|  |  |
| --- | --- |
| 3GPP TR 26.812 V0.3.0 (2023-02) | |
| Technical Report | |
| 3rd Generation Partnership Project;  Technical Specification Group SA;  Study on QoE Metrics for AR/MR Services (Release 18) | |
|  | |
|  |  |
|  | |
| The present document has been developed within the 3rd Generation Partnership Project (3GPP TM) and may be further elaborated for the purposes of 3GPP. The present document has not been subject to any approval process by the 3GPPOrganizational Partners and shall not be implemented. This Specification is provided for future development work within 3GPPonly. The Organizational Partners accept no liability for any use of this Specification. Specifications and Reports for implementation of the 3GPP TM system should be obtained via the 3GPP Organizational Partners' Publications Offices. | |

|  |
| --- |
|  |
| ***3GPP***  Postal address  3GPP support office address  650 Route des Lucioles - Sophia Antipolis  Valbonne - FRANCE  Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16  Internet  http://www.3gpp.org |
| ***Copyright Notification***  No part may be reproduced except as authorized by written permission. The copyright and the foregoing restriction extend to reproduction in all media.  © 2022, 3GPP Organizational Partners (ARIB, ATIS, CCSA, ETSI, TSDSI, TTA, TTC).  All rights reserved.  UMTS™ is a Trade Mark of ETSI registered for the benefit of its members  3GPP™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners LTE™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners  GSM® and the GSM logo are registered and owned by the GSM Association |

Contents

Foreword 4

Introduction 5

1 Scope 6

2 References 6

3 Definitions of terms, symbols and abbreviations 7

3.1 Terms 7

3.2 Symbols 8

3.3 Abbreviations 8

4 Current state of AR/MR QoE metrics outside 3GPP 8

4.1 AR/MR QoE related work in ITU-T 8

4.2 AR/MR QoE related work in IEEE 11

4.3 AR/MR QoE related work in MPEG 11

5 Relevant VR metrics 12

5.1 Head-motion aware viewport quality metric (HMAVQ) 12

5.1.1 Background 12

5.1.2 Metric description 12

6 Identification of AR/MR QoE Metrics for 3GPP 13

6.1 General 13

6.2 AR/MR QoE reference model 14

6.2.1 Observation Point 1 14

6.2.2 Observation Point 2 15

6.2.3 Observation Point 3 15

6.2.4 Observation Point 4 15

7 QoE Metrics Subjective Assessment 16

8 Other QoE metrics and collaborations with other 3GPP groups 16

9 Conclusions and Recommendations 16

Annex <A> (informative): <Normative annex for a Technical Specification> 16

Annex <X> (informative): Change history 19

# Foreword

This Technical Report has been produced by the 3rd Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

x the first digit:

1 presented to TSG for information;

2 presented to TSG for approval;

3 or greater indicates TSG approved document under change control.

y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.

z the third digit is incremented when editorial only changes have been incorporated in the document.

In the present document, modal verbs have the following meanings:

**shall** indicates a mandatory requirement to do something

**shall not** indicates an interdiction (prohibition) to do something

The constructions "shall" and "shall not" are confined to the context of normative provisions, and do not appear in Technical Reports.

The constructions "must" and "must not" are not used as substitutes for "shall" and "shall not". Their use is avoided insofar as possible, and they are not used in a normative context except in a direct citation from an external, referenced, non-3GPP document, or so as to maintain continuity of style when extending or modifying the provisions of such a referenced document.

**should** indicates a recommendation to do something

**should not** indicates a recommendation not to do something

**may** indicates permission to do something

**need not** indicates permission not to do something

The construction "may not" is ambiguous and is not used in normative elements. The unambiguous constructions "might not" or "shall not" are used instead, depending upon the meaning intended.

**can** indicates that something is possible

**cannot** indicates that something is impossible

The constructions "can" and "cannot" are not substitutes for "may" and "need not".

**will** indicates that something is certain or expected to happen as a result of action taken by an agency the behaviour of which is outside the scope of the present document

**will not** indicates that something is certain or expected not to happen as a result of action taken by an agency the behaviour of which is outside the scope of the present document

**might** indicates a likelihood that something will happen as a result of action taken by some agency the behaviour of which is outside the scope of the present document

**might not** indicates a likelihood that something will not happen as a result of action taken by some agency the behaviour of which is outside the scope of the present document

In addition:

**is** (or any other verb in the indicative mood) indicates a statement of fact

**is not** (or any other negative verb in the indicative mood) indicates a statement of fact

The constructions "is" and "is not" do not indicate requirements.

# Introduction

In Rel-18, simple QoE Metrics for AR media will be specified in MeCar WI, which focus on which VR QoE metrics can be reused and enhanced for AR media. This study intends to complement MeCar WI by studying additional and new information related to QoE metrics which are different from VR QoE metrics. QoE metrics typically reflect the expected user experience based on measurable parameters. XR (e.g. AR/MR) user experience is impacted by different aspects, such as the media quality, the rendering capabilities of the device, the tracking capabilities of the devices, etc. In order to identify the impact of different QoE factors, as well as in order to identify how QoE related parameters can be measured in devices, additional detailed study is necessary.

Based on this, the present document collects available information in other organizations such as the ITU-T, MPEG-I groups, and provides recommendation on normative work for new XR QoE metrics in such that XR (e.g. AR/MR) user experience can be enhanced with the complete defined QoE metrics.

# 1 Scope

This document addresses information collection and QoE metrics definition aspects when the XR (e.g. AR/MR) devices are based on the architecture defined in MeCAR WI. The following details are in scope:

- Collect relevant external information on QoE Metrics for AR and XR services, for example taking into account information in ITU-T, MPEG or other groups dealing with quality measurements, include device related QoE metrics, network transmission related QoE metrics, content handling related QoE metrics, and other immersiveness/presence related QoE metrics.

- Documentation of subjective tests results on XR QoE metrics, if considered relevant.

- Identification of relevant XR QoE Metrics and their impacts on the user experience.

- Identification of relevant observation points and define the measurement and derivation of relevant XR QoE metrics in the device architecture based on MeCAR.

- Collaboration with relevant groups or specifications on NWDAF, RRC-based metrics configuration and collection.

- Provide recommendation on normative work for new XR QoE metrics based on the findings in this study.

# 2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non‑specific.

- For a specific reference, subsequent revisions do not apply.

- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[2] 3GPP TR 26.998: "Support of 5G glass-type Augmented Reality / Mixed Reality (AR/MR) devices".

[3] 3GPP TR 26.928: "Extended Reality (XR) in 5G".

[4] 3GPP TR 26.926: "Traffic Models and Quality Evaluation Methods for Media and XR Services in 5G Systems".

[5] 3GPP TS 26.119: "Media Capabilities for Augmented Reality".

[6] ITU-T G.1036, “Quality of experience (QoE) influencing factors for augmented reality (AR) services”.

[7] ITU-T P.1320, “QoE assessment of extended reality (XR) meetings”.

[8] Technical Report ITU-T GSTR-5GQoE, QoE requirements for real-time multimedia services over 5G networks.

[9] ITU-T H.430.1 Recommendation ITU-T H.430.3 (2018), Service scenario of immersive live experience (ILE).

[10] ITU-T RGM-Q8-DOC10-R2 (2022-06) H.IIS-reqts: “Requirements of Interactive Immersive Services”

[11] IEEE P3333.1.1/D2, May 2022 - IEEE Draft Standard for Quality of Experience (QoE) and Visual-Comfort Assessments of Three-Dimensional (3D) Contents Based on Psychophysical Studies

[12] IEEE P3333.1.2/D3, Apr 2022 - IEEE Draft Standard for the Perceptual Quality Assessment of Three Dimensional (3D), Ultra High Definition (UHD) and High Dynamic Range (HDR) Contents

[13] IEEE Std 3333.1.3-2022 “IEEE Standard for the Deep Learning-Based Assessment of Visual Experience Based on Human Factors”

[14] ISO/IEC 23090-6:2021 Information technology — Coded representation of immersive media — Part 6: Immersive media metrics

[15] ISO/IEC JTC 1/SC 29/WG 03 N0777 Technologies under Consideration for ISO/IEC 23090-6 WG 03, MPEG Systems

[16] ISO/IEC JTC 1/SC 29/WG 03 N00710. “Text of ISO/IEC 23090-6 FDAM 1 Immersive Media Metrics for V3C data and OMAF”.

[17] ISO/IEC JTC 1/SC 29/WG 03 N00826. “Preliminary WD of ISO/IEC 23090-6 AMD 2 Additional latencies and Other Improvements”.

[18] Kara, B., Akcay, M. N., Begen, A. C., Ahsan, S., Curcio, I. D., & Aksu, E. B. (2022). Could Head Motions Affect Quality When Viewing 360-Degree Videos?. IEEE MultiMedia.

[19] Van der Hooft, J., Vega, M. T., Petrangeli, S., Wauters, T., & De Turck, F. (2019). Quality assessment for adaptive virtual reality video streaming: A probabilistic approach on the user’s gaze. In 2019 22nd conference on Innovation in Clouds, Internet and Networks (ICIN) (pp. 19-24). IEEE.

[20] Hu, Z., Bulling, A., Li, S., & Wang, G. (2021). Ehtask: Recognizing user tasks from eye and head movements in immersive virtual reality. IEEE Transactions on Visualization and Computer Graphics.

[21] Rai, Y., Le Callet, P., & Guillotel, P. (2017, May). Which saliency weighting for omni directional image quality assessment?. In 2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX) (pp. 1-6). IEEE.

[22] S4-221567 MeCAR Permanent Document v4.0.0

# 3 Definitions of terms, symbols and abbreviations

This clause and its three subclauses are mandatory. The contents shall be shown as "void" if the TS/TR does not define any terms, symbols, or abbreviations.

## 3.1 Terms

For the purposes of the present document, the terms given in 3GPP TR 21.905 [1] and the following apply. A term defined in the present document takes precedence over the definition of the same term, if any, in 3GPP TR 21.905 [1].

Definition format (Normal)

**<defined term>:** <definition>.

**example:** text used to clarify abstract rules by applying them literally.

## 3.2 Symbols

For the purposes of the present document, the following symbols apply:

Symbol format (EW)

<symbol> <Explanation>

## 3.3 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905 [1] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in 3GPP TR 21.905 [1].

AR Augmented Reality

MR Mixed Reality

# 4 Current state of AR/MR QoE metrics outside 3GPP

## 4.1 AR/MR QoE related work in ITU-T

In ITU-T, there are many study groups focusing on the quality of experience (QoE) for XR services, e.g. Study Group 12 (SG12), Study Group 16 (SG16). Typically, Study Group 12 (SG12) majoring in the P​erformance, QoS & QoE​ is quite relevant to the study of AR/MR QoE in SA4.

In SG12, there are several ongoing or completed work about the AR/MR QoE.

1) ITU-T G.1036 “Quality of experience (QoE) influencing factors for augmented reality (AR) services” [6].

This Recommendation has been consented by the recent ITU-T Study Group 12 meeting, June 2022. It lists typical use cases of augmented reality services and identifies the key QoE factors in it, and also gives a suggested scheme for AR QoE assessment in future works.

Different from Virtual Reality, which aims to completely replace the user’s real world environment with a simulated one, Augmented Reality keeps the real world and puts additional information to the natural environment.

Due to addition of new ways to locate the self-position of users and new display mode of perceptual information, a set of new requirements to QoE assessment to characterize AR’s immersive video, spatial audio, and interactivity are emerging. Besides, it’s important to address the requirements and basic factors affecting the VR QoE before benchmarking work is executed. This document identifies the key factors affecting user-perceived experience.

The identified AR QoE metrics can be shown as following:

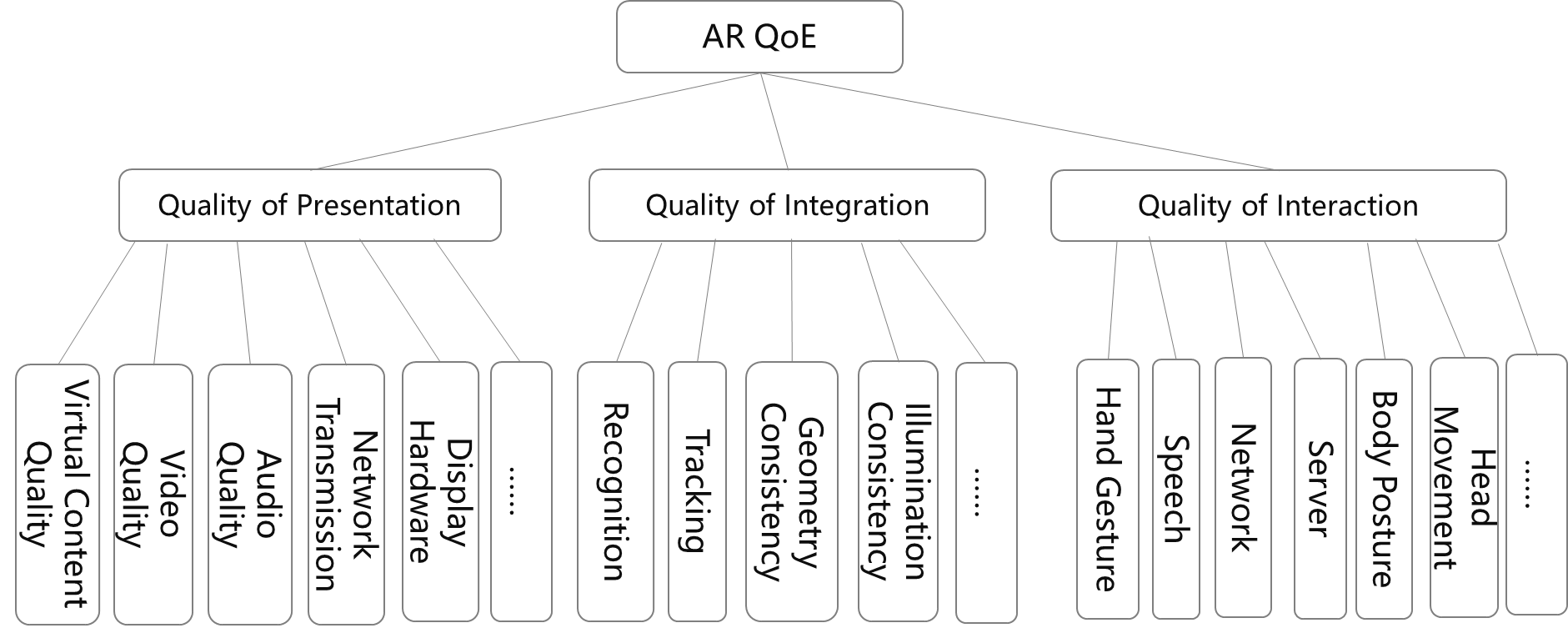


Figure 4.1-1: AR QoE assessment scheme [6]

Table 4.1-1: Identified QoE factors for AR services [6]

|  |  |  |
| --- | --- | --- |
| Human Factors | Vision | myopia, hyperopia, astigmatism, and amblyopia. |
| Hearing | the audio perceptibility |
| Touch/Force Sensation | tactile/force and other types of perceptual enhancement. |
| System Influence Factors | Content Related | Real Environment, virtual content, Superimposition of Real Environment and Virtual Content |
| Media and Coding Related | Codec, Bitrate, Resolutionl, Framerate, coding delay, streaming quality, stordage and transport. |
| Network and Transmission related | Bandwidth, latency, packet loss, jitter. |
| Recognition related | Marker attributes, Recognition response. |
| Consistency | Geometry consistency, Lighting consistency, Time consistency |
| Harware Related | Comfortableness, FoV, stereo, depth range |
| Interaction Related | Hand Gesture, Speech, Body Posture, Tangible Interface, Eye/head based interfaces, Brain-computer interfaces |
| Content Influence Factors | Physical context | physical environment condition |
| Colloboration context | collaborative environment for multi-person AR services. |
| Task context | tasks and goals users are aiming to in their minds. |
| Temporal context | the frequency and duration is to use AR services. |
| General factors | Human factors | Olfactory, Multisensory integration, Simulator sickness, Static and Dynamic Human Factor |
| Hardware related factors | Colour Space, Dynamic Range, Refresh Rate |

2) ITU-T P.1320 “QoE assessment of extended reality (XR) meetings” [7].

This Recommendation has been consent by the recent ITU-T Study Group 12 meeting, June 2022. It advises on aspects of importance for QoE assessment of telemeetings with extended reality elements. The goal is to define the human, context, and system factors that affect the choice of the QoE assessment procedure and metrics when extended reality telemeeting systems are under evaluation.

This Recommendation focuses on aspects of importance for the assessment of Quality of Experience (QoE) of different types of eXtended Reality (XR) telemeetings, which may comprise a combination of telemeetings taking place in virtual reality (VR), augmented reality (AR), or mixed reality (MR) environments. It targets XR services aiming to immerse the user and augment the exchange of information by delivering interactive real-time uni- or multimodal sensory information for two-party and multiparty communication. The services include telemeetings taking place in a virtual location or a combination of virtual and real locations; telemeetings with mixtures of real and virtual participants; telemeetings with augmented elements for collaboration; virtual conferences; and joint teleoperation of equipment.

While multiple QoE evaluation methodologies for telemeetings have been developed, novel XR telemeeting services may result in cognitive effects that are not covered by the existing Recommendations. These effects may include simulator sickness, fatigue, immersion, or presence, for example. This document advises on the key QoE factors affecting the user experience of an XR telemeeting service.

The identified QoE metrics for XR meetings can be found as the following.

Table 4.1-2: Factors influencing QoE of XR telemeetings [7]

|  |  |  |
| --- | --- | --- |
| Human Influence Factors | Perceptual and congnitive characteristics | Visual, audio, olfactory, tactile acuity. |
| Internal state of individual participants | Simulator sickness, immersion, level of expertise, spatial intelligence and introversion, etc. |
| Conversation behaviour | Conversation behaviour, degrees of involvement. |
| Relations between participants | A closer (e.g., romantic) relationship may be preferred for this kind of tactile stimulation. |
| Language and body language aspects | linguistic social cues and body language. |
| Content Influence Factors | Communication environment | Real Environment, virtual content, Superimposition of Real Environment and Virtual Content |
| Communication scenarios | Codec, Bitrate, Resolution, Framerate, coding delay, streaming quality, storage and transport. |
| Time aspects | Bandwidth, latency, packet loss, jitter. |
| System Influence Factors | Human/world-related factors | Degree of freedom, representation of users, realism/style, locamotion, position, proxemics |
| Rendering | Rendering per client, multi-party effects on rendering, redering errors, resolutions, foveated rendering, overlaying rendering images. |
| Network and compression | Media encoding/decoding, latency, bandwidth, synchronization |
| Temporal context | the frequency and duration is to use AR services. |

3) Technical Report ITU-T GSTR-5GQoE, QoE requirements for real-time multimedia services over 5G networks [8]

This report has been produced by the 5G-KPI working group of the Video Quality Experts Group (VQEG) and defines a scope for the analysis of QoE in 5G services and several use cases where this scope is applicable.

This Technical Report defines a scope for the analysis of QoE in 5G services and several use cases where this scope is applicable. Such use cases are: Tele-operated Driving, Wireless Content Production, Mixed Reality Offloading, and First Responder Networks. For all the targeted use cases, the work item aims to study the specific QoE requirements, as well as the required performance and features from the network. By addressing them in parallel, it is possible to find synergies between them and, more relevantly, extract the common information that can be used to also analyse other use cases that may arise outside the scope of this work item.

Specially, the MR Offloading case assumes that the immersive environment where the application takes place is either the real world around the user (AR) or a completely virtual environment (VR). 5G-network capabilities (i.e. Edge Computing) are used to enable running state-of-the-art MR algorithms while wearing a lightweight HMD.

The identified QoE factors for MR Offloading can be found in the following:

- Motion-to-photon latency

- Responsiveness in human-virtual interaction

- Embodiment feeling

- RTT time

- Peak uplink/downlink throughputs

- Coding/processing delay

- Mean network throughput

- Type of MR application

- Temporal context

Study Group 16 (SG16) majoring in the “immersive live experience (ILE) ” and “interactive immersive services (IIS)” shares some views on the requirements of immersive services, e.g. extended reality service. The related details are listed as below:

1) ITU-T H.430.1 “Requirements for immersive live experience (ILE) services” [9] provide some requirements of ILE as follows:

General requirements: Displaying real-sized objects, Direction of sounds, Reconstruction of stage effects, Spatial environment, Synchronous media representation of multiple assets, Augmented information attachment ability.

High-level requirements: Real-time object extraction, Synchronous transmission of multiple media streams, Media processing, High-realistic auditory lateralization, Video stitching.

2) ITU-T H.430.6 H.IIS-reqts: “Requirements of Interactive Immersive Services (IIS)” work item is under study [10]. The baseline text provides some service and capability requirements of IIS service as follows:

Service and capability requirements: Support for interactive capabilities, Support for synchronous transmission of concurrent streams, Intelligent distribution of massive multimedia data, Media processing, Network status awareness and QoS scheduling, Multimedia stream embellishing.

Both the ILE and IIS work item in ITU-T SG 16 are mainly focusing on the requirements of immersive service without any information related to how to map the requirements to the AR/MR QoE metrics or factors.

## 4.2 AR/MR QoE related work in IEEE

In IEEE, P3333.1 series standards focus on the experience assessment of 3D contents, and visual experience assessment based on human factors. The details are listed as below.

1) IEEE P3333.1.1/D2, May 2022 - IEEE Draft Standard for Quality of Experience (QoE) and Visual-Comfort Assessments of Three-Dimensional (3D) Contents Based on Psychophysical Studies [11] captures the subjective assessment for quantifying the visual discomfort and quality of experience (QoE) of 3D image and video.

2) IEEE P3333.1.2/D3, Apr 2022 - IEEE Draft Standard for the Perceptual Quality Assessment of Three Dimensional (3D), Ultra High Definition (UHD) and High Dynamic Range (HDR) Contents [12] defines quality metrics for 3D, UHD and HDR quality assessment.

3) IEEE Std 3333.1.3-2022 “IEEE Standard for the Deep Learning-Based Assessment of Visual Experience Based on Human Factors” [13] defines deep learning-based metrics of content analysis and quality of experience (QoE) assessment for visual contents. The standards mainly cover quality assessment of visual contents and cybersickness assessment of visual contents for VR.

The above IEEE standards firstly provide the methods of video quality assessment (e.g. 3D contents quality), and then cover deep learning models considering human factors for various QoE assessments from an objective point of view. However, those information are not related to AR/MR QoE metrics definition, and even the IEEE Std 3333.1.3 only covers QoE assessment for VR cybersickness.

## 4.3 AR/MR QoE related work in MPEG

MPEG has defined a series of standards for immersive media with a project of MPEG-I (ISO/IEC 23090 Coded Representation of Immersive Media). It contains 26 parts related with immersive media components. In the part 6, the immersive media metrics and the measurement framework are specified in ISO/IEC 23090-6:2021 [14]. This standard also includes a VR client reference model with observation and measurement points for collection of the metrics. The immersive media metrics in [14] are listed as below:

- Rendered FOV set metric

- Display information set metric

- Rendered viewports metric

- Comparable quality viewport switching latency metric

In addition, Viewpoint Mismatch Duration Metric is also agreed to be added into the specification ISO/IEC 23090-6 [15].

Furthermore, the following metrics are defined in ISO/IEC 23090-6 AMD1 [16] and proposed in ISO/IEC 23090-6 AMD2 [17]:

- Omnidirectional Viewpoint Switching Latency metric [16]

- V3C Viewpoint Switching Latency [16]

- Viewpoint Switching Latency [17]

All the 8 immersive media metrics are defined for the VR service based on VR client model, not relevant to AR/MR QoE metrics.

Editor’s Note:

* Collect relevant external information on QoE Metrics for AR and XR services, for example taking into account information in ITU-T, MPEG or other groups dealing with quality measurements, include device related QoE metrics, network transmission related QoE metrics, content handling related QoE metrics, and other immersiveness/presence related QoE metrics.

# 5 Relevant VR metrics

## 5.1 Head-motion aware viewport quality metric (HMAVQ)

### 5.1.1 Background

Perceptual quality is defined as a user’s degree of satisfaction while viewing a video. Various objective quality metrics (e.g. average viewport quality, motion-to-high-quality delay) were proposed to estimate the quality of a 360-degree video. TS 26.118 specified a metric where the viewport quality is calculated by multiplying the quality ranking of each tile in the viewport by the percentage of the viewport it covers.

According to some studies [19, 20], the methods that best estimate the actual user experience are the ones that take into account the human visual system (HVS). While these studies considered user’s eye gaze, they largely ignored the effects of the head motion velocity, direction and duration. For instance, the content rendered in the viewport may be unfocused, or even blurry, to a human eye if the head is turning too fast (that is, the user can’t focus on the content during a fast head motion). In this case, this individual viewport’s quality may have only a slight, if any, effect on the overall user experience since the low-quality video is shown to the user only for a brief amount of time.

Also, the head motion may affect the region-of-interest (ROI) the user is paying attention to inside the viewport. Since the head and eye gaze mainly move in the same direction [20], the ROI shifts from the viewport center to the edges of the viewport, in the direction of the head motion. Consequently, the quality of the tiles in the opposite direction are expected to matter less and have a smaller effect on the overall user experience. Studies also show that users tend to view the content with little focus during a fast head motion and the actual quality of the viewport during this motion is not as relevant [20, 21]. This implies that not every viewport a user has viewed has an equal weight in determining the user’s overall experience.

A QoE metric that takes into account the above considerations can facilitate more advanced algorithms for viewport-dependent streaming of XR experiences by evaluating these algorithms more accurately. In the next section, a metric that brings “head motion awareness” in quality assessment is described. Inclusion of head motion exemplified in this metric can be useful for development of more effective AR/MR metrics.

### 5.1.2 Metric description

Computation of the head-motion aware viewport quality metric (HMVAQ) comprises two steps. In the first step, the quality of an individual viewport is calculated during head motion. In the second step, the overall viewport quality over a period of time is calculated from the individual viewport qualities. Details of these two steps are described below.

**Individual viewport qualities**

Quality of the individual viewports is calculated using sampled points that are shifted and weighted based on head motion speed and direction. Each viewport is sampled with n circles, where each circle has m sample points. The circles are indexed from 1 to n, and each circle’s diameter grows from the viewport center towards the viewport edges. Depending on head movements, the circles are shifted by  degrees horizontally and  degrees vertically, where these shift amounts depend on the radius  of each circle, head motion velocities in horizontal and vertical axes  and a speed threshold  beyond which the viewport starts becoming unfocused (blurry). The shift amounts are calculated as:



where w and h are the width and height of the viewport, respectively.The circle weights are assigned linearly in descending order from viewport center towards the edges. The weight of circle i ( is calculated by dividing its reversed order by the sum of the circle indexes:



Subsequently, the points on each circle are distributed with an angular distance of . The coordinates of each point are calculated using the parametric equation of the circle. These coordinates are then used to find the tile T with the projection of the point. The quality of the tile T, on which a point is projected, is recorded as the quality of that point. The same operation is applied for all points, and the average across all points determines the quality of the circle i. The circle quality is then multiplied by the weight of the circle , and the weighted sum of all weight circle qualities yields the individual viewport quality.

Example sampled viewports at different head motion velocities are shown in Figure 5.1.2-1, where vertical and horizontal lines represent the tile boundaries. Fig. 5.1.2-1 (a) shows the case here the viewport is stationary (no head movement). Figures 5.1.2-1 (b) and 5.1.2-1 (c) show the shifted circles for slow and fast head movements, respectively. Fig. 5.1.2-1 (d) shows the case where the head movement is faster than the speed threshold () in the horizontal axis.

Chart, diagram

Description automatically generated

**Figure 5.1.2-1: Sampled viewports at different head motion speeds**

**Overall viewport quality**

Overall viewport quality is calculated as the weighted average of the individual viewport qualities over a period of time. The weights are assigned based on head motion velocity. The idea is to assign smaller weights to the viewports viewed at faster head motions, whereas those viewed in stationary head positions or at slower head motions are assigned a larger weight. Weight assignment is not a trivial task and the best one can do is to find a good method empirically. After tests, a thresholding approach is adopted such that a weight equal to 1 is used for viewports where  is reached or exceeded, and 2 for the other viewports. More sophisticated weight assignment functions might perform better, and this deserves further testing and validation.

For further details on the metric, please refer to [18].

# 6 Identification of AR/MR QoE Metrics for 3GPP

## 6.1 General

There are 22 core use cases identified for AR/MR devices in TR 26.998 [2] and they are further clarified into the several categories based on the similar requirements for media flow and device functional structure:

- Immersive media downlink streaming

- 5G interactive immersive service

- 5G cognitive immersive service

- AR conversational service

- Shared AR conversational service

The AR/MR QoE metrics can be studied based on the following aspects:

1) Content part: study needs to be conducted on factors of content part which would help analyse user experience.

2) Delivery part: changing network conditions may lead to problems in user experience, especially the impact of transmission latency on user experience.

3) Device part: device capabilities also have impact on user experience.

QoE metrics relevant with the above aspects need to be studied under this study item, and based on the result of this study, user experience of AR/MR service could be evaluated.

## 6.2 AR/MR QoE reference model

According to the clause 4.1.2 of MeCAR PD [22], a defined AR/MR QoE framework and the observation points can be reused as baseline for the AR QoE reference model, which is illustrated in Figure 6.2-1.

NOTE: The observation points can also be used to identify the advanced AR/MR QoE metrics.

**Figure 6.2-1: AR/MR QoE reference model and Metrics Observation Points [22]**

It’s also noted that the above observation points may be further updated based on the agreements of the AR/MR QoE metrics identification and definition.

### 6.2.1 Observation Point 1

XR Runtime is a set of functions that interface with a platform to perform commonly required operations, such as accessing the controller/peripheral state, getting current and/or predicted tracking positions, performing spatial computing, and submitting rendered frames to the display processing unit. The XR Runtime provides the viewer pose and projection parameters needed to render each view for use in a composition projection layer.

XR Source Management addresses the management of data sources provided through the XR runtime such as microphones, cameras, trackers, etc. The XR Source Management may expose information to the application or may provide a subset to the media access function to be sent remote.

Presentation Engine is a set of composite renderers, rendering the component of the scenes, based on the input from the Scene Manager. The Scene Manager together with the Presentation Engine that includes functions such as scene composition and possible complex audio or visual rendering.

Observation point 1 is derived from the XR Runtime API which exchanges information between XR Runtime and XR Source Management/Presentation Engine and is defined to monitor.

[Editor’s Note: the applicability of these parameters is TBD

- Viewer pose

- Projection parameters

- Camera information

- Gesture

- Body action

- Tracking position prediction error

- Mapping latency for reconstructing the surrounding space]

Editor’s Note: Additional parameters to be monitored in OP1 are FFS.

### 6.2.2 Observation Point 2

Scene Manager is a set of functions that supports the application in arranging the logical and spatial representation of a multisensorial scene based on support from the XR Runtime. XR Scene Manager has access to the latest pose and tracking information from the XR Runtime which is then provided. Based on this information, the Scene Manager may for example determine the objects visible to the user at a given point in time or more generally the objects that may be needed to be rendered in the next rendering cycles.

Media Access Function is a set of functions that enables access to media and other XR-related data that is needed in the Scene manager or XR Runtime to provide an XR experience. The media access function accesses the network resources or sends data to the network using the established media pipelines.

Observation point 2 is derived from the API which exchanges information between Scene Manager and Media Access Function and is defined to monitor:

[Editor’s Note: the applicability of these parameters is TBD

- Scene update latency

- FOV

- Viewport

- Viewport error for rendered objects]

Editor’s Note: Additional parameters to be monitored in OP2 are FFS.

### 6.2.3 Observation Point 3

Observation point 3 is derived from the API which exchanges information between Media Access Function and 5G System and is defined to monitor:

[Editor’s note: the applicability of these parameters is TBD

- Media resolution

- Media codec

- Media decoding time

- Average throughput]

Editor’s Note: Additional parameters to be monitored in OP3 are FFS.

### 6.2.4 Observation Point 4

Observation point 4 is derived from the API which exchanges information between XR Source Management and Metrics Access Functions and is defined to monitor:

[Editor’s note: the applicability of these parameters is TBD

- Poses

- Sensor information]

Editor’s Note: whether observation point 4 shares the same information with the observation point 1 is FFS.

Editor’s Note: Additional parameters to be monitored in OP4 are FFS.

Editor’s Note:

* Identification of relevant XR QoE Metrics and their impacts on the user experience.
* Define the relevant observation points and define the measurement and derivation of relevant XR QoE metrics in the device architecture based on MeCAR.

# 7 QoE Metrics Subjective Assessment

Editor’s Note:

* Documentation of subjective tests results on XR QoE metrics, if considered relevant.

# 8 Other QoE metrics and collaborations with other 3GPP groups

Editor’s Note

* Collaboration with relevant groups or specifications on NWDAF, RRC-based metrics configuration and collection.

# 9 Conclusions and Recommendations

Editor’s Note

* Provide recommendation on normative work for new XR QoE metrics based on the findings in this study

Annex <A> (informative):  
<Normative annex for a Technical Specification>

Start each annex on a new page.

Annexes are labelled A, B, C, etc. and designated either "normative" or "informative" depending on their content.

Normative annexes only to appear in Technical Specifications. Use style "Heading 8".

Annex <X> (informative):  
Change history

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Change history** | | | | | | | |
| **Date** | **Meeting** | **TDoc** | **CR** | **Rev** | **Cat** | **Subject/Comment** | **New version** |
| 2022-08 | SA4#120e | S4-221164 |  |  |  | Initial version | V0.0.1 |
| 2022-08 | SA4#120e | S4-221165 |  |  |  | S4-221163: Collect information from ITU-T | V0.1.0 |
| 2022-11 | SA4#121 | S4-221564 |  |  |  | Add QoE-related activity in the external standardizations | V0.2.0 |
| 2023-02 | SA4#122 | S4-230294 |  |  |  | Add QoE information collection from MPEG, relevant VR metrics, basic architecture and observation points. | V0.3.0 |