Response to reviewers' comments
EGUSPHERE-2025-1110
Curvature-based pebble segmentation for reconstructed surface meshes

Aljoscha Rheinwalt, Benjamin Purinton, Bodo Bookhagen

Dear all,

We appreciate the thoughtfulness of the reviews and thank the reviewers for making time to provide in-depth and constructive critiques of our work. We think that their comments have helped to improve the manuscript considerably. In order to address the reviewer's points, we have substantially changed the structure, figures, and text of the manuscript. In short, we have:

- Changed the structure by having the Material and Methods section ordered by workflow. The reconstruction accuracy assessment was moved to its own section in the Results. Also, discussion text fragments were moved to the Discussion.
- Added a new figure (cf. Fig. 3) and improved two figures (cf. Fig. 1 and 8) to address some of the reviewer's concerns.
- Added additional references as requested (cf. Introduction).

In the following, we provide a detailed and point-by-point response to the various issues raised by the two reviews. Note that we use bold, italicized typeface for the reviewers' comments and normal typeface for our responses. In addition to this, we also provide a latex-diff output that tracks all changes made to the original submission.

Reviewer #1

- ***However, the presentation of the work is currently hard to follow, due to: 1) A convoluted and occasionally confusing method section that presents some results (the table tennis ball experiment) and some discussion (e.g., of table tennis ball results, camera recommendations). For more details and specific suggestions, please refer to the in-line comments in the attached pdf.***

  We thank the reviewer for the critical evaluation of our work and for supporting the manuscript for publication. The suggestions were well taken.

  We restructured the method section in the order of the workflow and moved the table tennis results to the results section. Also, the discussion about that result, as well as about the choice of camera and photo taking, was moved to the discussion section.

- ***2) A too-short results section with more bullet points than text (see in-line comments in attached pdf).***

  We rephrased the results section into text without bullet points and also extended it with the reconstruction accuracy results. We also addressed the PDF in-line comments. Most importantly, we incorporate the F1-score as a measure for the pebble detection performance instead of IoU. This likely also reduces confusion with the IoU that we estimate for segment surface area evaluations. Regarding this, we also modified Fig. 1 and 8 so that it is more obvious that these are side views or cross sections of pebbles and not top-down views.

PDF in-line comments:

- ***Acronym needs explanation; better to avoid in the abstract***

  We spelled it out.

- ***1) No real comparision with photo-sieving approaches is undertaken. 2) It is arguable if topographic data from SfM represents "true" 3D because of occlusion. "True" would suggest something like CT data or a tif stack. Better rephrase to something like: "our results emphasize the potential of 3D analysis".***

  We changed the sentence to: "Our findings emphasize the potential of 3D analysis over traditional 2D approaches and suggest future improvements through refined segmentation algorithms and enhanced surface reconstructions."

- ***Would suggest rephrasing; While it is important to highlight that 2D data are inhernetly different than 3D data, this is highlighted in most of the cited studies. Would propose something along: "Furthermore, and most importantly, any pebble shape measured on a projected plane ..." Also, Fig. 2 is referenced before Fig. 1. The figure oder should be switched here.***

  We changed the sentence to: "Furthermore, we note that any pebble shape determination on a 2D projected plane will likely suffer from an additional bias introduced by tilting the pebble, such that the c-axis is not pointing directly upwards (cf. Fig. 1)."

  We also fixed the order of the figures.

- ***Fig.1. It seems unintuitive to start with panel A in the upper right.***

  This is a bit subjective, but we are coming from a circle like in trigonometry, with 0° pointing East and then going counterclockwise. In any case, the order is clear with arrows and A, B, C, and D.

- ***would suggest rephrasing to something like "a more complete 3D model" because topographic point clouds are not completely 3d (cf. Fig. 2).***

  We changed that to: "An exciting prospect is the segmentation of individual grains from a mesh or point cloud, allowing for a partial 3D model of every visible pebble down to a measurement limit determined by the model accuracy." The term full 3D model is now only used for the 3D models that reconstruct the object from all sides.

- ***please specifay what is meant by "network-based". Also, the statement currently suggests that Steer et al. 2022 would use a machine learning approach. As far as I remember, they use geometric ellipse fitting and a watershed approach.***

  We changed that to: "Walicka and Pfeifer (2022) presented a method based on random forest classification and then clustering to segment grains, and Steer et al. (2022) presented a method that segments the point cloud into watersheds to obtain grain boundaries."

- ***too specific for popping in up in the last line of the introduction; would suggest removing.***

  Removed.

- ***Major point: The method section is long, convoluted and sometimes confusingly structured. This in parts is caused switiching back-and-fort between higher level workflow statements and specific technical details. In its current form the methods read like summaries of several convoluted, small-sclale experiments (e.g., curvature 3D reconstruction in section 2.2, different camera settings in 2.5...). While all the information therein seems of value, it is convoluted with discussion statements that are just too much to process for readers. For example, in line 278ff is stated: "We have used additional cameras to reconstruct scenes with similar results (not shown). We would like to de-emphasize the importance of a specific camera to reconstruct scenes and instead make some general observations to improve scene reconstructions. A full-format sensor with a fixed lens will provide excellent results." Another example is the text in lines 119-127; these are results and a short discussion of thereof.***
  I do support the intention of providing readers with guidance at each step. However in the current form, it is hard to follow which setting, step and workflow was applied at which point; E.g., which camera setup was used for the full pebble setup, which for the table tennis balls. Which acquisition strategy was used for the balls? Which of the 3 image acquisition configuirations was used? Was the setup for the ground truth pebbles different? How many SfM models were created? The answer to these questions is currently spread over the entire method section.
  To improve this, I would suggest to: 1) add a higher level workflow overview to allow readers to orientate themselves at which stage they are in workflow. 2) re-ordering the section in more logical way (i.e., following a workflow order as described above). It is utterly confusing to have a section on camera acquisition after the sections on segmentation end evaluation. Section 2.3 and 2.4 should be the last, since they are your main contribution and how you evaluate the quality of your segmentation. 3) a brief descritption of the main experiment(s?) of the paper. E.g., the ball experiment and the pebble scene experiments. 4) Details on each step that are more of a sensitivity analysis or not entirely relevant for the main experiments could be moved to an appendix section (which is typesetted in ESurf). This would allow readers to focus on

the main experiments, while still conveying the helpful informations for each step.

Yes, we reordered the sections in the order of the workflow and moved the table tennis results to the Results. Also, a few text fragments were moved to the Discussion. We added a few sentences explaining the different setups, but also the new structure should clarify this further.

- ***Could you provide a bit more details on rounding, sorting, the lithologies and light conditions?***
  Also, is this depicted in Fig. 1A?

  We added a bit more detail (cf. Sec. 2.1 P5, L95). Yes, this is depicted in the mentioned figure. We mention that in the figure caption.

- ***Did you change the pebble arrangements or do you mean your one setup represents a wide range of narurally occuring conditions? Please clarify.***

  This was just one random placement without repetition. Something that could be further explored in future studies. We added a sentence explaining that more (cf. Sec. 2.1 last sentence).

- ***At this point it would be good to specify why you consider these 2 tools.***

  We just mention these two tools because we use them. We also mention that differences are negligible considering our results, it's rather a matter of preference. Some prefer a GUI, some a command-line interface, useful for remote logins. It's beyond the scope of this manuscript to provide a review of all available software for surface reconstruction.

- ***Can you provide more specific information here or maybe cite some other work? Otherwise, I would propose dropping the statement beause of it i) not really being within the scope of the study, and ii) being hard to corroborate.***

  Removed.

- ***This should be reported in the result section.***

  Done.

- ***Please specify here if you mean your 10 ground truth pebbles or something else.***

  Done. We also avoid the term "ground truth pebbles" and use "full 3D models" instead.

- ***just to confirm, you calculate the 2D IoU here (or do you mean the recintructed 3D surface? Please clarify if not 2D. Are you calculating an IoU for each object(triangle?)-area or an IoU for the area of all detections vs ground truth? The latter would better be called Jaccard score or index, because in object detection IoUs are usually done on individual objects.***

  We calculate the IoU of 2D surfaces in 3D, i.e., without projection bias. We do this on the basis of triangles, so we have IoU values for each pebble (object). We added a few sentences explaining this more (cf. Sec. 2.5 P10, L226 and P12, L266). Also, Fig. 8 is slightly modified in order to clarify that this is a side view or cross section of a schematic pebble.

- ***This reads like a result discussion.***

  Moved to the Discussion.

- ***Are your referring to the models of the 10 pebbles or are you refferring here to the best model of the full scene? Please specify.***

  This is done on a one-by-one basis. One full 3D model of a pebble is globally matched into the sandbox scene, and the retrievable part of that full 3D model is determined, i.e., which part of the full 3D model is visible in the sandbox scene. This is what we call the best-case segment. All of it is done for each of the ten numbered pebbles. We tried to clarify this in the new version (cf. Sec. 2.5 P12, L260).

- ***I am not sure if this coukld be called IoU here. This read more like a mAP (mean average precision) score in object detection (e.g., https://learnopencv.com/mean-average-precision-map-object-detection-model-evaluation-metric/).***
  Maybe using the F1 score would be less ambigous. The F1 score for your results would be also very high, so I am not challenging the accuracy of the results.

  As mentioned in one of the previous points, we compute the IoU for 2D areas in 3D on the basis of triangles. One can also call this measure Jaccard score, and we mention that now. For the detection of pebbles, we use the above-mentioned F1-score now.

- ***This could be moved to the digital supplement. A few sentences that summarize the steps would be more helpful to every reader that is not using the OpenMVS CLI.***

  Removed.

- ***can you specify what constitutes "sufficient filtering", "clean tie points"? Do you refer here to the manual filtering option in Metashape via "gradual selection"?***

  The filtering is done by reconstruction uncertainty and projection accuracy, as stated in the sentence before that sentence. Sufficient in this context is very debatable. We just want to create awareness for the possibility of tie point filtering.

- ***What is the ground truth you are using here to obtain the 318 pebbles and their are for calculating these numbers? I went back the method section but I am still unsure. Could you do a figure where you plot ground truth next (or over) your reconstructed pebbles?***

  We manually counted the 318 pebbles visible in the sandbox scene reconstruction. We showed a photo of the sandbox scene (cf. Fig. 2) and a virtual photo of the reconstructed scene (cf. Fig. 4). Since we argue for sub-millimeter accuracy of the reconstruction, differences between a photo and the model are very subtle and hard to spot. The model and also the photos are available (cf. https://doi.org/10.5281/zenodo.14987825).

- ***Did you use a threshold (e.g., an IoU value) to gauge when a pebble (not triangle?) is considered a true positive?***

  No thresholding. As stated in the second sentence of that paragraph, we consider a segment to be a true positive if it is the only segment on a pebble. There is no minimum IoU required. As soon as there is a second segment on that pebble, only one is the true positive, and the other is a false positive. If no segment is touching that pebble, we have a false negative.

- ***Do you mean here the average IoU or the mean average precision at a certain IoU? Please specify.***
  Maybe consider renaming to Jaccard index here (if not object-based; see related comment above) and only keep IoU for obejcts in section 3.4 to avoid

confusions with mAPs. Maybe also consider calculating the F1 score instead (see also above).

Yes, we use the F1-score instead now (see also above).

- ***This would be better placed in the discussion section.***

  Done.

- ***These are really encouraging results.***

  Thank you. Much appreciated. As said before, one should redo the analysis with more pebbles and variable settings (bootstrapping ensemble) in order to get uncertainties on these numbers as well as better constraints in more general settings.

- ***This would also be better placed in the discussion section; or, it could be expanded; then the first sentence could go in the method section (section 2.4). In this case there could be one or two more sentence summarizing the represented surface area ratio.***

  We moved it to the Discussion.

- ***Adding some additional information on the computational cost, processing time and hardware demand might be helpful for users. Not a full benchmark, just a few informations on your setup; e.g., do you use standard desktop, something like a gmainc PC, high-end or even a cluster? GPU or mainly CPU? Does the segmentation take seconds, minutes, hours with said setup?***

  Done (cf. Sec. 4.5 P21, L405).

- ***On what result is this statement based? It could be worth to add a subsection in the result section where you describe the different meshs you obtained together with their quality. Some of this information is currently included in the method section.***

  We added a "likely" and provided reasoning in the sentences that follow (cf. Sec. 4.1 P20, L350).

- ***please spell the acronym out here once.***

  We replaced UAV with drone for all instances.

- ***This statement would need a reference. What "dangers" are you referring to here? Current European UAV to only deploy safe UAVs in adequate distance to people. Please specify or remove, especially because this seems out of scope of this study. The previous sentences already warn users that UAVs might be not ideal, points them to pole alternative that you woul redommend.***

  Removed.

- ***This has already been stated in the introduction (and in the next section). Since you are not directly comparing to 2D methods, I suggest removin this sentence here.***
  in cas of "true 3D estimations" topographic surfaces are not "full" or "true" 3D (see comments above; Fig. 2). Please rephrase to something like "high-quality 3D"

  Removed.

- ***Bounding box methods also exist for 2D segmentation, and the bouding box could alos underestimate the true lenghts. Hence, I would keep the advantage of almost fully 3D separated from the bounding box results (which are nice for themslves).***

  Removed.

- ***I think this would be better placed in the Appendix to increase the readability of the discuission (in ESurf its typesetted and part of the main pdf; so it wouldn't be relegated to a digital supplement).***

  We prefer to have it as a sub-section in the Discussion. This avoids the overhead of separating it out of the main text, which also would require referencing it somewhere else in the main text. We also believe it nicely connects the discussion about our results with the final Discussion sub-section, Implications, and Future Directions.

- ***Could you plot these results either as histograms, pdfs or cdfs?***

  These are histograms. We also say so in the figure caption now.

- ***These results should be reported in the result section, if they prominently picked up in the conclusions (see related comments above)***

  This was just a rephrasing in percent instead of absolute numbers, but we now also added the percentages to the results section (cf. Sec. 3.1 P15, L292).

- ***While it is true that 3D data has inherently more information, your study does not compare your results to 2D methods (or other 3D methods). Hence, I'd suggest to focus ion the high quality of your results by rephrasing to something like "High quality pebble segmentation in 3D provides more information than 2D data on pebble orientation, primary axes, and potentially volumentric parameters"***

  We rephrased this to: "Compared to 2D approaches, 3D pebble segmentation provides a more complete representation of individual pebbles. 3D pebble segments allow for the estimation of all three primary axes and orientations, including the possibility of retrieving surface or volumetric parameters."

Reviewer #2

- ***However, the structure of the paper needs improvement to emphasize the results better and provide a clearer line of thought. Now, the reader needs to jump forward and backward in the text. The appearance of the references of the figures does not follow the figure order (e.g. Fig. 2 is referenced sooner than Fig. 1). Terms and methods are used before they are explained, and some conclusions are drawn before presenting the supporting data. Secondly, the main findings of the manuscript are not well separated from existing techniques. Section 2 contains forward and backward references, and the subsections do not follow the order of the workflow (taking photos, reconstructing the surface, evaluating the mesh), while Section 4 is redundant and contains repeated information from Section 3.***

  We fixed the order of figures and also added the requested figure showing vertex normals and divergence. We moved all discussions to the Discussion and all results to the Results (including the table tennis ball results). We also structured the method section in the order of the workflow.

- ***The main result of the manuscript is the segmentation algorithm. The techniques to obtain the axes and orientation from the segments are possibly new, but their novelty is emphasized neither in the abstract nor the text. The abstract and the introduction should be revised accordingly.***

  Yes, the segmentation algorithm is our main contribution. Retrieving yaw, pitch, and roll from 3D objects is done using existing techniques. Although these might be uncommon in the context of pebble segmentation, we would like to establish the use of orientations. 3D orientations are mentioned in the abstract and Introduction. The same holds true for the bounding-box method.

General comments:

- ***The phrase "ground truth pebble" is unclear. Some parts of the text suggest that ground truth is a 3D shape of the pebble, while others indicate that they are high-quality scanning of the pebbles sitting in the sandbox. If ground truth means the separated 3D shape of the pebble, the manuscript lacks an explanation of the method of matching the ground truth to the scene.***

  We removed that phrase and replaced it with "full 3D model" in order to make it more clear that these are 3D models that have the pebble reconstructed from all sides. It is explained how these models are created, and it is also explained how these are globally matched into the sandbox scene using the FPFH method (cf. Sec. 2.5 P10, L207).

- ***The validation of the orientations is unclear. How were the reference orientations of the "ground truth pebbles" determined?***

  Yes, we explain that now in more detail (cf. Sec. 2.5 P10, L220 and P11, L250).

- ***The tennis balls were evaluated by fitting spheres to their segments. What is the reason of not using the mean curvature for validation determined during the segmentation?***

  Yes, that would be an additional option. For a perfect sphere, both principal curvatures are equal to the mean curvature, and it would be proportional to the radius. However, we think it is more intuitive to directly work with the radius and discuss deviations from a perfect sphere in terms of these.

- ***Many referenced papers aim to determine the roundness of pebbles. Approximating the roundness of the total pebble from its segments would be worth mentioning in the part listing future directions. I also suggest referencing the paper by Ludmány et al. 2023, which also handled the problem of analyzing the geometry of a partly reconstructed body.***

  Yes, this is one of our main motivations for doing pebble segmentation in 3D. We want to go beyond common metrics like a-axis lengths. We believe this point is well conveyed in the current discussion section and even in the introduction (cf. P2, L50).

- ***The text does not clarify how the pebbles were placed in the sandbox (distributing them by human hand one by one or by some other technique). There are multiple possible ways of putting them into the box, which could lead to different results. It would be very interesting to repeat the sandbox experiments multiple times in the future.***

  Yes, we agree it would be very interesting, but this was just one random placement without repetition. Something that could be further explored in future studies. We added a sentence explaining that more (cf. Sec. 2.1 last sentence).

Specific comments:

- ***Page 2, line 53: I suggest including Fehér et al. 2023 next to the two referenced papers, which aim to retrieve shape properties from 3D models.***

  Added.

- ***Page 3, line 61: The last sentence should be followed by a statement that this manuscript aims to measure the accuracy of the SfM for small pebbles.***

  Added (cf. Introduction P3, L62).

- ***Page 3, lines 64-69: These sentences aim for clarity but cause confusion here. It is clear from the previous sentence what "mesh" means and it is also clear what the "volume of an individual pebble" means without further explanation.***

  Removed.

- ***Page 4, line 84: There should be some additional sentences stating the paper's goals to motivate the reader.***

  Added (cf. last paragraph of the Introduction).

- ***Page 5, Fig. 3: The word "ground truth" is unclear in the caption. The small images show partly reconstructed pebbles sitting on other pebbles, while the table contains properties (e.g. sphericity) suggesting that these pebbles were fully reconstructed. The matching technique also needs further explanation.***

  The term "ground truth model" is replaced by "full 3D model". The fact that these full 3D models are globally matched into the sandbox scene is now also stated in the mentioned figure caption. The matching is done using the FPFH method (see previous points).

- ***Page 6, line 116: How was the sphere fitting achieved?***

  It's a standard least-squares fitting now stated also in the text and the figure caption (cf. Fig. 10 and Sec. 3.1 P13, L282).

- ***Page 7, Fig. 5, caption line 3: It is unclear how fitting a sphere to a single triangle is possible.***

  We removed that sentence. We don't fit to individual triangles, but to mesh segments.

- ***Page 8, first paragraph of Section 2.3: Instead of referencing Fig. 7B, I suggest referencing the new Figure 2.***

  Done. However, the new figure is Fig. 3.

- ***Page 9, Fig.7: The colors of subfigures C and D are not explained in the caption.***

  Done.

- ***Page 10, line 169: How are the concave, convex, and flat regions distinguished from the mean curvature?***

  A positive mean curvature corresponds to convex parts in the mesh. This is made more clear in text (cf. Sec. 2.4 P7, L170) and also well conveyed by the new Fig. 3 and its caption, as suggested by reviewer 2.

- ***Page 10, lines 170-186: The text mentions concave and convex triangles that are quite confusing without a clear definition. I suggest either avoiding these terms or defining them at the end of line 169.***

  Yes, triangles themselves are planes and neither concave or convex. However, we extend vertex curvatures to triangles via averaging (cf. Sec. 2.4 P7, L167).

- ***Page 11, lines 202-203: It is unclear how the alignment of a 3D model can be compared to a pebble segment of a scene.***

  Again, we explain that now in more detail (cf. Sec. 2.5 P10, L220 and P11, L250).

- ***Page 11, lines 210-231: The description of calculating the axes and orientation of the 3D models and the segments should be separated. Now, the descriptions of manual measurements of the physical object, calculations of the 3D model and pebble segments are mixed, which is confusing.***

  It is. We first provide an overview of the three different types of validation (primary axes, orientations, and surface areas), and then describe in more detail how we do it. This section changed a lot, hopefully it is clearer now (cf. Sec. 2.5).

- ***Page 12, Fig. 8: Instead of referencing another figure, I suggest using the number of the pebble (2). The last sentence should be removed because it is not visible from this figure and corresponds to a later part of the manuscript.***

  Done.

- ***Page 12, Fig. 9: The caption should reference pebble 2.***

  Done.

- ***Page 13, lines 234-236: The last two sentences jump forward in the text and should be removed from here.***

  Done.

- ***Page 13, lines 241-244: I did not understand how to determine the retrievable part. The clarification of the phrase "ground truth model" might also clarify this part. Is the 3D model aligned somehow to the reconstructed surface?***

  Yes, the full 3D models are globally aligned or matched into the sandbox scene. We clarify this at multiple positions (see previous points).

- ***Page 17, Fig. 11: What do the red circles correspond to?***

  They highlight three examples of failures in the segmentation. We explain this in the figure caption now.

- ***Page 19, Fig. 13: The last two sentences of the caption should be moved into the body of the text.***

  Done (cf. Sec. 4.3 P20, L378).

- ***Page 19: I would also mention that Fig. 12 and Fig. 14. show that the errors in the axis and the orientation are related: the bottom five and the top five pebbles are the same in both cases.***

  Done (cf. Sec. 4.3 P20, L370).

- ***Pages 22-23: Many sentences repeat sentences from Section 3. These repeated sentences can be removed after combining old Sections 3 and 4 in the new Section 4.***

  This has been resolved during the restructuring of the text.

- ***Page 25-26: The conclusion should emphasize the main result, the segmentation technique, and how to calculate the axes and orientations.***

  We believe it does (cf. Conclusion P24, L442).