

# Low-Cost 3D Reconstruction of Caves

João Marcelo Teixeira<sup>a</sup>, Narjara Pimentel, Eder Barbier, Enrico Bernard, Veronica Teichrieb  
and Gimena Chaves

*Universidade Federal de Pernambuco, Recife, PE, Brazil*

**Keywords:** RGB-D Sensors, 3D Reconstruction, Cave Surveying.

**Abstract:** Caves are spatially complex environments, frequently formed by different shapes and structures. Capturing cave's spatial complexity is often necessary for different purposes – from geological to biological aspects – but difficult due to the challenging logistics, frequent absence of light, and because the necessary equipment is prohibitively expensive. Efficient and low-cost mapping systems could produce direct and indirect benefits for cave users and policy-makers, enabling from non-invasive research of fragile structures (like speleothems) to new forms of interactive experiences in tourism, for example. Here we present a low-cost solution that combines hardware and software to allow capturing cave spatial information through RGB-D sensors and the later interpretation of the processed data. Our solution allows the navigation in a 3D reconstructed cave, and may be used to estimate volume and area information, frequently necessary for conservation or environmental licensing. We validated the proposed solution by partially reconstructing one cave in Northeastern Brazil. Although some challenges have to be overcome, our approach showed that it was possible to retrieve relevant information despite using low-cost RGB-D sensors.

## 1 INTRODUCTION

Caves are spatially complex tridimensional environments, mainly composed by irregular shapes. Such complexity can manifest from a macro (i.e., conduits, ceiling, walls, floor) to a smaller scale (i.e., stalactites, stalagmites and other speleothems of reduced dimension). Documenting the spatial complexity of caves is a challenging task, because besides their inherent characteristics, the frequent lack of light in the cave's interior make such task even more difficult. Although caves have been subject of study for decades, traditional topography techniques are not always capable of accurately representing the tridimensional complexity of such environments (Gallay et al., 2015).

Recently, 3D reconstruction techniques have been applied as an alternative for cave mapping and documentation. Such approaches are usually based on laser scanning systems (i.e., Terrestrial Laser Scanning - TLS, Light Detection and Ranging - LiDAR or Mobile Laser Scanning - MLS) (Ullman et al., 2023) (Li et al., 2023), where laser beams with thousands of pulses per second are used to generate point clouds, which will be recorded together in sequence. Such approach allows to construct highly detailed 3D rep-

resentations of the environment captured, in high resolution (dense), enabling a refined view of the cave (Grohmann, 2019).

The use possibilities of 3D reconstructed representations of caves are vast and go beyond scientific goals, including from cave documentation to a more precise estimate of cave volume and area, all useful information for the environmental licensing, ecotourism and environmental education. However, despite being very promising, the use of laser scanning techniques in Brazil is still immature and hindered by some factors:

- **Cost:** Laser scanning equipment is very expensive, with prices varying from thousands to hundreds of thousands of Reais (R\$);
- **Difficult Access:** the necessary equipment has to be imported requiring time-consuming and bureaucratic processes. Moreover, some laser scanning solutions are rated restrict due to military or security use;
- **Handling:** Data volume generated by such systems is usually very large and requires high performance computers.

Such factors make the diffusion of laser-based 3D mapping techniques very restrict in Brazil, lim-

<sup>a</sup>  <https://orcid.org/0000-0001-7180-512X>

iting their access and restricting their use to a few big companies, some specialized consulting agencies and an even smaller number of research laboratories (Grohmann, 2019). Such restrictive scenario in Brazil is very concerning, considering the huge speleological potential of the country, estimated at about 310,000 caves (?). This situation is aggravated by the lack of information about Brazilian caves and, at the same time, by the high human pressure they experience.

A low-cost 3D alternative for cave reconstruction could increase the possibilities of documentation of Brazilian caves, contributing to reduce the current information gap. Moreover, documenting caves and their characteristics is supported by the current policies, like the National Program for the Conservation of the Brazilian Speleological Heritage (item MMA 358/2009). Such initiative provides support for data production which may help the Brazilian government to achieve their conservation and research goals. Besides that, protected areas harboring caves could benefit from cave surveying initiatives, enabling, for example, non-invasive studies of fragile speleothems, as well as by new forms of virtual interactions between tourists and the caves in those areas. Here we present a low-cost solution that combines hardware and software to allow capturing cave spatial information through RGB-D sensors and the later interpretation of the processed data for the extraction of additional data such as cave volume and area.

This paper is therefore structured in the following sections. Section 2 briefly describes some related studies with different solutions for 3D reconstruction of caves, highlighting the technologies used and the direct application of the results obtained. Section 3 explains the methodology we used, with the steps for the proposed process. In Section 4 we discuss the limitations found in the proposed methodology together with some insights on how to improve the results obtained with the proposed solution. At last, in Section 5 we provide future directions that could make the proposed solution more complete.

## 2 RELATED WORKS

The work from Oguchi et al. (Oguchi et al., 2011) divides the main data sources used in earth surface processes into analogic and digital ones. Analog data sources comprise text descriptions, hand-drawn illustrations, analog photographs and videos for visual interpretation, data from classical ground surveying, topographic data from plane-table and analog photogrammetry and topographic maps and thematic

maps. Regarding digital data sources, they include digital ground/aerial photographs and videos for visual interpretation, digital satellite imagery, digital aerial imagery, topographic data from modern ground surveying (GPS, total station, laser range finder, terrestrial laser scanning), analytical and digital photogrammetry, height data from airborne LiDAR and Airborne/Satellite InSAR, com-piled height information and digital topographic maps and thematic maps.

It is important to note that none of the aforementioned data sources include the use of depth (RGB-D) sensors. One could argue that Oguchi's work was published in 2011, while the first low cost mainstream depth sensor (Kinect V1) was released in 2010. A different work, from Idrees and Pradhan (Mohammed Oludare and Pradhan, 2016), which presents the panorama of cave surveying solutions for 10 years (2006-2016), also does not mention any low-cost approaches. From the almost fifty works published in various international journals related to mapping caves in their true 3D geometry with focus on sensor design, methodology, data processing and application development, all of them were based on laser scanning technologies (time-of-flight and phase shift ones).

According to Escalera (Escalera, 2012), it is possible to acquire depth information (to be further used for 3D reconstructions) using different technologies. Figure 1 illustrates different depth acquisition technologies, highlighting the light wave-based ones that do not make use of laser measurements (in other words, the ones that are low cost alternatives). The solution proposed in this work makes use of an RGB-D sensor based on the emission of structured light (camera + light pattern projection) to infer depth information, the ASUS Xtion PRO<sup>1</sup>. By performing a search in Google Scholar using the "3D reconstruction of caves" string, we noticed that most works found make use of laser scanning to acquire cave data (Gallay et al., 2016) (Pukanská et al., 2018) (Gallay et al., 2015) (Beraldin et al., 2006). Despite this tendency, the work from Sellers and Chamberlain (Sellers and Chamberlain, 1998) uses ultrasound reflections while the work from Lee (Lee, 2018) presents the use of RGB cameras (photogrammetry) as a low-cost alternative to the reconstruction of caves, but the work is mostly focused on underwater caves.

Additionally, some works utilize more than one type of sensor at the same time to improve data acquisition (Azpúrua et al., 2023). This happens in the work of MacFarlane et al. (MacFarlane et al., 2013), in which they combine data acquired from aerial photogrammetry using an autonomous drone, three-

<sup>1</sup><https://www.asus.com/3D-Sensor/XtionPRO>

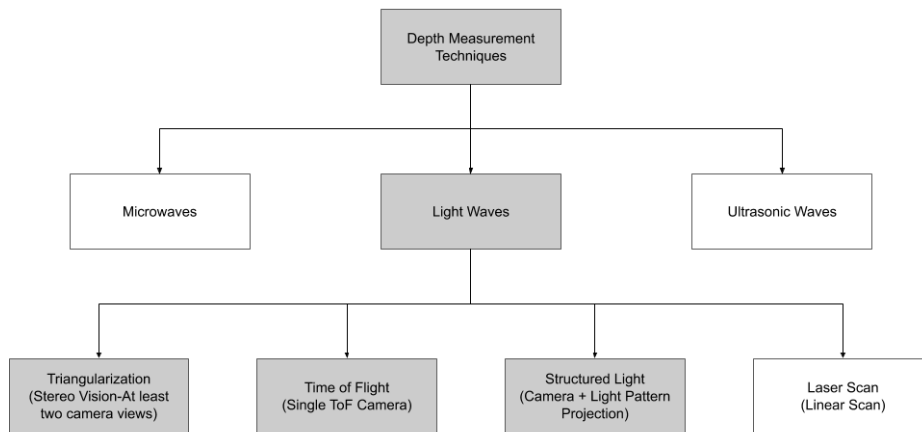


Figure 1: Different technologies for the acquisition of depth maps. Modified from (Escalera, 2012).

dimensional cave laser scanning at millimeter resolution, differentially corrected geodetic GPS, and conventional compass-based cave surveying techniques.

The applications resulting from cave reconstructions go from preservation of the caves themselves (Aiello et al., 2019), to make them accessible to disabled people that cannot visit them “in loco” (Tometzová et al., 2020) and art (Idrees and Pradhan, 2017a). Besides that, it is possible to obtain valuable information such as the cave volume and channel surface area.

### 3 METHODOLOGY

Independently of the data acquisition technology used, the methodologies used for capturing data inside the cave and visualizing the final post processed reconstruction are very similar. Figures 2, 3, 4 and 5 illustrate some of them.

The methodology we adopted can be simplified in three macro stages: data capture, 3D reconstruction of the cave, and visualization/extraction of information based on the generated point cloud.

#### 3.1 Data Capture

The state-of-the-art in capturing tridimensional data (specially for caves) is focused on laser scanning technologies (Grohmann, 2019). The problem with such technologies is that their cost is directly proportional to the quality of the obtained results. A high-resolution laser scanner/LiDAR may reach dozens of thousands of dollars, which restricts their acquisition and use. Low cost alternatives do exist, and it is possible to obtain equivalent results with the use of photogrammetry (RGB-only) or low resolution RGB-D sensors, by combining many captured images. We

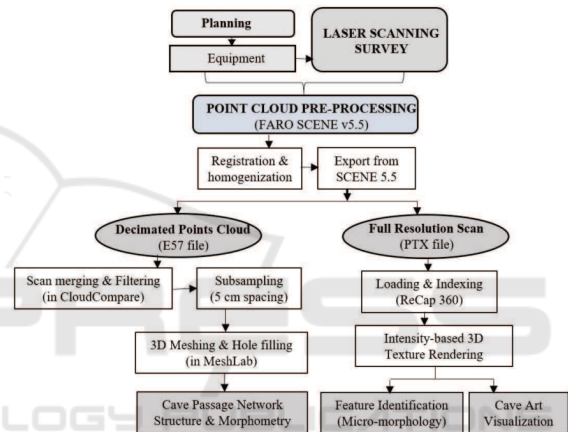


Figure 2: Data processing and analysis workflow from the work of Idrees and Pradhan (Idrees and Pradhan, 2017a).

analyzed several low cost RGB-D sensors focusing different aspects such as size, weight, available APIs, power consumption, resolution and field of view. The ten different sensors analyzed were: Orbecc Astra, Orbecc Astra Pro, Microsoft Kinect V1, Microsoft Kinect V2, Asus Xtion PRO, Realsense D435, Realsense F200, Realsense SR300, Realsense R200, Realsense ZR300. While both Kinect models allowed very good reconstruction results, they need an external power source, which makes their use difficult in the field, inside caves. The Asus Xtion PRO showed similar quality to Kinect V1, with the advantage of no necessity to an external power source, and that was the reason behind it was selected for our work.

The sensor was attached to a MacBook Air (11 inches version) with 4GB RAM, 128GB SSD drive and a Core i5 1.4GHz processor. Despite being a computer with more than five years, its battery still keeps it running for about 2 hours, enough for data capture. External drives were also used to free Mac-

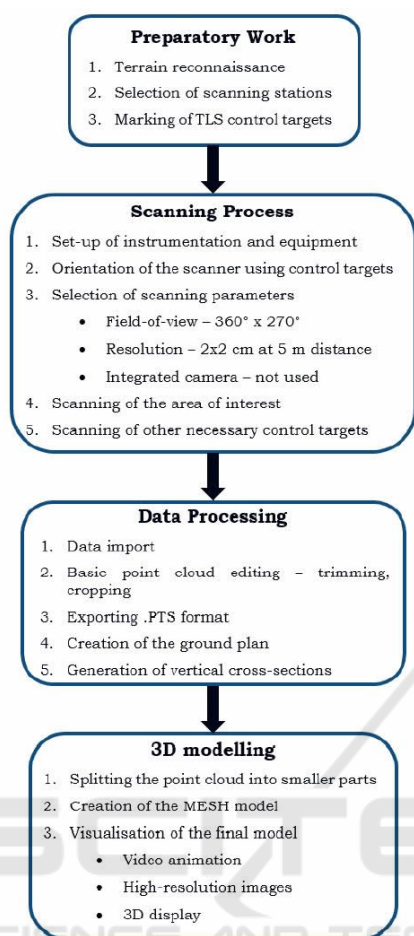


Figure 3: General workflow diagram from the work of Tometzova et al. (Tometzová et al., 2020).

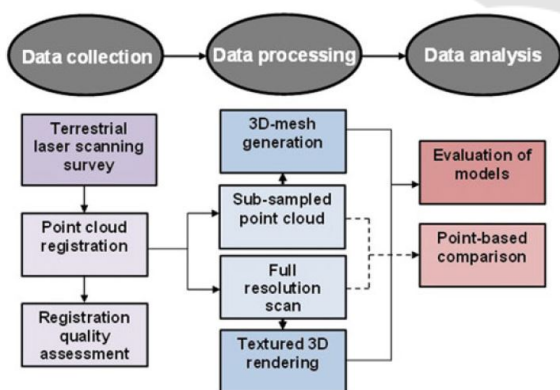


Figure 4: Methodological workflow diagram from the work of Idrees and Pradhan (Idrees and Pradhan, 2017b).

Book’s storage space. In the future, we intend to use a more portable and dedicated solution such as a Raspberry Pi 4.

For performing both capture and processing of the data, the Real-Time Appearance-Based Mapping

(RTAB-Map<sup>2</sup>) was used. It is an RGB-D, Stereo and LIDAR Graph-Based SLAM software based on an incremental appearance-based loop closure detector. RTAB-Map being a loop-closure approach with memory management as its core, is independent of the odometry approach used, meaning that it can be fed with visual odometry, LIDAR odometry, or even just wheel odometry. This means that RTAB-Map can be used to implement either a visual SLAM approach, a LIDAR SLAM approach, or a mix of both, making it possible to reconstruct a 3D point cloud based on RGB-D data.

Using the aforementioned hardware/software combination, three biologists went to field-test the solution in a real cave (Carrapateira), located at the municipality of Felipe Guerra, in Rio Grande do Norte state, northeastern Brazil. They started scanning the cave’s entrance with RTAB-Map’s main application running, capturing the data and viewing in real time if image registration was working. The entire cave was scanned in parts and the output for each capture session corresponded to a .db file, which is an SQLite file format that stores RGB and depth information, and will be later used in the 3D reconstruction process.

### 3.2 3D Reconstruction

The RTAB-Map’s 3D reconstruction pipeline can be fed with real time data from sensors, or with image frames of a previously recorded database, which was used in this work. The default options were used and the only modification was the feature detector/descriptor chosen, set to SIFT (Lowe, 2004). In order to perform the reconstruction process, a more powerful computer was used. Even using a computer with a Core i7 CPU 2.9GHz processor, with 32GB of RAM memory, 2TB SSD storage and an NVIDIA GTX 1080 GPU, the entire reconstruction takes more than 15 minutes to be completed.

Since the cave data was captured by parts, i.e., there were different .db files corresponding to the same cave, it was possible to split the reconstruction process, processing each file at a time. RTAB-Map is capable of receiving all .db files at once and process all of them together. In fact, we tried that option, but the reconstruction took too long and the software was not able to correctly register the different point clouds captured. Therefore, we opted to generate a point cloud for each .db file captured and later perform a manual registration of all point clouds.

<sup>2</sup><http://introlab.github.io/rtabmap>

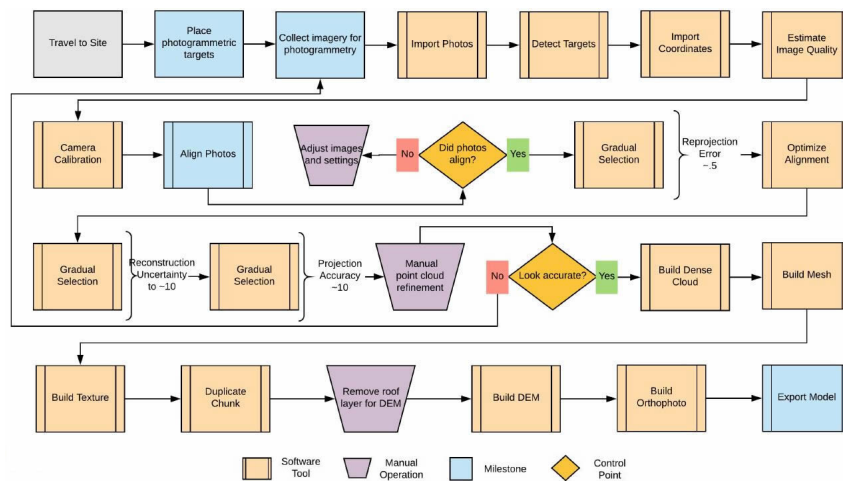


Figure 5: Protogrammetry workflow from the work of Lee (Lee, 2018).

### 3.3 Data Extraction/Visualization

An important part of the full 3D reconstruction of a cave lies in the combination of its point clouds. If there is more than one point cloud to process, a registration operation is necessary, i.e., an alignment between different points clouds will be necessary and special attention must be paid to precisely overlap intersections. This alignment can be done manually or automatically, and this is a computationally demanding process since its goal is to find a 3D to 3D transform that brings the points from a point cloud coordinate system to a different one, in a coherent way.

After obtaining the isolated point clouds, we used Meshlab<sup>3</sup> for the manual point cloud registration. This process was done in the presence of the persons that captured the original data, so they could solve questions about the cave shape. Figure 6 illustrates the process of manually registering two point clouds using Meshlab. While one point cloud remains still, the second one is rotated/moved to match the previous one. Since we use an RGB-D sensor to capture data, and their reference scale is the same, there is no need to change scale for all points clouds. Sometimes, photos taken inside the cave were used to make sure the point clouds being registered was correct. An example of such photos can be visualized in Figure 7. In future situations, using some landmarks (e.g. a solid object set at the cave floor or walls) when capturing data would make this matching process easier and more precise. After performing registration for all generated point clouds, we were able to see the entire path reconstructed, as shown in Figure 8.

After generating the final point cloud based on the combination of the partial ones, data visualization

is a concern. Considering that in the case of caves points clouds will be usually dense (made of millions of points close to each other), optimized approaches must be used to avoid wasting the available computational resources. Idrees and Pradhan (Idrees and Pradhan, 2017b) describe different possible visualizations for data captured.

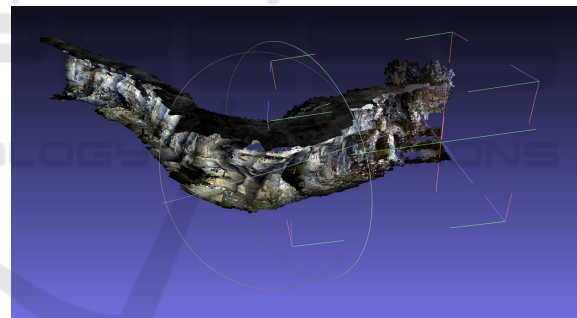


Figure 6: Example of two different point clouds being manually registered using Meshlab.

## 4 DISCUSSION

3D mapping has been a valuable support tool in several areas but just recently it became accessible for speleologists and other cave-related professionals. 3D mapping may help those professionals to acquire reliable information on the cave's physical structure, and even on biological features, like deposits of guano or areas with higher animal concentration. However, this is a relatively new technology and many cave professionals are not familiar with, so capturing such data in field conditions may not be a simple task. The major difficulties may be associated to the correct equipment manipulation, good quality illumination

<sup>3</sup><https://www.meshlab.net>



Figure 7: Photo taken by biologist to illustrate an interior passage of the cave.

and even to how well-trained are the professionals to capture data. All these characteristics are fundamental for an adequate data registration. Initial patience and training are necessary.

The equipment used (Macbook + Xtion sensor) showed to be very sensible to rapid movements and to abrupt direction changes during the scanning process. In situations like that, the software used for capturing the data (RTAB-Map) frequently freezes due to imprecise tentatives to find consecutive matched frames. When the baseline is too large, the frames are different enough to result in less matches, and consequently the slam algorithm used in RTAB-Map failed. Moreover, there are difficult parts of the cave to access due to their irregularity, making it harder for the team of biologists to reach them.

The need for regular illumination and the use of the sensor coupled to the laptop demanded more than one professional for capturing the data. In our approach, three professionals fully equipped with lanterns and reflectors were required to work synchronously so that the data capture could be successful. In order to shorten the wide baseline between frames, the field team started moving as slow as possible. One suggestion was to adapt the equipment to a stabilization system to minimize the impact caused

by the terrain irregularity. The biologists were capable of using the equipment (Macbook + Xtion), but they were just testing it. Refinements will be necessary to improve both usability and ergonomics, in case a better precision is needed.

Considering a laptop is used for visualization of the captured data in real time, the cable connecting the laptop to the RGB-D sensor has to be long enough so the field team can keep distance while capturing data inside the cave. A shorter cable would make the capturing process less continuous and unstable, contributing to the interruption of the process.

The proposed approach may generate different database files for the same physical space (with much data intersection). While this is good for registering the point clouds, it is advisable to use some kind of marker on the cave in order to facilitate further manual registration of the captured point clouds. In this work, the capture was performed without the use of such visual markers, and it made the manual registration process take longer than expected.

Battery availability for the notebook used in field work may also be an issue. A good battery autonomy may allow a single shot mapping, depending on the extension of the cave and the objective of the field team. Since our field team was testing the solution and wanted to evaluate the quality of the captured data, they had no problem in this sense. However, longer and deeper caves will certainly require multiple scanning sessions or frequent battery replacements.

During our field test, the system was freezing after some minutes of use due to memory restrictions of the notebook used. We suspect that this happens due to RTAB-Map's high memory consumption during the capture. Our field team recommend that caves with larger spaces should be scanned in different parts so that the point clouds could be joined later. The longer is the capture time, the higher is the possibility of the software to freeze and to lose the captured data so far. Therefore, the field team must previously analyze which parts of the cave will be recorded before moving to a new section.

In summary, the equipment proved to be satisfactory in the sense of allowing the acquisition of data needed, but there is still room for improvements in both hardware and software.

In the software side, we are currently working on porting RTAB-Map to the Android platform while adapting it to work with Intel Realsense RGB-D sensors. This way, any Android phone could work as a portable capture/reconstruction station. There is a RTAB-Map version for iOS devices that have LiDAR sensors, and we intend to do the same with Android

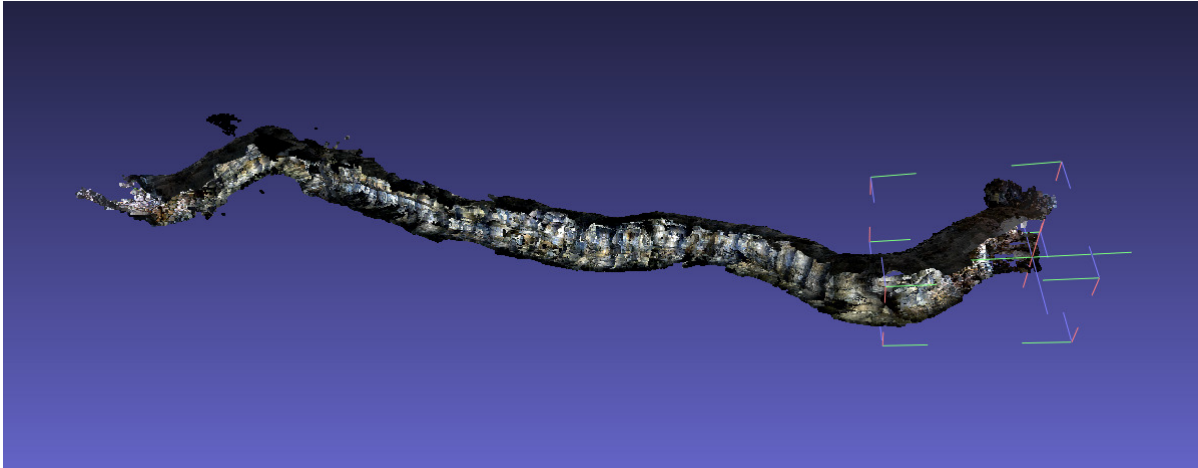


Figure 8: Final reconstructed point cloud obtained from manual registration using Meshlab.

devices that use external RGB-D sensors. We also intend to do some comparison of the obtained results with the 3D reconstruction that comes as native output of an iPhone 14 pro max, since it already has a LiDAR sensor and provides 3D reconstruction capabilities natively. We have performed some experiments on controlled environments with the iPhone and the results were quite impressive, but tests in the real cave scenario are still being planned. Efforts regarding improving the reconstruction processes are also on the go. We are performing some experimentation with Neural Radiance Fields (NeRFs), which seems to be the state of the art in 3D reconstruction at the time this paper was written. Even for NeRFs, the cave environment may be challenging due to lack of or irregular illumination and should be carefully investigated.

## 5 CONCLUSION

This work proposed a low-cost solution for 3D reconstruction of caves, by means of the use of RGB-D sensors. Our approach is an interesting alternative to the well-known and commonly used laser scanning technologies, which are prohibitively expensive and hard to acquire in Brazil. The proposed approach should allow cave users to scan parts or even complete cave structures, with emphasis on smaller parts like specific speleothems or details. The used sensor showed enough resolution to enable a full 3D reconstruction based on the combination of a sequence of captured RGB-D images. Our proposed solution could enable, for instance, the non-destructive documentation of cave shapes, forms and structures with the 3D models being manipulated from different points of view in the computer. Besides, the 3D maps produced may

complement, correct and refine already existing topographical and cartographic mappings or new ones in unmapped caves.

Complementary, information such as volume and area calculations from the final cave 3D point cloud may be used for additional research and environmental licensing. As future directions, we intend to develop a miniaturized hardware platform based on a Raspberry Pi 4 so it would be possible to have higher mobility while capturing data inside the caves. This would solve some of the problems detected by us in the field. We also intend to improve the capture process by using a customized data recorder tool from RTAB-Map. This would allow less processing while capturing data, and consequently less power consumption.

Finally, we intend to improve the 3D reconstruction process so that less human intervention is necessary to record the point clouds. We plan to develop an optimized solution to allow cave visualization via web browser, using for example the Three.js<sup>4</sup> library. However, this will be challenging due to the high number of points present in final caves reconstructions. Optimizations will be necessary in order to allow real time navigation on the browser inside the virtual caves. Comparison with existing methods regarding performance is also planned in a near future.

## ACKNOWLEDGEMENTS

The authors would like to thank CECAV/ICMBio for financially supporting project “Modelagem 3D de cavidades naturais subterrâneas”.

<sup>4</sup><https://threejs.org/>

## REFERENCES

- Aiello, D., Basso, A., Spena, M., D'Agostino, G., Montedoro, U., Galizia, M., Grasso, R., and Santagati, C. (2019). The virtual batcave: A project for the safeguard of a unesco whl fragile ecosystem. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*.
- Azpúrua, H., Saboia, M., Freitas, G. M., Clark, L., Aghamohammadi, A.-a., Pessin, G., Campos, M. F., and Macharet, D. G. (2023). A survey on the autonomous exploration of confined subterranean spaces: Perspectives from real-world and industrial robotic deployments. *Robotics and Autonomous Systems*, 160:104304.
- Beraldin, J.-A., Blais, F., Cournoyer, L., Picard, M., Gamache, D., Valzano, V., Bandiera, A., Gorgoglione, M. A., et al. (2006). Multi-resolution digital 3d imaging system applied to the recording of grotto sites: the case of the grotta dei cervi. In *VAST*, pages 45–52.
- Escalera, S. (2012). Human behavior analysis from depth maps. In *International Conference on Articulated Motion and Deformable Objects*, pages 282–292. Springer.
- Gallay, M., Hochmuth, Z., Kaňuk, J., and Hofierka, J. (2016). Geomorphometric analysis of cave ceiling channels mapped with 3-d terrestrial laser scanning. *Hydrology and Earth System Sciences*, 20(5):1827.
- Gallay, M., Kaňuk, J., Hochmuth, Z., Meneely, J. D., Hofierka, J., and Sedlák, V. (2015). Large-scale and high-resolution 3-d cave mapping by terrestrial laser scanning: a case study of the domica cave, slovakia. *International Journal of Speleology*, 44(3):6.
- Grohmann, C.H., Z.-G. G. B. F. A. F. A. I. (2019). Mapeamento e modelagem 3d do setor turístico da caverna do diabo (sp) com laser scanner: resultados preliminares. In *Anais do 35º Congresso Brasileiro de Espeleologia*.
- Idrees, M. O. and Pradhan, B. (2017a). Characterization of macro-and micro-geomorphology of cave channel from high-resolution 3d laser scanning survey: case study of gomantong cave in sabah, malaysia. In *Cave Investigation*. IntechOpen.
- Idrees, M. O. and Pradhan, B. (2017b). Frontier in three-dimensional cave reconstruction—3d meshing versus textured rendering. In *Global Civil Engineering Conference*, pages 1029–1038. Springer.
- 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.
- Li, Y., Zhao, L., Chen, Y., Zhang, N., Fan, H., and Zhang, Z. (2023). 3d lidar and multi-technology collaboration for preservation of built heritage in china: A review. *International Journal of Applied Earth Observation and Geoinformation*, 116:103156.
- Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110.
- McFarlane, D., Buchroithner, M., Lundberg, J., Petters, C., Roberts, W., and Van Rentergen, G. (2013). Integrated three-dimensional laser scanning and autonomous drone surface-photogrammetry at gomantong caves, sabah, malaysia. In *Proceedings of the 16th International Congress of Speleology, Brno*, volume 1, pages 317–319.
- Mohammed Oludare, I. and 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.
- Oguchi, T., Yuichi, S. H., and Wasklewicz, T. (2011). Data sources. In *Developments in Earth Surface Processes*, volume 15, pages 189–224. Elsevier.
- Pukanská, K., Bartoš, K., Sabová, J., and Tometzová, D. (2018). High-resolution 3-d mapping for the survey of valuable inaccessible medvedia cave in the national park of slovenský raj. In *Advances and Trends in Geodesy, Cartography and Geoinformatics: Proceedings of the 10th International Scientific and Professional Conference on Geodesy, Cartography and Geoinformatics (GCG 2017), October 10-13, 2017, Demänovská Dolina, Low Tatras, Slovakia*, page 89. CRC Press.
- Sellers, W. and Chamberlain, A. (1998). Ultrasonic cave mapping. *Journal of archaeological science*, 25(9):867–873.
- Tometzová, D., Hlavňová, B., Drevko, S., Pukanská, K., Bartoš, K., Bieda, A., and Kseňák, L. (2020). The development of geotourism with the use of 3d digital cave model. a case study of the bear cave, slovakia. *Geomatics and Environmental Engineering*, 14(2).
- Ullman, M., Laugomer, B., Shicht, I., Langford, B., Ya'aran, S., Wachtel, I., Frumkin, A., and Davidovich, U. (2023). Formation processes and spatial patterning in a late prehistoric complex cave in northern israel informed by slam-based lidar. *Journal of Archaeological Science: Reports*, 47:103745.