13 points). They are marked with black nails, and most of them bear a corresponding station number (Fig. 3). Therefore, it was necessary to select these points in the photographs and provide their XYZ coordinates in detail so that the software could subsequently calculate a matrix enabling 1) translation, 2) rotation and 3) scaling to best align the model with the Total Station reference frame.

The following steps aim at generating a high-resolution mesh and here again the chunk strategy was employed to reduce processing time. Meshes were calculated in Metashape by using a depth map method consisting of 1) generating image channels containing depth information for the overlapping image pairs considering their orientation parameters and 2) filtering and triangulation based on the depth information to generate meshes. The final triangulated envelopes were finally cleaned of noise, topological errors and potential artefacts (using Meshlab and Blender open-source programs) before being reimported in Metashape for the calculation of textures and normal maps in order to obtain final photorealistic renders. All sub-meshes were merged together and gave a final 35M triangle mesh representing the entire cave geometry (as seen here in lower resolution: https://sketchfab.com/3d-models/grotte-du-mas-des-caves-acces-prive-63438f3425f9415fb1508d1b1b6a0698).

3.2. Terrestrial laser scanning

The laser acquisition was carried out using a Faro Focus3D 120s scanner. This device emits an invisible laser beam with a wavelength of 905 nm (infrared) that enables the acquisition of highly accurate 3D point clouds (measurement uncertainty of +- 2mm between 10 and 25m given by the constructor). The principle of the device is to record the surface that the laser hits as a point cloud referenced in X, Y, and Z coordinates, with the scanner as the reference origin. Each scene contains a large number of points, and these point clouds are assembled into a single dataset in post-processing. The acquisition parameters (250kpoints/second = ~ 45 M points per scan) were set to facilitate fast acquisition with an average time of ~ 5min per scan scene. An integrated camera allowing capturing 360-degree photographs of the scanned areas is present on this model but was not used in this work due to the difficulty of homogeneously illuminating the cave scene and the insufficient sensitivity and optical quality of this sensor in low light conditions. Each acquired point was however assigned a hue providing additional information to the "intensity" values measured by the laser, which are stored as grayscale.

The future alignment of the point clouds is prepared in the field. Artificial markers (A4-sized checkerboard targets and highly reflective 145mm spheres) are positioned in a way that at least five markers are visible in each scan, and at least three spheres are common between two scans. Regardless of the scanner positions, the software can detect the sphere as long as it is sufficiently close and not occluded. For this project, we utilized 15 spheres and 6 checkerboard targets distributed throughout the entire site. The number of scans performed is influenced by several factors. At first, the range at which the spheres can be recognized, which must be below 20m with our acquisition parameters. Secondly, the need to minimize blind spots in order to provide the most accurate description of the area and terrain topography. The only way to reduce the potentially shadowed areas is to increase the number of scan angles and, consequently, the number of stations. To achieve this, the number of stations in the excavation area has been multiplied while in the southern part the scans were more spaced out and had fewer connections (Fig. 4). In total, fifty-one scans were performed which, accounting for ~ 15min per scan position, represents more than 12 hours of effective work (two days of fieldwork). The assembly of the fifty-one point clouds was done using the Faro SCENE software. We chose to perform target-based assembly, which combines speed and accuracy by utilizing the spheres and checkerboards (~ 5min per scan for a total of ~ 5hours for the 51 scans). Figure 4 shows the location of the various scans on the site. The average value of the target positioning error given by the program is less than 3 mm (2.7 mm min 0.3; max 10.6) and corresponds to the positional shift of the same sphere in different scans. The spacing and low angular distribution decrease the assembly quality. In fact, to effectively constrain two point clouds, it is necessary to rely on targets distributed along a star pattern around the scanner station which is not always possible due to the morphology and number of available targets.

Once aligned, a merged cloud of ~ 2B points was finally exported to CloudCompare software for noise filtering and resampling to a final 800M high resolution points cloud. (1.5mm min distance between points). As for the photogrammetry data, the TLS data was also aligned to the total station referential by selecting the topographic markers “seen” in the clouds and by adjusting accordingly a matrix relative to the Total Station coordinates. Among the 13 reference points labelled as 'station', 12 were successfully identified in the scans. Point number 10, however, lacked sufficient definition for inclusion in the scans. Among the 12 identified points in the scans, point number 3 was only recorded in a single scan, while the others were detected in at least 2 scans. Notably, reference point number 4 was identified and marked in 8 different scans.

4. Results and model cross-comparisons

4.1. 3D Model resolutions

After post-processing, each of the two methods allowed for the generation of a 3D model representing about 4500 m^2 of scanned wall along ~ 215m of developing galleries. The final full resolution mesh obtained by photogrammetry counts ~ 35M triangles and ~ 16M vertices with an average density of ~ 4k points/m^2 while the final TLS point cloud exported after noise filtering counts ~ 50 times more points. As such it is difficult to compare the data quality with these numbers alone as the final resolution for both methods remain a subjective choice which depends on the objects of interest. When not occluded, the TLS data always give higher point concentration but we observe that both datasets allow visualizing pluri-centimetric objects (ex. small pebbles along the excavation fronts, as in Fig. 5), at least in the well resolved areas. The representative resolution is indeed by nature variable along the cave as it depends mostly on 1) the sensor to surface distance and 2) the presence of occlusions. This representative resolution for the SfM model, which lies around the centimetric barrier in average, tends to decrease and becomes pluri-centimetric in the less constrained portion of the cave where pictures were taken from a more distant position or in the case of occlusion, particularly visible in the upper parts towards the gallery ceiling (Fig. 6).

4.2. Geometric accuracy

All available reference points of photogrammetry data were registered and we observe (Table 1) that the residual difference once aligned on the wall/ground control points (W/GCPs in the text) falls within a satisfactory range (min: 6 mm, max: 24 mm, median: 16 mm; N = 12). The positioning errors observed on the TLS data are better than those observed for photogrammetry (Table 1) while still falling within similar orders of magnitude (min: 5 mm, max: 23 mm, median: 13 mm; N = 12). For both datasets, the majority of the error is expressed on the X component, which on average constitutes 90% (SFM) and 85% (LiDar) of the norms of the residual vectors (columns Abs. ΔXYZ in Table 1). The significance of this heterogeneity in residuals across axes would be explained by the distribution of reference markers, which is much more spread along the X-axis (35.5 m) than along the Y and Z axes (8.6 m and 2.9 m, respectively), implying significantly different degrees of freedom for the three axes in the adjustment calculation. We also observe a similar distribution of errors along the X-axis between the two acquisition methods. Indeed, points P1 to P8 exhibit positive residuals (except for P5 in photogrammetry) while points P9 to P13 (excluding P10) are characterized by negative residuals. This similarity in the variations of residuals between the two independent methods indicates a probable bias in the measurement of the referencing nails (at least for P9 to P13, possibly due to a weak tacheometer base change?). Accordingly, by aligning the SFM mesh on the LiDar data this time (using the LiDar/GCP coordinates and excluding P6 in the SFM data and P10 in the TLS data), the final positioning errors lower down to the sub-centimetric order (min: 3mm, max: 20mm, median: 8mm; N = 11) and the anomaly on the errors on the X axis is not present anymore.

Table 1
Residual differences after alignment of the SfM and LiDar data on the W/GCPs referential (SfM/GCP and LiDar/GCP columns) and after alignment of the SfM LiDar data (SfM/LiDar columns). P6 and P10 could not be picked successfully, respectively in the SfM and Lidar data, and were discarded from the alignmer

W/GCP n°	SfM/GCP ΔX (m)	SfM/GCP ΔY (m)	SfM/GCP ΔZ (m)	Sfm/GCP Abs. ΔXYZ (m)	LiDar/GCP ΔX (m)	LiDar/GCP ΔY (m)	LiDar/GCP ΔZ (m)	LiDar/GCP Abs. ΔXYZ (m)	SfM/LiDar ΔX (m)	SfM/LiDar ΔY (m)	SfM/LiDa ΔZ (m)
P1	0.008	-0.007	-0.003	0.011	0.007	0.001	-0.001	0.007	0.005	-0.006	-0.002
P2	0.016	-0.006	-0.002	0.017	0.009	0.007	-0.001	0.011	0.008	-0.011	0
P3	0.013	-0.005	0.001	0.014	0.010	0.005	0.001	0.011	0.003	-0.007	-0.001
P4	0.003	0.005	-0.001	0.006	0.005	0.001	0	0.005	0.005	0.007	-0.001
P5	-0.024	0.001	0.001	0.024	0.004	-0.004	-0.001	0.006	-0.018	0.008	0.001
P6	-	-	-	-	0.004	-0.009	-0.002	0.01	-	-	-
P7	0.015	0.006	0.001	0.016	0.011	0.010	0.001	0.015	0.002	-0.002	0
P8	0.018	-0.001	0.003	0.019	0.012	0.005	0.003	0.013	0.005	-0.002	-0.001
P9	-0.012	-0.002	-0.004	0.013	-0.012	-0.006	-0.003	0.014	-0.009	0.008	0.001
P10	0.010	0.015	0.003	0.018	-	-	-	-	-	-	-
P11	-0.012	-0.004	-0.001	0.013	-0.018	0.004	0.001	0.018	0.003	-0.006	-0.001
P12	-0.015	-0.003	0.001	0.016	-0.023	-0.002	0.00	0.023	0.003	0.004	0.001
P13	-0.019	0.001	0.003	0.020	-0.018	-0.002	0.003	0.018	-0.008	0.007	0.002
	0.014	0.005	0.002	**0.016**	0.011	0.005	0.001	**0.013**	0.005	0.007	0.001

As the registration constraints are only present in the excavation area, it is challenging at this stage to quantify accuracy across the entire 3D model. For this purpose, a series of orthogonal sections to the primary axis of the cavity were extracted from both datasets, spaced at 5-meter intervals (Fig. 7). In the excavation zone, the two meshes closely align, with the majority of surfaces differing by less than 5mm from each other. The few areas where the two models significantly deviate represent zones with minor occlusions characterized by overhangs or recesses in the ceiling and on the walls. These localized discrepancies are better captured by photogrammetry due to its superior viewing angles compared to LiDAR. In areas outside the excavation area, where lower resolutions are expected due to the limited number of source data (number of photos and scan scenes), the higher laser resolution is occasionally offset by significant occlusions, especially in narrow areas at the base of the walls where it would be challenging to acquire additional LiDAR data, even with more scans (section 50 in Fig. 7).

Similarly, a mesh-to-mesh distance calculation performed on the two full size models (Fig. 8) shows a strong consistency between the two methods with a distance standard deviation of ~ 0.5 cm in the well constrained excavation area (one quarter of the total cave 3D area with the highest data density, A1 in Fig. 4) and ~ 2.4 cm on the rest of the development (A2 in Fig. 4).

5. Discussion

5.1 Key benefits and drawback of the two approaches

TLS is widely acknowledged as the state-of-the-art methodology for acquiring high-density and precise 3D data in constrained environments, such as caves, underground spaces and heritage or archaeological sites. With its remarkable measurement density, TLS is also often established as a good choice for quick investigating of infra-centimetric objects and structures on localized panels. For such small-scale settings (macro to microscopic recording), it has however been shown that passive methods like photogrammetry together with photometric stereo and reflectance transformation imaging, could also compete against active sensor-based techniques (Peña-Villasenín et al., 2019; Plisson & Zotkina, 2015). The efficiency of static Lidar campaigns diminishes rapidly when applied to the scanning of developing galleries in larger scale cave, primarily due to: 1) the frequent base and spheres repositioning, 2) the multiplication of scanning positions to mitigate occlusions and 3) the inability to scan in narrow sections. Furthermore, TLS tends to generate heavy datasets (approximately ~ 45 million points per scan multiplied by 51 scans equals 2.3 billion points for this project) and consistently necessitates substantial subsampling to enable post-processing on standard consumer-grade computers. In the context of extended underground networks spanning more than a few hundred meters, employing TLS mandates allocating a considerable amount of digital storage—ranging from hundreds of gigabytes to terabytes for high-density point clouds. While such measurement density is needed for robust cloud-to-cloud registrations, it then becomes superfluous in the final aligned point cloud considering the lower resolutions typically required for most applications. On the other hand, SfM only requires photographs as input and we show in this work that even images of relatively low quality (GoPro 2.7k, 2.3 MB per image) can produce satisfactory 3D data as long as the overlap and shooting distance are adapted to the resolution of interest. The data acquisition process is more versatile with this method and requires less storage capacity, but the bulk of the work to generate 3D data is carried out in a post-processing step that can be computationally and time demanding in comparison with TLS approach.

Recently, the emergence of mobile scanning devices based on SLAM algorithms has led to envision of new types of fast and mid to high-resolution spatial data acquisition in constrained contexts (Menna et al., 2016; Dewez et al., 2017; Gautier et al., 2020; Breu & lapierre, 2020; Di Stefano et al., 2021; Giordan et al., 2021). The ability to continuously record an entire cave by using such approach is appealing and it would be interesting to conduct a similar large-scale comparison to quantify its precision for long underground cavities. However, due to the inherent tendency of SLAM based devices to drift with the acquisition time and the need to perform loop closures which would often be difficult to realise in most caves, it is still difficult to conceive the acquisition of spatial data with the similar levels of precision as presented in this work with such methods.

Furthermore, one significant advantage of photogrammetry over static and mobile laser scanning lies in its capability to generate textured and photorealistic 3D models (Fig. 9). This feature enhances the visual representation of the captured environment, providing a more immersive and detailed depiction of the studied sites together with the ability of conveying innovative concepts to a broad public. These realistic models offer valuable insights in fields such as karstology, archaeology, cultural heritage preservation and virtual tourism, where an accurate and visually appealing representation is crucial for analysis, documentation, and public engagement.

To summarize, Table 2 provides a brief overview of the main characteristics, advantages and disadvantages of both methods when used in underground conditions.

Table 2
This table outlines the main characteristics and differences between photogrammetry and static LiDAR when used for surveying caves.

Criteria	Photogrammetry	Static LiDAR
Data Acquisition method	Uses photographs to create 3d models through image correlation and Sfm algorithms	Uses laser pulses to measure distances to the surface and create 3d models
Equipment cost	Low-cost action-cameras can be used (400–500$) with a powerful LED floodlight (400–500$) High-end computer (~ 3k$) and Sfm software (free to a few k$)	Typically involves expensive static laser scanners (15k$ to 60k$) and referential markers (500$) High-end computer (~ 3k$) and software (generally comes with the scanner)
Field work efficiency	Fast and versatile. A few hundred meters per day can be recorded in underground settings	Slower due to the need for multiple scanning positions to mitigate occlusions and the inability to scan narrow sections
Adapted to underground conditions	Yes – cheap and shock/water- proof equipment	Not always – expensive and fragile electronic that cannot be used in wet, muddy, constrained or vertical sections of caves
Scale and level control	Requires external control points or a special rig to scale and level the 3D data	Provides scaled and levelled data inherently without the need for external control points
Post-processing	Computationally intensive. The bulk of the work is carried out in post-processing. Even if most of the calculation can be automatised, human inputs are still needed at different steps	A few minutes per scans and then automatic
Data density and Accuracy	Lower measurement density (but adjustable during acquisition based on the capture distance) with comparable (few cms) geometric accuracy at the cave scale.	High density and millimetric accuracy at the scale of a single scan but final full size 3D model similar in accuracy with Sfm
End products	Textured and photorealistic 3D models, enhancing visual representation	High resolution 3D point cloud
Emerging technologies	Depth cameras with visual SLAM algorithms. Cheap and versatile but lower resolution and accuracy	Mobile SLAM based laser scanners. Expensive (a few 10k$) and fragile but fast and high-resolution data acquisition

5.2 Scientific applications in the Mas des Caves

The morphology and dimensions of a cave provide crucial insights into its scientific understanding, encompassing its formation process, sedimentary history, the infillings it contains, and the fossils found within. The cave' topography, and its opening(s), condition the geological and biological nature of the deposits, in close relation to the general external environment. A better understanding of karst subterranean geometry enables us to envision the spatio-temporal frame, the dynamics and setting up of sediments, as well as the biological and anthropic agents that have occupied the cave-site that constitutes the natural host structure. They also explain the taphonomical history (origin, conservation, modification, destruction) and (paleo)biodiversity of fossil, archaeological and/or paleontological, assemblages. For example, the detailed, spatial 3D topography of the cave itself allows us to better understand the presence of bone refits between layers.

For the current excavations in LVI, the use of a tacheometer enabled to precisely record and visualize the excavated remains. Likewise, all the data, from EB and JPB s excavations were converted in a same database and allows to spot them in a common GIS (Geographic Information System). As part of an ongoing doctoral project (C. Giuliani), the spatial distribution of bone remains and lithic artefacts in the Mas des Caves site will be examined thoroughly. This study is based on the horizontal and vertical projection of all these remains within the 3D model of the cave (Fig. 10), supported by physical reassembly work (*e.g.*, joint reconstruction, refitting). The aim is to better define the fossil packages, and then the different levels of occupation that took place through time. Such analysis is replaced in the sedimentary sequence and confronted with the topographical constraints of the cave. The combination of all these factors and variables enables a high-resolution reconstruction of the biological and anthropic events present in this infill over a period on the scale of the MIS 7 isotopic stage, *i.e.* a period estimated at 30–50 thousand years.

6. Conclusion

This work aimed at evaluating the viability of low-cost cameras based handheld photogrammetry in addition to more traditional and expensive static laser scanning for 3D mapping in challenging underground environments. In the “grotte du Mas des caves”, this approach allowed to survey the entire underground area with a total ~ 220m passages in only 2.5h of effective field work. While the static laser scanning provided a higher final data density, the geometric accuracy of both methods was comparable with centimetric scale average co-registering errors.

Cross-comparisons of the 3D models revealed that both photogrammetry and laser scanning accurately represented and detailed the cave's morphology, with minor discrepancies in regions with occlusions favouring photogrammetry due to its superior viewing angles and increased flexibility. The study highlights photogrammetry as a cost-effective and efficient method for cave mapping, offering valuable insights for scientific applications in contexts where TLS could not be used efficiently. This research contributes to the exploration of alternative underground surveying techniques, opening avenues for future investigations into visual methods for cave exploration and geomorphological studies.

Declarations

Acknowledgments

This work is part of a thesis project (PhD G. Cazes) focussed on the development of 3D modelling in caves based on visual methods and funded by a public/private partnership involving the company CENOTE, the Géosciences Montpellier laboratory, and the Occitanie region. The research project on Mas des caves/LV I ("People and Environments in the Middle Pleistocene", PI J.-Ph.B) is supported by the DRAC-SRA Occitanie, the region and Hérault Dpt. We would like to express our gratitude to the property owners for granting access to the cave.

Competing interests

The authors declare no competing interests related to the work submitted for publication and no financial relationships with organizations that might have an interest in the submitted work

References

1. Alessandri, L., Alessandri, L., Baiocchi, V., Pizzo, S. D., Ciaccio, F. D., Onori, M., Rolfo, M. F., Rolfo, M. F., & Troisi, S. (2019). Three-dimensional survey of Guattari cave with traditional and mobile phone cameras. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. https://doi.org/10.5194/isprs-archives-xlii-2-w11-37-2019.
2. Beraldin, J.-A., Blais, F., Cournoyer, L., Picard, M., Gamache, D., Valzano, V., Bandiera, A., & Gorgoglione, M. A. (2006). Multi-Resolution Digital 3D Imaging System Applied to the Recording of Grotto Sites: The Case of the Grotta dei Cervi. VAST, 45-52.
3. Bonifay, E. (1968). Stratigraphie et industries lithiques de la grotte n^{o} 1 du Mas des Caves à Lunel-Viel (Hérault). La Préhistoire : Problèmes et tendances, 37-46.
4. Bonifay, E. (1976). Grottes du Mas des Caves. Provence et Languedoc Méditerranéen Sites Paléolithiques et Néolithiques, Actes du IX Congrès de l'Union Internationale des Sciences Préhistoriques et Protohistoriques, 197-204.
5. Breux, Y., & Lapierre, L. (2020). Elevation Angle Estimations of Wide-Beam Acoustic Sonar Measurements for Autonomous Underwater Karst Exploration. Sensors, 20(14), Article 14. https://doi.org/10.3390/s20144028.
6. Brugal J.P., Giuliani C., Fosse P., Fourvel J.B., Magniez P., Pelletier M., Uzunidis A., (2021). Preliminary data on the middle Pleistocene site of Lunel-Viel I (Hérault, France). Alpine and Mediterranean Quaternary, 34 (1): 1-13, doi.org/10.26382/AMQ.2021.08.
7. Brugal J.P., Giuliani C., Fosse P., Fourvel J.B., Magniez P., Pelletier M., Uzunidis A., (2022). Le site archéo-paléontologique du Mas des Caves à Lunel-Viel (Hérault) : Historique des recherches et nouveaux travaux. Karstologia, Mém.25, vol. V / VI : 339-342 [Actes du 18e Congrès International de Spéléologie, Symposium 8.
8. Burens-Carozza, A., Grussenmeyer, P., Guillemin, S., Carozza, L., Leveque, F., & Mathé, V. (2013). Methodological developments in 3D scanning and modelling of archaeological French heritage site: The Bronze Age painted cave of" Les Fraux", Dordogne (France). International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 40, 131-135. https://doi.org/10.5194/isprsarchives-XL-5-W2-131-2013.
9. Crete, F., Dolmiere, T., Ladret, P., & Nicolas, M. (2007). The blur effect: Perception and estimation with a new no-reference perceptual blur metric. Human Vision and Electronic Imaging XII, 6492, 196-206. https://doi.org/10.1117/12.702790.
10. De Waele, J., Fabbri, S., Santagata, T., Chiarini, V., Columbu, A., & Pisani, L. (2018). Geomorphological and speleogenetical observations using terrestrial laser scanning and 3D photogrammetry in a gypsum cave (Emilia Romagna, N. Italy). Geomorphology, 319, 47-61. https://doi.org/10.1016/j.geomorph.2018.07.012.
11. De Serres, M. (1828). Recherches sur les ossements fossiles des cavernes de Lunel-Vieil (Hérault). Mém. Du Muséum d'Hist. Nat.., XVII, 380-463.
12. De Serres, M. (1839). Recherches sur les ossements humatiles des cavernes de Lunel-Viel. Montpellier : Boehm et Cie (éds), pp. 256.
13. Falguères C., Lahaye C., Tombret O., Garbé L., Lebrun B., Bahain J.J., Frerebeau N., Giuliani C., Brugal J.P., (2024). ESR/U-series and IRSL dating of the Middle Pleistocene site of Lunel Viel (LV I), Hérault, Southern France. Quaternary Geochronology, 81 : 101516, https://doi.org/10.1016/j.quageo.2024.101516.
14. Fosse, P. (1996). La grotte n° 1 de Lunel-Viel (Hérault, France) : Repaire d'hyènes du Pléistocène moyen. Étude taphonomique du matériel osseux. Paléo, Revue d'Archéologie Préhistorique, 8(1), 47-79.
15. Fryer, J. G., Chandler, J. H., & El-Hakim, S. F. (2005). Recording and modelling an aboriginal cave painting: With or without laser scanning. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36(5/W17), 1-8.
16. Furlani, S., Antonioli, F., Colica, E., D'Amico, S., Devoto, S., Grego, P., & Gambin, T. (2023). Sea Caves and Other Landforms of the Coastal Scenery on Gozo Island (Malta): Inventory and New Data on Their Formation. Geosciences, 13(6), Article 6. https://doi.org/10.3390/geosciences13060164.
17. Gautier, Q. K., Garrison, T. G., Rushton, F., Bouck, N., Lo, E., Tueller, P., Schurgers, C., & Kastner, R. (2020). Low-cost 3D scanning systems for cultural heritage documentation. Journal of Cultural Heritage Management and Sustainable Development, 10(4), 437-455.
18. Giordan, D., Godone, D., Baldo, M., Piras, M., Grasso, N., & Zerbetto, R. (2021). Survey Solutions for 3D Acquisition and Representation of Artificial and Natural Caves. Applied Sciences, 11(14), 6482. https://doi.org/10.3390/app11146482.
19. Grussenmeyer, P., Landes, T., Alby, E., & Carozza, L. (2020). High resolution 3D recording and modelling on the Bronze Age cave" les Fraux" in Perigord (France). Conference ISPRS Commission V Symposium, 38, 262-267.
20. Heeb, B. (2008). Paperless caving-An electronic cave surveying system. IV European Speleological Congress.

21. Jaillet, S., Delannoy, J.-J., Génuite, K., Hobléa, F., & Monney, J. (2019). L'image topographique du karst et des grottes : Représentations 2D et technologies 3D, entre réalité et imaginaire. Géomorphologie : relief, processus, environnement, 25(3), Article 3. https://doi.org/10.4000/geomorphologie.13488.
22. Jordan, J. (2017). Modeling Ozark Caves with structure-from-motion photogrammetry: An assessment of stand-alone photogrammetry for 3-dimensional cave survey. [PhD Thesis]. University of Arkansas.
23. Kruger, A., Randolph-Quinney, P., & Elliott, M. (2016). Multimodal spatial mapping and visualisation of Dinaledi Chamber and Rising Star Cave. South African journal of science, 112(5-6), 1-11. http://dx.doi.org/10.17159/sajs.2016/20160032.
24. Le Grand, Y. (1994). Processus de formation des dépôts archéologiques du Pléistocène moyen de Lunel-Viel 1 (Hérault) : L'utilisation des remontages. Préhistoire anthropologie méditerranéennes, 3, 57-63.
25. Lee, T. O. (2018). An Examination of Close-Range Photogrammetry and Traditional Cave Survey Methods for Terrestrial and Underwater Caves for 3-Dimensional Mapping [PhD Thesis]. University of Southern California.
26. Lerma, J. L., Navarro, S., Cabrelles, M., & Villaverde, V. (2010). Terrestrial laser scanning and close-range photogrammetry for 3D archaeological documentation: The Upper Palaeolithic Cave of Parpalló as a case study. Journal of Archaeological Science, 37(3), 499-507. https://doi.org/10.1016/j.jas.2009.10.011.
27. Magnani, M., Douglass, M., Schroder, W., Reeves, J., & Braun, D. R. (2020). The Digital Revolution to Come: Photogrammetry in Archaeological Practice. American Antiquity, 85(4), 737-760. https://doi.org/10.1017/aaq.2020.59.
28. Martel, E.-A. (1894). Les Abimes : Les eaux souterraines, les cavernes, les sources. La spéléologie. Book éd. Delagrave.
29. Menna, F., Nocerino, E., Fassi, F., & Remondino, F. (2016). Geometric and Optic Characterization of a Hemispherical Dome Port for Underwater Photogrammetry. Sensors. https://doi.org/10.3390/s16010048.
30. Nocerino, E., Menna, F., Farella, E., & Remondino, F. (2019). 3D Virtualization of an Underground Semi-Submerged Cave System. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, KLII(2/W15), 857-864. https://doi.org/10.5194/isprs-archives-XLII-2-W15-857-2019.
31. Núñez, M. A., Buill, F., & Edo, M. (2013). 3D model of the Can Sadurní cave. Journal of Archaeological Science, 40(12), 4420-4428.
32. Oludare, I. M., & Pradhan, B. (2016). A decade of modern cave surveying with terrestrial laser scanning: A review of sensors, method and application development. International Journal of Speleology, 45(1), 8. https://doi.org/10.5038/1827-806x.45.1.1923.
33. Peña-Villasenín, S., Gil-Docampo, M., & Ortiz-Sanz, J. (2019). Professional SfM and TLS vs a simple SfM photogrammetry for 3D modelling of rock art and radiance scaling shading in engraving detection. Journal of Cultural Heritage, 37, 238-246. https://doi.org/10.1016/j.culher.2018.10.009.
34. Plisson, H., & Zotkina, L. V. (2015). From 2D to 3D at macro- and microscopic scale in rock art studies. Digital Applications in Archaeology and Cultural Heritage, 2(2), 102-119. https://doi.org/10.1016/j.daach.2015.06.002.
35. Pukanská, K., Bartoš, K., Bella, P., Gašinec, J., Blistan, P., & Kovanič, Ľ. (2020). Surveying and High-Resolution Topography of the Ochtiná Aragonite Cave Based on TLS and Digital Photogrammetry. Applied Sciences, 10(13), Article 13. https://doi.org/10.3390/app10134633.
36. Pukanská, K., Bartoš, K., Gašinec, J., Pašteka, R., Zahorec, P., Papčo, J., Kseňak, Ľ., Bella, P., Andrássy, E., Dušeková, L., & Bobíková, D. (2023). Measurement of spatio-temporal changes of cave ice using geodetic and geophysical methods: Dobšiná Ice Cave, Slovakia. Frozen ground/Field Studies. https://doi.org/10.5194/tc-2023-110.
37. Rahtu, E., Heikkilä, J., Ojansivu, V., & Ahonen, T. (2012). Local phase quantization for blur-insensitive image analysis. Image and Vision Computing, 30(8), 501-512. https://doi.org/10.1016/j.imavis.2012.04.001.
38. Remondino, F., & Campana, S. (2014). 3D recording and modelling in archaeology and cultural heritage. BAR international series, 2598, 111-127.
39. Silvestre, I., Rodrigues, J. I., Figueiredo, M., & Veiga-Pires, C. (2015). High-resolution digital 3D models of Algar do Penico Chamber: Limitations, challenges, and potential. International Journal of Speleology, 44(1), 25-35. https://doi.org/10.5038/1827-806x.44.1.3.
40. Telling, J., Lyda, A., Hartzell, P., & Glennie, C. (2017). Review of Earth science research using terrestrial laser scanning. Earth-Science Reviews, 169, 35-68. https://doi.org/10.1016/j.earscirev.2017.04.007.
41. 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). Journal of Structural Geology, 120, 55-66. https://doi.org/10.1016/j.jsg.2019.01.001.
42. Trimmis, K. P. (2018). Paperless mapping and cave archaeology: A review on the application of DistoX survey method in archaeological cave sites. Journal of Archaeological Science: Reports, 18, 399-407. https://doi.org/10.1016/j.jasrep.2018.01.022.
43. Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., & Reynolds, J. M. (2012). 'Structure-from-Motion' photogrammetry: A low-cost, effective tool for geoscience applications. Geomorphology, 179, 300-314. https://doi.org/10.1016/j.geomorph.2012.08.021.

Programs

AgiSoft Metashape Professional (Version 1.8.2) (Software). (2022). Retrieved from http://www.agisoft.com

Community, B. O. (2022). *Blender - a 3D modelling and rendering package*. Stichting Blender Foundation, Amsterdam. Retrieved from http://www.blender.org

Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F., & Ranzuglia, G. (2008). Meshlab: An open-source mesh processing tool. *Eurographics Italian chapter conference*, *2008*, 129-136. Retrieved from https://www.meshlab.net/

Figures



Figure 1

The image acquisition is performed using two action cameras mounted on a powerful LED flood light (Photograph R. Muller)

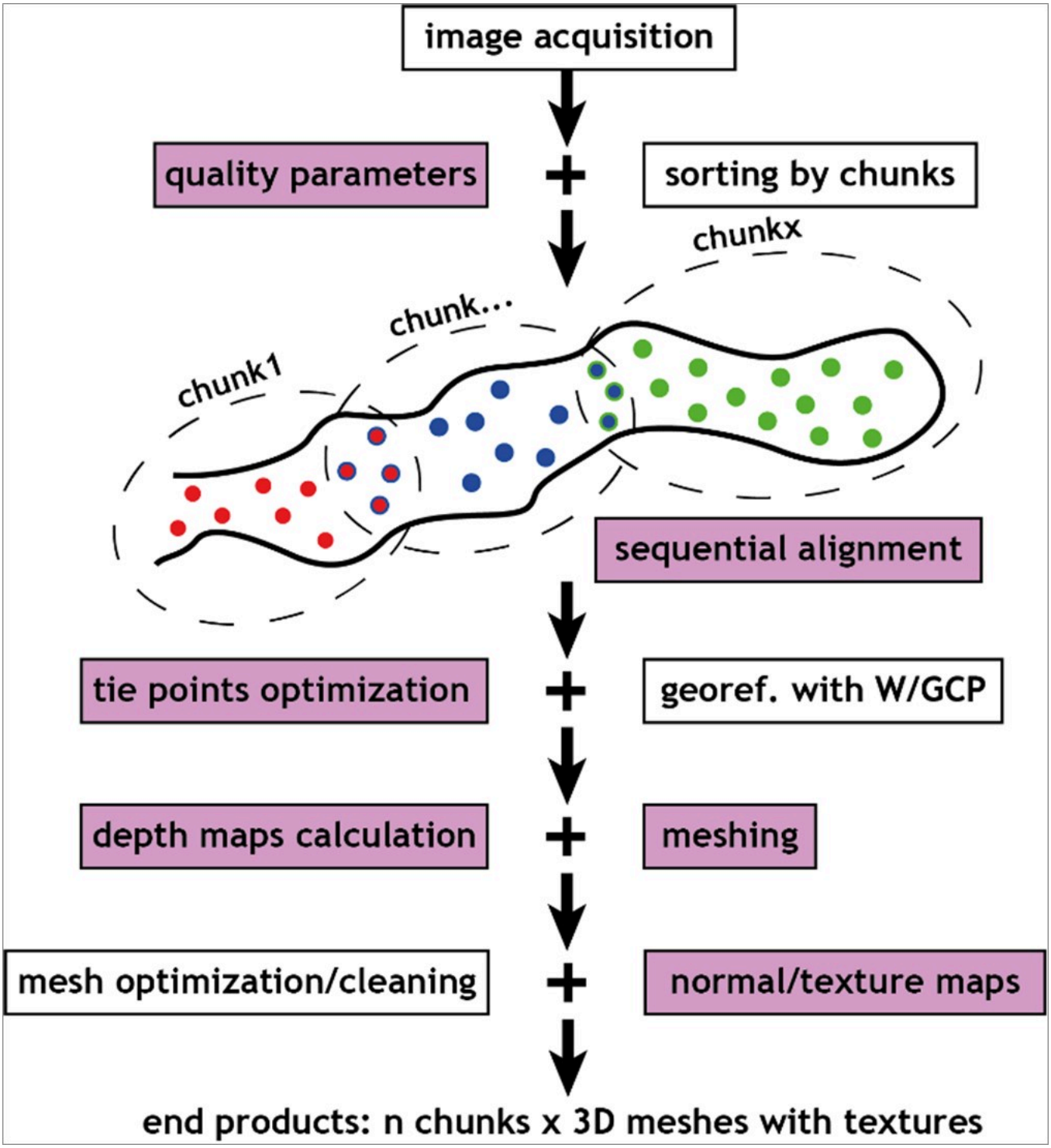


Figure 2

Postprocessing schematic view for image analysis and photogrammetric calculations to generate textured and scaled 3D meshes from sequences of images. White rectangles illustrate human inputs while pink ones show automatically processed steps.



Figure 3

Reference points (topographical nails with diameter of 15 mm) measured with the total station clearly visible in the photographs allowing a precise registration of the 3D models.

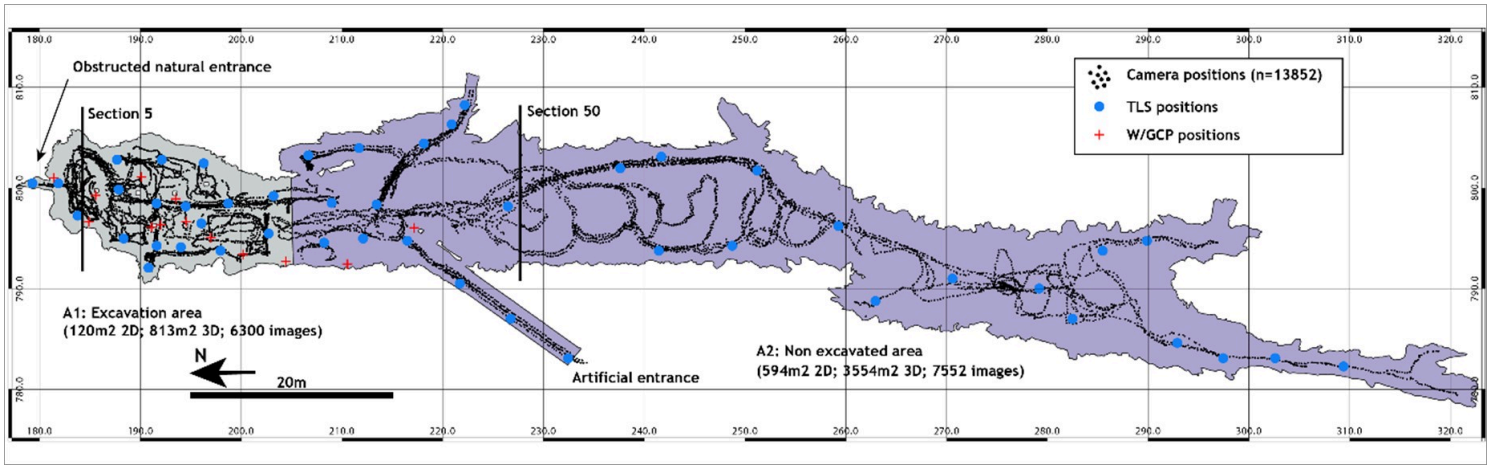


Figure 4

Plan view of the cave (orthographic top view of the Sfm model) with positions of 1) the camera poses (dark dots); 2) the Lidar scans (light blue large dots) and 3) the W/GCP. The two vertical black lines show the positions of the sections described in Fig. 8.

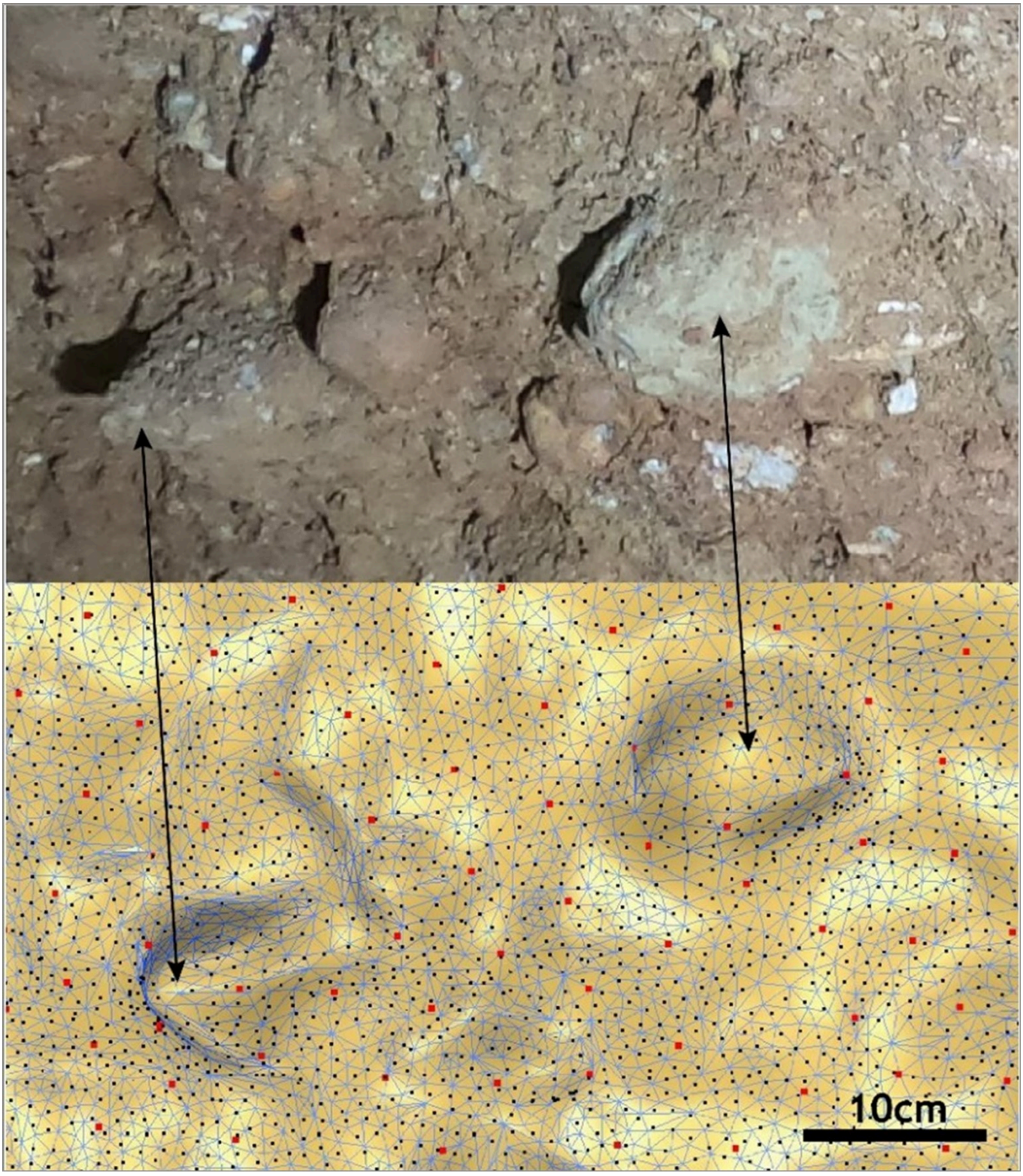


Figure 5

The same ~10cm long pebbles represented by the final SFM model (yellow surface and blue triangles); a subsampled 1cm/25M pts TLS cloud (black points) and a subsampled 5cm/1.5M pts TLS cloud (red points). While the TLS data exhibits a better density, pluri-centimetric objects are well represented by both methods.

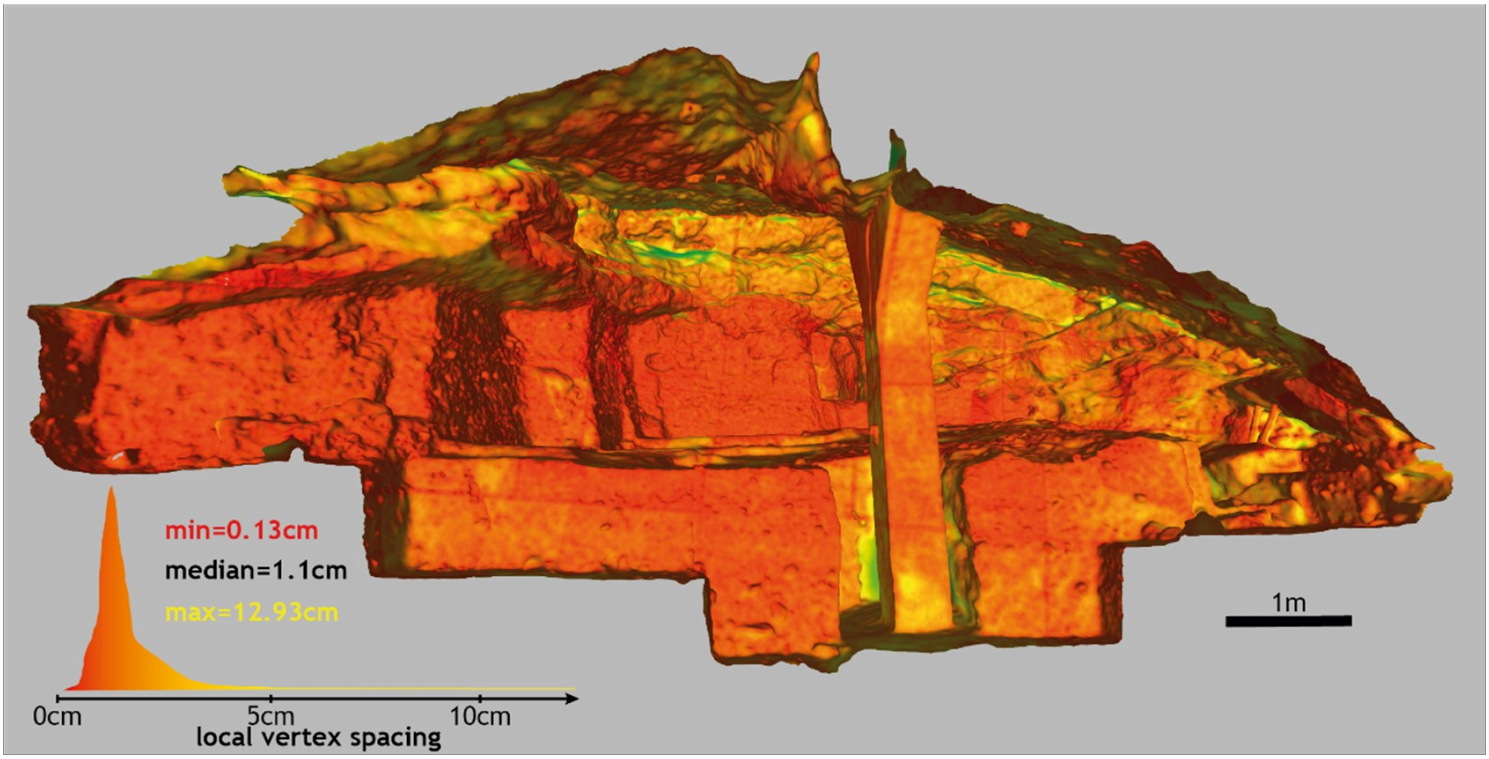


Figure 6

A portion of the full resolution Sfm mesh with colour coded local average neighbour distance (a visual representation of point density made by calculating the average distance to the nearest n points for each vertex). While the neighbour distance is in most part sub-centimetric (toward red colour), some portions of the model show lower pluri-centimetric resolution due to increased capture distance (upper parts with ledge) and/or occlusions (behind the concrete pillars)

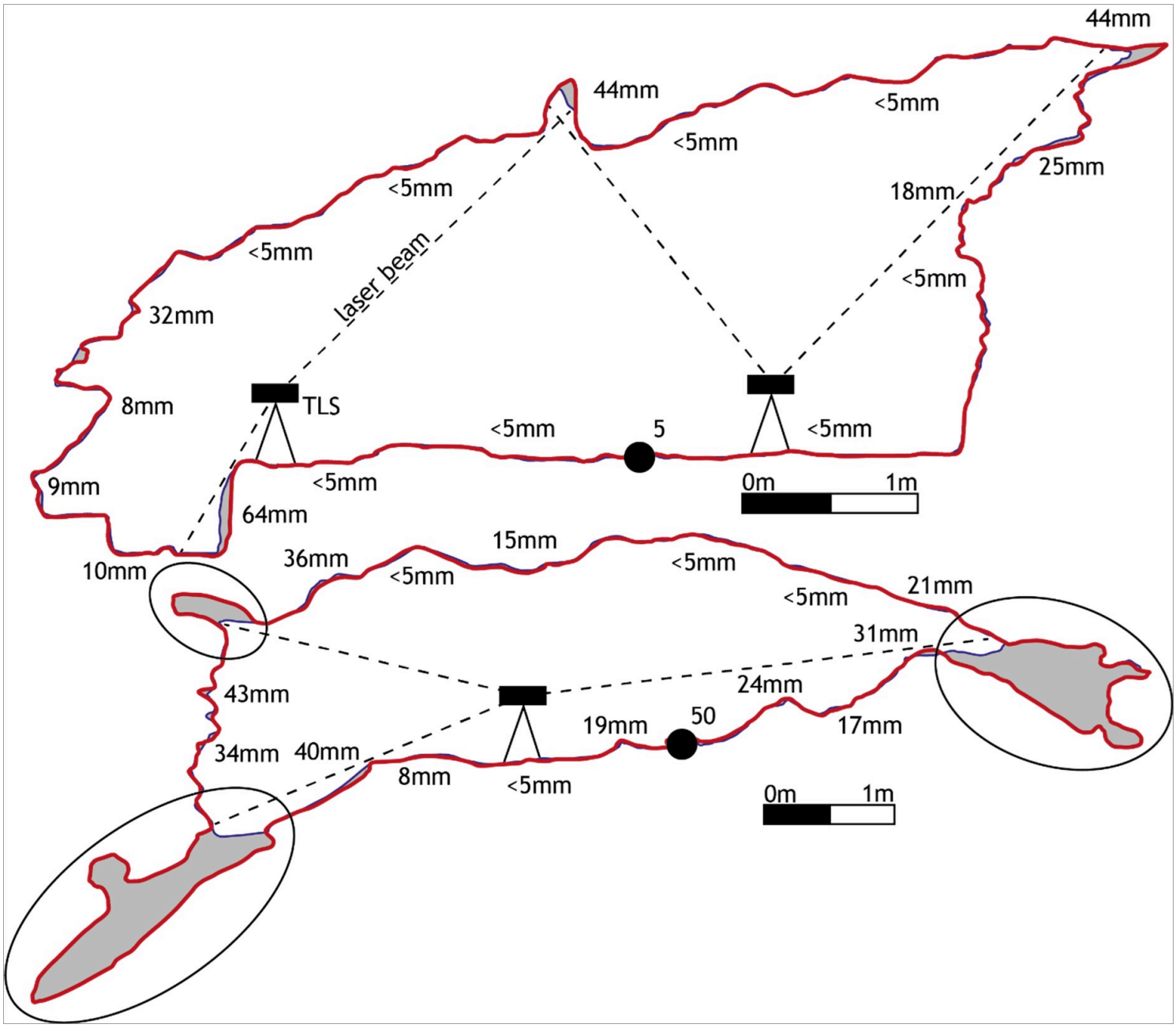


Figure 7

SFM (red)/TLS (blue) comparison along vertical cross-sections in the excavation zone (section 5) and in the more distant area (section 50). The cross-section locations are given in Figure 4. The major visible differences (greyed areas) are mainly explained by occlusions in the TLS data.

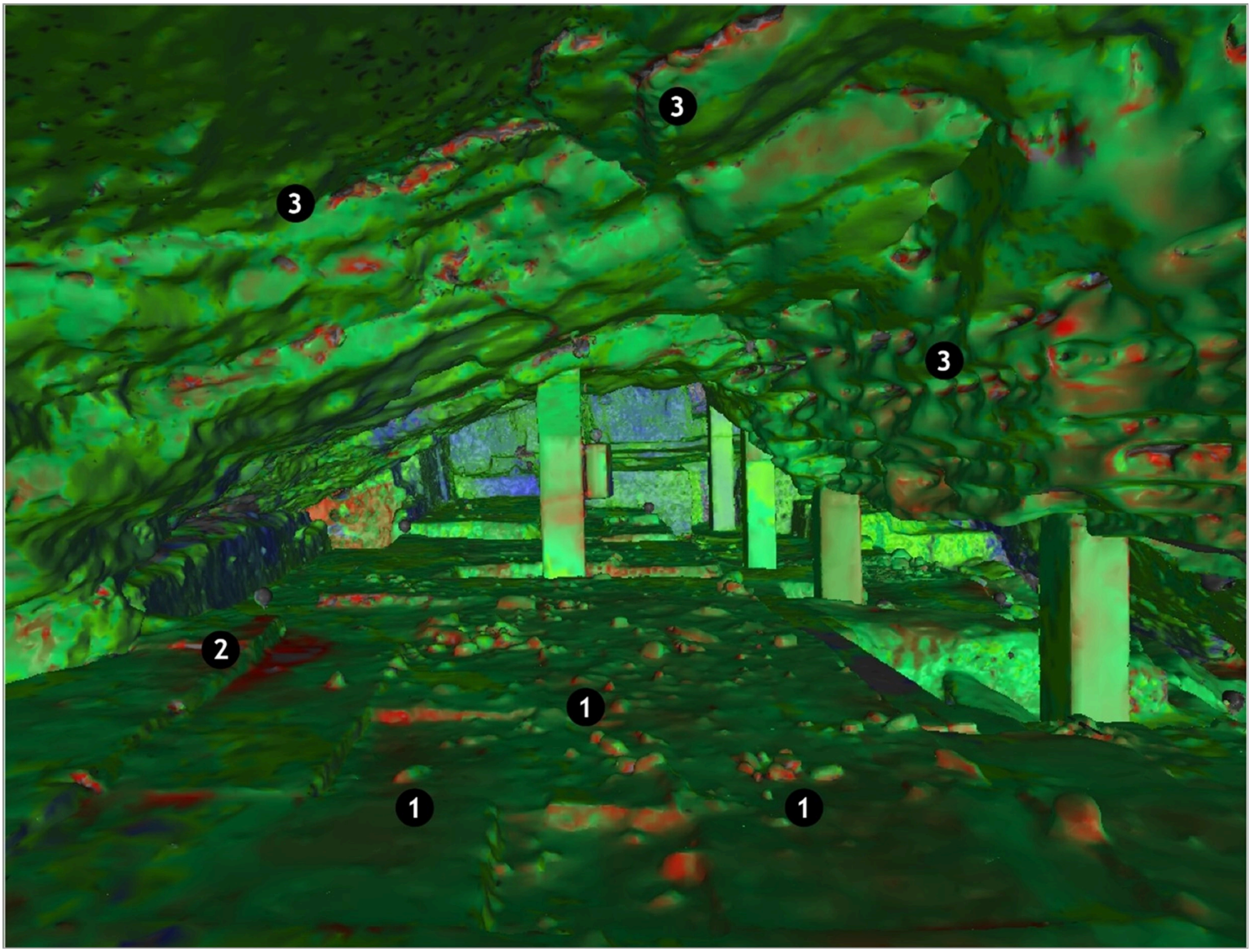


Figure 8

SfM/TLS mesh to mesh comparison (colour ramp saturated to +1cm=red and -1cm= blue). The most important differences are often explained by occluded areas in the TLS data (one side of pebbles on the ground (1); shade area under the TLS station (2) and sharp depressions/pockets on the ceiling (3)). Note: the cement pillars were built during EB excavations to support the roof

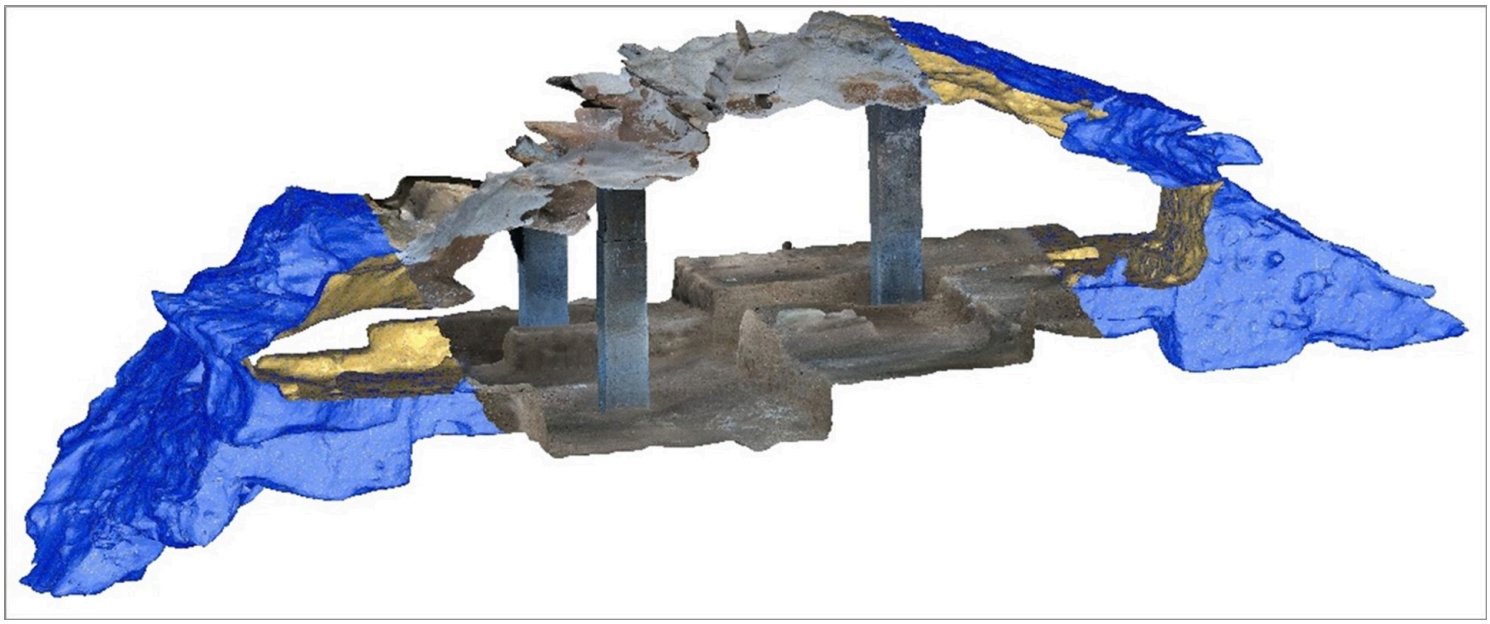


Figure 9

A portion of the cave 3D mesh with and without photorealistic texture

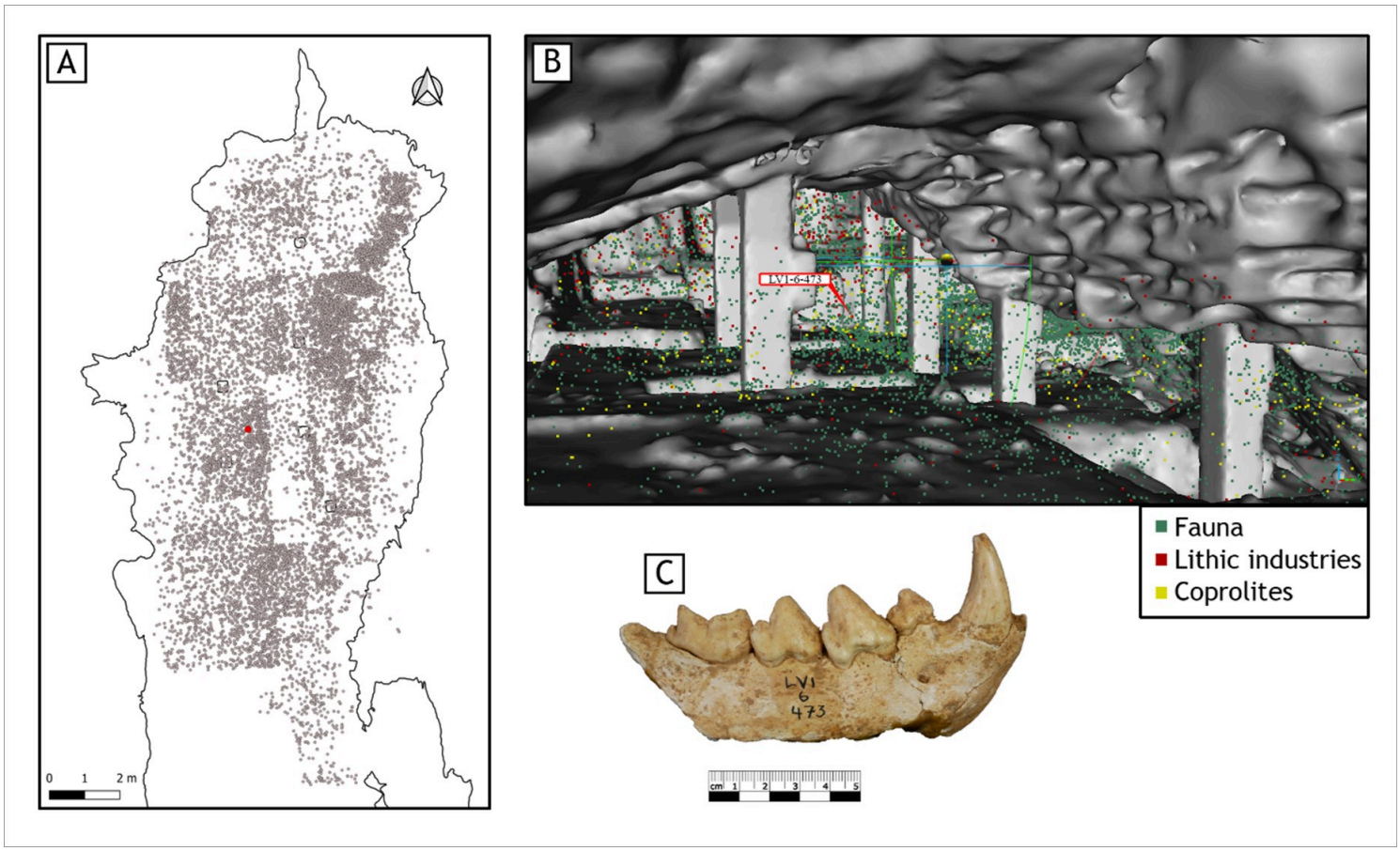


Figure 10

All combined remains (fauna, lithic industry coprolite) from EB and JPB excavations, converted into a similar and common reference and replaced within the 3D envelope, seen in a top plan view (A) and inside the model with a perspective view (B). (C) typical fauna remain: hyena jawbone LVI-6-473, located in B).