

Dear editor and reviewers,

I thank sincerely for the chance to submit a revised draft of "Towards a Mobile 3D Documentation Solution. Video Based Photogrammetry and iPhone 12 Pro as Fieldwork Documentation Tools". I thank as well for all the valuable comments and suggestions I received from the reviewers. I have included below a list of reviewers' comments and concerns and my reactions to them, with references in changes by line numbers. Additionally, some additional comments regarding the errors in the RealityCapture dataset were included in lines 291–298.

The revised article (both as a text document and a pdf-file) with relevant tables, figures and data have been resubmitted to Zenodo.

## Reviewers comments to authors:

### Reviewer 1:

1. The abstract is good, but suggests a positive outcome for the use of these mobile alternatives, which is not the case in the conclusions.

*Author response: Thank you very much for the comment! The abstract has been harmonized with the conclusions (especially lines 15–20).*

2. Only one previous work by Luetzenburg, Kroon and Bjørk is cited that has already done some preliminary testing. More comparable articles like the one discussed here should be included, as the comparison of different scanning methods is nothing new.

*Author response: Other relevant articles have been referred to now (e.g. lines 40--48).*

3. Here, I would have wished for more explanation concerning the M3C2 comparisons, as this part was not very clear (see also in Tables and Figures). The author's statement, that 2cm are within acceptable margins of error might be true for the site he was investigating, but this certainly isn't true for every site. I would advise to discuss more the application of these scanning results with regard to the type of excavation (Paleolithic excavations for example works quite differently).

*Author response: I have expanded the text regarding these concerns, which I find reasonable. Motivation to use M3C2 has been added in 202--206). Explanation for the comparisons has been added in the results section LINES, with a comment regarding their suitability for different contexts.*

4. The discussion is much clearer than the beginning of the paper as it clearly states that the technology tested is not capable of achieving acceptable results yet. I agree with the author that at the moment, these technologies are not sufficient and reliable enough, but this might as well change in due time. Again, here the author could refer to other works that have done similar things in the field, albeit not with the exact same technology. But comparisons between Structure-from-Motion and Laserscanning have been done many times for example. These references are missing.

*Author response: Several references were added, as noted above. I also tried to sharpen the results in the discussion section.*

5. I think Figure 7 should be three images, but I only see one. Figure 6 is very good and comprehensible. Figure 5 needs more explanation, so readers can understand what this graph actually means. Figures 3 & 4 are very good. Again, I think Figure 3 is missing images, as I only see two overlapping in the center. Figure 1 is fine. Tables are useful and comprehensive. Captions are good.

*Author response: Figures have been fixed and explanations for the graphs added. I included a pdf version of the article in the repository with the images correctly displayed, so now they should all be correctly visible. Images are also available in a separate zip file with corresponding filenames..*

## **Reviewer 2:**

1. In general terms, it is a well-written paper of an interesting topic. The language is adequate, clear, precise and consistent. Also, it is very positive that the use of innovative documentation technologies is pointed out quite frequently in the text, defining well the scope of such applications within the archaeological prospection.

*Author response: Thank you!*

2. it is generally recommended to be integrated with a reference link in the text, and especially for the figures presenting results.

*Author response: References to figures and tables have been added.*

3. but video-based photogrammetry cannot be considered as novelty, as it dates back to more than a decade now, and is used massively in cultural heritage documentation.

*Author response: Thank you for this comment. My intention was to emphasize that the method has not been widely adopted for archaeological fieldwork. I have now added references to*

*applications in other cultural heritage fields and also on the general history of "videogrammetry" in lines 65–69.*

4. but neither in this section nor in the next one (Data acquisition) you mention the distance of the sensor (camera) from the object that defines the resolution (pixel size) of your model and the overlap within the sequential frames. Time interval is an important parameter when working with videos for photogrammetry, but it is also important to mention the overlap that results from this frame rate. And of course the spatial resolution of the images.

*Author response: This important comment is now attested to in lines 134–137.*

5. In Data acquisition section, line 166, it would be preferable to add the actual dimensions of the area of interest.

*Author response: Actual dimensions have been added to 2.2 Data acquisition, under lines 187–188.*

6. In line 182, there is a very interesting tool mentioned, the MSA. Maybe it would be useful to add a comment on the (almost apparent) reason why you used this tool (in order to align the different datasets, coming from different sensors which means different georeferencing etc). The fusion of data from different sensors (in this case from different GPS antennas) is an active research topic of high interest.

*Author response: Thank you for the comment. Discussion and explanations regarding MSA have been expanded in 209–219.*

7. In Volumetric comparisons section, lines 270-271, the statement that Lidar point clouds "tend" to be more noisy than the photogrammetric ones, needs documentation. Maybe add one or two references from similar works that show this tendency?

*Author response: Thank you for pointing this out – in the original the thought was expressed unclearly. It is true that there is no such universal tendency for LiDAR data to be more noisier – indeed, depending on the acquisition methodology and the equipment used, photogrammetric measurements can similarly exhibit significant errors due to "noise" in the data. However, in this particular case the tendency of LiDAR data (that had not been filtered by the user to maintain the idea of performing only rudimentary postprocessing) to contain scattered points around the actual surface of the object seemed to inflate the volumes, whereas this phenomenon does not occur with the photogrammetric measurements. As stated in the text, this is only one*

*possible explanation to this – it is possible, that the error is instead due to inaccuracies of the RealityCapture result.*

8. In Conclusions section, lines 293-296, when referring to "many developers", add references.

***Author response:** This part has been removed, since I have not seen any developer explicitly state this – instead, there has been a tendency of not focusing on developing specialized handheld devices that could excel in this kind of applications (for instance, FARO Freestyle 2, which I have personally done experiments with, has not seen much upgrades after its launch, and competitors have not been active in making similar devices). The section was replaced with some additional comments beginning from the line 344.*

9. Lines 303-312 are stating something completely inaccurate.

***Author response:** Thank you for the comment. Whereas I respectfully disagree with the section being completely inaccurate, I decided to cut this part out of the article due it not being entirely relevant to the actual topic at hand. The same problem was raised by another reviewer as well.*

10. In line 319, define "traditional methodology". Is it surveying with total stations? Taking tape measurements? And are the results of "traditional methodology" actually comparable with photogrammetry? (Define also which results could be compared among traditional surveying and photogrammetry. Is it a distance? an area? a volume?).

***Author response:** This has now been expanded in 360–363.*

11. To sum up, this is an interesting article about state-of-the-art and low-cost methods of documentation, but it needs to be slightly enhanced technically.

***Author response:** Thank you for all the accurate and point-on comments, which I hope I have adequately reacted to.*

### **Reviewer 3:**

1. It is somewhat surprising that the conclusion, particularly in the first paragraph, seems to - at least partially - deviate from the general tone of the paper, especially in comparison to the general tone of the abstract. I think the conclusion of a research paper should ideally provide a succinct summary of the findings and their implications, tying together the main arguments and evidences presented in the body of the paper. If the conclusion appears to diverge from the overall narrative or

argument developed in the paper, it might lead to confusion or misinterpretation for readers (at least it seems somewhat like that to me in the present case). If the divergence in the conclusion is intentional and serves a specific purpose, such as introducing a new perspective or highlighting future research directions, I think it would be best that it should be done in a manner that still aligns with the overall theme and argument of the paper. If it isn't intentional, then revising the conclusion to better reflect the paper's main points and findings would be recommended to maintain coherence and continuity. Especially regarding methodology, it is worth questioning whether the evaluation of precision and accuracy values, which are nevertheless at a very high level, may be excessively meticulous

*Author response: I am grateful for these comments and agree with them. The conclusions and the abstract have been altered accordingly. Additionally, the arguments for retaining the meticulousness have been honed (lines 358–363).*

2. Considering the everyday practice of archaeology (as the authors have already acknowledged in the text), it's clear that laboratory conditions cannot be assumed. Consequently, the conclusions should be tempered when considering field conditions, which are often influenced by extreme weather conditions, time pressure, and restrictive budgetary constraints. In light of these considerations, it may be more appropriate to view smart devices in a somewhat more favorable light, as they make a significant contribution to the further democratization of technology. Specifically, they make LiDAR and photogrammetry technologies more accessible to a wider audience. The use of smart devices can streamline these processes, making them more practical and feasible in field conditions where traditional, more expensive equipment might be impractical or unaffordable.

*Author response: Thank you. I agree with the above, but simultaneously I think that the capabilities of these devices should be scrutinized systematically. I have now toned down the criticisms and made it more obvious that there are already many uses for these devices and that they are an important step on the democratization of these technologies (e.g. line 364–365).*

3. The exploration of dominance and supremacy within geopolitical 'blocks' presents a complex and multifaceted research topic, encompassing elements such as political power, economic influence, and cultural hegemony. This subject inherently involves the dynamics of global power structures, regional hierarchies, and the strategic maneuvers of different nation-states. However, the scope and nature of this topic diverge significantly from a research paper focusing on the utilization of smart devices for digital recording. [...] Hence, an attempt to incorporate an in-depth exploration of

geopolitical dominance into a paper about digital recording with smart devices would likely result in a lack of cohesion and focus. [...]

***Author response:** I am grateful for these comments regarding the section about the wider geopolitical discussion, and have decided to remove the section to improve the focus of the article.*