

I n t e r n a t i o n a l T e l e c o m m u n i c a t i o n U n i o n

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

P.918

(01/2020)

SERIES P: TELEPHONE TRANSMISSION QUALITY,
TELEPHONE INSTALLATIONS, LOCAL LINE
NETWORKS

Audiovisual quality in multimedia services

Dimension-based subjective quality evaluation for video content

Recommendation ITU-T P.918

ITU-T

ITU-T P-SERIES RECOMMENDATIONS

TELEPHONE TRANSMISSION QUALITY, TELEPHONE INSTALLATIONS, LOCAL LINE NETWORKS

Vocabulary and effects of transmission parameters on customer opinion of transmission quality	P.10–P.19
Voice terminal characteristics	P.30–P.39
Reference systems	P.40–P.49
Objective measuring apparatus	P.50–P.59
Objective electro-acoustical measurements	P.60–P.69
Measurements related to speech loudness	P.70–P.79
Methods for objective and subjective assessment of speech quality	P.80–P.89
Voice terminal characteristics	P.300–P.399
Objective measuring apparatus	P.500–P.599
Measurements related to speech loudness	P.700–P.709
Methods for objective and subjective assessment of speech and video quality	P.800–P.899
Audiovisual quality in multimedia services	P.900–P.999
Transmission performance and QoS aspects of IP end-points	P.1000–P.1099
Communications involving vehicles	P.1100–P.1199
Models and tools for quality assessment of streamed media	P.1200–P.1299
Telemeeting assessment	P.1300–P.1399
Statistical analysis, evaluation and reporting guidelines of quality measurements	P.1400–P.1499
Methods for objective and subjective assessment of quality of services other than speech and video	P.1500–P.1599

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

Recommendation ITU-T P.918

Dimension-based subjective quality evaluation for video content

Summary

Recommendation ITU-T P.918 presents guidelines for undertaking subjective experiments for the quality of experience (QoE) assessment of perceptual video quality dimensions. In addition to scores for the overall video quality, the methodology yields scores for five perceptual dimensions. Each perceptual dimension scores are based on the ratings of the amount of degradation present in one system/test condition. The method is designed to be used with naïve subjects. The dimension scores can be used to provide diagnostic information on what may cause the degradation.

The perceptual dimensions are described in this Recommendation as well as the method to conduct a subjective experiment. Furthermore, information is provided about possible test environment and setup, participant instructions, and test material.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T P.918	2020-01-13	12	11.1002/1000/14153

Keywords

Multi-dimensional quality assessment, subjective testing, video quality evaluation.

* 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 not 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 2020

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

Table of Contents

		Page
1	Scope.....	1
2	References.....	1
3	Definitions	1
	3.1 Terms defined elsewhere	1
	3.2 Terms defined in this Recommendation.....	1
4	Abbreviations and acronyms	2
5	Conventions	2
6	Introduction to video quality analysis.....	2
7	Test methodology	3
	7.1 Dimension rating scales.....	3
	7.2 Quality rating scales	3
	7.3 Test design and rating scheme.....	4
	7.4 Instructions and training.....	4
8	Test environment	4
	8.1 Participants	4
	8.2 Test material	4
	8.3 Data analysis and reporting of the results.....	5
	Appendix I – Test instructions for the video quality – Video dimension test	6
	Appendix II – Results of an evaluation experiment – Investigating the video quality and the five perceptual video quality dimensions.	11
	Bibliography.....	17

Recommendation ITU-T P.918

Dimension-based subjective quality evaluation for video content

1 Scope

This Recommendation describes a subjective test methodology, which can assess and diagnose the video quality based on underlying perceptual quality dimensions.

Traditional quality tests, as described in [ITU-T P.800] and [ITU-T P.910], provide valid methods for the overall video quality but do not give insights into reasons for possible quality losses.

This Recommendation describes a test methodology that not only yields overall quality scores for video but specifically allows participants to rate five underlying perceptual quality dimensions for video and video transmission in general. The perception-based assessment of video QoE is important when planning and implementing services, as well as for the development of instrumental quality prediction models, especially when the service on hand is aimed for usage by humans.

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 P.800] Recommendation ITU-T P.800 (1996), *Methods for subjective determination of transmission quality*.
- [ITU-T P.806] Recommendation ITU-T P.806 (2014), *A subjective quality test methodology using multiple rating scales*.
- [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*.

3 Definitions

3.1 Terms defined elsewhere

None.

3.2 Terms defined in this Recommendation

This Recommendation defines the following term:

3.2.1 direct scaling: The rating of perceptual dimensions by a test participant without any additional mathematical procedure, like principal component analysis (PCA) or multidimensional scaling (MDS).

4 Abbreviations and acronyms

This Recommendation uses the following abbreviations and acronyms:

ACR	Absolute Category Rating
DIC	Discontinuity
FRA	Fragmentation
LUM	Suboptimal Luminosity
MDS	Multidimensional Scaling
MOS	Mean Opinion Score
NOI	Noisiness
PCA	Principal Component Analysis
QoE	Quality of Experience
SD	Semantic Differential
UCL	Unclearness
VQD	Video Quality Dimensions

5 Conventions

None.

6 Introduction to video quality analysis

To provide information about the quality of transmitted video, ITU-T recommends several different experimental designs. The approach presented here targets assessing perpetual dimensions to give a deeper insight into possible quality loss. The recommended method refers to a passive perception scenario and gives additional information about the overall video quality mean opinion scores (MOS) absolute category rating (ACR) experiments, as recommended in [ITU-T P.800].

To identify the video quality relevant perceptual dimensions, a pairwise similarity experiment with a subsequent multidimensional scaling (MDS) and a semantic differential experiment with a subsequent principal component analysis (PCA), were analyzed. Applying both test paradigms in separate experiments [b-Schiffner1] resulted in the set of perceptual dimensions for transmitted video given in Table 1.

Table 1 – Overview of the five identified and proposed perceptual video quality dimensions

VQD	Name	Description	Example impairment
I	Fragmentation (FRA)	Fallen apart, torn and disjointed	Packet loss
II	Unclearness (UCL)	Unclear and smeared image	Low coding bitrate
III	Discontinuity (DIC)	Interruptions in the flow of the video	Buffer delay and limitations
IV	Noisiness (NOI)	Random change in brightness and colour	Quantization, Circuit noise
V	Suboptimal luminosity (LUM)	Too high or low brightness	Over- and under-exposure