

3D Vision: Evaluation and Documentation

Saip-Can Hasbay, 01428723
saipcanhasbay@gmail.com



Fig. 1. Two photographs from two angles of the scanned bunny object.

Abstract—In this report, we discuss our work for the 3D Vision course (Spring 2024) at the Technical University of Vienna. We summarize the 3D reconstruction of a bunny model obtained through a Structured Light Scanner (SLS) and from Structure from Motion (SfM). Alongside discussing our workflow for 3D reconstruction, we document, and compare our measurements obtained from the real world, SfM and SLS.

I. STRUCTURED LIGHT SCANNER

We begin by briefly discussing our workflow for the 3D reconstruction of a bunny object (Fig. 1) using a Structured Light Scanner (SLS).

Background. An SLS emits patterns such as stripes onto the scanned object (Fig. 2). The beams of patterns are distorted and reflected by the scanned object’s surface. The SLS captures these reflected distortions for multiple frames. The software of the SLS analyzes the captured information and tries to reconstruct the 3D object accurately. For further literature on SLS, we refer the reader to [Fofi et al., 2004; McMillion, 2022].

A. Scan Session

The scan session was held on the 22nd of March 2024 and took approximately one hour and thirty minutes encompassing both group members. We used a [Hexagon, 2024] AICON PRIMESCAN SLS and the corresponding Optocat software.

We began the scan session with the highly specular banana object (Fig. 2). Note that this object belongs to our group member, and a detailed analysis can be found in their report.

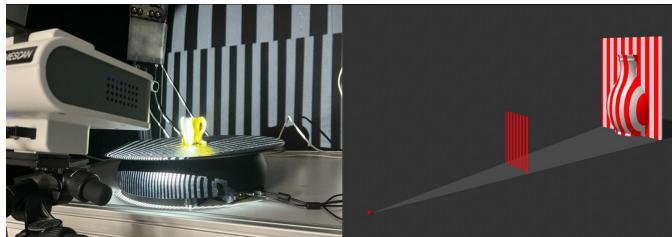


Fig. 2. An SLS projects patterns and captures the distortions. Left: Taken from the scan session by Maggie Wang. Right: Taken from [McMillion, 2022].

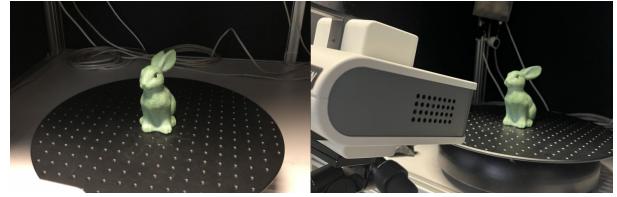


Fig. 3. Bunny object throughout the SLS scan session from two angles.

First, the object was held vertically and scanned according to the instructions of our tutor (Sonja Klement, and [Robert and Marvin, 2024]). Due to the object’s specular nature, some of the regions of the object were not obtained accurately. Moreover, the complexity of the shape made the measurements and the alignment of multiple scans difficult. We used six scans per rotation throughout the session, but larger numbers were also experimented. To obtain a full 3D reconstruction, the object was also rotated and scanned horizontally. Overall, more than thirty scans were necessary.

We then continued with the scanning of the bunny object (Fig. 3). Due to some previous hardships [Hasbay and Hahn, 2023], a very diffuse material object was selected for an efficient scan session. However, the complexity of the shape still provided enough difficulty. As previously, we first scanned the bunny object vertically. The initial results were highly successful. However, the ears of the bunny caused some occlusion and shadows (Fig. 4), which caused some holes in the scans.

To address the hole artifacts, we placed the object horizontally where the face was looking at the scanner. However, due to the larger width of the torso of the bunny object—compared to its head width—the horizontal placement made the object unstable. While we scan the object, the turntable rotates it. Therefore an unstable placement would not only move the object but also block one side of its face. Consequently, we used black playdough to stabilize the object from one side of its face. However, the playdough occluded one side of the face. Therefore, we scanned the object one more time to capture the occluded regions (i.e., reflection on the y-axis).

In total we scanned the object in three parts: vertically, and two times horizontally. For each part, we only used six scans per rotation. As a result, we finished the scanning session in total with eighteen scans (i.e., eighteen scan files in Breuckman file format, BRE [Filext, 2024]). Overall, we obtained relatively acceptable results (Fig. 4). However, we still needed to do some post-processing to address the inaccuracies and tiny holes around the object.

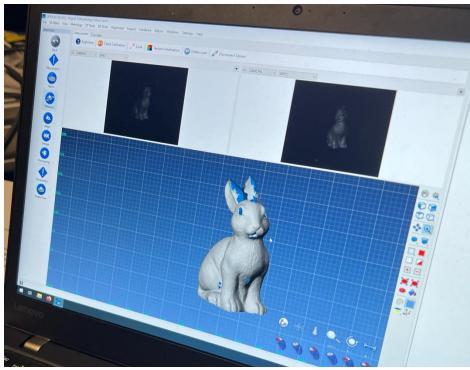


Fig. 4. Initial results of the scan on the Optocat software. Observe the holes around the ears of the bunny object.

B. Post-processing

The post-processing was done on the 17th of May 2024 and took approximately three hours and thirty minutes. We used Geomagic Wrap [Oqton, 2024] for post-processing. Note that Meshlab [Cignoni et al., 2008] also allows the import of BRE files. However, to our surprise, when we opened the scan files both on Geomagic Wrap and Meshlab, the scans were **not aligned** (Fig. 5) as we previously saw in the scan sessions (Fig. 4).

As a consequence, we had to re-register each scan. Our first attempt was using the *Global Registration* feature to align the scans. However, our attempt failed as the optimization did not converge. Hence, we had to manually align each scan using the *Manual Registration* feature. For each scan, we performed the following. (1) Pick two scans that portray two different views. (2) Manually pick at least three common points from step 1. (3) Apply manual registration. (4) Perform global registration on the result of step 3 to further optimize the alignment. The final result looked more like the originally aligned object (Fig. 5). Unfortunately, unlike our results from *photogrammetry* (Sec. II), the colors of the bunny object were lost as they were not part of the scan files. To remedy that, one can utilize physics-based differentiable rendering [Hasbay and Hahn, 2023]. Note material estimation is beyond the scope of this report.

After obtaining the fully aligned bunny object in a point cloud, we continued with further post-processing steps. As in many 3D reconstruction works, the result contained noise and artifacts (e.g., from the turntable). We used the lasso tool to remove redundant points. Next, we used the *Merge Scans*

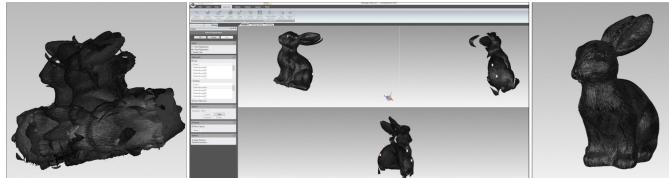


Fig. 5. After opening scan files the bunny object was not aligned (left). Hence, we manually aligned each scan (middle). Result of alignment (right).

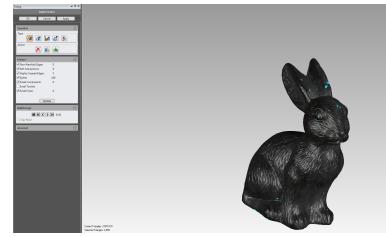


Fig. 6. Merging point clouds and meshing resulted in a mesh that contained holes and spikes.

feature to combine all the points from the aligned scans. The merge feature performs noise reduction as well as deleting small components and limiting resulting triangle count. Note that we used the default settings. This step resulted in an imperfect polygon mesh (Fig. 6).

To fix the holes on the resulting polygon, we used the *Fill All* feature. Next, we also used the *Mesh Doctor* feature to fix the highly creased edges and spikes. To get a higher quality mesh, we used the *Remesh* feature with default settings. The resulting mesh consisted of over 2 million triangles. As this is not practical in many use cases, we duplicated the mesh and applied the *Decimate* feature targeting 50 thousand triangles. We remeshed the low-resolution mesh to improve the triangulation quality. Finally, we used the *Mesh Doctor* feature to ensure that both high-resolution and low-resolution meshes had no holes or spikes. In total, we obtained highly accurate 3D reconstruction of the bunny object in two versions (Fig. 7): high-resolution (over 2 million triangles) and low-resolution (around 50k triangles).

Our results indicate that the SLS approach is highly accurate (Sec. III). However, we observed multiple drawbacks. First, it requires a laboratory environment which makes the SLS approach impractical for a variety of fields (e.g., 3D reconstruction in archeology or architecture). This also indicates that the SLS approach is more suited for industry professionals rather than novice end-users. Second, it requires an expensive machine as well as expensive software. With all this equipment, we still observed that the SLS struggled with specular materials. Moreover, we also had to manually align scans. In practice, this is undesirable and preferably automated. We will see that some of these drawbacks are less of an issue in the photogrammetry approach. Next, we divert our attention to the reconstruction of the bunny object through SfM.

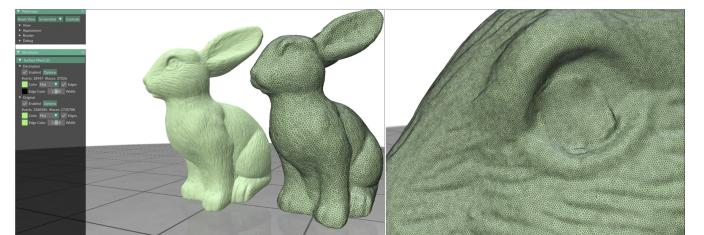


Fig. 7. The high-resolution and low-resolution 3D reconstruction of the bunny object using SLS (left). Close-up of high-resolution bunny's eye (right).



Fig. 8. The bunny and banana objects from different views throughout the photogrammetry session.

II. STRUCTURE FROM MOTION

In this section, we discuss the 3D reconstruction of the bunny object using *Structure from Motion* (SfM). We begin with a brief introduction to SfM. Then, we discuss our workflow and the practical aspects of photogrammetry (Sec. II-A). Finally, we delve into the post-processing steps (Sec. II-B) using Metashape [AgiSoft, 2022].

Background. Photogrammetry is a technique to acquire measurements usually taken from stereoscopic photographic images [CICES, 2024]. In other words, it is the science of making measurements from photographs. *Structure from Motion* (SfM) is a subset of photogrammetry. This technique utilizes corresponding image points in multiple views [Cipolla and Torr, 2008]. In SfM, 3D points and projection matrices are computed simultaneously using corresponding points in each view.

As a consequence, we also used multiple views (or, photographs) to obtain the 3D reconstruction of the bunny object. For additional background details, we refer the reader to [Cipolla and Torr, 2008; MathWorks, 2024]. Now, we divert our attention to the practical aspects of SfM.

A. Photogrammetry Session

The photogrammetry session was held on the 12th of April 2024 and took approximately one hour encompassing both group members. We used a Canon EOS 800D DSLR camera and a Canon 18-55mm lens (Fig. 8). To further ease the process, we also used the Benro Slim CF tripod, Revopoint portable turntable, and Rollei remote release. Note also, *unlike* the bunny object, we photographed the banana object inside of a Puluz lightbox. This was due to the specular nature of the banana object and to cover it from direct sunshine. Note that the details about the banana object can be found in our group members' report.

We started the photogrammetry session with the bunny object. Due to its diffuse material, we did not use the lightbox while photographing the object. As light rays scatter diffusely



Fig. 9. Bunny from different angles during the photogrammetry session.

from the object, it makes the object uniformly lit. Similarly, photographing the object in an open environment helped avoid shadows due to uniform and global illumination. Nevertheless, we could not completely avoid shadows on the object (Fig. 9). A thin layer of cloud would further circumvent shadows.

Throughout the session, the camera remained *static* with the following settings. We set our camera to *Aperture Priority* (Av) as well as in *auto-focus* mode. The *focal length* was fixed at 18mm. To get an overall good focus of the bunny, we picked a relatively high *F-Stop* value (f/16). In the Av mode, the camera automatically chooses an adequate *Exposure Time* to the corresponding F-Stop value. For all of our photographs, the exposure time ranged between 1/60 to 1/100 seconds. Due to overall good lighting conditions, we did not have to increase how much our camera let light in. Hence, we left the *ISO* setting fixed at 100. Lastly, all of our photographs are taken with the 6024x4020 *resolution*.

The remote release allows one to photograph without manual intervention (i.e., automated photography, Fig. 8). Hence, we had to pick some settings for our remote control. We set 5 seconds *delay time* to start to photograph. This feature allows some time of delay after activating the remote control. Similarly, we set 1 second of *interval time* in between each photograph. Note that the exposure time was overridden by the camera's Av mode. More importantly, the *number of photographs* to take was set to 60.

Similar to the SLS scan session (Sec. I-A), we started the session where the object was held vertically (Fig. 9). While the object rotated on the turntable, we activated our remote control to photograph the object. To capture the object robustly and have enough resources in the post-processing, we activated the remote control 4 times. Hence, we have taken 240 photographs of the object vertically. Next, we placed the object horizontally. Using a similar approach, we have taken 127 photographs of the object horizontally. Note that unlike in the SLS scan session, we have not further stabilized the object. But the *speed of the turntable*—default and fix—was slow enough and did not require additional stabilization. Therefore, we did not have to flip the object and take additional photographs. Note that overall, we could have used fewer photographs. Ultimately, the speed of the process (367 photos in approximately 240 seconds) and the need for robustness justify the excessive number of photographs. Having discussed the practical aspects, we now move on to the post-processing steps.

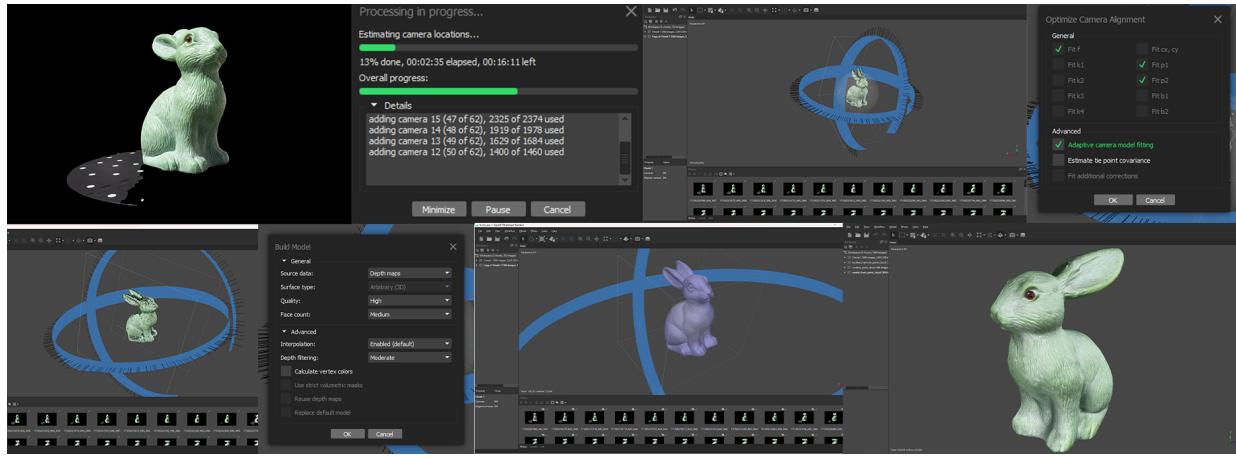


Fig. 10. (From left to right) Our input is mostly background removed and lens-corrected images. Next, we aligned photographs and obtained the initial camera positions and point cloud. We then optimized cameras and removed unnecessary points using gradual selection. Finally, we build a model and its texture to obtain the final reconstructed bunny object.

B. Post-processing

We continue with the post-processing steps using Metashape [AgiSoft, 2022]. The post-processing steps are summarized in (Fig. 10). We followed an introductory tutorial [Mallison, 2015] to get a successful 3D reconstruction of the bunny object.

Before we aligned all 367 photographs in Metashape, we did the following pre-post-processing steps. First, we uploaded the photographs—in RAW, or CR2 format—into Lightroom Classic CC [Adobe Inc., 2024a]. Next, we selected a photograph, and under *Develop*, *Lens Corrections*, we enabled *Remove Chromatic Aberration* and *Enable Profile Corrections*. Note that Lightroom automatically detected our lens profile. Hence, we did not have to change the lens profile manually. Next, we selected all photographs and used the *Sync...* feature to apply the lens correction settings to all images. Then, we exported all the images in JPEG format.

Our last pre-post-processing step was *background removal*. We used a script in Photoshop [Adobe Inc., 2024b] to remove the background as much as possible from all lens-corrected photographs. Consequently, our input to Metashape was 366 lens-corrected and background-removed photographs (Fig. 10).

We now briefly summarize our workflow in Metashape. The reader should refer to [Mallison, 2015] for a more in-depth discussion. We started by loading all of our photographs into Metashape in *one* chunk. Alternatively, one could create separate chunks for each different position (e.g., horizontal, or vertical placement). Next, under the *Workflow* tab, we used the *Align Photos* feature with the following settings (Fig. 10). This step ended successfully, and we obtained 366 different views surrounding the object (Fig. 10).

Next, we optimized the alignment using the *Gradual Selection* under the *Model* tab. We selected 10 as the *Reconstruction uncertainty* and removed all the tie points that had a low probability of being correct (Fig. 10). This reduced the unnecessary

points surrounding the object (e.g., points from the turntable). Then, under the *Tools* we used the *Optimize Cameras* feature with these settings (Fig. 10). After camera optimization, we once again used the *Gradual Selection* feature. This time, we used it by selecting 0.64 as the *Reprojection error*, 10 as the *Projection accuracy*, and deleted all the redundant points. These steps resulted in a relatively clean and well-aligned point cloud (Fig. 10).

Subsequently, under the *Workflow* tab, we used the *Build Model* feature using *Depth Maps* (Fig. 10). We also experimented using *Masks* and building from *point clouds*. However, in our experience building from depth maps resulted in better models.

Finally, as in the SLS post-processing (Sec. I-B), we improved the mesh and obtained it in different resolutions. First, under the *Tools*, *Model*, we used the *Close Holes* feature. Next, we performed *Build Texture* feature under the *Workflow* tab. This step was the last action for our ultra-high model. As ultra-high models are not usable in practice, we duplicated the model. Before we *decimated* the model we used the *Smooth Model* feature. In total, we obtained three textured models of the bunny in three different qualities (Fig. 11): medium (\approx 50 thousand triangles), high (\approx 1 million triangles), and ultra-high (\approx 22 million triangles).

Having discussed both SLS and SfM techniques, we now move on to evaluation and comparisons.

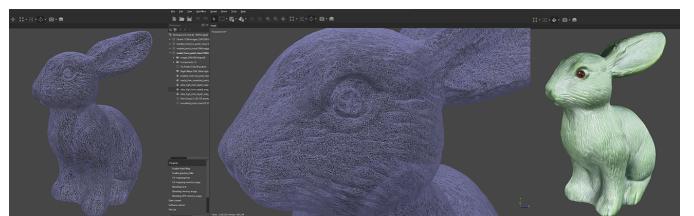


Fig. 11. The reconstructed bunny in medium (left, \approx 50k triangles), high quality (middle, \approx 1 mil. triangles), and textured in medium quality (right).

TABLE I
MEASUREMENTS OF THE 3D RECONSTRUCTED BUNNY OBJECT

Method	Volume (m^3)	distance leg (m)	distance eye (m)	distance ear (m)	bounding box X (m)	bounding box Y (m)	bounding box Z (m)
SfM	0.045	0.047	0.011	0.058	[0.024, 0.034]	[0.020, 0.035]	[-0.079, -0.080]
Reference	0.00009	0.04	0.015	0.055	-	-	-
SLS	0.0004	0.042	0.011	0.056	[-0.046, 0.041]	[-0.070, 0.085]	[-0.105, 0.008]

III. EVALUATION AND COMPARISONS

In this section, we examine the measurements of the 3D reconstructed bunny obtained via SLS and SfM. The summary of our results can be found in Table I and (Fig. 12, 13, 14). All of our resources can be found at [Hasbay, 2024a].

The reference lengths are taken using a simple ruler. Specifically, we measured the eye width, leg width, and ear length of the bunny object. As our ruler is not fully suited for this use case our reference measurements may contain small errors. We also measured the volume of the object using water displacement [wikiHow Contributors, 2024] (Fig. 14git). We see that our results from SfM and SLS do not closely match the reference volume measurements. This indicates a limitation of our work and necessitates more precise measurements in the future. However, for the task at hand, we believe our results are adequate for making comparison measurements against SfM and SLS.

For the SLS, all the measurements are done in [Oqton, 2024] using the Windows client in the PongLab [Technische Universität Wien, 2024]. We obtained our measurements using the *Distance* or *Compute* feature under the *Analysis* tab. The SLS approach is known to be robust concerning making measurements. Hence, our results—except for the volume—are relatively close to reference measurements (Table I). The differences for the bounding box are expected due to the obtained coordinate system. Note also that due to the size and uniform appearance of the object, the reference measurements may also contain inaccuracies.

On the other hand, the measurements for the SfM are done in Metashape Professional edition [AgiSoft, 2022] using our PC (Win 11, 32GB RAM, 13th Gen Intel(R) Core(TM) i7-13700K, GPU(s) NVIDIA GeForce RTX 4080). Note that all previous steps concerning SfM were done using Metashape Standard edition. Unfortunately, during the SfM session, we did not own a scale bar to include in our photographs. Hence, we were not able to use the automatic scale bar detection feature. This unused feature automates the measurement process. Therefore the reconstructed object is scaled according to the reference length of the scale bars. Nevertheless, we were able to obtain measurements for all the reference parameters (Table I). To obtain the measurements, we created virtual scale bars [AgiSoft Support, 2023]. For three distances we measured (i.e., eye, leg, and ear width), we selected two points on the model. Then, on three photographs, we adjusted the positions for better precision of the virtual scale bar. To create the virtual scale bar, we selected each pair of points (e.g., point 5 and 6 in Fig. 12) and used the *Create Scale Bar* feature. Finally, we entered the reference distances for each virtual

scale bar and pressed the *update* button to apply the changes. For further details on how to use virtual scale bars, we refer the reader to [AgiSoft Support, 2023]. For additional obtained measurements, please see also [Hasbay, 2024b].

After applying the virtual scale bars, our model is updated relatively closely to the reference scale. Overall, for each measured parameter, we see a close resemblance to the reference measurements. However, we believe that we would have obtained even higher accuracy and efficiency in the process if we used a reference scale bar during the photogrammetry session. In general, the SLS approach provides better robustness for 3D reconstruction and measurements. However, we observed that the SfM approach not only worked better in terms of obtaining a model but was also more efficient than SLS. Overall, assuming one includes the buyout and operational cost of an SLS machine, we recommend the reader use the SfM approach, especially for smaller-scale projects, and noting the recent advancements in this field (e.g., NeRF [Mildenhall et al., 2020; Müller et al., 2022] and Gaussian splats [Kerbl et al., 2023]).

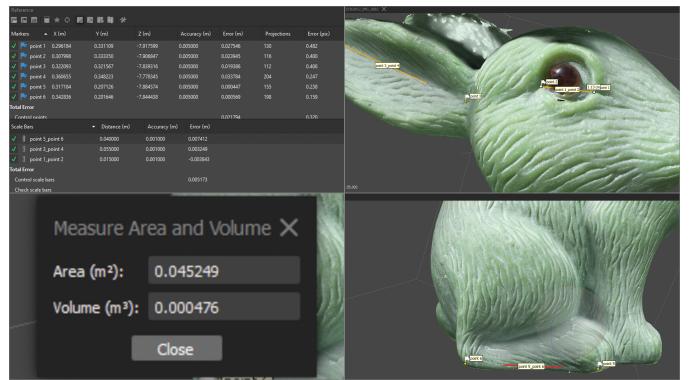


Fig. 12. Measurements obtained through SfM.

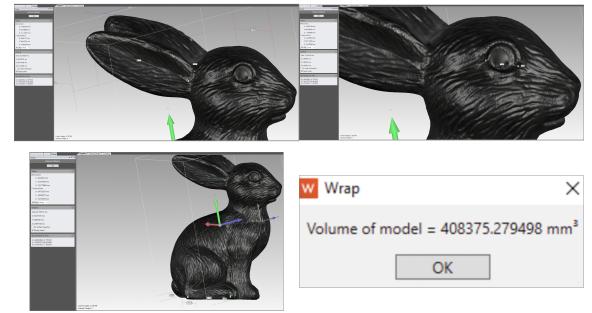


Fig. 13. Measurements obtained through SLS.

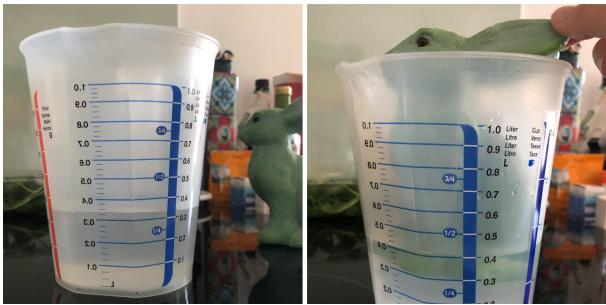


Fig. 14. We measured the volume of the object using water displacement.

REFERENCES

- Adobe Inc. (2024a). *Adobe lightroom cc*. <https://www.adobe.com/products/photoshop-lightroom.html>
- Adobe Inc. (2024b). *Adobe photoshop*. <https://www.adobe.com/products/photoshop.html>
- Agisoft. (2022). *Metashape* (Version 1.8.4) [<https://www.agisoft.com/>].
- Agisoft Support. (2023). *Creating scale bars in the project without coded targets* [Last Accessed 06.06.2024]. Retrieved June 6, 2024, from <https://www.agisoft.freshdesk.com/support/solutions/articles/31000162602-creating-scale-bars-in-the-project-without-coded-targets>
- CICES. (2024). Photogrammetry and Remote Sensing [Chartered Institution of Civil Engineering Surveyors, Last accessed: 28.05.2024]. <https://web.archive.org/web/20170830062535/https://www.cices.org/pdf/P%26RSinformation.pdf>
- Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F., & Ranzuglia, G. (2008). MeshLab: an Open-Source Mesh Processing Tool. In V. Scarano, R. D. Chiara, & U. Erra (Eds.), *Eurographics italian chapter conference*. The Eurographics Association. <https://doi.org/10.2312/LocalChapterEvents/ItalChap/ItalianChapConf2008/129-136>
- Cipolla, R., & Torr, P. H. (Eds.). (2008). *Multi-view geometry in computer vision* [ISBN: 9783642128486]. Springer.
- Filext. (2024). What is a BRE file? [Last Accessed 18.05.2024]. <https://filext.com/file-extension/BRE>
- Fofi, D., Sliwa, T., & Voisin, Y. (2004). A comparative survey on invisible structured light. *SPIE Electronic Imaging-Machine Vision Applications in Industrial Inspection XII, San José, USA, 5303, 90–97*. <https://doi.org/10.1117/12.525369>
- Hasbay, S.-C. (2024a). 3D Vision Resources [Accessed: 2024-06-07]. <https://github.com/sapo17/3DVision>
- Hasbay, S.-C. (2024b). SfM Metashape Report [Accessed: 2024-06-07]. <https://drive.google.com/file/d/1PDxWKSFlmTjP4f6E-SVmVCazAHHrHcqc/view?usp=sharing>
- Hasbay, S.-C., & Hahn, D. (2023). Translucent material parameter estimation. *Central European Seminar On Computer Graphics*. https://cescg.org/cescg_submission/translucent-material-parameter-estimation/
- Hexagon. (2024). Hexagon AICON PRIMESCAN Structured Light Scanner [Last Accessed 27.03.2024]. <https://hexagon.com/products/primescan>
- Kerbl, B., Kopanas, G., Leimkühler, T., & Drettakis, G. (2023). 3d gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics, 42(4)*. <https://repo-sam.inria.fr/fungraph/3d-gaussian-splatting/>
- Mallison, H. (2015, October). Photogrammetry tutorial 11: How to handle a project in agisoft metashape (photoscan) [EDIT June 3, 2022]. <https://dinosaurpalaeo.wordpress.com/2015/10/11/photogrammetry-tutorial-11-how-to-handle-a-project-in-agisoft-photoscan/>
- MathWorks. (2024). *What is structure from motion?* [<https://de.mathworks.com/help/vision/ug/what-is-structure-from-motion.html>]. MathWorks.
- McMillion, M. (2022, November). How does structured-light 3D scanning work? [Last Accessed 27.03.2024]. <https://www.artec3d.com/learning-center/structured-light-3d-scanning>
- Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., & Ng, R. (2020). Nerf: Representing scenes as neural radiance fields for view synthesis. *ECCV*.
- Müller, T., Evans, A., Schied, C., & Keller, A. (2022). Instant neural graphics primitives with a multiresolution hash encoding. *ACM Trans. Graph., 41(4)*, 102:1–102:15. <https://doi.org/10.1145/3528223.3530127>
- Oqton. (2024). *Geomagic Wrap* [<https://oqton.com/geomagic-wrap/>].
- Robert, S., & Marvin, B. (2024). TU Wien, 3D Scanning Instructions [Last Accessed 18.05.2024]. https://drive.google.com/file/d/1s9u5YJ0IP_7HJIJXD-RH8mWhn2yloJR/view?usp=sharing
- Technische Universität Wien. (2024). Ponglab at informatics innovation center (i2c) [Accessed: 2024-06-06]. <https://www.inflab.tuwien.ac.at/>
- wikiHow Contributors. (2024). How to calculate volume and density [Accessed: 2024-06-07]. <https://www.wikihow.com/Calculate-Volume-and-Density>