

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.1201.1

(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 – Lower resolution application area

Recommendation ITU-T P.1201.1



International
Telecommunication
Union

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.1201.1

Parametric non-intrusive assessment of audiovisual media streaming quality – Lower resolution application area

Summary

Recommendation ITU-T P.1201.1 specifies the algorithmic model for the lower resolution (LR) application area of Recommendation ITU-T P.1201. The ITU-T P.1201 series of Recommendations specifies models for monitoring the audio, video and audiovisual quality of IP-based video services based on packet-header information. The lower resolution application area of the ITU-T P.1201.1 part of ITU-T P.1201 can be applied to the monitoring of performance and quality of experience (QoE) of video services such as mobile TV. The algorithm for the higher resolution (HR) case is specified in ITU-T P.1201.2. See ITU-T P.1201 for details and respective application ranges and limitations of use.

This Recommendation includes an electronic attachment for testing compliance.

History

Edition	Recommendation	Approval	Study Group
1.0	ITU-T P.1201.1	2012-10-14	12

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.....	1
3	Definitions	2
4	Abbreviations and acronyms	2
5	Conventions	2
6	Model description	2
6.1	Parameter extraction modules	3
6.2	Aggregation of basic parameters into audio and video parameters.....	12
6.3	Aggregation of basic parameters into quality estimation model	27
6.4	Quality estimation modules	31
7	Model compliance	40

Electronic attachment: Test vectors for testing compliance.

www.itu.int

Recommendation ITU-T P.1201.1

Parametric non-intrusive assessment of audiovisual media streaming quality – Lower resolution application area

1 Scope

Recommendation ITU-T P.1201.1¹ describes an algorithmic model (ITU-T P.1201-LR: Parametric non-intrusive assessment of audiovisual media streaming quality – Lower resolution application area) for monitoring the audio, video and audiovisual quality of IP-based video services based on packet-header information. ITU-T P.1201.1 addresses the lower resolution (LR) application area, including services such as mobile TV. The ITU-T P.1201-LR model is a no-reference (i.e., non-intrusive) model which operates by analysing packet header information. 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 algorithm out-of-band, for example in the form of stream-specific side information. As output, the model provides individual estimates of audio, video and audiovisual quality in terms of the five-point absolute category rating (ACR) mean opinion score (MOS) scale. This Recommendation is to be used with media payload (audio, video) encapsulated in RTP/UDP/IP packets.

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 (2011), *Advanced video coding for generic audiovisual services*.
- [ITU-T P.1201] Recommendation ITU-T P.1201 (2012), *Parametric non-intrusive assessment of audiovisual media streaming quality*.
- [IETF RFC 3267] IETF RFC 3267 (2002), *Real-Time Transport Protocol (RTP) Payload Format and File Storage Format for the Adaptive Multi-Rate (AMR) and Adaptive Multi-Rate Wideband (AMR-WB) Audio Codecs*.
- [IETF RFC 3550] IETF RFC 3550 (2003), *RTP: A Transport Protocol for Real-Time Applications*.
- [IETF RFC 3640] IETF RFC 3640 (2003), *RTP Payload Format for Transport of MPEG-4 Elementary Streams*.
- [IETF RFC 3984] IETF RFC 3984 (2005), *RTP Payload Format for H.264 Video*.
- [IETF RFC 4352] IETF RFC 4352 (2006), *RTP Payload Format for the Extended Adaptive Multi-Rate Wideband (AMR-WB+) Audio Codec*.
- [IETF RFC 4566] IETF RFC 4566 (2006), *SDP: Session Description Protocol*.

¹ This Recommendation includes an electronic attachment for testing compliance.

3 Definitions

None.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

ACR	Absolute Category Rating
HR	Higher Resolution
HVGA	Half Video Graphics Array
IP	Internet Protocol
LR	Lower Resolution
MOS	Mean Opinion Score
P-E	Parameter Extraction
QCIF	Quarter Common Intermediate Format
QVGA	Quarter Video Graphics Array
RTP	Real-time Transport Protocol
SDP	Session Description Protocol
UDP	User Datagram Protocol

5 Conventions

None.

6 Model description

Figure 6-1 shows a block diagram for the ITU-T P.1201.1 model. The model takes packet headers and side information (e.g., encoder and decoder (codec), client behaviour, etc.) as input. The streams may be encrypted or unencrypted. The parameter extraction modules (P-E) extract audio and video-related parameters and output them to parameter calculation modules that consist of modules calculating basic parameters and modules calculating input parameters of quality estimation modules. Finally, quality estimation modules output audio, video and audiovisual MOSs at predetermined time intervals (e.g., 10 seconds) upon request of the performance monitoring system. The model details are described in the following clauses. Note that a prefix is added to the name of parameters (i.e., *A_* for audio parameters, *V_* for video parameters, *AV_* for audiovisual parameters, and no suffix for rebuffering parameters).

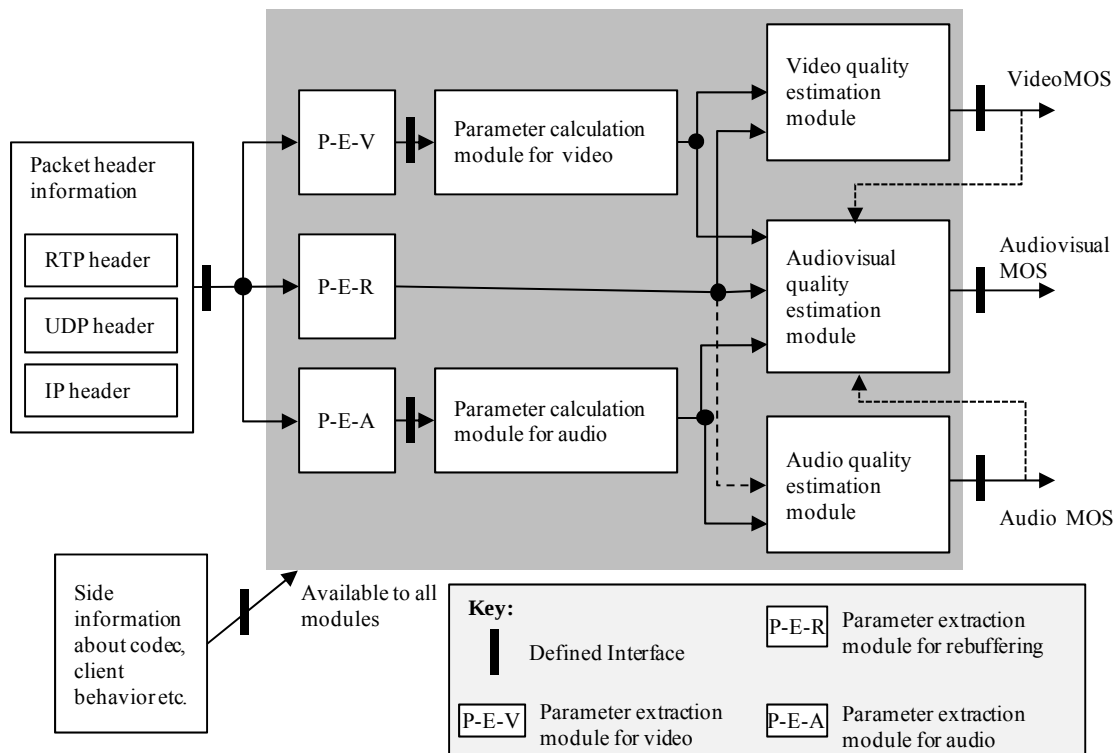


Figure 6-1 – Block diagram for the ITU-T P.1201.1 model

6.1 Parameter extraction modules

6.1.1 Parameter extraction module for audio

The parameter extraction module for audio consists of three parts on the basis of input information: 1) the side information, 2) RTP timestamp, and 3) RTP sequence number and payload, as shown in Figure 6-2.

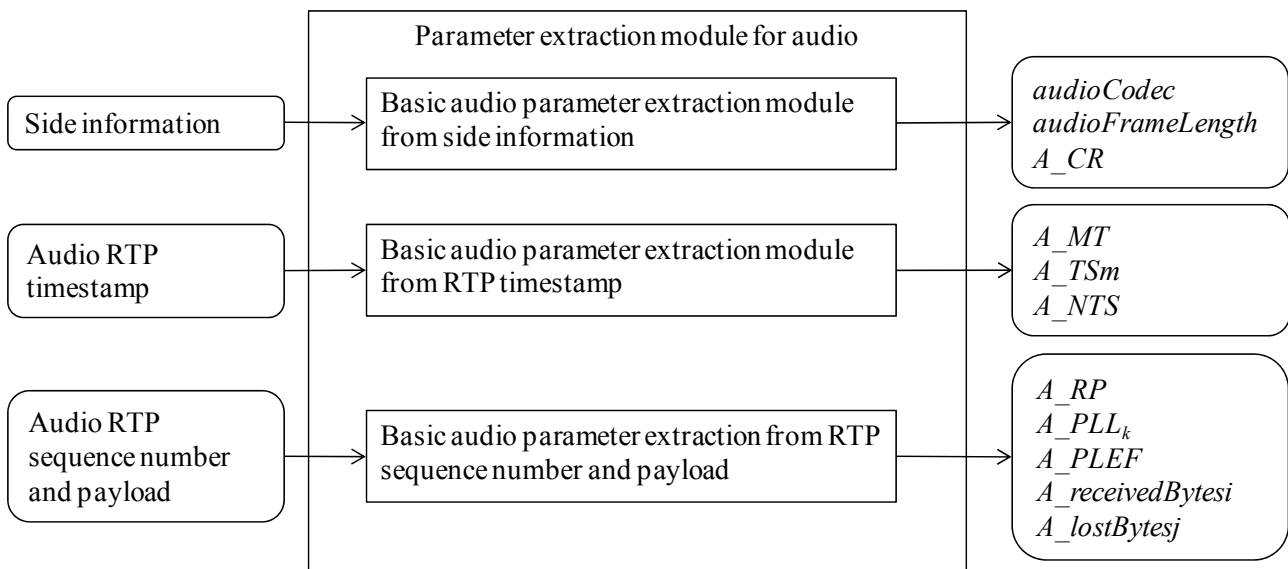


Figure 6-2 – Parameter extraction module for audio

1) Basic audio parameter extraction module from side information

The basic audio parameter extraction module from side information extracts the payload type (*audioRTPPayloadType*), destination port number (*audioDestPort*), audio codec (*audioCodec*) and audio frame length (*audioFrameLength*) in milliseconds from side information (e.g., pre file). The payload type (*audioRTPPayloadType*), destination port number (*audioDestPort*) and audio codec (*audioCodec*) may also be extracted from the session description protocol (SDP), as defined in [IETF RFC 4566], if SDP is available. The audio frame length (*audioFrameLength*) value can also be parsed from the audio elementary stream directly if the stream is not encrypted.

If the audio RTP timestamp clock rate (*A_CR*) is not input by the side information or SDP, the module calculates the audio RTP timestamp clock rate (*A_CR*) using audio frame length (*audioFrameLength*) as follows:

$$A_CR = \frac{1024 \times 10^3}{\text{audioFrameLength}} \quad (6-1)$$

Note that in AMR-NB and AMR-WB+, this step is not needed because the RTP timestamp clock rate (*A_CR*) is 8000 or 72000, respectively, as defined in [IETF RFC 4867] and [IETF RFC 4352].

The inputs and outputs of the basic audio parameter extraction module from side information are listed in Tables 6-1 and 6-2.

Table 6-1 – Inputs for the basic audio parameter extraction module from side information

Input parameter	Data type	Example value	Input from
<i>audioRTPPayloadType</i>	int	97	Side info. (e.g., pre file), or SDP if available
<i>audioDestPort</i>	int	1234	Side info. (e.g., pre file), or SDP if available
<i>audioCodec</i>	char	AAC-LC	Side info. (e.g., pre file), or SDP if available
<i>audioFrameLength</i>	double	21.33	Side info. (e.g., pre file)

Table 6-2 – Outputs of the basic audio parameter extraction module from side information

Output parameter	Data type	Example value	Output to
<i>audioCodec</i>	Char	AAC-LC	Audio quality estimation module
<i>audioFrameLength</i>	double	21.33	Audio parameter calculation module
<i>A_CR</i>	double	48000.00	Basic audio parameter extraction module from RTP timestamp Basic audio parameter calculation module

Note that in the following modules audio RTP packets should be identified by the given *audioRTPPayloadType* and/or *audioDestPort*.

2) Basic audio parameter extraction module from RTP timestamp

The basic audio parameter extraction module from RTP timestamp extracts audio RTP timestamp (*A_TS*) for the start and end of the audio stream (*A_TSs* and *A_TSe*) within predetermined time intervals (e.g., 10 seconds). The module calculates the minimum value of the difference between consecutive audio RTP packets (*A_TSm*) when the difference between consecutive audio RTP packets is greater than 0. The measurement time for audio (*A_MT*) is calculated as follows:

IF ($A_TSe > A_TSs$)

THEN

$$A_MT = \frac{A_TSe - A_TSs + A_TSm}{A_CR} \quad (6-2)$$

ELSE

$$A_MT = \frac{A_TSe + 4294967296 - A_TSs + A_TSm}{A_CR} \quad (6-3)$$

ENDIF

The module counts the number of different RTP timestamp (A_NTS) that is incremented by one when the audio RTP timestamp changes.

The inputs and outputs of the basic audio parameter extraction module from RTP timestamp is shown in Tables 6-3 and 6-4.

Table 6-3 – Inputs for the basic audio parameter extraction module from RTP timestamp

Input parameter	Data type	Example value	Input from
A_CR	double	48000.00	Basic audio parameter extraction module from side information
A_TS	int	0-4294967295	Packet (e.g., PCAP)

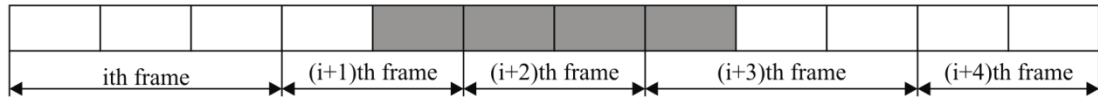
Table 6-4 – Output of the basic audio parameter extraction module from RTP timestamp

Output parameter	Data type	Example value	Output to
A_MT	double	10.06776	Audio quality estimation module
A_TSm	int	1024	Basic audio parameter calculation module
A_NTS	int	472	Basic audio parameter calculation module

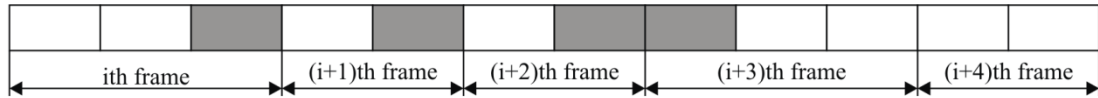
3) Basic audio parameter extraction module from RTP sequence number and payload

The basic audio parameter extraction module from RTP sequence number and payload extracts the packet-loss length (A_PLL_k) per packet-loss event on the basis of audio RTP sequence number (A_SN_i) and the audio packet-loss-event frequency (A_PLEF) at the same time. Some examples of calculating A_PLL_k and A_PLEF can be found in Figure 6-3.

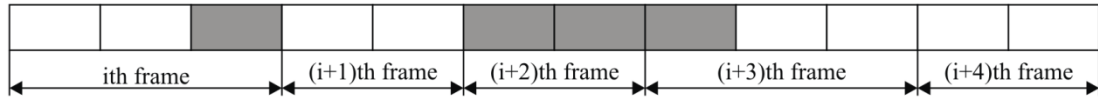
Case 1: $A_PLL1 = 4, A_PLEF = 1$



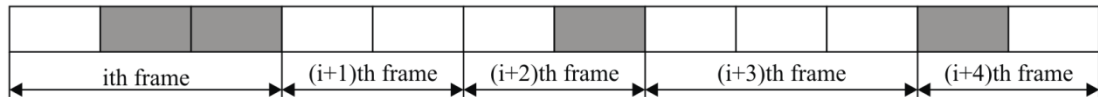
Case 2: $A_PLL1 = 1, A_PLL2 = 1, A_PLL3 = 2, A_PLEF = 3$



Case 3: $A_PLL1 = 1, A_PLL2 = 3, A_PLEF = 2$



Case 4: $A_PLL1 = 2, A_PLL2 = 1, A_PLL3 = 1, A_PLEF = 3$



P.1201.1(12)_F6-3

Figure 6-3 – Examples of calculating A_PLL_k and A_PLEF

A respective pseudocode is shown below:

IF ($A_SN_t > A_SN_{t-1}$)

THEN

$$A_PLL_k = A_SN_t - A_SN_{t-1} - 1 \quad (6-4)$$

ELSE

$$A_PLL_k = A_SN_t + 65536 - A_SN_{t-1} - 1 \quad (6-5)$$

ENDIF

where A_SN_t represents the sequence number of the current received audio RTP packet and A_SN_{t-1} is the sequence number of the previous received audio RTP packet.

The basic audio parameter extraction module from RTP sequence number and payload counts the number of received audio RTP packets (A_RP) that is incremented by one when the model receives an audio RTP packet. The module also counts the audio payload per RTP packet ($A_receivedBytes_i$) in bytes. When packet loss occurs, the module estimates the audio payload per RTP packet ($A_lostBytes_j$) for the lost packet in bytes using the current and previous received audio RTP payloads (i.e., $A_receivedBytes_i$ and $A_receivedBytes_{i-1}$) as follows:

$$A_lostBytes_j = \frac{A_receivedBytes_i + A_receivedBytes_{i-1} + 1}{2} \quad (6-6)$$

where $A_lostBytes_j$ is integral quotient.

Note that when RTP packets for audio are consecutively lost (i.e., burst packet loss), $A_lostBytes_j = A_lostBytes_{j+1} = \dots = A_lostBytes_{j+k}$, as shown in Figure 6-4.

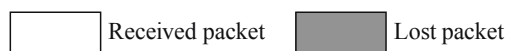
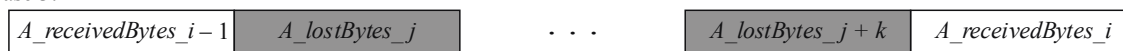
Case 1:



Case 2:



Case 3:



P.1201.1(12)_F6-4

Figure 6-4 – Examples of estimating $A_lostBytes_j$

The inputs and outputs of the basic audio parameter extraction from RTP sequence number and payload is shown in Tables 6-5 and 6-6.

Table 6-5 – Inputs for the basic audio parameter extraction module from RTP sequence number and payload

Input	Data type	Example value	Input from
A_SN	int	0-65535	Packet (e.g., PCAP)
$A_RTP\ payload$	int	1024	Packet (e.g., PCAP)

Table 6-6 – Outputs of the basic audio parameter extraction module from RTP sequence number and payload

Output	Data type	Example value	Output to
A_RP	int	472	Basic audio parameter calculation module Audio parameter calculation module
A_PLL_k	int	3	Audio parameter calculation module
A_PLEF	int	6	Audio parameter calculation module
$A_receivedBytes_i$	int	1024	Audio parameter calculation module
$A_lostBytes_j$	int	756	Audio parameter calculation module

6.1.2 Parameter extraction module for video

The parameter extraction module for video consists of two parts on the basis of input information: 1) the side information, 2) video RTP header, as shown in Figure 6-5.

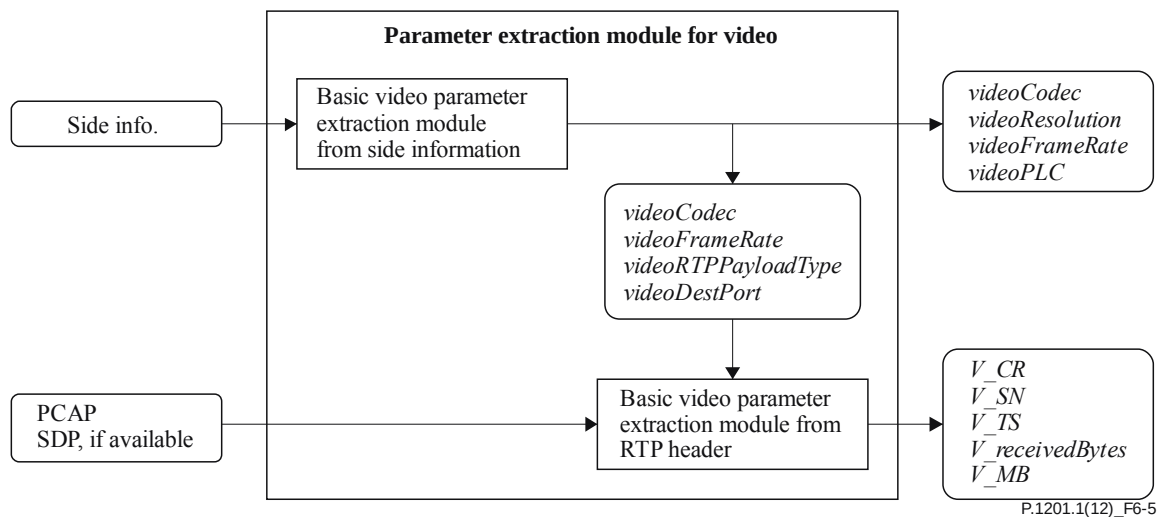


Figure 6-5 – Parameter extraction module for video

1) Basic video parameter extraction module from side information

The following parameters are extracted from side information (e.g., pre file):

- *videoCodec*: video codec.
- *videoResolution*: video resolution.
- *videoRTPPayloadType*: payload type of RTP packet for video stream.
- *videoDestPort*: destination port number of RTP packet for video stream.
- *videoFrameRate*: video frame rate, number of frames per second.
- *videoPLC*: type of packet loss concealment used by the decoder.

The video codec (*videoCodec*), payload type (*videoRTPPayloadType*), and destination port number (*videoDestPort*) may also be extracted from the session description protocol (SDP), as defined in [IETF RFC 4566], if SDP is available. The *videoResolution* value can also be parsed from the video elementary stream directly if the stream is not encrypted.

The inputs and outputs of the basic video parameter extraction module from side information is shown in Tables 6-7 and 6-8.

Table 6-7 – Inputs of the basic video parameter extraction module from side information

Input parameter	Data type	Example value	Input from
<i>videoCodec</i>	char	ITU-T H.264, MPEG-4	Side information (e.g., pre file), or SDP if available
<i>videoResolution</i>	char	QCIF, QVGA, HVGA	Side information (e.g., pre file)
<i>videoRTPPayloadType</i>	int	96	Side information (e.g., pre file), or SDP if available
<i>videoDestPort</i>	int	1232	Side information (e.g., pre file), or SDP if available
<i>videoFrameRate</i>	double	1.0-30.0	Side information (e.g., pre file)
<i>videoPLC</i>	char	SLICING, FREEZING	Side information (e.g., pre file)

Table 6-8 – Outputs of the basic video parameter extraction module from side information

Input parameter	Data type	Example value	Output to
<i>videoCodec</i>	char	ITU-T H.264, MPEG-4	Basic video parameter extraction module from RTP header Video/audiovisual quality estimation module
<i>videoResolution</i>	char	QCIF, QVGA, HVGA	Video/audiovisual quality estimation module
<i>videoRTPPayloadType</i>	int	96	Basic video parameter extraction module from RTP header Basic video parameter calculation module
<i>videoDestPort</i>	int	1232	Basic video parameter extraction module from RTP header Basic video parameter calculation module
<i>videoFrameRate</i>	double	1.0-30.0	Basic video parameter extraction module from RTP header Basic video parameter calculation module Video parameter calculation module Video/audiovisual quality estimation module
<i>videoPLC</i>	char	SLICING, FREEZING	Video parameter calculation module Video/audiovisual quality estimation module

Note that in the following modules, video RTP packets should be identified by the given *videoRTPPayloadType* and/or *videoDestPort*.

2) Basic video parameter extraction module from RTP header

The following parameters are extracted from the RTP header:

- V_SN : the sequence number of a video RTP packet
- V_TS : the timestamp of a video RTP packet
- $V_receivedBytes$: the received payload size of a video RTP packet
- V_MB : the marker bit of a video RTP packet

The video RTP timestamp clock rate (V_CR) of ITU-T H.264/AVC must be 90000, as defined in [IETF RFC 3984].

The video RTP timestamp clock rate (V_CR) of MPEG-4 is 90000, as recommended in [IETF RFC 3640]. The following two methods can be used to get the V_CR :

Method 1:

If SDP is available, the V_CR can be extracted from it.