Designing Real-time, Continuous QoE Score Acquisition Techniques for HMD-based 360° VR Video Watching

Tong Xue
Beijing Institute of Technology
Centrum Wiskunde & Informatica (CWI)
Beijing, China
xuetong@bit.edu.cn

Abdallah El Ali
Centrum Wiskunde & Informatica (CWI)
Amsterdam, The Netherlands
abdallah.el.ali@cwi.nl

Irene Viola
Centrum Wiskunde & Informatica (CWI)
Amsterdam, The Netherlands
irene.viola@cwi.nl

Gangyi Ding
Beijing Institute of Technology
Beijing, China
dgy@bit.edu.cn

Pablo Cesar
Centrum Wiskunde & Informatica (CWI)
Delft University of Technology (TU Delft)
Amsterdam, The Netherlands
P.S.Cesar@cwi.nl

***Abstract*—Watching HMD-based 360° video has become increasing popular as a medium for immersive viewing of photo-realistic content. To evaluate subjective video quality, researchers typically prompt users to provide an overall Quality of Experience (QoE) score after viewing a stimulus. However, since users can adjust their viewport throughout a 360° video, a higher level of spatiotemporal granularity is needed for adaptive 360° video streaming. To address this, we design several real-time, continuous QoE annotation input and peripheral visualization techniques, with the goal of minimizing mental workload and distraction during score acquisition. Drawing on two parallel co-design sessions with seven experts, we find that touchpad and joystick are most suitable for continuous input, with DotMorph (circle with tick label that varies in filling) for peripheral state feedback. We contribute design findings for testing QoE score acquisition techniques during HMD-based 360° video watching, which enable more precise optimization of adaptive video streaming quality.**

***Index Terms*—Quality of Experience, 360° video, real-time, continuous, acquisition, peripheral visualization**

I. Introduction

With the development of 3D acquisition and rendering technologies, 360° videos have become increasingly popular as medium to immerse users with photo-realistic Head-Mounted Display (HMD)-based immersive virtual reality experiences [1]. Such experiences are characterized by having more degrees of freedom with respect to traditional media, as users can freely navigate through the 360° content [2]. This comes at the cost of larger expenditures in terms of bandwidth in order to transmit the contents in high quality. Thus, adaptive streaming solutions have been devised to spatially partition the 360° videos into segment (tiles), which are encoded with varying quality parameters depending on the saliency of the content or the user's viewing angle [3], [4]. Such adaptive tiling solutions aim at reducing the bitrate requirements while maintaining high levels of Quality of Experience (QoE).

To assess the QoE of 360° videos, subjective questionnaires are commonly employed to collect visual quality scores of a given stimulus. However, current methodologies and tools for 360° video quality evaluations rely on ITU-defined, post-stimulus scoring methodologies, such as Absolute Category Rating (ACR) or Double Stimulus Impairment Scales (DSIS), which asks the subjects to rate the experience of the video as a whole, averaging across the variations in quality that might occur in the video [5]. However, these are spatiotemporally imprecise, since one can experience different perceived quality across different parts of the 360° video, and is especially critical when tiling with varying quality levels is involved. Moreover, prior research [6] has shown that interactivity could lead to larger confidence intervals and less precise ratings since people do not experience the same content. ITU-R Recommendation BT.500-14 [7] defines a continuous methodology for video quality assessment, namely Single Stimulus Continuous Quality Evaluation (SSCQE) [8]. However, no tool has been designed for continuous evaluation of immersive contents, for which the annotation needs to be performed unobtrusively and in real-time.

To address this gap, we ask: **RQ:** How can we design annotation input and visualization techniques that are suitable for collecting continuous QoE self-reports while users are wearing an HMD and experiencing 360° VR content? We scope our work to the commonly used HTC Vive Pro Eye HMD (which advantageously integrates an eye tracker), and develop five annotation input techniques using the Vive controller, including the position of the controller, the touchpad and controller trigger button. We draw on stimuli with different encoding qualities from VQEG, a validated public QoE database of immersive VR videos [5]. For annotation state feedback, two initial peripheral visualizations (ArcMorph and DotMorph) based on arc and dot peripheral element shapes

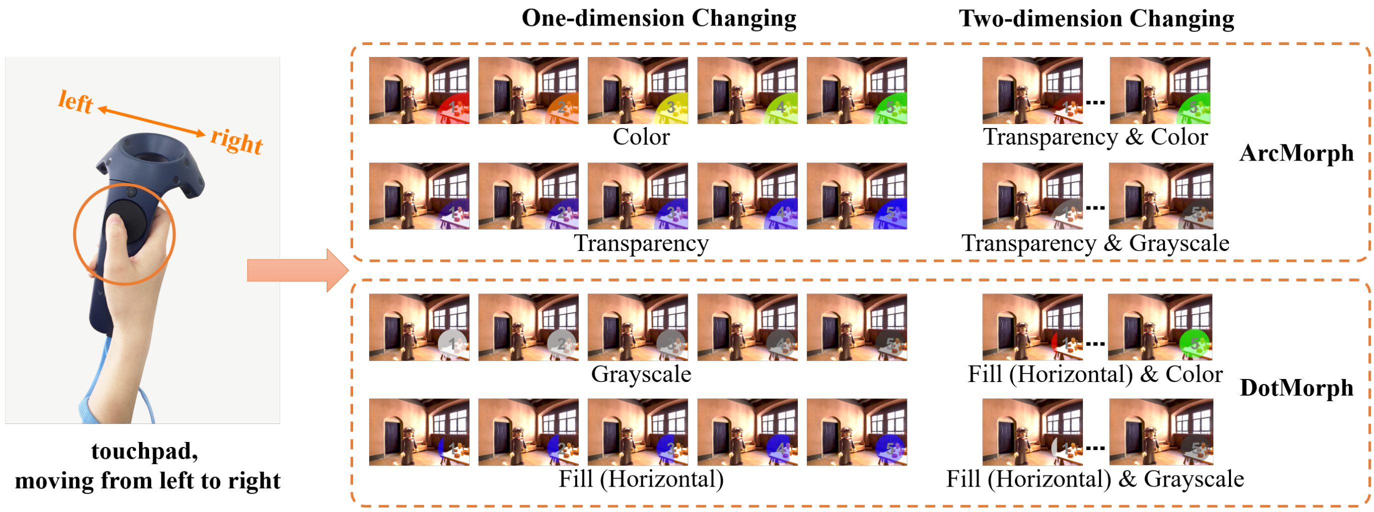

Fig. 1: ArcMorph and DotMorph, two annotation feedback visualizations. The element attributes include color, transparency and fill, which vary with the annotation state.

with different attributes (transparency, color, fill / volume) are proposed. The UI elements are fixed in the right-bottom corner of the viewport and the element attribute as well as tick label on the element shows the annotation feedback. We evaluate and iterate on our designs in two parallel co-design sessions with experts and designers, where we contribute two suitable annotation input techniques (touchpad up/down, joystick up/down), and the DotMorph peripheral state feedback, for annotating QoE continuously while watching 360° VR content. Moreover, these techniques and an introductory video are publicly available at https://github.com/cwi-dis/RC-QoE-A-Techniques. Our work enables future explorations of real-time and continuous QoE score acquisition for HMD-based 360° videos, which can be used not only for improved processing, coding, delivering and rendering techniques, but also as a means to dynamically adapt quality based on implicit user behavior during 360° viewing experiences.

II. Related Work

In this section, we review research on QoE studies, and prior work on real-time and continuous annotation techniques.

A. *QoE assessment for 360° video*

Subjective assessment of 360° videos has received a lot of attention in recent years [9], [10], and tools have been proposed to perform QoE assessment using the HMD [11], [12]. Gutierrez et al. [5] perform a large scale experiment involving multiple labs and test conditions in order to define a subjective methodology for the ITU-T Recommendation P.919 [13]. However, the methodologies they evaluated consisted of post-stimulus evaluation using ACR and DSIS. To capture QoE ratings in real-time for traditional video contents, Alpert et al. [8] presented a single stimulus continuous quality evaluation (SSCQE) method, which allowed viewers to dynamically rate the quality of traditional video sequences using a slider mechanism. Pinson et al. [14] later conducted subjective tests to compare different continuous quality evaluation methods in video streaming applications and validate the usability of these continuous assessments. Given the spatial and temporal variations of 360° videos, by which users could change their viewing direction interactively at any point in time, it remains a challenge to develop real-time continuous QoE annotation techniques.

B. *Real-time Continuous Annotation Techniques*

Performing multiple tasks simultaneously can result in divided attention [15]. There has been some research considering the usage of auxiliary devices to lower mental workload while annotating. Girard et al. [16] developed a one-dimensional emotion slider CARMA for users to report basic emotion valence (positive or negative) by sliding it up and down. Lopes et al. [17] presented the RankTrace tool, which was implemented by a physical radial controller to specify a single, continuous dimension such as emotional intensity. Buchinger et al. [18] presented and analyzed the usage data gloves for time-continuous subjective multimedia assessment in mobile contexts. Recently, Xue et al. [19] proposed RCEA-360VR, which is suitable for continuous emotion annotation via valence and arousal while watching 360° videos. Given that in our case video quality would be rated as a numerical scale (1-5), we need to design novel methods for real-time continuous QoE annotation in immersive HMD-based 360° video contexts.

III. Designing Real-time Continuous QoE Score Acquisition Prototypes for 360° VR

Below we discuss the design principles and development of our continuous 360° VR video QoE annotation prototypes that draw on Mean Opinion Scores (MOSs). We consider three design principles, which served as heuristics to narrow down the design space, and are based on considerations for designing VR HMD-based interactions [20]:

D1: Design for HMD-based VR Interactions. Users will experience 360° videos through the HMD and report their QoE

(a) Hand with controller, moving left/right.

(b) Hand with controller, moving down/up.

(c) Squeezing trigger stick.

(d) Thumb on touchpad, moving left/right.

(e) Thumb on touchpad, moving down/up.

Fig. 2: QoE annotation input techniques for HMD-based 360° VR videos.

scores in real-time without attending to the annotation device. We first need to ensure the viewing experience is comfortable and minimizes motion sickness and latency [21]. We also need to ensure the annotation device is ergonomic, precise, and easily operable within 360° VR [22].

D2: Design for Multitasking. Given that users will annotate video quality continuously while watching 360° videos, this can lead to divided attention [15]. Our design draws on peripheral visualization feedback [23] to reduce the burden of annotating and convey state feedback without interfering too much with the viewing experience. We introduce one translucent GUI element fixed in the right-bottom corner of the viewport to visualize the user's annotation state, which can lower interruptions and help users keep awareness of the primary task [22], [24].

D3: Design for Numerical Scales. Typically QoE scores are annotated on a one-dimensional numerical scale, for example, from 1 to 5 [5]. The visualization of numerical scale generally adopts the form of number axis, with interval properties and anchor points [25]. We design different changing states of the element attribute to visualize the annotation tick numbers, which occupies less viewport space. To make the user's current annotation feedback clear, we also consider adding the current tick label on the element.

A. Real-time, Continuous QoE Score Acquisition Prototypes

Our 360° VR continuous QoE annotation prototype consists of two major components: (a) the HTC Vive Pro Eye HMD[1] with a resolution of 2880 x 1600 pixels, a 110° field of view and a refresh rate of 90Hz [9] (b) the Vive Controller input device for QoE annotation, which is bundled with the HMD. Previous research [26] indicated that the Vive controller allows a faster and more accurate interaction than other input devices like Myo armband in immersive VR scenarios. A custom scene was constructed in the Unity Engine[2] to display 360° videos at 25 fps and show the annotation feedback based on users' continuous ratings. In parallel, the audio signal was sent to the HMD.

1) 360° Video Stimuli: Drawing on D1 and D2, we consider short 360° video clips with uniform (using homogeneous QPs) and non-uniform (using different configurations of tiles) coding degradations from the VQEG database [5] as stimuli, which contains MOS scores from ten laboratories and more than 300 participants. In this early work, we select the $< VSenseLuther >$ video clip with animation content and a main character [27]. Two uniform encodings ($QP = 22$, $QP = 42$) and two non-uniform encodings ($QP = 22\_6x3\_gradual$, $QP = 22\_6x3\_abrupt$) for this sequence are captured from the VQEG dataset.

2) Annotating Video Quality Continuously: Drawing on D1, we designed five initial prototypes based on the Vive controller, as shown in Figure 2. These methods consider the hand with controller movement, and the interaction between the user's hand and controller components such as trigger button and touchpad, which have been shown to be intuitive and easy for users to handle [26]. The first two methods map the relative position of the controller to the annotation values. Here, a user moves their hand with controller from left to right (Figure 2a) or from down to up (Figure 2b) indicating that the annotation data changes from one to five. The third method is reporting QoE values by squeezing the trigger stick of the controller (Figure 2c), and the QoE level increases with squeezing further. The fourth and fifth methods map the user's thumb on the touchpad from left to right (Figure 2d) or down to up (Figure 2e) to QoE data from one to five.

3) Visual Annotation Feedback: Drawing on D2 and D3, we presented two annotation visualizations, ArcMorph and DotMorph, using translucent arc and dot fixed to the right-bottom corner of the HMD viewport based on [22], and the attributes with the tick number on the element visualizes the quality levels users annotate. In ArcMorph and DotMorph, we first considered the color attribute, a common red to green scale (i.e, traffic light model), including grayscale for people with color impairments [28], transparency, a fully transparent to translucent scale, which has been validated in previous studies [24]. According to the progress bar research for univariate data representation [29], we also designed a Fill method with three directions: horizontal, vertical, and radial 360. To enhance the visualization feedback, we included two-dimension changing methods by combining Transparency and Color, Transparency and Grayscale, Fill and Color, and Fill and Grayscale, respectively. With the user's thumb on the touchpad from left to right, we show the changing states of the UI element in Figure 1.

[1]https://enterprise.vive.com/us/product/vive-pro-eye/
[2]https://unity3d.com/

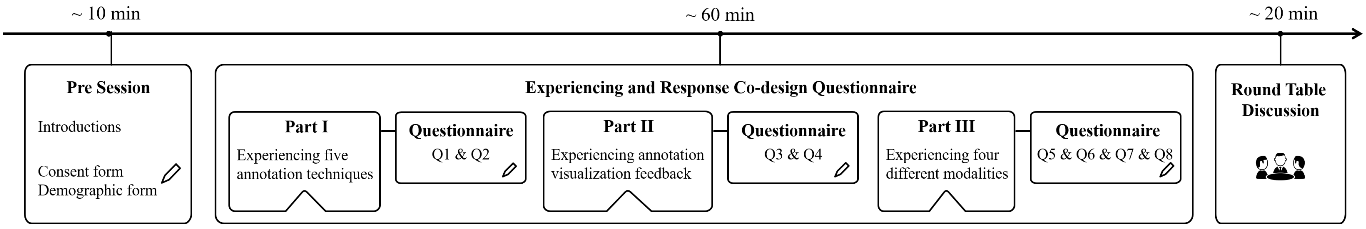

Fig. 3: Co-design session procedure.

IV. Co-design Session

A. Session and Procedure

We follow a user-centric approach with multiple iterative design rounds based on expert co-design sessions [30] to evaluate and iterate on our initial prototypes. Given current COVID-19 restrictions, and our need to evaluate with experts from diverse research backgrounds, including QoE, VR, design, and engineering, we conducted two co-design sessions at X and Y institutes in parallel across two countries. The X session involved three participants (1f, 2m) aged between 30-35 (M=32, SD=2.6), and the Y session involved four participants (4m) aged between 25-34 (M=29, SD=4.9). All participants are right-handed, have no visual impairments, and their research backgrounds are listed in Table I.

TABLE I: Research background of participants.

	ID	Research Background
X Session	**P1**	QoE;VR
	P2	VR
	P3	QoE
Y Session	**P4**	Software Engineering
	P5	Design
	P6	Software Engineering, VR
	P7	Design

Our co-design session procedure is shown in Figure 3, where each session lasted approximately 90 minutes. We first explained the research objectives and shared a web survey link to participants, which contained a consent form, demographic questions, and the design preference questions. We asked participants to carefully read and digitally sign the consent form, and fill in the demographic form. Participants freely volunteered for this study, and therefore did not receive monetary compensation. Participants experienced each prototype individually in turn, and thereafter answered the corresponding questions in the questionnaire. All questions are listed in Table II, and for each question, we followed up with their reasons for responding as such (by inputting responses in an open text field). For each prototype, participants were instructed to focus on evaluating the quality of the video, and not solely focus on the visualization feedback. The session consisted of three parts: (1) Five initial annotation techniques. In this part, we additionally considered how to show the annotation feedback once the user's finger leaves the touchpad, which is presented as $Q2$, and discussed in Section IV-B1. (2) Two element shapes with nine one-dimension changing and four two-dimension changing visualizations. (3) Four modalities: continuous or discrete changing of element, continuous or discrete changing or hiding of element text, four types of video quality, annotating both video and audio quality. Lastly, we invited participants for a round table discussion, where audio was recorded and later transcribed. The questionnaire data and discussion were analyzed qualitatively using an open coding approach [31].

TABLE II: Co-design session questions.

ID	Question
Q1	*Which annotation input method using the Vive controller did you find most suitable for rating visual quality continuously?*
Q2	*If the user's hand leaves the touchpad, should the element state reset back to neutral (level 3) or leave on the last state?*
Q3	*Which element shape did you find most suitable for providing rating feedback?*
Q4	*Which element changing states did you find most suitable for providing rating feedback?*
Q5	*Which type of element changing state is most suitable for providing continuous rating feedback?*
Q6	*Which type of tick text in the element is most suitable for providing continuous rating feedback?*
Q7	*Which type of stimuli are most suitable to be assessed for quality using the continuous annotation techniques?*
Q8	*Which type of element state is most suitable for rating both video and audio quality while watching 360 videos through HMD?*

B. Key Findings and Design Considerations

Below we list the key findings from the co-design sessions:

1) Annotation Input Techniques: Arm movement was criticized by all the experts. There was no clear range and it was difficult to maintain, especially while rotating their head to view the stimuli. While this method may be suitable for stationary settings, however not so for 360° videos. We additionally discarded the trigger stick method, though P3 liked as it was less related to head movement. However P1 argued *"...too sensitive...required more tension in my hand to go from 1 to 5 and back, with a lot of cognitive load for precision"*. P7 also stated *"Vive controller's trigger stick is often used as a throttle in racing games, it's difficult for most users to get used to the squeezing strength and force feedback"*. Touchpad for input was intuitive, precise and natural. Regarding the direction of thumb movement, P7 preferred touchpad left/right *"when I hold the Vive controller, thumb moves horizontally more comfortably than vertical"*. However, P1 and P2 mentioned left/right created a mismatch

Fig. 4: Resulting prototypes. Two annotation input techniques (left): Joy-Con joystick controller (up/down), touchpad (up/down). DotMorph (right): a circle with tick label fixed to the bottom-right corner of the viewport, which varies in Fill (up/down) with annotation state.

and became messy when turning their head left to right against the annotation direction. Thus up/down seems more suitable, as users typically spend more time in 360° video turning their head left/right than up/down [32].

For cases of users' finger leaving the touchpad, P1 and P4 expressed it was easier to associate lifting with no score or neutral, and keeping finger on touchpad when reporting a specific score. While P6 mentioned it could be related to the magnitude of viewport rotation, *"if the user rotates slightly, leave on the last state, but if the user would view very different things, resetting back to neutral is safer"*. We did not consider because different settings would bring confusion to users. On the other hand, two designers argued for leaving on the last state, *"...it's easier to consider current ratings compared with last state (increase or decrease) then giving a new state, especially during real-time evaluation"*. P3 added *"No news is the same news...I don't want the system to decide for me"*.

At the end of the discussion, experts recommended that given the nature of continuous annotation, the annotation value should be no score or zero if user leaves the touchpad, which additionally aids data cleaning later. Regarding input device, the final recommendation was to use physical joystick like the Joy-Con wireless controller, which has force feedback, and overall more controllable than a trigger stick. The joystick head would re-align to center position under no force and the annotation value would be thus neutral.

2) Peripheral Visualization Feedback: Visualization Attributes. Participants generally liked DotMorph. However P3 and P6 preferred ArcMorph, as they found it occupied a smaller part of the viewport. However, others believed DotMorph was clearer overall and less distracting, and better generalizes across participant evaluators in the future. For element intensity changing, transparency was not suitable as it was not clear and easily blended with the video content, especially when transparency was high. P1 said *"Color helps in understanding what rating is given while having it in the periphery"*. However, it was discarded later as flashing colors overall interfered with video watching. P2 liked the Grayscale as it did not interfere with the content, which was designed for people with color impairments (Section III-A3). All participants expressed Fill was the best visualization method due to being intuitive and unambiguous, where most mentioned using Fill only was sufficient for the video quality rating task. Moreover, they highlighted the direction of Fill should be consistent with the annotation movement.

Continuous vs. Discrete Feedback. A key design aspect is finding whether the changing of element and text tick marks should be continuous or discrete. P5 liked discrete changing as it matched the text changing. However, given that users' subjective quality self-reports change continuously, participants suggested the element changing continuously was more comfortable. Also, all participants stressed that without tick text, it would be confusing while annotating in real-time. However, including the decimal scale incurred higher mental workload, which can increase reaction times, *"...it's much harder to differentiate between a 2.3 and 2.4"*. P1 and P3, who are QoE experts, suggested the possibility of modifying the scale to maintain the continuous numerical changes, *"...maybe conduct some tests with 1-5, 1-7 and 1-9 and then make decisions"*.

3) Annotating Video and Audio Quality Simultaneously: Lastly, all participants argued that annotating both video and audio simultaneously was too cognitively demanding, where the annotations influenced one another. Therefore, this should be skipped for future tests, with testing of a single modality quality at a time.

C. Resulting Prototypes

Overall, we find that joystick and touchpad annotation input techniques, and using DotMorph visualization feedback, to be the most suitable for later controlled QoE tests. These are shown in Figure 4.

While a physical joystick (up/down) could be naturally suited for continuous and real-time QoE score acquisition, we note that it requires a specific level of force to maintain the same quality. Due to this, we also consider the touchpad (up/down) method, which is precise and allows the user to rest their finger during continuous annotation. If the user's finger leaves the touchpad, the annotation value should be recorded as zero. For state feedback, we aim to later test DotMorph, a solid, translucent and blue circle fixed to the bottom-right corner of the viewport. To match the two input techniques, DotMorph should use a continuous Fill in vertical direction to visualize the quality score. Furthermore, a tick label as an integer is embedded in the center of DotMorph to relay the current annotation score.

V. Next Steps and Research Agenda

In this work, we designed techniques for collecting more precise QoE scores while watching HMD-based 360° VR videos. Our future work comprises different facets: First, we plan to investigate the usability of both touchpad (up/down) and joystick (up/down) annotation input techniques, with DotMorph visualization for peripheral feedback. Moreover, we want to assess whether the visualization technique, which naturally covers a small portion of the audiovisual stimulus to be assessed, has any influence on the collected QoE scores. Thus, we aim to test video stimuli with different encoding qualities from the VQEG dataset [5]. For usability measures [19], we aim to consider: subjective workload, physiological measures, presence questionnaires, and VR simulator sickness. Second, we plan on collecting head and eye movement data through the Vive Pro Eye HMD, and analyze the relationship between behavioral data and continuous QoE scores, which has downstream implications for adaptive tiling mechanisms for optimizing delivery of 360° video.

Our work enables future explorations of real-time and continuous QoE score acquisition for HMD-based 360° videos and other media (e.g., light field, point clouds), which can be used not only for improved processing, coding, delivering and rendering techniques, but also as a means to dynamically adapt streaming quality (cf., [33]) based on implicit user behavior during such viewing experiences.

References

[1] C.-L. Fan, W.-C. Lo, Y.-T. Pai, and C.-H. Hsu, "A survey on 360 video streaming: Acquisition, transmission, and display," *ACM Computing Surveys (CSUR)*, vol. 52, no. 4, pp. 1–36, 2019.

[2] A. MacQuarrie and A. Steed, "Cinematic virtual reality: Evaluating the effect of display type on the viewing experience for panoramic video," in *2017 IEEE Virtual Reality (VR)*. IEEE, 2017, pp. 45–54.

[3] M. Hosseini and V. Swaminathan, "Adaptive 360 vr video streaming: Divide and conquer," in *ISM '16*. IEEE, 2016, pp. 107–110.

[4] X. Corbillon, G. Simon, A. Devlic, and J. Chakareski, "Viewport-adaptive navigable 360-degree video delivery," in *2017 IEEE international conference on communications (ICC)*. IEEE, 2017, pp. 1–7.

[5] J. Gutierrez, P. Perez, M. Orduna, A. Singla, C. Cortes, P. Mazumdar, I. Viola, K. Brunnstrom, F. Battisti, N. Cieplinska, D. Juszka, L. Janowski, M. I. Leszczuk, A. Adeyemi-Ejeye, Y. Hu, Z. Chen, G. Van Wallendael, P. Lambert, C. Diaz, J. Hedlund, O. Hamsis, S. Fremerey, F. Hofmeyer, A. Raake, P. Cesar, M. Carli, and N. Garcia, "Subjective evaluation of visual quality and simulator sickness of short 360 videos: Itu-t rec. p.919," *IEEE Transactions on Multimedia*, 2021.

[6] I. Viola, M. Řeřábek, and T. Ebrahimi, "Impact of interactivity on the assessment of quality of experience for light field content," in *QoMEX '17*. IEEE, 2017, pp. 1–6.

[7] ITU-R BT.500-14, "Methodology for the subjective assessment of the quality of television pictures," International Telecommunications Union, Oct 2019.

[8] T. Alpert and J.-P. Evain, "Sscqe (single stimulus continuous quality evaluation): A new subjective assessment method introduced in itu-r recommendation 500-7: Presentation and results," in *International Workshop on HDTV '96*, 1996, pp. 417–427.

[9] A. Singla, S. Fremerey, W. Robitza, and A. Raake, "Measuring and comparing qoe and simulator sickness of omnidirectional videos in different head mounted displays," in *QoMEX '17*. IEEE, 2017, pp. 1–6.

[10] G. Regal, R. Schatz, J. Schrammel, and S. Suette, "Vrate: A unity3d asset for integrating subjective assessment questionnaires in virtual environments," in *QoMEX '18*. IEEE, 2018, pp. 1–3.

[11] E. Upenik, M. Řeřábek, and T. Ebrahimi, "Testbed for subjective evaluation of omnidirectional visual content," in *2016 Picture Coding Symposium (PCS)*. IEEE, 2016, pp. 1–5.

[12] P. Pérez and J. Escobar, "Miro360: A tool for subjective assessment of 360 degree video for itu-t p. 360-vr," in *QoMEX '19*. IEEE, 2019, pp. 1–3.

[13] ITU-T P.919, "Subjective test methodologies for 360º video on head-mounted displays," International Telecommunications Union, Oct 2020.

[14] M. H. Pinson and S. Wolf, "Comparing subjective video quality testing methodologies," in *Visual Communications and Image Processing 2003*, vol. 5150. International Society for Optics and Photonics, 2003, pp. 573–582.

[15] D. Kahneman, *Attention and effort*. Citeseer, 1973, vol. 1063.

[16] J. M. Girard and A. G. Wright, "Darma: Software for dual axis rating and media annotation," *Behavior research methods*, vol. 50, no. 3, pp. 902–909, 2018.

[17] P. Lopes, G. N. Yannakakis, and A. Liapis, "Ranktrace: Relative and unbounded affect annotation," in *Proc. ACII '17*. IEEE, 2017, pp. 158–163.

[18] S. Buchinger, W. Robitza, M. Nezveda, E. Hotop, P. Hummelbrunner, M. C. Sack, and H. Hlavacs, "Evaluating feedback devices for time-continuous mobile multimedia quality assessment," *Signal Processing: Image Communication*, vol. 29, no. 9, pp. 921–934, 2014.

[19] T. Xue, A. El Ali, T. Zhang, G. Ding, and P. Cesar, "Rcea-360vr: Real-time, continuous emotion annotation in 360 vr videos for collecting precise viewport-dependent ground truth labels," in *Proc. CHI '21*, 2021, pp. 1–15.

[20] J. Jerald, *The VR book: Human-centered design for virtual reality*. Morgan & Claypool, 2015.

[21] S. Davis, K. Nesbitt, and E. Nalivaiko, "A systematic review of cyber-sickness," in *Proc. IE '14*. ACM, 2014, pp. 1–9.

[22] T. Xue, S. Ghosh, G. Ding, A. El Ali, and P. Cesar, "Designing real-time, continuous emotion annotation techniques for 360 vr videos," in *Proc. CHI '20 EA*, 2020, pp. 1–9.

[23] S. Bakker, D. Hausen, and T. Selker, *Peripheral Interaction: Challenges and Opportunities for HCI in the Periphery of Attention*. Springer, 2016.

[24] D. Lindlbauer, K. Lilija, R. Walter, and J. Müller, "Influence of display transparency on background awareness and task performance," in *Proc. CHI '16*. ACM, 2016, pp. 1705–1716.

[25] R. Heiberger and N. Robbins, "Design of diverging stacked bar charts for likert scales and other applications," *Journal of Statistical Software*, vol. 57, no. 1, pp. 1–32, 2014.

[26] L. T. De Paolis and V. De Luca, "The impact of the input interface in a virtual environment: the vive controller and the myo armband," *Virtual Reality*, vol. 24, no. 3, pp. 483–502, 2020.

[27] S. Knorr, C. Ozcinar, C. O. Fearghail, and A. Smolic, "Director's cut: a combined dataset for visual attention analysis in cinematic vr content," in *Proc. SIGGRAPH CVMP '18*, 2018, pp. 1–10.

[28] P. L. Rheingans, "Task-based color scale design," in *28th AIPR Workshop: 3D Visualization for Data Exploration and Decision Making*, vol. 3905. International Society for Optics and Photonics, 2000, pp. 35–43.

[29] C. Harrison, B. Amento, S. Kuznetsov, and R. Bell, "Rethinking the progress bar," in *Proc. UIST '07*, 2007, pp. 115–118.

[30] E. B.-N. Sanders and P. J. Stappers, "Co-creation and the new landscapes of design," *CoDesign*, vol. 4, no. 1, pp. 5–18, 2008.

[31] G. R. Gibbs, "Thematic coding and categorizing," *Analyzing qualitative data*, vol. 703, pp. 38–56, 2007.

[32] A. T. Nasrabadi, A. Samiei, A. Mahzari, R. P. McMahan, R. Prakash, M. C. Farias, and M. M. Carvalho, "A taxonomy and dataset for 360 videos," in *Proc. MMSys '19*, 2019, pp. 273–278.

[33] S. Xie, Y. Xu, Q. Shen, Z. Ma, and W. Zhang, "Modeling the perceptual quality of viewport adaptive omnidirectional video streaming," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 9, pp. 3029–3042, 2019.