QoE Assessment Aspects for Virtual Reality and Holographic Telepresence Applications

Georgios Kougioumtzidis[1](✉), Vladimir Poulkov[1], Zaharias Zaharis[2], and Pavlos Lazaridis[3]

[1] Technical University of Sofia, Sofia, Bulgaria
{gkougioumtzidis,vkp}@tu-sofia.bg
[2] Aristotle University of Thessaloniki, Thessaloniki, Greece
zaharis@auth.gr
[3] University of Huddersfield, Huddersfield, UK
p.lazaridis@hud.ac.uk

Abstract. The cutting-edge evolution of mobile communication systems and Internet technologies in nowadays transitional period from the information age to the experience age has brought attention to the evolving virtual reality (VR) and augmented reality (AR) applications and moves towards the development of holographic telepresence systems. Since these applications are devoted in creating immersive and interactive experiences, the quality of experience (QoE) as it is perceived by the end-users will become fundamental constituent in their performance evaluation process. In this paper, the significance of QoE in the development and implementation of the emerging technologies of VR and holographic telepresence systems is analyzed. Moreover, the QoE influencing factors for VR applications and the distinction among this evolving technology and the conventional 2D video content are outlined. Furthermore, a classification of the QoE assessment methods, together with an analysis of the more significant metrics with regard to VR applications is presented.

Keywords: Quality of experience (QoE) · Virtual reality · Augmented reality · Holographic telepresence

1 Introduction

The interest in video streaming services is ever-increasing and the mobile video traffic has seen an exponential growth in recent years, accounting for the bulk of the mobile data traffic. As video resolution competency along with the capabilities of terminal equipment is upsurging at a fast pace and the ultra-high channel bandwidths and novel physical layer practices of the forthcoming beyond 5G (B5G) and sixth generation (6G) of mobile communication systems are underway, the research attention is being drawn towards the virtual reality (VR) and augmented reality (AR) applications.

The interest in such technologies is envisioned to grow exponentially in the years to come, which will lead forward to the next step of the extended reality (XR) implementations that will be focused on holographic telepresence applications. Holograms

D. Perakovic and L. Knapcikova (Eds.): FABULOUS 2022, LNICST 445, pp. 171–180, 2022.
https://doi.org/10.1007/978-3-031-15101-9_12

constitute the evolution of video communications, offering a significantly more complete experience to the users. Holographic telepresence will allow remote users to be rendered as a local presence, enabling a substantial number of innovating interactive usage scenarios, including tele-conferencing, tele-surgery and tele-education.

The fact that these technologies rely on interactivity and immersiveness, renders the users' quality of experience (QoE) a particularly crucial parameter during their development. In the course of implementing VR and holographic applications and designing future usage scenarios, it is very significant to comprehend the experiences and expectations of users as they are expressed and conveyed through QoE. QoE is a multidisciplinary measure affected by a wide range of factors belonging to various fields. It has become an essential element in the assessment of network services and operations in recent years, enabling for a greater understanding of how network technical aspects influence the quality of service as it is perceived by end-users.

In this paper, we discuss and analyze the significance of QoE in the development and implementation of the emerging technologies of VR and holographic telepresence systems. Moreover, we examine the QoE influencing factors for VR applications and underline the distinction among this evolving technology and the conventional 2D video content. Furthermore, we classify the QoE assessment methods with regard to VR applications and provide analysis of the more significant metrics both subjective and objective. Last, we examine the parameters of holographic telepresence systems that effect QoE, as embedding QoE-awareness in their design is still an open research issue.

2 QoE Influencing Factors

QoE influencing factors (IFs) can be characterized as the true state or adaptation of any attribute of a user, system, service, application or context that might impact the user's perceived quality [1]. The IFs contain attributes such as the form and properties of an application or service, the utilization context, the fulfillment of the user's expectations, the user's cultural background, socioeconomic status and psychological portrait, and ultimately, the user's emotional state [2]. QoE IFs can be classified as human-related, system-related and context-related [3] (Table 1).

2.1 Human-Related Factors

Any variant or invariant attribute or trait of a user is referred to as a human IF. Human IFs may have an impact on the development of a particular experience, as well as its overall quality. Because of their subjectivity and connection to internal processes of the human, they are extremely complicated [4]. Human IFs in VR applications include the physiological features of the user such as age and gender, anomalies in the human visual system (HVS) [5], as well as impairments in the human auditory system (HAS) [6], the simulator sickness which is caused by visual stimuli and can induce symptoms like fatigue, perspiration, vertigo or nausea [7], the level of user's immersion and also the user's expectations and expertise with VR applications [3].

Table 1. QoE influencing factors.

Influence factor	Feature	Description
Human factors	Physiological features	User's age & gender
	HVS & HAS	Impairments in human vision and hearing system
	Simulator sickness	Symptoms like fatigue, perspiration, vertigo & nausea
	Immersion	The user's propensity in experiencing immersion
	Expectations & expertise	The level of user's experience with VR systems
System factors	Content-related	Spatial audio, spatial depth & spatiotemporal complexity
	Media/codec-related	Compression, video codecs, audio, storage & transport, bitrate, resolution, frame rate, audio sample rate & coding delay
	Network/transmission-related	Delay, bandwidth & packet loss
	Hardware-related	HMD, headphones, decoder performance, head-tracking, FoV, display resolution & refresh rate
Context factors	Physical context	The environment in which a user interacts with a VR system
	Temporal context	The frequency & duration of utilizing a VR service
	Social context	Interaction with other users & consideration on popularity of a VR content & the way a VR service is accessed
	Task context	The purpose of using a VR service

2.2 System-Related Factors

System IFs are qualities and features that have an impact on the overall performance of a VR application or service with regard to technical parameters. VR system IFs can be further classified into four categories, which are as follows: content-related, media/codec-related, network/transmission-related and hardware-related [3, 4]:

- Since different content properties may need distinct system features, the content and its nature have a significant impact on the overall QoE of the system. The content-related IFs include parameters such as spatial audio, spatial depth and spatiotemporal complexity.

- The media/codec-related IFs refer to media configuration features such as compression, video codecs, audio, storage and transport, bitrate, resolution, frame rate, audio sample rate and coding delay.
- The network/transmission-related IFs are influenced by errors that occur during network transmission and are inextricably linked to network quality of service (QoS). The network/transmission-related IFs include parameters such as delay, bandwidth and packet loss.
- Hardware-related IFs pertain to terminal systems and equipment of the transmission channel and refer to components such as head-mounted display (HMD), headphones, decoder performance, head-tracking, field of view (FoV), display resolution and refresh rate.

2.3 Context-Related Factors

Context IFs encompass any situational feature to define the user's surroundings. Context IFs may arise at various degrees of magnitude, behavior, and occurrence patterns, both individually and in combinations, and they can be classified as physical context factors, temporal context factors, social context factors and task context factors.

3 QoE Assessment

There are two approaches for assessing QoE: the subjective and objective assessment. In subjective models human assessors are subjected to a range of tests or stimuli in order to provide information about quality [8]. Objective models on the other hand are regarded as a way for assessing subjective quality using solely objective quality metrics [9].

3.1 Subjective Methods

Quantitative methodologies from adjacent fields such like psychophysics and psychometrics are used in the subjective methods, by employing ratings that represent the assessors' perception of the level of quality they experienced. Additionally, qualitative methods including focus groups, interviews and profile surveys are also utilized to discover which IFs contribute to QoE and in what extent [8]. Ordinarily, assessors rank a series of perceived quality characteristics on a mean opinion score (MOS) scale, which ranges from 1 to 5 (i.e., bad to excellent) and indicate their degree of satisfaction with a particular service [10]. Owing to direct data obtainment from end-users, the subjective assessment approach offers the most accurate outcomes. These outcomes are used as reference in model training and affirmation. The fundamental drawbacks of the subjective methods on the other hand, stem from the fact that they are expensive, time-consuming, incapable for use in real time and unrepeatable [11].

In the case of VR applications, the subjective QoE assessments should be carried out with use of HMDs to preserve the immersive traits and ensure exact perception of the quality of panoramic videos [12]. Moreover, the spatial and temporal perceptual information of the test sequence are vital factors, since they determine the feasible amount of video compression and, as a result, the degree of impairment that is afflicted when the test sequence is broadcasted over a fixed-rate communication channel [13]. The subjective VR QoE assessment methods should value the assessors' ratings on a series of perceptual scales, including audiovisual quality, simulator sickness symptoms and exploration behavior [14] (Table 2):

- The most frequently used metrics to evaluate the audiovisual quality include the absolute category rating (ACR) and the degradation category rating (DCR). The ACR metric is a single stimulus approach that can be utilized in case testing time is critical, because it generates a large number of ratings in a short amount of time. The DCR metric, also known as the double stimulus impairment scale (DSIS), is a double stimulus approach that generates smaller number of ratings in the same amount of time in comparison with ACR, but is statically more reliable [15].
- Simulator sickness is an unpleasant condition induced by a sensory imbalance among the visual and vestibular systems. For measuring simulator sickness symptoms, the simulator sickness questionnaire (SSQ) is indicated [7]. The virtual reality sickness questionnaire (VRSQ) is a condensed version of the SSQ that only examines the subsequent symptoms: overall discomfort, tiredness, eyestrain, poor concentration, headache, head heaviness, blurred vision, dizziness with eye closed and vertigo [16]. When simulator sickness requires to be examined often and due to the objective of the testing quick self-reporting from the assessors is essential, it is advised to utilize the vertigo score rating (VSR) [17].
- The participant's exploration behavior is measured by tracking the position of the head rotation throughout the active viewing session. Head rotation position is required to be monitored at regular intervals and timed to the beginning of each test sequence, such that exploration behavior can be linked to the content of the test sequence. Likewise, eye movements can be captured utilizing eye trackers embedded in the HMDs [18].

Table 2. Subjective assessment metrics.

Perceptual feature	Metric	Description
Audiovisual quality	ACR	Single stimulus ratings of test sequence on a five-level class scale
	DCR/DSIS	Double stimulus ratings of test & reference sequence on a five-level class scale
Simulator sickness	SSQ	Questionnaire method for simulator sickness evaluation
	VRSQ	Compacted version of SSQ
	VSR	Rapid self-reporting method
Exploration behavior	Head rotation	Head position tracking
	Eye movement	Eye trackers embedded in HMDs

3.2 Objective Methods

Due to the constraints of subjective methods, there has been a considerable push to implement objective models that predict the subjective perceived quality based solely on physical traits. The objective methods are projected to generate a prediction of the QoE value that approximates the ratings of the subjective assessment methods. The objective approach has the advantage of being simple to apply and modify, as the assessment process has only to account on the observable QoS parameters and corresponding correlation mathematical models. The objective assessment's drawback is its inaccuracy, as the computed QoE is only an approximation instead of an exact value of the quality as perceived by the end-users [19].

Objective quality assessment methods can be classified in five categories [20]: 1) media-layer models, which take actual media audiovisual signals as input and incorporate codec compression and channel features; 2) packet-layer models, which calculate QoE only based on packet header data; 3) bitstream models, which handle encoded bitstream data as well as packet header data as input; 4) hybrid models, which are a fusion of the aforementioned models that use as much information and data as feasible to evaluate QoE; and 5) planning models, which use network or terminal quality planning features to calculate their input.

Table 3. Objective assessment metrics.

Source information	Metric	Description
Full reference	SSIM	Still image sequence quality metric based on luminance, contrast & structure
	VQM	Video quality metric based on structural & temporal features
	VIF	Still image quality metric based on the ratio of the distorted image to reference image
	PSNR	Image & video quality metric that assesses the variance between source & distorted signals
	VMAF	Video quality metric that evaluates the influence of compression & rescaling
Reduced reference	STRRED	Calculation of the impairment among a distorted & a reference video sequence
	SpEED-QA	Mean-subtracted pixel values of frames and frame variances in video sequences
No reference	NR-P	Video quality metric based on the decoded representation to compute the quality of a received stream
	NR-B	Video quality metric based on parameters extracted from the encoded bitstream

The objective metrics employ audio, image and video attributes to assess quality and they are classed as full reference, reduced reference or no-reference, depending on the amount of source information available [21] (Table 3):

- The reference and result video sequences are both supplied in the full reference (FR) metrics, allowing for thorough comparison of the videos [22]. Paradigms of such metrics include the structural similarities (SSIM), video quality model (VQM), visual information fidelity (VIF), peak signal to noise ratio (PSNR) and video multimethod assessment fusion (VMAF).
- The reference and result video sequences are generated by employing the same selection of attributes in reduced reference (RR) metrics. Just a selection of partial features from the source input sequence are necessary for the quality assessment [23]. Examples of RR metrics include the spatio-temporal reduced-reference entropic differencing (STRRED) and spatial efficient entropic differencing for quality assessment (SpEED-QA).
- No reference (NR) metrics have access only in the outcome video sequence and the quality evaluation is achieved by employing information comprised in a related image or video media stream. The computing demands of NR techniques are the lowest when compared to the other methods and they also have the fastest temporal response, but they are unable to provide an accurate evaluation across a wide range of video states [24]. NR metrics include the pixel-based methods (NR-P) and bitstream methods (NR-B).

Recent works focusing on the VR QoE assessment have provided enhanced variants of objective metrics [25], including PSNR of spheres (S-PSNR) and perceptual PSNR (P-PSNR) [26], weighted PSNR (W-PSNR) [27] and spherical SSIM (S-SSIM) [28].

4 Holographic Telepresence Applications

Holographic telepresence communications will allow users from remote locations to interact with holographic data over a communication channel. Consequently, hologram-based applications will impose considerable demands on network capacity, including the ability to provide ultra-low latency, very high bandwidth, and the coordination, synchronization and dynamic adaptation of numerous data streams. The addition of parallax in a holographic material denotes that the viewer may interact with the images in a way that depends on the viewer's location. This shifts the user's participation from being passive in 2D and 3D video to become active and engaged in holograms, significantly boosting the network's requirements. As a result, it necessitates solutions for data capture, transmission and interactivity [29].

The network parameters that are critical in the development of holographic communications and impact significantly QoE include the following [30]: 1) data rates, which are contingent on the hologram's structure, the type of the display and the amount of images to be synchronized; 2) latency, which needs to be ultra-low for truly immersive experiences and avoidance of simulator sickness; 3) synchronization, which is essential for the coordinated provision of data streams from multiple sources; 4) security, which is contingent on how critical is the type of application; 5) resilience, which decreases packet loss, jitter and latency at the system level, whereas the corresponding QoE metrics at the service level would be availability and reliability; and 6) computation, which is contingent on the real-time requirements for hologram generation and reception.

The next breakthrough in XR communications include holograms and multisense communications. While the senses of sight and hearing are already included in audio, video and VR applications, holographic communications integrating all five senses is gaining significant research interest. In principle, holograms may be used to combine a range of sensory experiences and digital encounters can be enhanced with scents and tastes. The emotion-sensing wearable equipment adept of monitoring mental health, enabling social connections and enhancing the user's experience will become the foundations of future mobile networks, employing the holograms as the means of communication [31].

Since the perceived quality of a hologram is determined by a variety of factors such as frame resolution and rate, as well as degree resolution, a comprehension of the implications of each parameter on user perception is critical for reducing bandwidth utilization while preserving QoE. In holograms, the spatial resolution gives an additional dimension for trade-offs that can be used to maximize QoE, similar to how quality of individual images can be sacrificed for frame rates in conventional video. This, nevertheless, necessitates the capacity to adapt parameters in a dynamic manner [29].

5 Conclusions

QoE has a crucial role in the development of the emerging VR and holographic telepresence systems, as these technologies are committed in providing immersive and interactive experiences to their users, thus the quality as it is perceived by the end-users is of outmost importance. In this paper we presented and analyzed the QoE influencing factors and examined the aspects of QoE assessment with regard to VR applications. Moreover, we outlined the elements in the holographic telepresence systems architecture that have an impact on QoE of the end-users.

Acknowledgement. This work was supported by the European Union, through the Horizon 2020 Marie Skodowska-Curie Innovative Training Networks Programme "Mobility and Training for beyond 5G Ecosystems (MOTOR5G)" under grant agreement no. 861219.

References

1. Brunnström, K., et al.: Qualinet white paper on definitions of quality of experience. In: Fifth Qualinet Meeting, Novi Sad (2013)
2. ITU-T, Rec. ITU-T P.10/G.100: Vocabulary for performance, quality of service and quality of experience (2017)
3. ITU-T. Rec. ITU-T G.1035: Influencing factors on quality of experience for virtual reality services (2020)
4. Reiter, U., et al.: Factors influencing quality of experience. In: Möller, S., Raake, A. (eds.) Quality of Experience. TSTS, pp. 55–72. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-02681-7_4

5. Burbeck, C.A., Kelly, D.H.: Spatiotemporal characteristics of visual mechanisms: excitatory-inhibitory model. J. Opt. Soc. Am. **70**(9), 1121–1126 (1980)
6. Greenberg, S., Ainsworth, W.A.: Speech processing in the auditory system: an overview. In: Speech Processing in the Auditory System. Springer Handbook of Auditory Research, vol. 18. Springer, NY (2004). https://doi.org/10.1007/0-387-21575-1_1
7. Kennedy, R.S., Lane, N.E., Berbaum, K.S., Lilienthal, M.G.: Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness. Int. J. Aviat. Psychol. **3**(3), 203–220 (1993)
8. Schatz, R., Hoßfeld, T., Janowski, L., Egger, S.: From packets to people: quality of experience as a new measurement challenge. In: Biersack, E., Callegari, C., Matijasevic, M. (eds.) Data Traffic Monitoring and Analysis. LNCS, vol. 7754, pp. 219–263. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-36784-7_10
9. Takahash, A.: Framework and standardization of quality of experience (QoE) design and management for audiovisual communication services. NTT Tech. Rev. **7**(4), 1–5 (2009)
10. ITU-T, Rec. ITU-T P.800.1: Mean Opinion Score (MOS) Terminology (2006)
11. Alreshoodi, M.A., Woods, J.C.: Survey on QoE/QoS correlation models for multimedia services. Int. J. Distrib. Parallel Syst. **4**(3) (2013)
12. Zhang, Y., et al.: Subjective panoramic video quality assessment database for coding applications. IEEE Trans. Broadcast. **64**(2), 461–473 (2018)
13. ITU-T, Rec. ITU-T P.910: Subjective video quality assessment methods for multimedia applications (2008)
14. ITU-T, Rec. ITU-T P.919: Subjective test methodologies for 360º video on head-mounted displays (2020)
15. Singla, A., Robitza, W., Raake, A.: Comparison of subjective quality test methods for omnidirectional video quality evaluation. In: Proceedings of the MMSP 2019, Kuala Lumpur (2019)
16. Kim, H.K., Park, J., Choi, Y., Choe, M.: Virtual reality sickness questionnaire (VRSQ): motion sickness measurement index in a virtual reality environment. Appl. Ergon. **69**, 66–73 (2018)
17. Perez, P., Oyaga, N., Ruiz, J.J., Villegas, A.: Towards systematic analysis of cybersickness in high motion omnidirectional video. In: Proceedings of the QoMEX 2018, Cagliari (2018)
18. David, E.J., Gutiérrez, J., Coutrot, A., Da Silva, M.P., Le Callet, P.: A dataset of head and eye movements for 360° videos. In: Proceedings of the MMSys 2018, Amsterdam (2018)
19. Wang, Y., Zhang, P.: QoE Management in Wireless Networks, Springer, NY (2016). https://doi.org/10.1007/978-3-319-42454-5
20. ITU-T, Rec. ITU-T G.1011: Reference guide to quality of experience assessment methodologies (2016)
21. Fiedler, M., Hossfeld, T., Tran-Gia, P.: A generic quantitative relationship between quality of experience and quality of service. IEEE Netw. **24**(2), 36–41 (2010)
22. Juluri, P., Tamarapalli, V., Medhi, D.: Measurement of quality of experience of video-on-demand services: a survey. IEEE Commun. Surv. Tutor. **18**(1), 401–418 (2015)
23. Song, W., Tjondronegoro, D.W., Docherty, M.J.: Understanding user experience of mobile video: framework, measurement, and optimization. In: Mobile Multimedia - User and Technology Perspectives, InTech, Rijeka, pp. 3–30 (2012)
24. Vega, M.T., Perra, C., De Turck, F., Liotta, A.: A review of predictive quality of experience management in video streaming services. IEEE Trans. Broadcast. **64**(2), 432–445 (2018)
25. Ruan, J., Xie, D.: A survey on QoE-oriented VR video streaming: some research issues and challenges. Electronics **10**(17), 2155 (2021)
26. Liu, Y., Yang, L., Xu, M., Wang, Z.: Rate control schemes for panoramic video coding. J. Vis. Commun. Image Represent. **53**, 76–85 (2018)
27. Zakharchenko, V., Choi, K.P., Park, J.H.: Quality metric for spherical panoramic video. In: Proceedings of the SPIE Optics + Photonics 2016, San Diego, CA (2016)

28. Chen, S., Zhang, Y., Li, Y., Chen, Z., Wang, Z.: Spherical structural similarity index for objective omnidirectional video quality assessment. In: Proceedings of the IEEE ICME 2018, San Diego, CA (2018)
29. Tataria, H., Shafi, M., Molisch, A.F., Dohler, M., Sjöland, H., Tufvesson, F.: 6G wireless systems: vision, requirements, challenges, insights, and opportunities. Proc. IEEE **109**(7), 1166–1199 (2021)
30. Clemm, A., Vega, M.T., Ravuri, H.K., Wauters, T., De Turck, F.: Toward truly immersive holographic-type communication: challenges and solutions. IEEE Commun. Mag. **58**(1), 93–99 (2020)
31. Shu, J., Chiu, M., Hui, P.: Emotion sensing for mobile computing. IEEE Commun. Mag. **57**(11), 84–90 (2019)