Subjective QoE assessment method for 360° videos

Taichi Kono¹ and Takanori Hayashi¹ a)
¹Graduate School of Science and Technology, Hiroshima Institute of Technology
2-1-1 Miyake, Saeki-Ku, Hiroshima-shi, Hiroshima 731-5193, Japan
a) t.hayashi.xk@it-hiroshima.ac.jp

Abstract: This paper proposes a subjective assessment method for video quality and sense of presence for 360° videos. Since it is necessary to view 360° videos with a wide field of view, we adopted a subjective assessment method that allows repeated viewing of 10 s long videos. First, by analysis of the subject’s observation behavior (i.e., head movement), it was clarified that the appropriate number of observation repetitions was 2. The relationship between the video quality and the sense of presence was then quantified using the proposed method. Finally, we determined the number of subjects necessary to derive stable evaluation results.

Keywords: QoE, 360-degree video, subjective assessment, video quality, sense of presence, head movement

Classification: Multimedia Systems for Communication

References

[1] Huiyu Duan, Guangtao Zhai, Xiaokang Yang, Duo Li, and Wenhan Zhu, “IVQAD 2017: An Immersive Video Quality Assessment Database,” IWSSIP 2017, 2157-8702, May 2017.
[2] Huyen T. T. Tran, Cuong T. Pham, Nam Pham Ngoc, Anh T. Pham, and Truong Cong Thang, “A Study on Quality Metric for 360 Video Communications” IEICE Transaction on Information and Systems, vol. E101-D, no. 1, pp. 28-36, January 2018.
[3] Hak Gu Kim, Heoun-Taek Lim, Sangmin Lee, and Yong Man Ro, “VRSA Net: VR Sickness Assessment Considering Exceptional Motion for 360° VR Video,” IEEE Transactions on Image Processing, vol. 28, no. 4, pp. 1646-1660, April 2019.
[4] ITU-T Recommendation P.910, “Subjective video quality assessment methods for multimedia applications,” April 2008.
[5] ITU-T Recommendation P.915, “Subjective assessment methods for 3D video quality,” March 2016.

1 Introduction

In recent years, with the development of virtual reality (VR) video technology, various services and applications using VR videos have become widespread, especially in the fields of entertainment, architecture, medical science, education,
tourism, etc. Designing VR systems based on quality of experience (QoE) is important for users’ comfort. To do this, it is necessary to understand the QoE characteristics of VR videos from various viewpoints.

Conventional studies have mainly focused on video quality, sense of presence, immersion, and viewing safety as QoE factors for VR videos. Duan et al. [1] analyzed the effects of the coding rate, video resolution, and frame rate on VR video quality. Tran et al. [2] studied appropriate objective quality metrics for 360° videos. Kim et al. [3] proposed an objective VR sickness assessment network. However, subjective quality assessment methods for VR videos are not well discussed and established. In the absolute category rating (ACR) method, which is a typical subjective quality assessment method for 2D/3D videos, the video quality is assessed using a test sequence of approximately 10 s [4]. In ITU-T and VQEG, however, it has been argued that 10 s may be too short of a viewing time for evaluating the quality of 360° video.

This paper proposes a subjective assessment method in which the same video sequence is repeatedly viewed so that the video can be viewed from various directions and evaluated. First, by analyzing the subject's head movement to measure observation behavior, the appropriate number of observation repetitions was clarified. The relationship between the video quality and the sense of presence was then quantified using the proposed method. Finally, we determined the number of subjects necessary to derive stable evaluation results.

2 Head movement analysis during subjective assessment

2.1 Head movement measurement

It is necessary to secure sufficient time to view 360° videos from various directions to stably assess the video quality and sense of presence. Therefore, a QoE assessment method that observes a 10-second video sequence multiple times is a candidate. To clarify the subject's viewing behavior characteristics when the number of viewings was changed, the subject's head movement during the assessment test was measured.

In this study, an environment that allows viewing a 360° video was constructed using Unity. HTC VIVE was used for a head-mounted display (HMD), and head movement was measured using the HMD accelerometer. As video contents, nine kinds of 10-second video sequences were taken with a fixed 360-degree camera. The quality of each video was varied to create three grades and a total of 27 test conditions.

For subjective assessment, we used three methods in which the number of viewings of the test video sequence was changed from 1 to 3, as shown in Fig. 1. If the subject views the same video content multiple times, the viewing behavior may change. For this reason, each subject assessed different video contents with different qualities when the number of viewings was changed. A total of 9 conditions were assessed in three experiments with different viewing times. Therefore, the results for one assessment method included a total of 54 conditions. The subjects were 18 non-experts (16 male and 2 female college students) in video quality.
In the head movement measurement, pitch, yaw, and roll angles were sampled at 90 Hz from the HMD accelerometer and acquired to analyze the head movements of the subject during viewing. In this analysis, however, roll angles that had little influence on the analysis of the viewing behavior were excluded. From this data, the average amount of movement of the subject’s head for 10 s was derived. After these experiments were completed, we asked the subjects to answer a questionnaire about the appropriate number of viewings to assess for 360° videos.

2.2 Experimental results
As a result of the measurements, the maximum amount of head movement of the subjects was 180° for the pitch angle and 430° for the yaw angle. Because the yaw angle of many subjects was larger than the pitch angle, only the yaw angle data was used to analyze head movement.

As a result, when the number of viewings was 1, 2, and 3, the average amount of movement was 417.5°, 370.5°, and 373.3°, respectively. There was a statistically significant difference at the 5% level between the result when the number of viewings was 1 and the result when the number of viewings was 2 or 3. This is because head movement during viewing was stable because of the time allowed to view the video sequence multiple times. Moreover, according to the results of a questionnaire asking about the appropriate number of viewings for assessment of 360° video, the average number of viewings was 2.28. Therefore, it is appropriate to set the number of viewings to 2.

3 Evaluation of relationship between video quality and sense-of-presence for 360-degree videos
In this section, the relationship between video quality and sense of presence of 360° videos is derived using the proposed subjective assessment method described in the previous section. Moreover, the number of subjects required to obtain stable evaluation results is determined.

3.1 Subjective assessment test
The environment for subjective assessment was the same as that described in the previous section, except that a SteamVR Media Player was used as the application to play the 360° videos. The test sequences were 10-second videos shot by a fixed camera, and the following three types were prepared by considering the spatial definition and movement of the videos (video coding method: H.264/MPEG-4 AVC, video resolution: 3840 × 1920p, video Frame rate: 30 fps):

- Penguins: A crowd of penguins all around,
- Elephants: One of the elephants in a meadow approaching,
- Seal: A seal swimming in the sea.

These videos were re-encoded to change their qualities in four grades. Therefore, the number of test conditions was 12.

The subjective assessment method proposed in the previous section was used, and the number of viewings of the test sequence was 2. The video quality was evaluated on a five-grade quality scale (5: excellent, 4: good, 3: fair, 2: poor, 1: bad). Moreover, the sense of presence was defined as "the feeling of seeing the real thing and/or the feeling of being there," and evaluated on a five-grade scale (5: extremely present, 4: present, 3: neither, 2: absent, 1: not at all). The subjects were 25 non-experts (22 male and 3 female college students) in video quality. The subjective video quality and sense of presence were represented as a mean opinion score (MOS) calculated by averaging the scores of 25 subjects.

### 3.2 Relationship between video quality and sense of presence

Figure 2 shows that the MOS for sense of presence tend to be higher than that of video quality, regardless of the type of video content. That is, it was found that the sense of presence decreases as the video quality deteriorates, but the amount of decrease in the sense of presence tends to be less than that of the video quality. This is because even if the video quality is slightly lower, the subject can view the 360-degree video from a free-viewpoint, and a certain sense of presence is maintained.

![Fig. 2. Relationship between video quality and sense of presence.](image)

### 3.3 Number of subjects required for subjective assessment

The relationship between the number of subjects and the stability of the MOS was analyzed using the results obtained in subsection 3.2. The stability was expressed...
as the mean of the 95% confidence interval (MCI) for the MOS.

The stability of the MOS when the number of subjects was less than 25 was derived by averaging the MCI obtained from randomly selected subjects in three patterns. Figure 3 shows the relationship between the number of subjects and the average MCI. It was found that the average MCI decreased as the number of subjects increased.

According to ITU-T Recommendation P.915 [5], which defines the quality assessment method for 3D videos, the required number of subjects for assessing 3D videos was derived based on an MCI of about 0.32 when 2D videos were assessed by 24 subjects. Therefore, we determined that the required number of subjects for 360° videos should be 0.32 or less. As a result, it was found that the number of subjects assessing the video quality and sense-of-presence was at least 17 and 23, respectively. Therefore, it is more difficult to assess the sense of presence than the video quality.

![Fig. 3. Relationship between the number of subjects and the averaged MCI.](image)

4 Conclusion

This paper proposed a subjective assessment method for 360° videos in which the same video sequence is viewed repeatedly. The results of measuring the subject's head movements and the questionnaire, showed that the number of repetitions should be 2. Using the proposed assessment method, we quantified the relationship between video quality and sense of presence for 360° videos and analyzed the number of subjects required to obtain stable evaluation scores. As a result, it was found that the tendency of the sense of presence to decrease was slower than that of the video quality for 360° videos. Moreover, from the viewpoint of stability of evaluation scores, it was shown that 18 or more subjects are required for video quality assessment, and 23 or more subjects are required for sense of presence assessment. In the future, we need to clarify the evaluation characteristics of video quality and sense of presence when 360° videos are distributed over a network and to establish a subjective assessment method for operability in virtual space.

Acknowledgments

We would like to thank Editage (www.editage.com) for English language editing.