https://doi.org/10.2352/ISSN.2470-1173.2020.11.HVEI-091

Quality of Experience Assessment of 360-degree video

Anouk van Kasteren[a,c], Kjell Brunnström[a,b], John Hedlund[a], Chris Snijders[c]
[a]RISE Research Institutes of Sweden AB, Kista, Sweden
[b]Mid Sweden University, Sundsvall, Sweden
[c]Eindhoven University of Technology, Eindhoven, The Netherlands

Abstract

The research domain on the Quality of Experience (QoE) of 2D video streaming has been well established. However, a new video format is emerging and gaining popularity and availability: VR 360-degree video. The processing and transmission of 360-degree videos brings along new challenges such as large bandwidth requirements and the occurrence of different distortions. The viewing experience is also substantially different from 2D video, it offers more interactive freedom on the viewing angle but can also be more demanding and cause cybersickness. Further research on the QoE of 360-videos specifically is thus required. The goal of this study is to complement earlier research by (Tran, Ngoc, Pham, Jung, and Thank, 2017) testing the effects of quality degradation, freezing, and content on the QoE of 360-videos. Data will be gathered through subjective tests where participants watch degraded versions of 360-videos through an HMD. After each video they will answer questions regarding their quality perception, experience, perceptual load, and cybersickness. Results of the first part show overall rather low QoE ratings and it decreases even more as quality is degraded and freezing events are added. Cyber sickness was found not to be an issue.

Introduction

Multimedia streaming has been gaining great popularity amongst consumers everywhere. The number of streaming services (e.g. Netflix, Amazon prime, and HBO) has been growing, and the available content even more so. The Majority of the offer is 2D video streaming such as movies and TV shows. However, a new format, "VR 360-degree video", omnidirectional content that can be viewed through a head-mounted display (HMD), is growing in popularity. This format offers a much more immersive viewing experience compared to 2D video. Popular streaming platforms such as YouTube, Vimeo and Fox Sports are increasing their offer of 360-degree content. Additionally, HMD's are becoming more affordable and available for the general public allowing for a larger audience to access these omnidirectional videos. For acceptance and application of this media format in people's everyday life, an understanding of how to provide a pleasant viewing experience while efficiently utilizing network resources is needed. In multimedia research the Quality of Experience is an important measure, which is defined by the International Telecommunication Union (ITU) standard or Recommendation as follows: "Quality of Experience (QoE) refers to the overall acceptability of an application or service, as perceived subjectively by the end-user." (ITU-T 2017)[1]. There are many factors that can influence the QoE. An influencing factor (IF) is any characteristic of a user, system, application, or context whose actual state or setting may influence on the QoE (Callet et. al., 2012)[2]. As with 2D video, studying the QoE is important in the development and improvement of 360-video technology. Even more so as applying methods and theory on 2D video streaming to 360-video is not trivial. The processing and transmission of 360-degree format brings along new challenges such as large bandwidth requirements and the occurrence of different distortions. Some are new to 360 video e.g. stitching artifacts, warped blocking artifacts and motion-to-photon delay[3]. The viewing experience is also substantially different from 2D video; it offers more interactive freedom on the viewing angle but can also be more demanding and cause cybersickness. This paper is based on a M.Sc thesis study also involving the influence visual attention based on eye-tracking[4]. The goal of the M. Sc. thesis was to complement earlier research by gathering subjective data on the effects of different influence factors such as quality degradations, freezing events and content on the QoE, cybersickness and the perceptual load of 360-videos. Results may be used to define quality recommendations and contribute to the improvement of objective QoE metrics.

Now let's first go back and discuss QoE a bit more. As mentioned earlier, many factors could influence the QoE. There are Human IFs, System IFs, and Context IFs. Human IFs are characteristics of the user such as demographics, socio-economic background or the physical and mental/emotional state. System IFs are properties of characteristics influencing the technical quality of a product and can be content-, media-, network-, or device-related. Context IFs are situational properties and the physical, temporal, social, economic and technical characteristics of a user's environment. In the end, a combination of influence factors defines the QoE. (Reiter et al., 2014)[5]. QoE research focusses on studying how IFs are related to the user's experience to support the development of media technologies and has become an important domain with the growing availability of multimedia. Important system IFs on the QoE of 2D video streaming are among others: viewing distance, display size, resolution, bitrates, network performance (Kuipers, Kooij, Brunnstrom, 2018)[6]. Furthermore, distortions and artifacts that occur during the different processing steps also have negative influence on the QoE,(Möller, Raake and Küpper 2014)[7] with distortions in salient regions having even more so. (Engelke, Pepion, Callet, and Zepernick, 2010)[8]. To simply apply the theory and methods of 2D video to the 360-video domain is not trivial and requires more specific research on how new challenges and a different viewing experience that come with 360-video relate to the QoE.

The following section will first discuss these challenges and differences in a bit more detail. 360-Videos are recorded with multiple dioptric cameras each capturing a different angle. The input from these cameras is then stitched together by a mosaicking algorithm. Artifacts may occur due to inconsistency between the cameras. For example, the light could be different in different viewpoints. As the material is stitched together, issues could arise causing quality artifact and distortions. How these new artifacts affect the QoE and users viewing behavior has yet to be determined. The transmission of 360-videos is a great challenge due to the large required bandwidth. Today, 4K resolution is accepted as a minimum functional resolution but requirements are increasing with prospects of 8K, and even 16K resolutions to be stored and transmitted efficiently (Azevedo et al., 2019)[3]. To lower the bandwidth requirements, video material has to be compressed to lower qualities which causes compression artifacts which may negatively affect the experience (Azevedo et al., 2019)[3]. The right balance between compression, latency and user satisfaction has thus to be found. Watching VR 360-videos through an HMD offers a much more immersive experience. The level of immersiveness and presence a person experiences is related to the video quality and influences the QoE (Tran, Ngoc, Pham et al., 2017)[9]. Furthermore, delays, rebuffering and quality fluctuations due to networks reaching their limits, could cause confusion and cybersickness by disrupting the course of movements which negatively influences the viewing experience (e.g. Hettinger and Riccio, 1992; Porcino, Clua, Trevisan, Vasconcelos and Valente, 2017) [10, 11]. The HMD is also much closer to the eye compared to conventional 2D displays, which may increase the visibility of distortions and induce more eye strain and fatigue which makes the experience more demanding. Space between pixels may also be visible due to the closeness of the screen to the eyes (Azevedo et al., 2019)[3]. The quality perception and experience may, therefore, be different in VR 360-videos compared to regular 2D video. Additionally, the 360-degree material allows for a new way of interacting with media. In contrary to 2D video, users have more freedom in deciding from what angle to watch the content. The total image is also larger than the users' field of view, therefore, different users may view different parts and proportions of the content and each explores it in a different order. Visual attention would thus be an interesting IF in 360-video and its relation to the QoE should be studied. In summary, QoE research is important for the development of video streaming technology to most efficiently handle the trade-off between providing good quality and limiting network burden. Additionally, 360-video offers a substantially different experience compared to regular 2D, therefore, it would be prudent to do more research on the QoE in 360-video specifically.

So far not many studies have been conducted, and subjective methods have yet to be standardized. Nevertheless, some studies have been performed adapting methodologies from 2D video quality assessments. An elaborate study by Tran, Ngoc, Pham et al. (2017)[9] tested the effects of several IFs such as resolution, bitrate, content, and camera motion and viewing device on the perceptual quality, presence, cybersickness and acceptability of 360-video. In general, their results show that overall acceptance is high but drops significantly for resolutions below 2K. They also found that videos with QP 40 or higher the acceptance level drops below 30%, but that QP values 22 and 28 did not significantly differ in acceptance. Furthermore, cybersickness was found to be a serious problem as 93% felt dizzy or nauseous while watching 360-videos. Their results are a step towards understanding the QoE in 360-video and will additionally be used in the development of objective QoE metrics. Another study looked at the effects of stalling events on the QoE and annoyance (Schatz, Sackl, Timmerer and Gardlo, 2017)[12]. Their results show that even a single stalling event leads to a significant increase in annoyance and should thus be avoided. However, different patterns did not necessarily result in different annoyance scores.

Method

Design

The study had a 4 x 3 x 3 within-subject design. The main dependent variable of the study is the QoE of 360-videos, which is measured through four subjective post-hoc evaluations. After viewing a video, participants were asked to rate their perceived video quality, cybersickness, overall experience, and perceptual load. While viewing the videos the eye-movements of the participants were tracked. Based on the studies. To test the hypothesis, the manipulated independent variables are:

- the video content: there were 3 different;
- the video quality: degrading the videos in 4 levels;
- freezing events in no freezing, low and high frequency.

Participants

Based on the a priori method proposed in a recent paper by Brunnström and Barkowsky (2018)[13] on sample size estimation in QoE studies, the number of minimum required participants was estimated as a function of the planned comparisons, the target MOS difference and the desired power. For the final calculation the expected MOS difference was based on results of the effect of resolution on perceived quality which was the smallest effect found resulting in the largest required sample. Based on this analysis, 33 participants would be needed to obtain 90% power, as was required by the Human Technology Interaction group. Participants were recruited internally at the RISE Kista office or via known affiliates. Participants were required to have good vision (through aids that do not interfere with the HMD), have no other severe eye or vision impairments and do not have progressive glasses or contact lenses. In total 33 people did also participate in the experiment. One participant had to be excluded due to such severe software issues that most data was missing and including the participant would not be possible. Due to some last-minute cancelations, a number of 33 participants was not reached, leaving data of 32 participants available for analysis, 22 of which were male and 10 were female. Participants were aged 21 to 44 (Mean=28.8, SD=6.3), 9 of them had glasses and 3 had contact lenses, one suffered from minor colorblindness. Acuity was measured before the experiment and results were between 0.65 and 1.5 (Mean=1.1, SD=0.2).

Experiment Setup and Stimulus Materials

The experiments took place at the RISE office in Kista in one of the labs designated for subjective testing. The lab was 2 by 3 meter. Inside the lab room were 2 cabinets, a high table to fill out forms, a chair for the participant to sit on and a chair and desk with a computer used by the researcher. The participant's chair allowed the participant to turn, allowing them to be able to fully explore the 360 environments. Hardware used in the experiments was an ASUS ROG ZEPHYRUS (GX501) laptop running on Windows 10 and an HTC VIVE VR headset with integrated Tobii eye trackers. For this headset the maximum recommended resolution of 360 video material was 4096 x 2048 @ 30 fps. Supported file formats included for images: jpeg, png, and bmp; for videos: mp4 and avi with codecs H264, DIVX, or XVID. The software used to conduct the experiments was Tobii pro lab v1.111. In the program video or image material can be uploaded and prepared in timelines. Seven separate timelines were created for the current study. One timeline for the practice trial, and for each video a reference and degradations timeline. In addition to the videos, images with the questions were added to the timeline after each video. The order of the videos was randomized within the software; however, the order of the timelines cannot be randomized automatically for which then a dice was used to determine the order on the spot.

Overview of the video stimuli.

Three different videos were selected to be used in the experiment. These videos were selected as the source was known, and there was access to the pristine version of the video. This limited the number of available videos. However, it was attempted to find distinctive material. Table 1 shows the properties of each of the original video files. Figure 1 shows still images of each of the videos. Of each of the videos, the motion vectors were calculated as well as the spatial and temporal index (SI and TI) (Figure 2 & 3)[14]. The motion vector example script of FFmpeg was used to generate the motion vectors, and for the SI and TI a python program called “siti” written by W. Robitza (2017-2019) was used. These can be used to classify the videos based on their motion activity. However, with the limiting number of available videos this distinction is less prominent as aimed for initially. Looking at the SI, TI and motion vectors it can be seen however the TI is slightly lower for the intersection video. Comparing the Dojo and Flamenco video, the movements in the dojo video are more intense, hence, larger spikes in the motion vector graphs (Figure 2). The spatial index is slightly higher for the flamenco video as there are simply more moving objects (people in this case). Another distinction between videos is the lighting. The intersection video is recorded outside with natural lighting, the flamenco video is inside but with a reasonable number of windows and the Dojo video is inside in a more closed of space with mainly artificial lighting. All videos have a static camera position in common. With FFmpeg the videos were cut, encoded to H.264 .mp4 with resolution 3840x2160 format and degraded to lower qualities by changing the Constant Rate Factor (CRF). CRF is a quality setting of the encoder, which tries to output video with constant quality. The values are ranging between 0 and 51. Higher values result in more compression. To add the freezing events, the program “bufferer” written by W. Robitza (2017-2019) was used. The bufferer program will freeze the video and add a spinner at desired moments for a set time. In total, it resulted in 36 videos that was used in the experiment.

Table 1: Properties of the original videos used in the experiment.

Name	Original resolution	Original encoder	fps	Bitrate	Size	Duration	Source	Setting	AOI	Bottom logo
Dojo	3840x2160	h264	30	40 mb/s	1.2GB	4:19min	Nantes/ Nokia	inside	People	Yes
Flamenco	3840x2160	h264	30	40 mb/s	0.92GB	3:10min	Nantes/ Nokia	inside	People	Yes
Inter-section	7680x3840	h265	30	157 mb/s	5.5GB	5min	Inter-digital	outside	Cars	No

Figure 1: Still images of the video stimuli. From left to right: Dojo, Intersection, Flamenco.

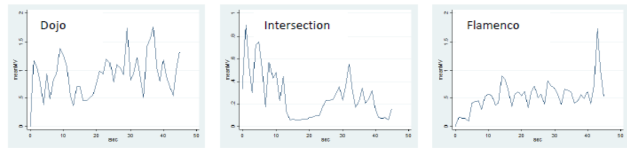

Figure 2: Motion vector graphs over time for the three videos. Left to right: Dojo, Intersection, Flamenco.

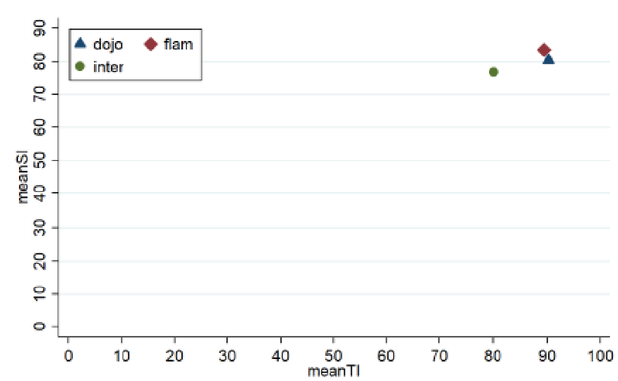

Figure 3: The mean Spatial (SI) and Temporal (TI) index of the three videos.

Measurements

Variables used in the study can be divided into four categories: subjective measurements (dependent variables), manipulation variables (independent variables), eye-tracking data (both dependent and independent variables) and the other variables.

Subjective measures

The subjective measurements were collected by asking four questions after each video. The questions were chosen based on the hypotheses and previous research by Tran, Ngoc, Pham et al. (2017)[9] and Schatz et al. (2017)[12]. These questions were accompanied with a 5-point scale and participants verbally answered by stating the number corresponding to their answer. These four questions are:

- “Did you experience nausea or dizziness?”
  1 (not at all) – 5 (a lot)
- “How would you rate the video quality?”
  1 (very low) – 5 (very high)
- “How much cognitive and perceptual load did you experience?”
  1 (none) – 5 (a lot)
- “How would you rate your overall experience?”
  1 (unpleasant) – 5 (pleasant)

To verify whether participants understood how to answer the questions, they were asked to describe them prior to the experiment. When needed they would be corrected, and additional explanation would be given. Especially the cognitive and perceptual load was found to be complicated. All participants were instructed to consider it as the effort needed to visually and cognitively perceive the video.

These four questions resulted in the four variables in the dataset perceived video quality (VQ), overall experience, cybersickness, and perceptual and cognitive load (PCL).

Manipulations.

The second category was the manipulations that served as independent variables. The CRF (see explanation above) value is a quality parameter. Increasing the CRF value will result in more compression and thus lower video quality. Videos were generated with CRF values of 15 (visible lossless), 20, 28, and 36. CRF was added at categorical variable. The second manipulation variable was the freezing event frequency. Freezing events of three seconds were added in different frequencies to the videos. Videos either had no freezing (none), 2 freezing events (low), or 4 freezing events (high). The third manipulation was the content of the video. There were 3 different videos: Dojo, Intersection, and Flamenco. These are described in the stimulus material section above.

Procedure

Participants were received and welcomed at the lab after which a short introduction of the experiment was given. Participants were sent instructions beforehand, which were discussed to make sure the participants have understood the procedure and the meaning of the subjective questions.. If everything was clear a consent form was signed after which a vision test was done to measure participant's acuity and possible color deficiencies. The participant then filled out a form asking their demographics. Once the form was completed the participant was seated and instructed on how to adjust the headset. First, a training round was done for the participant to get acquainted with the VR environment and the question format. The training sequence consisted of one omnidirectional image to set up the lenses, an example of the calibration, and videos acquitting participants with different quality and freezing conditions. All four questions were asked once during the training to give the participant an idea of how to answer these. If everything was clear after the training sequence, the first experiment sequence was shown. The experiment consisted of a total of 3 video contents with each 11 degraded videos and one reference video. The order of the sequences was determined by rolling a dice. The order of the videos within a sequence was randomized within the software, except for the reference, this video was shown first. After each video the 4 rating questions were shown one by one inside the headset and the participants were asked to answer verbally on a 1 – 5 scale. After each video sequence the participant could take off the headset for a short break. This was repeated for all 3 video contents. After all videos were viewed the participant handed the HMD back over to the researcher and they were given a movie ticket as compensation.

Results

QoE was measured through the perceived quality and the overall experience and expressed in a Mean Opinion Score (MOS). Values range from 1 to 5 (mean = 3.10; SD = 0.97). Figure 4 shows the QoE MOS as an interaction plot of the different manipulations. Inspecting these results none of the videos received a MOS above 4. The highest MOS was received by the flamenco reference video without freezing (MOS = 3.92). For the none of the freezing conditions all but videos with CRF = 28 or higher had a score above the acceptable level of 3.5. All the freezing conditions fell below 3.5. The lowest MOS was received by the flamenco video with CRF = 36 and high freezing frequency.

Looking at the individual responses (Figure 5), it can be seen that there are differences in baseline between participants. Some would rate everything high as for example participant 31, while others would rate everything low as for example participant 1. Therefore, to test the significance of these results a multilevel regression was performed, with video condition as the first level and participant as the second. An empty model shows that 45% of the variance is on participant level, confirming the need for multi-level analysis.

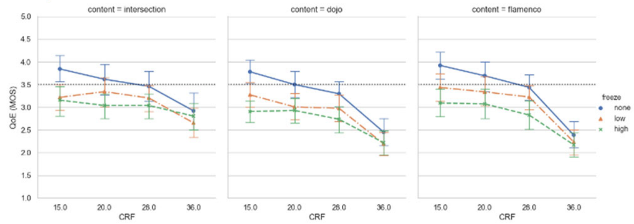

Figure 4: MOS Quality of Experience interaction plot for the different manipulation conditions. Reference line of the acceptable level of 3.5 included

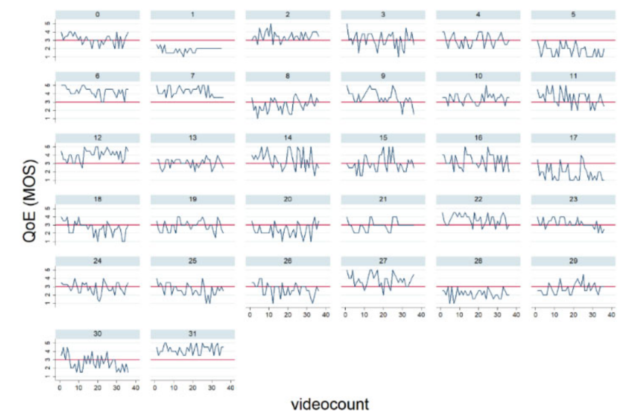

Figure 5: Quality of Experience rating per participant for consecutive videos.

The resulting model on QoE after backward elimination can be seen in Table 2 column "QoE part 1" (within-R2= 0.42; between-R2=0.02; rho=0.57). Included in the model are CRF, Freeze, Content, and the interactions between CRF and both freeze and content. The results show a negative effect of CRF in the no freezing condition where higher CRF values result in lower QoE. This effect increases as the CRF value increases showing a larger than linear effect. As can be seen in Figure 6, the effect of CRF in the low and high freezing frequency conditions was found to be smaller compared to the no freezing condition. The effect of freezing is larger for lower CRF values compared to higher ones. Furthermore, it was found that the effect of CRF 36 on the QoE was more negative in the Dojo and Flamenco video compared to the Intersection video. Additionally, a small but significant effect of the video order was found.

Secondly, the effects of the manipulations on the PCL were evaluated. Answers were expressed in a MOS. answers range from 1 to 4 (mean = 1.97; SD = 0.91). Figure 6 shows the interaction plot of the effects in the different conditions. PCL scores were rather low in all conditions, never surpassing MOS=2.5. The resulting model of the multi-level regression on PLC after backward elimination can be seen in Table 3 (within-R2= 0.06; between-R2=0.02; rho=0.58). Included in the model are CRF, Freeze, Content and the video order. Results show that the PCL was larger in the CRF=36 condition compared to and the other CRF values. Furthermore, increasing the freezing frequency significantly increases the PCL. Comparing the different contents, the PCL was larger in the Dojo video compared to the Flamenco and Intersection video. Video order also had a small but significant negative effect

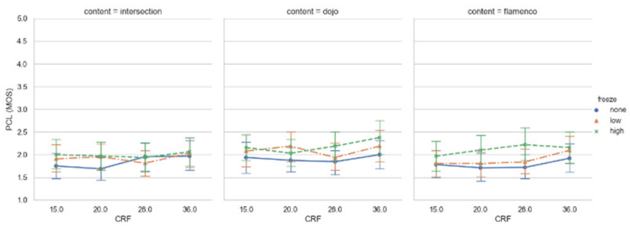

Figure 6: MOS Perceptual and Cognitive load interaction plot for the different manipulation conditions

Table 2: Regression coefficients of the multi-level regressions on the subjective metric

b	QoE	PCL	Cybersick
CRF (15)			
20	-0.156	-0.039	0.058
28	-0.282**	-0.025	0.036
36	-0.891***	0.125*	0.117***
Freeze	-0.398***	0.124***	0.046***
Content			
Dojo	-0.086	0.147***	0.039
Flamenc	0.062	0.014	-0.032
Video Order	-0.004*	0.004*	0.003**
CRF*Freeze (15)			
20	0.109		
28	0.134*		
36	0.302***		
Content*CRF (Intersection | 15)			
Dojo | 20	-0.101		
Dojo | 28	-0.147		
Dojo | 36	-0.419***		
Flamenc	-0.002		
Flamenc	-0.123		
Flamenc	-0.581***		
_cons	3.860***	1.703	1.067***
*Note: * = level of significance*			

Third, the effect of the manipulations on cybersickness was tested. Answers were again expressed through a MOS, ranging between 1 and 4 (mean=1.25; SD=0.52). The MOS never exceeded 1.5 indicating low levels of cybersickness. Figure 7 displays the interaction plots for the different conditions and shows no clear trends. Scores are low as the mean is never above 1.5. Looking at Figure 8, it can be seen that it is quite dependent on the participant whether cybersickness occurs at all. The resulting model of the multi-level regression on PLC after backward elimination can be seen in Table 3 (within-R2= 0.04; between-R2=0.00; rho=0.50). Results show that for CRF only 36 has a significant positive effect, increasing the cybersickness slightly. Although rather small, freezing was also found to have a significant positive effect on the cybersickness. Finally, a very small effect of video order was found, increasing the cybersickness as more videos have been watched. The model has a rather low R^2, indicating that the manipulations do not explain the variance in the data well.

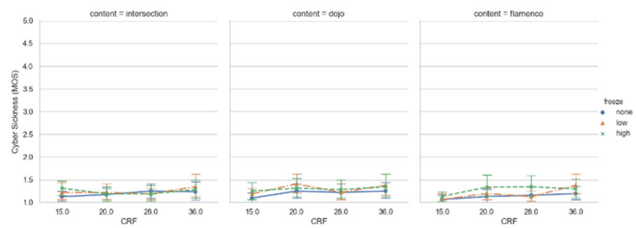

Figure 7: MOS Cybersickness interaction plot for the different manipulation conditions

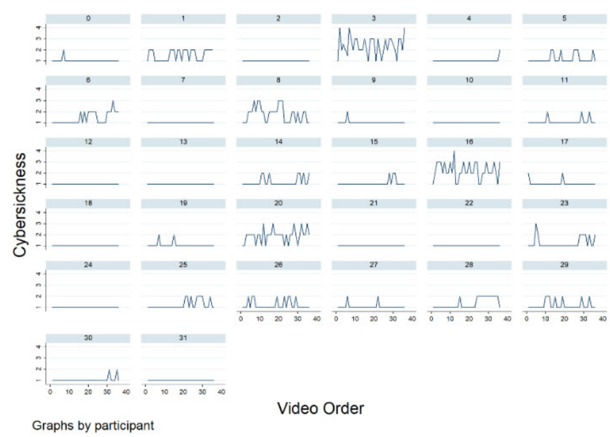

Figure 8: Cybersickness MOS per participant.

Conclusions

360-video is a video format that is currently growing in popularity and even though QoE research on 2D video streaming is well established, little is done yet on 360-video. 360-video brings new challenges and experiences along. Applying QoE theory and methods from 2D video streaming is not trivial and more research is one 360-video specifically is needed. QoE assessment metrics are still in development and subjective methods have yet to be standardized. The goal of the current study was to provide subjective data on the effects of quality degradations, freezing, and content on the QoE.

The overall quality and experience were rated rather low, and answers varied a lot among participants. Degrading the quality does not have much impact on the QoE up until the threshold (here CRF 28), after which the QoE dropped. Freezing events drop the QoE below acceptable level in any condition. In general, higher qualities for 360-video would be required to meet with users' expectations. However, if network resources are limited, more compression would be desired to avoid freezing. Furthermore, the results of this study show that PCL and cybersickness do not cause any serious issues for the QoE and are not much affected by the manipulations.

References

[1]. ITU-T. (2017). *Vocabulary for performance, quality of service and quality of experience* (ITU-T Rec. P.10/G.100). International Telecommunication Union (ITU), ITU Telecommunication Standardization Sector: Place des Nations, CH-1211 Geneva 20.

[2]. Le Callet, P., S. Möller, and A. Perkis, eds. (2012). *Qualinet White Paper on Definitions of Quality of Experience (2012). European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003)*. 2012: Lausanne, Switzerland.

[3]. Azevedo, R.G.d.A., N. Birkbeck, F.D. Simone, I. Janatra, B. Adsumilli, and P. Frossard, (2019). *Visual Distortions in 360-degree Videos.* IEEE Transactions on Circuits and Systems for Video Technology, 2019: p. 1-14, DOI: 10.1109/TCSVT.2019.2927344.

[4]. van Kasteren, A. *The Contribution of Eye Tracking to Quality of Experience Assessment of 360-degree video*. (Doc nr: 0845668 and acr062449), Human Technology Interaction and Visual Media Quality, Eindhoven University of Technology and RISE Research Institutes of Sweden AB, Eindhoven, The Netherland and Kista, Sweden, M. Sc. thesis. 2019

[5]. Reiter, U., K. Brunnström, K. De Moor, L. Mohamed-Chaker, M. Pereira, A. Pinheiro, J. You, and A. Zgank (2014). *Factors Influencing Quality of Experience*, in *Quality of Experience: Advanced Concepts, Applications and Methods*, S. Möller and A. Raake, Editors. Springer. p. 45-60, DOI: https://doi.org/10.1007/978-3-319-02681-7_4.

[6]. Kuipers, F., R. Kooij, D. De Vleeschauwer, and K. Brunnström (2010). *Techniques for Measuring Quality of Experience*, in *Wired/Wireless Internet Communications, Lecture Notes on Computer Science, Volume 6074*. Springer-Verlag. p. 216-227.

[7]. Möller, S. and A. Raake, (2014). *Quality of Experience - Advanced Concepts, Applications and Methods*. T-Labs Series in Telecommunication Services. Switzerland: Springer International Publishing.

[8]. Engelke, U., R. Pepion, P.L. Callet, and H.-J. Zepernick, (2010). *Linking distortion perception and visual saliency in H.264/AVC coded video containing packet loss*. Visual Communications and Image Processing 2010. Vol. 7744. SPIE.

[9]. Tran, H.T.T., N.P. Ngoc, C.T. Pham, Y.J. Jung, and T.C. Thang. (2017). *A subjective study on QoE of 360 video for VR communication*. in *2017 IEEE 19th International Workshop on Multimedia Signal Processing (MMSP)*. 2017.

[10]. Hettinger, L.J. and G.E. Riccio, (1992). *Visually Induced Motion Sickness in Virtual Environments.* Presence: Teleoperators and Virtual Environments, 1992. **1**(3): p. 306-310, DOI: 10.1162/pres.1992.1.3.306.

[11]. Porcino, T.M., E. Clua, D. Trevisan, C.N. Vasconcelos, and L. Valente. (2017). *Minimizing cyber sickness in head mounted display systems: Design guidelines and applications*. in *2017 IEEE 5th International Conference on Serious Games and Applications for Health (SeGAH)*. 2017.

[12]. Schatz, R., A. Sackl, C. Timmerer, and B. Gardlo. (2017). *Towards subjective quality of experience assessment for omnidirectional video streaming*. in *2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX)*. 2017.

[13]. Brunnström, K. and M. Barkowsky, (2018). *Statistical quality of experience analysis: on planning the sample size and statistical significance testing.* Journal of Electronic Imaging, 2018. **27**(5): p. 11, DOI: 10.1117/1.JEI.27.5.053013.

[14]. ITU-T. (1999). *Subjective video quality assessment methods for multimedia applications* (ITU-T Rec. P.910). International Telecommunication Union, Telecommunication standardization sector.

Author Biography

Anouk van Kasteren *received a BSc in Industrial Design with a minor in cognition and social psychology from the University of Technology Eindhoven (TU/e) (2017) and a MSc in Human Technology Interaction from the TU/e (2019) with an exchange in Data Science at the Kungliga Tekniska Högskolan (KTH). She completed her master thesis project at RISE, Research Institutes of Sweden in Stockholm*

Kjell Brunnström*, Ph.D., is a Senior Scientist at Acreo Swedish ICT AB and Adjunct Professor at Mid Sweden University. He is an expert in image processing, computer vision, image and video quality assessment having worked in the area for more than 25 years. Currently, he is leading standardization activities for video quality measurements as Co-chair of the Video Quality Experts Group (VQEG). His current research interests are in Quality of Experience for visual media in particular video quality assessment both for 2D and 3D, as well as display quality related to the TCO requirements.*

John Hedlund*, is a Usability Researcher at RISE. He works with usability and UX evaluations in all stages of product development helping clients to deliver the best possible user experience. John holds an MSc in Industrial Design Engineering, with focus on product design*

Chris Snijders*, is full professor of technology and innovation at the Human-Technology Interaction group of Eindhoven University of Technology. His research interests include human and model-based decision making, virtual reality and persuasion, trust in technology, and in general the way in which humans interact with technological artefacts.*

.