Jannis Weil, Yassin Alkhalili, Anam Tahir, Thomas Gruczyk, Tobias Meuser, Mu Mu, Heinz Koeppl, Andreas Mauthe. *Modeling Quality of Experience for Compressed Point Cloud Sequences based on a Subjective Study*. In Proceedings of the 15th International Conference on Quality of Multimedia Experience (QoMEX), 2023, pp. 135-140.

Modeling Quality of Experience for Compressed Point Cloud Sequences based on a Subjective Study

Jannis Weil[1]*, Yassin Alkhalili[1]*, Anam Tahir[1]*, Thomas Gruczyk[1], Tobias Meuser[1], Mu Mu[2], Heinz Koeppl[1], Andreas Mauthe[3]

[1]Technical University of Darmstadt, [2]University of Northampton, [3]University of Koblenz
*Equal contribution {jannis.weil, yassin.alkhalili, anam.tahir}@tu-darmstadt.de

***Abstract*—There is growing interest in point cloud content due to its central role in the creation and provision of interactive and immersive user experiences for extended reality applications. However, it is impractical to stream uncompressed point cloud sequences over communication networks to end systems because of their high throughput and low latency requirements. Several novel compression methods have been developed for efficient storage and adaptive delivery of point cloud content. However, these methods primarily focus on data metrics and neglect the influence on the actual Quality of Experience (QoE). In this paper, we conduct a user study with 102 participants to analyze the QoE of point cloud sequences and develop a QoE model that can enhance the quality of point cloud content distribution under dynamic network conditions. Our analysis is based on user opinions regarding two representative point cloud sequences, three different frame rates, three viewing distances, and two state-of-the-art point cloud compression libraries, Draco and V-PCC. The results indicate that the proposed models can accurately predict the users' quality perception, with frame rate being the most dominant QoE factor.**

***Index Terms*—point clouds, subjective user study, quality of experience, machine learning**

I. Introduction

Ultra-high-definition video content and modern end devices are making it possible to experience 2D videos in excellent quality. The next wave of innovation in multimedia technology is expected to enable fully immersive experiences with increased interactivity and a higher degree of freedom (DoF) compared to video content [1]. In the context of extended reality, DoF refers to the number of ways in which a user can interact with and move within a virtual environment [2]. Content types that provide a higher DoF usually require higher data rates and come with many challenges, such as efficient storage, transmission, and visualization [3].

Point clouds and meshes are both commonly used to represent 3D data in computer graphics and enable immersive applications with a higher DoF [4]. A point cloud is a set of points representing an object's surface or environment, often produced by 3D scanning devices such as LiDAR or synchronized RGB-D cameras. Each point has x, y, and z coordinates and may carry additional attributes such as color, surface normal, or intensity information. A high-quality 3D representation using a point cloud of a complex object, such as a human figure, requires a high density cloud of points [3]. The sheer size and the level of interaction of point cloud data create unique challenges for modern communication systems. The primary challenges are related to handling and transmitting large quantities of data efficiently, as well as adjusting the data rate to accommodate fluctuating network conditions. As a consequence, efficient compression techniques, such as MPEG point cloud compression based on video (V-PCC) or geometry (G-PCC), have been developed to facilitate the storage and transmission of point cloud data [5].

979-8-3503-1173-0/23/$31.00 ©2023 IEEE

(a) (b) (c) (d) (e)

Fig. 1. Example visual views of the point cloud objects used in this study. Figures (a), (b), and (c) show the *Dancer* point cloud at near, medium, and far distances, respectively. Figures (d) and (e) show the *Thaidancer* point cloud encoded using V-PCC with m5 and mb quantization parameters, respectively.

Despite the growing interest in point cloud streaming, there is limited research on optimizing the content distribution of point cloud data for the best QoE. In this paper, we design and conduct a PC-based subjective user study to examine how users perceive quality degradations introduced by different point cloud compression techniques. Specifically, we study how variations in quantization level, frame rate, and the objects' distance from the camera affect the perceived quality of two point cloud sequences of human figures. Overall, we experiment with five levels of degradation per encoding method, three camera distances, and three distinct frame rates. We investigated two state-of-the-art point cloud compression techniques, V-PCC by MPEG [5] and Draco by Google [6]. The perceptual quality of the compressed point clouds is subjectively evaluated by 102 participants using an Absolute

Jannis Weil, Yassin Alkhalili, Anam Tahir, Thomas Gruczyk, Tobias Meuser, Mu Mu, Heinz Koeppl, Andreas Mauthe. *Modeling Quality of Experience for Compressed Point Cloud Sequences based on a Subjective Study*. In Proceedings of the 15th International Conference on Quality of Multimedia Experience (QoMEX), 2023, pp. 135-140.

Category Ranking (ACR) scale. Based on the users' assessment, we developed QoE models for point cloud sequences using common machine learning algorithms. The evaluation indicates that the proposed models can accurately capture and predict the users' quality perception. The results of our study are available at https://github.com/jw3il/point-cloud-qoe.

The remainder of the paper is organized as follows. In the next section, we introduce related work. In Section III, we describe our user study setup and the involved parameters. In Section IV, we analyze the results of the user study, and in Section V, we model and evaluate QoE models with various machine learning algorithms. Lastly, we give the conclusion.

II. Related Work

A recent summary of previous research on subjective evaluation and objective metrics for point clouds is provided by Dumic et al. [7]. Furthermore, there exist extensive surveys on the use of machine learning models for QoE assessment and prediction [8], [9]. In the following, we describe selected works on QoE assessment in more detail.

Van der Hooft et al. [10] evaluated the QoE of volumetric 6DoF streaming using objective and subjective methods. They investigated the impact of different factors on the perceived quality of PCC-DASH, a method for adaptive streaming of scenes with multiple dynamic point cloud objects over HTTP. Their study considers the available bandwidth, different rate adaptation algorithms, viewport prediction strategies, and user motion. They conclude that there is a need for more representative metrics and QoE models for point clouds sequences. This also applies to static point clouds [11].

A recent study [12] assessed the QoE of point clouds in an immersive 6DoF environment, specifically examining the impact of geometry and texture parameters on a compression mechanism. The authors proposed objective evaluation metrics and showed that they improve the correlation with the subjective user scores compared to previous works.

Zerman et al. [13] studied the perceptual quality of compressed 3D sequences using point clouds and mesh-based representations. Their experiments suggest that meshes provide superior visualization for high-bandwidth applications, while point clouds also allow for good QoE with limited bandwidth. This was also confirmed by Cao et al. [14], who compared the perceptual quality of point cloud and mesh-based representations with respect to the available bitrate and observation distance to the viewer. They found that the perceived quality increases with higher distances and proposed a model to estimate the quality using bit rate quality and distance correction factors. Both works considered a fixed frame rate in their experiment setups.

While many works investigate the effect of different quality settings on the perceived quality, a study on the joint impact of varying frame rates and viewing distance levels on point cloud QoE is missing. In this work, we analyze the joint effect of multiple factors on the QoE of point cloud sequences, namely frame rate, quantization level, and distances.

III. Experiment Setup

This section describes the setup of our user study, the chosen independent variables, and the corresponding parameters.

A. Selected Point Cloud Sequences

We selected two naturalistic full-body human figures from publicly available data sets as the objects of our study, namely *Dancer* and *Thaidancer*. These objects have stark differences in color, texture and movement dynamics, allowing to investigate how much object differences affect the perceived quality.

Fig. 1(a) shows the *Dancer* from the Owlii dynamic human mesh sequence dataset [15] when the distance parameter is set to near. The corresponding point cloud sequence was recorded at 30 frames per second (fps) and consists of 600 frames, each containing approximately 2.6 million points. Fig. 1(d) shows the second object, which is the *Thaidancer* from the 8i Voxelized Surface Light Field dataset [16] at near distance. This point cloud sequence was captured at 30 fps and consists of 300 frames, each containing more than 3 million points. Each point of the point cloud sequences contains 3D coordinates and red, green, and blue color channels. As commodity hardware like RealSense LiDAR and RGB-D cameras does not directly return normals, we remove all attributes besides point position and color information before compression.

B. Point Cloud Compression and Video Generation

We consider two state-of-the-art point cloud compression methods, V-PCC [5] and Draco [6], [17]. V-PCC projects point clouds onto frames and leverages an advanced 2D video codec, while Draco achieves compression through two primary stages: 1) quantizing the information that constitutes each individual point and 2) conducting mesh compression. V-PCC was selected instead of the geometry-based G-PCC because of its higher coding efficiency for dense point clouds [5].

Both compression methods can be customized, resulting in file size and quality variations. In our study, we have defined five quantization levels for each compression method applied to the raw point cloud sequences. These levels are labelled as (mb, m0, m1, m3, m5) for V-PCC and (d8, d9, d10, d11, d16) for Draco. To obtain different levels of V-PCC compression, we modify the geometry quantization parameter g, attribute quantization parameter a, and occupancy precision o parameter. The levels m0, m1, m3, m5 are equivalent to the standard V-PCC levels r0, r1, r3, r5 as described in [18]. The degraded level mb was generated with $g = 51$, $a = 51$, and $o = 4$. Fig. 1(d) shows the *Thaidancer* object compressed using V-PCC to the quantization level of m5 and Fig. 1(e) is for level mb. For Draco, only the position quantization parameter, p, was set to the desired quantization level. For example, to achieve the d8 quantization level, p was set to 8.

Each frame of the point cloud sequences is encoded at each quantization level. The frames were encoded independently for each method due to the high sequence encoding time for V-PCC and the lack of built-in support for temporal compression in Draco. Real-time playback of sequences compressed with V-PCC was not feasible on commodity hardware. Thus,

Jannis Weil, Yassin Alkhalili, Anam Tahir, Thomas Gruczyk, Tobias Meuser, Mu Mu, Heinz Koeppl, Andreas Mauthe. *Modeling Quality of Experience for Compressed Point Cloud Sequences based on a Subjective Study*. In Proceedings of the 15th International Conference on Quality of Multimedia Experience (QoMEX), 2023, pp. 135-140.

we used lossless encoding to render the point cloud sequences from fixed camera perspectives and stored the results as H.264 video files. Note that the rendering process itself did not add compression loss or any other artifacts to the content. In addition to the five quantization levels, we rendered the point cloud sequences at three different distances: near, medium, and far. Due to the perspective projection, the point cloud sequence appears smaller at greater distances. Figures 1 (a), (b), and (c) show the same object at near, medium, and far distances, respectively. For V-PCC, we also explored different frame rates, rendering the point cloud object with frame rates of 30, 15, and 10 fps. For Draco, the frame rate was fixed at 30 fps, resulting in 120 different experiment configurations. The intention behind limiting the number of configurations was to maintain the participants' engagement throughout the experiment, improving the reliability of our findings.

C. Execution of the User Study

In order to assess the quality of the generated videos using a large number of participants, we embedded them in a webpage-based survey. The user study was run offline and locally on each participant's computer to minimize the impact of video buffering delay. Participants were given the instructions to download and extract an archive containing all our study files. Each participant started the experiment by opening the webpage using a browser. This webpage provided a step-by-step guideline for completing the experiment.

The quality assessment was split into two parts, one for each point cloud object. Each part contained 45 videos according to our experiment configurations (encoding, quantization level, distance, and frame rate). The order of the objects and the order of presented configurations for each object was random for each participant. At the beginning of each part, the participant was shown the ideal reference video generated using uncompressed point cloud data at 30 fps and near distance. After watching a point cloud video for at least 3 seconds, the participants could choose one of the five quality levels {bad, poor, fair, good, excellent} with their corresponding score $\{1, 2, 3, 4, 5\}$, respectively. Once a quality level was selected, the survey advanced to the next video. The data was collected locally, and at the end of the survey, the results of the study were saved in the participant's download directory as a JSON file. We provided a link to anonymously upload these results to our servers. Participation in the study required approximately 15 to 20 minutes for the entire process. We used the Prolific[1] crowdsourcing tool and had 102 users whose participation was incentivized with around 5€ each.

IV. User Study Results

In this section, we first provide an analysis of the results of our user study and then focus on selected configurations in the following subsections. Fig. 2 shows the MOS for each experiment configuration in ascending order at the bottom, along with the corresponding independent variables in the subplots

[1]https://www.prolific.co/ [Accessed in February 2023]

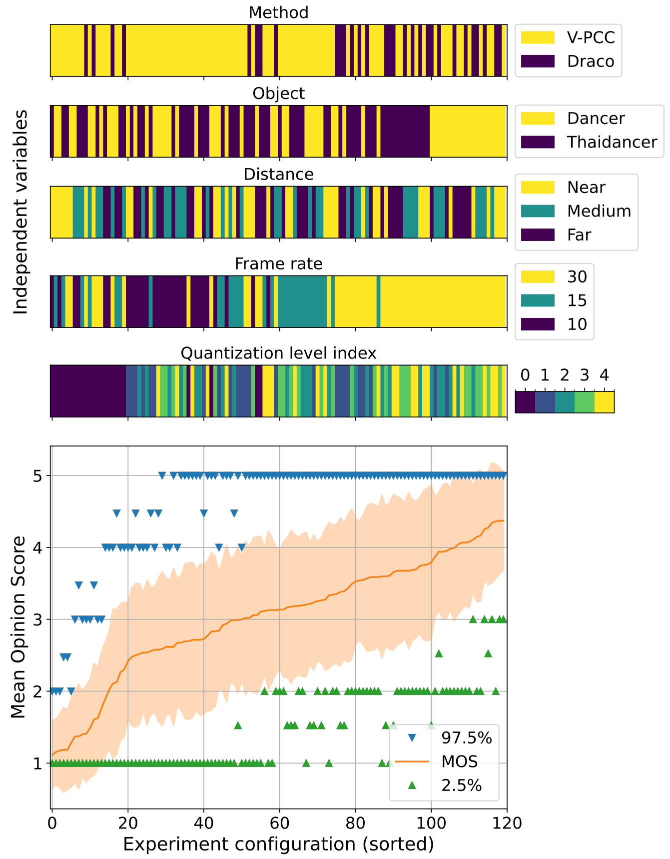

Fig. 2. Mean Opinion Score (MOS) per experiment configuration, sorted in ascending order. The shaded area indicates the standard deviation of all scores for each configuration. The markers show the 97.5% and 2.5% quantiles.

above. The minimum and maximum MOS are 1.12 and 4.37, respectively. The first twenty configurations use the lowest quality level. In these configurations it can be noticed that the MOS strongly increases with higher distance and higher frame rates. Beyond configuration 20, the increase in MOS is approximately linear. The frame rate has a noticeable impact on the MOS. All experiments in the upper half either use 15 or 30 fps, the latter predominately leads to higher scores. The effect of the distance is inverted, as a higher distance appears to lead to a lower MOS. The participants generally assigned higher scores to the *Dancer* object compared to the *Thaidancer* object, as seen in configurations 100 to 119. The configurations with Draco are predominantly in the upper half of the experiment configurations since only 30fps was used, and therefore have a comparatively high MOS.

A. V-PCC Compression Format

Fig. 3 visualizes the aggregated MOS for both point cloud sequences using the V-PCC compression format. Although individual ratings of the object's quality are in the interval $[1, 5]$, we observe that the maximum MOS over all configurations is 4.06 (near, 30 fps, m5). As all participants have seen the ideal reference video, this could indicate that they are either hesitant to select the highest rating 5 due to hardly distinguishable quality or that the content itself had an underwhelming visual

Jannis Weil, Yassin Alkhalili, Anam Tahir, Thomas Gruczyk, Tobias Meuser, Mu Mu, Heinz Koeppl, Andreas Mauthe. *Modeling Quality of Experience for Compressed Point Cloud Sequences based on a Subjective Study*. In Proceedings of the 15th International Conference on Quality of Multimedia Experience (QoMEX), 2023, pp. 135-140.

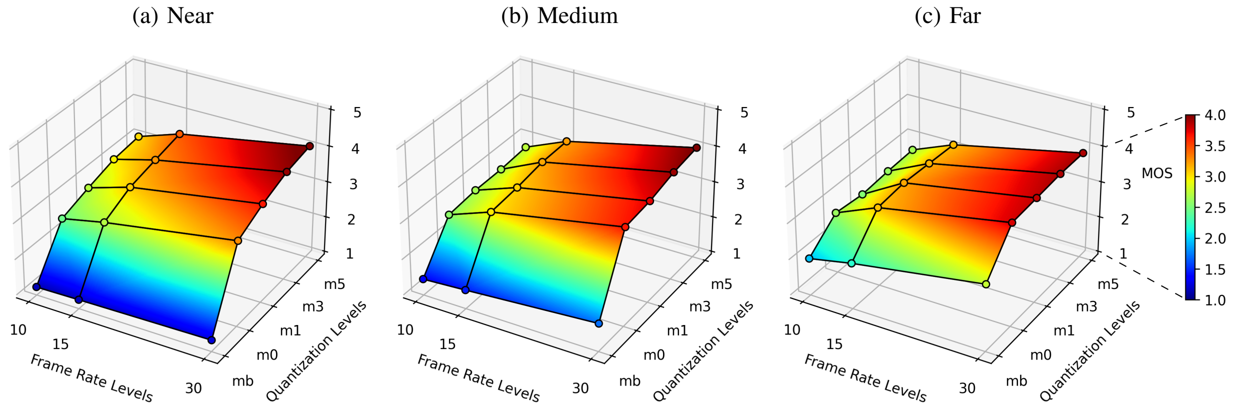

Fig. 3. Aggregated MOS for the V-PCC compression format with distance levels: (a) near, (b) medium, and (c) far. Each subplot shows the MOS over both objects for individual frame rate and quantization levels. The bilinearly interpolated color intensity visualizes the MOS in the interval $[1, 4]$.

fidelity. When the users are close to the object in Fig. 3 (a), the MOS monotonously increases with the quantization level. This is consistent with previous studies [13], [14]. It can be observed that the MOS increases with higher frame rates for quantization levels m0 to m5, whereas it has little effect on the lowest quantization level mb.

For a medium distance shown in Fig. 3 (b), the surface starts to flatten. The maximum MOS is 4.02 at (medium, 30 fps, m5), and the MOS of most quantization levels at 30 fps are similar to the near setting, except for slightly higher values for the lower mb and m0 quantization levels. Note that the effect of the frame rate is higher than in the near setting. The lower frame rates 10 and 15 in the medium distance setting lead to a noticeable lower MOS than in the near setting.

The surface flattens further with far objects in Fig. 3 (c). High quantization levels result in a lower MOS than in the other two distance settings, with a maximum of 3.87 at (far, 30 fps, m5). The mb and m0 quantization levels primarily lead to a higher MOS than for the near or medium distance levels. The results for quantization levels m1 to m5 are similar to the medium distance level for a fixed frame rate.

Our results indicate that improvements and degradations in an object's encoding quality become less noticeable when it is further away. In contrast, different frame rates are still noticeable and clearly affect the MOS. For example, increasing the frame rate of a far object with quantization level m0 from 10 to 30 has a much higher positive effect on the MOS than increasing its quantization level to m5. These findings can help optimize future point cloud streaming strategies when the current distance to the viewers is known.

B. Draco Compression Format

We now provide the results for Draco and compare them to V-PCC. All Draco configurations used 30 fps and the resulting MOS for different quantization levels is illustrated in Fig. 4 (a). We observe a similar trend as for V-PCC. When the viewing distance increases, users are less likely to give either high or low scores. Hence the impact of the quantization is reduced.

The comparison of Draco with V-PCC at 30 fps in Fig. 4 (b) shows that at the first quantization level d8-mb, Draco

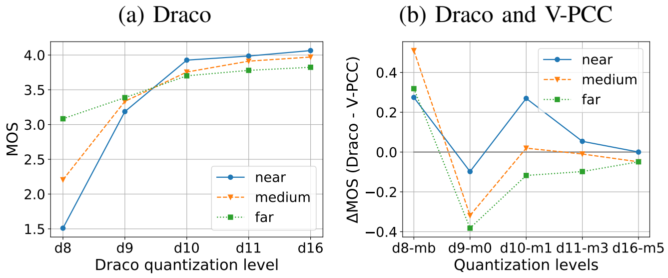

Fig. 4. Overview on Draco MOS with a) aggregated MOS for both objects using the Draco compression format on the distance levels near, medium and far for 30 objects per second and b) MOS difference between Draco and V-PCC quality settings using the same configurations. A value greater zero indicates that Draco received a higher MOS.

achieved a much higher MOS than V-PCC, while V-PCC shows better results at the the following quantization levels. The difference in MOS approaches zero at higher quality levels. Excluding the pair d8-mb, we can see that users favor V-PCC as the viewing distance increases. In the near setting, Draco is the preferred method over V-PCC.

C. Object Differences

Based on Fig. 2, we hypothesized that the *Dancer* object receives higher scores than the *Thaidancer* object. Fig. 5 shows that this predominantly holds for both compression formats. While the *Thaidancer* receives higher scores for medium and far distances using the Draco compression format and lower compression quality settings d8 and d9 (only far), the *Dancer* object receives higher scores in all other configurations. Interestingly, the difference increases for higher qualities in Draco.

We believe this trend could be due to the higher level of detail in the *Thaidancer* object leading to more recognizable artifacts for both compression formats. For quality settings d8 and d9, fine details are hard to notice, and participants might favor the *Thaidancer* object due to smoother movements. While we noticed differences between the objects, we leave an in-depth investigation of object-related perceived quality differences with more objects to future work. In the following, we continue with the scores averaged over both objects.

Jannis Weil, Yassin Alkhalili, Anam Tahir, Thomas Gruczyk, Tobias Meuser, Mu Mu, Heinz Koeppl, Andreas Mauthe. *Modeling Quality of Experience for Compressed Point Cloud Sequences based on a Subjective Study*. In Proceedings of the 15th International Conference on Quality of Multimedia Experience (QoMEX), 2023, pp. 135-140.

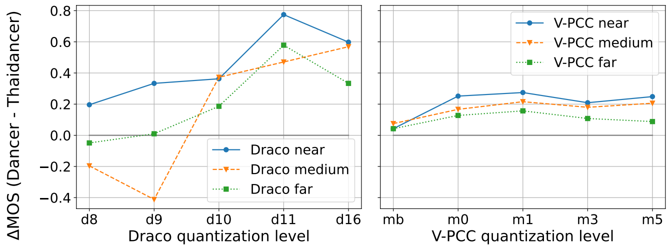

Fig. 5. Difference in MOS between *Dancer* and *Thaidancer* objects for Draco and V-PCC for all distance and quantization levels. A value greater than zero indicates that the dancer object received a higher MOS.

Method	FR	Quantization level index				
		0	1	2	3	4
V-PCC	10	1.3	2.6	3.2	7.5	21.0
	15	1.9	3.9	4.8	11.3	31.6
	30	3.8	7.8	9.6	22.5	63.1
Draco	30	703.8	905.8	1144.3	1409.2	2713.6

TABLE I
AVERAGE BIT RATE IN MBIT/S OVER BOTH OBJECTS FOR V-PCC AND DRACO AT DIFFERENT FRAME RATES (FR).

D. Tradeoff between Quality and Resources

In point cloud compression, there is a tradeoff between the visual quality and the resources required to achieve that quality [19]. Contributing factors include the bit rate of the point cloud stream, the time it takes to encode and decode each frame, and the processing power of the end devices.

While the MOS of our selected configurations is similar for Draco and V-PCC (see Section IV-B), the bit rate differs. Our raw point cloud sequences have an average data rate of approximately 3.6 Gbit/s for 10 fps, 5.5 Gbit/s for 15 fps, and 10.9 Gbit/s for 30 fps. Table I shows that the bit rates of the compressed Draco streams are magnitudes higher than the ones for the V-PCC streams, but still lower than transmitting the raw point cloud sequences. We want to highlight that the encode and decode times for Draco are much lower than for V-PCC on our system with an AMD Ryzen 7 5800X processor. The running times averaged over both objects for V-PCC vary from 118.8 to 140.0 seconds for encoding and from 4.2 to 5.1 seconds for decoding per frame, whereas the Draco running times range from 1.0 to 1.7 seconds for encoding and from 0.25 to 0.30 seconds for decoding per frame. To allow for live streaming, there is a need for compression methods that combine low data rates with fast encoding and decoding times.

The previous sections showed that higher quantization levels and frame rates lead to a higher MOS. However, this also results in higher data rates. There are cases where a stream with a lower bit rate is preferred by the participants. Let us ignore the timing constraints for en- and decoding and assume a client with 50 Mbit/s at a near distance setting. With V-PCC, the client could e.g. stream with quantization level m5 at 15 fps or with quantization level m3 at 30 fps (see Tab. I). According to Fig. 3 (a), m3 at 30 fps corresponds to a higher MOS, even though it has a lower bit rate. Combining the expected resource requirements for different configurations with an estimation of the MOS is therefore essential for good adaptation schemes.

V. MODELING QOE WITH MACHINE LEARNING

The results from our user study can be utilized to adjust the content based on the available bandwidth, e.g. by choosing the configuration with the highest expected MOS that still satisfies the given constraints. However, the results should be generalized beyond the experimental configurations to benefit wider scenarios. Previous research has shown promising results in using machine learning techniques to model QoE [8] and their ability to generalize to new data, as seen in the context of 360-degree video [20]. In this section, we explore the use of supervised machine learning to preserve the results from our experiments and to extend their applicability to configurations that were not included in the study. Such QoE models could then be employed in an adaptive streaming system that considers a broader range of configurations.

The learning and prediction task was achieved through three steps: feature engineering, training the machine learning models to predict the participants' quality ratings, and evaluation of the predictions. We focus on V-PCC in this section, but want to highlight that the approaches can also be used to model QoE for Draco. After applying outlier filtering by interquantile range, our collected dataset for V-PCC contains approximately 9,000 samples. Each sample consists of 5 features: frame rate, distance, and the V-PCC parameters as detailed in Section III. The corresponding user opinion scores range from one to five and serve as the learning targets. We analyzed feature importance to determine which features have the most significant influence on the model's predictions, and found that the occupancy precision is not a significant contributor to the accuracy of our model. This feature has only two different values across our configurations, which might not be sufficient to make it a useful predictor of the target.

We leverage leave-one-out cross-validation to assess different QoE model types. This method divides the data set into k folds, where $k-1$ folds are used for training and one fold is used for testing. This is repeated for k rounds, with each fold serving as the test set once. As the total number of V-PCC configurations aggregated over both objects is 45, we have $k = 45$ and group all individual user study responses by these folds. This approach provides an unbiased evaluation of the models' performance and helps to detect overfitting, thereby ensuring the robustness of the model.

Table II presents the results of the top five performing models that were tested, implemented using the Python library scikit-learn [21]. The table shows the R-squared (R2) score and the mean squared error (MSE) to the MOS, averaged over all rounds. A high R2 score and a low MSE indicate that the learned models can accurately predict the MOS of configurations that were not seen during training. Our examination revealed that the Gradient Boosting Regression model achieved the best performance. While regression models directly predict the MOS, we also explored the potential benefits of using classification models, where the probability

Jannis Weil, Yassin Alkhalili, Anam Tahir, Thomas Gruczyk, Tobias Meuser, Mu Mu, Heinz Koeppl, Andreas Mauthe. *Modeling Quality of Experience for Compressed Point Cloud Sequences based on a Subjective Study*. In Proceedings of the 15th International Conference on Quality of Multimedia Experience (QoMEX), 2023, pp. 135-140.

QoE Model Type	R2 score	MSE
Gradient Boosting Regressor	0.9754	0.0175
Polynomial Regression (Degree 2)	0.9677	0.0229
Random Forest Regressor	0.9565	0.0310
Decision Tree Regressor	0.9457	0.0386
Decision Tree Classifier	0.9440	0.0398

TABLE II
PERFORMANCE OF VARIOUS MACHINE LEARNING MODELS EVALUATED USING THE R2 SCORE (HIGHER IS BETTER) AND THE MEAN SQUARED ERROR (LOWER IS BETTER), SORTED BY DECREASING PERFORMANCE.

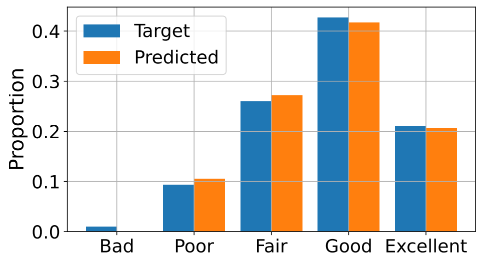

Fig. 6. Actual and predicted QoE distributions for the test fold with medium distance, 30 fps, and m1 quantization level using the Decision Tree Classifier.

per class prediction serves as a QoE distribution. Fig. 6 shows the comparison between the actual QoE distribution from the participants' votes and the predicted distribution for a selected configuration. The Decision Tree Classifier can accurately predict the perceived quality distribution, although none of the corresponding samples were seen during training. Predicting the quality distribution instead of the MOS provides additional insights that can be used for adaptation. For example, the prediction in Fig. 6 suggests that less than 11% of the participants are likely to perceive this particular configuration as poor or bad, similar to the actual QoE distribution.

VI. Conclusion

We have designed and conducted a subjective user study to systematically study the effect on point cloud objects of varying frame rates, quantization levels and distance to the camera. For compression, we investigated Draco and V-PCC. Our results confirm that when increased independently, frame rate and quantization level lead to a higher perceived quality. However, we found that a higher bitrate does not necessarily imply a higher MOS. We also observed that the degradation of the visual quality based on the quantization levels becomes less noticeable at a higher viewer distance. Lastly, we developed QoE models using machine learning algorithms for an accurate quality assessment of compressed point cloud sequences. Such models can be used for future QoE-aware adaptive point cloud streaming solutions or intelligent network resource management. Additionally, we plan to conduct further studies with immersive end devices, such as VR headsets, and compare the results to those obtained in this study using regular screens.

Acknowledgments

This work has been co-funded by the German Research Foundation (DFG) as part of Collaborative Research Center (CRC) 1053 MAKI subprojects C3 and B1. The authors thank Dr. Antonio Fernández Anta for his valuable comments.

References

[1] S. Hudson, S. Matson-Barkat, N. Pallamin, and G. Jegou, "With or without you? Interaction and immersion in a virtual reality experience," *Journal of Business Research*, vol. 100, pp. 459–468, 2019.
[2] S. Rossi, I. Viola, and P. Cesar, "Behavioural Analysis in a 6-DoF VR System: Influence of Content, Quality and User Disposition," in *Proc. 1st Workshop on Interactive eXtended Reality*. ACM, 2022.
[3] Z. Liu, Q. Li, X. Chen, C. Wu, S. Ishihara, J. Li, and Y. Ji, "Point cloud video streaming: Challenges and solutions," *IEEE Network*, vol. 35, no. 5, pp. 202–209, 2021.
[4] W. Lin, S. Lee *et al.*, "Visual Saliency and Quality Evaluation for 3D Point Clouds and Meshes: An Overview," *APSIPA Transactions on Signal and Information Processing*, vol. 11, no. 1, 2022.
[5] H. Liu, H. Yuan, Q. Liu, J. Hou, and J. Liu, "A comprehensive study and comparison of core technologies for MPEG 3-D point cloud compression," *IEEE Transactions on Broadcasting*, vol. 66, no. 3, pp. 701–717, 2019.
[6] Google, *Draco: 3D Data Compression*. GitHub Repository, 2018. [Online]. Available: https://github.com/google/draco
[7] E. Dumic, C. R. Duarte, and L. A. da Silva Cruz, "Subjective evaluation and objective measures for point clouds — State of the art," in *First International Colloquium on Smart Grid Metrology*, 2018, pp. 1–5.
[8] G. Kougioumtzidis, V. Poulkov, Z. D. Zaharis, and P. I. Lazaridis, "A Survey on Multimedia Services QoE Assessment and Machine Learning-Based Prediction," *IEEE Access*, vol. 10, pp. 19 507–19 538, 2022.
[9] A. Ahmad, A. B. Mansoor, A. A. Barakabitze, A. Hines, L. Atzori, and R. Walshe, "Supervised-learning-Based QoE Prediction of Video Streaming in Future Networks: A Tutorial with Comparative Study," *IEEE Communications Magazine*, vol. 59, no. 11, pp. 88–94, 2021.
[10] J. van der Hooft, M. T. Vega, C. Timmerer, A. C. Begen, F. De Turck, and R. Schatz, "Objective and Subjective QoE Evaluation for Adaptive Point Cloud Streaming," in *Proc. 12th International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 2020, pp. 1–6.
[11] Q. Liu, H. Su, Z. Duanmu, W. Liu, and Z. Wang, "Perceptual Quality Assessment of Colored 3D Point Clouds," *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–1, 2022.
[12] X. Wu, Y. Zhang, C. Fan, J. Hou, and S. Kwong, "Subjective Quality Database and Objective Study of Compressed Point Clouds With 6DoF Head-Mounted Display," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 31, no. 12, pp. 4630–4644, 2021.
[13] E. Zerman, C. Ozcinar, P. Gao, and A. Smolic, "Textured Mesh vs Coloured Point Cloud: A Subjective Study for Volumetric Video Compression," in *Proc. 12th International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 2020, pp. 1–6.
[14] K. Cao, Y. Xu, and P. Cosman, "Visual Quality of Compressed Mesh and Point Cloud Sequences," *IEEE Access*, vol. 8, pp. 171 203–171 217, 2020.
[15] Y. Xu, Y. Lu, and Z. Wen, "Owlii Dynamic Human Textured Mesh Sequence Dataset," https://mpeg-pcc.org/index.php/pcc-content-database/owlii-dynamic-human-textured-mesh-sequence-dataset/, 2017.
[16] M. Krivokuća, P. A. Chou, and P. Savill, "8i Voxelized Surface Light Field (8iVSLF) Dataset," https://mpeg-pcc.org/index.php/pcc-content-database/8i-voxelized-surface-light-field-8ivslf-dataset/, 2018.
[17] J. de Hoog, A. N. Ahmed, A. Anwar, S. Latré, and P. Hellinckx, "Quality-Aware Compression of Point Clouds with Google Draco," in *Proc. 16th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*. Springer, 2022, pp. 227–236.
[18] L. Li, Z. Li, S. Liu, and H. Li, "Efficient Projected Frame Padding for Video-Based Point Cloud Compression," *IEEE Transactions on Multimedia*, vol. 23, pp. 2806–2819, 2020.
[19] E. Alexiou, I. Viola, T. M. Borges, T. A. Fonseca, R. L. de Queiroz, and T. Ebrahimi, "A comprehensive study of the rate-distortion performance in MPEG point cloud compression," *APSIPA Transactions on Signal and Information Processing*, vol. 8, p. e27, 2019.
[20] M. S. Anwar, J. Wang, W. Khan, A. Ullah, S. Ahmad, and Z. Fei, "Subjective qoe of 360-degree virtual reality videos and machine learning predictions," *IEEE Access*, vol. 8, pp. 148 084–148 099, 2020.
[21] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion *et al.*, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.