International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

G.1072
(01/2020)

SERIES G: TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS

Multimedia Quality of Service and performance – Generic and user-related aspects

---

Opinion model predicting gaming quality of experience for cloud gaming services

Recommendation ITU-T G.1072

ITU-T

ITU-T G-SERIES RECOMMENDATIONS

TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS

INTERNATIONAL TELEPHONE CONNECTIONS AND CIRCUITS	G.100–G.199
GENERAL CHARACTERISTICS COMMON TO ALL ANALOGUE CARRIER-TRANSMISSION SYSTEMS	G.200–G.299
INDIVIDUAL CHARACTERISTICS OF INTERNATIONAL CARRIER TELEPHONE SYSTEMS ON METALLIC LINES	G.300–G.399
GENERAL CHARACTERISTICS OF INTERNATIONAL CARRIER TELEPHONE SYSTEMS ON RADIO-RELAY OR SATELLITE LINKS AND INTERCONNECTION WITH METALLIC LINES	G.400–G.449
COORDINATION OF RADIOTELEPHONY AND LINE TELEPHONY	G.450–G.499
TRANSMISSION MEDIA AND OPTICAL SYSTEMS CHARACTERISTICS	G.600–G.699
DIGITAL TERMINAL EQUIPMENTS	G.700–G.799
DIGITAL NETWORKS	G.800–G.899
DIGITAL SECTIONS AND DIGITAL LINE SYSTEM	G.900–G.999
MULTIMEDIA QUALITY OF SERVICE AND PERFORMANCE – GENERIC AND USER-RELATED ASPECTS	**G.1000–G.1999**
TRANSMISSION MEDIA CHARACTERISTICS	G.6000–G.6999
DATA OVER TRANSPORT – GENERIC ASPECTS	G.7000–G.7999
PACKET OVER TRANSPORT ASPECTS	G.8000–G.8999
ACCESS NETWORKS	G.9000–G.9999

For further details, please refer to the list of ITU-T Recommendations.

Recommendation ITU-T G.1072

Opinion model predicting gaming quality of experience for cloud gaming services

Summary

Recommendation ITU-T G.1072 presents an opinion model that predicts the overall gaming quality of experience (QoE) of non-expert gamers for cloud gaming services. The model uses an impairment factor approach in which the impact of typical Internet protocol (IP) network parameters and video encoding parameters on the video and input quality is estimated. The knowledge summarized in Recommendations ITU-T G.1032 and ITU-T P.809 serves as a basis for the development of the model. The model is a network planning tool which can be used by stakeholders to manage resource allocation and to configure IP-network transmission settings such as the selection of encoding framerates, resolutions and bitrates, under the assumption that the network is prone to packet loss and latency. Depending on whether the respective stakeholder has a priori knowledge of the type of game being offered through the cloud gaming service, either a default mode, which assumes the game to be highly sensitive towards delays and frame losses as well as having high encoding complexity, or an extended mode, which uses an adjusted model coeffiecient to increase the prediction accuracy, can be used.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T G.1072	2020-01-13	12	11.1002/1000/14151

Keywords

Cloud gaming, game, mean opinion score (MOS), modelling, quality of experience (QoE).

* To access the Recommendation, type the URL http://handle.itu.int/ in the address field of your web browser, followed by the Recommendation's unique ID. For example, http://handle.itu.int/11.1002/1000/11830-en.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure, e.g., interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had not received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database at http://www.itu.int/ITU-T/ipr/.

Table of Contents

Recommendation ITU-T G.1072

Opinion model predicting gaming quality of experience for cloud gaming services

1 Scope

This Recommendation describes a model that delivers predicted mean opinion scores (MOS) on a 5-point absolute category rating (ACR) scale, see [ITU-T P.800.1], [ITU-T P.910], based on the impact of impairments introduced by typical Internet protocol (IP) networks on the quality experienced by players using a cloud gaming system. This Recommendation targets cloud gaming services that perform video streaming over real-time transport protocol (RTP) (over user datagram protocol (UDP)) and which select various video encoding parameters to adapt to the network throughput, packet loss, and end-to-end delay.

The focus of the described model is to predict gaming quality of experience (QoE) by considering relevant factors that are identified and discussed in [ITU-T G.1032]. The impairment factors are derived based on network and encoding parameters. By analysing the suitability of a variety of quality features for the prediction of the overall gaming QoE, an impairment model inspired by the E-model [b-ITU-T G.107] was developed.

The model is a network planning tool which can be used by various stakeholders for purposes such as resource allocation and configuration of IP-network transmission settings such as the selection of resolution and bitrates, under the assumption that the network is prone to packet loss and latency.

The model offers two different modes: a default mode where no information about the game type is considered, and an extended mode, for which various impairment factors based on a content classification with respect to the encoding complexity as well as the delay and frame loss sensitivity of a game are considered. Depending on whether the respective stakeholder (cloud gaming service provider or the network planner) has a priori knowledge of the type of game being offered through the cloud gaming service, the appropriate mode can be used. More information on the modes is given in clause 6.2 and Annex A.

Virtual reality games requiring 3D rendering devices, mobile input, and output devices, as well as input devices other than keyboard and mouse are not within the scope of this model. Nevertheless, the model might apply to such systems as well. This is an item for further study. The model is also not designed to predict the influence of the game design or the motivation of users to play them. The subjective ratings collected to develop the model are primarily derived from non-expert gamers and hence, the model predictions might not be accurate for highly experienced gamers due to their different expectations and sensitivity towards the various degradations. Furthermore, the influence of social factors are not considered in this model. While the focus of the described model is on cloud gaming, some parts may also be relevant for online gaming (where the game is primarily executed on the client) or passive gaming video streaming (where only the video content is streamed to passive viewers of the game) applications. With respect to the technologies considered, the model addresses cloud gaming services using graphics processing unit (GPU) hardware accelerator engines for video compression, and H.264 [ITU-T H.264] as the video compression standard.

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T G.1032] Recommendation ITU-T G.1032 (2017), *Influence factors on gaming quality of experience.*

[ITU-T G.1071] Recommendation ITU-T G.1071 (2016), *Opinion model for network planning of video and audio streaming applications.*

[ITU-T H.264] Recommendation ITU-T H.264 (2019), *Advanced video coding for generic audiovisual services.*

[ITU-T P.800.1] Recommendation ITU-T P.800.1 (2016), *Mean opinion score (MOS) terminology*.

[ITU-T P.809] Recommendation ITU-T P.809 (2018), *Subjective evaluation methods for gaming quality*.

[ITU-T P.910] Recommendation ITU-T P.910 (2008), *Subjective video quality assessment methods for multimedia applications.*

[ITU-T P.1201.2] Recommendation ITU-T P.1201.2 (2012), *Parametric non-intrusive assessment of audiovisual media streaming quality – Higher resolution application area.*

[ITU-T P.1401] Recommendation ITU-T P.1401 (2020), *Methods, metrics, and procedures for statistical evaluation, qualification, and comparison of objective quality prediction models.*

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 cloud gaming [ITU-T G.1032]: Cloud gaming is characterized by game content delivered from a server to a client as a video stream with game controls sent from the client to the server. The execution of the game logic, rendering of the virtual scene, and video encoding is performed at the server, while the client is responsible for video decoding and capturing of client input.

3.1.2 game [b-Juul]: A game is a rule-based system with a variable and quantifiable outcome, where different outcomes are assigned different values, the player exerts effort in order to influence the outcome, the player feels emotionally attached to the outcome, and the consequences of the activity are optional and negotiable.

3.1.3 model, model algorithm [b-ITU-T P.1203]: An algorithm with the purpose of estimating the subjective (perceived) quality of a media sequence.

3.1.4 slicing artefacts [ITU-T G.1071]: Artefacts that are introduced when packet losses are concealed through the use of a packet loss concealment (PLC) scheme to repair erroneous frames.

3.1.5 freezing artefacts [ITU-T G.1071]: Artefacts that are introduced when the packet loss concealment (PLC) scheme of the receiver replaces the erroneous frames (either due to packet loss or error propagation) with the previous error-free frame until a decoded picture without errors have been received. Since the erroneous frames are not displayed, this type of artefact is also referred to as freezing with skipping.

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.2.1 input quality: Describes the playability of a game scenario in terms of the components responsiveness, immediate feedback, and controllability.

3.2.2 spatial video quality: Spatial component of video quality composed of the dimensions unclearness and fragmentation as presented in [b-Schiffner], which can be influenced by artefacts such as blockiness and blur.

3.2.3 temporal video quality: Temporal component of video quality represented by the dimensions discontinuity as presented in [b-Schiffner], which can be influenced by artefacts such as freezing and jerkiness.

3.2.4 average frames per second: The average number of successfully transmitted frames per second during a video stream.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

3D	Three-dimensional space
ACR	Absolute Category Rating
Avg_FPS	Average Frames Per Second
FEC	Forward Error Correction
FR_enc	Encoding Framerate
GoP	Group of Pictures
GPU	Graphics Processing Unit
HEVC	High Efficiency Video Coding
IP	Internet Protocol
IPQ	Input Quality
llhq	Low Latency, High Quality
MOS	Mean Opinion Score
PLC	Packet Loss Concealment
PLCC	Pearson Linear Correlation Coefficient
QoE	Quality of Experience
RMSE	Root-Mean-Square Error
RTP	Real-time Transport Protocol
SVQ	Spatial Video Quality
TVQ	Temporal Video Quality
UDP	User Datagram Protocol

5 Conventions

Within the scope of this Recommendation, the person interacting with a game is referred to as the player, whereas the software, which is used in a cloud gaming set-up to display a remotely rendered game video stream, is referred to as the client. Furthermore, when referring to a delay, the roundtrip time for transmitting a video and command stream is considered, excluding additional delays such as system tick rate, rendering delay, encoding/decoding time, refresh rate of the display, and any other processing between client and servers.

6 Areas of application

6.1 Application range for the model

The described model was developed based on a dataset of subjective assessment studies using the interactive and passive viewing-and-listening paradigms proposed in [ITU-T P.809]. The passive viewing-and-listening paradigm was used to cover a broader range of video encoding parameters and games. It must be noted that packet loss in the conducted interactive tests resulted in freezing artefacts, while in the passive viewing-and-listening tests, slicing artefacts occured. The factors and application range summarized in Table 1 are considered.

Table 1 – Factors and application ranges of the model

Application information	Interactive	Passive viewing-and-listening
	Value range, unit	
Sequence duration	90 seconds	30 seconds
Screen size	24"	24"
Input devices	Mouse and keyboard	–
Packetization	RTSP (over RTP/UDP/IP)	RTSP (over RTP/UDP/IP)
Video codec	H.264 using NVENC	H.264 using NVENC
Resolution	1080p	720p, 1080p
Coded video bitrate (mbps)	1-50	0.3-50
Frame rate (fps)	10, 20, 30, 60	20, 30, 60
Group of pictures (*Note 1*)	Infinite	Infinite and 0.5, 1, 2 seconds
Pre-set	llhq (low latency, high quality)	llhq (low latency, high quality)
Encoding mode	CBR	CBR
Video compression	Standard H.264, Main 4.0	Standard H.264, Main 4.0
Audio codec	AC3	–
Coded audio bitrate (kbps)	192 (stereo)	–
Audio sample rate (Hz)	48,000	–
Packet loss degradation (Note 2)	uniform loss (0-5%)	uniform loss (0-2%)
Delay range	0-400 ms	–

NOTE 1 – The low latency, high quality (llhq) preset does not use B frames and it uses an infinite GoP length by default. In case of a corrupted frame, no spatial artefacts will be visible but instead the FEC will lead to a replacement of the corrupted frame resulting in jerkiness of the video (freezing artefacts). However, for conducted passive tests, we also considered different GoP sizes, especially when considering different packet loss scenarios.

NOTE 2 – The assumption of a uniform probability of packet loss is a hypothetical one which does not reflect the situation in real-life networks, where losses would typically occur in bursts. The modelling of bursty packet loss is an item for further study.

The following factors were not considered during model design:

– Audio/video sync distortions

– Different levels of audio quality

– Packet loss and delay distribution
– Video codecs for which the model is not validated (MPEG2, HEVC, VP9, AV1, etc.)
– Transcoding solutions
– The effects of noise and color correctness in a video
– Different ranges of parameters than the ones tested, e.g. delay of 800 ms, display size lower/higher than 24".

6.2 Modes of operation

The impact of network and encoding distortions on various quality features perceived by a player depends strongly on the sensitivity of a game towards these degradations. In order to reach a higher accuracy of the model, two modes of operation are defined depending on whether a network planner or service provider can make any assumption on the type of game that is targeted or not. If the network planner or cloud gaming provider has knowledge of the type of the targeted game with respect to its encoding complexity and sensitivity towards delay and frame losses, the quality of experience (QoE) can be predicted more accurately by considering the content classes. If the network planner cannot make any assumption on the game type, the highest content class will be assumed, which is referred to as the default mode in the rest of this document. Therefore, in the default mode, the model will result in a pessimistic quality prediction for games that are not of high complexity and sensitivity. Considerations on how to quantify game characteristics to derive the classes considered for the model are given in Annex A.

7 Gaming QoE prediction

7.1 Model structure

The model structure, as illustrated in Figure 1, is composed of two main modules, namely input quality (*IPQ*) and video quality (*VQ*).

The model considers two types of input parameters: network and encoding parameters. For the network parameters, delay and packet loss (*PL*) are used. For the encoding parameters the video resolution (*Res*), encoding framerate (*FR_enc*), and the bitrate (*Br*) used for the video stream are used.

Furthermore, five different estimations of quality impairments expressed on the R-scale, namely *I_VQ_cod*, *I_VQ_trans*, *I_TVQ*, *I_IPQ_frames*, and *I_IPQ_delay*, are calculated based on the previously mentioned input parameters. Their calculation is described in the following subsections.

To predict the overall gaming QoE (*MOS_QoE*), the estimated quality impairments are weighted with the coefficients *a*, *b*, *c*, *d*, and *e*, which depend on the game type. Next, their sum is subtracted from a reference value, *R_max*, resulting in *R_QoE*. Finally, the predicted *MOS_QoE* is calculated using a conversion to the MOS-scale.

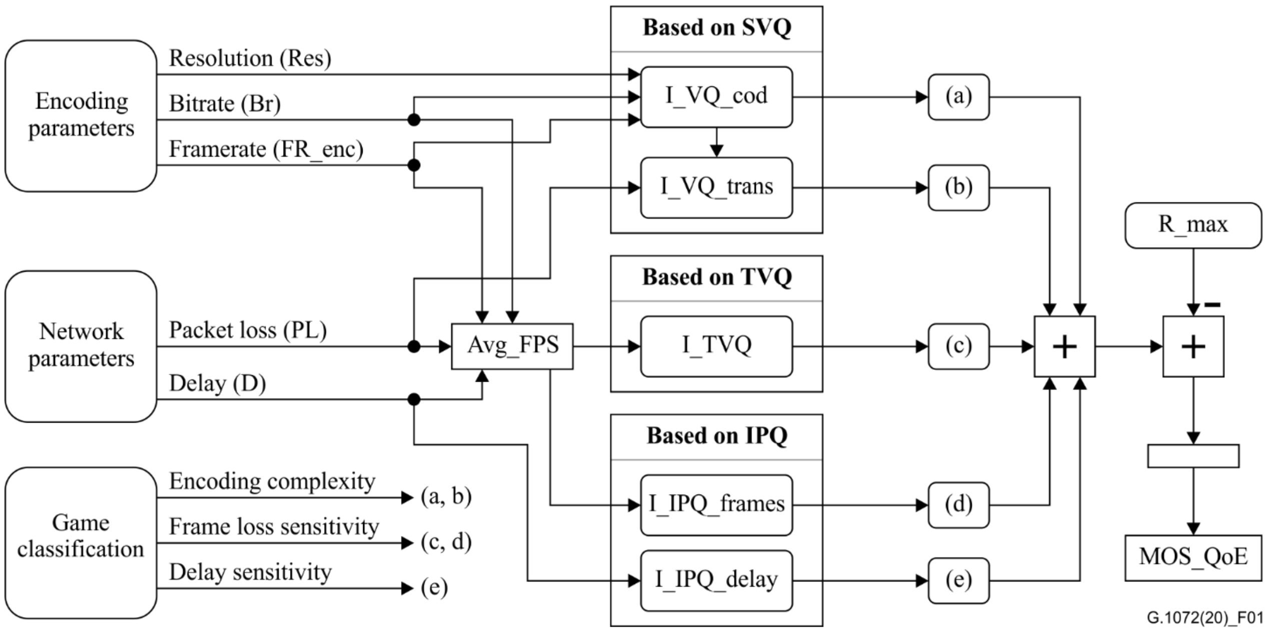

Figure 1 – Model structure

7.2 Core model

The core model predicting gaming QoE is defined as:

$$R_{QoE} = R_{max} - a \cdot I_{VQ_{cod}} - b \cdot I_{VQ_{trans}} - c \cdot I_{TVQ} - d \cdot I_{IPQ_{frames}} - e \cdot I_{IPQ_{delay}} \quad (1)$$

$$MOS_{QoE} = \text{MOS\_from\_R} \left(R_{QoE} \right) \quad (2)$$

where,

R_{QoE} is the overall estimated gaming QoE expressed on the R-scale, where 0 is the worst quality and 100 the best quality;

MOS_{QoE} is the overall estimated gaming QoE expressed on the MOS-scale, where 1 is the worst quality and 5 is the best quality;

R_{max} is the reference value indicating the best possible gaming QoE (= 100) on the R-scale;

$I_{VQ_{cod}}$ is the estimated spatial video quality impairment for video compression artefacts on the R-scale;

$I_{VQ_{trans}}$ is the estimated spatial video quality impairment for video transmission errors on the R-scale;

I_{TVQ} is the estimated temporal video quality impairment for frame rate reductions on the R-scale;

$I_{IPQ_{frames}}$ is the estimated input quality impairment for frame rate reductions on the R-scale; and

$I_{IPQ_{delay}}$ is the estimated input quality impairment for network delay degradations on the R-scale;

with the constant coefficients *a*, *b*, *c*, *d*, and *e* values summarized in Table 2.

Table 2 – Weighting factors of the final model

Coefficient	*a*	*b*	*c*	*d*	*e*
Value	0.788	0.896	0.227	0.625	0.848

MOS_max, the MOS value corresponding to R_max, is 4.64, whereas the lowest possible MOS value, *MOS_min*, is 1.3. The following transformation from [ITU-T P.1201.2] is used to derive *MOS_QoE*:

```
function MOS_QoE = MOS_from_R(R_QoE)

set MOS_MAX = 4.64;
set MOS_MIN = 1.3;

if (R_QoE > 0 & R_QoE < 100),
  MOS_QoE = (1+(MOS_MAX-MOS_MIN)/100×R_QoE+R_QoE×(R_QoE-60)×(100-R_QoE)×7.0E-6);
elseif (R_QoE >= 100),
  MOS_QoE = MOS_MAX;
else
  MOS_QoE = MOS_MIN;
end
```

8 Calculation of quality impairment estimations

This clause describes the five estimations of quality impairments.

8.1 Spatial video quality modeling

The spatial video quality is modelled based on [ITU-T G.1071]. Since different encoding settings, as well as datasets, are used here, the model proposed in [ITU-T G.1071] is retrained.

8.1.1 Video quality impairment factor (I_VQ_cod) due to video compression artefacts

For the calculation of the impairment factor *I_VQ_Cod*, the parameters bitrate (*BR*), resolution (*Res*), and encoding framerate (*FR_enc)* are used, where the number of pixels per video frame (*NumPixelPerFrame*) is calculated as the product of height and width of the video, e.g., 1920 × 1080 for 1080p resolution). It must be noted that the *ContentComplexity* here is not the same as the proposed encoding complexity classification, which is used for the model coefficients for each class separately. Based on models proposed in [ITU-T G.1071] and [ITU-T P.1201.2], *I_VQ_Cod* is defined as:

$$I_{VQ_{cod}} = a_{1V} \cdot \exp(a_{2V} \cdot BitPerPixel) + a_{3V} \cdot ContentComplexity + a_{4V} \quad (3)$$

where a_{1V}, a_{2V}, a_{3V} and a_{4V} are the constant coefficients and

$$ContentComplexity = a_{31} \cdot \exp(a_{32} \cdot BitPerPixel) + a_{33} \quad (4)$$

$$BitPerPixel = \frac{Bitrate \cdot 10^6}{NumPixelPerFrame \cdot Framerate} \quad (5)$$

where, $\boldsymbol{a_{31}, a_{32}\ and\ a_{33}}$ are constant coefficients.

8.1.2 Video quality impairment factor (I_VQ_trans) due to video transmission errors

For the calculation of the impairment factor *I_VQ_Trans*, the parameters packet loss (*PL*) and results from *I_VQ_Cod* are used, following the equations below which are based on the models proposed in [ITU-T G.1071] and [ITU-T P.1201.2].

$$I_{VQ_{trans}} = c_{1V} \cdot \log(c_{2V} \cdot LossMagnitudeE + 1) \quad (6)$$

$$LossMagnitudeE = q_1 \cdot \exp(q_2 \cdot LossMagnitudeNP) - q_1 \quad (7)$$

$$LossMagnitudeNP = \frac{(c_{21} - I_{codn}) \cdot PL}{c_{23} \cdot I_{codn} + PL} \quad (8)$$

$$I_{codn} = \begin{cases} I_{VQ_{cod}}, & if\ I_{VQ_{cod}} \leq 65 \\ 65, & else \end{cases} \quad (9)$$

where, c_{1V}, c_{2V}, $q_1, q_2, c_{21}, and\ c_{23}$ are the coefficients derived based on the training dataset. The list of coefficients for each class of encoding complexity is given in Table 3. It must be noted that packet loss (PL) in Eqn. 8 should be zero if the packet loss concealment (PLC) scheme is causing freezing artefacts.

Table 3 – Coefficients of retrained G.1071 impairments for each encoding complexity in which classes 1, 2 and 3 represent low, medium and high complexity classes respectively

Coefficient	Class 1	Class 2	Class 3 (default mode)
a_{1V}	52.5052	37.9882	47.7463
a_{2V}	–28.017	–13.7208	–12.07
a_{3V}	–2.68405	8.57837	9.05168
a_{4V}	5.46648	3.26581	3.41919
a_{31}	12.4214	6.83276	7.62306
a_{32}	–28.0192	–127.997	–167.838
a_{33}	0.215799	0.479595	0.0760333
c_{1V}	19.7092	0.612879	1.57176
c_{2V}	3358.31	0.00139396	3.68596
c_{21}	28.3699	56.2893	74.0571
c_{23}	0.0234973	0.0047567	0.00406
q_1	0.0016474	0.0581327	2.58892e-08
q_2	0.0895914	2.38014	0.868407

8.2 Temporal video quality impairment factor (I_TVQ) due to frame rate reduction(s)

For the impairment factor *I_TVQ*, the parameters *Avg_FPS* and *FR_enc* are used. The following equation for *I_TVQ* is obtained.

$$I_{TVQ} = d_1 + d_2 \cdot {FR_{enc}}^2 + d_3 \cdot FR\_enc + d_4 \cdot Log(FrameLossRate) \quad (10)$$

where,

$$FrameLossRate = \frac{FR_{enc} - Avg_{FPS}}{FR_{enc}} \cdot 100 \quad (11)$$

where, if Delay < 16 ms, Avg_FPS = FR_enc, else:

$$\mathrm{Avg}_{\mathrm{FPS}} = FR\_enc \cdot \exp(-(d_5 + d_6 \cdot FR_{enc} + d_7 \cdot Bitrate \cdot FR_{enc}) \cdot (d_8 \cdot Delay - d_9) \cdot PL) \quad (12)$$

where, $d_5 = 0.08526$, $d_6 = 0.00073$, $d_7 = 1.425\text{e-}07$, $d_8 = 0.09656$, $d_9 = 1.5$. It must be noted that packet loss (PL) in Eqn. 12 should be zero if the packet loss concealment (PLC) scheme is causing slicing artefacts. The constant coefficients $d_1, d_2, d_3, and\ d_4$ with respective values for each frame loss sensitivity class mentioned in Table 4.

Table 4 – Coefficients of I_TVQ for each frame loss sensitivity class

Coefficient	Low sensitive	High sensitive (default mode)
d_1	29.13	47.03
d_2	0.01344	0.01747
d_3	–1.283	–1.823
d_4	6.724	10.7

8.3 Input quality modeling

8.3.1 Input quality impairment factor (I_IPQ_frames) due to frame rate reduction(s)

For the calculation of the impairment factor *I_IPQ_frames*, the parameters encoding framerate (*FR_enc*) and the average frames per second (*Avg_FPS*), see Eqn. 12, are used. The coefficients used are given for each frame loss sensitivity class in Table 5. The impairment is modeled as:

$$I_{IPQ_{Frames}} = e_1 + e_2 \cdot FrameRate^2 + e_3 \cdot FrameRate + e_4 \cdot Log(FrameLossRate) \quad (13)$$

Table 5 – Coefficients of I_IPQ_frames for each frame loss sensitivity class

Coefficient	Low sensitive	High sensitive (default mode)
e_1	23.43	54.71
e_2	0.008574	0.02589
e_3	–0.9253	–2.485
e_4	5.855	9.306

8.3.2 Input quality impairment factor (I_IPQ_delay) due to network delay degradations

For the calculation of the impairment factor *I_IPQ_delay*, the parameter network delay (*Delay*) is used. The impairment is modeled as:

$$I_{IPQ_{Delay}} = \frac{f_1}{1 + \exp(f_2 - f_3 \cdot Delay)} + f_4 \quad (14)$$

The coefficients obtained for each delay sensitivity class are summarized in Table 6.

Table 6 – Coefficients of I_IPQ_delay for each delay sensitivity class

Coefficient	Low sensitive	High sensitive (default mode)
f_1	47.97	90
f_2	2.097	1.191
f_3	0.01073	0.009775
f_4	–4.567	–18.73

9 Performance of the model

On the evaluation dataset, unknown to the model, the performance of the model is reported in Table 7 in terms of root-mean-square error (RMSE) and Pearson linear correlation coefficient

(PLCC), see [ITU-T P.1401]. The prediction of the gaming QoE in comparison with the subjective ratings is shown as a scatter plot in Figure 2 for the default mode, and for the extended mode. It must be noted that due to the same range of parameters used in the training and test set, the result could be optimistic.

Table 7 – Final prediction performance of the model on the test set

	Considering classification (extended mode)		Without considering classification (default mode)	
	R-scale	MOS-scale	R-scale	MOS-scale
RMSE	8.03	0.33	12.19	0.47
PLCC	0.89	0.90	0.80	0.82

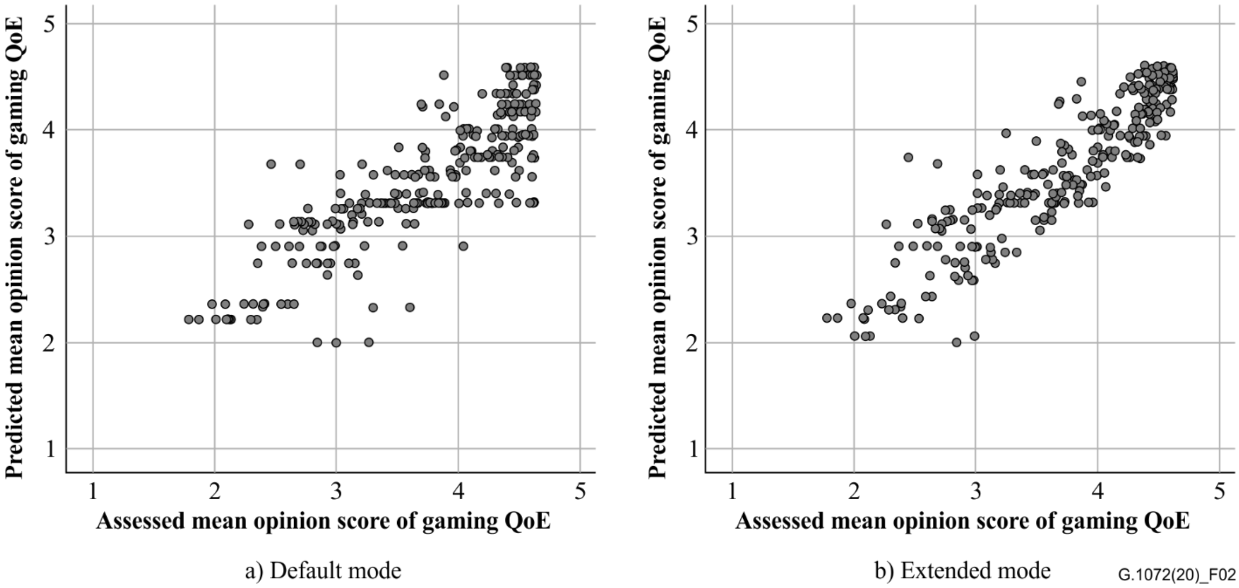

Figure 2 – Scatter plot of the predicted and assessed MOS ratings on the test dataset without using content classification (default mode) on the left, and using the content classification (extended mode) on the right

It can be observed that the extended mode due to content information, results in a higher accuracy than the default mode where no content information is used. Nevertheless, the model results in a good performance for both modes.

Annex A

Description of the modes used for the model considering content classes

(This annex forms an integral part of this Recommendation.)

The model described in this Recommendation has two modes, a default mode and an extended mode, depending on the game type. The game type is determined by encoding complexity, and sensitivity towards delay and frame losses.

Encoding complexity is determined, amongst others, by:

- Movements of a virtual camera
- Texture details
- Frequency of movements of game objects

Sensitivity towards delay is determined, amongst others, by:

- Continuous (e.g., mouse) or discrete type of input (e.g., keyboard)
- Number of possible input directions in the virtual scene
- Minimum number of required actions
- Available time interval for a player to perform a desired interaction
- Predictability of game events

Sensitivity towards frame losses is determined, amongst others, by:

- Movements of a virtual camera
- Frequency of game object movements
- Pace of the interaction with the game

As long as no information on these characteristics is available to the user of the model, the default mode should be used. In case that it is known that the model shall be applied to games of low or medium encoding complexity, the "low complexity" or "medium complexity" class regarding this aspect can be used instead of the default mode, see Table 3. In case that it is known that the model shall be applied to games with low sensitivity towards frame losses, the "low sensitivity" class regarding this aspect should be used instead of the default mode, see Tables 4 and 5. In case that it is known that the model shall be applied to games with low sensitivity towards delay, the "low sensitivity" class regarding this aspect can be used instead of the default mode, see Table 6.

The classification of game types according to the cited characteristics is currently left at the discretion of the model user, and examples can be found in [b-Zadtootaghaj]. The development of a method for accurately classifying games into these classes is an object of further study by Study Group 12 of ITU-T.

Bibliography

[b-ITU-T G.107] Recommendation ITU-T G.107 (2015), *The E-model: a computational model for use in transmission planning*.

[b-ITU-T P.1203] Recommendation ITU-T P.1203 (2017), *Parametric bitstream-based quality assessment of progressive download and adaptive audiovisual streaming services over reliable transport.*

[b-Juul] Juul, J. (2005), Video games between real rules and fictional worlds, *Video Games Real Rules Fict. Worlds*.

[b-Schiffner] Schiffner, F. and Möller, S. (2018), Direct Scaling & Quality Prediction for perceptual Video Quality Dimensions, *Tenth International Conference on Quality of Multimedia Experience (QoMEX)*, pp. 1-3, IEEE.

[b-Zadtootaghaj] Zadtootaghaj, S., Schmidt, S., Barman, N., Möller, S., and Martini, M. G. (2018), A classification of video games based on game characteristics linked to video coding complexity, *16th Annual Workshop on Network and Systems Support for Games (NetGames)*, pp. 1-6.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	Tariff and accounting principles and international telecommunication/ICT economic and policy issues
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	**Transmission systems and media, digital systems and networks**
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Environment and ICTs, climate change, e-waste, energy efficiency; construction, installation and protection of cables and other elements of outside plant
Series M	Telecommunication management, including TMN and network maintenance
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling, and associated measurements and tests
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks, open system communications and security
Series Y	Global information infrastructure, Internet protocol aspects, next-generation networks, Internet of Things and smart cities
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2020