

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.1201

Amendment 2

(12/2013)

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

**Amendment 2: New Appendix III – Use of
ITU-T P.1201 for non-adaptive, progressive
download type media streaming**

Recommendation ITU-T P.1201 (2012) – Amendment 2

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

Amendment 2

New Appendix III – Use of ITU-T P.1201 for non-adaptive, progressive download type media streaming

Summary

Amendment 2 to Recommendation ITU-T P.1201 (2012) introduces Appendix III.

Appendix III to ITU-T P.1201 describes a method allowing ITU-T P.1201 to be used for quality predictions of TCP-based, non-adaptive streaming typically referred to as progressive download (PD). The appendix can be used standalone, since it reproduces the required portions of model algorithms from ITU-T P.1201.1 and ITU-T P.1201.2.

Here, the following changes or additions have been applied to the ITU-T P.1201 scope:

- Longer duration source sequences of between 30 s and 60 s compared to 10 s to 16 s long source sequences as in the case of the ITU-T P.1201.1, ITU-T P.1201.2, ITU-T P.1202.1 and ITU-T P.1202.2 models;
- TCP-based transport instead of UDP-based transport;
- Inclusion of longer stalling events as they may occur in case of playout-buffer underrun;
- Inclusion of initial buffering, before the actual playout starts.

It is noted that the model has been developed based on a relatively small subjective test dataset, as compared to the large database used for the ITU-T P.1201.x series model training and validation procedure.

What is not addressed by this model:

- Adaptive streaming;
- Sequences longer than 60 s;
- User-interaction with the player (stop, play, rewind, fast forward, jump to specific part of sequence, etc.).

The goal of this appendix is to provide the market with an initial model capable of estimating quality related with PD-type streaming applications but with a limited scope.

The appendix is targeted to be usable standalone, and therefore the required parts from ITU-T P.1201 and the model descriptions in ITU-T P.1201.1 and ITU-T P.1201.2 are provided here for the implementers' convenience.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T P.1201	2012-10-14	12	11.1002/1000/11727
1.1	ITU-T P.1201 (2012) Amd. 1	2013-03-28	12	11.1002/1000/11940
1.2	ITU-T P.1201 (2012) Amd. 2	2013-12-12	12	11.1002/1000/12109

Keywords

Audio, audiovisual, mean opinion score (MOS), monitoring, multimedia, progressive download, QoE, stalling, 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 2014

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

Recommendation ITU-T P.1201

Parametric non-intrusive assessment of audiovisual media streaming quality

Amendment 2

New Appendix III – Use of ITU-T P.1201 for non-adaptive, progressive download type media streaming

1) Appendix III

Introduce Appendix III after Appendix II of this Recommendation, as shown below.

Appendix III

Use of ITU-T P.1201 for non-adaptive, progressive download type media streaming

(This appendix does not form an integral part of this Recommendation.)

III.1 Scope

This appendix describes an objective parametric quality assessment model that predicts the impact of observed IP network impairments on quality experienced by the end user in multimedia mobile streaming and fixed network applications using progressive download.

The model described is restricted to information provided to it by an appropriate packet- or bitstream-analysis module. The model is applicable for the effects due to audio- and video-coding as well as initial buffering and re-buffering (which are both perceivable as stalling of the media) as the typical degradations associated with progressive download.

The model predicts mean opinion scores (MOS) on a 5-point ACR scale (see [ITU-T P.910]) as a global multimedia MOS score (as defined in [ITU-T P.911], for instance). In addition, audio-, video-only and audio-visual MOS according to ITU-T P.1201 as well as a perceptual stalling quality indicator (MOS) are provided for diagnostic purposes.

The primary applications for this model are monitoring of transmission quality for operations and maintenance purposes. The ITU-T P.1201 model for non-adaptive, progressive download type media streaming may be deployed both in end-point locations and at mid-network monitoring points. The location of the model, together with the location of the measurement probe determines the mode of operation. Note, however, that the present document only describes the model that maps input parameters obtained from a probe located at a specific point in the network or in the client to the global multimedia MOS-scores and related quality diagnostic indicators, as described above.

This model cannot provide a comprehensive evaluation of transmission quality as perceived by a particular end user because its scores reflect the impairments on the IP network being measured, which may only be part of the end-to-end connection. Furthermore, the scores predicted by a parametric model necessarily reflect an average perceptual impairment. Note also that the model is developed with a specific encoder and decoder pair. If a different encoder and decoder pair is used in a monitoring situation the scores will not reflect that.

The effects of audio level, noise, delay (and corresponding similar video factors) and other impairments related to the payload are not reflected in the scores computed by this model. Therefore, it is possible to have high scores with this model, yet have a poor quality stream overall. 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 model still enables estimation of payload-related information, and thus allows for the provision of valid and in most cases accurate predictions, presuming that it is 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. The implementations that can be assessed with ITU-T P.1201 for non-adaptive, progressive download type media streaming include the audio and video encoding bitrates, the employed video GOP-structure, frame-rate, resolution and the audio codec type.

The application areas of the ITU-T P.1201 model for non-adaptive, progressive download type media streaming are summarized in Tables III.1, III.2 and III.3:

Table III.1 – Application areas, test factors and coding technologies where ITU-T P.1201 for non-adaptive, progressive download type media streaming has been verified and are known to produce reliable results^a

Applications for which the model is intended	
In-service monitoring of audiovisual, TCP-based video and audio. Both so called over-the-top (OTT) services (for example YouTube), and operator managed video services (over TCP), using the protocols HTTP/TCP/IP and RTMP/TCP/IP.	
This model is intended to be used for video services typically using container formats such as Flash (FLV), MP4, WebM and 3GP. Note that this model is agnostic to the type of container format	
Performance and quality assessment of live networks (including codecs) including the effect due to encoding bitrate, and transmission problems causing long initial buffering and stalling events	
Test factors for which the model has been validated	
Encoding (compression) degradation of audio and video with a variety of bitrates: Video: 200-6000 kbit/s (HVGA), 2-16 Mbit/s (HD) Audio: 24-128 kbit/s	
Stalling (re-buffering) degradation (audio-only and video-only re-buffering not validated) and initial buffering	
Video contents of different spatio-temporal complexity	
Different video keyframe and frame-rates and GOP lengths for HVGA resolution	
• Frame rate: 12-30 Hz • GOP lengths: 2 s For HD resolution: • Frame rate: 24 and 30 Hz • GOP lengths: M3N24 (about 1 s)	
Different video resolutions: HVGA, HD (1080i50, 1080p24, 1080i60, 1080p30)	
Interlaced and progressive scan for HD resolution	

Table III.1 – Application areas, test factors and coding technologies where ITU-T P.1201 for non-adaptive, progressive download type media streaming has been verified and are known to produce reliable results^a

Coding technologies on which the model has been trained
ITU-T H.264 (MPEG4 Part 10)
Audio for HD resolution sequences: HE-AACv1/v2 Audio for HVGA resolution sequences: AMR-WB, AAC-LC, HE-AACv1
^a For details about the settings, cf. clause III.8.

Table III.2 – Application areas, test factors and coding technologies for which further investigation of ITU-T P.1201 for non-adaptive, progressive download type media streaming is needed

Applications where the model can be used, but the results may not be reliable
–
Test factors where the model can be used but the results may not be reliable (conditions not included in subjective tests underlying the model development)
Video resolutions: QCIF, CIF, SD (PAL/NTSC), HD (720p50, 720p60)
Encoding (compression) degradation of audio and video with bitrates slightly wider range than what has been verified: Video: 200 kbit/s-30 Mbit/s Audio: 4.75-576 kbit/s
Additional keyframe, frame-rates and GOP structures for lower resolutions: • Frame rate: 5-30 Hz • GOP lengths: 2-10 s For HD resolution: • Different video group-of-pictures (GOP) structures and video frame-rates
Coding technologies where the models can be used but the results may not be reliable
MPEG4 Part 2
For lower resolution videos: AMR-NB, HE-AACv2 For higher resolution videos: AAC-LC, AC3, MPEG1-LII
Specific implementations of video en- and decoders other than the codecs used in the development, cf. clause III.8