

Recommendation

ITU-T P.940 (03/2025)

SERIES P: Telephone transmission quality, telephone installations, local line networks

Audiovisual quality in multimedia services

Computational model used for the monitoring and quality assessment of videotelephony services

ITU-T P-SERIES RECOMMENDATIONS

Telephone transmission quality, telephone installations, local line networks

Vocabulary and effects of transmission parameters on customer opinion of transmission quality	P.10-P.19
Voice terminal characteristics	P.30-P.39
Reference systems	P.40-P.49
Objective measuring apparatus	P.50-P.59
Objective electro-acoustical measurements	P.60-P.69
Measurements related to speech loudness	P.70-P.79
Methods for objective and subjective assessment of speech quality	P.80-P.89
Voice terminal characteristics	P.300-P.399
Objective measuring apparatus	P.500-P.599
Measurements related to speech loudness	P.700-P.709
Methods for objective and subjective assessment of speech and video quality	P.800-P.899
Audiovisual quality in multimedia services	P.900-P.999
Transmission performance and QoS aspects of IP end-points	P.1000-P.1099
Communications involving vehicles	P.1100-P.1199
Models and tools for quality assessment of streamed media	P.1200-P.1299
Telemeeting assessment	P.1300-P.1399
Statistical analysis, evaluation and reporting guidelines of quality measurements	P.1400-P.1499
Methods for objective and subjective assessment of quality of services other than speech and video	P.1500-P.1599

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

Recommendation ITU-T P.940

Computational model used for the monitoring and quality assessment of videotelephony services

Summary

Recommendation ITU-T P.940 proposes a computational model that can be used for assessing the combined effects of network, media stream and terminal device related parameters on perceived quality. This model can be used for the supervision of videotelephony services and can also be used by videotelephony service providers to self-test to improve the quality of experience (QoE) and quality of service (QoS). This model provides estimates of multimedia quality, interaction experience and comprehensive videotelephony quality perceived by users.

The application of this Recommendation is limited to QoE/QoS monitoring. Other applications such as quality benchmarking and QoE/QoS planning are outside the scope of this Recommendation.

History*

Edition	Recommendation	Approval	Study Group	Unique ID
1.0	ITU-T P.940	2025-03-01	12	11.1002/1000/16210

Keywords

Audiovisual quality, conversational quality, monitor model, QoE, QoS, videotelephony.

* To access the Recommendation, type the URL <https://handle.itu.int/> in the address field of your web browser, followed by the Recommendation's unique ID.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, and 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/software copyrights, 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 appropriate ITU-T databases available via the ITU-T website at <https://www.itu.int/ITU-T/ipr/>.

© ITU 2025

All rights reserved. No part of this publication may be reproduced, by any means whatsoever, without the prior written permission of ITU.

Table of Contents

		Page
1	Scope	1
2	References.....	3
3	Definitions	3
	3.1 Terms defined elsewhere	3
	3.2 Terms defined in this Recommendation.....	4
4	Abbreviations and acronyms	4
5	Conventions	4
6	Overview of the model	5
	6.1 Model inputs	6
	6.2 Model outputs	6
7	Description of the computational model.....	6
	7.1 Audio quality assessing block	7
	7.2 Video quality assessing block	8
	7.3 Audiovisual quality assessing block.....	9
	7.4 Audiovisual interaction delay assessing block.....	9
	7.5 Audiovisual media synchronization assessing block	10
	7.6 Videotelephony quality assessing block.....	10
8	Performance figures.....	10
	Annex A – Coefficients in the parametric computational models	12
	A.1 Coefficients in the video quality assessing block.....	12
	A.2 Coefficients in the audiovisual quality assessing block	13
	A.3 Coefficients in the audiovisual interaction delay and audiovisual media synchronization assessing blocks	13
	A.4 Coefficients in the videotelephony quality assessing block	14
	Appendix I – Equipment impairment factor parameters in the audio quality assessing block	15
	Appendix II – Overview of databases used for model development	16
	Bibliography.....	18

Recommendation ITU-T P.940

Computational model used for the monitoring and quality assessment of videotelephony services

1 Scope

The model described in this Recommendation consists of a set of objective quality assessment modules that predict the quality of single-channel bidirectional videotelephony calls comprising both audio and video components. The model described in this Recommendation does not assess anything other than the audio and video components (i.e., it does not cover document sharing, chat, augmented reality). The focus is on pure audiovisual communications and on one single dimension of quality: the quality experienced by the end user (availability and maintainability of services are out of the scope).

This Recommendation describes a parametric monitoring model. This model uses network and stream related parameters obtained at the endpoint either during or at the end of a call to analyse the status of the communication in terms of QoS. The main factors affecting the quality of videotelephony services include video and audio quality factors, terminal factors and network transmission delay and loss factors.

From the perspective of human subjective perception, the model described in this Recommendation predicts the interactive audiovisual mean opinion score (MOS) on a 5-point ACR scale [ITU-T P.910] as a global multimedia MOS score, together with underlying quality estimations for audio, video, delay and synchronization between audio and video.

The application areas for this Recommendation are:

- Videotelephony services based on RTP/UDP/IP, which mainly focus on personal usage applications.
- Video-related terminals include PC, TV and mobile phones.
- Video over both mobile and fixed connections.
- It is intended to be used for videotelephony services typically using the following video codecs: H.264, H.265
- OTT services and operator managed videotelephony services, starting from, but not necessarily restricted to WebRTC services. The impact of network transmission impairment factors and audiovisual interaction information presented by the model described in this Recommendation was verified using test results on a WebRTC-based platform [b-BigBlueButton]. Verification and updates on other platforms using technologies other than WebRTC are for further study

The model described in this Recommendation requires input data from the endpoint and is not suitable for midpoint monitoring applications.

The QoE influencing factors of videotelephony services are generally divided into three main categories: human influence factors, system influence factors and context influence factors [ITU-T G.1092]. The model described in this Recommendation focuses on the system influence factors and evaluates videotelephony service quality by taking multiple groups of factors into consideration, including audio and video quality factors, network transmission impairment factors, audiovisual interaction information and terminal factors. Table 1 gives a list of influencing factors covered by the model.

Supplementary non-normative information can be found in [b-ITU-T P Suppl.31], [b-ITU-T TR CMVTQS1] and [b-ITU-T TR CMVTQS2].

Table 1 – List of influencing factors for videotelephony quality covered by this Recommendation

Dimensions	Factors	Description	Range
Video quality factors	Video codec	The video codec format used in the videotelephony service	H.264: Baseline Profile H.265: Main profile
	Video resolution	The resolution of the received video stream	QVGA, VGA, 720p, 1080p, 2160p (4K)
	Video frame rate	The frame rate of the received video stream	≤ 60fps
	Video bit rate	The bit rate of video stream in network transmission	From 200 kbit/s to 30 Mbit/s
Audio quality factors	Audio codec	The audio codec format used in the videotelephony service	Opus, in Variable Bit Rate (VBR) mode and SWB/FB audio mode
	Audio bit rate	The bit rate of audio stream in network transmission	From 32 kbit/s to 80 kbit/s (stereo)
Network transmission impairment factors	Audio packet loss rate	Packet loss rate of the audio stream at the application level	≤ 50%
	Video packet loss rate	Packet loss rate of the video stream at the application level	≤ 3%
	Video interarrival jitter	The interarrival jitter of the received video stream	< 100 ms
Audiovisual interaction information	Audio delay	End-to-end delay of audio stream in the transmission network	≤ 1000 ms
	Video delay	End-to-end delay of video stream in the transmission network	≤ 1000 ms
Terminal factors	Terminal type	Terminal types include mobile phone, PC and TV. It is assumed that the terminal device works without causing any additional degradation of video or speech signals	Mobile phone, PC, TV
	Screen size	Screen size corresponds to the diagonal length of the displaying screen, usually in inches	Mobile phone: smaller than or equal to 10 inches PC and TV: larger than 14 inches and smaller than 80 inches
	Screen resolution	Resolution of the displaying screen	Mobile phone: FHD PC: HD, FHD, UHD TV: UHD
NOTE 1 – The audio quality estimation module adopts the E-model recommended in [ITU-T G.107.x]. The applicable range of audio quality factors corresponds to that of E-model. The audio quality estimation results of other audio codecs except for Opus [IETF RFC 6716] [IETF RFC 8251] were not verified here. NOTE 2 – The applicable range of audio packet loss rate is validated for uniform packet loss and Opus codec, which has embedding forward correction mechanisms.			

Table 1 – List of influencing factors for videotelephony quality covered by this Recommendation

Dimensions	Factors	Description	Range
NOTE 3 – The applicable range of video packet loss is validated for uniform packet loss, ITU-T H.264 and ITU-T H.265 codecs. In these cases, the video packet retransmission mechanisms were disabled to focus on the packet loss rates at application level.			

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.107.1] Recommendation ITU-T G.107.1 (2019), *Wideband E-model*.
- [ITU-T G.107.2] Recommendation ITU-T G.107.2 (2023), *Fullband E-model*.
- [ITU-T G.113] Recommendation ITU-T G.113 (2024), *Transmission impairments due to speech processing*.
- [ITU-T G.1092] Recommendation ITU-T G.1092 (2023), *Taxonomy of telemeetings from a quality of experience perspective*.
- [ITU-T P.910] Recommendation ITU-T P.910 (2023), *Subjective video quality assessment methods for multimedia applications*.
- [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*.
- [ITU-R BT.1359] Recommendation ITU-R BT.1359 (1998), *Relative timing of sound and vision for broadcasting*.
- [IETF RFC 6716] IETF RFC 6716 (2012), *Definition of the Opus Audio Codec*.
- [IETF RFC 8251] IETF RFC 8251 (2017), *Updates to the Opus Audio Codec*.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 conversational quality [b-ITU-T P.10]: The quality of a bi- or multidirectional conversation as perceived by a communication partner.

3.1.2 quality of service (QoS) [b-ITU-T P.10]: The totality of characteristics of a telecommunications service that bear on its ability to satisfy stated and implied needs of the user of the service.

3.1.3 quality of experience (QoE) [b-ITU-T P.10]: The degree of delight or annoyance of the user of an application or service.