

# Rebuttal - “Towards Spherical Robots for Mobile Mapping in Human Made Environments”

July 13, 2021

*Manuscript Number: OPHOTO-D-21-00020 Towards Spherical Robots For Mobile Mapping In Human Made Environments*

*Dear Mr. Bredenbeck,*

*Thank you for submitting your manuscript to ISPRS Open Journal of Photogrammetry and Remote Sensing.*

*I have completed my evaluation of your manuscript. The reviewers recommend reconsideration of your manuscript following revision. I invite you to resubmit your manuscript after addressing the comments below. Please resubmit your revised manuscript by Jul 13, 2021. When revising your manuscript, please consider all issues mentioned in the reviewers' comments carefully: please outline in a cover letter every change made in response to their comments and provide suitable rebuttals for any comments not addressed. Please note that your revised submission may need to be re-reviewed.*

*To submit your revised manuscript, please log in as an author at <https://www.editorialmanager.com/ophoto/>, and navigate to the "Submissions Needing Revision" folder under the Author Main Menu.*

*ISPRS Open Journal of Photogrammetry and Remote Sensing values your contribution and I look forward to receiving your revised manuscript.*

*Kind regards,*

*George Vosselman Editor-in-Chief ISPRS Open Journal of Photogrammetry and Remote Sensing*

*Editor and Reviewer Comments:*

*Editor: As you will see below both reviewers appreciate your work and don't have concerns regarding the methodology or results. Reviewers, in particular Reviewer 2, however, do have a large number of recommendations to improve the readability of your manuscript. Given the number of recommendations, I ask for a major revision. The recommendations should, however, be easy to implement.*

The authors thank the reviewers for their time to review the paper and to provide valuable input. The following lists how we improved the manuscript based on their thoughtful comments.

*Reviewer 1: The manuscript presents the concept of a spherical mobile mapping system, and a data registration method utilizing the geometric constraints available in human-made environments, e.g. planar surfaces.*

*The manuscript has novel content and is well written, and can be accepted after minor revision. My comments and suggestions to be considered by the authors are as follows:*

*Lines 27-31: Spherical shell can protect the sensors inside, but this does not prevent the external disturbances such as dust or other impurity staining the sphere itself, affecting the performance of the system. So in principle this does not change the situation from using protective covers on "traditional" mobile mapping systems.*

We clarified that the shell can enclose all sensor without requiring a number of points of connection as individual enclosing for all sensors would. Further we added the requirement, that the shell is very durable.

*Lines 111-116: Modern kinematic scanning systems collect point cloud data with high speed, so also kinematic point clouds are often dense enough for robust plane determination. And on the other hand, sparse point clouds are not optimal for point-to-point matching due to the long distances between adjacent points.*

We clarified that also the limited FOV plays a role in not finding planes for robust correspondences.

*Lines 244-246: The Livox ranging precision is given twice in this sentence.*

Corrected.

Figure 4: The x-scale is given in figure caption, but I would suggest adding it to the figure also.

We added a label to the x-scale

*Conclusion: Whereas the title, Abstract and Introduction concentrate mostly on the spherical mobile mapping, Conclusion deals almost entirely with the topic of registration, thus feeling somewhat separated from the rest of the manuscript. Reading the Conclusion, one could not expect that the title of the manuscript is "Towards Spherical Robots For Mobile Mapping In Human Made Environments".*

We added more context about spherical robots and put the registration method into context.

*Reviewer 2: The authors have written a very interesting paper on the topic of mobile mapping with a spherical robot in human-made environments. In a rather slim paper and pleasantly concise form, they present their approach and show first results based on simulated and real measurement data. The reading flow is mostly pleasant and no irrelevant information is given. Conversely, the authors leave out important details in several places, which would be very helpful for the reader's understanding. In the following, there are several points for improvement and critical questions that should be taken into account by the authors. Due to the large number and the importance of central points, I recommend a major revision.*

*- A short outline at the end of the introduction is missing and would give the reader a better overview of the further content and should therefore be added.*

We added a short outline, introducing you to the next sections.

*- In the entire article, no reference is made to the targeted or required accuracy potential with regard to the quality of the pose determination and the final point cloud, which in my view requires at least a rough classification.*

The accuracy potential depends heavily on the motion profile, the environment and the hardware being used, thus is not easy to generalize. However, the simulation results show that the algorithm has the potential to improve the map quality based on point-distances by approximately a factor of ten. We added this to the conclusions.

*- For the short state of the art of existing mobile systems, which is mainly based on the authors' own systems, in my opinion, at least the NavVis VLX system could be added, which in comparison is newer and as a human-operated system for mapping is one of the benchmarks.*

We added the NavVis VLX as a reference in the state of the art section.

*- The authors mention unknown and dangerous environments as a possible field of application, in which the advantages of the spherical system should come into play. However, only one space application is mentioned as a specific example, which may be valid in itself, but does not fit the topic of the paper very well. Application scenarios should therefore be supplemented by terrestrial areas and in particular by human-made environments, which better motivate the future use and justification of such systems for the reader.*

We added four more terrestrial, human-made environments as dangerous examples where a compact, spherical robot would really unfold its potential.

*- The advantages of a spherical mobile platforms are discussed, but nothing is said about possible disadvantages or limitations. Moreover, I wonder whether the "unique advantages" mentioned in the abstract are really so remarkable. For example, what about the recording constellation, which in my view is relatively unfavourable on the ground and leads to restricted fields of view and weak angles of incidence when mapping the ground. Another advantage stated by the authors is that the shell protects the sensors, which is reasonable. However, this protection is only necessary because you are moving directly on the ground. In addition, can dust and dirt accumulate on the outer shell and influence the acquisition?*

We added disadvantages (including your example) in the introduction. They show the need for a reliable, robust registration procedure. There is an approach to prevent dust and dirt to accumulate on the outer shell. We added a note / reference to that.

*- Section 2.1 discusses existing laser scanner-based SLAM approaches, concluding with the statement (page 2, column 1, line 30-33) that all of these approaches do not include the structural properties of human-made environments in the calculations. This statement is not true. There are a variety of SLAM approaches (e.g. Jung et al. 2015; Nguyen et al. 2006), but also other applications with recursive state estimation, which make use of this additional information in the context of georeferencing. This should be investigated and reworked in more detail. Jung, J., Yoon, S., Ju, S., and Heo, J. (2015). Development of Kinematic 3D Laser Scanning System for Indoor Mapping and As-Built BIM Using Constrained SLAM. In: Sensors (Basel, Switzerland), 15(10), DOI: <https://doi.org/10.3390/s151026430> Nguyen, V., Harati, A., Martinelli, A., Siegwart, R., and Tomatis, N. (2006). Orthogonal SLAM: A Step Toward Lightweight Indoor Autonomous Navigation. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, DOI: <https://doi.org/10.1109/IROS.2006.282527>*

The authors thank the reviewer for spotting this false statement. The state of the art has been extended with approaches that consider the structural properties of human-made environments.

- *The introduction in Section 3 should be revised as it is not very intuitive for the reader and something about the objectives and motivation for this section is missing. However, this is also directly related to the general lack of outline beforehand.*

The introduction in Section 3 has been revised and includes now an outline.

- *In section 3, the authors introduce their own algorithm for registration, but they do not give any further information about the temporal context in which this takes place. This is probably done in the context of each epoch, although it is also not clear how an epoch is defined. Is it a single full rotation of the laser scanner or that of the sphere? And does the registration always take place only locally in relation from the current epoch to the last epoch or with a window approach or even globally?*

The registration is performed globally. The explanation has been adjusted to clarify this.

- *Section 3.2 deals with the extraction of planes, but the following questions remain unclear: Are the planes always searched for only in the context of the observational data of the current epoch or with regard to the already registered global point clouds? How many planes are needed at all and what demands are made on their spatial distribution? The abstract mentioned parallel and perpendicular planes, but this is not discussed further here. Are only these used and if so, why? Furthermore, what about planes that are not necessarily parallel or perpendicular (e.g. doors, cupboards)? How is the flatness of a plane checked? How are outliers dealt with, especially with regard to existing measurement and motion noise?*

We added more detail in the plane extraction subsection.

- *Throughout Section 3, numerous variables are not introduced and are not explained in more detail, so that a more precise interpretation of the individual equations is unfortunately very difficult.*

We clarified the equations.

- *What orders of magnitude were finally used for the range noise model in Section 4.2 and why was this not also done for the angular resolution?*

The distributions we sample from are specified in the caption of figure 3.

- *In the context of the real experiment in Section 5, a major simplification is made. Even if this simplification is certainly suitable for the beginning, this special case comes quite surprising here...especially in comparison to the simulation and the previously described topic. Have other measurements not yet been carried out or what is the background? At least this simplification should be mentioned beforehand, otherwise the reader will have false expectations.*

We added a further experiment that uses nearly all DoF: A sphere rolling on a ground plane. This was not done in the previous manuscript because we had issues getting the hardware shipped, assembled and the experiment carried out in time.

- *Wouldn't it also be interesting to compare the estimated pose parameters with the nominal values from the simulation within the framework of Section 6.1? So far, the focus has only been on the point cloud.*

Whenever the scanner is looking into the ground the number of points is too small and thus neglected in the optimization step. This leads to jumps in the trajectory, which make the comparison results less meaningful than expected. Due to time constraints this issue could not be resolved.

- *Why is there no qualitative validation for the results of the real data in Section 6.2? A reference scan based on a terrestrial laser scanner would be a possibility or otherwise based on the flatness of known planes in the registered point cloud?*

We couldn't access the rooms the dataset was acquired in for a terrestrial reference scan in the time given, therefore the quantitative analysis was done with the simulation and the real data was analyzed qualitatively.

- *In the Conclusion, the parallelism and perpendicularity within the environment is mentioned again, although this did not play a role in the main part of the paper. Therefore, the question arises why do the planes have to be parallel or perpendicular at all and why is this mentioned at the beginning? This is not clear, since no additional constraints on this additional information have been introduced so far. There are also already other methods available that take this idea of geometric constraints for georeferencing into account.*

We removed the initial mention of parallelism and perpendicularity of the planes as it is of no further interest in the rest of the manuscript.

*Further open questions/comments are: Page 1, Line 15: Double use of "protect", which should be replaced by another word.*

We removed first instance.

*Page 1, Line 21: The link for the "second option" to the use of an IMU is not necessarily obvious.*

We added IMU as sensor for coarse estimation.

*Page 1, Column 2, Line 52-56: Complicated sentence which should be reformulated.*

We have split into two sentences

*Page 2, Column 2, Line 15: The term "long stop" is not to be interpreted clearly in this context. How is "long" defined in this sense in terms of time or space?*

The specified stop is required to find enough points for plane detection.

*Page 3, Column 1, Line 6-8: A specific reference should be given for the first half of the sentence, as reference 19 probably only refers to the second part.*

Two references were added as examples for the first half of the sentence.

*Page 3, Column 1, Line 23-26: Not all variables were introduced accurately*

We have introduced missing variables and clarified rho.

*Page 3, Column 1, Line 33-34: Selection of references for this?*

We have added plane segmentation sources.

*Page 3, Column 1, Line 34-35: But what demands are made on the assignment? This is unclear to the reader, also in connection with the general quality requirements of the whole approach.*

We have added a specification for "good enough for plane detection".

*Page 3, Column 1, Line 38: How can the threshold value be specified, since it is probably not identical for all applications and environments? Later, there is also no indication of the choice of the value in the own simulated and real data sets.*

We indicated that we observed the noise level around the expected planes by the naked eye, as an educated guess. You find that part in the 'Local Planar Clustering' section.

*Page 3, Column 1, Line 50: What does the index  $k$  describe and what is the difference to the index  $i$  from Section 3.2?*

We have clarified indices  $k$ =planes  $i$ =points.

*Page 3, Equation (1): Again, some variables are not specified. For example, the source of  $t_x$   $t_y$   $t_z$  is unclear.*

We have introduced  $t_x$   $t_y$   $t_z$  properly.

*Page 3, Equation (3): Equation unclear regarding the use of the 2 in the exponent as well as the index.*

We have specified that we're using the square of the L2 norm.

*Page 3, Equations (4)-(5): Here I don't understand how to get to the elements of the normal vector through the transformation. Moreover, these individual elements were not introduced in advance.*

We have corrected indices and introduced the elements.

*Page 3, Equation (9): The variable  $\alpha$  is not described and explained*

An explanation of  $\alpha$  was added.

*Page 4, Equation (10): Description of the individual variables is missing and, moreover, rho has already been used as the distance parameter of the Hesse-normal form*

we have replaced rho with zeta and introduced the variables.

*Page 4, Equation (12): Is epsilon here the same variable as the threshold in Section 3.3? Probably not.*

We have renamed epsilon and clarified the difference to sec 3.3.

*Page 4, Column 2, Line 15: "some dimensions": Which ones and why exactly are they chosen? Are they the same ones all the time?*

We chose a better example to clarify the use case.

*Page 4, Column 2, Line 20-25: Reference is missing*

We added a citation.

*Page 4, Column 2, Line 55: Doubled information regarding the distance measurement precision*

We removed duplicated information.

*Page 5, Column 1, Equation (18): Delta t is not introduced and there is no indication of how large the time steps are in the simulation.*

We have introduced delta t and give a default value.

*Page 5, Column 2, Line 14: Should it be "out" or "our"?*

We have corrected this.

*Page 5, Column 2, Line 27-28: What are the specifications of the used IMU with regard to its uncertainty budget?*

We have added precision of IMU and added the source.

*Page 6, Figure 1: How are these values to be interpreted and from where are they derived? This is unclear.*

We have specified this. It is chosen to still allow plane detection.

*Page 8, Column 1, Table 1: What is missing is a classification of the values in relation to the expectations / demands...which are, however, still missing in advance*

This table was given to quantify the improvement of the map via the algorithm. In particular they are put forward as evidence of the improvement.

*Page 8, Column 1, Line 37-43: At this point, an interpretation of the significant second peak in the right-hand diagram in Figure 4 should be made. This has not been discussed so far.*

We redid the experiment with a larger amount of points (all that have a point-to-point distance of less than 3000 cm). The second significant peak was a artifact of the previous thresholding and hence is now not significant anymore.

*Page 8, Column 1, Line 51: "ten scans" means ten individual scan lines or 10 epochs or is this identical? What kind of temporal offset is there in between and wouldn't a motion model have to be taken into account for this or is an identical pose assumed for the ten scans?*

We have specified which scans and that we assume constant pose.

*Page 8, Column 1, Line 55-56: What could be the cause of this?*

We have specified the outlier dependence.

*Page 8, Column 2, Line 44: This sentence always applies and should not remain here.*

We have removed this sentence.

*Page 9, Figure 4: The two point clouds in the middle of figure 4 are difficult to interpret in terms of their orientation. Could it be an option to add a coordinate system?*

It was specified in the figure label that lateral images have the same orientation.