ITUPublications

International Telecommunication Union
Standardization Sector

ITU-T Technical Report

(01/2025)

PSTR-CMVTQS2

Alternative parametric computational model used as a quality monitor to assess videotelephony services

Technical Report ITU-T PSTR-CMVTQS2

Alternative parametric computational model used as a quality monitor to assess videotelephony services

Summary

This Technical Report introduces a set of parametric blocks composing an alternative model for quality assessment of videotelephony services to Recommendation ITU-T P.940. This model uses the same building block for audio quality assessing block and audiovisual quality assessing block as in clause 7 of Recommendation ITU-T P.940. Therefore, this Technical Report only describes other parts of the model: the video quality assessing block, audiovisual interaction delay assessing block, audiovisual media synchronization assessing block and videotelephony quality assessing block. Clauses of this report address the description of the various building blocks of the model and their formulae, as well as the performance of the model on the databases used for the validation of Recommendation ITU-T P.940.

Keywords

QoE, QoS, videotelephony.

Note

This is an informative ITU-T publication. Mandatory provisions, such as those found in ITU-T Recommendations, are outside the scope of this publication. This publication should only be referenced bibliographically in ITU-T Recommendations.

Change Log

This document contains Version 1 of the ITU-T Technical Report PSTR-CMVTQS2 on "*Alternative parametric computational model used as a quality monitor to assess videotelephony services*" approved at the ITU-T Study Group 12 meeting held in Geneva, 14-23 January 2025.

Editor: Mengying Liu
China Mobile Communication Co. Ltd.
China
Email: liumengying@chinamobile.com

Jose Joskowicz
Universidad de la República
Uruguay
Email: josej@fing.edu.uy

Rafael Sotelo
Universidad de Montevideo
Uruguay
Email: rsotelo@ieee.org

Shijun Zhang
China Mobile Communication Co. Ltd.
China
Email: zhangshijun@chinamobile.com

Lei Yang
China Mobile Communication Co. Ltd.
China
Email: yangleiyj@chinamobile.com

Table of Contents

Technical Report ITU-T PSTR-CMVTQS2

Alternative parametric computational model used as a quality monitor to assess videotelephony services

1 Video quality assessing block

The video quality is estimated by considering video coding factors, terminal capabilities, and network transmission conditions. The video quality is expressed as:

$$Q_v = 1 + 4I_cI_t \tag{1}$$

where I_c represents the encoding video quality and I_t represents the transmission video quality factor due to network impairments. These components are detailed in the following sections. Q_v is bounded between 1 and 5.

1.1 Encoding video quality

The encoding video quality I_c represents the video quality achievable under ideal network conditions, without being influenced by any degradation related to transmission. It mainly depends on the frame rate (Fr_v) and the video bit rate (Br_v), and is calculated as follows:

$$I_c = f_{Fr} \cdot \left(1 - \frac{1}{1+\left(\frac{f_{res} \cdot Br_v}{c_1}\right)^{c_2}}\right) \tag{2}$$

The factor f_{res} is related to the video and screen resolution and can be calculated as:

$$f_{res} = c_3(rh \cdot rw)^{c_4} \tag{3}$$

where rh and rw are the maximum values between the video and screen resolution (height and width, respectively).

Additionally, the factor f_{Fr} influences the perceived video quality as the frame rate varies. It is calculated as follows:

$$f_{Fr} = (1 - e^{c_5 Fr_v})\left(1 + (60 - Fr_v)\left(c_6 + c_7 e^{c_8 \cdot f_{res} \cdot Br_v \cdot (60 - Fr_v)}\right)\right) \tag{4}$$

where 60 fps represents the maximum frame rate considered.

1.2 Transmission video quality

The transmission video quality factor I_t accounts for the degradation introduced by the transmission network, primarily the video packet loss rate (Plr_v). It is calculated as:

$$I_t = e^{-\frac{Plr_v}{c_9}} \tag{5}$$

2 Audiovisual interaction delay assessing block

The audiovisual interaction delay quality (Q_{delay}) reflects the user experience of interaction delay during video calls. It is expressed as:

$$Q_{delay} = (w_1 - w_1 e^{w_2 Q_{av}}) \cdot f_{delay} + w_3 \tag{6}$$

where:

$$f_{delay} = 1 - e^{\frac{w_4}{\sqrt{T_a^2 + T_v^2}}} \tag{7}$$

Q_{delay} is bounded between 1 and 5.

3 Audiovisual media synchronization assessing block

The audiovisual media synchronization quality (Q_{sync}) reflects the user experience of media synchronism during video calls. [b-ITU-R BT.1359] and [b-Saidi] showed that users have different levels of sensitivity to the changes in audiovisual media synchronization when audio is ahead of video and when video is ahead of audio. Therefore, the audiovisual media synchronization quality (Q_{sync}) is expressed as:

$$Q_{sync} = (w_5 - w_5 e^{w_6 Q_{av}}) \cdot f_{sync} + w_7 \tag{8}$$

where:

$$f_{sync} = \frac{1}{1+(w_8(T_v - T_a))^{w_9}} \quad \text{if } T_v > T_a \tag{9}$$

$$f_{sync} = \frac{1}{1+(w_{10}(T_a - T_v))^{w_{11}}} \quad \text{if } T_v \leq T_a \tag{10}$$

Q_{sync} is bounded between 1 and 5.

4 Videotelephony quality assessing block

The videotelephony quality (Q_{vt}) is the final videotelephony score that represents the overall quality during video calls. It comprehensively considers the audiovisual quality, the impact of audiovisual interaction delay and media synchronization. The videotelephony quality is expressed as:

$$Q_{vt} = (n_1 - n_1 e^{n_2 Q_{av}})\left(n_3 \cdot f_{delay} + (1 - n_3) \cdot f_{sync}\right) \tag{11}$$

Q_{vt} is bounded between 1 and 5.

5 Coefficients in the non-normative parametric models

This clause provides the coefficients to be used for the parametric models for quality of experience (QoE) / quality of service (QoS) assessment of videotelephony services. Table 1 summarizes the conditions and corresponding coefficients for calculating the video quality.

Table 1 – Conditions and coefficients for calculating the video quality

Coefficients	Mobile phone		PC		TV	
	H.264 (baseline)	**H.265 (main)**	**H.264 (baseline)**	**H.265 (main)**	**H.264 (baseline)**	**H.265 (main)**
c_1	1.3858×10^{-3}	1.2015×10^{-4}	5.1880	2.4674	2.3744×10^{-3}	2.1431
c_2	1.2048	0.8816	1.11631	0.7731	1.1096	0.5869
c_3	15.9693	10.4425	7.1162	4.1372	14.4589	14.8975
c_4	−1.1194	−1.2118	−0.5449	−0.4567	−1.0590	−0.5240
c_5	−0.2191	−0.1604	−1.1571	−0.1617	−1.8098	−0.1257
c_6	-2.5017×10^{-3}	-3.7178×10^{-3}	-1.7913×10^{-4}	4.30×10^{-5}	-3.4699×10^{-3}	6.6041×10^{-4}
c_7	1.6652×10^{-2}	5.9589×10^{-3}	6.1047×10^{-2}	4.5546×10^{-4}	5.0390×10^{-2}	2.99×10^{-15}
c_8	−11.6690	−11.7717	-8.7327×10^{-3}	−5.9106	−16.1914	−17.4160
c_9	1.0905	1.0905	1.0905	1.0905	1.0905	1.0905

NOTE 1 – The values for video displayed on mobile phone and PC have been obtained for cases where video resolution is QVGA/VGA/720p/1080p/4K and video codec is H.264 (baseline), and for cases where video resolution is VGA/720p/1080p/4K and video codec is H.265 (main).

NOTE 2 – The values for video displayed on TV have been obtained for cases where video resolution is 720p/1080p and video codec is H.264 (baseline), and for cases where video resolution is 720p/1080p/4K and video codec is H.265 (main).

Table 2 provides the coefficients to be used for audiovisual interaction delay quality estimation and audiovisual media synchronization quality estimation.

Table 2 – Coefficients for audiovisual interaction delay and audiovisual media synchronization assessing blocks

Coefficients	values
w_1	9.5983×10^{3}
w_2	-1.0090×10^{-4}
w_3	0.9828
w_4	-1.2230×10^{3}
w_5	8.8051×10^{3}
w_6	-1.3654×10^{-4}
w_7	0.1336
w_8	1.5544×10^{-3}
w_9	9.0791
w_{10}	1.1352×10^{-3}
w_{11}	2.6180

NOTE 3 – These coefficients for estimation of audiovisual media synchronization quality have been obtained for a delay difference smaller than or equal to 500 ms between audio and video.

NOTE 4 – These coefficient values for estimation of audiovisual interaction delay quality have been obtained for audio and video delays smaller than or equal to 1 000 ms. These values should be used within the specific range.

Table 3 summarizes the coefficients to be used for the final videotelephony quality assessing block to estimate the overall quality.

Table 3 – Coefficients for videotelephony quality assessing block

Coefficients	Values
n_1	9.4571
n_2	−0.1659
n_3	0.5096

6 Performance figures

The performance of the model on the databases described in Appendix II of [b-ITU-T P.940] is summarized in Table 4 below.

Table 4 – Performance of non-normative parametric model

Quality evaluation perspective	PLCC	RMSE
Video quality (with both passive and active test data)	0.859	0.476
Video quality (with active test data)	0.935	0.667
Audio quality	0.915	0.785
Audiovisual interaction delay quality	0.922	0.426
Audiovisual media synchronization quality	0.955	0.351
Final videotelephony quality	0.951	0.337

NOTE 1 – The performance figures were calculated after final training on all available subjective test databases.

NOTE 2 – The performance figures of video quality assessing block were calculated using both passive and active test data, and only active test data, separately.

NOTE 3 – The impact of delay is considered separately in the audiovisual interaction delay assessing block and the audiovisual media synchronization assessing block. Therefore, the validation of the video quality assessing block is conducted based on subjective test data with video delay smaller than 100 ms, and the validation of the audio quality assessing block is conducted based on subjective test data with audio delay smaller than 100 ms.

Bibliography

[b-ITU-T P.940] Recommendation ITU-T P.940 (2025), *Computational model used for the monitoring and quality assessment of videotelephony services.*

[b-ITU-R BT.1359] Recommendation ITU-R BT.1359 (1998), *Relative timing of sound and vision for broadcasting*.

[b-Saidi] Saidi, I., Zhang, L., Barriac, V., and Deforges, O. (2016), *Interactive vs. non-interactive subjective evaluation of ip network impairments on audiovisual quality in videoconferencing context.* Eighth International Conference on Quality of Multimedia Experience (QoMEX).