FEASIBILITY STUDY OF USING VIRTUAL REALITY FOR INTERACTIVE AND IMMERSIVE SEMANTIC SEGMENTATION OF SINGLE TREE STEMS

C. R. Fol*, A. Murtiyoso, V. C. Griess

Forest Resources Management Group, Institute of Terrestrial Ecosystems, Department of Environmental Systems Science, ETH Zürich, Switzerland - (cyprien.fol, arnadi.murtiyoso, verena.griess)@usys.ethz.ch

Commission II

KEY WORDS: Virtual Reality, Point Cloud Segmentation, Training Dataset Creation, Tree Stem Labelling, Microhabitats

ABSTRACT:

Forest digitisation is one of the next major challenges to be tackled in the forestry domain. As a consequence of tremendous advances in 3D scanning technologies, broad areas of forest can be mapped in 3D dramatically faster than 20 years ago. Consequently, capturing 3D forest point clouds with the use of 3D sensing technologies - such as lidar - is becoming predominant in the field of forestry. However, the processing of 3D point clouds to bring semantics to the 3D forestry data - e.g. by linking them with ecological values - has not seen similar advancements. Therefore, in this paper we consider a novel approach based on the use of VR (Virtual reality) as a potential solution for deriving biodiversity from 3D point clouds acquired in the field. That is, we developed a VR labelling application to visualise forest point clouds and to perform the segmentation of several biodiversity components on tree stems e.g., mosses, lichens and bark pockets. Furthermore, the VR segmented point cloud was analysed with standard accuracy and precision metrics. Namely, the proposed VR application managed to achieve an IoU (Intersection over Union) rate value of 98.74% for the segmentation of bark pockets and resp. 93.71% for the moss and lichen classes. These encouraging results reinforce the potential for the proposed VR labelling method for other purposes in the future, for example for AI (Artificial Intelligence) training dataset creation.

1. INTRODUCTION

Recent technological advances in TLS (Terrestrial Laser Scanner) and airborne lidar allow scientists to develop 3D models of forests that closely reflect the morphological and physical traits of individual trees (Lines et al., 2022). This opened up a completely new interpretation of forest ecology based on direct measurements in the 3D digital twin. The three dimensional approach already allowed researchers to increase the understanding of radiative transfer modelling (Calders et al., 2018), canopy microclimate (Zellweger et al., 2019), and biomass estimation (Calders et al., 2015).

The most widespread forest 3D model data are point clouds acquired by TLS or airborne lidar. However, thanks to progress made in low-cost 3D scanning technology, the source of forest 3D point clouds is likely to progressively diversify (Mokroš et al., 2018, Fol et al., 2022, Kükenbrink et al., 2022). Nevertheless, point clouds are certainly going to remain the preferred 3D format in the forestry field for the time being due to its heterogeneous and complex environment which is quite difficult to model using other 3D representations, such as meshes or geometric primitives. For this reason, the focus of the VR application as proposed in this study is exclusively put on the segmentation of 3D point clouds.

By reviewing the literature for VR and its interaction with point cloud data, another precedent was found on the subject of forest point cloud handling: the PointCloudXR project documented on the SLU (Swedish University of Agricultural Sciences) 3D remote sensing lab webpage (https://www.rslab.se/pointcloud-xr/, accessed 10 October 2022). This VR application was designed for forest point clouds acquired from TLS and allows to

both visualise point clouds based on colour, intensity or class information stored and perform basic operations on the point clouds e.g, distance measurements or noise removal. However, this application does not allow live labelling in VR and only handles point clouds up to a maximum of 10 to 15 million points (https://github.com/mhzse/PointcloudXR, accessed 17 October 2022). Nevertheless, several solutions exist in other domains. For example, Kharroubi et al. (2019) managed to load 2.3 billion points in real time and in a continuous manner in VR by making use of the computer's GPU (graphical processing unit) to store the point clouds and optimise the rendering process. In parallel, the use of the paint brush approach in VR can also be seen in several studies to segment point cloud (Zhao et al., 2019, Wirth et al., 2019, Virtanen et al., 2020). In this paper we propose a proof-of-concept for a VR application to help the point cloud segmentation task by taking into account the state-of-art of the technology as identified from these previous studies. The following research questions will be addressed in this paper: (1) how to load and optimise the visualisation of a large size point cloud, and (2) how to use VR tools to precisely segment point clouds of single trees.

Finally, in recent years, with the assistance of machine and deep learning algorithms, the identification and segmentation of biodiversity from tree point clouds has become possible as can be seen in (Rehush et al., 2018) who identified tree microhabitats (TreMs) from TLS point clouds. TreMs (tree-related microhabitats) are a popular biodiversity indicator that represents geometric features from trees that can serve as a shelter for highly specific species in the forest such as insects, bats, birds, or mammals (Bütler et al., 2020). However, this novel AI (Artificial Intelligence) assisted approach to segment forest biodiversity has not yet been able to unlock its full potential due to a major bottleneck: the availability of training datasets for forest

---

* Corresponding author

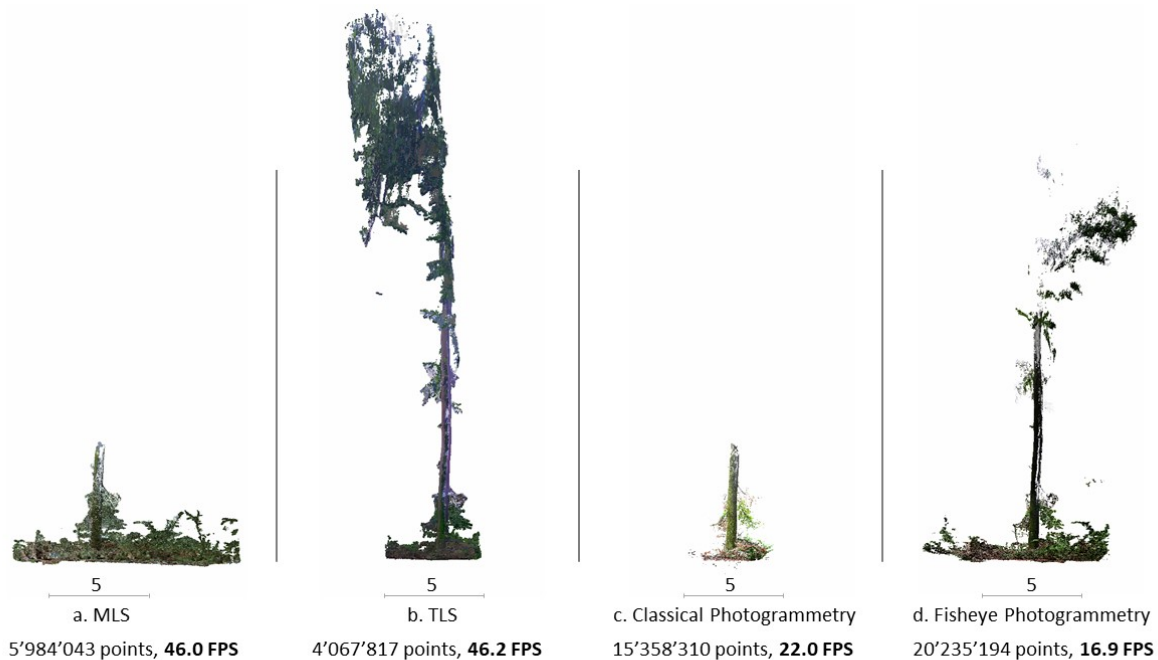| 5 | 5 | 5 | 5 |
| a. MLS | b. TLS | c. Classical Photogrammetry | d. Fisheye Photogrammetry |
| 5'984'043 points, **46.0 FPS** | 4'067'817 points, **46.2 FPS** | 15'358'310 points, **22.0 FPS** | 20'235'194 points, **16.9 FPS** |

Figure 1. Rendering performance comparison for forest terrestrial 3D scanning methods. Scale bar units are in meters.

environments. Indeed, the classical labelling approach makes use of dedicated computer software to manually segment the 3D data, a task which is time and resource-consuming for the operator. Therefore, the potential to use VR in aiding this task also constitutes one of the objectives in this research, in order to develop a more user-friendly, virtually immersive, and interactive system. This will in turn help fill the gap in the creation of training datasets for AI purposes.

2. METHOD

First of all, research question (1) will be addressed by analysing the FPS (frames per seconds) value for four different point clouds, acquired by terrestrial scanning techniques: TLS, classical photogrammetry, MLS (mobile laser scanning) and fisheye photogrammetry (c.f. Figure 1). FPS is the default value used in the domain of VR and more broadly in gaming to evaluate the rendering performance of a device. Hence, we will use this value to investigate the risk of encountering cybersickness issues while the user is immersed in our VR application (Poux et al., 2020, Virtanen et al., 2020). For that purpose, the Unity game engine has a dedicated profiler plugin providing the corresponding FPS value while a point cloud was loaded into the scene. Due to the disparity in the extent of these point clouds, a common section was segmented from the different point clouds to guarantee that only overlapping point cloud zones will be compared during the analyses (c.f. Figure 2) using the open source software CloudCompare (https://www.danielgm.net/cc/ accessed on 16 September 2022). This procedure allowed us to conduct a preliminary visual analysis to determine whether or not the level of detail on the tree stems was sufficient to extract ecologically valuable information.

This created the link to research question (2). Based on this analysis only the tree stems obtained by classical photogrammetry were labelled to extract ecological valuable information and answer question (2). More precisely, two types of TreMs,

namely bark pockets and mosses/ lichen, were segmented from the tree stems. In Switzerland, the inventory of TreMs is performed by a 2 person crew visually inspecting select trees from forest plots. In this regard, the inventory of TreMs is deemed to be a potentially useful application of the proposed VR application. Finally, to evaluate the quality of the segmentation task, a ground truth dataset was generated using manual screen-based segmentation in CloudCompare. Afterwards, metrics were computed to assess the quality of the labelling result, by computing the confusion matrix, Precision, Recall, IoU (Intersection over Union or the Jaccard index) and OAcc (overall accuracy) for each class.

3. RESULTS AND DISCUSSIONS

3.1 Point cloud visualisation

Figure 1 illustrates the difference in size and density of point clouds acquired by different terrestrial 3D scanning methods. FPS values are inversely correlated to the number of points and the size of the point cloud does not have an impact on the FPS value. TLS (Figure 1 b) point cloud was rendered with the highest FPS values despite its large size around 30 metres. However, each tree had FPS values between 16.9 and 46.2, which is below the minimum requirement for immersive VR application of 90 FPS (Virtanen et al., 2020). Despite the fact that such lower FPS are normally associated with jitters latency artefacts, the point cloud visualisation was still very smooth. This can be explained by the fact that the point cloud is static in the middle of the scene and that the GPU is handling all the computation related to the display. Using this setting, it is therefore assumed that the risk of user cybersickness may be limited. It is worth mentioning that there exists a way to meet the FPS required for immersive VR applications which implies the sub-sampling of the point clouds and the employment of algorithms such as real-time CLOD (Continuous Level of Details) (Schütz

1	1	1	1
a. MLS	b. TLS	c. Classical Photogrammetry	d. Fisheye Photogrammetry

Figure 2. Transversal section of tree stem acquired by terrestrial remote sensing methods. Scale bar units are in meters.

et al., 2019) or the implementation of the octree structure (Kharroubi et al., 2019). Regarding the levels of detail, in Figure 2 a very clear contrast between point clouds acquired by laser scanning (TLS in (a) and MLS in (b)) and photogrammetry (in (c) and (d)) can be observed. While in the two image-based point clouds the tree surface is faithfully represented which allows for a clear segmentation of the different features on the stem, in the TLS point cloud it is only possible to determine the presence of large features such as moss, lichen and cavities. This is even more evident in the MLS point cloud where colour information is practically non-existent and the level of noise around the bark is high. Based on these visual results, the photogrammetry point clouds (c) will be used as a reference for the segmentation of TreMs on the tree stem.

3.2 Biodiversity segmenation

For the evaluation of the segmentation in virtual reality, four metrics widely used in the field of machine learning and computer vision were assessed: OAcc, Recall, Precision and IoU. In addition, to better understand the performance of our VR application, we calculated the confusion matrix for each of the separate classification performed on the tree stem. The overall accuracy is a suitable metric to start the discussion about the evaluation of the segmentation because it indicates the probability of correctly labelling a point.

The overall accuracy for the labelling of mosses and cavities on tree barks are very high, respectively 93.91% and 98.42%. This result supports the claim that the accuracy of VR labelling is comparable to classical (screen based) labelling methods. To further evaluate the accuracy of the VR labelling results, the precision and recall rates are also useful. In fact, both values provide additional information on the quality of label assignment for point clouds in VR. Figure 3 shows that the assignment of moss/lichen label (P=89.91%, R=87.72%) was performed with better quality than for the bark pockets class

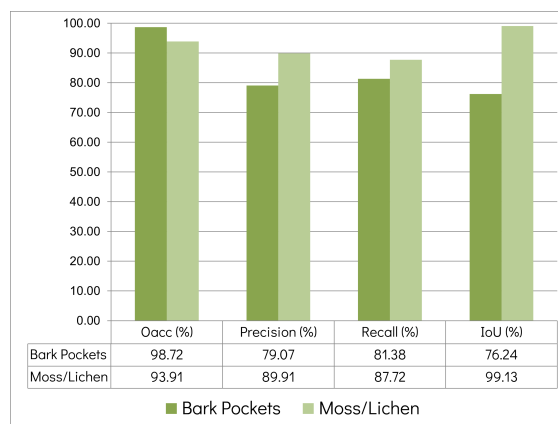|	Oacc (%)	Precision (%)	Recall (%)	IoU (%)
Bark Pockets	98.72	79.07	81.38	76.24
Moss/Lichen	93.91	89.91	87.72	99.13

Figure 3. Overview of VR segmentation performance metrics.

(R=81.38%,P=79.07%). This could be explained by the shape and size of the TreMs class. The mosses and lichens are part of a single instance spread over a larger section of the tree stems, whereas the bark pockets are small patches covering a small section of the bark (around 3% of total area).

They are further divided into several separate instances. Thus the difficulty of accurately labelling small patches compared to homogeneous and larger areas is evident and may explain the difference in quality between these two classes. Lastly, the IoU (Intersection over Union) or the Jaccard index metric is used to assess the quality of the segmentation task since it illustrates not only the accuracy but also the precision of such tasks. The IoU value is very high for the mosses and lichens class. This confirms clearly the relevance and efficiency of labelling mosses and lichen on tree stems using the proposed VR-based method. However, the IoU rate for the bark pocket class is lower, just above the third quartile. Even though there is a significant difference between the labelling of bark pockets and mosses and lichen, these values are in line with the results obtained by Re-

hush et al. (2018) using AI. We conclude that the prototype VR-based tool developed in this project provides great promise. It may also prove to be an interesting solution for the specific use case of the creation of AI training datasets.

3.3 VR proof of concept
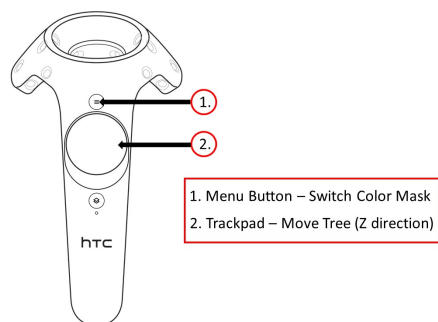

Figure 4. Tree stems before labelling process.

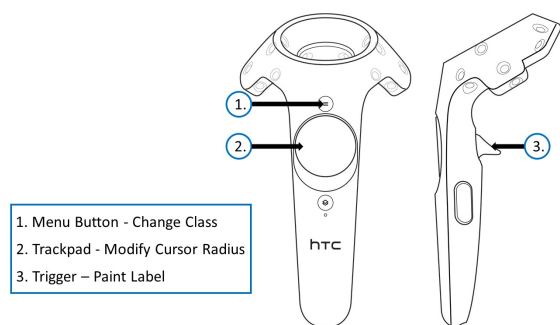Figure 5. Schematic of the VR functionnalities associated to the left controller.

Figure 6. Schematic of the VR functionnalities associated to the right controller.

In this section, preliminary feedback on the usability of the VR prototype tool as well as the possible improvements will be presented. Figure 4 illustrates the VR proof of concept application, Figure 5 and 6 show the links between different features in the virtual world with their corresponding controller buttons. This implementation relied on the most recent projects in the field such as PointCloudXR and Kharroubi et al., (2019). One notable observation is the spherical cursor used for the labelling. Although this solution was easy to implement and its applicability has already been proven in Wirth et al (2019) and Virtanen et al.(2020), the accuracy of the segmentation task is constrained by the volume of the sphere, leading to difficulties

for users labelling components of smaller size when using the spherical cursor, as it was confirmed during the labelling of bark pockets. In fact, it was very difficult to correctly label the bark pockets, despite the possibility of letting the user adjust the radius of the cursor between 0.5 and 1.5 cm. Consequently, a new shape for the labelling object should be developed to create a better user experience. To achieve this goal, inspiration could be taken from the diverse applications already existing in the field of virtual reality art, e.g. XR-painting.

Another relevant remark to improve the user friendliness of the app is the inclusion of text information. If the application requires the use of multiple labels, it can become difficult for the user to remember which colour corresponds to which class label. Furthermore, a GUI (Graphical User Interface) containing information about the tools available on the controller as well as the current size of the cursor and the current name of the label in use has the potential to improve the user friendliness of the proposed tool. Nevertheless, the VR app at this state is still in its early stages of development but has already shown a large potential in helping with the point cloud labelling task.

4. CONCLUSIONS

Here, we introduced a new VR application for the segmentation of biodiversity information from dense and large point clouds. The newly developed VR application can handle point clouds coming from various sources including files generated by classical photogrammetry with high level of detail in the texture. A proof of concept was carried out for the segmentation of single-tree stems using two types of TreMs: (1) bark pockets and (2) mosses and lichens. The results obtained were compared to a ground truth point cloud labelled using a traditional screen-based segmentation approach. This comparison demonstrated satisfactory quality and reinforced its possible application to generate AI training datasets.

Further work is required to improve remaining shortcomings. Our results are based on data from a single tree. The presented findings are therefore not yet statistically sound, although they already serve as a proof of concept. Repeating the experiment on a significantly larger dataset is planned in the near future to validate results presented in this paper. Furthermore, for the adoption of this VR labelling approach by the scientific community working on TreMs segmentation, it would be necessary to diversify the classes with more types of TreMs . However, the extension of classes would require the correct definition of the different categories as well. For this reason, a discussion between experts is planned in order to compile their knowledge into a standardised user guide or a well-defined usage protocol. Finally, a tutorial or an informative GUI (graphical user interface) would help users to understand the application and the various VR labelling functionalities at their disposal.

REFERENCES

Bütler, R., Lachat, T., Krumm, F., Kraus, D., Larrieu, L., 2020. Field guide to tree-related microhabitats descriptions and size limits for their inventory.

Calders, K., Newnham, G., Burt, A., Murphy, S., Raumonen, P., Herold, M., Culvenor, D., Avitabile, V., Disney, M., Armston, J., Kaasalainen, M., 2015. Nondestructive estimates of above-ground biomass using terrestrial laser
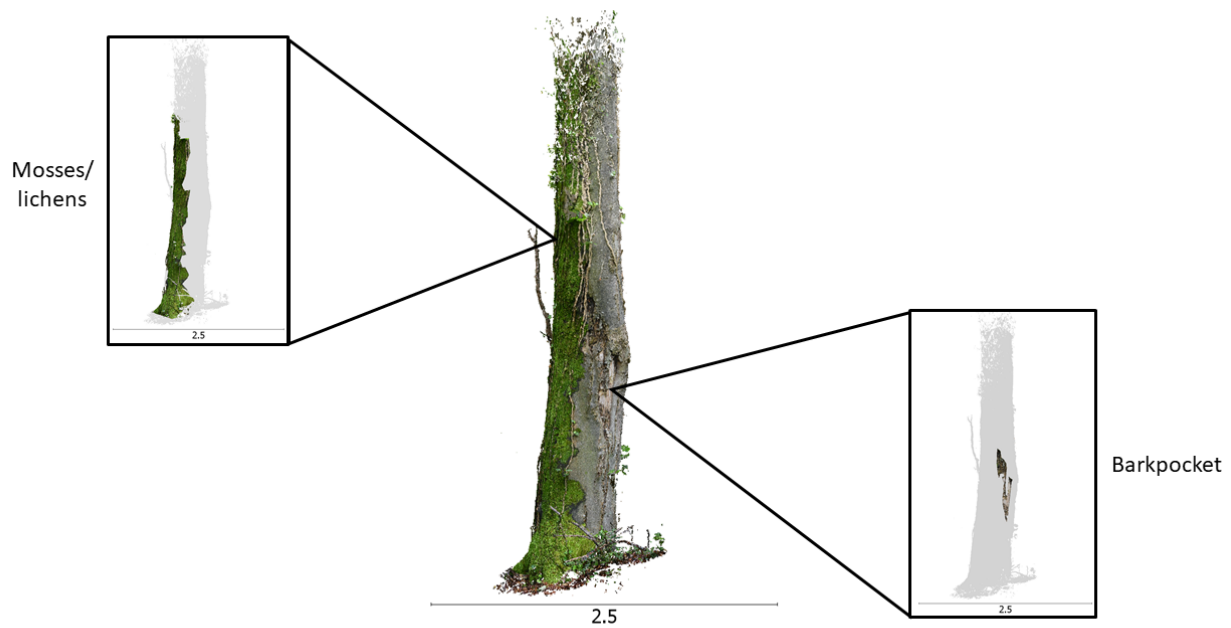
Figure 7. Different layers visible in VR for facilitating TreMs identification. Scale bar unit are in meters.

scanning. *Methods in Ecology and Evolution*, 6(2), 198-208. https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.12301.

Calders, K., Origo, N., Burt, A., Disney, M., Nightingale, J., Raumonen, P., Åkerblom, M., Malhi, Y., Lewis, P., 2018. Realistic Forest Stand Reconstruction from Terrestrial LiDAR for Radiative Transfer Modelling. *Remote Sensing*, 10(6). https://www.mdpi.com/2072-4292/10/6/933.

Fol, C., Murtiyoso, A., Griess, V., 2022. Evaluation of Azure Kinect Derived Point Cloud to Determine the Presence of Microhabitats on Single Trees Based on the Swiss Standard Parameters. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B2-2022, 989-994.

Kharroubi, A., Hajji, R., Billen, R., Poux, F., 2019. Classification and Integration of Massive 3D Points Clouds in a Virtual Reality (VR) Environment. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W17, 165–171. https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLII-2-W17/165/2019/.

Kükenbrink, D., Marty, M., Bösch, R., Ginzler, C., 2022. Benchmarking laser scanning and terrestrial photogrammetry to extract forest inventory parameters in a complex temperate forest. *International Journal of Applied Earth Observation and Geoinformation*, 113, 102999. https://www.sciencedirect.com/science/article/pii/S1569843222001881.

Lines, E. R., Fischer, F. J., Owen, H. J. F., Jucker, T., 2022. The shape of trees: Reimagining forest ecology in three dimensions with remote sensing. *Journal of Ecology*, 110(8), 1730-1745. https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/1365-2745.13944.

Mokroš, M., Liang, X., Surový, P., Valent, P., Čerňava, J., Chudý, F., Tunák, D., Saloň, , Merganič, J., 2018. Evaluation of close-Range photogrammetry image collection methods for estimating tree diameters. *ISPRS International Journal of Geo-Information*, 7. https://www.mdpi.com/2220-9964/7/3/93.

Poux, F., Valembois, Q., Mattes, C., Kobbelt, L., Billen, R., 2020. *Remote Sensing*, 12(16). https://www.mdpi.com/2072-4292/12/16/2583.

Rehush, N., Abegg, M., Waser, L. T., Brändli, U. B., 2018. Identifying tree-related microhabitats in TLS point clouds using machine learning. *Remote Sensing*, 10.

Schütz, M., Krösl, K., Wimmer, M., 2019. Real-time continuous level of detail rendering of point clouds. *2019 IEEE Conference on Virtual Reality and 3D User Interfaces*, IEEE, 103–110.

Virtanen, J.-P., Daniel, S., Turppa, T., Zhu, L., Julin, A., Hyyppä, H., Hyyppä, J., 2020. Interactive dense point clouds in a game engine. *ISPRS Journal of Photogrammetry and Remote Sensing*, 163, 375-389. https://www.sciencedirect.com/science/article/pii/S0924271620300691.

Wirth, F., Quehl, J., Ota, J., Stiller, C., 2019. Pointatme: Efficient 3d point cloud labeling in virtual reality. *2019 IEEE Intelligent Vehicles Symposium (IV)*, 1693–1698.

Zellweger, F., De Frenne, P., Lenoir, J., Rocchini, D., Coomes, D., 2019. Advances in Microclimate Ecology Arising from Remote Sensing. *Trends in Ecology Evolution*, 34(4), 327-341. https://www.sciencedirect.com/science/article/pii/S0169534718303045.

Zhao, J., Wallgrün, J. O., LaFemina, P. C., Normandeau, J., Klippel, A., 2019. Harnessing the power of immersive virtual reality - visualization and analysis of 3D earth science data sets. *Geo-spatial Information Science*, 22(4), 237-250. https://doi.org/10.1080/10095020.2019.1621544.