

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

J.246

(08/2008)

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

Measurement of the quality of service

**Perceptual visual quality measurement
techniques for multimedia services over digital
cable television networks in the presence of a
reduced bandwidth reference**

Recommendation ITU-T J.246

Recommendation ITU-T J.246

Perceptual visual quality measurement techniques for multimedia services over digital cable television networks in the presence of a reduced bandwidth reference¹

Summary

The term multimedia as defined in Recommendation ITU-T J.148 is the combination of multiple forms of media such as: video, audio, text, graphics, fax, and telephony in the communication of information. A three stage approach has been adopted to recommending objective assessment methods for multimedia. The first two stages will identify perceptual quality tools appropriate for measuring video and audio individually. The third stage will identify objective assessment methods for the combined audiovisual media. This Recommendation contains the first stage video only used in multimedia applications.

Recommendation ITU-T J.246 provides guidelines on the selection of appropriate objective perceptual video quality measurement methods when a reduced reference signal is available. The following are example applications that can use this Recommendation:

- 1) Internet multimedia streaming
- 2) Video telephony and conferencing over cable and other networks
- 3) Progressive video television streams viewed on LCD monitors over cable networks including those transmitted over the Internet using Internet Protocol. (VGA was the maximum resolution in the validation test.)
- 4) Mobile video streaming over telecommunications networks
- 5) Some forms of IPTV video payloads (VGA was the maximum resolution in this validation test.)
- 6) Video quality monitoring at the receiver when side-channels are available.

Source

Recommendation ITU-T J.246 was approved on 13 August 2008 by ITU-T Study Group 9 (2005-2008) under Recommendation ITU-T A.8 procedures.

¹ The name was changed to reflect the fact that this Recommendation covers video only.

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 2010

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

CONTENTS

Page

1	Scope	1
1.1	Application	2
1.2	Limitations.....	2
2	References.....	3
2.1	Normative references.....	3
2.2	Informative references.....	3
3	Definitions	4
3.1	Terms defined elsewhere	4
3.2	Terms defined in this Recommendation.....	4
4	Abbreviations and acronyms	5
5	Conventions	6
6	Description of the reduced reference measurement method.....	6
7	Findings of the Video Quality Experts Group (VQEG)	7
	Annex A – Yonsei University Reduced Reference Method	10
A.1	Introduction	10
A.2	The EPSNR Reduced Reference Models	10
A.3	Conclusions	22
	Appendix I – Optimal side-channel bandwidths.....	23
	Appendix II – Excerpts from the Synopsis from the Video Quality Experts Group on the validation of objective models of multimedia quality assessment, phase I.....	26
II.1	Introduction	26
II.2	Model performance evaluation techniques.....	27
II.3	RR model performance.....	27
II.4	Data analysis executed by ILG.....	29
	Appendix III – Equations for Model Evaluation Metrics	30
III.1	Evaluation Metrics.....	30
	Bibliography.....	33

Copyright International Telecommunication Union/ITU Telecommunication Sector

Recommendation ITU-T J.246

Perceptual visual quality measurement techniques for multimedia services over digital cable television networks in the presence of a reduced bandwidth reference

1 Scope

This Recommendation provides guidelines and recommendations on the selection of appropriate perceptual video quality measurement equipment for use in multimedia applications when the reduced reference measurement method can be used.

The reduced reference measurement method can be used when features extracted from the unimpaired reference video signal are readily available at the measurement point, as may be the case of measurements on individual equipment or a chain in the laboratory or in a closed environment such as a cable television head-end. The estimation methods are based on processing video in VGA, CIF, and QCIF resolution.

The validation test material contained both multiple coding degradations and various transmission error conditions (e.g., bit errors, dropped packets). In the case where coding distortions are considered in the video signals, the encoder can utilize various compression methods (e.g., MPEG-2, H.264, etc.). The models proposed in this Recommendation may be used to monitor the quality of deployed networks to ensure their operational readiness. The visual effects of the degradations may include spatial as well as temporal degradations (e.g., frame repeats, frame skips, frame rate reduction). The models in this Recommendation can also be used for lab testing of video systems. When used to compare different video systems, it is advisable to use a quantitative method (such as that in [ITU-T J.149]) to determine the models' accuracy for that particular context.

This Recommendation is deemed appropriate for telecommunications services delivered at 4 Mbit/s or less presented on mobile/PDA and computer desktop monitors. The following conditions were allowed in the validation test for each resolution:

- PDA/Mobile (QCIF): 16 kbit/s to 320 kbit/s
- CIF: 64 kbit/s-2 Mbit/s (C01 has several 2 Mbit/s)
- VGA: 128 kbit/s-4 Mbit/s (V13 has one HRC with 6 Mbit/s)

Table 1 – Factors for which J.246 has been evaluated

Test factors
Transmission errors with packet loss
Video resolution QCIF, CIF and VGA
Video bit rates • QCIF: 16 kbit/s to 320 kbit/s • CIF: 64 kbit/s-2 Mbit/s • VGA: 128 kbit/s-4 Mbit/s
Temporal errors (pausing with skipping) of maximum 2 seconds
Video frame rates from 5 fps to 30 fps
Coding technologies
H.264/AVC (MPEG-4 Part 10), VC-1, Windows Media 9, Real Video (RV 10), MPEG-4 Part 2. See Note 1.

Table 1 – Factors for which J.246 has been evaluated

Applications
Real-time, in-service quality monitoring at the source
Remote destination quality monitoring when side-channels are available for features extracted from source video sequences
Quality measurement for monitoring of a storage or transmission system that utilizes video compression and decompression techniques, either a single pass or a concatenation of such techniques
Lab testing of video systems
NOTE 1 – The validation testing of models included video sequences encoded using 15 different video codecs. The five codecs listed in this table were most commonly applied to encode test sequences and any recommended models may be considered appropriate for evaluating these codecs. In addition to these five codecs a smaller proportion of test sequences were created using the following codecs: Cinepak, DivX, H.261, H.263, H.263+ (Note 2), JPEG-2000, MPEG-1, MPEG-2, Sorenson, H.264 SVC, Theora. It can be noted that some of these codecs were used only for CIF and QCIF resolutions because they are expected to be used in the field mostly for these resolutions. Before applying a model to sequences encoded using one of these codecs the user should carefully examine its predictive performance to determine whether the model reaches acceptable predictive performance. NOTE 2 – H.263+ is a particular configuration of H.263 (1998).

1.1 Application

This Recommendation provides video quality estimations for video classes TV3 to MM5B, as defined in Annex B of [ITU-T P.911]. Note that the maximum resolution was VGA and the maximum bit rate covered well in the test was 4 Mbit/s. The applications for the estimation models described in this Recommendation include but are not limited to:

- 1) potentially real-time, in-service quality monitoring at the source;
- 2) remote destination quality monitoring when side-channels are available for features extracted from source video sequences;
- 3) quality measurement for monitoring of a storage or transmission system that utilizes video compression and decompression techniques, either a single pass or a concatenation of such techniques;
- 4) lab testing of video systems.

1.2 Limitations

The estimation models described in this Recommendation cannot be used to fully replace subjective testing. Correlation values between two carefully designed and executed subjective tests (i.e., in two different laboratories) normally fall within the range 0.95 to 0.98. If this Recommendation is utilized to make video system comparisons (e.g., comparing two codecs), it is advisable to use a quantitative method (such as that in [ITU-T J.149]) to determine the models' accuracy for that particular context.

The models in this Recommendation were validated by measuring video that exhibits frame freezes up to 2 seconds.

The models in this Recommendation were not validated for measuring video that has a steadily increasing delay (e.g., video which does not discard missing frames after a frame freeze).

It should be noted that in case of new coding and transmission technologies producing artifacts which were not included in this evaluation, the objective models may produce erroneous results. Here a subjective evaluation is required.

NOTE – The structure and content of this Recommendation have been organized for ease of use by those familiar with the original source material; as such, the usual style of ITU-T recommendations has not been applied.

2 References

2.1 Normative 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 J.143] Recommendation ITU-T J.143 (2000), *User requirements for objective perceptual video quality measurements in digital cable television.*
- [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.*

2.2 Informative references

- [ITU-T H.261] Recommendation ITU-T H.261 (1993), *Video codec for audiovisual services at $p \times 64$ kbits.*
- [ITU-T H.263] Recommendation ITU-T H.263 (1996), *Video coding for low bit rate communication.*
- [ITU-T H.263+] Recommendation ITU-T H.263 (1998), *Video coding for low bit rate communication (H.263+).*
- [ITU-T H.264] Recommendation ITU-T H.264 (2003), *Advanced video coding for generic audiovisual services.*
- [ITU-T J.144] Recommendation ITU-T J.144 (2001), *Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference.*
- [ITU-T J.148] Recommendation ITU-T J.148 (2003), *Requirements for an objective perceptual multimedia quality model.*
- [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.244] Recommendation ITU-T J.244 (2008), *Calibration methods for constant misalignment of spatial and temporal domains with constant gain and offset.*
- [ITU-T P.931] Recommendation ITU-T P.931 (1998), *Multimedia communications delay, synchronization and frame rate measurement.*
- [ITU-R BT.500-11] Recommendation ITU-R BT.500-11 (in force), *Methodology for the subjective assessment of the quality of television pictures.*

[VQEG] Final report from the video quality experts group on the validation of objective models of multimedia quality-Phase I, 2008.
ftp://vqeg.its.bldrdoc.gov/Documents/Projects/multimedia/MM_Final_Report/VQEG_MM_Report_Final_v2.6.pdf

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 objective perceptual measurement (picture): [ITU-T J.144]

3.1.2 Proponent: [ITU-T J.144]

3.1.3 subjective assessment (picture): [ITU-T J.144]

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.2.1 anomalous frame repetition: An event where the HRC outputs a single frame repeatedly in response to an unusual or out of the ordinary event. Anomalous frame repetition includes but is not limited to the following types of events: an error in the transmission channel, a change in the delay through the transmission channel, limited computer resources impacting the decoder's performance, and limited computer resources impacting the display of the video signal.

3.2.2 constant frame skipping: An event where the HRC outputs frames with updated content at an effective frame rate that is fixed and less than the source frame rate.

3.2.3 effective frame rate: Number of unique frames (i.e., total frames – repeated frames) per second.

3.2.4 frame rate: Number of (progressive) frames displayed per second (fps).

3.2.5 intended frame rate: Number of video frames per second physically stored for some representation of a video sequence. The intended frame rate may be constant or may change with time. Two examples of constant intended frame rates are a BetacamSP tape containing 25 fps and a VQEG FR-TV Phase I compliant 625-line YUV file containing 25 fps; these both have an intended frame rate of 25 fps. One example of a variable intended frame rate is a computer file containing only new frames; in this case the intended frame rate exactly matches the effective frame rate. The content of video frames is not considered when determining intended frame rate.

3.2.6 live network conditions: Errors imposed upon the digital video bit stream as a result of live network conditions. Examples of error sources include packet loss due to heavy network traffic, increased delay due to transmission route changes, multi-path on a broadcast signal, and fingerprints on a DVD. Live network conditions tend to be unpredictable and unrepeatable.

3.2.7 pausing with skipping: Events where the video pauses for some period of time and then restarts with some loss of video information. In pausing with skipping, the temporal delay through the system will vary about an average system delay, sometimes increasing and sometimes decreasing. One example of pausing with skipping is a pair of IP Videophones, where heavy network traffic causes the IP Videophone display to freeze briefly; when the IP Videophone display continues, some content has been lost. Another example is a videoconferencing system that performs constant frame skipping or variable frame skipping. Constant frame skipping and variable frame skipping are subsets of pausing with skipping. A processed video sequence containing pausing with skipping will be approximately the same duration as the associated original video sequence.

3.2.8 pausing without skipping: Any event where the video pauses for some period of time and then restarts without losing any video information. Hence, the temporal delay through the system must increase. One example of pausing without skipping is a computer simultaneously downloading and playing an AVI file, where heavy network traffic causes the player to pause briefly and then continue playing. A processed video sequence containing pausing without skipping events will always be longer in duration than the associated original video sequence.

3.2.9 refresh rate: The rate at which the computer monitor is updated.

3.2.10 simulated transmission errors: Errors imposed upon the digital video bit stream in a highly controlled environment. Examples include simulated packet loss rates and simulated bit errors. Parameters used to control simulated transmission errors are well defined.

3.2.11 source frame rate (SFR): The intended frame rate of the original source video sequences. The source frame rate is constant. For the VQEG MM Phase I test the SFR was either 25 fps or 30 fps.

3.2.12 transmission errors: Any error imposed on the video transmission. Example types of errors include simulated transmission errors and live network conditions.

3.2.13 variable frame skipping: An event where the HRC outputs frames with updated content at an effective frame rate that changes with time. The temporal delay through the system will increase and decrease with time, varying about an average system delay. A processed video sequence containing variable frame skipping will be approximately the same duration as the associated original video sequence.

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

ACR	Absolute Category Rating
ACR-HR	Absolute Category Rating with Hidden Reference
AVI	Audio Video Interleave
CIF	Common Intermediate Format (352 × 288 pixels)
DMOS	Difference Mean Opinion Score
FR	Full Reference
FRTV	Full Reference TeleVision
HRC	Hypothetical Reference Circuit
ILG	VQEG's Independent Laboratory Group
LCD	Liquid Crystal Display
MM	Multimedia
MOS	Mean Opinion Score
MOSp	Mean Opinion Score, predicted
NR	No (or Zero) Reference
PDA	Personal Digital Assistant
PSNR	Peak Signal to Noise Ratio
PVS	Processed Video Sequence
QCIF	Quarter Common Intermediate Format (176 × 144 pixels)

RMSE	Root Mean Square Error
RR	Reduced Reference
SFR	Source Frame Rate
SRC	Source Reference Channel or Circuit
VGA	Video Graphics Array (640 × 480 pixels)
VQEG	Video Quality Experts Group
VQM	Video Quality Metric
YUV	Color Space and file format

5 Conventions

None.

6 Description of the reduced reference measurement method

The double-ended measurement method with reduced reference, for objective measurement of perceptual video quality, evaluates the performance of systems by making a comparison between features extracted from the undistorted input, or reference, video signal at the input of the system, and the degraded signal at the output of the system (Figure 1).

Figure 1 shows an example of application of the reduced reference method to test a codec in the laboratory.

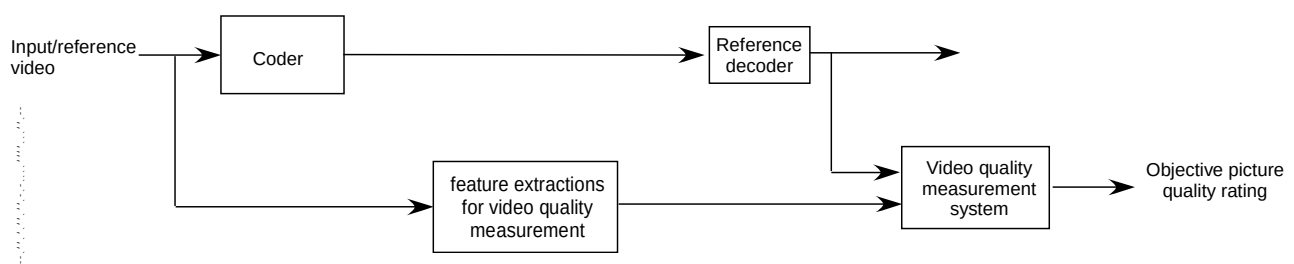


Figure 1 – Application of the reduced reference perceptual quality measurement method to test a codec in the laboratory

The comparison between input and output signals may require a temporal alignment or a spatial alignment process, the latter to compensate for any vertical or horizontal picture shifts or cropping. It also may require correction for any offsets or gain differences in both the luminance and the chrominance channels. The objective picture quality rating is then calculated, typically by applying a perceptual model of human vision.

Alignment and gain adjustment is known as registration. This process is required because most reduced reference methods compare the features extracted from reference pictures and processed pictures on what is effectively a pixel-by-pixel basis. The video quality metrics described in Annex A include registration methods.

As the video quality metrics are typically based on approximations to human visual responses, rather than on the measurement of specific coding artefacts, they are in principle equally valid for analogue systems and for digital systems. They are also in principle valid for chains where analogue and digital systems are mixed, or where digital compression systems are concatenated.

Figure 2 shows an example of the application of the reduced reference method to test a transmission chain.

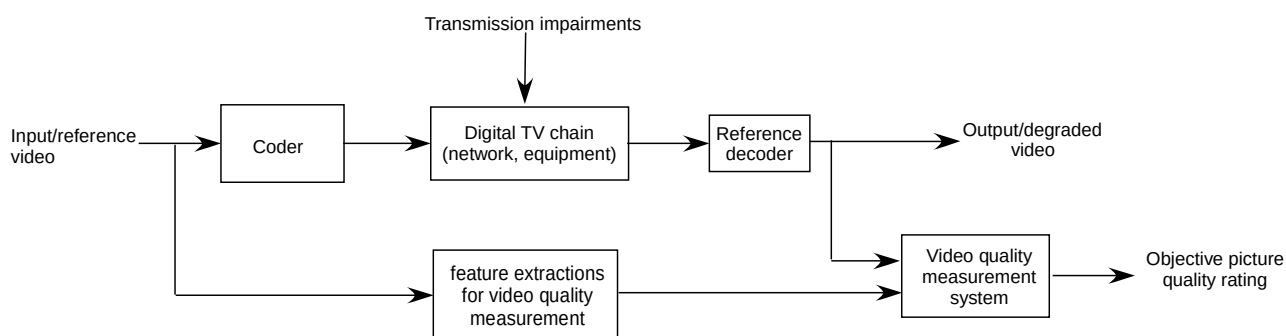


Figure 2 – Application of the reduced reference perceptual quality measurement method to test a transmission chain

In this case, a reference decoder is fed from various points in the transmission chain, e.g., the decoder can be located at a point in the network, as in Figure 2, or directly at the output of the encoder as in Figure 1. If the digital transmission chain is transparent, the measurement of objective picture quality rating at the source is equal to the measurement at any subsequent point in the chain.

It is generally accepted that the full reference method provides the best accuracy for perceptual picture quality measurements. The method has been proven to have the potential for high correlation with subjective assessments made in conformity with the ACR-HR methods specified in [ITU-T P.910].

7 Findings of the Video Quality Experts Group (VQEG)

Studies of perceptual video quality measurements are conducted in an informal group, called Video Quality Experts Group (VQEG), which reports to ITU-T Study Groups 9 and 12 and ITU-R Study Group 6. The recently completed Multimedia Phase I test of VQEG assessed the performance of proposed reduced reference perceptual video quality measurement algorithms for QCIF, CIF, and VGA formats.

Based on present evidence, the following method can be recommended by ITU-T at this time:

Annex A – VQEG Proponent: Yonsei University, Korea

The technical descriptions of this model can be found in Annex A.

Table 2 below provides informative details on the models' performances in the VQEG Multimedia Phase I test.

Table 2 – VGA resolution: Informative description on the models' performances in the VQEG Multimedia Phase I test: Averages over 13 subjective tests

Statistic	Yonsei RR10k	Yonsei RR64k	Yonsei RR128k	PSNR
Correlation	0.803	0.803	0.803	0.713
RMSE	0.599	0.599	0.598	0.714
Outlier Ratio	0.556	0.553	0.552	0.615