

International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.1201

(10/2012)

**SERIES P: TERMINALS AND SUBJECTIVE AND
OBJECTIVE ASSESSMENT METHODS**

Models and tools for quality assessment of streamed
media

Parametric non-intrusive assessment of audiovisual media streaming quality

Recommendation ITU-T P.1201



International
Telecommunication
Union

Documents
Document Sharing Hub

ITU-T P-SERIES RECOMMENDATIONS

TERMINALS AND SUBJECTIVE AND OBJECTIVE ASSESSMENT METHODS

Vocabulary and effects of transmission parameters on customer opinion of transmission quality	Series	P.10
Voice terminal characteristics	Series	P.30
		P.300
Reference systems	Series	P.40
Objective measuring apparatus	Series	P.50
		P.500
Objective electro-acoustical measurements	Series	P.60
Measurements related to speech loudness	Series	P.70
Methods for objective and subjective assessment of speech quality	Series	P.80
		P.800
Audiovisual quality in multimedia services	Series	P.900
Transmission performance and QoS aspects of IP end-points	Series	P.1000
Communications involving vehicles	Series	P.1100
Models and tools for quality assessment of streamed media	Series	P.1200
Telemeeting assessment	Series	P.1300
Statistical analysis, evaluation and reporting guidelines of quality measurements	Series	P.1400

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

Recommendation ITU-T P.1201

Parametric non-intrusive assessment of audiovisual media streaming quality

Summary

Recommendation ITU-T P.1201 provides an overview of algorithmic models for non-intrusive monitoring of the audio, video and audiovisual quality of IP-based video services based on packet-header information. The ITU-T P.1201-series of Recommendations addresses two application areas:

- ITU-T P.1201.1 specifies the model algorithm for the lower resolution (LR) application area, including services such as mobile TV.
- ITU-T P.1201.2 specifies the model algorithm for the higher resolution (HR) application area, including services such as IPTV.

The two ITU-T P.1201 model algorithms are no-reference (i.e., non-intrusive) models which operate by analysing packet header information as available from respective packet trace data, provided to the model algorithms in the packet capture format (PCAP). Further input information on more general aspects of the stream, such as the video resolution, which may not be available from packet header information, is provided to the model algorithm out-of band, for example in the form of stream-specific side information.

As output, the model algorithms provide individual estimates of audio, video and audiovisual quality in terms of the five-point absolute category rating (ACR) mean opinion score (MOS) scale. Further, diagnostic information on causes of quality degradations can also be made available.

Complementary to the ITU-T P.1201 models, there are further models described in the ITU-T P.1202-series of Recommendations. These so-called ITU-T P.1202 models are bitstream-based video quality models. The main differences with ITU-T P.1201 can be summarized as follows:

- The ITU-T P.1201 models provide audio, video and audiovisual quality estimates, while the ITU-T P.1202-only models provide video quality estimates.
- The ITU-T P.1201 models use packet header information, while the ITU-T P.1202 models exploit further bitstream information, such as coding-related information. As a consequence, the ITU-T P.1202-models can be more accurate in their quality predictions. In turn, they require non-encrypted streams, to enable the access to payload information. Since the ITU-T P.1202 models are more complex, they also require more computational power to estimate the video quality.

History

Edition	Recommendation	Approval	Study Group
1.0	ITU-T P.1201	2012-10-14	12
1.1	ITU-T P.1201 (2012) Amd. 1	2013-03-28	12

Keywords

Audio, audiovisual, IPTV, mean opinion score (MOS), mobile TV, monitoring, multimedia, QoE, video.

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 2013

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.....	4
3	Definitions	5
3.1	Terms defined elsewhere	5
3.2	Terms defined in this Recommendation.....	5
4	Abbreviations and acronyms	5
5	Conventions	7
6	Areas of application.....	7
6.1	Application range for the models	7
6.2	Modes of operation.....	11
7	Model input interfaces	14
8	Model output information and performance details	19
9	Description of the ITU-T P.1201 model algorithms.....	20
Appendix I – Detailed performance figures for the ITU-T P.1201.1 and ITU-T P.1201.2 algorithms		21
Appendix II – Performance figures for the ITU-T P.1201.2 algorithm in the case of MPEG2-TS/UDP		23



Documents

Rec. ITU-T P.1201 (10/2012)

iii

Document Sharing Hub

Recommendation ITU-T P.1201

Parametric non-intrusive assessment of audiovisual media streaming quality

1 Scope

This Recommendation provides an overview of two recommended objective parametric quality assessment models that predict the impact of observed IP network impairments on quality experienced by the end user in multimedia mobile streaming and IPTV applications over transport formats such as: RTP (over UDP), MPEG2-TS (over UDP or RTP/UDP), 3GPP-PSS (over RTP).

As an umbrella Recommendation, ITU-T P.1201 provides the framework for the actual two algorithmic models described in [ITU-T P.1201.1] and [ITU-T P.1201.2]. These algorithms are aimed at monitoring the audio, video and audiovisual quality of IP-based video services based on packet-header information, with two application areas:

- [ITU-T P.1201.1] specifies the model algorithm for the lower resolution application area, including services such as mobile TV.
- [ITU-T P.1201.2] specifies the model algorithm for the higher resolution application area, including services such as IPTV.

These models are restricted to input information contained in packet headers, prior and static knowledge about the media stream and dynamic buffering information from the client. Information from decoding the bit-stream or parsing the packet payload is not used.

These models predict mean opinion scores (MOS) on a five-point ACR scale (see [ITU-T P.910]) for audio and video parts of the stream, as well as a global audiovisual MOS score (as defined in [ITU-T P.911], for instance).

The primary application for these models is the monitoring of transmission quality during in-service operation or for maintenance purposes. The ITU-T P.1201 model may be deployed both in end-point locations and at mid-network monitoring points. The location of the model and the location of the measurement probe together determine the mode of operation, as described in more detail in clause 6.1.

The primary quality prediction made by such models is not based on the information in the audio and video coding layer (payload) of the stream being analysed, but on information on the payload as it is available from packet headers and additional side-information.

This Recommendation cannot provide a comprehensive evaluation of audiovisual quality as perceived by a particular end user because its scores reflect the impairments due to encoding and the subsequent IP network being assessed, which may only be one part of the end-to-end connection. An explicit inclusion of processing steps, such as content contribution from, for example, satellite networks, display properties, etc., are not considered. Hence, also the effects of audio level, noise, delay (and corresponding similar video factors) and other impairments related to the media payload are not reflected in the scores computed by such a model. Further, the quality impact due to a specific audio or video encoder implementation or a specific decoder-side packet loss concealment implementation is not explicitly addressed. Instead, the models have been developed for a set of dedicated service implementations, which are assumed to be meaningful representations of today's IP-based streaming video services. As a consequence, however, in case of significant deviations of a given service being assessed from the service configurations used for developing this standard, it is possible to obtain high-quality scores with this Recommendation, but yet to have a poor quality of the stream as it is perceived by actual users, or vice versa. Moreover, the scores predicted by a parametric model (i.e., without access to payload information) necessarily reflect a somewhat simplified representation of the perceptual impairment of the considered stream.



However, with only using packet-header information, the models still enable estimation of payload-related information, and thus allow to provide valid and, in most cases, accurate predictions, presuming that they are applied in an appropriate manner, following this Recommendation.

As a consequence, this Recommendation can be used for applications such as:

- in-service quality monitoring for specific IP-based audiovisual services, as specified in more detail below;
- benchmarking of different service implementations. However, it cannot be used for direct benchmarking of different encoder implementations, but only the effect of different encoding bitrates and transmission errors for a given decoder-based packet loss concealment. The implementations that can be assessed with ITU-T P.1201 include the audio and video encoding bitrates, the employed video GOP-structure, frame rate, resolution, potential packet loss and the audio codec type.

The audio model algorithm is considered as a part of ITU-T P.1201 that may be employed in combination with other than the ITU-T P.1201 video model algorithm, for example, with the bitstream-based ITU-T P.1202 model algorithms (ITU-T P.1202-series of Recommendations), or be used as a stand-alone audio model. Consequently, care is taken that the application range for the audio model algorithm is designed to be slightly larger than necessary for a component of an audiovisual model applied to services as considered for the ITU-T P.1201 activity.

The application areas of the ITU-T P.1201 model algorithms are summarized in Tables 1, 2, and 3 below:

**Table 1 – Application areas, test factors and coding technologies where [ITU-T P.1201.1] and [ITU-T P.1201.2] have been verified and are known to produce reliable results.
For details about the settings, see clause 6**

[ITU-T P.1201.1] – Lower resolution (LR)	[ITU-T P.1201.2] – Higher resolution (HR)
Applications the models are intended for	
In-service monitoring of audiovisual, video, and audio UDP-based streaming	(same for both models, see entry on the left)
Performance and quality assessment of live networks (including codecs) including the effect due to encoding bitrate, and transmission errors	(same for both models, see entry on the left)
Test factors the models have been validated for	
Encoding (compression) degradation of audio and video with a variety of bitrates Video: 40-6000 kbit/s Audio: 4.75-576 kbit/s	Encoding (compression) degradation of audio and video with a variety of bitrates Video: 0.5-30 Mbit/s Audio: 16-384 kbit/s
Packet loss degradation of audio and video (both random and bursty packet loss patterns)	Packet loss degradation of audio and video (both random and bursty packet loss patterns)
Re-buffering degradation (audio-only re-buffering not validated)	–
Video contents of different spatio-temporal complexity	(same for both models, see entry on the left)

Table 1 – Application areas, test factors and coding technologies where [ITU-T P.1201.1] and [ITU-T P.1201.2] have been verified and are known to produce reliable results.
For details about the settings, see clause 6

[ITU-T P.1201.1] – Lower resolution (LR)	[ITU-T P.1201.2] – Higher resolution (HR)
Different video keyframe and frame rates Frame rates: 5-30 Hz GOP lengths (1/keyframe rate): 2-10 s	Different video group of picture (GOP) structures and video frame rates Supporting variable GOP structures with all frame types
Different video resolutions: HVGA, QVGA, QCIF	Different video resolutions: SD (PAL/NTSC), HD (720p50, 720p60, 1080i50, 1080p25, 1080i60, 1080p30)
Different decoder-side packet loss concealment strategies (freezing with skipping, one slice per RTP packet/frame)	Different decoder-side packet loss concealment strategies (freezing with skipping, one/multiple slices per frame – based on the slicing settings chosen in the encoder)
–	Interlaced and progressive scan
Coding technologies the models have been trained on	
Video: MPEG4 Part 2, ITU-T H.264 (MPEG4 Part 10)	Video: ITU-T H.264 (MPEG4 Part 10)
Audio: AMR-NB/WB+, AAC-LC, HE-AACv1/v2	Audio: AAC-LC, HE-AACv2, AC3, MPEG-LII

Table 2 – Application areas, test factors and coding technologies for which further investigation of ITU-T P.1201.1 and ITU-T P.1201.2 models is needed

[ITU-T P.1201.1] – Lower resolution (LR)	[ITU-T P.1201.2] – Higher resolution (HR)
Applications where the models can be used, but the results may not be reliable	
In-service monitoring of live network audiovisual, video, and audio TCP-based streaming (assuming that parameter extraction from TCP-based streaming is implemented)	In-service monitoring of live network audiovisual, video, and audio TCP-based streaming without re-buffering (assuming that parameter extraction from TCP-based streaming is implemented)
Test factors where the models can be used but the results may not be reliable	
–	–
Coding technologies where the models can be used but the results may not be reliable	
Specific implementations of video encoders and decoders other than the codecs used in the development, see clause 6	(same for both models, see entry on the left)

Table 3 – Application areas, test factors, and coding technologies for which [ITU-T P.1201.1] and [ITU-T P.1201.2] are not intended to be used

[ITU-T P.1201.1] – Lower resolution (LR)	[ITU-T P.1201.2] – Higher resolution (HR)
Applications the models are not intended for	
Direct comparison/benchmarking of encoder and decoder implementations, and thus of services that employ different encoder or decoder implementations	(same for both models, see entry on the left)
Evaluation of audiovisual quality including display/device properties	(same for both models, see entry on the left)
Test factors the models are not intended for	
Audiovisual streaming with significant rate adaptation (such as used in dynamic adaptive streaming over HTTP (DASH) streaming)	(same for both models, see entry on the left)
Transcoding situations	(same for both models, see entry on the left)
The effects of audio level, noise, delay (and corresponding similar video factors)	(same for both models, see entry on the left)
–	Re-buffering degradation of audio, video and audiovisual
Coding technologies the models are not intended for	
ITU-T H.261, MPEG-2, ITU-T H.263, ITU-T H.265, etc. (Note)	
NOTE – For the exact set of codecs for which the models have been validated, see clause 6.	

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 H.264] Recommendation ITU-T H.264 (2012), *Advanced video coding for generic audiovisual services*.
- [ITU-T P.800.1] Recommendation ITU-T P.800.1 (2006), *Mean Opinion Score (MOS) terminology*.
- [ITU-T P.910] Recommendation ITU-T P.910 (2008), *Subjective video quality assessment methods for multimedia applications*.
- [ITU-T P.911] Recommendation ITU-T P.911 (1998), *Subjective audiovisual quality assessment methods for multimedia applications*.
- [ITU-T P.1201.1] Recommendation ITU-T P.1201.1 (2012), *Parametric non-intrusive assessment of audiovisual media streaming quality – Lower resolution application area*.