

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

J.341

(03/2016)

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

Measurement of the quality of service – Part 3

**Objective perceptual multimedia video quality
measurement of HDTV for digital cable
television in the presence of a full reference**

Recommendation ITU-T J.341

ITU-T



International
Telecommunication
Union

Document Sharing Hub

Copyright International Telecommunication Union/ITU Telecommunication Sector

Recommendation ITU-T J.341

Objective perceptual multimedia video quality measurement of HDTV for digital cable television in the presence of a full reference

Summary

Recommendation ITU-T J.341 provides an objective perceptual video quality measurement method for HDTV when a full reference signal is available. The following list shows example applications that can use this Recommendation:

- 1) Potentially real-time, in-service quality monitoring at the source.
- 2) Remote destination quality monitoring when a copy of the source is available at the point of measurement.
- 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.

This Recommendation includes an electronic attachment containing several short high definition (HD) video sequences, their corresponding predicted mean opinion scores (MOS), and the predicted MOS for five HD databases available via Video Quality Experts Group (VQEG).

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T J.341	2011-01-13	9	11.1002/1000/11033
2.0	ITU-T J.341	2016-03-15	9	11.1002/1000/12771

* 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 2016

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	Application	2
1.2	Limitations.....	2
2	References.....	2
3	Definitions	2
3.1	Terms defined elsewhere	2
3.2	Terms defined in this Recommendation.....	3
4	Abbreviations and acronyms	3
5	Conventions	3
6	Description of the full reference methodology	4
7	Findings of the Video Quality Experts Group	5
Annex A	– Description of model VQuad-HD	6
A.1	Preprocessing.....	7
A.2	Time alignment.....	8
A.3	Spatial frame alignment.....	9
A.4	Computation of local similarity and local difference features	10
A.5	Analysis of the distribution of local features.....	10
A.6	Computation of the blockiness feature	11
A.7	Computation of the jerkiness feature (temporal quality).....	13
A.8	Aggregation to MOS	14
A.9	Handling of heavily spatially misaligned video sequences	19
A.10	Reference source code implementation.....	20
Annex B	– Description of model VQuad-HD-fast.....	21
Annex C	– Conformance testing	25
Bibliography		26

www.itu.int

Recommendation ITU-T J.341

Objective perceptual multimedia video quality measurement of HDTV for digital cable television in the presence of a full reference

1 Scope

This Recommendation¹ provides a perceptual video quality measurement method for use in high definition television (HDTV) non-interactive applications when the full reference (FR) measurement method can be used. The model VQuad-HD was developed to estimate subjective quality scores obtained using [b-ITU-T P.910].

The full reference measurement method can be used when the unimpaired reference video signal is 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 method includes both calibration and objective video quality estimations.

The validation test material contained both [b-ITU-T H.264] and MPEG-2 coding degradations and various transmission error conditions (e.g., bit errors, dropped packets). The model 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. The model 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 [b-ITU-T J.149]) to determine the model's accuracy for that particular context.

This Recommendation is deemed appropriate for telecommunication services delivered at between 1 Mbit/s and 30 Mbit/s. The following resolutions and frame rates were considered in the validation test:

- 1080i 60 Hz (29.97 fps);
- 1080p (25 fps);
- 1080i 50 Hz (25 fps);
- 1080p (29.97 fps).

The following conditions were allowed in the validation test for each resolution:

Test factors
Video resolution: 1920 x 1080 interlaced and progressive
Video frame rates 29.97 and 25 frames per second
Video bitrates: 1 to 30 Mbit/s
Temporal frame freezing (pausing with skipping) of maximum 2 seconds
Transmission errors with packet loss
Conversion of the SRC from 1080 to 720p, compression, transmission, decompression, and then conversion back to 1080.
Coding technologies
ITU-T H.264/AVC (MPEG-4 Part 10)
MPEG-2

¹ This Recommendation includes an electronic attachment containing several short high definition (HD) video sequences, their corresponding predicted mean opinion scores (MOS), and the predicted MOS for five HD databases available via VQEG.

Note that 720p was considered in the validation test plan as part of the test condition hypothetical reference circuit (HRC). As 720p is currently commonly up-scaled as part of the display, it was felt that 720p HRCs would more appropriately address this format.

1.1 Application

The applications for the estimation model 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 a copy of the source is available at the point of measurement.
- 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 video quality estimation model described in this Recommendation cannot be used to 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 [b-ITU-T J.149]) to determine the model's accuracy for that particular context.

When frame freezing was present, the test conditions typically had frame freezing durations less than two seconds. The model in this Recommendation was not validated for measuring video quality in a re-buffering condition (i.e., video that has a steadily increasing delay or freezing without skipping). The model was not tested on other frame rates than those used in TV systems (i.e., 29.97 frames per second and 25 frames per second, in interlaced or progressive mode).

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

Note that the model in this Recommendation was not evaluated on talking-head content typical of video-conferencing scenarios.

2 References

This Recommendation does not incorporate any normative provisions from external references.

3 Definitions

3.1 Terms defined elsewhere

This Recommendation uses the following terms defined elsewhere:

3.1.1 subjective assessment (picture) [b-ITU-T J.144]: The determination of the quality or impairment of programme-like pictures presented to a panel of human assessors in viewing sessions.

3.1.2 objective perceptual measurement (picture) [b-ITU-T J.144]: The measurement of the performance of a programme chain by the use of programme-like pictures and objective (instrumental) measurement methods to obtain an indication that approximates the rating that would be obtained from a subjective assessment test.

3.1.3 proponent [b-ITU-T J.144]: An organization or company that proposes a video quality model for validation testing and possible inclusion in an ITU Recommendation.

3.2 Terms defined in this Recommendation

This Recommendation defines the following terms:

3.2.1 frame rate: The number of unique frames (i.e., total frames – repeated frames) per second.

3.2.2 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.

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

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
DMOS	Difference Mean Opinion Score
FR	Full Reference
FRTV	Full Reference Television
HD	High Definition
HDTV	High Definition Television
HRC	Hypothetical Reference Circuit
ILG	VQEG's Independent Laboratory Group
MOS	Mean Opinion Score
MOSp	Mean Opinion Score, predicted
NR	No (or zero) Reference
PSNR	Peak Signal-to-Noise Ratio
PVS	Processed Video Sequence
RMSE	Root Mean Square Error
RR	Reduced Reference
SFR	Source Frame Rate
SRC	Source Reference Channel or Circuit
VQEG	Video Quality Experts Group

5 Conventions

This Recommendation uses the following convention:

YUV Colour Space and file format

6 Description of the full reference methodology

The double-ended measurement method with full reference (FR), for objective measurement of perceptual video quality, evaluates the performance of systems by making a comparison between the undistorted input, or reference, video signal at the input of the system, and the degraded signal at the output of the system (Figure 6-1).

Figure 6-1 shows an example of application of the full reference method to test a codec in the laboratory.

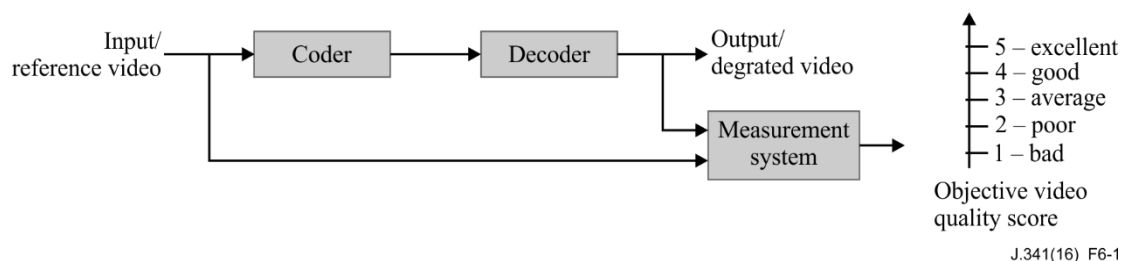


Figure 6-1 – Application of the full reference perceptual quality measurement method to test a codec in the laboratory

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 may also 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 full reference methods compare reference and processed pictures on what is effectively a pixel-by-pixel basis. The video quality metrics described in Annexes A and B 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 and 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 6-2 shows an example of the application of the full reference method to test a transmission chain.

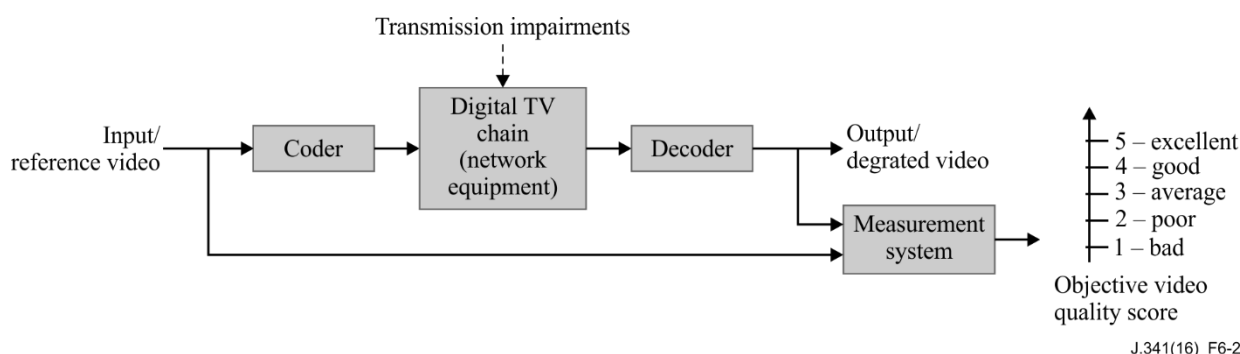


Figure 6-2 – Application of the full reference perceptual quality measurement method to test a transmission chain

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 6-2, or directly at the output of the encoder, as in Figure 6-1. 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 absolute category rating with hidden reference (ACR-HR) methods specified in [b-ITU-T P.910].

7 Findings of the Video Quality Experts Group

Studies of perceptual video quality measurements are conducted in an informal group, called the Video Quality Experts Group (VQEG), which reports to ITU-T Study Groups 9 and 12 and ITU-R Study Group 6. The recently completed high definition television phase I test of VQEG assessed the performance of proposed full reference perceptual video quality measurement algorithms.

The following statistics are taken from the final VQEG HDTV report [b-VQEG Report]. Note that the body of the VQEG HDTV report includes other metrics including Pearson Correlation and root mean square error (RMSE) calculated on individual experiments, confidence intervals, statistical significance testing on individual experiments, analysis on subsets of the data that include specific impairments (e.g., ITU-T H.264 coding-only), scatter plots, and the fit coefficients.

Primary analysis

The performance of the FR model is summarized in Table 7-1. The peak signal-to-noise ratio (PSNR) is calculated according to [b-ITU-T J.340] and included in this analysis for comparison purposes. "Superset RMSE" identifies the primary metric RMSE computed on the aggregated superset (i.e., all six experiments mapped onto a single scale). "Top Performing Group Total" identifies the number of experiments (0 to 6) for which this model was either the top performing model or statistically equivalent to the top performing model. "Better Than PSNR Total" identifies the number of experiments (0 to 6) for which the model was statistically better than PSNR. "Better Than Superset PSNR" lists whether each model is statistically better than PSNR on the aggregated superset. "Superset Correlation" identifies the Pearson Correlation computed on the aggregated superset.

Table 7-1 – Performance of the FR model VQuad-HD

Metric	PSNR	VQuad-HD
Superset RMSE	0.71	0.56
Top performing group total	1	5
Better than PSNR total	–	4
Better than superset PSNR	–	Yes
Superset correlation	0.78	0.87