International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

J.343.4

(11/2014)

SERIES J: CABLE NETWORKS AND TRANSMISSION OF TELEVISION, SOUND PROGRAMME AND OTHER MULTIMEDIA SIGNALS

Measurement of the quality of service - Part 3

Hybrid-RR objective perceptual video quality measurement for HDTV and multimedia IP-based video services in the presence of a reduced reference signal and non-encrypted bitstream data

Recommendation ITU-T J.343.4

Recommendation ITU-T J.343.4

Hybrid-RR objective perceptual video quality measurement for HDTV and multimedia IP-based video services in the presence of a reduced reference signal and non-encrypted bitstream data

Summary

Recommendation ITU-T J.343.4 provides Hybrid-reduced reference (Hybrid-RR) objective perceptual video quality measurement methods for HDTV and multimedia when a reduced reference signal and non-encrypted bitstream data are available. The following list shows example applications that can use this Recommendation:

- potentially real-time, in-service quality monitoring at the head-end;
- video television streams over cable/IPTV networks including those transmitted over the Internet using Internet protocol;
- video quality monitoring at the receiver when a reduced reference signal and non-encrypted bitstream data are available;
- video quality monitoring at measurement nodes located between point of transmission and point of reception when a reduced reference signal and non-encrypted bitstream data are available;
- quality measurement for monitoring of a transmission system that utilizes video compression and decompression techniques, either a single pass or a concatenation of such techniques;
- lab testing of video transmission systems.

This Recommendation includes an electronic attachment containing test vectors, including video sequences, bitstream files and predicted objective model scores.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T J.343.4	2014-11-29	9	11.1002/1000/12319

* 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 2015

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
1.1	Applications	2
1.2	Limitations	2
2	References	3
3	Definitions	3
4	Abbreviations and acronyms	3
5	Conventions	4
6	Performance metrics	4
7	Description of the hybrid reduced-reference methodology	4
8	Models	5
Annex A – The YHyRR model		6
A.1	Introduction	6
A.2	Hybrid-RR video quality metric computation	6
Bibliography		12

Electronic attachment: Test vectors, including video sequences, bitstream files and predicted objective model scores.

Recommendation ITU-T J.343.4

Hybrid-RR objective perceptual video quality measurement for HDTV and multimedia IP-based video services in the presence of a reduced reference signal and non-encrypted bitstream data

1 Scope

This Recommendation[1] describes algorithmic models for measuring the visual quality of IP-based video services.

The models are hybrid-reduced reference (Hybrid-RR) models, which use features extracted from the unimpaired reference video signal, the non-encrypted bitstream data and the video image data captured at the video player.

As output, the models provide an estimate of visual quality on the [1,5] difference mean opinion score (DMOS) scale, derived from five-point absolute category rating (ACR) as in [ITU-T P.910]. The models address low resolution (VGA/WVGA) application areas, including services such as mobile TV, as well as high resolution (HD) application areas, including services such as IPTV.

This Recommendation is to be used with video encoded using ITU-T H.264 and media payload encapsulated in RTP/UDP/IP packets for low resolution video and encapsulated in MPEG-TS/RTP/UDP/IP for high resolution video.

The models in this Recommendation measure the visual effect of spatial and temporal degradations as a result of video coding, erroneous transmission, or video rescaling. The models may be used for applications such as to monitor the quality of deployed networks to ensure their operational readiness or to benchmark service quality. The models in this Recommendation can also be used for lab testing of video transmission systems.

The models identified in this Recommendation have limited precision. Therefore, directly comparing model results can be misleading. The accuracy of models has to be understood and taken into account (e.g., using [ITU-T J.149]).

The validation test material consisted of video encoded using different implementations of [ITU-T H.264]. It included media transmitted over wired and wireless networks, such as WIFI and 3G mobile networks. The transmission impairments included error conditions, such as dropped packets and packet delay, from simulations and from transmission over commercially operated networks.

The following source reference channel (SRC) conditions were included in the validation test:

- 1080i 60 Hz (29.97 fps);
- 1080p (25 fps);
- 1080i 50 Hz (25 fps);
- 1080p (29.97 fps);
- SRC duration: HD: 10 s, VGA/WVGA: 10 s or 15 s (rebuffering);
- VGA at 25 and 30 frames-per-second;
- WVGA at 25 and 30 frames-per-second.

[1] This Recommendation includes an electronic attachment containing test vectors, including video sequences, bitstream files and predicted objective model scores.

The following hypothetical reference circuit (HRC) conditions were included in the validation test for each resolution:

Test factors
Video resolution: 1920 × 1080 interlaced and progressive
Video frame rates 29.97 and 25 frames per second
Video bitrates: 1 to 30 Mbit/s (HD), 100 kbit/s to 3 Mbit/s (VGA/WVGA)
Temporal frame freezing (pausing with skipping) of up to 50% of video duration
Transmission errors with packet loss
Rebuffering (VGQ/WVGA only): up to 50% of SRC
Coding technologies
ITU-T H.264/AVC (MPEG-4 Part 10)
Tandem coding

1.1 Applications

The applications for the estimation model described in this Recommendation include, but are not limited to:

- Potentially real-time, in-service quality monitoring at the headend.
- Video television streams over cable/IPTV networks including those transmitted over the Internet using Internet Protocol.
- Video quality monitoring at the receiver when side-channels and non-encrypted bitstream data are available.
- Video quality monitoring at measurement nodes located between the point of transmission and the point of reception when side-channels and non-encrypted bitstream data are available.
- Quality measurement for monitoring of a transmission system that utilizes video compression and decompression techniques, either as a single pass or as a concatenation of such techniques.
- Lab testing of video transmission systems.

1.2 Limitations

The video quality estimation model described in this Recommendation cannot be used to fully replace subjective testing.

When frame freezing was present, the test conditions had frame freezing durations up to 50% of SRC duration. The models in this Recommendation were validated for measuring video quality in a re-buffering condition (i.e., video that has a steadily increasing delay or freezing without skipping) only for VGA/WVGA. The models were not tested on frame rates other than those used in TV systems (i.e., 29.97 frames per second and 25 frames per second, in interlaced or progressive mode).

The models in this Recommendation may not be used if forward error correction techniques are employed.

It is important that no additional transmission errors occur between the collection point of the bitstream data and the capture point of the processed video sequence (PVS).

It should be noted that the objective model may produce erroneous results if new coding and transmission technologies produce artifacts that were not included in the validation tests. Here, a subjective evaluation is required.

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 (2014), *Advanced video coding for generic audiovisual services*.

[ITU-T J.149] Recommendation ITU-T J.149 (2004), *Method for specifying accuracy and cross-calibration of Video Quality Metrics (VQM)*.

[ITU-T J.342] Recommendation ITU-T J.342 (2011), *Objective multimedia video quality measurement of HDTV for digital cable television in the presence of a reduced reference signal*.

[ITU-T J.343] Recommendation ITU-T J.343 (2014), *Hybrid perceptual/bitstream models for objective video quality measurements*.

[ITU-T P.910] Recommendation ITU-T P.910 (2008), *Subjective video quality assessment methods for multimedia applications*.

3 Definitions

This Recommendation uses the following term defined elsewhere:

3.1 hybrid reduced reference model [ITU-T J.343]: An objective video quality model that predicts subjective quality using the decoded video frames, packet headers, video payload and features extracted from the reference video. Such models can be deployed in-service but cannot analyse encrypted video.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

CODEC	Coder-Decoder
DMOS	Difference Mean Opinion Score
EPSNR	Edge Peak Signal-to-Noise Ratio
ES	Elementary bitStream
HRC	Hypothetical Reference Circuit
Hybrid-RR	Hybrid Reduced Reference model
LUT	Look-Up Table
MPEG	Moving Picture Experts Group
PES	Packetized Elementary bitStream
PVS	Processed Video Sequence
QP	Quantization Parameter
RR	Reduced Reference
STB	Set-Top Box
SRC	Source Reference Channel or Circuit

VQEG Video Quality Experts Group

VQM Video Quality Metric

VGA Video Graphics Array

WVGA Wide Video Graphics Array

5 Conventions

None.

6 Performance metrics

A summary of this and other Hybrid models may be found in [ITU-T J.343]. See [b-VQEG Hybrid] for the complete analysis of the models included in this Recommendation.

7 Description of the hybrid reduced-reference methodology

This Recommendation specifies objective video quality measurement methods which use both processed video sequences and bitstream data. The bitstream data may be provided in the forms of elementary bitstream (ES), packetized elementary bitstream (PES) or packet video (Figure 1).

In addition, Hybrid-RR models also use features extracted from source video sequences. Figure 2 shows a Hybrid-RR model. In addition to the non-encrypted bitstream data, the Hybrid-RR model uses the features extracted from the SRC.

Figure 1 – Block-diagram depicting the core concept of hybrid perceptual/bitstream models

DMOSp: predicted DMOS by the model
STB: Set-Top Box

Figure 2 – Block-diagram depicting the Hybrid-RR model

8 Models

Annex A contains a full disclosure of all models included in this Recommendation. That model is referred to as YHyRR.

Annex A

The YHyRR model

(Hybrid-RR Model)

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

A.1 Introduction

The YHyRR model first computes a video quality metric (VQM) value using the quantization parameter (QP) and the error area (log) using a pre-defined look-up-table. Then, the EPSNR value is computed using [ITU-T J.342]. Finally, post-processing is applied to reflect various impairments due to transmission errors.

A.2 Hybrid-RR video quality metric computation

A.2.1 Feature computation

A.2.1.1 Quantization parameter (QP) and error area

In ITU-T H.264 advanced video coding (AVC) the QP is an important factor for video quality. The averaged QP and I frame QP are computed as follows:

$$QP_{ave} = \frac{1}{NumFrame} \sum_{<VideoFrame>} QP[i]$$

$$QP_{Iframe} = \frac{1}{NumIframe} \sum_{<I\ frame>} QP[i]$$

where *NumFrame* is the total number of frames and *NumFrame* is the number of *I* frames.

Figure A.1 – Error areas due to transmission errors

To reflect the effects of transmission errors, an error area can be calculated using an H.264/AVC decoder (Figure A.1). The transmission error pixel may be categorized into two types. One is an error pixel due to erroneous packets and the other is the propagation of error pixels. The error pixels due to erroneous packets can be directly calculated during decoding and the propagation error pixels are calculated using the reference frame index and motion vectors. First, an error map is produced as follows:

$$ErrorMap[i, j, k] = \begin{cases} 1 & if\ Video[i, j, k] = Error \\ 0 & otherwise \end{cases}$$

where (i, j) is a spatial index and k is temporal index.

$$ErrorFrame[k] = \sum_{(i,j)\in frame} ErrorMap[i, j, k]$$

Then, isolated error frames are removed, which may not affect overall perceptual quality as follows:

$$ErrorFlag[k] = \begin{cases} 1 & if \;\; ErrorFrame[k] \times (\sum_{l=k-searchrange}^{k+searchrange} ErrorFrame[l] - ErrorFrame[k]) > 0 \\ 0 & otherwise \end{cases}$$

The error area is calculated as follows:

$$ErrorArea = \frac{1}{NumFrame \times NumPixelImage} \sum_{\{k|ErrorFlag[k]=1\}} ErrorFrame[k]$$

A.2.1.2 Green block feature

Some videos may contain mono-color blocks due to severe transmission errors. A feature (*Greenblk*) reflecting this impairment is computed as follows:

$$Uzero_{pixel}(i, j, k) = \begin{cases} 1 & U(i, j, k) = 0 \\ 0 & otherwise \end{cases}$$

$$Vzero_{pixel}(i, j, k) = \begin{cases} 1 & V(i, j, k) = 0 \\ 0 & otherwise \end{cases}$$

$$Uzero_{line}(j, k) = \sum_{i=1}^{width} Uzero_{pixel}(i, j, k)$$

$$Vzero_{line}(j, k) = \sum_{i=1}^{width} Vzero_{pixel}(i, j, k)$$

$$Uzero_{flag}(j, k) = \begin{cases} 1 & Uzero_{line}(j, k) > width/8 \\ 0 & otherwise \end{cases}$$

$$Vzero_{flag}(j, k) = \begin{cases} 1 & Vzero_{line}(j, k) > width/8 \\ 0 & otherwise \end{cases}$$

$$Uzero = \sum_{k=1}^{NumFrame} \sum_{j=1}^{height} Uzero_{flag}(j, k)$$

$$Vzero = \sum_{k=1}^{NumFrame} \sum_{j=1}^{height} Vzero_{flag}(j, k)$$

$$Greenblk = \frac{1}{NumFrame}(Uzero + Vzero)$$

Here, *U* is the *u* channel and *V* is the *v* channel in the yuv video format.

A.2.1.3 Freeze features

To compute a freeze feature, the frame difference is calculated using the luminance channel:

$$FrameDiff(k) = \frac{1}{FramePixelSize} \sum_{(i,j)} |Y(i, j, k) - Y(i, j, k-1)|$$

$$FreezeFlag(k) = \begin{cases} 1 & if \quad FrameDiff(k) < Th_{frz} \\ 0 & otherwise \end{cases}$$

$$FRZ_{total} = \sum_k FreezeFlag(k)$$

$$FreezeFlag_{occur}(k) = \begin{cases} 1 & if \quad FreezeFlag(k) = 1 \ and \ FreezeFlag(k-1) = 0 \\ 0 & otherwise \end{cases}$$

$$FRZ_{num} = \sum_k FreezeFlag_{occur}(k)$$

The maximum freeze length "$FRZ_{\max}$" is also calculated, which is the longest freeze interval. Thus three freeze features are calculated: FRZ_{total}, FRZ_{num}, $FRZ_{\max}$.

A.2.2 Video quality metric computation

A.2.2.1 Video quality metric computation using non-encrypted bitstream data and look-up table

A predefined look-up table (LUT) is used for VQM calculation. The two-dimensional LUT has the x-axis for QP values and the y-axis for the log error area. The inputs to LUT are computed as follows:

$$X = QP_{ave} + QP_{Iframe}$$
$$Y = \log_{10}(ErrorArea + 1)$$

Using a LUT, $HRR1_{enc}$ is computed as follows:

$$HNR1 = LUT(X, Y)$$

$$HNR1 = \max(HNR1, 1.0)$$

The function *LUT* uses the bilinear interpolation. The LUT for HD and the LUT for VGA/WVGA are provided electronically.

A.2.2.2 Resizing

To consider resized video, the video resolution of the bitstream data is also considered (Figure A.2). If the video is reduced before encoding, $HNR1$ is adjusted as follows:

$$ImageSize = width \times height$$
$$ImageSize_{Bitstream} = width_{codec} \times height_{codec}$$

$$HNR_2 = \begin{cases} b_1 \times log_{10}(HNR_1) + 1 & if \ ImageSize_{Bitstream} < ImageSize \times r_1 \\ b_2 \times log_{10}(HNR_1) + 1 & else\ if \ ImageSize_{Bitstream} < ImageSize \times r_2 \\ b_3 \times log_{10}(HNR_1) + 1 & else\ if \ ImageSize_{Bitstream} < ImageSize \times r_3 \\ b_4 \times log_{10}(HNR_1) + 1 & else\ if \ ImageSize_{Bitstream} < ImageSize \times r_4 \\ HNR_1 & otherwise \end{cases}$$

where ($r_1 = \frac{1}{8}, r_2 = \frac{1}{4}, r_3 = \frac{4}{9}, r_4 = \frac{1}{2}, b_1 = 2.5, b_2 = 3.5, b_3 = 4.0, b_4 = 4.5$).

Figure A.3 shows the mapping function for resizing.

Figure A.2 – Resizing examples

Figure A.3 – Mapping function for resizing

A.2.2.3 EPSNR

The EPSRR value ($EPSNR$) is computed without the post-processing using [ITU-T J.342] at the corresponding side channel bandwidth.

A.2.3 Post-processing

A.2.3.1 Post-processing for video quality metric computed using non-encrypted bitstream data and look-up table

The VQM value computed using non-encrypted bitstream data and LUT ($HNR1$) is used as input (vqm_{in}) for the post-processing in this section.

First, the green block impairment is reflected as follows:

$$vqm_1 = \begin{cases} MIN(vqm_{in}, 1.6) & if \quad Greenblk > 1.0 \\ MIN(vqm_{in}, 2.5) & if \quad Greenblk > 0.0 \\ vqm_{in} & otherwise \end{cases}$$

Second, the frame rate is considered only for VGA/WVGA as follows:

$$vqm_2 = \begin{cases} MIN(vqm_1, 3.2) & if \quad fps < 6 \\ MIN(vqm_1, 3.5) & if \quad fps < 10 \\ vqm_1 & otherwise \end{cases}$$

Finally, the freeze impairment is reflected as follows:

$$FRZ_{temp} = \begin{cases} FRZ_{total} - \left(1 - \frac{fps}{fps_{original}}\right) \times fps_{original} \times VideoSec & if \;\; Resolution = VGA \;\; or \;\; Resolution = WVGA \\ FRZ_{total} & otherwise \end{cases}$$

$$FRZ_{\log} = MIN(\log_{10}(FRZ_{temp} + 1.0), \; 2.3)$$

$$vqm_{out} = \begin{cases} MIN(vqm_2, 4 - \log_{10}(FRZ_{\log} - 0.3) \times 3.8) & if \;\; FRZ_{\log} > 1.3 \\ vqm_2 & otherwise \end{cases}$$

A.2.3.2 Post-Processing for the EPSNR value

The EPSNR value ($EPSNR$) is used as input (vqm_{in}) for the post-processing in this section.

First, upper and lower bounds are applied with scaling as follows:

$$vqm_1 = \begin{cases} 1.4 & if \;\; vqm_{in} < 23 \\ 4.4 & if \;\; vqm_{in} > 42 \\ \frac{3}{19}(vqm_{in} - 23) + 1.4 & otherwise \end{cases}$$

Second, the error rate is considered as follows:

$$ErrorRate = \frac{TotalPacket_{Loss}}{TotalPacket + TotalPacket_{Loss}}$$

$$vqm_2 = \begin{cases} vqm_1 - 1 & if \;\; ErrorRate > 0.01 \\ vqm_1 & otherwise \end{cases}$$

Third, the error area is considered as follows:

$$vqm_3 = \begin{cases} MIN(vqm_2, 1.5) & if \;\; ErrorArea > 20 \\ MIN(vqm_2, 2.0) & if \;\; ErrorArea > 10 \\ MIN(vqm_2, 3.0) & if \;\; ErrorArea > 5 \\ vqm_2 & otherwise \end{cases}$$

Finally, the number of freeze, maximum freeze length and the total freeze are considered as follows:

$$vqm_4 = \begin{cases} MIN(vqm_3, 2.2) & if \;\; FRZ_{\max} > 100 \\ MIN(vqm_3, 2.5) & if \;\; FRZ_{\max} > 50 \\ vqm_3 & otherwise \end{cases}$$

$$vqm_5 = \begin{cases} MIN(vqm_4, 2.5) & if \;\; FRZ_{total} > 100 \\ MIN(vqm_4, 3.0) & if \;\; FRZ_{total} > 60 \\ vqm_4 & otherwise \end{cases}$$

$$EPSNR_{out} = \begin{cases} MIN(vqm_5, 2.5) & if \;\; FRZ_{num} > 12 \\ vqm_5 & otherwise \end{cases}$$

A.2.3.3 Final video quality metric

The final VQM is computed by averaging the EPSFR value ($EPSNR_{out}$) and vqm_{out} of the previous sections as follows:

$$VQM_{final} = \frac{1}{2}(EPSNR_{out} + vqm_{out})$$

Bibliography

[b-VQEG Hybrid] Video Quality Experts Group (2014), *Hybrid Perceptual/Bitstream Validation Test Final Report*.

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	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