

Supplement **ITU-T P Suppl. 31 (01/2025)**

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

Supplements to ITU-T P-series Recommendations

Subjective quality evaluation of audiovisual communication in 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.

Supplement 31 to ITU-T P-series Recommendations

Subjective quality evaluation of audiovisual communication in videotelephony services

Summary

Supplement 31 to ITU-T P-series Recommendations describes subjective methods to evaluate the quality experienced by end-users during audio-visual communication using videotelephony services. It includes both non-interactive subjective evaluation methods for one-way communication, and interactive subjective evaluation methods for two-way communication. The focus is on audio quality, video quality, audiovisual quality, interaction delay experience, media synchronization experience, and overall videotelephony quality perceived by end-users.

The subjective evaluation methods described in this Supplement were applied to construct the subjective test database required for the training and validation of the quality of experience (QoE) / quality of service (QoS) monitor model for videotelephony services recommended in Recommendation ITU-T P.940. These subjective evaluation methods can also be used for purposes including, but not limited to, subjective quality evaluation of videotelephony services, comparison of quality on different devices or in different environments, and comparison of quality between different videotelephony services.

History*

Edition	Recommendation	Approval	Study Group	Unique ID
1.0	ITU-T P Suppl. 31	2025-01-23	12	11.1002/1000/16323

Keywords

Audio quality, audiovisual quality, conversational quality, interaction experience, subjective quality evaluation, video quality, videotelephony quality, videotelephony services.

* 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

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.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this publication 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 standards development process.

As of the date of approval of this publication, ITU had not received notice of intellectual property, protected by patents/software copyrights, which may be required to implement this publication. 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.....	1
3	Definitions	1
	3.1 Terms defined elsewhere	1
	3.2 Terms defined in this Supplement	2
4	Abbreviations and acronyms	2
5	Conventions	2
6	Non-interactive audiovisual material subjective tests	3
	6.1 Source stimuli	3
	6.2 Factors to be investigated and coding simulation.....	3
	6.3 Test methods.....	3
7	Interactive conversational subjective tests.....	5
	7.1 Experiment setting.....	5
	7.2 Factors to be investigated in transmission network.....	5
	7.3 Test methods.....	6
8	Test result analysis and processing.....	6
	Appendix I – Use case of the subjective evaluation methods in the construction of subjective database for objective quality assessment models.....	8
	I.1 Subjective test design	8
	I.2 Construction and use of subjective test datasets.....	11
	Appendix II – Reference command for coding simulation.....	13
	Appendix III – Example of instructions for subjects participating in non-interactive audiovisual material subjective tests	14
	Appendix IV – Reference command for network distortion simulation.....	16
	Appendix V – Example of instructions for subjects participating in interactive conversational subjective tests.....	17
	Bibliography.....	20

Supplement 31 to ITU-T P-series Recommendations

Subjective quality evaluation of audiovisual communication in videotelephony services

1 Scope

As specified in [ITU-T G.1092], the quality of experience (QoE) influencing factors of videotelephony services are generally divided into three main categories: human influence factors, system influence factors, and context influence factors. This Supplement introduces subjective evaluation methods which were designed to focus on user experience of real-time audiovisual communication mainly affected by system influence factors while context influence factors are specified. These subjective evaluation methods were defined for quantifying the impact of media coding and network impairments on audiovisual communication in videotelephony services. Depending on the purpose of subjective testing, these subjective evaluation methods can also be used to analyse the impact of human influence factors and context influence factors on the quality of audiovisual communication experience.

Two types of subjective evaluation methods are included in this Supplement. One is the audiovisual material subjective test, which is non-interactive and simulates one-way communication. The other is a conversational subjective test, which is interactive and emulates two-way communication.

2 References

- [[ITU-T F.700](#)] Recommendation ITU-T F.700 (2000), *Framework Recommendation for multimedia services*.
- [[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.920](#)] Recommendation ITU-T P.920 (2000), *Interactive test methods for audiovisual communications*.
- [[ITU-T P.940](#)] Recommendation ITU-T P.940 (2025), *Computational model used for the monitoring and quality assessment of videotelephony services*.
- [[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.500-15] Recommendation ITU-R BT.500-15 (2023), *Methodologies for the subjective assessment of the quality of television images*.

3 Definitions

3.1 Terms defined elsewhere

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

3.1.3 quality of service (QoS) [b-ITU-T P.10]: The collective effect of service performance which determines the degree of satisfaction of a user of the service.

3.1.4 spatial information; spatial perceptual information (SI) [ITU-T P.910]: A measure that indicates the amount of spatial detail in a picture.

3.1.5 temporal information; temporal perceptual information (TI) [ITU-T P.910]: A measure that indicates the number of temporal changes of a video sequence.

3.1.6 videotelephony service [b-ITU-T F.700]: An audiovisual conversational teleservice providing bidirectional symmetric real-time transfer of voice and moving colour pictures between two locations (person-to-person) via the networks involved. The minimum requirement is that under normal conditions the picture information transmitted is sufficient for the adequate representation of fluid movements of a person displayed in the head and shoulder's view.

3.2 Terms defined in this Supplement

None.

4 Abbreviations and acronyms

This Supplement uses the following abbreviations and acronyms:

ACR	Absolute Category Rating
ACR-HR	Absolute Category Rating – Hidden Reference
CI	Confidence Interval
DMOS	Differential Mean Opinion Score
FHD	Full High Definition
MOS	Mean Opinion Score
PC	Personal Computer
PVS	Processed Video Sequence
QoE	Quality of Experience
QoS	Quality of Service
RTCP	RTP Control Protocol
RTP	Real-time Transport Protocol
RTT	Round-Trip Time
SI	Spatial Information
TI	Temporal Information
TV	Television
UDP	User Datagram Protocol
UHD	Ultra High Definition
WebRTC	Web Real-Time Communication

5 Conventions

None.

6 Non-interactive audiovisual material subjective tests

6.1 Source stimuli

The content of the source stimuli for non-interactive audiovisual material subjective tests should cover a wide range of scenarios in real videotelephony services according to the purpose of the quality evaluation and related research. The duration of each source stimuli can range from 4 to 20 seconds according to [ITU-T P.910].

In terms of devices for source stimuli recording, video recording devices include but are not limited to webcams, mobile phone front/back cameras, and high-end cameras. Audio recording devices include but are not limited to built-in microphones, microphones on headsets, and microphones on earphones.

From the perspective of video information, according to the target users and scenarios of videotelephony services, typical recording environments include but are not limited to homes, offices, restaurants and outdoors. Video content should have different degrees of motion and texture complexity, which can be measured by temporal information (TI) and spatial information (SI) as defined in [ITU-T P.910]. The scene content can also be diverse, including but not limited to one person with head and shoulders, one person with half-body or full body, multiple persons, person presenting graphics or posters. The motion in the video can be caused by the body movements of the person or the movement of the recording camera.

From the perspective of audio information, different environments have different levels of background noise. The audio recording capabilities and noise compression capabilities are also different for different audio recording devices. There may be one speaker or multiple speakers on one side. These aspects should be reasonably selected according to the purpose of subjective tests and the common scenarios of the videotelephony service.

From the perspective of audiovisual information, when the purpose of non-interactive tests does not include investigating the impact of media synchronization on user experience or multimedia quality, the synchronization between audio and video of source stimuli and related test sequences should be guaranteed.

6.2 Factors to be investigated and coding simulation

The non-interactive audiovisual material subjective tests mainly investigate the influencing factors in the media coding and playback processes. Comprehensive coverage of typical media codec configurations is also recommended when the purposes of the audiovisual material subjective test include studying the impact of human influence factors and context influence factors on user experience.

In the media encoding and decoding process, the key factors that affect audio quality mainly include audio codec format, audio bit rate and audio channels. The key factors that affect video quality mainly include video codec format, video resolution, video bit rate and video frame rate. In the media playback process, the key factors that affect audio quality and video quality mainly include the type of terminal device, the type of audio speaker, screen size, and screen resolution.

In the non-interactive audiovisual material subjective tests, the high-quality source stimuli are transcoded via software or hardware to generate processed video sequences (PVS) with varying media coding configurations for subjective evaluation.

6.3 Test methods

The subjective quality in audiovisual material subjective tests is evaluated using the absolute category rating with hidden reference (ACR-HR) method according to [ITU-T P.910]. The ACR-HR method includes the evaluation of the source stimuli. During the data analysis, it is feasible to calculate mean opinion scores (MOS) as usual in the absolute category rating (ACR) method and differential mean