

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.1204

(01/2020)

SERIES P: TELEPHONE TRANSMISSION QUALITY,
TELEPHONE INSTALLATIONS, LOCAL LINE
NETWORKS

Models and tools for quality assessment of streamed
media

Video quality assessment of streaming services over reliable transport for resolutions up to 4K

Recommendation ITU-T P.1204

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.1204

Video quality assessment of streaming services over reliable transport for resolutions up to 4K

Summary

Recommendation ITU-T P.1204 is the introductory document for a set of documents that describe model algorithms for monitoring the video quality for streaming using reliable transport (e.g., adaptive streaming based on the hypertext transfer protocol (HTTP) over the transmission control protocol (TCP), quick user datagram protocol internet connections (QUIC)).

The ITU-T P.1204 series of Recommendations comprises different variants of models for sequence-related (between 5 and 10 s) and per-1-second video-quality estimation. The variants differ in the type of input information they use: bitstream based, pixel based, and hybrid (using both bitstream and pixel information).

In principle, the per-1-second outputs of these video-quality models can be used together with an audio-quality model for integration into audiovisual quality and, together with information about initial loading delay and media playout stalling events, further into a final per-session model output, an estimate of integral per-session quality (see e.g., ITU-T P.1203, ITU-T P.1203.2, ITU-T P.1203.3).

Recommendation ITU-T P.1204 was developed in collaboration with the Video Quality Experts Group (VQEG).

The structure of the set of Recommendations reflects the different functionality of modules described in each document:

- ITU-T P.1204 (2020), *Video quality assessment of streaming services over reliable transport for resolutions up to 4K*.
- ITU-T P.1204.1, Video quality assessment of streaming services over reliable transport for resolutions up to 4K with access to transport information (in development)
- ITU-T P.1204.2, Video quality assessment of streaming services over reliable transport for resolutions up to 4K with access to video frame information (in development)
- ITU-T P.1204.3 (2020), *Video quality assessment of streaming services over reliable transport for resolutions up to 4K with access to full bitstream information*.
- ITU-T P.1204.4 (2020), *Video quality assessment of streaming services over reliable transport for resolutions up to 4K with access to full and reduced reference pixel information*.
- ITU-T P.1204.5 (2020), *Video quality assessment of streaming services over reliable transport for resolutions up to 4K with access to transport and received pixel information*.

The ITU-T P.1204-series of Recommendations addresses three application areas, which are respectively indicated in the module-related ITU-T P.1204.1, ITU-T P.1204.2, ITU-T P.1204.3, ITU-T P.1204.4 and ITU-T P.1204.5:

- large-screen presentation as with fixed-network video streaming;
- mobile streaming on handheld devices such as smartphones;
- presentation on tablet-type devices.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T P.1204	2020-01-13	12	11.1002/1000/14155

Keywords

Adaptive streaming, IPTV, mean opinion score (MOS), mobile video, mobile TV, monitoring, multimedia, OTT, progressive download, QoE, TV, video.

* 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 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/>.

© ITU 2020

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.....	2
3	Definitions	3
	3.1 Terms defined elsewhere	3
	3.2 Terms defined in this Recommendation.....	3
4	Abbreviations and acronyms	3
5	Conventions	4
6	Areas of application.....	5
	6.1 Application range for the models	5
	6.2 Model types	6
7	Building blocks.....	7
	7.1 Model input interfaces.....	8
	7.2 Specification of inputs I.GEN, I.13, I.15 and I.16.....	9
	7.3 Model output information.....	11
8	Overview of databases used for model development	11
9	Description of ITU-T P.1204 model algorithms.....	12
	Appendix I – Performance figures	13
	Bibliography.....	14

Recommendation ITU-T P.1204

Video quality assessment of streaming services over reliable transport for resolutions up to 4K

1 Scope

This Recommendation describes a set of objective video quality assessment modules that together with audio and integration modules can be used to form a complete model to predict the impact of audio and video media encodings and observed Internet protocol (IP) network impairments on quality experienced by the end-user in multimedia streaming applications. The streaming techniques addressed comprise progressive download as well as adaptive streaming, for both mobile and fixed network streaming applications. The video quality modules can also be used stand-alone as a video quality prediction model.

Five model types are defined to cover a range of use-cases, from monitoring bitstreams where the video payload is fully encrypted, unencrypted bitstreams, and where deep packet inspection is possible, or where the bitstream is available at the encoding premises, up to measurement using pixel information available e.g., from the client side. The models thus have a wide range of application, from encoding optimization over client-side quality of experience (QoE) assessment to network or service optimization, or benchmarking purposes. The models in the ITU-T P.1204 series of Recommendations are bitstream based, pixel based and hybrid based.

The models described here are applicable to progressive download and adaptive streaming or other streaming applications with reliable transport, where the quality experienced by the end user is affected by video degradations due to coding, spatial re-scaling or variations in video frame rates. Quality assessment of adaptive streaming includes aspects of media adaptation, which may be handled in integration modules such as [ITU-T P.1203.3], and not in the video modules. This Recommendation is able to handle various video codecs (i.e., H.264, H.265 high-efficiency video coding (HEVC), video payload type 9 (VP9), resolutions up to 4K or ultra-high definition-1 (UHD-1) and frame rates of up to 60 frames/s. The video-quality module Pv of [b-ITU-T P.1203], i.e., [ITU-T P.1203.1], only addresses ITU-T H.264 and full high definition (HD) with up to 30 frames/s.

The models predict a mean opinion score (MOS) on a five-point absolute category rating (ACR) scale (see [ITU-T P.910]) as an overall video quality MOS (5 to 10 s). In addition to the overall quality score, the video quality models produce a per-one-second quality score, suitable for diagnostics or integration into an integral quality score for longer sessions (cf. e.g., [ITU-T P.1203.3] for 1 to 5 min duration sessions).

The models associated with this Recommendation cannot provide a comprehensive evaluation of the video quality as perceived by an *individual end-user* because the scores reflect the perceived impairments due to coded video media data being transmitted over an IP connection with a certain performance and do not include a specific terminal device or user-specific information. The scores predicted by such a general quality model necessarily reflect *average perceptual quality*.

Effects due to source generations, such as signal noise, video shake, certain colour properties (and other similar video factors), as well as other impairments related to the payload, are not reflected in the scores computed by this model.

As a consequence, this Recommendation can be used for applications such as the following.

- In-service quality monitoring for specific IP-based audiovisual services, as specified in more detail in clause 6.1.
- Performance and quality assessment of live networks (including video encoding) considering the effect due to encoding bitrate, encoding resolution, and encoding frame rate.