2020 Twelfth International Conference on Quality of Multimedia Experience (QoMEX)

Different Points of View: Impact of 3D Point Cloud Reduction on QoE of Rendered Images

Michael Seufert*, Julian Kargl*, Johannes Schauer†, Andreas Nüchter†, Tobias Hoßfeld*
*Chair of Communication Networks, University of Würzburg, Würzburg, Germany
†Chair of Robotics and Telematics, University of Würzburg, Würzburg, Germany
{michael.seufert|johannes.schauer|andreas.nuechter|tobias.hossfeld}@uni-wuerzburg.de, julian@kargl.net

***Abstract*—Modern photogrammetric methods as well as laser measurement systems make it easy to collect large 3D point clouds that sample objects or environments. As the recorded point clouds can be used to render computer-generated images and models, they are of particular interest in the domains of geographical and architectural engineering, as well as for computer graphics (e.g., games or virtual reality). However, point clouds have a huge storage demand, thus, point clouds shall be reduced by removing some of the points. This will inevitably also reduce the Quality of Experience (QoE) of media, which is rendered from the reduced point clouds. In this work, the impact of two different reduction methods on the QoE of rendered images is investigated from two point of views, i.e., based on ratings from both naive crowdworkers as well as point cloud experts.**

***Index Terms*—3D Point Clouds, Quality of Experience, Reduction, Image Quality, Crowdsourcing, Experts.**

I. Introduction

A 3D point cloud is a set of 3D points, where each point has its coordinate $(x, y, z) \in \mathbb{R}^3$ that may be accompanied with additional attributes, such as reflectance or color. Point clouds are generated from sensor data, i.e., from triangulation and Time-of-Flight (ToF). Triangulation is the underlying principle in stereo vision and structure-from-motion, where 3D points are generated by matching corresponding image points. Also sensors like the Microsoft Kinect (version 1) and other structured light scanners use triangulation. In contrast, ToF is often used by laser scanners, which emit a light and measure the reflection by either the phase shift of modulated light, or the direct ToF of a pulse. There are ToF cameras such as Microsoft Kinect (version 2) and LiDAR (light detection and ranging) systems, which include a mechanism for steering the laser over the object of interest. Sometimes, the mechanism includes mobile vehicles or even drones and aircraft. In this case, one talks about mobile or airborne mapping. The amount of data such LiDAR systems acquire is huge. Typical measurement rates in mobile or airborne measurement campaigns are in the order of 100k to 1M points per second which need to be stored, processed and visualized, i.e., users directly interact by viewing renderings of the 3D point clouds.

The majority of the sensors work in a spherical way, i.e., the measurement is done from a central source. It is important to note, that this applies to triangulation based systems as well as to ToF systems. The result is, that objects closer to the sensor are measured with higher point densities than objects further away. This fact implies a lot of potential for point cloud reduction and compression.

The question arises how the different reduction methods are influencing the user perceived quality of the 3D point clouds. A commonly accepted definition of Quality of Experience (QoE) in multimedia systems is provided in [1], which defines QoE as the degree of delight or annoyance of the user of an application or service. The definition highlights that QoE is influenced by expectations with respect to the utility and/or enjoyment of the application or service in the light of the user's personality and current state. For various multimedia services like speech, images, video streaming or gaming, the influence factors on QoE are investigated in literature with the ultimate goal to provide QoE models for those applications. Thereby, four different categories of influence factors are distinguished, which are influence factors on context, user, system, and content level [1]. In this 3D point cloud study, content is explicitly addressed, since it may have significant impact on QoE. To be more precise, for different types of contents under the same system influence factors, users may rate QoE differently, which has already been shown, e.g., for gaming [2] or video streaming [3]. As a consequence, different types of contents will be considered in our 3D point cloud QoE study. Further, the user level was explicitly considered, which includes psychological factors like expectations of the user, aesthetic perception, recency effects, or the usage history of the application. For this, our experiments were conducted with naive crowdsourcing users as well as 3D point cloud experts. This is especially interesting, as 3D point clouds cannot be considered mainstream yet, so crowdworkers have little to no previous experience with such kind of media.

To this end, a subjective user study was conducted with crowdsourcing users as well as expert users on the QoE of images rendered from 3D point clouds, which were compressed using two different point reduction techniques. Static images are explicitly selected in the subjective tests to avoid additional influences, e.g., due to the interactivity of users with current 3D point cloud applications. The subjects rated the quality of the images on a 5-point Absolute Category Rating (ACR [4]) scale. In addition, the subjects were asked to provide double stimulus difference ratings on a 5-point scale.

The main contribution of this work is the QoE analysis of the 3D point cloud reduction methods, which are essential in

978-1-7281-5965-2/20/$31.00 ©2020 IEEE

practice, especially when point clouds have to be efficiently stored or transmitted over the internet, or –more importantly– when viewed interactively using a 3D point cloud viewing tool, which only renders a fixed number of 3D points within a given framerate. All known 3D point cloud viewers reduce the number of points to be displayed, thus performing compression. First of all, the results show significant differences in the QoE ratings between experts and naive users due to the familiarity with point clouds, but also due to the aesthetic appeal of the test stimuli. Furthermore, our results show a strong impact of the contents on the QoE and interaction with the reduction method. Those results are fundamental for future studies, e.g., when using interactive 3D point cloud apps with point reduction.

The remainder of this paper is structured as follows. Section II provides background on 3D point cloud compression techniques as well as QoE studies on 3D point clouds. Section III presents the test stimuli, the crowdsourcing study and the expert study. Section IV presents the QoE results in terms of mean opinion scores as well as user rating distributions for the different contents and reduction methods. Finally, Section V concludes this paper.

II. Background and Related Work

Driven by the huge amount of airborne laser scan data, the American Society of Photogrammetry and Remote Sensing (ASPRS) created a simple binary exchange format, namely the LAS format [5]. The LAZ format, a compressor for LAS, is a widely-used lossless, non-progressive, order-preserving compressor for LiDAR measurements.

Some approaches for reduction of 3D point cloud employ special data structures such as k-d trees [6], [7] and octrees [8]–[11]. Both data structures are well-suited, since they also support other tasks like nearest-neighbor search, which is often needed for registration. Point clouds can be encoded with images. For example, Neci et al. present a method that exploits H.264 compression to reduce the size of the data stream from sensors such as the Kinect [12]. Depending on the sensors, the images might be panorama images [13], which store range information as pixel values, forming so-called range images. There is a one-to-one correspondence between range images and the resulting 3D point clouds. The images can be downscaled, and thus, the point clouds are reduced. Additionally, conventional image compression methods are applicable and in case of lossy compression combined with filtering. There is a reduction as well [13], especially in combination with resizing the range image. The different compression schemes have an impact on the 3D points [14], e.g., using the H.265/HEVC video compression results showed good visual quality with lossless video compression, while lossy compression introduced additional noise in projection images.

Several works have focused on the QoE of (reduced) point clouds. In [15], the quality of colored 3D point clouds under different resolution and noise was evaluated in a subjective study, but without reduction. The authors observed a linear correlation between resolution and human perception, and found that color had less impact than noise. The authors of [16] developed a framework for design, implementation, and evaluation of point cloud compression algorithms. In [17], a performance assessment methodology and benchmark of point cloud compression was presented based on the framework. Although no subjective study was conducted, a subjective testing procedure was proposed.

In [18], objective metrics and subjective feedback were compared for several point clouds of smaller objects, including point clouds altered by Gaussian noise and octree-pruned. The authors further investigated the same problem in [19] using augmented reality for the subjective survey. Also in [20], subjective evaluation methods and objective measures for point clouds were surveyed. Another specialized objective quality metric was proposed in [21], which is based on a local analysis of curvature.

In [22], the most similar work, also the quality of both an octree-based and a projection-based method is investigated. A cross-lab study was conducted on videos rendered from reduced point clouds. The authors conclude that projection-based reduction is inferior to octree-based reduction. Unfortunately, the authors used the projection-based reduction on datasets that this reduction method is ill-suited for. Just as in [22], here, the panorama reduction as a projection-based reduction method was used from 3DTK [23]. But instead of applying it on a fully registered scene with the scene center as the origin, it was applied on each individual scan with the scanner position as the origin, as this method was intended to be used.

III. Methodology

A. Test Stimuli

Given a large number of points from a laser scan, we propose to uniformly subsample the entire point cloud to reduce the number of points. This is achieved by first binning the point cloud in a regular 3D grid and then randomly selecting a fixed number of points in each voxel. Both the number of points and the side length of a voxel may be adjusted to allow for many different point densities. An additional advantage of the uniformity of the subsampling is that differences in density caused by the data acquisition process are reduced. Surfaces closer to the scanner are more densely sampled than surfaces further away. Selecting a fixed number of points from each voxel removes more points in voxels close to the scanner than further from the scanner. Afterwards, the points are uniformly distributed across the scanned object or environment.

The datasets were acquired using a Riegl VZ-400 terrestrial laser scanner with around 15 to 20 million points per individual 3D scan. The test stimuli "Chapel" and "Humans" come from a scan of a small chapel. The complete dataset is comprised of 11 individual scans with 194 million points overall. The test stimuli "Church" and "Text" come from a scan of a city center. The complete dataset is comprised of 215 million points in 13 scans. Both datasets combine multiple individual terrestrial scans into a larger dataset by registering them using 3DTK into a coherent point cloud using simultaneous localization and mapping (SLAM).

TABLE I: Comparison of Point Clouds and Reductions

Reduction	Church/Text Data Size	Church/Text Points	Chapel/Humans Data Size	Chapel/Humans Points
Original	10,346 MB	215,652,400	11,375 MB	194,754,633
OCV4	1,587 MB	35,985,142	769 MB	16,075,352
OCV8	657 MB	14,913,197	345 MB	7,196,991
OCV30	93 MB	2,113,011	89 MB	1,851,189
R3600x1000	1,629 MB	34,228,877	1,383 MB	29,008,210
R2400x667	727 MB	15,288,306	619 MB	12,983,283
R1200x333	183 MB	3,846,119	156 MB	3,275,927

Fig. 1: Original test stimuli. Chapel (top-left), Church (top-right), Human (bottom-left), Text (bottom-right)

Octree reduction was used on the dataset with different voxel sizes and random subsampling of points per voxel. The voxel sizes of 4, 8 and 30 cm were chosen, which corresponds to the reductions OCV4, OCV8 and OCV30, respectively. Projection-based panorama reduction was applied to the datasets, such that the resulting dataset sizes would roughly correspond to the Octree reduction. The reduction names R3600x1000, R2400x667, and R1200x333 correspond to a panorama size of 3600×1000, 2400×667 and 1200×333 pixels, respectively. The parameters were chosen to represent a wide variety of reductions and were applied to diverse perspectives, so that differences in more and less detailed data can be seen. See Table I for a numerical comparison of the different point clouds and reductions, and Figures 1 and 2 for a visual comparison of the generated test-stimuli.

B. Crowdsourcing Study

The online test framework was set up based on jsPsych[1], which is a JavaScript library for running behavioral experiments in a web browser. The framework was customized to follow the best practices of crowdtesting [24], and thus, included prior downloading of test content to the browser cache to exclude network influences, monitoring of test execution, and reliability checks. The study was advertised on the crowdsourcing platform Microworkers[2] for the top performers of the platform. The time consumption was roughly seven minutes and the monetary compensation for participation was US$ 0.25 upon completion of the test.

In the beginning, the participants were familiarized with the study and the test content by presenting the instructions and two exemplary images. One of the images was rendered from an original point cloud, and the other image showed the same content, but was rendered from a reduced point cloud. The participants were instructed to rate the difference in quality between the subsequently presented original and the reduced picture, and to not switch tabs or programs during the study. Moreover, they were instructed to memorize a striking red letter, which was included in the top right corner of each picture and which served as a simple check whether the participants had actually watched the presented image.

Fig. 2: Reduced test stimuli for Chapel. OCV4 (top-left), OCV8 (mid-left), OCV30 (bottom-left), R3600x1000 (top-right), R2400x667 (mid-right), R1200x333 (bottom-right)

After the instructions, rendered pictures were shown to the participants in blocks of four subsequent pages. Here, each block represented one test condition. The first page contained an image that was rendered from the original, uncompressed point cloud, a button to continue to the next page, and a small five second countdown. After the countdown expired, the button was activated, and the users could continue to the rating page, which asked the quality of the last picture on a 5-point Absolute Category Rating (ACR) scale [4] (bad, poor, fair, good, excellent). Additionally, the users had to specify the red letter that was contained in the last image. The third page also contained a single image, which had the same content as the original image on the first page, but this time, it was rendered from a reduced point cloud. After five seconds, the users could again click the button and continue to the final rating page of the block. On this page, the participants were asked to rate the difference between the original and the reduced image on a 5-point ordinal scale (no difference, small difference, visible difference, large difference, huge difference). Moreover, they were required to rate the quality of the reduced image on the 5-point ACR scale and to submit the displayed letter.

Every participant watched and rated ten blocks, which were randomly selected from the 28 test conditions in randomized order. Afterwards, participants were asked for a textual feedback about the study and the verification code was displayed, so that participants could earn the monetary compensation. In total, 337 users participated in the crowdsourcing study. 24 users were excluded because they had left the study tab more

[1]https://www.jspsych.org/ - Accessed: March 24, 2020.

[2]https://www.microworkers.com/ - Accessed: March 24, 2020.

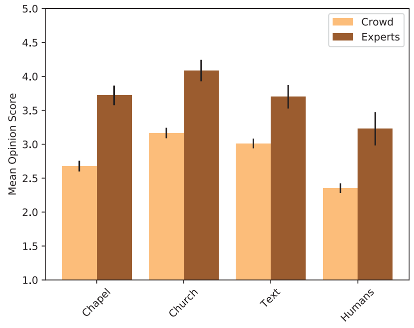

Fig. 3: Quality of unperturbed test stimuli (Original), i.e., images rendered from uncompressed point clouds

than four times, which was violating the task description, and 25 users were excluded because they had not recognized the red letter correctly more than four times. This results in 288 reliable users (85.46%). Moreover, those ratings were filtered out for which the tab was switched or the letter was incorrect. This results in 2386 (63.29%) remaining ratings. This gives on average 84.67 ratings per condition, having a minimum of 70 ratings and a maximum of 99 ratings per condition.

C. Expert Study

The same online study was advertised to a group of 45 manually selected point cloud experts, out of which 30 participated. A single user was excluded due to the letters. This results in 29 reliable users (96.67%). From the 290 ratings that were given by the reliable participants, again, ratings were filtered out for which the tab was switched or the letter was incorrect. This results in 232 (77.33%) remaining ratings. This gives on average 8.5 ratings per condition, having a minimum of five ratings and a maximum of 15 ratings per condition.

IV. Results

Before investigating the QoE of the images rendered from reduced point clouds, the internal validity of the crowdsourcing test is confirmed. First, it is evaluated whether repeated ratings of the same content lead to the same rating distribution. As the participants watched ten stimuli with four different contents, the four images rendered from the original point clouds were rated several times. The Mann-Whitney-U test is used to compare the distributions as it can be applied to the ordinal rating data [25]. No significant difference could be found between the first and second rating of the same original image having a p-value of $p = 0.84$. Further, as a consistency check, one original image was presented also as a reduced image, i.e., in this stimulus both the original and the reduced image were the same image. Again, no significant difference could be found ($p = 0.40$). Finally, fatigue effects were tested by comparing the rating distributions of the workers for the first two stimuli with the ratings of the last two stimuli. Again, no difference could be found ($p = 0.15$), which shows that the crowdsourcing study can be considered internally valid.

A. Differences between Naive Crowd and Experts

Figure 3 illustrates the overall quality of the test stimuli. Therefore, the mean opinion score (MOS) of the ratings and

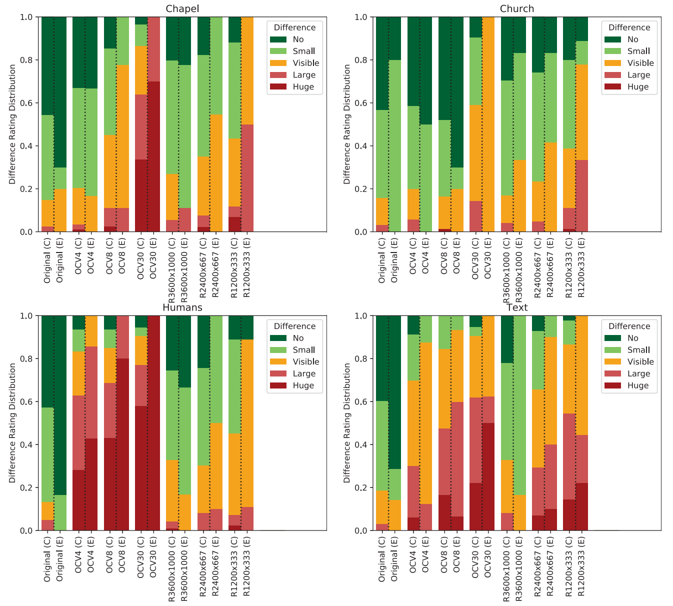

Fig. 4: Distributions of difference ratings

95% confidence intervals are shown for the unperturbed test stimuli, i.e., the images rendered from the four uncompressed point clouds. These images were already depicted in Figure 1 and will be referred to as “Original” in the remainder of this paper. The plot already suggests that there are significant differences between both the contents and the group of crowdworkers (yellow) and point cloud experts (brown). This was confirmed by comparing the ratings distributions for each content using the Kruskal-Wallis test [25], which gave a $p < 10^{-54}$. Moreover, the Mann-Whitney-U test confirmed that the experts rated the quality of the original images significantly better than the naive crowd ($p < 10^{-38}$). This still holds when considering not just the original images, but all presented stimuli ($p < 10^{-9}$). This is interesting as it is quite unusual in crowdsourcing. The reason could be that the point cloud experts are more familiar with the kind of images, which were used as test stimuli. In contrast, naive crowdworkers might find the unfamiliar images of the (compressed) 3D point clouds unnatural and less aesthetic due to the dark and grayscale style of the images, as was indicated by some of their comments submitted at the end of the test. This relationship of aesthetic perception and QoE is in line with previous results for visual stimuli [26] and web QoE [27], [28], which found a positive correlation between aesthetic perception and QoE.

Figure 4 shows the distributions of the difference ratings for each of the four contents, i.e., Chapel (top-left), Church (top-right), Humans (bottom-left), and Text (bottom-right), as a stacked bar plot. Here, the colors indicate the extent of the difference between the original and the reduced images as perceived by the participants on a five point scale, ranging from no difference (dark green) to visible difference (yellow) to huge difference (dark red). The x-axes show all test stimuli. For each stimulus, two bars are displayed, which are separated by a dashed line. The left bar shows the rating distribution

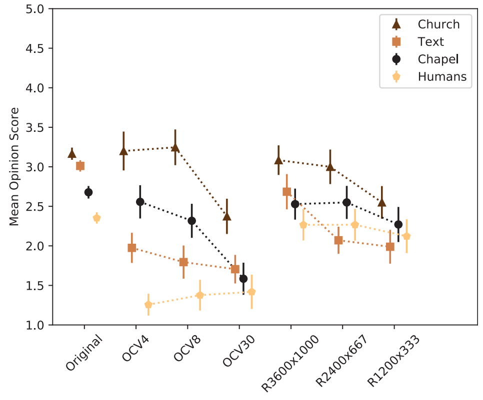

Fig. 5: Rating trends of crowd in terms of MOS.

of the crowd, while the right bar shows the corresponding rating distribution of the experts. It can be seen that the expert distribution is significantly shifted towards larger differences for all reduced stimuli of Chapel ($p < 10^{-3}$) and Humans ($p = 0.02$). This is expected since experts are familiar with point clouds, and thus, can more easily detect differences and are more discriminating. However, this cannot be observed for the reduced stimuli of Church ($p = 0.15$) and Text ($p = 0.08$), which clearly shows that the visibility of differences between original and reduced point clouds depends also on the content.

B. QoE Impact of Reduction Method

Figure 5 visualizes the general rating trends of the crowd. It shows the different reduction conditions on the x-axis, and the corresponding MOS and 95% confidence intervals on the y-axis. Different colors indicate the contents, and the reduction conditions of each method (Octree and Panorama) are connected with a dotted line to identify the trends more easily. It can be seen that the rating trends are generally decreasing, which is expected as an increasing point cloud reduction causes increasing quality degradation. However, the trends of the Panorama method do not drop as much as the trends of the Octree method. Moreover, the plot suggests that there is a strong content dependency for the different reduction levels. In the following, these hypotheses are investigated in full detail based on the rating distributions of each stimulus.

Figure 6 shows these rating distributions for all stimuli as a stacked bar plot. Similar to Figure 4, they are subdivided into four subplots by content, i.e., Chapel (top-left), Church (top-right), Humans (bottom-left), and Text (bottom-right), and for each stimulus, two bars are separated by a dashed line, which show both the rating distribution of the crowd (left) and the experts (right). As was shown above, experts rated the quality significantly better than the naive crowd, however, the general trends are confirmed by the expert ratings, which can be seen in Figure 6. Thus, as much more data was gathered from the crowd, only these results will be evaluated and discussed.

The rating distributions of the crowd show that reductions cause ratings to be shifted towards lower categories, which can be confirmed by Kruskal-Wallis tests for both Octree ($p < 10^{-82}$) and Panorama ($p < 10^{-24}$) method. Also

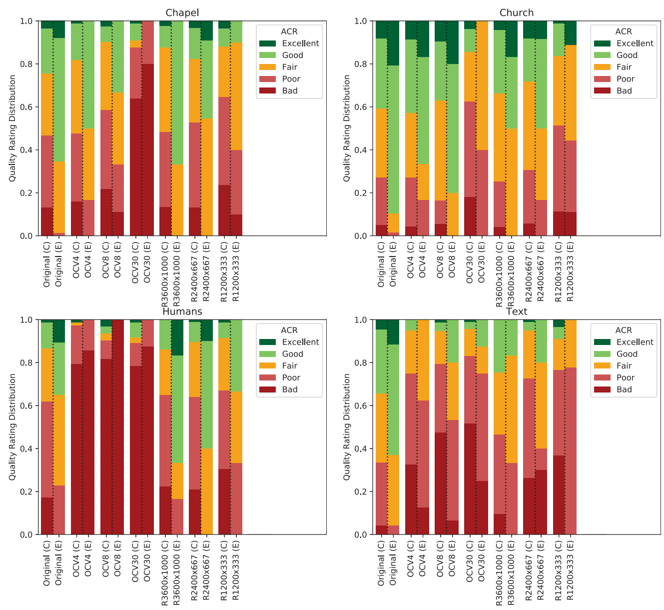

Fig. 6: Distributions of quality ratings

pairwise Mann-Whitney-U tests find significant differences between all reduction conditions of each method, except between the ratings of OCV4 and OCV8 ($p = 0.08$).

Looking at the Octree methods in detail for each content, two pairs of contents can be observed. For both, Human and Text only the first reduction from Original to OC4 reduces the ratings, while the rating distributions of OC4, OC8, and OC30 are not significantly different ($p = 0.87$ for Humans and $p = 0.053$ for Text). In contrast, for Chapel, OCV4 is not rated significantly worse than Original ($p = 0.19$), however, each further reduction reduces the quality significantly. For Church, there are no significant differences in the rating distributions of Original, OC4, and OC8 ($p = 0.83$), and only OC30 shows significantly lower quality. The overall impression is that already the first reduction level (OC4) visibly degrades the quality of the images for most of the contents. While OC4 and OC8 are on a par, the final reduction of OCV30 can again introduce another quality degradation.

For the Panorama method, again Chapel and Church show similar rating behavior. For both contents, Original, R3600x1000, and R2400x667 are rated similarly ($p = 0.30$ for Chapel and $p = 0.23$ for Church), and only R1200x333 has a significantly worse quality than all other reduction conditions. For Humans, the Kruskal-Wallis test cannot reject the hypothesis that all Original and all reduction conditions are rated similarly ($p = 0.17$), and also the pairwise Mann-Whitney-U tests only find a significantly lower ratings for R1200x333 compared to Original ($p = 0.02$). Finally, for Text, each reduction reduces the quality, except for the last. Here, the rating distribution of R1200x333 is not significantly worse than R2400x667 ($p = 0.12$). These results showed that, in contrast to Octree, the first reduction level is mostly not perceived as a quality degradation compared to the Original. Also further reductions do not degrade the quality as severe as with Octree.

When comparing Octree and Panorama directly over all ratings, Octree images received significantly lower quality ratings ($p < 10^{-21}$), which can also be observed when comparing the ratings within a quality level, i.e., OCV4 and R3600x1000 ($p < 10^{-7}$), OCV8 and R2400x667 ($p < 10^{-5}$), and OCV30 and R1200x333 ($p < 10^{-10}$). However, again, a strong content dependency can be observed. While Octree is rated worse on every level for Human and Text, this is not the case for Chapel and Church. Interestingly, for Church, OCV8 is rated significantly better than R2400x667 ($p = 0.04$), however, the quality difference is comparably small. For all other levels of Chapel and Church, Octree and Panorama were rated similarly, except for the most reduced level of Chapel, i.e., OCV30 is significantly worse than R1200x333 ($p < 10^{-6}$). To sum up, this shows that while the methods can perform on a par for some contents, overall, Panorama should be preferred as it results in equal or higher quality ratings compared to Octree.

V. Conclusion

This paper investigated the impact of 3D point cloud reduction on the QoE of images, which were rendered from these reduced point clouds. Therefore, a subjective study was conducted online with both naive crowdworkers as well as point cloud experts. After investigating the rating behavior of the participants, the study could be considered to provide internally valid ratings.

The results of the study showed that experts saw larger differences between the original and the reduced point clouds, which was expected as experts are more familiar with the kind of images, which were used as test stimuli. However, experts rated the quality of the images better than the crowd, which is quite unusual in crowdsourcing. The reason could be that naive crowdworkers found the unfamiliar images of the (compressed) 3D point clouds unnatural and less aesthetic due to the dark and grayscale style of the images, as was indicated by some of their comments submitted at the end of the test.

For the Octree reduction method, it could be observed that already the first reduction level visibly degraded the quality of the images for most of the contents, and the final reduction of OCV30 could again introduce another quality degradation. In contrast, for the Panorama method, the first reduction level was mostly not perceived as a quality degradation compared to the original. Also further reductions did not degrade the quality as severe as with Octree. For all evaluations, a strong content dependency could be observed, which might be due to the different visibility of artifacts. In future work, we plan to integrate and evaluate the MPEG methods V-PCC and G-PCC.

The results of this paper are fundamental not only for images, but also when presenting and watching other media generated from reduced point clouds, such as videos, e.g., from the church dataset[3], or interactive 3D point cloud apps. However, the detailed impact of 3D point cloud reduction on such media has to be investigated in future studies.

[3] https://youtu.be/DKNauiHKKhk - Accessed: March 24, 2020.

References

[1] P. Le Callet, S. Möller, and A. Perkis (eds), "Qualinet White Paper on Definitions of Quality of Experience," COST Action IC 1003 (Qualinet), Tech. Rep., 2013, version 1.2.

[2] M. Jarschel, D. Schlosser, S. Scheuring, and T. Hoßfeld, "Gaming in the Clouds: QoE and the Users' Perspective," *Mathematical and Computer Modelling*, vol. 57, no. 11-12, 2013.

[3] M. Seufert, S. Egger, M. Slanina, T. Zinner, T. Hoßfeld, and P. Tran-Gia, "A Survey on Quality of Experience of HTTP Adaptive Streaming," *Communications Surveys & Tutorials*, vol. 17, no. 1, 2015.

[4] International Telecommunication Union, "ITU-T Recommendation P.910: Subjective Video Quality Assessment Methods for Multimedia Applications," 2008.

[5] American Society for Photogrammetry and Remote Sensing, "Common Lidar Data Exchange Format," http://www.asprs.org/Committee-General/LASer-LAS-File-Format-Exchange-Activities.html (Accessed: March 24, 2020), 2011.

[6] O. Devillers and P. M. Gandoin, "Geometric Compression for Interactive Transmission," in *IEEE Visualization*, 2000.

[7] P.-M. Gandoin and O. Devillers, "Progressive Lossless Compression of Arbitrary Simplicial Complexes," *ACM Trans. Graph.*, vol. 21/3, 2002.

[8] R. Schnabel and R. Klein, "Octree-based Point-Cloud Compression," in *Symposium on Point-Based Graphics*, 2006.

[9] J. Peng and C.-C. J. Kuo, "Octree-based Progressive Geometry Encoder," in *SPIE Internet Multimedia Management Systems IV*, 2003.

[10] Y. Huang, J. Peng, C.-C. J. Kuo, and M. Gopi, "Octree-based Progressive Geometry Coding of Point Clouds," in *3rd Eurographics / VGTC Conf. on Point-Based Graphics*, 2006.

[11] A. Hornung, K. M. Wurn, M. Bennewitz, C. Stachniss, and W. Burgard, "OctoMap: An Efficient Probabilistic 3D Mapping Framework Based on Octrees," in *Autonomous Robots*, 2013.

[12] F. Nenci, L. Spinello, and C. Stachniss, "Effective Compression of Range Data Streams for Remote Robot Operations Using H.264," in *IROS*, 2014.

[13] H. Houshiar and A. Nüchter, "3D Point Cloud Compression Using Conventional Image Compression for Efficient Data Transmission," in *ICAT*, 2015.

[14] E. Dumic, A. Bjelopera, and A. Nüchter, "Projection Based Dynamic Point Cloud Compression Using 3DTK Toolkit and H.265/HEVC," in *SmaGriMet*, 2019.

[15] J. Zhang, W. Huang, X. Zhu, and J.-N. Hwang, "A Subjective Quality Evaluation for 3D Point Cloud Models," in *ICALIP*, 2014.

[16] R. Mekuria and P. Cesar, "MP3DG-PCC, Open Source Software Framework for Implementation and Evaluation of Point Cloud Compression," in *ACM MM*, 2016.

[17] R. Mekuria, S. Laserre, and C. Tulvan, "Performance Assessment of Point Cloud Compression," in *VCIP*, 2017.

[18] E. Alexiou and T. Ebrahimi, "On Subjective and Objective Quality Evaluation of Point Cloud Geometry," in *QoMEX*, 2017.

[19] E. Alexiou, E. Upenik, and T. Ebrahimi, "Towards Subjective Quality Assessment of Point Cloud Imaging in Augmented Reality," in *MMSP*, 2017.

[20] E. Dumic, C. R. Duarte, and L. A. da Silva Cruz, "Subjective Evaluation and Objective Measures for Point Clouds - State of the Art," in *SmaGriMet*, 2018.

[21] G. Meynet, J. Digne, and G. Lavoué, "PC-MSDM: A Quality Metric for 3D Point Clouds," in *QoMEX*, 2019.

[22] L. A. da Silva Cruz *et al.*, "Point Cloud Quality Evaluation: Towards a Definition for Test Conditions," in *QoMEX*, 2019.

[23] A. Nüchter *et al.*, "3DTK – The 3D Toolkit," http://slam6d.sourceforge.net/ (Accessed: March 24, 2020), 2011.

[24] T. Hoßfeld *et al.*, "Best Practices and Recommendations for Crowdsourced QoE - Lessons learned from the Qualinet Task Force Crowdsourcing," COST Action IC 1003 (Qualinet), Tech. Rep., 2014.

[25] M. Seufert, "Fundamental Advantages of Considering Quality of Experience Distributions over Mean Opinion Scores," in *QoMEX*, 2019.

[26] W. A. Mansilla, A. Perkis, and T. Ebrahimi, "Implicit Experiences as a Determinant of Perceptual Quality and Aesthetic Appreciation," in *ACM MM*, 2011.

[27] M. Varela, L. Skorin-Kapov, T. Mäki, and T. Hoßfeld, "QoE in the Web: A Dance of Design and Performance," in *QoMEX*, 2015.

[28] S. Baraković and L. Skorin-Kapov, "Multidimensional Modelling of Quality of Experience for Mobile Web Browsing," *Computers in Human Behavior*, vol. 50, pp. 314–332, 2015.