RAPID AND LOW-COST PHOTOGRAMMETRIC SURVEY OF HAZARDOUS SITES, FROM MEASUREMENTS TO VR DISSEMINATION

L. Perfetti [1], S. Teruggi [1], C. Achille [1], F. Fassi [1]

[1] Politecnico di Milano, Dept. of Architecture, Built environment and Construction engineering (ABC), Italy
(luca.perfetti, simone.teruggi, cristiana.achille, francesco.fassi)@polimi.it

Commission II

KEY WORDS: 3D Survey, Multi-camera, Photogrammetry, low-cost, VR, 3D assets

ABSTRACT:

In the field of geomatics, the development of 3D surveying techniques that can produce dense point clouds and high-resolution three-dimensional models has opened the possibility of using these techniques to generate digital replicas of a given asset to be used in contexts of cultural promotion or entertainment.
These non-traditional applications have been made easier by the spreading of low-cost but fast surveying techniques, allowing them to be used in contexts or projects with fewer resources. Close-range photogrammetry is nowadays the cheapest, most flexible, and fastest technique that can be used to produce high-resolution and high-quality digital copies of reality. In this paper, we present an application of the Ant3D multi-camera system, a novel photogrammetric device designed for the rapid 3D surveying of sites characterized by poor accessibility and narrow spaces. The paper describes the three-dimensional acquisition of underground tunnels in a hazardous abandoned mining site. The digitalisation work aims to create 3D models to be experienced safely in a virtual- or augmented- reality immersive navigation of the site for cultural promotion. A workflow is proposed to turn raw reality-based 3D models into refined and efficient 3D assets that can be exploited in a game engine. Applications for mobile devices, desktop, and VR headsets, using the produced 3D assets are described.

1. INTRODUCTION

In geomatics, we assisted to the development and widespread adoption of 3D surveying techniques capable of producing dense point clouds and high-resolution three-dimensional models. In recent years, the diffusion of these techniques exceeded the original application they were developed for: accurate 3D measurement. Nowadays, they are also used to create digital copies of different assets (architectural or naturalistic) for cultural promotion or entertainment.

This is the case of 3D models acquired/created focusing on visual quality and valorisation potential alongside or before the focus on metric accuracy. The purpose of such models ranges from documenting, preserving, and especially promoting cultural heritage to create reality-based 3D assets to be placed alongside artificial 3D models for the realisation of virtual scenes for films or video games.

In particular, low-cost rapid surveying techniques have further facilitated these non-traditional applications, spreading them to more resource-limited contexts or projects. This is the case of close-range photogrammetry employed from the ground or by a drone to create digital copies and 3D assets (Herrmann and Pastorelli, 2014; Comes et al., 2022; Cruz et al., 2022). As testimony of the growing use of reality-based photogrammetric 3D model for non-metrological application, there are productions such as the series "The Mandalorian" by Disney, and video games such as "The Vanishing of Ethan Carter" by the developer's studio "The Astronauts", both made use of image-based 3D assets to build their virtual scenes. This context of increasing demand for reality-based models is pushing the research towards improving the methods to acquire data, to process the models and to optimize them for a more effective display.

This paper presents a workflow to build a digital replica of a complex site of cultural interest considered dangerous to visitors and meant to be closed and experienced safely through VR (Virtual Reality) or AR (Augmented Reality). The paper covers: (i) the challenging three-dimensional acquisition of the hazardous site employing a photogrammetric multi-camera prototype (Chapter 2), (ii) the approach followed to turn the raw 3D model into a refined and optimised one (Chapter 3), and (iii) the opportunities associated with VR dissemination (Chapter 4).

1.1 Case study and objective

The case study is an abandoned mining site located in the Italian province of Sondrio, close to the municipality of Chiesa in Valmalenco, located on a mountainside that has been utilised for centuries – and up until a few years ago – for the extraction of serpentine. Although serpentine is nowadays mined using modern equipment by quarrying the rock on the surface, the mountainside still bears evidence of earlier tunnel-based underground mining operations. Many of these hollow tunnels are still preserved and reachable by foot pathways, a testimony of the history of the Valmalenco valley. Recently, it has been agreed that these tunnels, called "Giovelli" or "Giuél" should be closed due to the significant risk of collapse. A solution was therefore to be found to digitally preserve the location in its current state of conservation and, at the same time, to promote and disseminate the history of the serpentine quarrying in the valley, potentially allowing visitors to virtually tour the underground mines.

Within the context of an Interreg project (Italy-Switzerland 2014-2020 id. 631431), the goal was to develop a cultural installation along the mountain footpaths that would tell the story of serpentine quarrying in the area and offer the opportunity to explore the old mines virtually.
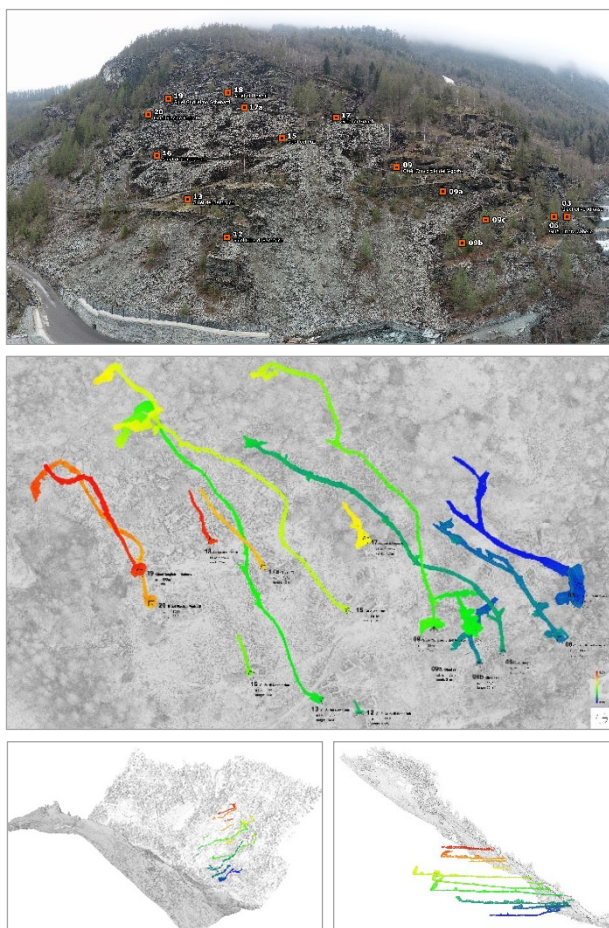
Figure 1. Panoramic view of the slope affected by the tunnel's entrances (top), and: plan (middle), axonometric (bottom left) and section (bottom right) views of the tunnels.

For these reasons, a thorough 3D digital study of the site was carried out before the tunnels' closure.

Fifteen serpentine quarrying tunnels were surveyed and modelled in this context. The tunnels, on average, measure about 1m x 1.5m in width and height respectively, and are about 150m long when not obstructed by the debris of previous collapses.

Figure 1 shows the tunnels' location on the mountainside and their extension underground while Figure 2 shows some pictures.

2. RAPID DATA ACQUISITION

2.1 Ant3D multi-camera system

The tool used for the three-dimensional data collection is the prototype of a photogrammetric multi-camera called Ant3D, which was developed for rapid on-the-go acquisition of confined spaces. The device consists of a small backpack housing a computer, a battery, and a handheld multi-camera. The device mounts five cameras employing fisheye lenses and a global shutter sensor and three illuminators. The multi-camera has an omnidirectional angle of view that is therefore able to frame in a single take the entire environment except for the portion occupied by the operator (Perfetti et al. 2022a). Ant3D has demonstrated in previous tests that it is capable of capturing complex and tortuous environments in short times and that it can guarantee architectural-range accuracy of 2-3cm with the use of Ground Control Points (GCPs) and drift of about 5-10cm for every 100m for unconstrained trajectories (Marotta et al., 2022; Perfetti et al., 2022a; Perfetti et al., 2022b).
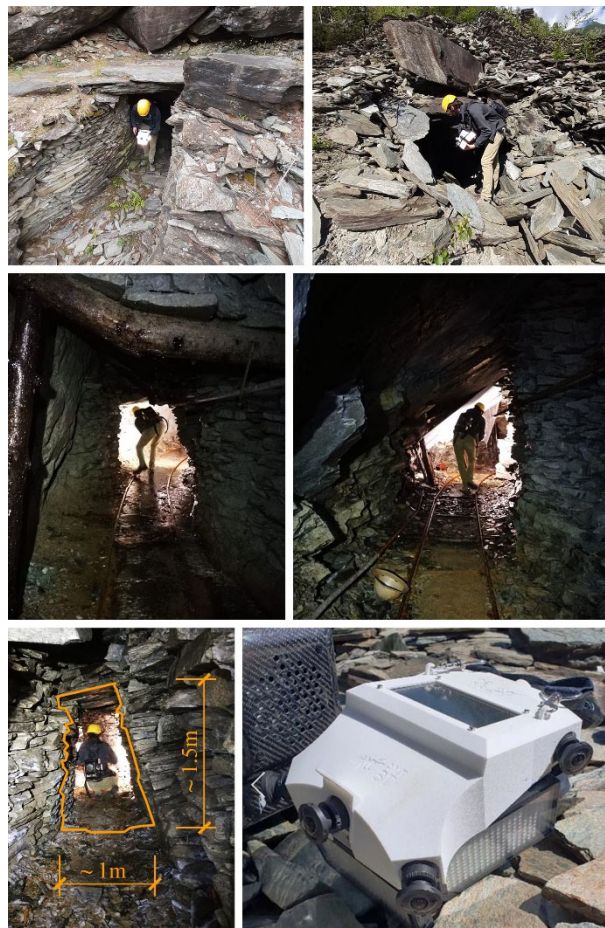

Figure 2. Photos of the entrance (top) and inside areas (middle) of the tunnels during the 3D acquisition carried out using the Ant3D multi-camera system (bottom right).

The proposed multi-camera device offers some key features that makes it ideal for the case study in hand:

2.1.1 Robust to drift error, even in the absence of GCPs, the multi-camera can deliver results free from noticeable distortions or bends. Based on previous tests (Perfetti et al., 2022a), for the longer tunnels (~ 150m) the expected maximum drift error in these environments is from 5cm to 15cm.

2.1.2 Manoeuvrability, it is portable and easily carried around even in such narrow environments, allowing one to complete the acquisition quickly and not exposing the operator to prolonged danger.

2.1.3 Good colour information, although the individual images are moderately low resolution (5 megapixels), the colour information provided by the photogrammetric approach is denser than the one supplied by coloured range scans.

2.2 Data acquisition and processing

The data acquisition spanned 3 days: one spent for a photogrammetric drone acquisition of the mountainside, while two days were necessary to complete the Ant3D multi-camera acquisition of the 15 tunnels. The survey started immediately outside the entrance of the tunnels with the purpose to collect some GCPs materialised with black and white targets placed on the ground. The constraint points were measured with a GNSS (Global Navigation Satellite System) receiver in RTK (Real Time

Kinematic) in order to georeference the external and internal acquisition together. No control points were used along the extent of the tunnels because of the impossibility of measuring them with a total station.

The extent of the underground environments was acquired in both the forward and backward directions to cover the undercuts of the irregular surface of the stone walls. Multi-images (five photographs each) were acquired at a 1 frame per second rate collecting sequences of ~3000 up to ~14000 images in 10 to 45 minutes.

Table 1 summarises the survey information of all the acquired image sequences. For all tunnels, the survey time was modest given the surrounding circumstances; in essence, the survey just required the time to walk along the mine undergrounds.

The image datasets were processed in full resolution using the software Agisoft Metashape. The calibrated distances between the images composing the multi-camera rig were constrained during the SfM (Structure from Motion) as described in Perfetti et al. (2022a). Following successful SfM, the mesh models of all tunnels were processed. Figure 3 shows some views of the obtained results. It can be noticed that the rapidly acquired datasets could register great details all around. Figure 4 shows two examples of the acquired multi-images.

Tunnel name	Tunnel length	Acquired images	Survey duration
Giuèl 03	84.5 m	12270	~ 42 min
Giuèl 06	73 m		
Giuèl 07	150 m	12000	~ 40 min
Giuèl 09b	33 m		
Giuèl 09a	31 m	4515	~ 16 min
Giuèl 09	156 m	10190	~ 38 min
Giuèl 10	15.5 m	5405	~ 18 min
Giuèl 17	42 m		
Giuèl 12	4.5 m	13800	~ 46 min
Giuèl 13	154 m		
Giuèl 15	158 m	10460	~ 45 min
Giuèl 16	15 m	3015	~ 11 min
Giuèl 18	22 m	4460	~ 17 min
Giuèl 19	66 m	3275	~ 11 min
Giuèl 20	95.5 m	6465	~ 21 min

Table 1. Summary of the main information regarding the tunnels acquired with the Ant3D multi-camera system.
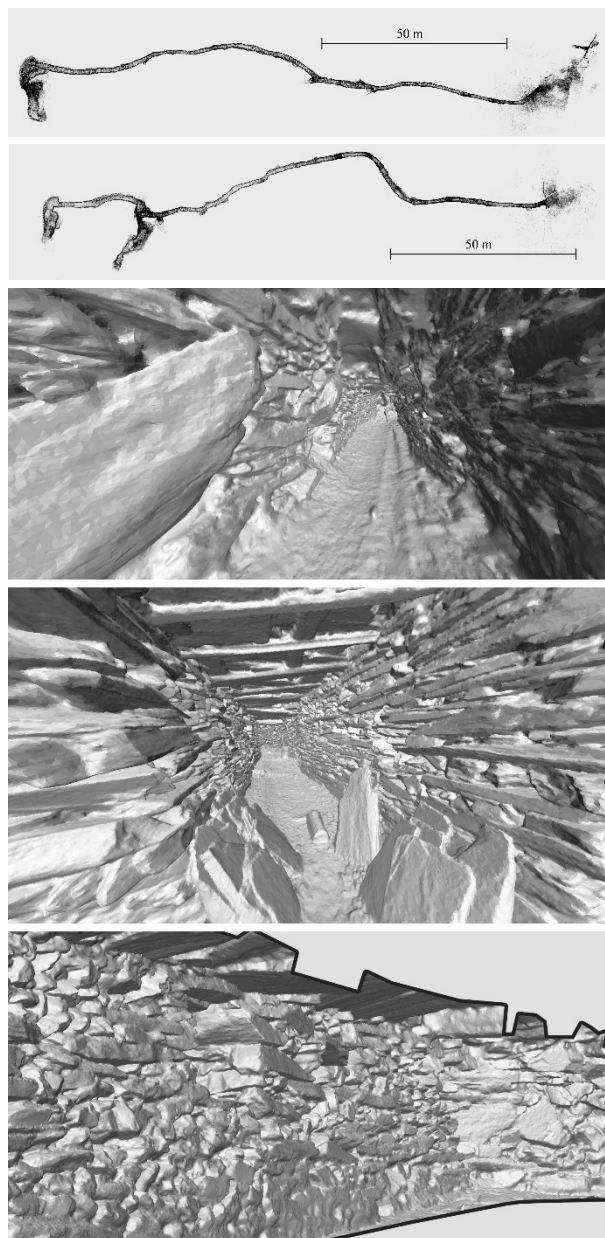

Figure 3. Sparse pointcloud of tunnels: 13 and 15 (top two) and view of the full resolution mesh model of tunnel 13 (bottom).
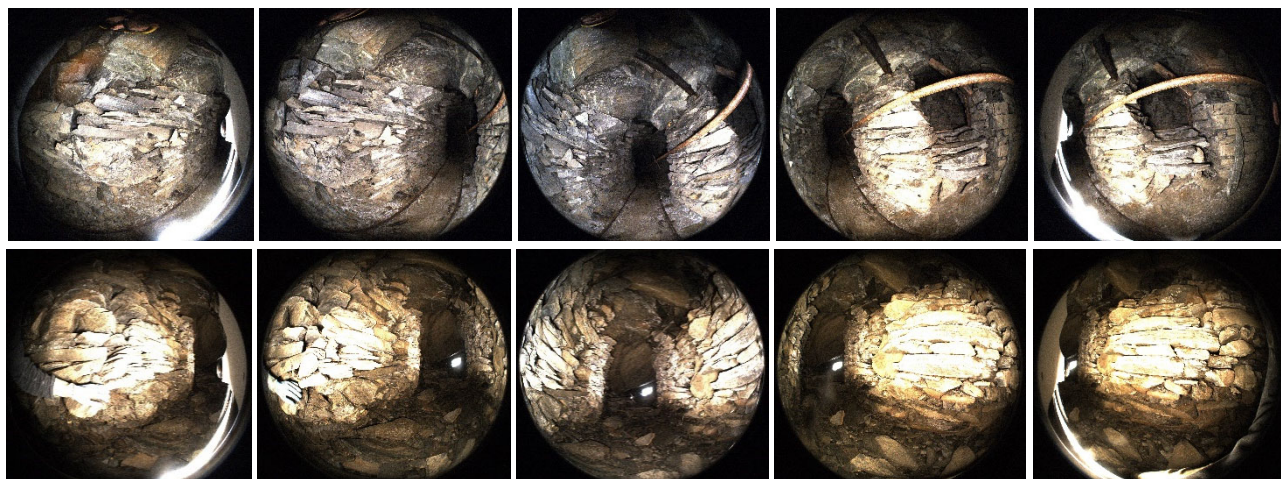

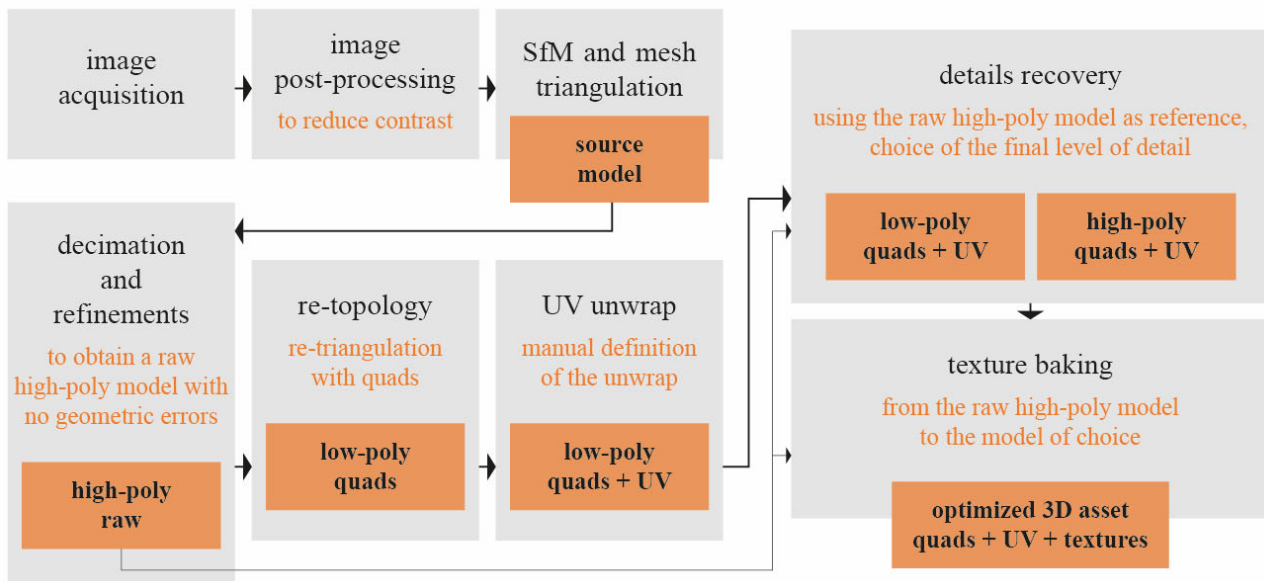**Figure 4.** Two multi-images taken by the Ant3D multi-camera device.

Figure 5. Scheme of the proposed workflow from image acquisition to 3D asset processing.

3. SOURCE 3D MODEL TO 3D ASSET

A pipeline of geometric reduction of the full-resolution models must be used to lower the number of polygons and adapt the model to the computational capabilities of mobile or VR devices to create a 3D model that is optimal for these platforms. Today's standard for the game and cinema industries is to balance the level of geometric simplification of the source 3D model with the creation of texture maps, such as the occlusion and normal maps, that, when applied to the simplified model in a game engine software (Unity, Unreal Engine) can mimic much of the details lost during simplification (Verhoeven, 2017). These refined 3D models, optimized for the real-time rendering of complex scenes, are commonly referred to "3D assets". They are ready-made elements that a 3D artist can handle, modify, and assemble into a virtual scene. The difference between the refined 3D model (3D asset) and the source one is crucial for effectively using the model with state-of-the-art software and standards.

3.1 Proposed workflow

The workflow described in this chapter and illustrated by Figure 5 and Figure 11 details the proposed method to turn the source, image-based mesh model, into a "3D asset". A similar pipeline was also presented by Verhoeven (2017). The pipeline is illustrated for the dataset of the tunnel named: "Giuèl 13".

The workflow envisages:

- The creation of a refined high-poly model;
- Re-triangulation or re-topology of the mesh into a strongly simplified low-poly model;
- Manual creation of the UV mapping, i.e.: the planar unwrap of the 3D model on the 2D cartesian plane of axes U and V;
- The recovering of geometric detail from the high-poly model to the low-poly one up to the desired level;
- Textures "baking", the creation of the photographic texture, occlusion, and normal maps.

High-poly model creation: The original image-based model is typically a high-poly one, with a number of triangles significantly higher than what modern gaming engines allow for consumer platforms. For Giuèl 13, the source model has 80+ Million faces. Moreover, the source model typically shows some errors in the triangulation (non-manifold geometry and blobs) that need to be solved. In the high-poly model creation phase, the source model is decimated to a lower polygon count but still preserving most of the high details and solving the geometric errors by manual or automated actions. The output of this phase is a high-polygon raw model free of flaws.

3.1.1 Re-topology: From the newly obtained high-poly model, the process of re-topology performs a re-triangulation that outputs a lower polygon model organized in quads (Figure 6), i.e., possessing an ordered triangle distribution in squared tiles. A quads model is easier to be altered by a 3D artist and is the industry standard for 3D assets. The re-topology in quads of Giuèl 13 was performed using the software InstantMeshes. A low target triangle count was specified in the software before the re-topology method was performed since it has been observed that a simpler output model results in fewer triangulation errors while making the creation of additional faces easier. As for the previous phase, any introduced triangulation errors must be resolved manually or with automatic algorithms. The output model of Giuèl 13 has 15k faces, which, considering the complexity of the tunnel and its details, can be considered highly simplified.
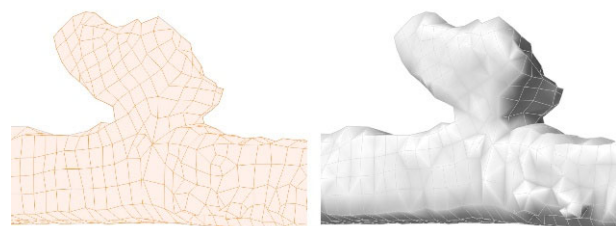

Figure 6. Piece of the tunnel 13 model after re-topology process. Wireframe (left) and solid view (right).

3.1.2 UV unwrap: the low-poly model organized in quads is ideal for manually defining the seam lines to cut and unwrap the 3D model on a plane, i.e. UV mapping. Manually defined seams allow unwrapping the 3Dmodel in a more rational and humanly understandable configuration. Moreover, the texture maps produced using UV parametrisation can be more easily inspected and modified. Figure 7 shows the UV parametrisation of Giuèl 13 divided into 8 planar maps. The unwrap has been manually defined by tracing seam crosscuts segmenting the tunnel in consecutive "cylindrical" shapes and then by tracing two longitudinal seam cut separating the ground, allowing the cylinder to unroll onto the plane.
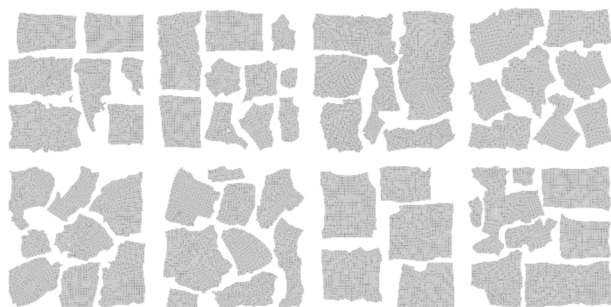

Figure 7. Regularized manual UV unwrap of tunnel 13 model in eight different maps.

3.1.3 Restoring details: The 3D model obtained so far is a low-poly model (15k faces), organized in quads and possessing a rational UV parametrization. This refined model can be used as it is for the following texture baking step. However, the level of geometric simplification of the re-topologized model may be considered insufficient. In this case, a higher level of details can be restored by: (i) subdividing the quads element of the low-poly model into smaller components, effectively increasing the number of polygons; and (ii) fitting the vertexes of the subdivided mesh model to adhere to a reference model e.g.: the high-poly model, effectively restoring details.
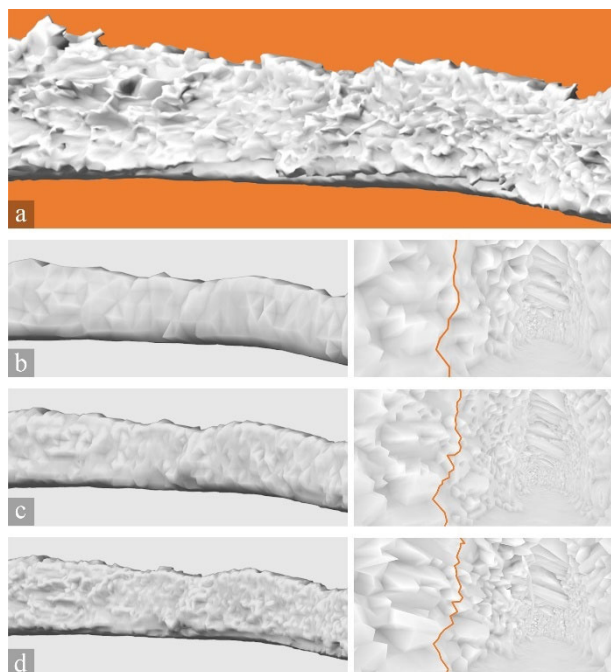

Figure 8. The gradual recovery of details applying the shrink wrap modifier in Blender. High-poly model raw with 20M faces (a), and re-topologized low-poly models with: ~ 15k faces (b), ~ 60k faces (c) and ~ 240k faces (d).

This process has been performed using the "Shrink wrap modifier" in the software Blender, Figure 8 shows the high-poly model used as reference (a), and three other models obtained with a subdivision level of 0 (b), 1 (c) and 2 (d). To a higher subdivision lever corresponds a higher level of restored details. This method allows us to obtain the desired level of detail from the input low-poly model up to a high-poly model maintaining the quads ordered layout of the mesh and the UV parametrisation obtained in the previous steps.

3.1.4 Textures baking: Once the refined 3D model has been obtained, the desired texture maps can be computed (texture baking). The most used maps are: the photographic texture (also known as diffuse map), the occlusion map and the normal map (Figure 10). Together they can represent a much higher level of detail than the one possessed by the actual 3D model itself, allowing for improved real-time rendering performance.
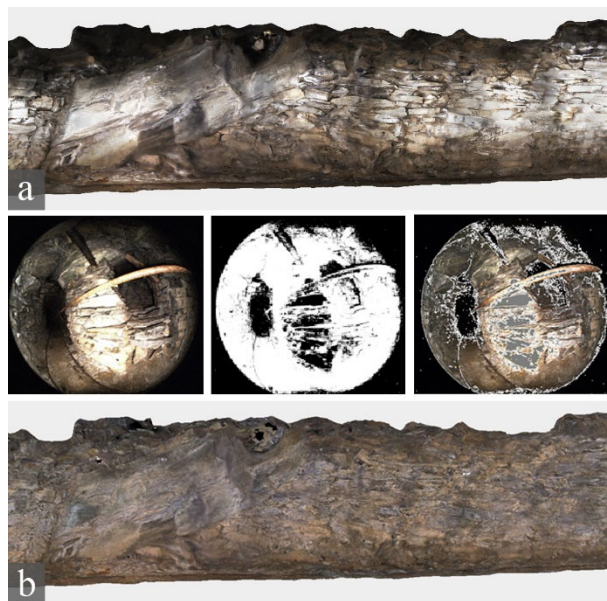

Figure 9. Images texture applied to the model: default outcomes (a), and result obtained after the images are masked (b). The central section shows an example image and the relative applied mask.
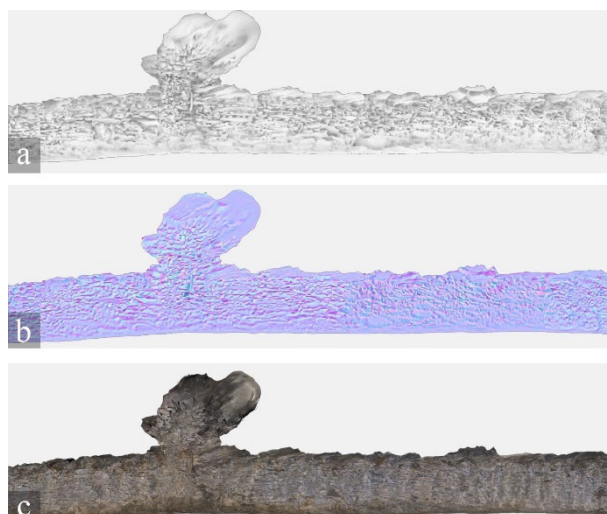

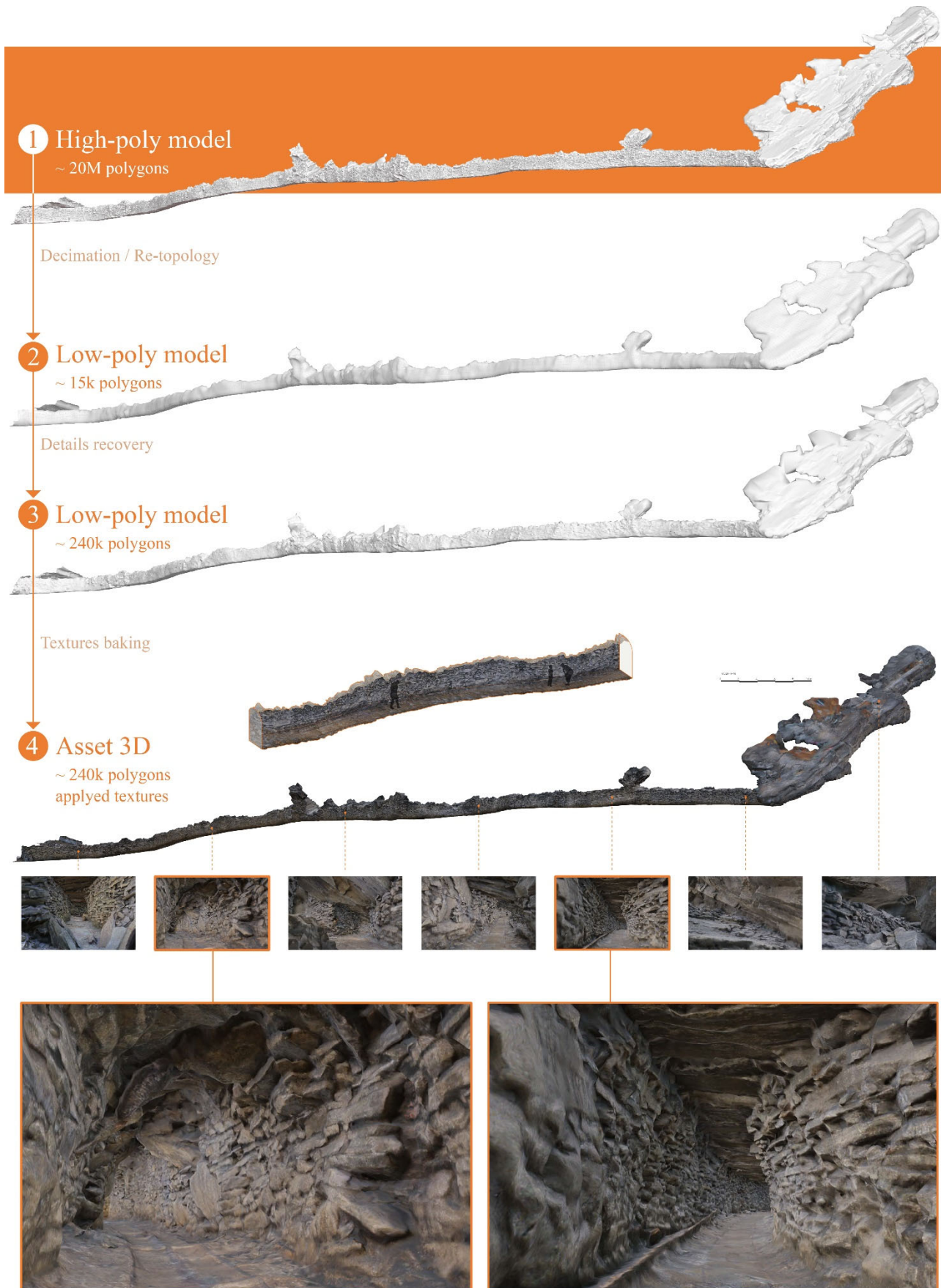**Figure 10.** Occlusion (a), normal (b) and diffuse (c) maps baked for the tunnel 13 model.

Figure 11. Visual synthesis of the proposed workflow. The high-poly reality-based model obtained at the end of the photogrammetric process is turned into a 3D asset usable for VR and AR application.

Figure 9 illustrates an issue that affects the image datasets of the tunnels acquired with the multi-camera system. The surfaces closer to the multi-camera are noticeably brighter than the rest of the image, especially the more distant part of the tunnel, due to how the multi-camera lights the surrounding with three illuminators on the device, as Figure 4 illustrates. For this reason, although the stone surfaces have similar albedo, the image brightness varies significantly, resulting in an un-homogenous diffuse texture. The problem has been tackled in two ways: (i) by processing the raw images to reduce the contrast and (ii) by masking the images in the areas where the brightness is below or higher than certain thresholds. Figure 9 shows that this step was crucial to creating a good, diffuse texture.

4. VR/AR ACCESSIBILITY

Using the game engine Unity (Unity, 2022), it has been possible to create a demo immersive cross-platform visualisation experience, starting from the more "traditional" monitor display (desktop), through a mobile display (smartphone), to immersive virtual and augmented reality (VR headset and smartphone) fruition. The advantage of using Unity is that it has functions to interface with most existing devices, either internally or using third-party libraries. The 3D model can be imported within the scene where all the possible user interactions can be defined.

The first step, and the most important for a high-quality fruition of the digital asset, is the definition of the object materials and especially how these are rendered through the screen of the selected device. Using a Unity material that uses a "Standard shader", it is possible to assign the three texture maps (diffuse, normal and occlusion maps) to the object. These maps are crucial in defining how the 3D objects will look independently of the fruition platform. Aside from the photographic texture ("albedo" in Unity), the normal and occlusion maps affect the shading of the low-poly model, rendering it more detailed than it is.
Having set up the asset materials, it is essential to define the interactions that the user can have with it. They strongly depend on the visualisation medium as each device requires specific libraries that interface with specific hardware.
VR fruition has been performed on the Meta Quest 2, a standalone VR headset that runs on the Android platform. Unity can dialogue with it using the "Oculus integration v46.0". This provides out-of-the-box functions that allow setting up a first-person tour of the selected digital asset. A spotlight has been tied to the movement of one of the two controllers to give a more realistic feeling to the VR tour, where the user can control a virtual flashlight (Figure 13).
Desktop and mobile visualization use the same libraries ("Starter assets") that provide first-person character controller functionality for mouse/keyboard and touchscreen controls.

The mobile application can be installed on a device through QR code scan and allows visitors that are hiking and touring the area to extend their visit virtually. The visitors can find an informative panel in front of the closed tunnels giving them information on the old mines and presenting a QR code. They can: (i) access non-accessible heritage while on-site, (ii) get a different (and impossible) point of view on the surrounding areas and (iii) virtually tour the closed mines safely (Figure 12). The user is presented with the general model of the area that they can orbit, pan, and zoom. Hotspots tie the position of relevant information (user position and images) to specific points. The visualization can be changed to isolate and highlight the tunnel of interest or, by selecting one of them, the user can be entered through the first-person point of view.

On the contrary, desktop and VR fruition allow bringing the site to the user, enriching installations experiences inside museum visits. This improves traditional media by enabling users to immerse themselves in hazardous environments in an entirely safe setting, from a fly-through of the area to a genuine first-person visit of the mine itself that would normally be impractical to carry out on-site.
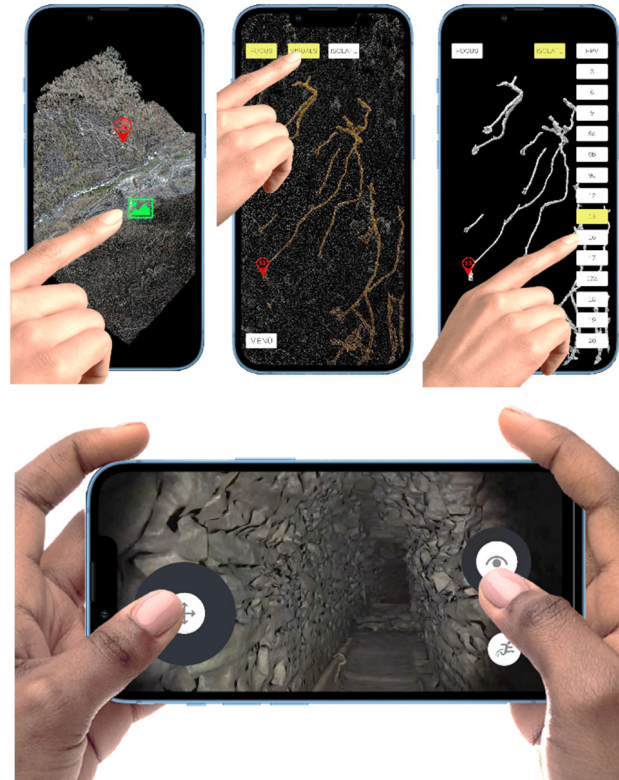

Figure 12. Interaction with the demo mobile application. The app could be downloaded and installed on-site allowing instant access to a virtual extension of the visit. The user can access information, have a comprehensive understanding of the area, and virtually tour the closed mines.

Figure 13. Interaction with the demo VR application. The VR headset allows the users to immerse themselves inside the closed mines and interact in the virtual environment the most similarly to how they would physically.

5. CONCLUSION

The process presented demonstrates an effective low-cost work pipeline from the rapid, complete, and accurate survey to its digital fruition, especially suited for those situations where access to the site is forbidden due to safety reasons.

The proposed multi-camera system made it possible to complete a challenging 3D survey in a very short time. Even without the use of GCPs, the obtained results do not present visible distortions, bends or drifts, and following the previous tests (Marotta et al., 2022; Perfetti et al., 2022a; Perfetti et al., 2022b) it is possible to estimate a maximum drift error in the acquisition of the tunnels to about 5-15cm. This ensures that the data can also be used for scientific purposes and geological analysis.

Moreover, the multi-camera system could record the tunnels' geometry to a high level of detail, as shown by Figure 3, even if the acquisition on-field had to be rapid and the image network simple and not ideal.

The image-based model was simplified and optimized using a pipeline, which substantially reduced the number of polygons from 80+ million to about 240k faces in the finished product while maintaining nearly all the original details.

Improvements can be made to the photographic texture quality limited by the low-resolution fisheye images that the multi-camera uses (5 megapixels per image) (Figure 4). The use of two higher-resolution cameras, specifically for the purpose of textures projection, will improve the VR immersive experience.

The proposed tools and methods proved effective in building the digital replica of inaccessible hazardous areas and allowed their fruition in VR across multiple platforms. In this way, areas of historical and cultural interest that would otherwise be inaccessible and inevitably destined for gradual abandonment could be enhanced and disseminated to the public.

ACKNOWLEDGEMENTS

REFERENCES

Comes, R., Neamțu, C. G. D., Grec, C., Buna, Z. L., Găzdac, C., & Mateescu-Suciu, L. 2022. Digital Reconstruction of Fragmented Cultural Heritage Assets: The Case Study of the Dacian Embossed Disk from Piatra Roșie. Applied Sciences, 12(16), 8131.

Cruz Franco, P.A.; Rueda Márquez de la Plata, A.; Gómez Bernal, E. 2022 Protocols for the Graphic and Constructive Diffusion of Digital Twins of the Architectural Heritage That Guarantee Universal Accessibility through AR and VR. Appl. Sci., 12, 8785.

Herrmann, H., Pastorelli, E. 2014. Virtual Reality Visualization for Photogrammetric 3D Reconstructions of Cultural Heritage. In: De Paolis, L., Mongelli, A. (eds) Augmented and Virtual Reality. AVR 2014. Lecture Notes in Computer Science(), vol 8853. Springer, Cham.

Marotta, F., Perfetti, L., Fassi, F., Achille, C., and Vassena, G. P. M., 2022: LIDAR iMMS vs handheld multicamera system: a stress-test in a mountain trailpath, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIII-B1-2022, 249–256.

Perfetti, L., and Fassi, F., 2022a. Handheld fisheye multicamera system: surveying meandering architectonic spaces in open-loop mode - accuracy assessment. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., 46(2/W1-2022), 435–442.

Perfetti, L., Elalailyi, A., and Fassi, F., 2022b: Portable multi-camera system: from fast tunnel mapping to semi-automatic space decomposition and cross-section extraction, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIII-B2-2022, 259–266.

Troisi, S., Baiocchi, V., Pizzo, S. D., & Giannone, F., 2017. A prompt methodology to georeference complex hypogea environments. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII-2/W3, 639–644.

Unity 3D, 2021. Unity 3D Software. https://unity.com/ (02 november 2022).

Verhoeven, G. J., 2017: Computer graphics meets image fusion: the power of texture baking to simultaneously visualise 3D surface features and colour, ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., IV-2/W2, 295–302.