

# Evaluating the Influence of the HMD, Usability, and Fatigue in 360VR Video Quality Assessments

Marta Orduna\* Pablo Pérez† César Díaz\* Narciso García\*

\*Grupo de Tratamiento de Imágenes, ETSI Telecomunicación, Universidad Politécnica de Madrid, Spain

†Nokia Bell Labs, Madrid, Spain

## ABSTRACT

VR communications present great challenges that should be addressed in subjective assessment methodologies. Thus, the HMD, the usability of controllers and touchpad, or even the fatigue produced by the test session are factors that highly influence the QoE. In addition, in a 360VR scenario, most of the experiments are focused on the evaluation of video quality or the evaluation of socioemotional concepts such as the sense of presence. However, no standardized methodology to evaluate both kinds of concepts is available nowadays. In this paper, we present an experiment where video quality and sense of presence are jointly assessed in two of the most popular HMDs. As a result, we have observed that while presence is affected by the evaluation mechanism or duration of the test, quality is mainly affected by the HMD. This statement implies that methodologies that take into account technical concepts, such as encoding and transmission, and socioemotional concepts are necessary to obtain reliable and robust results of QoE in a VR environment.

**Index Terms:** Virtual Reality—Video Quality—Presence—Subjective Assessment—Usability

## 1 INTRODUCTION

Virtual Reality (VR) applications are achieving wide interest for remote communications. The Quality of Experience (QoE) is typically determined by the video quality and socioemotional concepts, such as the sense of presence provided to users. Nevertheless, it is important to note that in 360VR environment, QoE is also dependent on factors such as the evaluation mechanism, the fatigue produced by the duration of the test or the HMD, elements that are not taken into account working with traditional contents.

Referring to the literature, on the one hand, we have noticed that most of the studies focused on QoE do not consider detailed socioemotional questionnaires [2,9]. On the other hand, we have found that there are several articles that make a great effort to analyze socioemotional concepts such as the sense of presence, attitude change, and perspective-taking, and the influence of the platform and headphones in the QoE [1, 11], avoiding the effects of technical conditions. Tran et al. [10] analyze the influence of both video quality and some socioemotional concepts with an adapted questionnaire. Singla et al. [7] present a QoE experiment where different socioemotional questionnaires are compared. Despite the contribution of knowledge that these experiments provide to the state of the art, they are still missing the evaluation of highly influential conditions that affect the QoE in a 360VR environment and the differences between the selected questionnaires and scales. For that, the design and implementation of new standardized methodologies are required to obtain robust and reliable results of QoE in 360VR applications.

\*e-mail: moc, cdm, narciso@gti.ssr.upm.es

†e-mail: pablo.perez@nokia-bell-labs.com

	1st Condition	2nd Condition
Test A	Samsung + touchpad	Lenovo + controller
Test B	Lenovo + controller	Samsung + touchpad

Table 1: Experiment conditions

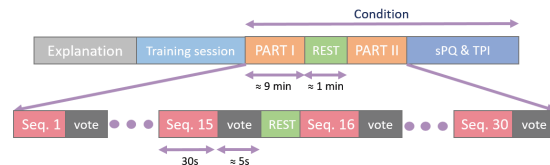


Figure 1: Condition structure

## 2 EXPERIMENT DESIGN

Based on the literature review, we present an experiment where the influence of different conditions, in terms of video quality, HMD, and evaluation mechanisms, on the sense of presence is evaluated. In addition, a comparative between two different questionnaires of presence is presented. Here, we explain in detail the subjective assessment carried out, considering two tests denoted A and B, as can be observed in Table 1. Detailed experiment results can be found in the supplementary material<sup>1</sup>.

**Test Material.** We used six representative sources with the original resolution 4K resolution and framerates to preserve the original characteristics. All sequences were selected with audio to facilitate the user immersion [11]. Also, they were encoded with HEVC using fixed Quantization Parameters (QPs): 22, 27, 32, 37 and 42 [3].

**Methodology.** The methodology applied in this experiment is the ACNR-HR (Absolute Category Rating with Hidden Reference) with a five-level rating scale, as recommended in ITU-T P.910. Regarding the sense of presence and usability were evaluated with the spatial presence scale of the Temple Presence Inventory (TPI) [4] and the subsampling of the Presence Questionnaire (sPQ) [12], both questionnaires rated on a 7-point Likert scale. The sPQ is composed by representative items that evaluate control, sensory, and realism factors. Specifically, items 3, 10, 12, 14, 26, 27, and 30 of the original questionnaire. Items 3, 27, and 30 were used to evaluate usability in the two conditions, controllers and touchpad.

**Environment.** Subjective tests were carried out in two popular HMDs: Samsung Galaxy S8 with Samsung Gear VR, which includes a touchpad on its right side, and Lenovo Mirage Solo with a Daydream handled controller (Table 1). The observers were located in the middle of a room, being able to spin around without any limitation while seated on a swivel chair.

**Observers.** A total of 48 observers (21 females, 27 males), with an age between 20 and 26, participated in this experiment.

**Test session.** Subjects received the instructions for the experiment at the beginning of the test. A training session to show the best and the worst qualities offered was carried out after the explanation. The quality was evaluated using an application not requiring

<sup>1</sup>www.gti.ssr.upm.es/data/

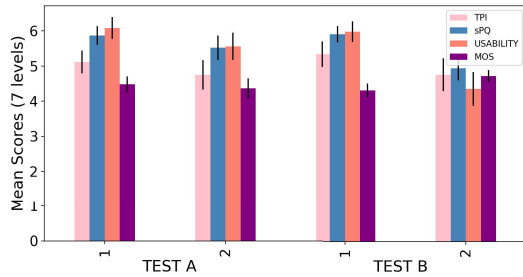


Figure 2: TPI - sPQ - Usability - MOS

to remove the goggles [6]. All videos were viewed and evaluated on both devices by every subject. Half of them took test A and the other half test B (see Table 1). The order of the conditions is the unique difference between them and the structure of each condition can be seen in Figure 1. The duration of the whole test was around 40 minutes (20 minutes for each condition session with a break in the middle). At the end of each condition session, the presence questionnaire was evaluated.

### 2.1 Hypotheses

- The HMD and the order in which conditions are evaluated have influence on: sense of presence (**H1**), quality (**H2**), and perceived usability (**H3**).
- sPQ and TPI provide similar measurements (scores) for spatial presence (**H4**).

### 2.2 Experimental Results

The quality evaluation was examined with the Mean Opinion Score (MOS) and the Differential Quality Score (DMOS) from the subject's rates, and the associated 95% Confidence Intervals (CI) [8]. In relation to the sPQ and TPI analysis, the Pearson & D'Agostino normality test was applied to validate the normal distribution of the collected data. Then, the 2-way Analysis of Variance (ANOVA) was applied to examine the differences between conditions [5]. Figure 2 presents the mean scores of TPI, sPQ, Usability, and quality (MOS). Note that MOS was rescaled to a 7-level scale, allowing the comparison between the evaluated assessments in both tests and conditions.

The first hypothesis sets out how the HMD and the evaluation order affect the sense of presence evaluated with TPI and sPQ independently. The TPI and sPQ results showed that there is not a significant effect on the test condition ( $F_{1,46} = 0.266$ ,  $p > .05$  and  $F_{1,46} = 3.896$ ,  $p > .05$  respectively), but there is a significant influence on the order of the conditions ( $F_{1,92} = 5.507$ ,  $p < .05$  and  $F_{1,92} = 17.761$ ,  $p < .05$ ). It makes sense because the sense of presence is mainly affected by the novelty of the 360VR experience.

The second hypothesis refers to quality evaluation. Unlike the sense of presence, the quality ratings showed that there is not a significant difference between the evaluation during the first or second condition of the test ( $F_{1,92} = 1.662$ ,  $p > .05$ ). In this way, we can assure that subjects perfectly discern between the evaluation of the sense of presence and quality. Likewise, quality evaluation is highly dependent on the used HMD ( $F_{1,92} = 5.102$ ,  $p < .05$ ) where Samsung HMD obtained the best quality evaluations.

The third hypothesis is related to the usability of touchpad or controller. It was measured with the aggregation of three items of the sPQ. This aggregated measure showed a statistically significant difference between conditions (touchpad or controller), ( $F_{1,92} = 8.351$ ,  $p < .001$ ). However, it showed a stronger effect taking into account the order of the evaluation, first or second condition of the test session ( $F_{1,92} = 31.517$ ,  $p < .001$ ), as well a significant effect on the interaction of both conditions ( $F_{1,92} = 11.976$ ,  $p < .001$ ), as presented in Figure 2.

Finally, the fourth hypothesis formulates the comparison between the obtained results with the sPQ and TPI. For that, we computed the Linear Pearson Correlation Coefficient (LPCC). We showed that there is a relevant correlation between both questionnaires ( $r = 0.5279$ ,  $p < .05$ ) but based on the differences, as can be observed in Figure 2, we can not assume that sPQ and TPI measures exactly the same concept.

## 3 CONCLUSIONS

We have carried out an experiment combining the video quality evaluation, well-known in the literature, with a novel characteristic of 360VR technology: the evaluation of the sense of presence. We will make public the database obtained from the experiment. We can conclude that the evaluation of presence is very dependent on the first or second condition of the test, regardless of the device. Instead, quality is always evaluated in the same way and it is mainly affected by the device. Also, there is a big difference in the evaluation mechanism used, which can affect the sense of presence. These factors need to be addressed when designing QoE evaluation methodologies for 360VR video. Finally, we have shown that different presence scales (sPQ, TPI) evaluate highly correlated, but different, socioemotional concepts.

### ACKNOWLEDGMENTS

This work has been partially supported by the Ministerio de Ciencia, Innovación y Universidades (AEI/FEDER) of the Spanish Government under project TEC2016-75981 (IVME).

### REFERENCES

- [1] T. Aitamurto, S. Zhou, S. Sakshuwong, J. Saldivar, Y. Sadeghi, and A. Tran. Sense of Presence, Attitude Change, Perspective-Taking and Usability in First-Person Split-Sphere 360 Video. In *Proceedings of the 2018 CHI Conf. on Human Factors in Comput. Syst.*, p. 545, 2018.
- [2] M. S. Anwar, J. Wang, A. Ullah, W. Khan, Z. Li, and S. Ahmad. User Profile Analysis for Enhancing QoE of 360 Panoramic Video in Virtual Reality Environment. In *2018 Int. Conf. on VR and Visualization (ICVRV)*, pp. 106–111, 2018.
- [3] F. Bossen. Common HM Test Conditions and Software Reference Configurations. *Joint Collaborative Team on Video Coding (JCT-VC)*, 11th meeting, output document M27343, Oct. Shanghai, China, 2012.
- [4] M. Lombard, T. B. Ditton, and L. Weinstein. Measuring Presence: the Temple Presence Inventory. In *Proceedings of the 12th Annual Int. Workshop on Presence*, pp. 1–15, 2009.
- [5] M. Narwaria, L. Krasula, and P. Le Callet. Data Analysis in Multimedia Quality Assessment: Revisiting the Statistical Tests. *IEEE Trans. on Multimedia*, 20(8):2063–2072, 2018.
- [6] P. Pérez and J. Escobar. MIRO360: A Tool for Subjective Assessment of 360 Degree Video for ITU-T P. 360-VR. In *2019 11th Int. Conf. on Quality of Multimedia Experience (QoMEX)*, pp. 1–3, 2019.
- [7] A. Singla, R. R. Rao, S. Göring, and A. Raake. Assessing Media QoE, Simulator Sickness and Presence for Omnidirectional Videos with Different Test Protocols. In *IEEE Conf. on VR and 3D User Interfaces*, pp. 1163–1164, 2019.
- [8] R. C. Streijl, S. Winkler, and D. S. Hands. Mean Opinion Score (MOS) Revisited: Methods and Applications, Limitations and Alternatives. *Multimedia Syst.*, 22(2):213–227, Mar. 2016.
- [9] J. Thompson, A. K. Karembai, and P. Seeling. Immersive Image QoE in Mobile Consumer Virtual Reality Settings. In *16th IEEE Consum. Commun. Networking Conf. (CCNC)*, pp. 1–4, 2019.
- [10] H. Tran, N. Ngoc, C. Pham, Y. Jung, and T. Thang. A Subjective Study on QoE of 360 Video for VR Communication. In *IEEE 19th Int. Workshop on Multimedia Signal Process. (MMSP)*, pp. 1–6, 2017.
- [11] A. Tse, C. Jennett, J. Moore, Z. Watson, J. Rigby, and A. L. Cox. Was I There?: Impact of Platform and Headphones on 360 Video Immersion. In *CHI Conf. Extended Abstracts on Human Factors in Comput. Syst.*, pp. 2967–2974, 2017.
- [12] B. G. Witmer and M. J. Singer. Measuring Presence in Virtual Environments: A Presence Questionnaire. 7(3):225–240, 1998.