# Performance Analysis for 3D Reconstruction Objects in Meshroom and Agisoft—A Comparative Study

https://doi.org/10.3991/ijoe.v19i05.37257

Indrit Enesi<sup>(⊠)</sup>, Anduel Kuqi Department of Electronic and Telecommunication, Polytechnic University of Tirana, Tirana, Albania ienesi@fti.edu.al

Abstract—3D reconstruction of objects is with interest nowadays, mainly in production industry. The main challenge of them is accuracy and processing time, especially for small and detailed objects. The field of photogrammetry realizes the 3D reconstruction of objects through 2D photos. Different software, free or nonfree exists, providing different quality and performance. Accurate 3D reconstruction is important in cloning objects, especially in the industry of spare parts or in the production of prostheses in medicine, etc. Determining accurately the sizes of the object, especially those with complex geometric shapes is very important in the 3D printing process. The purpose of this paper is the analysis of the performance of 3D reconstruction in terms of accuracy for objects of different sizes regarding the number of its photos and the time evaluation of this process. The 3D reconstruction will be performed by free software Meshroom, measurement will be done in MeshLab and non-free software Agisoft. Experimental results show that quality and performance of 3D reconstruction depends on the number of photos of the object, concluding in finding the optimal balance between these parameters. By comparing obtained results from MeshRoom and Meshlag versus Agisoft, it is claimed that Agisoft performs better than MeshRoom. It offers more optimization techniques, reduces processing time, more visual quality in the reconstructed 3D object as well as more accuracy in measurement.

**Keywords**—photogrammetry, 3D reconstruction, Agisoft, Meshroom, MeshLab, measurement accuracy, time analysis

### 1 Introduction

Reconstruction allows us to gain insight into qualitative features of the object which cannot be deduced from a single plane of sight, such as volume and the object relative position to others in the scene. 3D reconstruction is used in different fields of industry as well as in medical imaging equipment [1]. Input data is used to create a 3D reconstruction of the original object or objects of the scene. A scan of a person's body, for example, can be used to create a 3D model of that person in a computer system. The aim of the paper consists in the analysis of the performance of software that provides 3D reconstruction from images taken with usual camera, a mobile phone camera or a professional one. 3D imaging is a technique which develops or creates the impression of

depth within an image by deploying 2D data into 3-dimensional format. To aid in quality regulating processes for industrial purposes, 3D imaging has become an extremely valuable factor. The software used in the paper for 3D reconstruction based on photogrammetry are Meshroom [2] combined with MeshLab and Agisoft. Meshroom is free, open-source 3D Reconstruction Software based on the Alice Vision framework. Alice Vision is a photogrammetric computer vision framework which provides 3D reconstruction and camera tracking algorithms. Meshroom is designed as a nodal engine [3], [4], the parameters can be changed easily. The 3D point cloud is then exported to MeshLab to measure different sizes of the reconstructed object. It is open source and offers various tools for editing or clearing the message as well as the option to specify the size of the message [5], [6].

Agisoft Metashape is a stand-alone software product that performs photogrammetric processing of digital images and generates 3D spatial data to be used in GIS applications, cultural heritage documentation, and visual effects production as well as for indirect measurements of objects of various scales [20]. The whole Agisoft system is designed to deliver industry-specific results relying on machine learning techniques for post-processing and analysis tasks. The software allows to process images from RGB or multispectral cameras, including multi-camera systems, into the high-value spatial information in the form of dense point clouds, textured polygonal models, georeferenced true orthomosaics. Further, post-processing enables to eliminate shadows and texture artifacts from the models, calculate vegetation indices and extract information for farming equipment action maps, automatically classify dense point clouds, etc. [21].

A set of photos of an object is used as input in the Meshroom and Agisoft. The quality and processing time for the reconstructed object is analyzed. The object reconstructed in Meshroom will be exported to Meshlab to measure the dimensions. Mesh-Lab is used for processing and editing messages.

The same set of photos will be used for 3D reconstruction in Agisoft. The visual quality, accuracy and time processing will be compared with the results obtained from Meshroom and MeshLab. Agisoft offers more optimization parameters and camera calibration then Meshroom as well as the calculation of object volume.

Section two describes the methodology used in the paper, section three describes the 3D reconstruction process, section four describes the experimental part and its analysis of results, concluding with the conclusions of the paper.

## 2 Methodology

#### 2.1 Photogrammetry

Photogrammetry is a technique that enables the creation of 3D models from photos taken from real objects in different positions, possibly keeping the object static [7], [8]. It works by extracting 2D data and superimposing them. Since objects are of different sizes, photogrammetry is used in various fields and applications such as topographic maps or points clouds [9]. The process of obtaining 3D models would be much more complex if we did not use modern software today. But the accuracy and visual quality remain an important problem in this field which continue to have the attention of the researchers.

Software considered here are Meshroom combined with MeshLab and Agisoft.

### 2.2 Meshroom software

The photo inputs are placed below the left-top part, the right-top part serves to display the output of the photo processing, it serves to display the final output, the 3D object created but here can also be displayed the result of special nodes when executed. The bottom part is the most important part, the graph editor shows the nodes that participate in the execution workflow. The nodal environment makes Meshroom software flexible, each node is performed individually and easily configured. The photos used in Meshroom may be taken through usual cameras. The individual tasks are represented by nodes combined into directed acyclic dependency graphs as pipelines [10]. Interface is shown in Figure 1.



Fig. 1. Interface of Meshroom

Nodes considered in the pipeline are: camera initialization, feature extraction, image matching, feature matching, structure from motion, depth map, depth map filter, meshing, mesh filtering and texturing.

**CameraInit** loads image metadata, sensor information and generates viewpoints. **Feature extraction** extracts features from the images as well as descriptors for those features [8]. **Image matching** is a processing step which figures out which images make sense to match each other. **Feature matching** finds the correspondences between the images using feature descriptors. **Structure from motion** will reconstruct 3D points cloud from the input images. **Depth map** retrieves the depth value of each pixel for all cameras that have been resolved by SFM (structure from motion). Certain depth maps will claim to see areas that are occluded by other depth maps. The **Depth map filter** step isolates these areas and forces depth consistency. **Meshing** generates mesh from sfm point cloud or depth map. **Mesh filtering** filter out unwanted elements of the mesh. **Texturing** projects the texture change quality and size/file type of texture [11].

#### 2.3 Meshlab software

Meshlab is an open-source systems for 3D image processing and prepares models for 3D printing. It works based on point clouds or in meshes. A set of tools are provided from Meshlab software such as rendering, meshes, texturing, measurement of distances, cleaning, healing etc. [12], [13], [14].

#### 2.4 Agisoft software

Agisoft Metashape is a tool for a photogrammetry pipeline. Its main steps are: Align Photos, Build Dense Cloud points, Build Mesh, Build Texture. Align photos is based on a computer vision concept called Structure from Motion. Dense Cloud points are a useful starting point for 3D modeling and can be helpful in positioning 3D objects into a scene. Mesh is the process of turning irregular shapes into more recognizable volumes called element. The texture option serves to create a texture based on an analysis of uploaded photos. The software can run on different platforms.

### **3 3D reconstruction process**

In the paper a set of photos of a real object will be considered, the 3D object reconstruction is performed as well as the dimensions of the object will be measured and the accuracy is evaluated.

3D reconstruction of object will be realized in Meshroom and Agisoft. Its reconstruction is very sensitive from the input images, usually a considerable overlap between images is suggested to have a better reconstruction. The reconstructed object is obtained as a scaled version of the real one. To obtain the real dimensions, a known size element is needed. Rescaling the reconstructed object using the known dimension, enables the acquisition of the object in real size enabling the measurement of its other dimensions [15], [16], [17], [18], [19].

#### 3.1 Size measurements

Photogrammetry system needs camera calibration, which is very important in size measurements. Usually, it is done through a well-known object or using a known size of the target object to scale the 3D reconstruction. Markers are defined and scale between them is set. To obtain optimal results, markers are defined in the three axes.

### 4 **Experimental analysis**

In [22] the number of photos is analyzed to obtain optimal 3D reconstruction in terms of accuracy and time processing in Meshroom. Five sets of photos are used, 20,40, 60, 80 and 100 photos, experimental results show that optimal performance is obtained for 60 photos. The same set will also be used in Agisoft, to enable the comparison of the results obtained from the two software. The comparison will be made for accuracy and time processing.

#### 4.1 3D object reconstruction based on default pipeline in Meshroom

Based on experimental results as illustrated in Figure 2, the quality for the target object is at a high level, but the processing time of the nodes for this set and of course the total time of the object generation process is quite high, specifically the total measured time was around 100 minutes. Details of the time processing are given in the summary Table 1. Dimensions in MeshLab: The actual dimensions of the object are width = 27.5 cm, length = 48.5 cm and height = 39.5 cm. To find the dimensions of the reconstructed 3D object the MeshLab software will be used.



Fig. 2. 3D reconstruction based on default pipline

Meshing generated by Meshroom software is exported to MeshLab as illustrated in Figure 3.



Fig. 3. Measuring the box dimensions

Regarding the measurements illustrated in Figure 6 the dimensions obtained by Meshlab are: width = 27.08 cm (x axis), length = 48.68 cm (y axis) and height = 39.726 cm (z axis). By results yields that an accurate precision is achieved, the errors are: for the width an error of only 0.08 cm, for length an error of only 0.18 cm, and for the height is 0.226 cm.

#### 4.2 3D object reconstruction based on default pipeline in Agisoft

In 3D reconstruction in Meshroom the optimal performance is obtained for the set with 60 photos, the same set that will be used in Agisoft for experimental analysis.

The added photos are from the single scene and all the image metadata are loaded in the system. 3D object reconstruction in Agisoft is performed based on four steps: Align photos, Dense Cloud points, Mesh and Texture.

Align photos (camera alignment). Parameters tuned here are Detecting points, Matching points, and Estimating camera locations.

**Tie points limit** is the maximum number of points that Agisoft will match between photos, default is 4000. Increasing this number leads to a high accuracy but needs hours of processing. **Key points** are points of interest i.e. high contrast or interesting texture on the images that can be easily recognized. The number of key points depends on the size of the images, the visual content.

The Workflow begins with Align photos, an illustration is shown in Figure 4.



Fig. 4. The result of align photos

Statistical details of the process are shown on the console. It consists in three parameters:

**Detecting photos (points detected)**. For the key points limits = 40000, and for the set with 60 photos, the processing time is 13.78 seconds. **Matching points**, results show 764192 match points for the execution time 61.5 seconds. Continuing with matching combine and the filtering process, the measured time is 3.4 seconds.

**Estimating camera locations**, processing time is 33.24 seconds which includes 25.52 seconds for sfm and the optimization time 7.72 seconds.

The total processing time for Align photos is 111.93 seconds.

The obtained object so far has many erroneous points (noisy points), so it is very important to filter these points before moving to Dense cloud.

**Tie points Filtering** is performed to have a clear object before dense cloud, meshing and texture process.

Filtering process is performed by: Reprojection error, Reconstruction uncertainty, Image count and Projection accuracy.

On every step, camera calibration is performed to increase the accuracy of the object.

**Build mesh process.** Performing build mesh, the default parameters need to be tuned, to increase visual quality of the object. Based on experimental results, values 88–90 reconstruction uncertainty have the optimal visual quality as shown in Figure 5.



Fig. 5. Selection of the optimal value

By removing the unnecessary points, the results is shown in Figure 6.



Fig. 6. The result obtained

Optimize Camera Calibration at references is performed again to reobtain parameters. Execution time of this step lasts for 2 seconds.

In the **Reprojection Error** the two columns show the relations between projections and error (pixel) for every photo, in our case 60 rows since the set has 60 photos as shown in Figure 7.

Reference	+ A III II	×凸	Error (°)	Projections	Error (pix)
				2433	0.804
Error (°)	Projections	Error (pix)		2271	0.871
	3043	0.960		2434	0.892
	2807	1.022		2171	0.998
	3045	1.003		1842	1.026
	2496	1.229		1998	1.156
	1956	1.177		2414	1.083
	2163	1.379		2042	0.902
	2787	1.221		3043	0.092
	3366	1.037		2161	0.813
	2445	0.940		2707	0.912
	3185	1.017 🚽	-		• • • • • • • • • • • • • • • • • • •
		- IIII	Markers	Longitude	Latitude A

Fig. 7. The corresponding relation between projections and error (pixel) for each photo

We noticed that by default we have a value selected around 0.94. But we aim to optimize it, so at least we should try to have the number 0.5 (so half a pixel). We choose to continue with the value 0.49 and the result after execution is given as in the Figure 8.



Fig. 8. Output for the selected value of reprojection error

The results show that the errors are decreased. **Camera calibration** is performed again. Processing time for this step is 2 seconds.

Next, the **projection accuracy** option is performed. Based on experimental results, the optimal parameters are obtained for 8–10% of points, as shown in Figure 9.



Fig. 9. The view before execution for the selected value of projection accuracy

About 6275 points are selected or around 10% of total points. Camera calibration is performed and the errors are decreased as in the Figure 10.



Fig. 10. The output for the selected value of projection accuracy

If we compare the useless points around our object in focus in Agisoft with those given to us by Meshroom Software for the same case, those erroneous have been significantly reduced, but again starting from the last results as in the Figure 11 that Agisoft gave us after the optimizations done, it is noticed that there are unwanted points around our object. But since that was the optimal performance for this case that Agisoft gave us, then for these points the software offers the possibility of manual deletion.

And at the end of this manual deletion process, the result is given as in the Figure 11.



Fig. 11. The result in Agisoft after manual deletion

By comparing the number of points at this moment with the number of points before manual deletion, it is noticed that there is a reduction of about 10,000 points, this shows that the need for optimization is still high. At this moment a final optimization of camera calibration is made.

**Building the dense point cloud.** Based on the estimated camera positions and pictures themselves a dense point cloud is built by software. Build dense cloud is performed in the workflow and the Parameter Quality is set to Hight, but Depth filtering is set to moderate.



The total processing time was 18 minutes and 20 seconds (1099.56 sec), while the output after activating the dense point cloud is shown in Figure 12.

Fig. 12. Output after activating dense point cloud

It is seen some unwanted parts around the object, a deep filtering process is needed using Filter Dense Cloud with the option **"filter by confidence"**.



Based on experimental results, points in the interval [0–5] will be deleted. The result is shown in Figure 13.

Fig. 13. The result obtained after deleting points using filters

As noticed, the number of useless points has decreased significantly, but again there is a relatively low percentage of them around the object in focus. The interval for good quality points can be increased from [5-255] to [6-255] or [7-255] and so on, but in this way the quality of the reconstructed object would start to decrease, therefore [5-255] is precisely the interval which gave us the most optimal result.



A manual cleaning is needed to remove the unwanted points, the result is shown in Figure 14.

Fig. 14. The result obtained after deleting the remaining points manually



Build mesh. Default parameters will be used. The output is shown in Figure 15.

Fig. 15. Result after execution

We notice that the object is visually very clear, while the processing time was 5 minutes and 34 seconds (334.107 sec).

At this moment, we have an object with clear contours, but often there are objects with irregular shapes, which often after this state (at the end of the Mesh), may look visually unclear.

Texture Mesh is needed.

Texture mesh. Default parameters will be used.



Fig. 16. The generated output

The execution time was 141.27 sec. It is the final output, it is visually very clear, and from the comparison with the Meshroom analysis, a higher level of quality is possible by Agisoft, both from the side of the contours of the reconstructed object as well as from the content part as shown in Figure 16.

The total execution time was about 1686.867 seconds. Comparing with the time of Meshroom reconstruction, which was for this dataset 5861.8 seconds, it yields that there is a reduction of about 4174.93 seconds or 69.6 minutes.

An object with much better quality in a much shorter time is obtained using Agisoft.

### 4.3 Measurements in Agisoft

Four markers are defined as shown in Figure 17 and their specification are shown in Figure 18.



Fig. 17. Placement of markers in specific x,y,z positions

Markers	Longitude	Latitude	Altitude (m)	Accuracy (m)
📄 🏲 point 1	19.821961	41.316127	0.076437	0.005000
🗸 🏲 point 2	19.821961	41.316127	-0.007842	0.005000
point 3 🏲	19.821961	41.316126	-0.021370	0.005000
🗸 🏲 point 4	19.821960	41.316127	0.029559	0.005000

Fig. 18. Specifications for each marker

Points1 and 2 on the y-axis, points2 and 3 on the x-axis and points 2 and 4 on the z-axis.

Then we group points1 and points2 as in the Figure 19.



Fig. 19. Way of grouping markers

The value that will be obtained from the measurement of the real object between these specified points needs to scale the reconstructed object. In our case, distance = 39.5 m serves for the first scale.

We grouped points 2 and 3 according to the same logic and as a distance between them we will set the value that will be obtained from the measurement of the real object between these two specified points. In our case, distance = 48.5 m. This is the second scale.

We grouped points 2 and 4 according to the same logic and as the distance between them we will set the value that will be obtained from the measurement of the real object between these two specified points. In our case, distance = 27.5 m. This is the third scale.

It is important to highlight that the distances marked in the software are in meters, but the calibration is actually measuring them in centimeters.

At this moment, the object is scaled, so we can perform measurements.

After calibrating the system, the three sizes of the target object are measured. Figure 20 show the results.



Fig. 20. Measuring the box dimensions

The height is 39.5 cm converted to the real dimensions of the object. That was also the real length of the object, so the error in this case is 0.

The width is 27.4 cm converted to the real dimensions of the object. While the real length of the object turned out to be 27.5 cm, therefore the error in this case is only 0.1 cm (1 mm).

The length is 48.6 cm converted to the real dimensions of the object. While the real length of the object turned out to be 48.5 cm, therefore the error in this case is only 0.1 cm (1 mm).

### 5 Performance comparison

The results of the measurements from Meshroom are shown in Table 1 and from Agisoft in Table 2.

	Real Dimensions	Measured Dimensions 60 Photos	Difference	Time
width	27.5	27.08	0.08	5861.8 s (97.7 min)
length	48.5	48.68	0.18	
height	39.5	39.726	0.226	

Table 1. Measurement results in Meshroom

	Real Dimensions	Measured Dimensions 60 Photos	Difference	Time
width	27.5	27.4	0.1	1686.867 s (28.1 min)
length	48.5	48.6	0.1	
height	39.5	39.5	0	

Table 2. Measurement results in Agisoft

Also, the third element of the comparison between the reconstructed object in Meshroom and Agisoft is the quality of the reconstructed object. Agisoft is the software that gave the best visual result in this case study.

Analyzing the experimental results, it is concluded that measurements in agisoft are more accurate than Meshroom, and execution time offered by Agisoft are better compared to Meshroom software.

In conclusion we say that Agisoft compared to Meshroom

- Offers more optimization techniques
- Provides reduced processing time
- Offers much higher quality in the reconstructed 3D object
- Offers the highest accuracy in measurement

### 6 Conclusions

In this article it is studied the reconstructed a 3D object from different sets of photo. For set of 60 photos of a real object rebuilds the 3D object and respectively examins the processing time, quality and precision of the object dimensions in Meshroom and Agisoft.

The reconstructed object with a set of 60 photos resulted in a satisfactory recognizable, complemented by content and clear contours in Meshroom. Having a known

size element of the object, the reconstructed object is scaled. The dimensions of the reconstructed object from Meshroom are measured in MeshLab software. Results show that dimensions of the object are obtained with high accuracy in a range of errors between 0.08 cm to 0.226 cm in MeshLab.

From the experiment in this article it is seen that a very accurate precision has been achieved in Agisoft for the height, for the length an error of only 0.1 cm and for the height an error of only 0.1 cm.

In conclusion, the results provided by Agisoft which are better than results provided by Meshroom have an average error less than 0.1 cm and sufficiently recognizable in terms of quality.

# 7 Conflict of interest

The authors declare no conflict of interest.

### 8 Acknowledgment

This work was supported by the National Agency for Scientific Research and Innovation under the Contract no. 831.

### 9 References

- [1] Jiménez Moreno, R., Aviles, O., & Hernández Beleño, R. D. (2017). Humanoid robot cooperative system by machine vision. International Journal of Online and Biomedical Engineering (iJOE), 13(12): 162–173. <u>https://doi.org/10.3991/ijoe.v13i12.7594</u>
- [2] Enesi, I., Kuqi, A., & Zanaj, E. (2022). Quality of 3D reconstruction based on photogrammetry for small objects, a case study. Mater Sci Eng, 1254: 9. <u>https://doi.org/ 10.1088/1757-899X/1254/1/012039</u>
- [3] Marian, M., Krajcovic, M., & Gabajova, G. (2021). Creating 3D models of transportation vehicles using photogrammetry. Transp Res Proc, 55: 584–591. <u>https://doi.org/10.1016/j.trpro. 2021.07.025</u>
- [4] Erich, S., Oto, H., & Kučera, E. (2022). Low-cost method for 3D body measurement based on photogrammetry using smartphone. Eng, 11: 48. <u>https://doi.org/10.3390/electronics11071048</u>
- [5] Asadpour, A. (2021). Documenting historic tileworks using smartphone-based photogrammetry. Mersin Phot, 3: 15–20. <u>https://doi.org/10.53093/mephoj.899432</u>
- [6] Florent, P., Quentin, V., Christian, M., Leif, K., & Roland, B. (2020). Initial user-centered design of a virtual reality heritage system: Applications for digital tourism. Sci Eng, 12: 2583. <u>https://doi.org/10.3390/rs12162583</u>
- [7] Fidji, B., & Yann, B. (2020). Scyland3D: Processing 3D landmarks. Ope sou soft, 5: 1262. <u>https://doi.org/10.21105/joss.01262</u>
- [8] Lee, & Won, H. (2010). Bundle block adjustment with 3D natural cubic splines. Sci Eng, 9: 9629–9665. <u>https://doi.org/10.3390/s91209629</u>
- [9] Dias, B., Alencar, D., Carvalho, F., & Yaedú, R. (2019). Virtual planning of premaxilla replacement surgery using free software: A method proposal. J Oral Maxillofac, 48: 9. <u>https://doi.org/10.1016/j.ijom.2019.03.023</u>

- [10] Zuza, M. (2018). Photogrammetry 2-3D Scanning simpler, better than ever. Available online: <u>https://blog.prusaprinters.org/photogrammetry-23d-scanning-simpler-better-than-ever\_29393/</u>
- [11] Kraus, K. (2017). Photogrammetry: geometry from images and laser scans. Walter de Gruyter, Berlin, 459. ISBN 978-3-11-019007-6 Linder.
- [12] https://www.agisoft.com/forum/index.php?topic=9793.0
- [13] Metashape Wikipedia
- [14] Yang, G., Li, D., Ru, G., Cao, J., & Jin, W. (2018). Body height estimation system based on binocular vision. International Journal of Online and Biomedical Engineering (iJOE), 14(04): 177–192. Archeologists Wield Very Big Stick in Yucatan Exploration. Bill Giduz. <u>https://doi.org/10.3991/ijoe.v14i04.8400</u>
- [15] Agisoft PhotoScan. Geospatial Modeling and Visualization.
- [16] Mapping Vertical Terrain. Greg Hosilyk.
- [17] Brandolini, Filippo, Cremaschi, Mauro, Zerboni, Andrea, Degli Esposti, Michele, Mariani, Guido S., Lischi, & Silvia (2020). SfM-photogrammetry for fast recording of archaeological features in remote areas. Archeologia e Calcolatori, 31(2): 33–45. <u>ISSN 1120-6861</u>. <u>https://doi.org/10.19282/ac.31.2.2020.04</u>
- [18] 3D reconstruction and photogrammetry software PhotoScan. Geoscan.aero.
- [19] https://www.agisoft.com/
- [20] https://www.creativetools.se/software/metashape
- [21] Enesi, I., & Kuqi, A. (2022). Analyzing the measurement accuracy of 3D reconstruction for small objects using different software; A case study, Journal of Engineering Research and Sciences, 1(12): 9–14.
- [22] Enesi, I., & Kuqi, A. (2022). "Analyzing parameters in 3D reconstruction photogrammetry in Meshroom, a case study," 2022 11th Mediterranean Conference on Embedded Computing (MECO), Budva, Montenegro, pp. 1–8, <u>https://doi.org/10.1109/MECO55406.2022.9797170</u>

### 10 Authors

**Indrit Enesi** is professor from Polytechnic University of Tirana, Albania. He has PhD degree in 2009 from Polytechnic University of Tirana, more than 25 articles published in conferences and more than 10 articles in journals.

Anduel Kuqi is currently a PhD student in the Polytechnic University of Tirana, working in the field of computer vision, jas published in four conferences and in two journals.

Article submitted 2022-12-07. Resubmitted 2023-02-01. Final acceptance 2023-02-06. Final version published as submitted by the authors.