

International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.1202

(10/2012)

SERIES P: TERMINALS AND SUBJECTIVE AND
OBJECTIVE ASSESSMENT METHODS

Models and tools for quality assessment of streamed
media

Parametric non-intrusive bitstream assessment of video media streaming quality

Recommendation ITU-T P.1202



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.1202

Parametric non-intrusive bitstream assessment of video media streaming quality

Summary

Recommendation ITU-T P.1202 provides an overview of algorithmic models for non-intrusive monitoring of the video quality of IP-based video services based on packet-header and bitstream information. The ITU-T P.1202-series of Recommendations addresses two application areas:

- ITU-T P.1202.1 specifies the model algorithm for the lower resolution (LR) application area, including services such as mobile TV.
- ITU-T P.1202.2 specifies the model algorithm for the higher resolution (HR) application area, which includes services such as IPTV.

The ITU-T P.1202 model algorithms are no-reference (i.e., non-intrusive) models which operate by analysing packet header and bitstream 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, which may not be available from packet header and bitstream information, is provided to the model algorithm out-of-band, for example in the form of stream-specific side information.

ITU-T P.1202.1 describes one model, the model for the LR application area. ITU-T P.1202.2 describes two models for the HR application area corresponding to two modes: mode 1 and mode 2, which are both no-reference (i.e., non-intrusive) models. Mode 1 refers to a parsing mode; the model operates by analysing information in the video bitstream without fully decoding the bitstream (i.e., no pixel information is used) for MOS estimation. Mode 2 refers to a full decoding mode, in addition to the bitstream information which mode 1 uses, the model can also decode parts or all of the video bitstream (i.e., pixel information is used) for MOS estimation. Further client specific information, such as concealment type, is provided to the algorithm out-of-band, for example in the form of stream-specific side information. As output, the model algorithms provide individual estimates of video quality in terms of the five-point absolute category rating (ACR) mean opinion score (MOS). Further, diagnostic information on causes of quality degradations can be made available, too, since different types of performance parameters are derived during model calculations.

Complementary to the ITU-T P.1202 models, there are two further models described in Recommendations ITU-T P.1201.1 and ITU-T P.1201.2. The respective entry-Recommendation for these models is ITU-T P.1201. It describes packet-header-only-based video, and audio and audiovisual quality models. The main differences with ITU-T P.1202 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 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.1202	2012-10-14	12
1.1	ITU-T P.1202 (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	4
	3.1 Terms defined elsewhere	4
	3.2 Terms defined in this Recommendation.....	4
4	Abbreviations and acronyms	5
5	Conventions	5
6	Areas of application.....	6
	6.1 Application range for the models	6
	6.2 Modes of operation.....	8
7	Model input interfaces	10
8	Model output information and performance details	14
9	Description of the ITU-T P.1202 model algorithm	15
	Appendix I – Detailed performance figures for the ITU-T P.1202.1 algorithm.....	16
	Appendix II – Detailed performance figures for the ITU-T P.1202.2 mode 1 algorithm.....	17
	Appendix III – Detailed performance figures for the ITU-T P.1202.2 mode 2 algorithm	18

Recommendation ITU-T P.1202

Parametric non-intrusive bitstream assessment of video media streaming quality

1 Scope

This Recommendation describes recommended objective models for non-intrusive monitoring of the video quality of IP-based video services based on packet-header and bitstream information. This Recommendation addresses two application areas:

- [ITU-T P.1202.1] specifies the model algorithm for the lower resolution (LR) application area, including services such as mobile TV.
- [ITU-T P.1202.2] specifies the model algorithm for the higher resolution (HR) application area, which includes services such as IPTV. This application area is currently under study but not yet completed.

These models are restricted to information contained in packet headers, the packet bitstream (payload information), prior and static knowledge about the media stream and dynamic buffering information from the client.

[ITU-T P.1202.1] consists of one model. [ITU-T P.1202.2] consists of two models corresponding to two modes: mode 1 and mode 2, which both are no-reference (i.e., non-intrusive) models. Mode 1 refers to a parsing mode; the model operates by analysing information in the video bitstream without fully decoding the bitstream (i.e., no pixel information is used) for MOS estimation. Mode 2 refers to a full decoding mode, in addition to the bitstream information which mode 1 uses, the model can also decode parts or all of the video bitstream (i.e., pixel information is used) for mean opinion score (MOS) estimation. Further client specific information, such as concealment type, is provided to the algorithm out-of-band, for example in the form of stream specific side information.

These models predict video quality in terms of MOSs on a five-point ACR scale (see [ITU-T P.910]).

The primary application for these models is the monitoring of transmission quality during service operation or for maintenance purposes. The ITU-T P.1202 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 based on the payload of the stream being analysed. Therefore, this Recommendation can provide a comprehensive evaluation of quality as perceived by a particular end-user because its scores can reflect the impairments on the coding and the Internet Protocol (IP) network being measured, which differ from user to user. This Recommendation cannot provide a comprehensive evaluation of video 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 e.g., satellite networks, display properties etc. are not considered. See Table 3 for more information. Further, the quality-impact due to a specific 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 and yet to have a poor quality of the stream as it is perceived by actual users.



Documents

Document Sharing Hub

Rec. ITU-T P.1202 (10/2012)

1

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 in Tables 4 to 7;
- benchmarking of different service implementations. However, it cannot be used for direct benchmarking of different decoder implementations. The implementations that can be assessed with ITU-T P.1202 include the encoding aspects and potential packet loss.

The application areas of the ITU-T P.1202 model algorithms are summarized in Tables 1, 2, and 3 below:

**Table 1 – Application areas, test factors, and coding technologies for which [ITU-T P.1202.1] has been verified and is known to produce reliable results.
For details about the settings, see clause 6**

ITU-T P.1202.1 lower resolution (LR)	Higher resolution (HR)
Applications for which the models are intended	
In-service monitoring of video UDP-based streaming	Under study
Performance and quality assessment of live networks including the effect due to encoding and transmission errors	Under study
Test factors for which the models have been validated	
Encoding (compression) degradation of video with a variety of bitrates Video: 50 – 6000 kbit/s	Under study
Packet loss degradation of video (both random and bursty packet loss patterns)	Under study
Re-buffering degradation	–
Video contents of different spatio-temporal complexity	Under study
Different video keyframe and frame-rates Frame rates: 12.5-30 Hz GOP lengths (1/keyframe rate): 2-10 s	Under study
Different video resolutions: HVGA, QVGA, QCIF	Under study
Different decoder-side packet loss concealment strategies (freezing with skipping, one/multiple slices per frame)	Under study
Coding technologies on which the models have been trained	
Video: ITU-T H.264/AVC baseline profile	Under study