On the design of a framework for QoE Assessment for XR Applications

ERICA ROCCHI, Roma Tre University, Italy
MARCO CARLI, Roma Tre University, Italy

This paper proposes a research project that will focus on the evaluation of Quality of Experience for Extended Reality, which refers to all the emerging technologies that enable the creation of immersive experiences. However, the evaluation of user experience for these systems is still a relatively underdeveloped research area. To this aim, the main objective of the Ph.D. research project will be to develop a framework for evaluating the quality of user experience based on a systematic review of existing literature and an assessment of current solutions. Subsequently, the analysis will be conducted through a series of studies, utilizing a combination of qualitative and quantitive methods. Head Mounted Displays (HMDs) eye tracking accuracy for Extended Reality (XR) applications will be validated and, with the aid of user's data like movements and eye tracking pattern, a ground truth over user's subjective perception of the experience will be developed. The aim is to define metrics for assessing user experience quality in immersive media, such as mixed reality and virtual reality. Quality of Experience (QoE) is closely linked to users' satisfaction, can help to optimize network utilization. It is important to understand QoE in order to guide, in the right way, future XR applications and development.

CCS Concepts: • **Human-centered computing** → Heuristic evaluations; User studies.

Additional Key Words and Phrases: extended reality, virtual reality, augmented reality, quality of experience, human factors.

ACM Reference Format:
Erica Rocchi and Marco Carli. 2024. On the design of a framework for QoE Assessment for XR Applications. In *ACM International Conference on Interactive Media Experiences (IMX '24), June 12–14, 2024, Stockholm, Sweden.*. ACM, New York, NY, USA, 6 pages.

1 INTRODUCTION

The latest advancements in information technology, encompassing high-speed mobile internet, artificial intelligence, advancing computing capabilities, and high-resolution displays, are enabling new interest in users to explore immersive content [7]. While XR popularity and relevance is constantly increasing both in the academic and industrial worlds [4], we are far from a shared agreement on its proper definition and, consequently, there is still a lack of methods for the evaluation of QoE.

1.1 What is XR?

International Telecommunication Union (ITU) defines XR as "an environment containing real or virtual components or a combination thereof, where the variable X serves as a placeholder for any form of new environment", e.g. Augmented Reality (AR), Mixed Reality (MR), or Virtual Reality (VR) [9]. In [2], a review on different domains employing XR applications has revealed that the most used definition in this context is from Azuma's [1] where AR is outlined as any system that combines real and virtual, offering real-time interactivity and the combination of real and virtual objects is in 3D. Another popular definition is from Milgram and Kishino [15] where MR is defined as a system that is able to present real world and virtual world objects together within a single display.

The core limitation of these definitions is that they are restricted to visual features, whereas XR goes beyond the purely visual aspects. Moreover, in [23], five additional elements of reality, which may be introduced in virtual environments, have been identified:

- *Audio*: audio cues are critical for maintaining a high degree of realism, social connection, and spatial awareness. Examples range from immersive experiences in museums and theatres. Audio in a VR setting enables users to naturally look around virtual environments, reach out and interact with objects just like in the real world [12].
- *Motion*: motion is an important aspect for aligning physical and virtual realities.
- *Smell*: creating smells in VR is a vexing problem that has prevented consumer VR devices from offering a full sensory experience in most applications. As an example where smell was reproduced, in [11] two wearable interfaces with miniaturized odour generators and a grid of tiny containers filled with perfumed paraffin wax were developed. Entertainment, medical treatment, human emotion control and teaching prove the great potential of the soft olfaction interface in various practical applications.
- *Haptics*: this feature refers to the use of technology that stimulates the senses of touch, especially to reproduce, the sensations that would be felt by a user interacting directly with physical objects.
- *Taste/Flavour*: these features enable to simulate of the experiences of eating and tasting. In [16], the authors were able to create a virtual food texture through muscle stimulation.

XR holds immense potential, but it also faces several key challenges such as: high hardware costs, motion sickness, designing of intuitive and user-friendly interfaces, privacy and security, and content development. In fact, creating immersive XR content requires specialized skills, i.e. detailed 3D modelling or realistic physics.

Moreover, an unresolved research question concerns QoE assessment within this novel context. QoE, as defined by QUALINET, is *"the degree of delight or annoyance of the user of an application or service. It results from the fulfilment of his or her expectations with respect to the utility and/or enjoyment of the application or service in the light of the user's personality and current state"* [18]. Therefore, QoE is subjective by nature, because of its relationship with user's viewpoint and its own concept of quality. In XR applications QoE is critical for multiple reasons. As examples, poor image and rendering quality can lead to users experiencing nausea and discomfort, and the overall quality perceived by the user can be heavily influenced by the immersion and sense of presence provided. While immersion refers to the degree to which the range of sensory channel is engaged by the virtual simulation, the sense of presence is related to the sense of being in the virtual world [10].

In the industrial world, the discussion around QoE in XR focuses on performance, usability, and accessibility. In [6] it is highlighted that XR applications should have low latency, high reliability, and high data rates. Cortney Harding, CEO of Friends with Holograms, emphasizes the need for XR to be more intuitive and easier to use [5]. In [3], subjective testing is recommended as a critical approach: user testing is essential for identifying problem areas and optimizing XR experiences. Collecting user feedback and observing interactions help refine interfaces, improve usability, and address pain points. Iterative testing and ongoing user participation lead to more polished and user-centric XR applications.

In order to be able to develop XR systems and applications, we need to assert a precise framework of how QoE is perceived by users. This is the basis for the following research question (RQ):

RQ: *How can we develop a comprehensive framework to assess the quality of user experience in XR applications?*

---

2 PROJECT DESCRIPTION

In the proposed research project, the methodologies for the definition of a framework for QoE assessment for XR users will be investigated. To the best of our knowledge, such a framework does not exist. Furthermore, there is limited existing literature on assessing QoE in such context, and solely through qualitative analysis, specifically subjective questionnaires. Therefore, the aim of this project is to assess QoE using both quantitative and qualitative metrics. In [25], presence, embodiment, and overall QoE have been evaluated in a VR application where haptics have been involved for object manipulation. They have implemented a gamified application (an escape room game) as the vehicle for the evaluation. Objective performance data from game execution, combined with subjective questionnaire, have shown that haptic significantly improves key factors such as presence or embodiment. Two initial studies have preliminarily addressed the need for a framework to assess QoE in both the Metaverse and AR [19, 20]. In [19], it has been proposed that QoE in Metaverse should comprehend some traditional features, such as human factors (e.g., sex, age, discomfort), system factors (e.g., headset, content, image resolution), contextual factors (external stimuli) and specific factors for a Metaverse environment, like socio-economical features (e.g., human behaviour, attitude, interaction between humans, money, and job). However, in [20], the necessity to combine subjective with objective evaluations (time to complete a task) is shown. Objective evaluations can give performance rating. On the other hand, subjective ratings and questionnaires can better explain the source of performance rates obtained. Since QoE depends on a plethora of elements, there is a need to set some boundaries to define a framework for QoE in XR. An initial categorization of the factors that will be evaluated are listed hereby:

- **Human Factors:** at the current state, these are evaluated subjectively using state-of-the-art questionnaires:
  - *Cybersickness*: Simulator Sickness Questionnaire (SSQ) is employed to measure motion sickness and general discomfort. The questionnaire asks participants to score 16 symptoms on a four point scale (0-3).
  - *User eXperience*: User eXperience (UX) questionnaire address overall user experience in immersive virtual environments. As an example, in [24], a UX 10 subscales questionnaire is proposed. Their questionnaire evaluates presence, engagement, immersion, flow, usability, skill, emotion, experience consequence, judgement and technology adoption. Presence Questionnaire (PQ) is the most used questionnaire for evaluating sense of presence. This questionnaire is composed of 24 items divided in 5 subscales: involved, natural, auditory, resolution and interface quality. [29].
- **System Influence Factors:** for example, rendering system capabilities. In this context, in a typical subjective quality evaluation experiment, participants observe a series of stimuli and assess their quality using a numerical scale. Subjective quality is often expressed as a Mean Opinion Score (MOS), representing the quality attributed by an average observer to a given stimulus. MOS scores are collected using well-defined methods that aim to ensure identical experimental settings and conditions, facilitating experiment reproducibility. Recommendation ITU-R BT.500-13 [8] describes some of the most popular and widely adopted methods for subjective quality evaluation in 2D imaging. Subjective evaluation functions as a ground truth through which we can validate objective metrics. Some of the features to consider might be:
  - *Content-related*: we aim to evaluate the relationship between QoE and content complexity. We will deal with the application and adaptation of objective metrics for 2D image processing to XR content [17]. Visual complexity measures are categorized as either feature complexity-related (e.g., unstructured pixel-level variation, colour, luminance, and edges) or design complexity-related (e.g., structured design-level variation, number of objects, irregularity of object arrangement, and asymmetry of object arrangement). Metrics will

be validated via a subjective experiment as Single-Stimulus (SS) subjective study, using a 5-level Absolute Category Rating (ACR) scale, where participants are asked to evaluate the stimulus quality.
  - *Media-related*: to evaluate the relation between QoE and metrics about encoding, resolution, position of shown content. MOS can be obtained through a Double Stimulus Impairment Scale (DSIS) experiment with 5 level of distortion perception (from imperceptible to very annoying).
  - *Multi-modal application*: a comparative analysis on the overall user-experience in a multi-modal XR setting will be carried out.
- **Context Influence Factors:** are evaluated through objective metrics:
  - *Physical context*: analysis of the relation between QoE and movement, locomotion and space, through user's movement pattern.

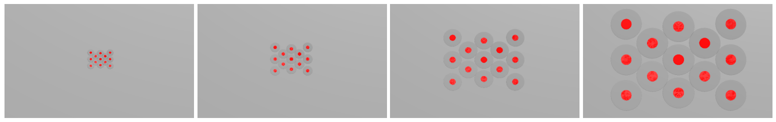
The goal of this Ph.D. is to evaluate QoE in an XR environment. Moreover, how it is affected by either a fully virtual reality or when real and virtual objects blend together. To this aim, the research project contemplates a thorough study of the state-of-the-art on QoE techniques and XR applications and systems. Next, HMD eye tracking accuracy for XR applications will be validated in order to develop a ground truth over user's subjective perception of the experience, with the aid of user's data like movements and eye tracking pattern. The project's final part will see the framework definition, obtaining objective metrics. Future development could include the aid of artificial intelligence algorithms.

2.1 Eye tracking accuracy validation

The initial phase of the research project involves validating the available sensors, specifically those allowing to collect eye tracking data. In today's HMD industry, there's a trend towards employing multiple cameras and sensors to gather additional data aimed at enhancing users' QoE. For instance, Meta's MR headset (Meta Quest Pro [13]) integrates eye tracking, face tracking, and hand tracking functionalities into a single device. Analysing the hardware's capabilities and limitations is crucial for achieving high accuracy in eye-tracking measurements for scientific research purposes [28]. Therefore, one of the first activities will undertake a comparative analysis of eye tracking accuracy and spatial precision for both the Meta Quest Pro and the HTC Vive Focus 3 [27]. In fact, these devices represent two of the most advanced consumer-available technologies integrating eye-trackers, and information regarding their accuracy and spatial precision is either missing, or related to ideal scenarios. In this project, users will test both headsets in two different evaluation settings: head-free and head-contained, using a chin rest support. Accuracy of VR eye-tracking HMD has been tackled in few works [21, 22, 28]. In [21, 22] HTC Vive Pro Eye eye-tracker accuracy has been analysed. From their study has been confirmed that discrepancies between official values and eye tracking performances are present. In [28] has been presented a preliminary user study to evaluate the spatial accuracy, spatial precision, under head-free and head-restrained conditions, in Meta Quest Pro. The average spatial accuracy for the head-free condition was 1.652°, meanwhile for the head-restrained condition was 2.162°. As far as we know, this is the first work on spatial accuracy for Meta Quest Pro. They have used 13 targets placed in a visual field spanning ±15° horizontally and vertically with a ±5° interval. Each target was shown as a red dot measuring around ±7° at a 1 m viewing distance. In our study we have evaluated to present the same stimuli distribution, at different distances, to examine the accuracy and spatial precision of the different headsets, across varying observation distances (0.5 m, 1 m, 2 m, and 4 m).

Therefore, we have developed both VR applications, for Meta Quest Pro and HTC Vive Focus 3, using Unity 2022.3.8f1 and, respectively, Movement SDK v4.2.1 [14] and Wave XR Plugin v5.6.0-r.10.2 [26] have been employed. VR design consists of a low reflectivity uniform grey background, which helps reduce visual distraction. The VR camera, as well

Fig. 1. Stimuli shown at different distances: 0.5 m, 1 m, 2 m and 4 m.



(a) Stimuli at 0.5 m. Target size: 3.42 cm. (b) Stimuli at 1 m. Target size: 5.80 cm. (c) Stimuli at 2 m. Target size: 11.10 cm. (d) Stimuli at 4 m. Target size: 22.54 cm.

as the user's starting point, were placed in Unity's origin world coordinate system. Users will view 13 stimuli in each scene, placed in the HMD Field of View (FOV) with an interval of ±15° horizontally and ±5° vertically, as shown in Figure 1. Each target could be shown as a red dot with increasing size with the distance. We will compute the accuracy of each sample as the angular offset (in degree) between the 3D vector $\vec{t}$ (from the centre of the eyes to the target position) and the normalised 3D vector $\vec{g}$ (from the centre of eyes to the reported gaze position) as formula:

$$\theta_{Offset} = \frac{1}{n}\sum_{i=1}^{n}\theta_i \quad \theta_i = \arccos\left(\frac{\vec{t}\cdot\vec{g}}{||\vec{t}||\cdot||\vec{g}||}\right) * \frac{180}{\pi}$$

We are going to evaluate both Standard Deviation (SD) precision and Root Mean Square (RMS) precision to evaluate eye-tracking performance.

2.2 VR movement pattern and eye tracking analysis

A subsequent phase of the research project tackles the analysis of user's movement and gaze pattern in VR. Utilizing HMD HTC Vive Focus 3, subjects will be asked to evaluate the perceived quality of 15 compressed 3D figures: 5 for training and 10 for testing. They would be able to move freely in a 4 x 4 m room, with no time restrictions. MOS, HMD position and gaze data would be recorded. Our goal could be to find if a correlation between the perceived quality, users' movement pattern and gaze data could be present.

3 CONCLUSIONS AND FUTURE WORKS

We have outlined a possible research project on the assessment of a framework for QoE evaluation in XR applications. QoE is a critical factor for the successful adoption of immersive technologies, which are likely to become the primary method for consuming multimedia content. QoE encompasses a diverse range of factors, involving both the quality of multimedia content and human factors. These factors necessitate a proper assessment through signal processing and, eventually, artificial intelligence techniques. With the aid of additional information given by HMD integrated sensors (e.g., movement pattern and eye tracking data). Future works could include the evaluation of the social impact on the user, that is increasingly becoming a focal point in the assessment of XR applications.

REFERENCES

[1] R. T. Azuma. 1997. A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments* 6, 4 (08 1997), 355–385. https://doi.org/10.1162/pres.1997.6.4.355 arXiv:https://direct.mit.edu/pvar/article-pdf/6/4/355/1623026/pres.1997.6.4.355.pdf

[2] A. Becker and C. Freitas. 2023. Evaluation of XR Applications: A Tertiary Review. *ACM Comput. Surv.* 56, 5, Article 110 (nov 2023), 35 pages. https://doi.org/10.1145/3626517

[3] CodeGama. 2024. Building Immersive Experiences with Extended Reality (XR) Technologies. Available from https://www.linkedin.com/pulse/building-immersive-experiences-extended-reality-xr-technologies. Last accessed: 2024-06-05.

[4] A. Cárdenas-Robledo, O. Hernández-Uribe, C. Reta, and A. Cantoral-Ceballos. 2022. Extended reality applications in industry 4.0. – A systematic literature review. *Telematics and Informatics* 73 (2022), 101863. https://doi.org/10.1016/j.tele.2022.101863

[5] Ericsson. 2024. Expert Insights: XR and the Experiences of the Future. Available from https://medium.com/watch-this-space/expert-insights-xr-the-experiences-of-the-future-f77574a5dfc6. Last accessed: 2024-06-05.

[6] Ericsson. 2024. Future network requirements for extended reality applications. Available from https://www.ericsson.com/en/reports-and-papers/ericsson-technology-review/articles/future-network-requirements-for-xr-apps. Last accessed: 2024-06-05.

[7] W. Hoyer, M. Kroschke, B. Schmitt, K. Kraume, and V. Shankar. 2020. Transforming the Customer Experience Through New Technologies. *Journal of Interactive Marketing* 51 (2020). https://doi.org/10.1016/j.intmar.2020.04.001

[8] ITU-R BT.500-13. 2012. Methodology for the subjective assessment of the quality of television pictures. *International Telecommunications Union* (2012).

[9] ITU-T P.1320. 2022. Quality of experience assessment of extended reality meetings. *International Telecommunications Union* (2022).

[10] G. Kim and F. Biocca. 2018. Immersion in Virtual Reality Can Increase Exercise Motivation and Physical Performance. In *Virtual, Augmented and Mixed Reality: Applications in Health, Cultural Heritage, and Industry: 10th International Conference, VAMR 2018, Held as Part of HCI International 2018, Las Vegas, NV, USA, July 15-20, 2018, Proceedings, Part II* (Las Vegas, FL, USA). Springer-Verlag, Berlin, Heidelberg, 94–102. https://doi.org/10.1007/978-3-319-91584-5_8

[11] Y. Liu, C. Yiu, Z. Zhao, W. Park, R. Shi, X. Huang, Y. Zeng, K. Wang, T. Wong, S. Jia, J. Zhou, Z Gao, L. Zhao, K. Yao, J. Li, C. Sha, Y. Gao, G. Zhao, Y. Huang, and X. Yu. 2023. Soft, miniaturized, wireless olfactory interface for virtual reality. *Nature Communications* 14 (2023). https://doi.org/10.1038/s41467-023-37678-4

[12] Medium. 2024. Creating Immersive Experiences with Audio-Visual Technology. Available from https://medium.com/@jamesespinosa926/creating-immersive-experiences-with-audio-visual-technology-5a89b39fc7ea. Last accessed: 2024-04-16.

[13] Meta. 2024. Meta Quest Pro. Available from https://www.meta.com/quest/quest-pro/. Last accessed: 2024-03-22.

[14] Meta. 2024. Movement SDK for Unity. Available from https://developer.oculus.com/documentation/unity/move-overview/. Last accessed: 2024-03-29.

[15] P. Milgram and F. Kishino. 1994. A Taxonomy of Mixed Reality Visual Displays. *IEICE Trans. Information Systems* E77-D, no. 12 (1994), 1321–1329.

[16] A. Niijima and T. Ogawa. 2016. Study on Control Method of Virtual Food Texture by Electrical Muscle Stimulation. In *Adjunct Proceedings of the 29th Annual ACM Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16 Adjunct)*. New York, NY, USA, 199–200. https://doi.org/10.1145/2984751.2984768

[17] G. Overgoor, W. Rand, W. van Dolen, and M. Mazloom. 2022. Simplicity is not key: Understanding firm-generated social media images and consumer liking. *International Journal of Research in Marketing* 39, 3 (2022), 639–655. https://doi.org/10.1016/j.ijresmar.2021.12.005

[18] A. Perkis, C. Timmerer, S. Baraković, J. B. Husić, S. Bech, S. Bosse, J. Botev, K. Brunnström, L. Cruz, K. De Moor, et al. 2020. QUALINET white paper on definitions of immersive media experience (IMEx). *arXiv preprint arXiv:2007.07032* (2020).

[19] S. Porcu, A. Floris, and L. Atzori. 2022. Quality of Experience in the Metaverse: An Initial Analysis on Quality Dimensions and Assessment. In *2022 14th International Conference on Quality of Multimedia Experience (QoMEX)*. 1–4. https://doi.org/10.1109/QoMEX55416.2022.9900897

[20] J. Puig, A. Perkis, F. Lindseth, and T. Ebrahimi. 2012. Towards an efficient methodology for evaluation of quality of experience in Augmented Reality. In *2012 Fourth International Workshop on Quality of Multimedia Experience*. 188–193. https://doi.org/10.1109/QoMEX.2012.6263864

[21] I. Schuetz and K. Fiehler. 2022. Eye tracking in virtual reality: Vive pro eye spatial accuracy, precision, and calibration reliability. *Journal of Eye Movement Research* 15, 3 (2022).

[22] A. Sipatchin, S. Wahl, and K. Rifai. 2021. Eye-Tracking for Clinical Ophthalmology with Virtual Reality (VR): A Case Study of the HTC Vive Pro Eye's Usability. *Healthcare* 9, 2 (2021). https://doi.org/10.3390/healthcare9020180

[23] M. Speicher, B. D. Hall, and M. Nebeling. 2019. What is Mixed Reality?. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. New York, NY, USA, 1–15. https://doi.org/10.1145/3290605.3300767

[24] K. Tcha-Tokey, E. Loup-Escande, O. Christmann, and S. Richir. 2016. A Questionnaire to Measure the User Experience in Immersive Virtual Environments. *VRIC '16: Proceedings of the 2016 Virtual Reality International Conference.* https://doi.org/10.1145/2927929.2927955

[25] A. Villegas, P. Perez, R. Kachach, F. Pereira, and E. Gonzalez-Sosa. 2020. Realistic training in VR using physical manipulation. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 109–118. https://doi.org/10.1109/VRW50115.2020.00025

[26] Vive. [n. d.]. VIVE Wave XR Plugin. Available from https://hub.vive.com/storage/docs/en-us/UnityXR/UnityXRSdk.html. Last accessed: 2024-03-29.

[27] HTC Vive. 2024. HTC Vive Focus 3 Eye Tracker. Available from https://business.vive.com/eu/product/vive-focus-3-eye-tracker. Last accessed: 2024-03-22.

[28] S. Wei, D. Bloemers, and A. Rovira. 2023. A Preliminary Study of the Eye Tracker in the Meta Quest Pro. In *Proceedings of the 2023 ACM International Conference on Interactive Media Experiences* (Nantes, France) *(IMX '23)*. Association for Computing Machinery, New York, NY, USA, 216–221. https://doi.org/10.1145/3573381.3596467

[29] B. G Witmer and M. J. Singer. 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence* 7, 3 (1998), 225–240.

Received XXXXXXX;; revised XXXXXXX;; accepted XXXXXXX;