# Perceptual Analysis of Compressed Dynamic Point Clouds in Virtual Reality Environments

**Josef Augusto Oberdan Souza Silva[1], Gustavo Rehbein[1], Adriane Borda[2], Marcelo Porto[1], Guilherme Corrêa[1]**

[1]Programa de Pós-Graduação em Computação (PPGC)
Video Technology Research Group (ViTech)
Universidade Federal de Pelotas (UFPel), Brazil
[2]Programa de Pós-Graduação em Arquitetura e Urbanismo (PROGRAU)
Universidade Federal de Pelotas (UFPel), Brazil

[1]{jaossilva, ghrehbein, porto, gcorrea}@inf.ufpel.edu.br,
[2]adribord@hotmail.com

***Abstract.*** *This study investigated the perceived quality of dynamic point clouds in virtual reality (VR) under different compression scenarios through an immersive subjective experiment. Six compression scenarios were analyzed, including one employing a machine learning-based solution for codec complexity reduction. A VR application, developed according to ITU-T recommendations, was used to conduct the experiments and collect user scores, response times, and visual attention data. The results indicated agreement between most objective and perceived quality metrics, with moderate variability across participants. Geometric metrics, particularly p2plane and p2point, showed stronger association with MOS than bitrate and luminance PSNR, while heatmaps revealed greater gaze concentration in central and visually informative regions of the point clouds. These findings support the development of perceptually optimized compression and rendering strategies for immersive point cloud applications.*

## 1. Introduction

In recent years, advances in Virtual Reality (VR) technologies have expanded the ways of interacting with three-dimensional (3D) content, especially through representations based on point clouds [Gutierrez et al. 2021]. This type of representation enables the capture and reconstruction of 3D scenes with a high level of geometric and visual detail, being applied in areas such as volumetric communication, scientific visualization, medical imaging, autonomous navigation, and digital preservation [Graziosi et al. 2020]. When combined with immersive devices such as head-mounted displays (HMDs), point clouds enable more realistic and explorable experiences across different perspectives and degrees of freedom (DoF) [Zhou et al. 2025].

Dynamic Point Clouds (DPCs), organized as temporal sequences of 3D content, require compression techniques to enable storage and transmission [Graziosi et al. 2020]. In this context, Video-based Point Cloud Compression (V-PCC), standardized by the Moving Picture Experts Group (MPEG), stands out [Preda 2020]. Its reference software, Test Model Category 2 (TMC2), uses High-Efficiency Video Coding (HEVC) and offers different trade-offs between compression rate and quality degradation [MPEG 2025] [Sullivan et al. 2012]. Since compression may introduce visual artifacts that affect the user experience [Viola et al. 2022], subjective evaluation has become essential for understanding its effects in immersive environments [Graziosi et al. 2020].

Several studies have investigated the perceived quality of compressed DPCs, mainly analyzing the impact of different codecs, compression levels, and viewing conditions on users' perception [Viola et al. 2022]. [Singla et al. 2023] investigated static and dynamic point clouds in remote experiments, analyzing distortions such as Gaussian noise, octree pruning, and compression with Geometry-based Point Cloud Compression (G-PCC) and Video-based Point Cloud Compression (V-PCC), with analysis based on Absolute Category Rating (ACR), Mean Opinion Score (MOS), 95% confidence intervals, and Standard deviation of Opinion Scores (SOS). In addition, [Zhou et al. 2023] and [Zhou et al. 2025] proposed and expanded the QAVA-DPC (Quality Assessment and Visual Attention) dataset for virtual reality evaluation with six degrees of freedom (6DoF), integrating compression codecs, eye-tracking, and the use of head-mounted displays (HMDs), while [Dumic and da Silva Cruz 2023] investigated packet loss in V-PCC transmissions with the Double Stimulus Impairment Scale (DSIS) method, both with analysis based on MOS.

Despite these advances, some aspects remain insufficiently explored, particularly how different geometry trade-offs from the TMC2 reference software are perceived in controlled immersive VR settings when subjective scores are analyzed together with response times and visual attention [Viola et al. 2022]. In addition, previous studies rarely investigate whether complexity-reduction strategies in the compression pipeline can be adopted with little perceived quality loss [Gutierrez et al. 2021].

Therefore, this work investigates the perceived quality of DPCs in immersive VR environments under different compression scenarios. A subjective evaluation methodology was implemented through a custom VR application designed according to ITU-T recommendations, enabling the collection of quality scores, response times, and visual attention data. The study also analyzes the relationship between subjective perception and objective quality metrics, highlighting the relevance of geometric metrics such as p2plane and p2point. In addition, gaze-based analysis provides insights into how users visually explore point cloud content, offering indications for perceptually driven compression and rendering strategies in immersive applications.

This paper presents the following contributions: (1) a subjective evaluation study of the perceived quality of compressed dynamic point clouds in a virtual reality environment; (2) the development of an immersive application for conducting subjective experiments with support for visualization, voting, and behavioral data recording; (3) the identification of relationships among perceived quality, objective metrics, and visual attention patterns through the joint analysis of subjective scores, response times, and heatmaps; and (4) the subjective analysis of a machine learning-based strategy aimed at reducing coding complexity, investigating its perceptual impact on the final quality of compressed point clouds.

This paper is organized as follows: Section 2 describes the adopted methodology;

Section 3 presents the developed immersive application; Section 4 discusses the experimental results; and, finally, Section 5 presents the conclusions and perspectives for future work.

## 2. Subjective Evaluation Methodology

This section describes the proposed methodology for designing the subjective evaluation tests of DPCs. The overall workflow, illustrated in Figure 1, summarizes the pipeline adopted in this study, beginning with a raw DPC that is encoded with V-PCC and then reconstructed for perceptual evaluation. In addition, the immersive VR application developed in this work, presented in Section 3, is responsible for managing and conducting the experiments according to the proposed methodology, including stimulus presentation, participant interaction, response collection, and data logging. The reconstructed content is then assessed through this application and a questionnaire, and the collected data are finally analyzed using quality-related indicators such as MOS, DMOS, $CI_{95}$, and average response time.

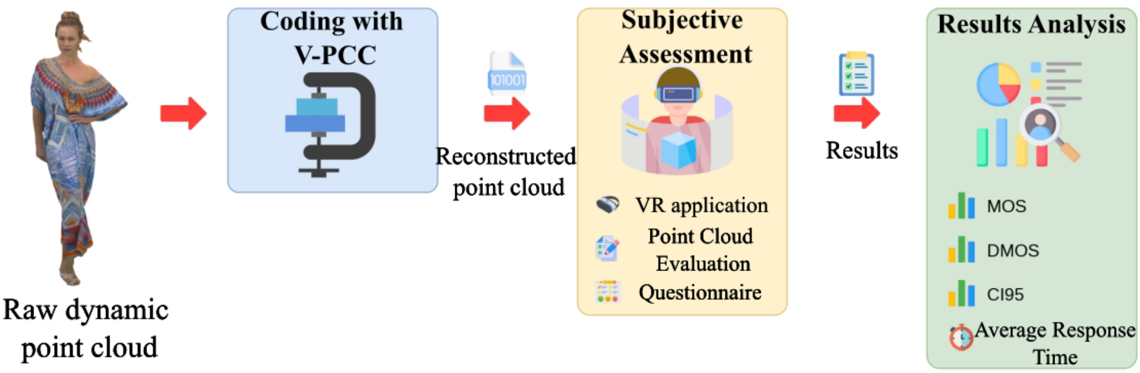



**Figure 1. Flowchart of the methodology used.**

### 2.1. V-PCC Experimental Setup

TMC2, version 22.1, was used to compress the DPC sequences employed in the experiments reported in this paper. Three sequences were selected from the seven sequences indicated in the common test conditions of V-PCC [ISO/IEC 2020], namely: Loot, Soldier, and Longdress. Table 1 presents the characteristics of each test sequence.

**Table 1. V-PCC test sequences [ISO/IEC 2020].**

<table>
<tr><th>Class</th><th>Sequence Name</th><th>Frames</th><th>FPS</th><th># Points</th><th>Precision</th><th>Maximum value</th><th>Attributes</th></tr>
<tr><td rowspan="4">A</td><td>Queen</td><td>250</td><td>50</td><td>~1,000,000</td><td rowspan="5">10 bit</td><td rowspan="5">1023</td><td rowspan="7">R, G, B</td></tr>
<tr><td>8i VFB – Loot</td><td rowspan="4">300</td><td rowspan="6">30</td><td>~780,000</td></tr>
<tr><td>8i VFB – Red and Black</td><td>~700,000</td></tr>
<tr><td>8i VFB – Soldier</td><td>~1,500,000</td></tr>
<tr><td>B</td><td>8i VFB – Long dress</td><td>~800,000</td></tr>
<tr><td rowspan="2">C</td><td>basketball_player_vox11</td><td rowspan="2">64</td><td>~2,900,000</td><td rowspan="2">11 bit</td><td rowspan="2">2047</td></tr>
<tr><td>dancer_vox11</td><td>~2,600,000</td></tr>
</table>

By default, TMC2 provides five bitrate configurations, namely R1, R2, R3, R4, and R5, where R1 presents the highest compression rate and R5 the highest quality. These configurations control, among other aspects, the quantization parameter (QP) used in the

**Table 2. Rate Points and Quantization Parameters (QP)**

| Rate Point | Geometry QP | Attribute QP |
|---|---|---|
| R1 | 32 | 42 |
| R2 | 28 | 37 |
| R3 | 24 | 32 |
| R4 | 20 | 27 |
| R5 | 16 | 22 |
| **R35** | 24 | 22 |
| **R53** | 16 | 32 |

video encoder for each sub-bitstream, as shown in Table 2. Two additional configurations were created to enable a better individual assessment of the impact of geometry and attributes on the final perceived quality. These configurations were named R35 and R53, and they alternate between the QP values used for geometry and attributes in R3 and R5.

**Table 3. Experimental results.**

| Sequence | Configuration | p2plane (dB) | C0 (dB) | Bitrate (kbps) |
|---|---|---|---|---|
| longdress | R3 | 71.80 | 30.18 | 686.16 |
| | R4 | 73.05 | 32.70 | 1,325.32 |
| | R4 [Rehbein et al. 2025] | 73.01 | 32.50 | 1,333.23 |
| | R5 | 74.93 | 35.70 | 3,059.13 |
| | R35 | 72.38 | 34.90 | 2,735.86 |
| | R53 | 74.93 | 30.30 | 1,050.93 |
| loot | R3 | 72.07 | 36.04 | 285.49 |
| | R4 | 73.31 | 38.51 | 472.95 |
| | R4 [Rehbein et al. 2025] | 73.28 | 38.39 | 487.12 |
| | R5 | 75.15 | 41.03 | 940.88 |
| | R35 | 72.64 | 40.61 | 699.92 |
| | R53 | 75.15 | 36.03 | 579.94 |
| soldier | R3 | 71.77 | 33.74 | 439.23 |
| | R4 | 73.08 | 36.16 | 762.30 |
| | R4 [Rehbein et al. 2025] | 73.04 | 36.00 | 773.12 |
| | R5 | 75.03 | 38.64 | 1,598.58 |
| | R35 | 72.43 | 38.01 | 1,251.44 |
| | R53 | 75.03 | 33.75 | 871.62 |

The experiments were conducted with TMC2 v22.1 [MPEG 2025], in Random Access mode, using the R3, R5, R35, and R53 configurations defined in Table 2, in addition to the R4 configuration in the standard implementation. Besides, a modified version of TMC2 with a state-of-the-art complexity reduction solution based on machine learning, proposed in [Rehbein et al. 2025], was included in the experiments. According to the authors, it reduces the computational cost of V-PCC by up to 60% without significant impact on objective quality metrics. However, subjective evaluations have not been reported; thus, the effect of the solution on perceived quality is assessed in this work.

For reference, the resulting objective metrics are presented in Table 3. The values were calculated according to the common test conditions of V-PCC [ISO/IEC 2020], including bitrate, geometric quality (p2plane), and attribute quality (C0), with the reconstructed point clouds being used in the subjective experiments. The C0 metric refers to the Peak Signal-to-Noise Ratio (PSNR) computed between the luminance channel of the observed and the original point clouds.

### 2.2. Experimental Setup

The experiment followed the Absolute Category Rating (ACR) methodology, according to the recommendations of ITU-T P.910 [ITU-T 2008], ITU-R BT.500-15 [ITU-R 2023], and ITU-T P.919 [ITU-T 2020], in which participants assigned scores from 1 to 5 (1 - bad, 2 - poor, 3 - fair, 4 - good, and 5 - excellent) to the perceived visual quality of the dynamic point cloud sequences. Before sessions, volunteers were informed about the procedures, signed the informed consent form, and underwent visual pre-screening with the Ishihara test [ITU-T 2020] [ITU-R 2023] [Ishihara 1960] [Gutierrez et al. 2021]. The average time per participant is shown in Table 4.

**Table 4. Summary of the average evaluation time per participant on each day of the experiment.**

| | Start | End | Total duration | Participants | Average time per participant |
|---|---|---|---|---|---|
| 10/02/2026 | 15:46:15 | 16:49:51 | 01:03:36 | 5 | 00:12:43 |
| 12/02/2026 | 15:30:00 | 17:00:00 | 01:30:00 | 8 | 00:11:15 |
| 24/02/2026 | 15:32:50 | 16:56:04 | 01:23:14 | 8 | 00:10:24 |
| 26/02/2026 | 09:49:52 | 12:27:03 | 02:37:11 | 7 | 00:22:27 |
| 27/02/2026 | 09:47:53 | 11:37:53 | 01:50:00 | 6 | 00:18:20 |
| **Total** | | | **08:24:01** | **34** | |

The tests were conducted in a controlled environment under the supervision of a moderator, using the Meta Quest® 3s HMD in a three degrees of freedom (3DoF) VR condition, and involved 34 participants, whose native language was predominantly Brazilian Portuguese, evaluating eighteen stimuli presented in 15-second visualization and voting stages, with the entire procedure recorded on video in 4K resolution. If the allotted voting time expired and the participant had not assigned a score to the visualized point cloud, no vote was recorded and the flow proceeded normally.

### 2.3. Subjective Metrics

Mean Opinion Score (MOS) was calculated as the average of the subjective ratings, while Standard Deviation of Opinion Scores (SOS) was used to evaluate the dispersion of the responses. Differential Mean Opinion Score (DMOS) corresponded to the difference between the MOS of the reference and that of the evaluated stimulus. Positive DMOS values indicate a decrease in perceived quality in comparison to the reference point cloud, whereas negative DMOS values indicate an increase in perceived quality. The 95% confidence intervals ($CI_{95}$), obtained using the Student's t-distribution [Zhou et al. 2025], were used to measure the statistical uncertainty of MOS [Dumic and da Silva Cruz 2023]. In addition, the mean reaction time ($rt_{mean}$) and its standard deviation ($rt_{sd}$) were calculated to analyze the participants' judgment time [Viola et al. 2022] [Zhou et al. 2023].

## 3. Proposed Application for Immersive Subjective DPC Evaluation

One of the contributions of this work is the development of an application to conduct the subjective quality assessment of DPCs in VR, following the methodology presented in Figure 2. The application was developed in the Unity engine to support a subjective quality evaluation experiment of DPCs in VR, using a GPU buffer-based renderer and the Pcx Point Cloud Importer plug-in[1] to load and convert point clouds [Neuman-Donihue et al. 2023]. PLY (Polygon File Format) frames in binary little-endian format are loaded through asynchronous I/O on a dedicated thread, using a sliding-window and prefetch strategy to maintain performance on the Meta Quest® 3S device.

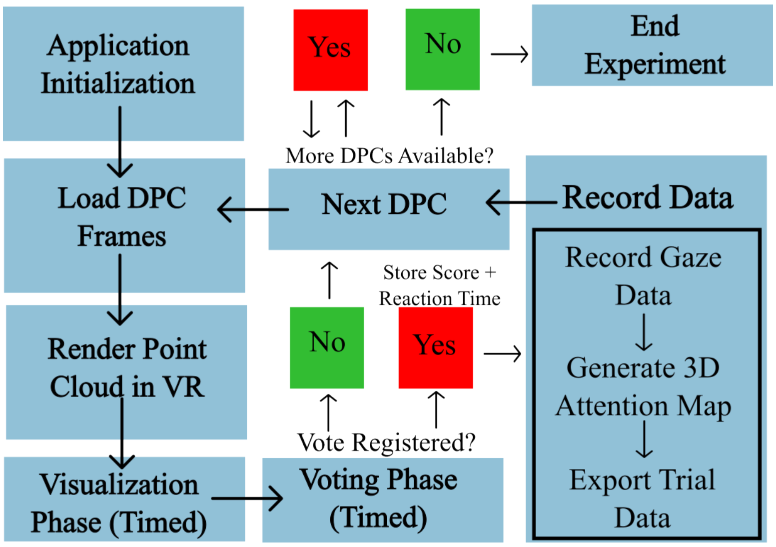
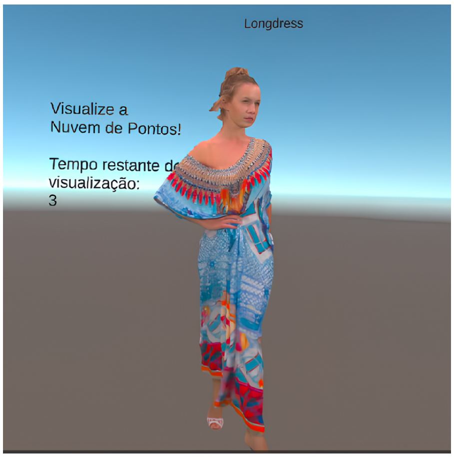



**Figure 2. Flowchart of the virtual reality application.**

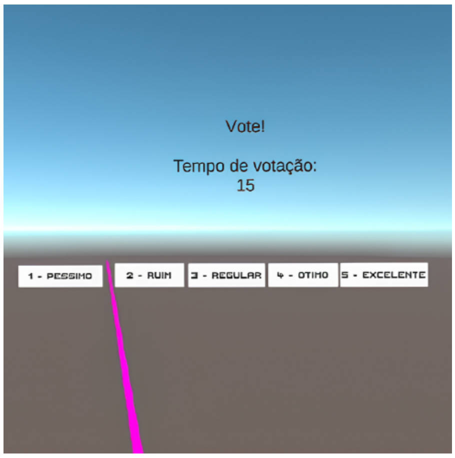
The application followed two timed phases: visualization, in which participants observed the point cloud, and voting, in which they rated its quality from 1 (bad) to 5 (excellent) [ITU-T 2020], with reaction time measured from the start of voting. The experiment was conducted in a 3DoF VR environment, allowing head rotation along the *yaw*, *pitch*, and *roll* axes, while point clouds remained fixed at a predefined distance for proper visibility. Interaction used *joystick*-based navigation with button highlighting, locked states during loading/transitions, and double-click protection. After each round, the "Next" button led participants to the next cloud or to the results screen (Figure 3). The source code is available online[2].

Finally, the experiment data were stored locally. In addition to the votes (score, reaction time, etc.), the application recorded attention by sampling the *rig* (position and rotation from the camera or eye center) and performing *raycasting* against a volume of interest. The volume of interest (surrounding the point cloud) was subdivided into a 3D grid with a defined resolution. Whenever the user's viewing ray intersected the volume of interest, the intersection position was converted into normalized coordinates within the grid. The corresponding cell was then identified, and the elapsed time since the last sampling was accumulated in that cell. This process was repeated throughout the visualization phase, resulting in a three-dimensional attention map. This map was exported for

[1]`https://github.com/keijiro/Pcx`
[2]`https://github.com/josefoberdan/PCQA_Test.git`




(a) Visualization screen.




(b) Voting screen.

**Figure 3. Virtual reality application interface: (a) timed visualization stage and (b) timed voting stage.**

each trial, enabling subsequent analyses of fixation and spatial distribution of attention [Nguyen et al. 2024].

## 4. Experimental Results and Discussion

This section presents the experimental results obtained in the study, as well as a discussion of the main observations arising from the evaluations carried out by the participants.

### 4.1. Participant Profile and Analysis of Questionnaire Responses

Initially, the participant profile was described based on demographic characteristics and relevant information collected through the questionnaire administered before the experiment. Next, the questionnaire responses were analyzed in order to identify patterns related to the participants' prior experience with subjective analysis and previous use of VR, as well as possible factors that may influence the evaluation process.

The questionnaire results suggested that male participants showed greater diversity in prior experience with virtual reality, while among female participants this experience was concentrated at intermediate levels. Despite this difference, intense exploration of the point cloud predominated in both groups, indicating that the experimental task was accessible and allowed active interaction regardless of the level of prior familiarity with VR, in agreement with [Zhou et al. 2023]. In addition, the results showed a predominance of young participants, especially those between 18 and 24 years old, and a low occurrence of motion sickness during the test, with only one positive case, which suggested good tolerability of the application and a lower probability of interference from physiological discomfort in the subjective evaluations [Viola et al. 2022].

### 4.2. Calculation of MOS, SOS, $CI_{95}$, DMOS and Response Time

The results of the subjective quality evaluations, including the calculation of MOS and DMOS metrics, were used to quantify the perceived quality of the analyzed point clouds. The 95% Confidence Interval ($CI_{95}$), indicating the reliability of the mean ratings assigned by the participants, is also presented in Table 5. The table also presents the number of participants who evaluated each point cloud (N), SOS values and the response time (mean and standard deviation) of users, in seconds.

**Table 5. MOS, SOS, $CI_{95}$, DMOS and response time (seconds) by point cloud.**

| CloudName | Group | Ref(Group) | N | MOS | SOS | $CI_{95}$ | $MOS_{ref}$ | DMOS | $RT_{mean}$ | $RT_{sd}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Longdress_R35 | Longdress | Longdress_R5 | 33 | 3.36 | 0.99 | 0.35 | 3.52 | 0.15 | 2.96 | 1.89 |
| Longdress_R3 | Longdress | Longdress_R5 | 33 | 3.09 | 0.77 | 0.27 | 3.52 | 0.42 | 4.80 | 2.24 |
| Longdress_R4_Rehbein | Longdress | Longdress_R5 | 33 | **3.18** | **0.88** | 0.31 | 3.52 | **0.33** | 3.15 | 1.71 |
| Longdress_R4 | Longdress | Longdress_R5 | 32 | **3.31** | **1.00** | 0.36 | 3.52 | **0.20** | 4.10 | 3.19 |
| Longdress_R53 | Longdress | Longdress_R5 | 33 | 3.36 | 0.96 | 0.34 | 3.52 | 0.15 | 3.93 | 2.48 |
| Longdress_R5 | Longdress | Longdress_R5 | 33 | 3.52 | 1.00 | 0.36 | 3.52 | 0.00 | 3.76 | 3.58 |
| Loot_R35 | Loot | Loot_R5 | 33 | 3.27 | 0.91 | 0.32 | 3.64 | 0.36 | 2.83 | 1.58 |
| Loot_R3 | Loot | Loot_R5 | 33 | 3.12 | 0.78 | 0.28 | 3.64 | 0.52 | 2.84 | 1.53 |
| Loot_R4_Rehbein | Loot | Loot_R5 | 33 | **3.48** | **0.67** | 0.24 | 3.64 | **0.15** | 4.10 | 3.16 |
| Loot_R4 | Loot | Loot_R5 | 33 | **3.42** | **0.79** | 0.28 | 3.64 | **0.21** | 2.69 | 1.68 |
| Loot_R53 | Loot | Loot_R5 | 33 | 3.45 | 0.87 | 0.31 | 3.64 | 0.18 | 3.08 | 2.14 |
| Loot_R5 | Loot | Loot_R5 | 33 | 3.64 | 0.74 | 0.26 | 3.64 | 0.00 | 3.03 | 1.64 |
| Soldier_R35 | Soldier | Soldier_R5 | 33 | 3.00 | 1.06 | 0.38 | 3.27 | 0.27 | 2.74 | 1.74 |
| Soldier_R3 | Soldier | Soldier_R5 | 32 | 3.06 | 0.91 | 0.33 | 3.27 | 0.21 | 3.44 | 2.61 |
| Soldier_R4_Rehbein | Soldier | Soldier_R5 | 33 | **3.09** | **0.98** | 0.35 | 3.27 | **0.18** | 3.17 | 2.19 |
| Soldier_R4 | Soldier | Soldier_R5 | 33 | **2.94** | **0.97** | 0.34 | 3.27 | **0.33** | 3.00 | 2.40 |
| Soldier_R53 | Soldier | Soldier_R5 | 33 | 3.15 | 1.03 | 0.37 | 3.27 | 0.12 | 2.59 | 1.67 |
| Soldier_R5 | Soldier | Soldier_R5 | 33 | 3.27 | 1.13 | 0.40 | 3.27 | 0.00 | 2.62 | 1.97 |

When comparing the R4 and R4_Rehbein versions (Table 5), it was observed that the effect of the proposed compression varied according to the content. In Longdress, the R4_Rehbein version showed a DMOS of 0.33, while the R4 version presented a DMOS of 0.20, meaning that the R4_Rehbein version presented a worse perceptual performance than R4. This can also be perceived by the MOS value, which achieved 3.18 for R4_Rehbein and 3.31 for R4, indicating a decrease of 0.13. In contrast, in the Loot sequence, the R4_Rehbein version presented an improved perceptual quality relative to R4, with a DMOS of 0.15 versus 0.21 and a MOS of 3.48 versus 3.42. This improvement was even more evident in the Soldier sequence, where R4_Rehbein obtained a DMOS of 0.18, substantially lower than that of R4, which presented a DMOS of 0.33. The results indicate that the compression approach proposed by [Rehbein et al. 2025] maintained perceptual quality very close to the R53 configuration and, in one case, it was even superior: in Loot, for example, R4_Rehbein presented a MOS of 3.48 and a DMOS of 0.15, while R53 obtained a MOS of 3.45 and a DMOS of 0.18. In Soldier, although R53 remained superior (MOS = 3.15; DMOS = 0.12), the difference relative to R4_Rehbein was small (MOS = 3.09; DMOS = 0.18), which reinforces the proximity between the results.

Table 5 showed that MOS varied at an intermediate level, with SOS close to $1.0$ and relatively low $95\%$ confidence intervals, indicating moderate dispersion and statistical consistency in the evaluations [Zhou et al. 2025]. In general, R5 tended to present the highest MOS values, while R3 recorded the lowest, and the hybrid configurations R35 and R53 exhibited content-dependent behavior, with performance close to or higher than R5 in Longdress, but not in Loot and Soldier [Zhou et al. 2025]. In addition, the small difference between R4_Rehbein and R4 suggested little perceived impact resulting from the complexity reduction strategy, while the low DMOS values indicated moderate perceptual differences. Taken together, the results suggested that the relative relevance of geometry and visual attributes varied according to the content, and that, especially at intermediate quality levels, judgment became less immediate and more dependent on the characteristics of the sequence [Zhou et al. 2023].

### 4.3. Correlation Metrics

In this subsection, the degree of agreement between the objective metrics presented in Table 3 and the subjective scores presented in Table 5 is discussed (Table 6). For this purpose, different correlation measures were employed, each capturing specific aspects of the relationship between variables, in addition to a nonlinear mapping procedure, which is a common practice in quality assessment studies. Table 6 presents the correlation results in terms of Spearman Rank-Order Correlation Coefficient (SROCC) [Sedgwick 2014], Kendall Rank-Order Correlation Coefficient (KROCC) [Noether et al. 1981], and Pearson Linear Correlation Coefficient (PLCC) [Schober et al. 2018]. Since the relationship between an objective metric and human perception is rarely strictly linear, PLCC was calculated after a nonlinear mapping using a logistic function, adjusting the predictions to the MOS domain and reducing saturation effects across the quality range. After this adjustment, prediction accuracy was evaluated through Root Mean Square Error (RMSE) and Mean Absolute Error (MAE), which indicate, respectively, the typical error and the mean absolute deviation between the predicted and observed values. The correlations were calculated both for the global set and separately for each sequence, in order to detect possible content-dependent effects.

**Table 6. MOS correlations**

| Point Cloud | Metric | N.Points | SROCC | KROCC | PLCC | RMSE | MAE |
|---|---|---|---|---|---|---|---|
| All | Bitrate | 18 | 0.328 | 0.226 | 0.454 | 0.199 | 0.164 |
| All | C0 | 18 | 0.099 | 0.040 | 0.377 | 0.207 | 0.173 |
| All | p2plane | 18 | 0.523 | 0.396 | 0.529 | 0.190 | 0.150 |
| All | p2point | 18 | 0.577 | 0.477 | 0.563 | 0.185 | 0.155 |
| Longdress | Bitrate | 6 | 0.406 | 0.276 | 0.615 | 0.171 | 0.140 |
| Longdress | C0 | 6 | 0.464 | 0.414 | 0.598 | 0.174 | 0.133 |
| Longdress | p2plane | 6 | 0.515 | 0.500 | 0.598 | 0.174 | 0.133 |
| Longdress | p2point | 6 | 0.250 | 0.357 | 0.598 | 0.174 | 0.133 |
| Loot | Bitrate | 6 | 0.600 | 0.467 | 0.754 | 0.108 | 0.072 |
| Loot | C0 | 6 | 0.314 | 0.200 | 0.478 | 0.144 | 0.127 |
| Loot | p2plane | 6 | 0.783 | 0.690 | 0.936 | 0.058 | 0.043 |
| Loot | p2point | 6 | 0.725 | 0.552 | 0.897 | 0.073 | 0.068 |
| Soldier | Bitrate | 6 | 0.543 | 0.467 | 0.781 | 0.066 | 0.052 |
| Soldier | C0 | 6 | 0.086 | 0.067 | 0.781 | 0.066 | 0.052 |
| Soldier | p2plane | 6 | 0.580 | 0.414 | 0.834 | 0.058 | 0.055 |
| Soldier | p2point | 6 | 0.638 | 0.552 | 0.834 | 0.058 | 0.055 |

In Table 6, N.Points indicates the number of samples used in each correlation analysis, rather than the number of geometric points in the 3D point clouds. Specifically, N.Points = 18 for the complete dataset and N.Points = 6 for each individual point cloud (Longdress, Loot, and Soldier). Thus, the reported SROCC, KROCC, PLCC, RMSE, and

MAE values were computed from the corresponding pairs of MOS and objective metric values for each sample set.

The results in Table 6 showed that the geometric metrics presented higher agreement with subjective opinion than the attribute metrics and bitrate. In the global scenario, p2plane and p2point achieved the highest correlations with MOS (SROCC of 0.523 and 0.577, respectively) and lower errors after nonlinear mapping, while C0 showed the weakest association (SROCC of 0.099), suggesting that, for the set of stimuli and for visualization in virtual reality, geometric distortions were more decisive for perceived quality than attribute variations. The sequence-based analysis reinforced the content-dependent effect: in Loot, the correlations of p2plane and p2point with MOS were high, with emphasis on an SROCC of 0.783 for p2plane, while in Longdress this metric also showed a consistent correlation. In Soldier, however, C0 showed an almost null correlation (SROCC of 0.086), whereas the geometric metrics maintained moderate correlations, indicating greater sensitivity to structural distortions than to color variations in the content.

### 4.4. Fixation Maps and Visual Exploration Analysis

This subsection analyzes the visual attention of the participants based on the gaze data collected during the exploration of the DPCs in virtual reality [Viola et al. 2022] [Gutierrez et al. 2021]. The gaze data collection and map generation followed approaches aligned with recent literature [Nguyen et al. 2024] [Zhou et al. 2023] [Zhou et al. 2025], integrating subjective data, rig position and rotation, and raycasting operations to support the interpretation of perceived quality (Figure 4).

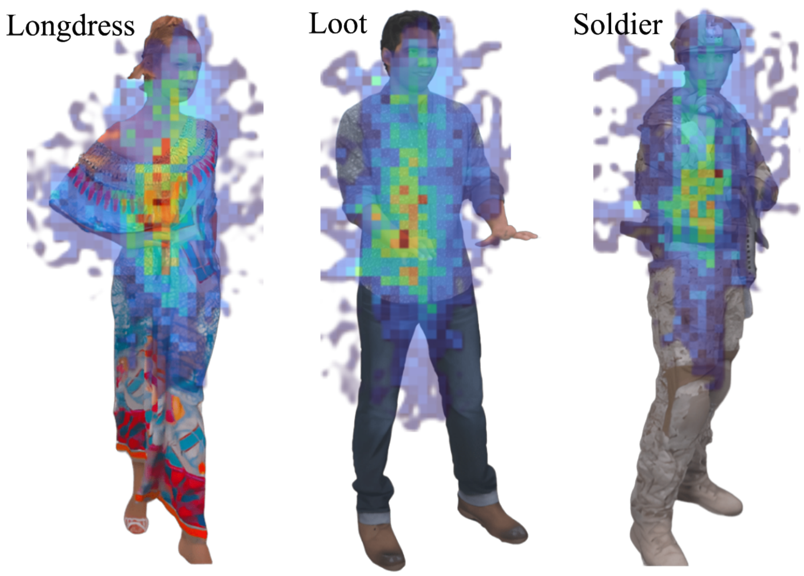



**Figure 4. Attention fixation map generated for each point cloud (*Longdress*, *Loot* and *Soldier*) during the experiments.**

Figure 4 shows that the heatmaps were mainly concentrated in the central regions of the stimuli, suggesting prioritization of structurally more informative areas for quality judgment. In the Longdress sequence, the heatmaps indicate that the gaze is mainly concentrated on the face, bust, and the hand placed on the waist, which moves the model's dress. In the Loot sequence, the heatmap shows a strong concentration of gaze on the character's right hand, which simulate piano playing. In the Soldier sequence, the heatmap indicates gaze concentration on the hands manipulating the soldier's weapon.

## 5. Conclusion

In this study, the perceived quality of DPCs in virtual reality was analyzed through the combination of subjective scores, response times, visual attention, and objective metrics. The results showed that the R5 quality level tended to present the highest MOS values, while R3 recorded the lowest means among the analyzed cases, and that hybrid configurations exhibited content-dependent behavior, suggesting a variable influence between geometry and visual attributes. In addition, the small difference between R4 and R4_Rehbein indicated that the complexity reduction strategy can be adopted with little perceived impact on visual quality.

The integrated analysis of subjective and objective scores provided a broader interpretation of the results. The low DMOS values, moderate confidence intervals, and SOS values indicate controlled perceptual differences and moderate variability among participants. In addition, the geometric metrics p2plane and p2point showed a stronger association with perceived quality than bitrate and C0 (luminance PSNR). These findings suggest that future work may focus on reducing computational effort in texture coding and placing greater emphasis on geometry information, which appears to be more sensitive to human visual perception.

Finally, the heatmaps showed greater gaze concentration in central and visually informative regions, indicating that perceived quality depends not only on distortion levels but also on how the stimulus is visually explored. This suggests that future work may adopt Region of Interest (RoI)-based techniques to prioritize compression or rendering quality in visually relevant areas of the point cloud.

## Acknowledgements

This research was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES), Finance Code 001. It was also financed in part by the Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul - Brasil (FAPERGS), and by the Conselho Nacional de Desenvolvimento Científico e Tecnológico – Brasil (CNPq).

## References

Dumic, E. and da Silva Cruz, L. A. (2023). Subjective quality assessment of v-pcc-compressed dynamic point clouds degraded by packet losses. *Sensors*, 23(12):5623.

Graziosi, D., Nakagami, O., Kuma, S., Zaghetto, A., Suzuki, T., and Tabatabai, A. (2020). An overview of ongoing point cloud compression standardization activities: Video-based (V-PCC) and geometry-based (G-PCC). *APSIPA Transactions on Signal and Information Processing*, 9:e13.

Gutierrez, J., Perez, P., Orduna, M., Singla, A., Cortes, C., Mazumdar, P., Viola, I., Brunnström, K., Battisti, F., Cieplińska, N., et al. (2021). Subjective evaluation of visual quality and simulator sickness of short 360º videos: Itu-t rec. p. 919. *IEEE transactions on multimedia*, 24:3087–3100.

Ishihara, S. (1960). *Test for color blindness*. Kanehara Shuppan, Co. Ltd., Tokyo, Japan.

ISO/IEC (2020). Common test conditions for v3c and V-PCC. 2020.

ITU-R (2023). BT.500 : Methodologies for the subjective assessment of the quality of television images. Recommendation ITU-R BT.500-15, Geneva, Switzerland.

ITU-T (2008). Itu-tp. 910. *Subjective video quality assessment methods for multimedia applications, Recommendation ITU-T*, 910.

ITU-T (2020). Subjective test methodologies for 360º video on head-mounted displays. *International Telecommunication Union*, 919.

MPEG (2025). Video point cloud compression - VPCC - mpeg-pcc-tmc2 test model candidate software. https://github.com/MPEGGroup/mpeg-pcc-tmc2.

Neuman-Donihue, E., Jarvis, M., and Zhu, Y. (2023). Fastpoints: A state-of-the-art point cloud renderer for unity. *arXiv preprint arXiv:2302.05002*.

Nguyen, M., Vats, S., Zhou, X., Viola, I., Cesar, P., Timmerer, C., and Hellwagner, H. (2024). Compeq-mr: Compressed point cloud dataset with eye tracking and quality assessment in mixed reality. In *Proceedings of the 15th ACM Multimedia systems conference*, pages 367–373.

Noether, G. E. et al. (1981). Why kendall tau. *Teaching Statistics*, 3(2):41–43.

Preda, M. (2020). V-PCC codec description. https://www.mpeg.org/wp-content/uploads/mpeg_meeti ngs/134_OnLine/w20352.zip.

Rehbein, G., Santos, C., Corrêa, G., and Porto, M. (2025). Machine learning-based block partitioning for V-PCC encoding acceleration. In *2025 International Conference on Visual Communications and Image Processing (VCIP)*, pages 1–5. IEEE.

Schober, P., Boer, C., and Schwarte, L. A. (2018). Correlation coefficients: appropriate use and interpretation. *Anesthesia & analgesia*, 126(5):1763–1768.

Sedgwick, P. (2014). Spearman's rank correlation coefficient. *Bmj*, 349.

Singla, A., Wang, S., Göring, S., Rao, R. R. R., Viola, I., Cesar, P., and Raake, A. (2023). Subjective quality evaluation of point clouds using remote testing. In *Proceedings of the 2nd international workshop on interactive eXtended reality*, pages 21–28.

Sullivan, G. J., Ohm, J.-R., Han, W.-J., and Wiegand, T. (2012). Overview of the high efficiency video coding (HEVC) standard. *IEEE Transactions on circuits and systems for video technology*, 22(12):1649–1668.

Viola, I., Subramanyam, S., Li, J., and Cesar, P. (2022). On the impact of vr assessment on the quality of experience of highly realistic digital humans: A volumetric video case study. *Quality and User Experience*, 7(1):3.

Zhou, X., Viola, I., Alexiou, E., Jansen, J., and Cesar, P. (2023). Qava-dpc: Eye-tracking based quality assessment and visual attention dataset for dynamic point cloud in 6 dof. In *2023 IEEE international symposium on mixed and augmented reality (ISMAR)*, pages 69–78. IEEE.

Zhou, X., Viola, I., Alexiou, E., Jansen, J., and Cesar, P. (2025). Subjective and objective quality assessment for dynamic point cloud with visual attention in 6 dof. *ACM Transactions on Multimedia Computing, Communications and Applications*, 21(8):1–24.