International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

G.1071
(06/2015)

SERIES G: TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS

Multimedia Quality of Service and performance – Generic and user-related aspects

---

Opinion model for network planning of video and audio streaming applications

Recommendation ITU-T G.1071

ITU-T

ITU-T G-SERIES RECOMMENDATIONS

TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS

INTERNATIONAL TELEPHONE CONNECTIONS AND CIRCUITS	G.100–G.199
GENERAL CHARACTERISTICS COMMON TO ALL ANALOGUE CARRIER-TRANSMISSION SYSTEMS	G.200–G.299
INDIVIDUAL CHARACTERISTICS OF INTERNATIONAL CARRIER TELEPHONE SYSTEMS ON METALLIC LINES	G.300–G.399
GENERAL CHARACTERISTICS OF INTERNATIONAL CARRIER TELEPHONE SYSTEMS ON RADIO-RELAY OR SATELLITE LINKS AND INTERCONNECTION WITH METALLIC LINES	G.400–G.449
COORDINATION OF RADIOTELEPHONY AND LINE TELEPHONY	G.450–G.499
TRANSMISSION MEDIA AND OPTICAL SYSTEMS CHARACTERISTICS	G.600–G.699
DIGITAL TERMINAL EQUIPMENTS	G.700–G.799
DIGITAL NETWORKS	G.800–G.899
DIGITAL SECTIONS AND DIGITAL LINE SYSTEM	G.900–G.999
MULTIMEDIA QUALITY OF SERVICE AND PERFORMANCE – GENERIC AND USER-RELATED ASPECTS	**G.1000–G.1999**
TRANSMISSION MEDIA CHARACTERISTICS	G.6000–G.6999
DATA OVER TRANSPORT – GENERIC ASPECTS	G.7000–G.7999
PACKET OVER TRANSPORT ASPECTS	G.8000–G.8999
ACCESS NETWORKS	G.9000–G.9999

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

Recommendation ITU-T G.1071

Opinion model for network planning of video and audio streaming applications

Summary

Recommendation ITU-T G.1071 provides algorithmic models for network planning of IP-based video services. This Recommendation addresses two application areas:

– the higher resolution (HR) application area, including services such as IPTV;

– the lower resolution (LR) application area, including services such as mobile TV.

The algorithmic model addressing the HR application area is described in Annex A.

The algorithmic model addressing the LR application area is described in Annex B.

The application of the models is limited to QoE/QoS planning. Other applications such as quality benchmarking and monitoring are outside the scope of this Recommendation.

As input, the models take network planning assumptions, for example, the video resolution, the audio and video codec types and profiles, the audio and video bitrates, the packet-loss-rate and the packet-loss distribution.

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.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T G.1071	2015-06-29	12	11.1002/1000/12512

Keywords

Audio, audiovisual, IPTV, mean opinion score (MOS), mobile TV, network planning, quality, QoE, 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

NOTE

INTELLECTUAL PROPERTY RIGHTS

Table of Contents

Recommendation ITU-T G.1071

Opinion model for network planning of video and audio streaming applications

1 Scope

This Recommendation provides models which deliver estimates of the impact of typical IP network impairments on the 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).

The models are network planning tools. They are of help in selecting IP-network transmission settings such as the audio and video format, the audio and video codecs and the audio and video bitrates, under the assumption that the network is prone to packet loss.

This Recommendation targets the same services as [ITU-T P.1201] and [ITU-T P.1202]. In particular, this Recommendation covers the same coding technologies and unreliable network mechanisms such as user datagram protocol (UDP). However, this Recommendation is limited to QoE/QoS planning, while [ITU-T P.1201] and [ITU-T P.1202] are dedicated to service monitoring and benchmarking.

The following Recommendations have also been developed for QoE/QoS planning, but for different applications:

– [ITU-T G.107]: speech (telephone band)

– [ITU-T G.107.1]: speech (wideband)

– [ITU-T G.1070]: videophone

– [ITU-T G.1030]: Appendix II: web browsing

For a summary of the services, encoding and network characteristics covered by the ITU-T G.1071 model algorithms, refer to Table 1 of 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 G.107] Recommendation ITU-T G.107 (2015), *The E-model: a computational model for use in transmission planning*.

[ITU-T G.1030] Recommendation ITU-T G.1030 (2014), *Estimating end-to-end performance in IP networks for data applications.*

[ITU-T G.1070] Recommendation ITU-T G.1070 (2012), *Opinion model for video-telephony applications.*

[ITU-T H.264] Recommendation ITU-T H.264 (2014), *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]	Recommendation ITU-T P.1201 (2012), *Parametric non-intrusive assessment of audiovisual media streaming quality*.
[ITU-T P.1202]	Recommendation ITU-T P.1202 (2012), *Parametric non-intrusive bitstream assessment of video media streaming quality*.
[ITU-T P.1401]	Recommendation ITU-T P.1401 (2012), *Methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models.*

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following term defined elsewhere:

3.1.1 mean opinion score (MOS) [ITU-T P.800.1]: The mean of opinion scores, i.e., of the values on a predefined scale that subjects assign to their opinion of the performance of the telephone transmission system used either for conversation or for listening to spoken material.

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.2.1 model, model algorithm: An algorithm used for estimating the subjective (perceived) quality of a media sequence.

3.2.2 sequence: A short decoded audio, video or audiovisual portion of a stream, typically shorter than 30 seconds.

3.2.3 compression artefacts: Artefacts that are introduced due to lossy compression of the encoding process.

3.2.4 slicing artefacts: Artefacts that are introduced when packet losses are concealed through use of a packet loss concealment (PLC) scheme to repair erroneous frames.

3.2.5 freezing artefacts: Artefacts that are introduced when the packet loss concealment (PLC) scheme of the receiver replaces the erroneous frames (either due to packet loss or error propagation) with the previous error-free frame until a decoded picture without errors has been received. Since the erroneous frames are not displayed, this type of artefact is also referred to as freezing with skipping.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

AAC	Advanced Audio Coding
AAC-LC	Advanced Audio Coding – Low Complexity
AC3	Audio Coding 3
ACR	Absolute Category Rating
AMR-NB	Adaptive Multi-Rate – Narrowband
AMR-WB	Adaptive Multi-Rate – Wideband
ARQ	Automatic Repeat Request
DASH	Dynamic Adaptive Streaming over HTTP

FB	Fullband
FEC	Forward Error Correction
GOP	Group Of Pictures
HD	High Definition (television)
HE-AAC	High-Efficiency Advanced Audio Coding
HR	Higher Resolution
HRC	Hypothetical Reference Circuit
HTTP	Hypertext Transfer Protocol
HVGA	Half Video Graphics Array
LR	Lower Resolution
MBMS	Multimedia Broadcast/Multicast Service
MOS	Mean Opinion Score
MPEG	Motion Pictures Expert Group
NB	Narrowband
NTSC	National Television Standard Committee
PAL	Phase Alternating Line
PCC	Pearson Correlation Coefficient
PES	Packetized Elementary Stream
PLC	Packet Loss Concealment
PVS	Processed Video Sequence
QCIF	Quarter Common Intermediate Format
QoE	Quality of Experience
QVGA	Quarter Video Graphics Array
RMSE	Root Mean Square Error
RTP	Real-time Transport Protocol
SD	Standard Definition
SRC	Source Reference Channel or Circuit
SWB	Superwideband
TS	Transport Stream
UDP	User Datagram Protocol
VSP	Visual Simple Profile
WB	Wideband

5 Conventions

None.

6 Areas of application

Table 1 below shows the application range of the models based on what the models have actually been developed for.

Table 1 – Factors and application ranges of the ITU-T G.1071 model algorithm

	ITU-T G.1071 – Lower resolution (LR)	**ITU-T G.1071 – Higher resolution (HR)**
Application information	**Value range, unit**	
Sequence duration (Ts)	It is expected that the model will give reliable prediction results for sequence durations within the range 8-24 seconds	It is expected that the model will give reliable prediction results for sequence durations of approximately 8-16 seconds
Packetization	3GPP MBMS, PSS or using RTSP directly (all three over RTP/UDP/IP)	MPEG2-TS/RTP/UDP/IP RTP/UDP/IP (Note 3) MPEG2-TS/UDP/IP (Note 3)
Video codec	MPEG4 visual simple profile (VSP) ITU-T H.264 baseline profile	ITU-T H.264 main profile, ITU-T H.264 high profile
Video size	QCIF, QVGA, HVGA	SD: PAL, NTSC HD: 720p, 1080p, 1080i (High profile: 1080; main profile: 720, SD)
Audio codec	AMR-NB, AMR-WB+, AAC-LC, HE-AAC (v1, v2)	MPEG-4 AAC-LC MPEG-4 HE-AAC (V1 and V2 = 3GPP enhanced AAC+) MPEG-1 Layer 2 AC3
Coded video bitrate	[ITU-T H.264] QCIF: 32-1000 kbit/s QVGA: 80-3000 kbit/s HVGA: 192-6000 kbit/s MPEG4 QCIF: 40-1500 kbit/s QVGA: 90-3500 kbit/s HVGA: 192-6000 kbit/s	HD [ITU-T H.264]: 0.5 up to 30 Mbit/s SD [ITU-T H.264]: 0.5 up to 9 Mbit/s
Coded audio bitrate	AMR-NB: 4.75-12.2 kbit/s AMR-WB+: 10.4-48 kbit/s AAC-LC: 16-128+ kbit/s HE-AAC (v1, v2): 32-128 kbit/s	AAC-LC:32-576 kbit/s HE AACv2: 16-96 kbit/s MPEG-1 Layer 2: 64-384 kbit/s AC3:64-384 kbit/s

Table 1 – Factors and application ranges of the ITU-T G.1071 model algorithm

	ITU-T G.1071 – Lower resolution (LR)	**ITU-T G.1071 – Higher resolution (HR)**
Application information	**Value range, unit**	
Video decoder packet loss concealment	Two types of assumed decoder behaviour are covered: 1) freezing with skipping; 2) slicing with: MPEG4: 1 slices/frame [ITU-T H.264]: 1 slice/packet Both MPEG4 and [ITU-T H.264]: Fixed PLC (using fixed decoder, details and settings)	Types of decoder behaviour: two dimensions: slicing, PLC 1) freezing with skipping (duration(source)=duration(processed sequence)); 2) slicing with 1 slice per frame; 3) slicing with 1 slice per macroblock row, PLC with zero-motion copy (temporal from same region of previous good frame).
Audio decoder packet loss concealment	Decoder default modes: Codec-implementation specific loss concealment	Decoder default modes: Codec-implementation specific loss concealment
Retransmission mechanisms (ARQ); forward error correction (FEC); client jitter buffer behaviour	Developed models assume that the dejitter buffer, ARQ and FEC mechanisms have already corrected the stream.	(No rebuffering) Developed models assume that the dejitter buffer, ARQ and FEC mechanisms have already corrected the stream.
Encoder implementation	The model has been developed using the following encoders (Note 1): Video: • MPEG4 Part 2: ffmpeg • ITU-T H.264 (MPEG4 Part 10): x264 Audio: • AMR-NB/WB+: According to standard • AAC-LC, HE-AAC(v1, v2): Nero	The model has been developed using the following encoders (Note 1): Video: • ITU-T H.264 (MPEG-4 Part 10): x264 Audio: • AAC-LC, HE-AAC v2: Nero • MPEG1-LII and AC3: ffmpeg
Decoder implementation	Reference decoder was a proprietary decoder provided by one proponent, which also performed de-packetization and audio-video-demultiplexing. The ITU-T H.264-decoding is standard-conformant, with the PLC as described above (Note 2).	Reference decoder was a proprietary decoder provided by one proponent, which also performed de-packetization and audio-video-demultiplexing. The ITU-T H.264-decoding is standard-conformant, with the PLC as described above (Note 2).
Group of pictures (GOP)	Typical GOP structure for which the model has been trained: M = 1, N = 40 (typically no B frames for mobile case) Length: fixed, variable, adaptive Structure (e.g., IPPP...PPPI)	Supporting default modes for typical GOP structures E.g., M = 3, N = 15 Length: fixed, variable, adaptive Structure (e.g., IBBPBB...PBBI) NOTE – GOP structure is explicitly estimated from stream.

Table 1 – Factors and application ranges of the ITU-T G.1071 model algorithm

	ITU-T G.1071 – Lower resolution (LR)	ITU-T G.1071 – Higher resolution (HR)
Application information	**Value range, unit**	
Frame rate	5, 8.33, 12.5, 15, 20, 25, 30 fps	SD: 50i (PAL), 59.94i (NTSC) HD: 50p, 59.94p, 60p, 50i, 59.94i, 60i, 25p, 29.97p, 30p
Audio channel number	1 (diotic mono), 2 (stereo)	2 (stereo)
# of Audio frames per RTP packet	1 to 5 audio frames	Bitrate-specific (depending on both the audio and video bitrate)
Audio-video multiplexed?	Default: No, at RTP-level; no audio-video asynchrony	In MPEG-2 TS/RTP/UDP and MPEG-2 TS/UDP: Supported
Packet loss degradation, video	Uniform loss:0-10% Burst loss:0-10% (4-state Markov model)	Uniform loss:0-2% Burst loss:0-2% (4-state Markov model)
Packet loss degradation, audio	Uniform loss:0-10% Burst loss:0-10% (4-state Markov model)	Uniform loss:0-6% Burst loss:0-6% (4-state Markov model)
Symmetrical versus asymmetrical handling of audio and video in audiovisual case	Model application: Symmetrical, but can handle asymmetric cases due to specific model development process	Model application: Symmetrical, but can handle asymmetric cases due to specific model development process

NOTE 1 – It is assumed that the model can be used for estimating quality when other encoder implementations for the given codec have been used. However, if the encoder performance is significantly worse or better than for the encoder used, the model prediction accuracy will be reduced.

NOTE 2 – One aspect not covered by decoder packet loss concealment is post-filtering. Guidance on how to adjust internal model parameters for specific other decoders, including set-top boxes, is for further study.

NOTE 3 – The ITU-T P.1201.2 model has been trained on MPEG-2 TS/RTP/UDP. However, due to the design of the ITU-T P.1201.2 algorithm, it is also applicable to MPEG2-TS/UDP/IP. Further, to the model's design, it is assumed to also work for RTP/UDP/IP transport with similar, but so far unverified, accuracy as compared to MPEG2-TS/RTP/UDP/IP.

Factors and applications not covered by the model are:

– Evaluation of audiovisual quality including display/device properties
– Audiovisual streaming with significant rate adaptation (such as that used in dynamic adaptive streaming over HTTP (DASH) streaming)
– Transcoding situations
– The effects of audio level, noise and delay (and corresponding similar video factors)
– 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.

7 Model framework

As shown in Figure 1, the ITU-T G.1071 models follow the same structure as the ITU-T P.1201 models. Indeed, they are composed of three modules: the audio module, the video module and the audiovisual module. However, in contrast to the ITU-T P.1201 models, the ITU-T G.1071 models take as inputs network planning assumptions instead of the encrypted bitstream. Some of the ITU-T P.1201 input parameters are available as planning assumptions as well. For the ITU-T P.1201 input parameters which are not available as planning assumptions, the ITU-T G.1071 models provide a set of rules to convert planning assumptions into these ITU-T P.1201 input parameters (see block "conversion rules").

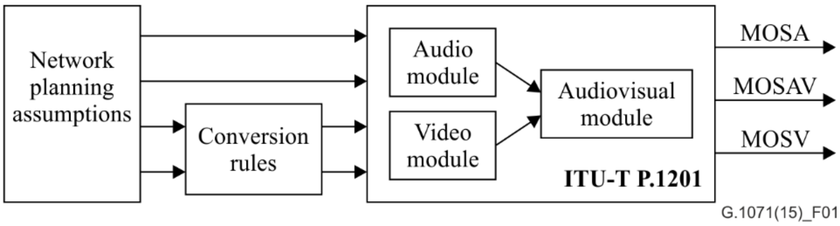

Figure 1 – Model framework

8 Model output information and performance details

ITU-T G.1071 has three output parameters:

1) Estimated audiovisual MOS on the 1 to 5 scale, which is an estimation of the perceived audiovisual quality.
2) Estimated video MOS on the 1 to 5 scale, which is an estimation of the perceived video quality (without audio present). Note that the model is able to give both a video score for a stream without audio and a stream including audio.
3) Estimated audio MOS on the 1 to 5 scale, which is an estimation of the perceived audio quality (without video present). Note that the model is able to give both an audio score for a stream without video and a stream including video.

The performance information for the ITU-T G.1071 models can be found in Table 2 and Table 3 for the HR and LR application areas respectively. The statistical metrics root mean square error (RMSE) and the Pearson correlation are used to describe the performance, see [ITU-T P.1401]. (Note that for these performance figures, the subjective ratings have been mapped to the model scores using a linear, i.e., 1st-order, mapping function at a per-database-level. This has been done in order to avoid misalignment due to bias in the different subjective tests, e.g., as a result of different test settings).

Table 2 – Performance information for ITU-T G.1071 (HR)
Samples were taken from the ITU-T P.1201.2 training and validation databases

	RMSE	Pearson correlation
Audiovisual	0.51 (based on 1595 samples)	0.87 (based on 1595 samples)
Video	0.53 (based on 3069 samples)	0.86 (based on 3069 samples)
Audio	0.37 (based on 680 samples)	0.93 (based on 680 samples)

Table 3 – Performance information for ITU-T G.1071 (LR)
Samples were taken from the ITU-T P.1201.1 training and validation databases

	RMSE	Pearson correlation
Audiovisual	0.50 (based on 1166 samples)	0.83 (based on 1166 samples)
Video	0.60 (based on 1430 samples)	0.78 (based on 1430 samples)
Audio	0.38 (based on 690 samples)	0.93 (based on 690 samples)

Annex A

Description of the ITU-T G.1071 model algorithm for HR application area

(This annex forms an integral part of this Recommendation.)

A.1 Audio module

The HR application area audio module is depicted in Figure A.1.

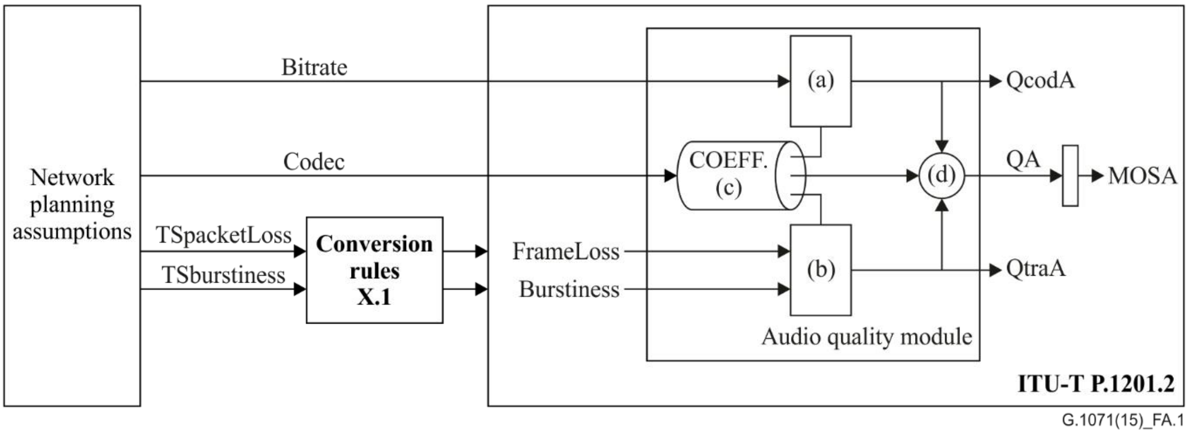

Figure A.1 – Audio module

$$QA = 100 - QcodA - QtraA \tag{1.1}$$

from [ITU-T P.1201.2]

$$MOSA = MOSfromR(QA) \tag{1.2}$$

from [ITU-T P.1201.2]

Where:

$$QcodA = a1A \cdot e^{a2A \cdot BitrateA} + a3A \tag{1.3}$$

from [ITU-T P.1201.2]

And

$$QtraA = (b1A - QcodA).\frac{FrameLossA}{FrameLossA + b2A \cdot BurstinessA + b3A} \tag{1.4}$$

from [ITU-T P.1201.2].

The following conversion rules are applied:

$$FramelossA = c1A \cdot BitrateA \cdot TSpacketLossA + c2A \cdot TSpacketLossA \tag{1.4a}$$

$$BurstinessA = d1A \cdot TSburstinessA + d2A \cdot BitrateA \cdot TSBurstinessA + d3A \tag{1.4b}$$

where:

- *QA* is the overall estimated audio quality, expressed on a 100-point scale [0,100], where 0 is the worst quality and 100 the best quality
- *MOSA* is the overall estimated audio quality, expressed on the MOS scale [1,5], where 1 is the worst quality and 5 the best quality
- *QcodA* is the estimated audio quality for audio compression artefacts
- *QtraA* is the estimated audio quality for audio transmission errors

BitrateA	is the audio bitrate (in kbit/s)
FramelossA	is the audio frame loss percentage
TSpacketLossA	is the percentage of audio-TS-packet loss
BurstinessA	is the audio frame loss "burstiness" (average number of consecutively lost audio frames),
TSburstinessA	is the audio-TS-packet "loss burstiness" (average number of consecutively lost audio-TS packets)

The coefficient values $a1A$, $a2A$, $a3A$, $b1A$, $b2A$, $b3A$, $c1A$, $c2A$, $d1A$, $d2A$ and $d3A$ depend on the audio codec. Coefficient values are provided in Tables A.1 and A.2.

MOSfromR transform quality scores expressed on the 100-point scale to the *MOS* scale; *MOSfromR* is expressed as follows (from [ITU-T P.1201.2]):

```
function MOS = MOSfromR(Q)

set MOS_MAX = 4.9;
set MOS_MIN = 1.05;

if (Q > 0 & Q < 100),
   MOS = (MOS_MIN +(MOS_MAX-MOS_MIN)/100×Q+Q×(Q-60)×(100-Q)×7.0E-6);
elseif (Q >= 100),
      MOS = MOS_MAX;
else
      MOS = MOS_MIN;
end
```

TSpacketLossA abd *TSburstinessA* are computed from RTP-layer parameters as follows:

If

$$\frac{BitrateA}{BitrateV \cdot 10^3} = \frac{(7 - NTSV)}{NTSV}$$

with:

$NTSV$	= [1,6]
$NTSV$	integer

then

$$TSpacketLossA = RTPpacketLoss \tag{1.4c}$$

$$TSburstinessA = 7 \cdot \frac{BitrateA}{BitrateA + BitrateV \cdot 10^3} \cdot RTPburstiness \tag{1.4d}$$

If

$$\frac{BitrateA}{BitrateV \cdot 10^3} = \frac{1}{6 + 7 \cdot D}$$

then

$$TSpacketLossA = RTPpacketLoss \tag{1.4e}$$

$$TSburstinessA = \frac{7 \cdot BitrateA}{BitrateV \cdot 10^3 + BitrateA} \cdot burstLengthA \cdot RTPburstiness \tag{1.4f}$$

If each RTP packet contain audio TS packets only, then

$$TSpacketLossA = RTPpacketLoss \tag{1.4g}$$

$$TSburstinessA = 7 \cdot RTPburstiness \tag{1.4h}$$

where:

BitrateA	is the audio bitrate (in kbit/s)
BitrateV	is the video bitrate (in Mbit/s)
TSpacketLossA	is the percentage of audio-TS-packet loss
RTPpacketLoss	is the percentage of RTP packet loss
TSburstinessA	is the audio-TS-packet "loss burstiness" (average number of consecutively lost audio-TS packets),
RTPburstiness	is the RTP packet loss "burstiness" (average number of consecutively lost RTP packets)
NTSV	is the average number of video TS packets into one RTP packet
D	is the number of RTP packets which contain video TS packets only between two RTP packets containing a single audio TS packet
burstLengthA	is the average number of audio TS packets in one RTP packet for RTP packets containing audio TS packets.

A.1.1 Coefficient values and parameter ranges

Table A.1 – Audio model coefficients for the different audio codecs (from [ITU-T P.1201.2])

Audio codec	`a1A`	`a2A`	`a3A`	`b1A`	`b2A`	`b3A`
Mp2	100.0	−0.02	15.48	100.0	1.51	1.64
AC3	100.0	−0.03	15.70	100.0	0.2	2.40
AacLC	100.0	−0.05	14.60	101.32	0.1	4.09
HeAac	100.0	−0.11	20.06	105.68	0.1	5.92

Table A.2 – Audio model coefficients to convert network planning parameters into ITU-T P.1201.2 parameters

Audio codec	`c1A`	`c2A`	`d1A`	`d2A`	`d3A`
Mp2	0.006	1.124	0.682	−0.001	0.908
AC3	0.016	0.973	0.277	−0.003	0.974
AacLC	0.005	0.976	0.486	−0.001	0.923
HeAac	0.026	0.482	−0.627	0.012	0.984

A.2 Video module

The HR application area video module is depicted in Figure A.2.

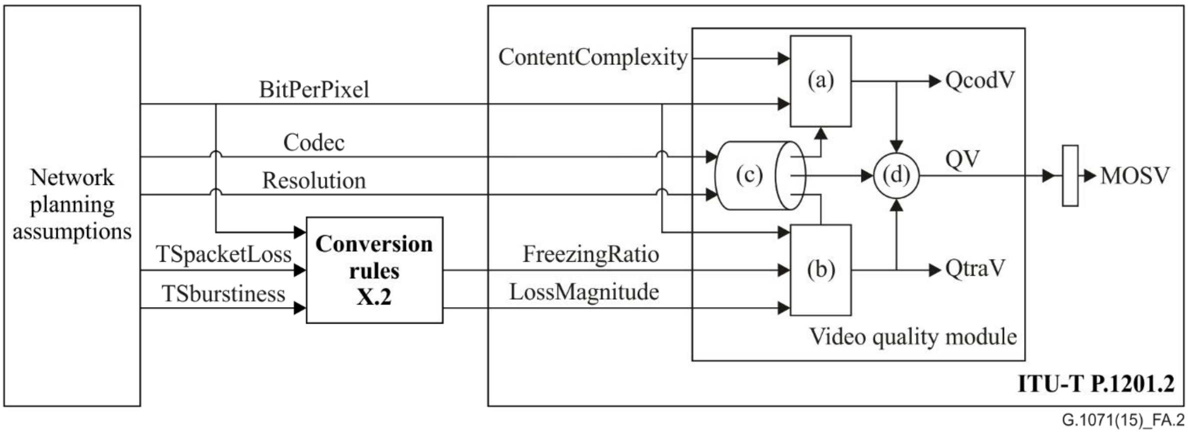

Figure A.2 – Video module

$$QV = 100 - QcodV - QtraV \tag{2.1}$$

from [ITU-T P.1201.2]

$$MOSV = MOSfromR(QV) \tag{2.2}$$

from [ITU-T P.1201.2]

with

$$QcodV = a1V \cdot e^{a2V \cdot BitPerPixel} + a3V \cdot ContentComplexity + a4V \tag{2.3}$$

from [ITU-T P.1201.2]

with

$$BitPerPixel = \frac{BitrateV \cdot 10^6}{NumPixelPerFrame \cdot FrameRate} \tag{2.3a}$$

from [ITU-T P.1201.2]

In the case of freezing:

$$QtraV = b1V \cdot \log(b2V \cdot FreezingRatioE + 1) \tag{2.4}$$

from [ITU-T P.1201.2]

(where $FreezingRatioE = FreezingRatio \cdot BitPerPixel$ in [ITU-T P.1201.2])

In the case of slicing:

$$QtraV = c1V \cdot \log(c2V \cdot LossMagnitudeE + 1) \tag{2.5}$$

from [ITU-T P.1201.2]

(where $sMagnitudeE = \frac{LossMagnitude}{QcodVn}$, with

$QcodVn = \begin{cases} 1 & , \ QcodV < 20 \\ 0.1125 \cdot QcodV - 1.25, & QcodV \geq 20 \end{cases}$ in [ITU-T P.1201.2])

QV	is the overall estimated video quality, expressed on a 100-point scale [0,100], where 0 is the worst quality and 100 the best quality
MOSV	is the overall estimated video quality, expressed on the MOS scale [1,5], where 1 is the worst quality and 5 the best quality
QcodV	is the estimated video quality for video compression artefacts
QtraV	is the estimated video quality for video transmission errors
BitPerPixel	is the average number of bits per pixel (see clause A.2.1)

BitrateV is the video bitrate in Mbps

NumPixelPerFrame is the number of pixels per video frame

FrameRate is the video frame rate

ContentComplexity captures the spatio-temporal complexity of the video sequence (see clause A.2.1)

FreezingRatioE captures the degradation when freezing is applied as packet loss concealment (see clause A.2.2)

LossMagnitudeE captures the degradation when slicing is applied as packet loss concealment (see clause A.2.3)

The coefficient values $a1V$, $a2V$, $a3V$, $a4V$, $b1V$, $b2V$, $c1V$, and $c2V$ depend on the video resolution. Coefficient values are provided in Table A.3. *MOSfromR* has been defined in clause A.1.

A.2.1 Video compressions

The *ContentComplexity* parameter of equation 2.3 is estimated from network planning parameters as follows:

$$ContentComplexity = a_{31} \cdot \exp(a_{32} \cdot BitPerPixel) + a_{33} \tag{2.3b}$$

where *BitPerPixel* is the already defined average number of bits per pixel.

If *BitPerPixel* > 0.1 OR medium content complexity is assumed for all contents:

$$ContentComplexity = a_{31} \cdot \exp(a_{32} \cdot BitPerPixel) + a_{33}$$

Else if *BitPerPixel* ≤ 0.1

If the network planner assumes that contents have high spatio-temporal (ST) complexity:

$$ContentComplexity > a_{31} \cdot \exp(a_{32} \cdot BitPerPixel) + a_{33} \tag{2.3c}$$

If the network planner assumes that contents have low to medium ST complexity:

$$ContentComplexity \leq a_{31} \cdot \exp(a_{32} \cdot BitPerPixel) + a_{33} \tag{2.3d}$$

The coefficient values a_{31}, a_{32}, and a_{33} depend on the video resolution. They are provided in Table A.4.

A.2.2 Video freezing

The *FreezingRatioE* parameter of equation 2.4 is estimated from network planning parameters as follows:

$$FreezingRatioE = \ p_1 \cdot \exp(p_2 \cdot FreezinRarioNP) \tag{2.4a}$$

where

$$FreezingRatioNP = (b_{21} - Icodn) \cdot \frac{TSpacketLossV}{Icodn \cdot (b_{22} \cdot TSburstinessV + b_{23}) + TSpacketLossV} \tag{2.4b}$$

and

$$Icodn = \begin{cases} QcodV, & QcodV \leq 65 \\ 65, & QcodV > 65 \end{cases} \tag{2.4c}$$

TSpacketLossV is the percentage of lost TS video packets in the measurement window, and

TSburtinessV is the average number of consecutively lost video TS packets in the measurement window

QcodV is expressed in equation 2.3

The coefficient values p_1, p_2, b_{21}, b_{22}, and b_{23} are the same for both SD and HD. Detailed values are provided in Table A.5.

TSpacketLossV abd *TSburstinessV* are computed from RTP-layer parameters as follows:

If

$$\frac{BitrateA}{BitrateV \cdot 10^3} = \frac{(7 - NTSV)}{NTSV}$$

with NTSV = [1,6], NTSV integer

then

$$TSpacketLossV = RTPpacketLoss \tag{2.4d}$$

$$TSburstinessV = \frac{7 \cdot BitrateV \cdot 10^3}{BitrateA + BitrateV \cdot 10^3} \cdot RTPburstiness \tag{2.4e}$$

If

$$\frac{BitrateA}{BitrateV \cdot 10^3} = \frac{1}{6 + 7 \cdot D}$$

then

$$TSpacketLossV = RTPpacketLoss \tag{2.4f}$$

$$TSburstinessV = RTPburstiness \cdot (7 - \frac{7 \cdot burstLengthA \cdot BitrateA}{BitrateV \cdot 10^3 + BitrateA}) \tag{2.4g}$$

If each RTP packet contains video TS packets only, then

$$TSpacketLossV = RTPpacketLoss \tag{2.4h}$$

$$TSburstinessV = 7 \cdot RTPburstiness \tag{2.4i}$$

where

BitrateA	is the audio bitrate (in kbit/s)
BitrateV	is the video bitrate (in Mbit/s)
TSpacketLossV	is the percentage of lost TS video packets in the measurement window, and
TSburtinessV	is the average number of consecutively lost video TS packets in the measurement window
RTPpacketLoss	is the percentage of RTP packet loss
RTPburstiness	is the RTP packet loss "burstiness" (average number of consecutively lost RTP packets)
NTSV	is the average number of video TS packets into one RTP packet
D	is the number of RTP packets which contain video TS packets only between two RTP packets containing a single audio TS packet
burstLengthA	is the average number of audio TS packets in one RTP packet for RTP packets containing audio TS packets.

A.2.3 Video slicing

The *LossMagnitudeE* parameter of equation 2.5 is estimated from network planning parameters as follows:

$$LossMagnitudeE = q_1 \cdot \exp(q_2 \cdot LossMagnitudeNP) - q_1 \tag{2.5a}$$

where

$$LossMagnitudeNP = (c_{21} - Icodn) \cdot \frac{TSpacketLossV}{Icodn \cdot (c_{22} \cdot TSburstinessV + c_{23}) + TSpacketLossV} \tag{2.5b}$$

and

$$Icodn = \begin{cases} QcodV, & QcodV \leq 65 \\ 65, & QcodV > 65 \end{cases} \tag{2.5c}$$

TSpacketLossV is the percentage of lost TS video packets in the measurement window, and

TSburstinessV is the average number of consecutively lost TS video packets in the measurement window,

QcodV is expressed in equation 2.3

The coefficient values q_1, q_2, c_{21}, c_{22}, and c_{23} are the same for SD and HD, but they depend on number of slices per frame. Detailed values are provided in Table A.6.

To compute *TSpacketLossV* abd *TSburstinessV* from RTP-layer parameters, the same conversion rules as for Video Freezing (equations 2.4d to 2.4i)) are used.

A.2.4 Coefficient values and parameter ranges

Table A.3 – Video model coefficients for the different video resolutions (from ITU-T P.1201.2)

Video resolution	a1V	a2V	a3V	a4V	b1V	b2V	c1V	c2V
SD	61.28	−11.00	6.00	6.21	12.70	907.36	17.73	123.08
HD	51.28	−22.00	6.00	6.21	12.70	907.36	17.73	123.08

Table A.4 – Coefficient values for the ContentComplexity parameter

	SD (PAL, NTSC)	HD (HD720, HD1080)
a_{31}	0.91	3.92
a_{32}	−9.39	−27.54
a_{33}	0.10	0.26

Table A.5 – Coefficient values for the FreezingRatioE parameter

	SD (PAL, NTSC), HD (HD720, HD1080)
p_1	0.0001661
p_2	0.1166
b_{21}	69.39
b_{22}	0.00019
b_{23}	0.00082

Table A.6 – Coefficient values for the LossMagnitudeE parameter

	1 slice / frame	> 1 slice/frame
q_1	0.018	0.018
q_2	0.040	0.040
c_{21}	80.61	67.15
c_{22}	0.00046	0.00144
c_{23}	0.00147	0

A.3 Audiovisual module

As shown in Figure A.3, the audiovisual module of ITU-T G.1071 is identical to the audiovisual module in [ITU-T P.1201.2].

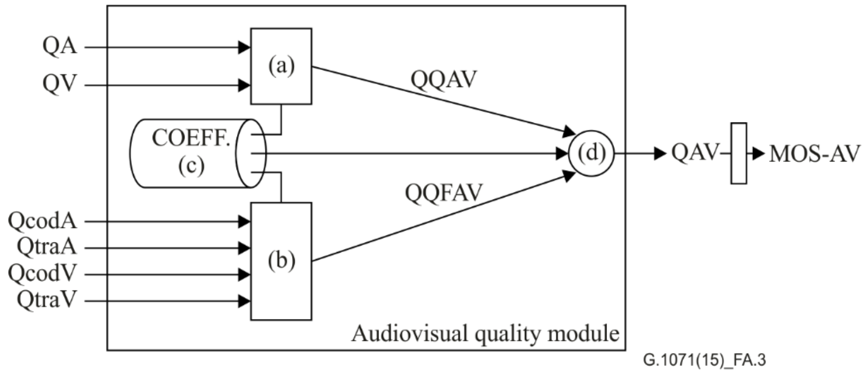

Figure A.3 – Audiovisual module

$$\mathrm{QAV} = 0.7 \times \mathrm{QQAV} + 0.3 \times \mathrm{QQFAV} \tag{3.1}$$

from [ITU-T P.1201.2]

$$MOSAV = MOSfromR(QAV) \tag{3.2}$$

from [ITU-T P.1201.2]

with

$$QQAV = \alpha + \beta \cdot QV + \gamma \cdot QA \cdot QV \tag{3.3}$$

from [ITU-T P.1201.2]

$$\begin{aligned} QQFAV &= a - b \cdot QcodA - c \cdot QcodV \\ &- d \cdot QtraA - e \cdot QtraV - f \cdot QtraA \cdot QtraV \\ &- g \cdot QcodV \cdot QtraA - h \cdot QcodA \cdot QtraV \end{aligned} \tag{3.4}$$

from [ITU-T P.1201.2]

where

QAV, *QA*, and *QV* are the overall estimated audiovisual, audio and video quality, expressed on an 100-point scale [0,100], where 0 is the worst quality and 100 the best quality

MOSAV, *MOSA*, and *MOSV* are the overall estimated audiovisual, audio and video quality, expressed on the MOS scale [1,5], where 1 is the worst quality and 5 the best quality

QcodA is the estimated audio quality for audio compression artefacts (see clause I.1)

QtraA is the estimated audio quality for audio transmission errors (see clause I.1)

QcodV is the estimated video quality for audio compression artefacts (see clause I.2)

QtraV is the estimated video quality for audio transmission errors (see clause I.2)

MOSfromR has been defined in clause I.1. Coefficient values α, β, γ, a, b, c, d, e, f, g, h are provided in Table A.7.

Table A.7 – Audiovisual model coefficients (from [ITU-T P.1201.2])

α	β	γ	a	b	c	d	e	f	g	h
5.89	0.52	0.0045	100.0	0.32	0.9	0.705	1.02	−0.007	−0.010	−0.008

Annex B

Description of the ITU-T G.1071 model algorithm for LR application area

(This annex forms an integral part of this Recommendation.)

B.1 Audio module

The LR application area audio module is depicted in Figure B.1.

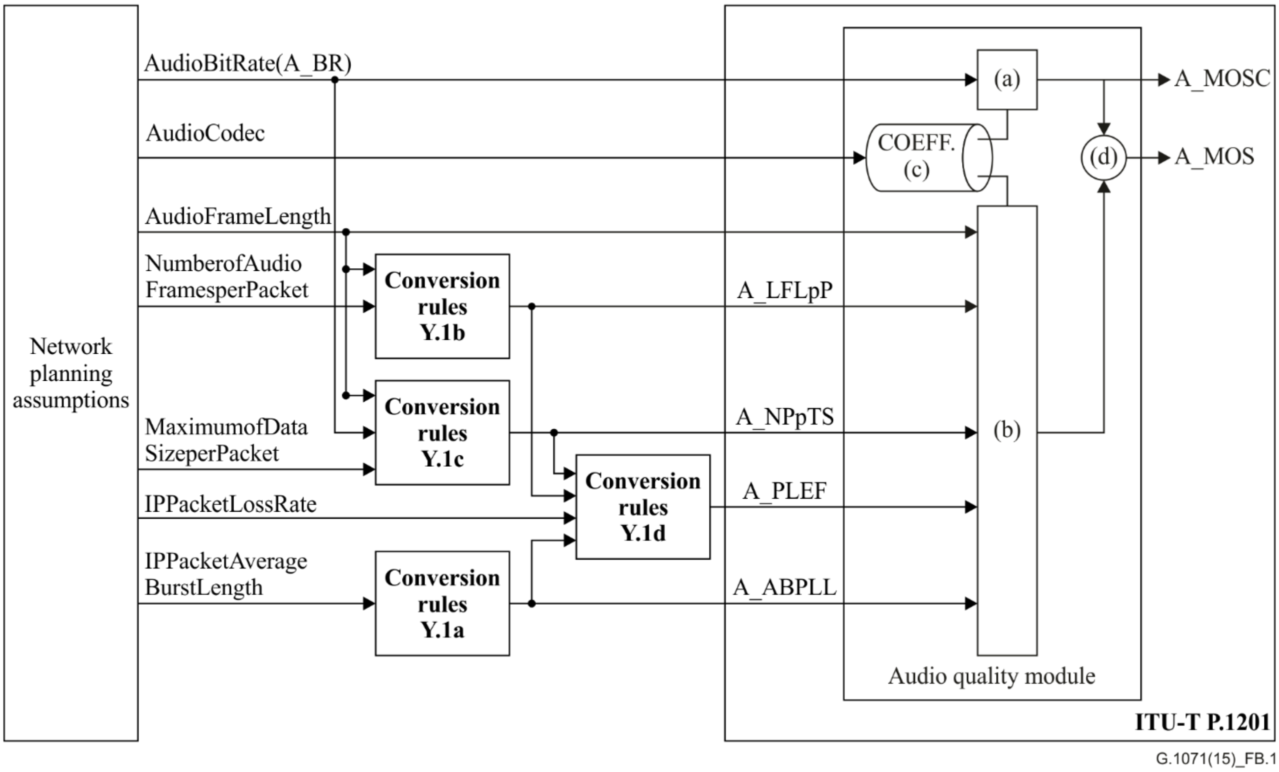

Figure B.1 – Audio module

$$A\_MOS = 1 + (A\_MOSC - 1) \times MA \qquad (4.1)$$

from [ITU-T P.1201.1]

where

$$A\_MOSC = 1 + \left(a1 - \frac{a1}{1+\left(\frac{A\_BR}{a2}\right)^{a3}}\right) \qquad (4.2)$$

from [ITU-T P.1201.1]

and

$$MA = (1 - a4)\exp\left(-\frac{10\times A\_LFL}{a5\times A\_MT}\right) + a4\exp\left(-\frac{10\times A\_LFL}{a6\times A\_MT}\right) \qquad (4.3)$$

from [ITU-T P.1201.1]

$$A\_LFL = A\_PLEF \times MAX\left(AudioFrameLength, A\_LFLpP \times \frac{A\_ABPLL + A\_NPpTS - 1}{A\_NPpTS}\right) \qquad (4.4)$$

from [ITU-T P.1201.1]

The following conversion rules are applied:

$$A\_ABPLL = IPPacketAverageBurstLength \quad (4.5a)$$

$$A\_LFLpP = AudioFrameLength \times NumberofAudioFramesperPacket \quad (4.5b)$$

$$A\_NPpTS = \left\lceil \frac{A\_BR \times AudioFrameLength}{8 \times MaximumofDataSizeperPacket} \right\rceil \quad (4.5c)$$

$$A\_PLEF = \frac{1000 \times A\_NPpTS \times IPPacketLossRate}{A\_LFLpP \times A\_ABPLL} \quad (4.5d)$$

The value of the following parameter is fixed

$$A\_MT = 1 \quad (4.5e)$$

where

A_*MOS* is the overall estimated audio quality, expressed on an MOS scale [1,5], where 1 is the worst quality and 5 is the best quality

A_*MOSC* is the estimated audio quality for audio compression artefacts

AudioBitRate(A_BR) is the audio bitrate (in kbit/s)

AudioCodec is the audio codec

MA is the degree of degradation for audio transmission errors on a scale [0,1], where 0 is no degradation for audio transmission errors and 1 is the maximum degradation for audio transmission errors

A_*MT* is the measurement time for audio (in seconds); this value is fixed to 1

A_*LFL* is the lost audio frame length per A_MT

A_*LFLpP* is the audio frame length per 1 audio RTP packet

A_*ABPLL* is the average burst IP packet-loss length for an audio stream

A_*NPpTS* is the number of audio RTP packets per 1 audio frame

A_*PLEF* is the number of audio packet-loss-events per A_MT

AudioFrameLength is the audio frame length (in milliseconds)

NumberofAudioFramesperPacket is the number of audio frames in 1 RTP packet (to be provided by a network/service planner)

IPPacketLossRate is the loss rate of IP packets

IPPacketAverageBurstLength is the average burst IP packet-loss length; this value is fixed to 1 if *IPPacketLossRate* is 0

MaximumofDataSizeperPacket is the maximum size for audio stream per 1 RTP packet (Bytes)

The coefficient values $a1$, $a2$, $a3$, $a4$, $a5$, and $a6$ depend on the audio codec. Coefficient values are provided in Table B.1.

B.1.1 Coefficient values and parameter ranges

Table B.1 – Audio model coefficients for the different audio codecs (from [ITU-T P.1201.1])

Audio codec	a1	a2	a3	a4	a5	a6
AAC-LC	3.36209	16.46062	2.08184	0.352	508.83419	37.78354
AAC-HEv1	3.19135	4.17393	1.28241	0.68955	6795.99773	186.76692
AAC-HEv2	3.13637	7.45884	2.15819	0.61993	3918.639	153.3399
AMR-NB	1.33483	6.42499	3.49066	0	723.3661	1
AMR-WB+	3.19158	5.7193	1.63208	0	826.7936	1

B.2 Video module

The LR application area video module is depicted in Figure B.2.

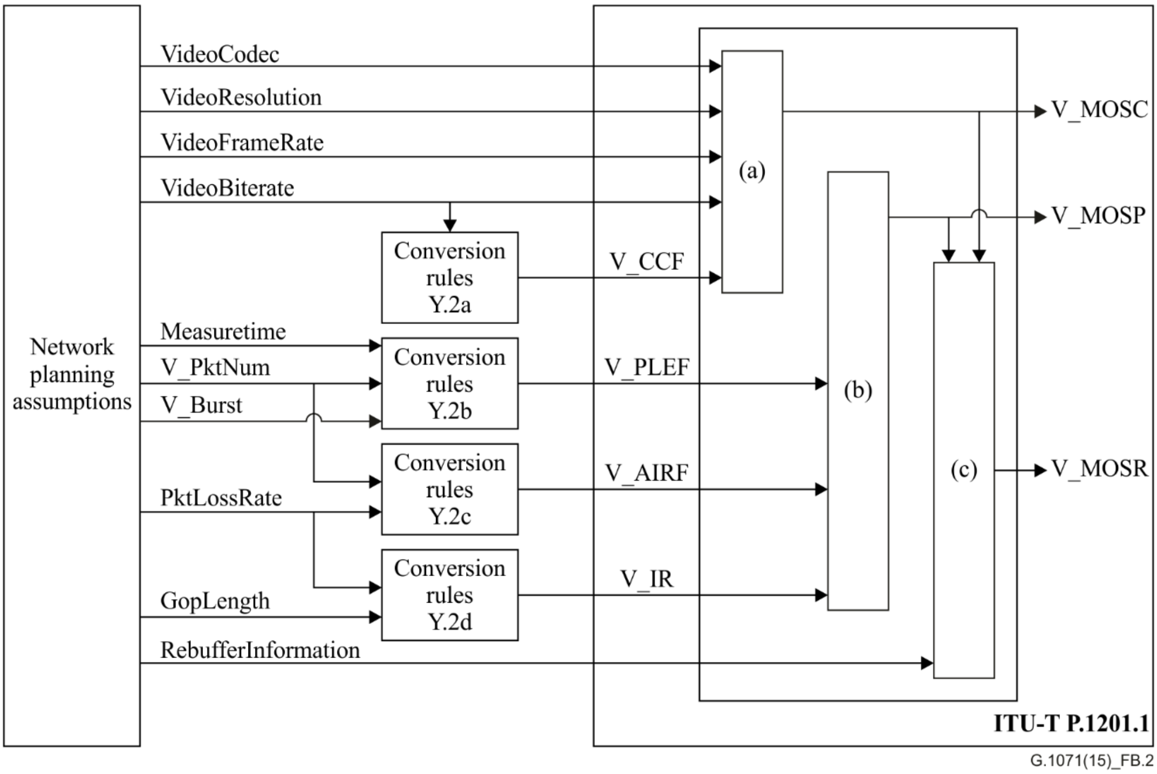

Figure B.2 – Video module

The final video quality is estimated as follows:

IF No Packet-loss AND No rebuffering THEN

$$V\_MOS = V\_MOSC \quad (5.1)$$

from [ITU-T P.1201.1]

ELSEIF Packet-loss AND No rebuffering

$$V\_MOS = V\_MOSP \quad (5.2)$$

from [ITU-T P.1201.1]

ELSE

$$V\_MOS = V\_MOSR \quad (5.3)$$

from [ITU-T P.1201.1]

ENDIF

V_MOSC is calculated as follows:

IF videoFrameRate >= 24 THEN

$$V\_MOSC = MOS\_MAX - V\_DC \quad (5.4)$$

from [ITU-T P.1201.1]

ELSE

$$V\_MOSC = (MOS\_MAX - V\_DC) \cdot \left(1 + v1 \cdot V\_CCF - v2 \cdot V\_CCF \cdot \log\left(\frac{1000}{videoFrameRate}\right)\right) \tag{5.5}$$

from [ITU-T P.1201.1]

ENDIF

V_DC is calculated as follows:

$$V\_DC = \frac{MOS\_MAX - MOS\_MIN}{1 + \left(\frac{V\_NBR}{v3 \cdot V\_CCF + v4}\right)^{(v5 \cdot V\_CCF + v6)}} \tag{5.6}$$

from [ITU-T P.1201.1]

V_MOSP is calculated as follows:

$$V\_MOSP = V\_MOSC - V\_DP \tag{5.7}$$

from [ITU-T P.1201.1]

IF videoPLC==SLICING

$$V\_DP = (V\_MOSC - MOS\_MIN) \cdot \frac{\left(\frac{V\_AIRF \cdot V\_IR}{v7 \cdot V\_CCF + v8}\right)^{v9} \cdot \left(\frac{V\_PLEF}{v10 \cdot V\_CCF + v11}\right)^{v12}}{1 + \left(\frac{V\_AIRF \cdot V\_IR}{v7 \cdot V\_CCF + v8}\right)^{v9} \cdot \left(\frac{V\_PLEF}{v10 \cdot V\_CCF + v11}\right)^{v12}} \tag{5.8}$$

from [ITU-T P.1201.1]

IF videoPLC==FREEZING

$$V\_DP = (V\_MOSC - MOS\_MIN) \cdot \frac{\left(\frac{V\_IR}{v7 \cdot V\_CCF + v8}\right)^{v9} \cdot \left(\frac{V\_PLEF}{v10 \cdot V\_CCF + v11}\right)^{v12}}{1 + \left(\frac{V\_IR}{v7 \cdot V\_CCF + v8}\right)^{v9} \cdot \left(\frac{V\_PLEF}{v10 \cdot V\_CCF + v11}\right)^{v12}} \tag{5.9}$$

from [ITU-T P.1201.1]

V_MOSR is calculated as follows:

$$V\_MOSR = Video\_Quality - V\_DR \tag{5.10}$$

from [ITU-T P.1201.1]

$$V\_DR = (Video\_Quality - MOS\_MIN) \cdot \frac{\left(\frac{NRE}{v13}\right)^{v14} \cdot \left(\frac{ARL}{v15}\right)^{v16} \cdot \left(\frac{MREEF}{v17}\right)^{v18}}{1 + \left(\frac{NRE}{v13}\right)^{v14} \cdot \left(\frac{ARL}{v15}\right)^{v16} \cdot \left(\frac{MREEF}{v17}\right)^{v18}} \tag{5.11}$$

from [ITU-T P.1201.1]

IF rebuffering AND packet-loss THEN

$$Video\_Quality = V\_MOSP \tag{5.12}$$

from [ITU-T P.1201.1]

ELSE

$$Video\_Quality = V\_MOSC \tag{5.13}$$

from [ITU-T P.1201.1]

ENDIF

The conversion rule of *V_CCF*

$$V\_CCF = \begin{cases} a \ln V\_BR + b & if\ V\_BR < Threshold \\ c & if\ V\_BR \geq Threshold \end{cases} \tag{5.5a}$$

$$Threshold = d \cdot Width \cdot Height$$

Table B.2 – Coefficient values for the V_CCF parameter

	ITU-T H.264			MPEG4	
	QCIF	**QVGA**	**HVGA**	**QCIF**	**QVGA**
a	0.1077	0.0975	0.0908	0.1155	0.1129
b	0.0207	0.0001	0.0001	0.0994	0.0931
c	0.91	0.85	0.86	0.90	0.90
d	0.02			0.012	

The conversion rule of *V_PLEF*

IF V_ratio<1 THEN

$$V\_PLEF = \frac{TotalPktNum \cdot V\_LossRate}{V\_Burst}$$

ELSE

$$V\_PLEF = \frac{TotalPktNum \cdot V\_LossRate}{V\_Burst} \cdot \frac{V\_Burst}{V\_PktpF} = \frac{TotalPktNum \cdot V\_LossRate}{V\_PktpF} \tag{5.5b}$$

Where *TotalPktNum* is to be provided by the network/service planner otherwise it can be derived from the video bitrate and measurement time as follows:

$$TotalPktNum_{tmp} = \frac{V\_BR*1000*MeasureTime}{1000*8}$$

IF

$$TotalPktNum_{tmp} < TotalFrameNum \text{ THEN}$$

$$TotalPktNum = TotalFrameNum + \frac{TotalPktNum_{tmp}}{10}$$

ELSE

$$TotalPktNum = TotalPktNum_{tmp}$$

Where *V_Ratio* can be derived from the *V_Burst* and *V_PktpF* as follows:

$$V\_ratio = \frac{V\_Burst}{V\_PktpF}$$

Where V_PktpF can be derived from *TotalPktNum* and *TotalFrameNum* as follows:

$$V\_PktpF = \frac{TotalPktNum}{TotalFrameNum} = \frac{TotalPktNum}{FrameRate \cdot MeasureTime}$$

Where V_AIRF can be derived from *V_LossRate* and *V_PktpF* as follows:

$$V\_AIRF = \frac{1}{1-(1-V\_LossRate)^{V\_PktpF}} - \frac{1-V\_LossRate}{V\_LossRate \cdot V\_PktpF} \qquad (5.5c)$$

Where *V_IR* can be derived from *V_LossRateFrame* and *GopLength* as follows:

$$V\_IR = 1 - \frac{(1-V\_LossRateFrame)}{V\_LossRateFrame \cdot GopLength} \cdot \left(1-(1-V\_LossRateFrame)^{GopLength}\right) \qquad (5.5d)$$

Where

$$V\_LossRateFrame = 1-(1-V\_LossRate)^{V\_PktpF}$$

Where

V_MOS	is the final video quality, expressed on an MOS scale [1,5], where 1 is the worst quality and 5 is the best quality
V_MOSC	is the video quality due to compression
V_MOSP	is the video due to packet loss
V_MOSR	is the video quality due to rebuffering
V_DC	is the video distortion quality due to compression
V_DP	is the video distortion quality due to packet-loss
V_DR	is the video distortion quality due to rebuffering
V_CCF	is the video content complexity
V_PLEF	is the video packet-loss event frequency
V_AIRF	is the average impairment rate of video frame
V_IR	is the impairment rate of video stream
VideoCodec	is the video codec
MeasureTime	is the measurement time in the pre-determined interval, in seconds
V_BR	is the bit rate of the video stream, in kbps
GopLength	is the length of GOP
V*_Burst*	is the average burst IP packet-loss length per video stream.
V*_PktpF*	is the average number of video RTP packets per video frame
TotalPktNum	is the total number of video RTP packets
Framerate	is the video frame rate, in frames per second
V*_LossRate*	is the loss rate of IP packets
V*_LossRateFrame*	is the loss rate of video frame
NRE	is the number of rebuffering events
ARL	is the average rebuffering length
MREFF	is the multiple rebuffering events effect factor
MOS_MAX	is the maximum value of video quality, i.e., 5.0
MOS_MIN	is the minimum value of video quality, i.e., 1.0
Width	is the video width
Height	is the video height

The coefficient values v1-v18 depend on the video codec and resolution, and are provided in Tables B.3 to B.6. They are identical to those in [ITU-T P.1201.1].

Table B.3 – Coefficient sets for *V_MOSC* and *V_DC* video quality estimation

	ITU-T H.264			MPEG4		
	QCIF	QVGA	HVGA	QCIF	QVGA	HVGA*
*v*1	3.4	2.49	2.505	2.43	1.6184	1.6184
*v*2	0.969	0.7094	0.7144	0.692	0.4611	0.4611
*v*3	104.0	324.0	170.0	0.01	280.0	280.0
*v*4	1.0	3.3	130.0	134.0	11.0	11.0
*v*5	0.01	0.5	0.05	0.01	1.69	1.69
*v*6	1.1	1.2	1.1	1.7	0.02	0.02

Table B.4 – Coefficient sets for *V_DP* for *SLICING* video quality estimation

	videoPLC = SLICING					
	ITU-T H.264			MPEG4		
	QCIF	QVGA	HVGA	QCIF	QVGA	HVGA*
*v*7	−0.63	−0.64	−0.05	−0.01	−0.01	−0.01
*v*8	1.4	0.81	0.42	0.99	0.76	0.76
*v*9	0.01	0.4	0.72	0.34	0.39	0.39
*v*10	−14.4	−9.0	−3.3	−0.1	−0.01	−0.01
*v*11	19.0	11.5	7.0	15.5	10.0	10.0
*v*12	1.04	0.4	0.49	0.66	0.86	0.86
* Provisional values, since this condition was not included in the test plan.						

Table B.5 – Coefficient sets for *V_DP* for *FREEZING* video quality estimation

	videoPLC = FREEZING					
	ITU-T H.264			MPEG4		
	QCIF	QVGA	HVGA	QCIF*	QVGA*	HVGA*
*v*7	−0.115	−0.05	−0.05	−0.115	−0.05	−0.05
*v*8	0.25	0.53	0.32	0.25	0.53	0.32
*v*9	2.05	0.6	0.24	2.05	0.6	0.24
*v*10	−0.7	−0.1	−0.1	−0.7	−0.1	−0.1
*v*11	1.5	11.5	1.0	1.5	11.5	1.0
*v*12	0.45	0.01	1.16	0.45	0.01	1.16
* Provisional values, since these conditions were not included in the test plan.						

Table B.6 – Coefficient sets for video rebuffering quality estimation

	Single rebuffering event			Multiple rebuffering event		
	QCIF	QVGA	HVGA	QCIF	QVGA	HVGA
$v13$	–	–	–	2.5	2.1	3.4
$v14$	0	0	0	1.1	1.8	0.79
$v15$	9.8	20.6	52.0	2.5	2.7	3.71
$v16$	0.85	0.37	0.42	0.15	0.55	0.39
$v17$	–	–	–	4.65	7.6	7.25
$v18$	0	0	0	0.35	0.05	0.1

B.3 Audiovisual module

As shown in Figure B.3, the LR application area audiovisual module of ITU-T G.1071 is identical to the audiovisual module for [ITU-T P.1201.1].

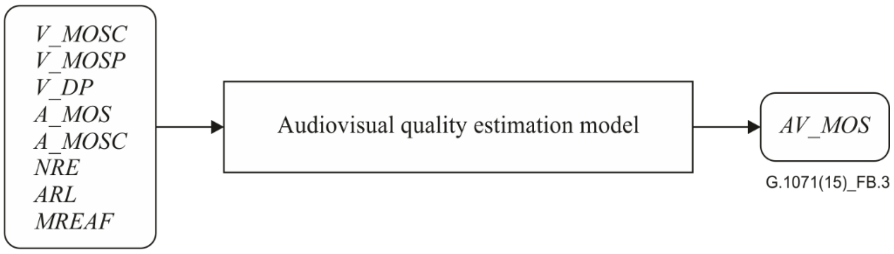

Figure B.3 – Audiovisual module

AV_MOS is calculated as shown in the pseudocode below:

IF No Packet-loss AND No rebuffering

THEN

$$AV\_MOS = AV\_MOSC \quad (6.1)$$

from [ITU-T P.1201.1]

ELSEIF Packet-loss AND No rebuffering

$$AV\_MOS = AV\_MOSP \quad (6.2)$$

from [ITU-T P.1201.1]

ELSE

$$AV\_MOS = AV\_MOSR \quad (6.3)$$

from [ITU-T P.1201.1]

ENDIF

AV_MOSC is calculated as follows:

$$AV\_MOSC = av1 \cdot V\_MOSC + av2 \cdot A\_MOSC + av3 \cdot V\_MOSC \cdot A\_MOSC + av4 \quad (6.4)$$

from [ITU-T P.1201.1]

AV_MOSP is calculated as follows:

$$AV\_MOSP = AV\_MOSC - AV\_DP \quad (6.5)$$

from [ITU-T P.1201.1]

Where *AV_DP* is calculated as follows:

$$AV\_DP = \left(AV\_MOSC - MOS\_MIN\right) \cdot AV\_DF \tag{6.6}$$

from [ITU-T P.1201.1]

where

$$AV\_DF = \frac{av5 \cdot AV\_DFV + av6 \cdot AV\_DFA}{1 + av5 \cdot AV\_DFV + av6 \cdot AV\_DFA} \tag{6.7}$$

from [ITU-T P.1201.1]

$$AV\_DFV = \frac{V\_DP}{V\_MOSC} \quad or \quad AV\_DFV = \frac{V\_MOSC - Video\_MOS}{V\_MOSC} \tag{6.8}$$

from [ITU-T P.1201.1]

$$AV\_DFA = \frac{A\_DP}{A\_MOSC} \quad or \quad AV\_DFA = \frac{A\_MOSC - Audio\_MOS}{A\_MOSC} \tag{6.9}$$

from [ITU-T P.1201.1]

IF video packet-loss of audiovisual stream has occurred THEN

$$Video\_MOS = V\_MOSP \tag{6.10}$$

from [ITU-T P.1201.1]

ELSE

$$Video\_MOS = V\_MOSC \tag{6.11}$$

from [ITU-T P.1201.1]

ENDIF

IF audio packet-loss of audiovisual stream has occurred THEN

$$Audio\_MOS = A\_MOS \tag{6.12}$$

from [ITU-T P.1201.1]

ELSE

$$Audio\_MOS = A\_MOSC \tag{6.13}$$

from [ITU-T P.1201.1]

ENDIF

AV_MOSR is calculated as follows:

$$AV\_MOSR = Audiovisual\_Quality - AV\_DR \tag{6.14}$$

from [ITU-T P.1201.1]

$$AV\_DR = \left(Audiovisual\_Quality - MOS\_MIN\right) \cdot \frac{\left(\frac{NRE}{av7}\right)^{av8} \cdot \left(\frac{ARL}{av9}\right)^{av10} \cdot \left(\frac{MREEF}{av11}\right)^{av12}}{1 + \left(\frac{NRE}{av7}\right)^{av8} \cdot \left(\frac{ARL}{av9}\right)^{av10} \cdot \left(\frac{MREEF}{av11}\right)^{av12}} \tag{6.15}$$

from [ITU-T P.1201.1]

where
IF rebuffering AND packet-loss THEN

$$Audiovisual\_Quality = AV\_MOSP \quad (6.16)$$

from [ITU-T P.1201.1]

ELSE

$$Audiovisual\_Quality = AV\_MOSC \quad (6.17)$$

from [ITU-T P.1201.1]

ENDIF

where

AV_MOS is final audiovisual MOS, expressed on the MOS scale [1,5], where 1 is the worst quality and 5 the best quality

AV_MOSC is the audiovisual quality due to compression

- AV_MOSP is the audiovisual quality due to packet-loss
- AV_MOSR is the audiovisual quality due to rebuffering
- V_MOSC is the estimated video quality due to compression
- V_MOSP is the estimated video quality due to packet loss
- V_DP is the video distortion quality due to packet-loss
- A_DP is the audio distortion quality due to packet-loss
- A_MOS is the estimated audio quality
- A_MOSC is the estimated audio quality for audio compression artefacts
- NRE is the number of rebuffering events
- ARL is the average rebuffering length
- $MREFF$ is the multiple rebuffering events effect factor
- AV_DP is the audiovisual distortion quality due to packet-loss
- AV_DFV represents the packet-loss distortion factor for video
- AV_DFA represents the packet-loss distortion factor for audio
- AV_DF represents the packet-loss distortion factor for audiovisual stream
- AV_DR is the audiovisual distortion quality due to rebuffering

Coefficient values $av1$-$av12$ are provided in Tables B.7 to B.9 and are identical to those in [ITU-T P.1201].

Table B.7 – Coefficients sets for *AV_MOSC* audiovisual quality estimation

	QCIF	QVGA	HVGA
*av*1	0.7977	0.7495	0.6419
*av*2	0.03732	0.09736	0.1362
*av*3	0.02472	0.006725	0.016
*av*4	0.1657	0.3186	0.5694

Table B.8 – Coefficient sets for *AV_DP* audiovisual quality estimation

	QCIF	QVGA	HVGA
*av*5	2.908	1.541	1.94
*av*6	0.4755	0.96	2.178

Table B.9 – Coefficient sets for audiovisual rebuffering estimation

	Single rebuffering event			Multiple rebuffering event		
	QCIF	QVGA	HVGA*	QCIF*	QVGA	HVGA*
*av*7	–	–	–	1.54	1.54	1.54
*av*8	0	0	0	0.85	0.85	0.85
*av*9	79.6	12.6	12.6	1.66	1.66	1.66
*av*10	0.3	0.26	0.26	0.45	0.45	0.45
*av*11	–	–	–	3.5	3.5	3.5
*av*12	0	0	0	0.31	0.31	0.31
* Provisional values, since these conditions were not included in the test plan.						

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	General tariff principles
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	**Transmission systems and media, digital systems and networks**
Series H	Audiovisual and multimedia systems
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Environment and ICTs, climate change, e-waste, energy efficiency; construction, installation and protection of cables and other elements of outside plant
Series M	Telecommunication management, including TMN and network maintenance
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Terminals and subjective and objective assessment methods
Series Q	Switching and signalling
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks, open system communications and security
Series Y	Global information infrastructure, Internet protocol aspects and next-generation networks
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2015