ETSI TR 102 493 V1.3.1 (2017-07)

Speech and multimedia Transmission Quality (STQ); Guidelines for the use of Video Quality Algorithms for Mobile Applications

Reference

RTR/STQ-00211m

Keywords

QoS, telephony, video

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00   Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

The present document can be downloaded from:
http://www.etsi.org/standards-search

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the only prevailing document is the print of the Portable Document Format (PDF) version kept on a specific network drive within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at
https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

If you find errors in the present document, please send your comment to one of the following services:
https://portal.etsi.org/People/CommiteeSupportStaff.aspx

Contents

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

Foreword

This Technical Report (TR) has been produced by ETSI Technical Committee Speech and multimedia Transmission Quality (STQ).

Modal verbs terminology

In the present document "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

1 Scope

The present document gives guidelines for the use of video quality algorithms for the different services and scenarios applied in the mobile environment.

2 References

2.1 Normative references

Normative references are not applicable in the present document.

2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long term validity.

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

[i.1] ETSI TS 126 233 (V13.0.0): "Universal Mobile Telecommunications System (UMTS); LTE; Transparent end-to-end Packet-switched Streaming service (PSS); General description (3GPP TS 26.233 version 13.0.0 Release 13)".

[i.2] ETSI TS 126 114: "Universal Mobile Telecommunications System (UMTS); LTE; IP Multimedia Subsystem (IMS); Multimedia telephony; Media handling and interaction (3GPP TS 26.114)".

[i.3] Void.

[i.4] Recommendation ITU-T P.1201.1: "Parametric non-intrusive assessment of audiovisual media streaming quality - Lower resolution application area".

[i.5] Recommendation ITU-T P.1201.2: "Parametric non-intrusive assessment of audiovisual media streaming quality - Higher resolution application area".

[i.6] Recommendation ITU-T P.1202.1: "Parametric non-intrusive bitstream assessment of video media streaming quality - Lower resolution application area".

[i.7] Recommendation ITU-T P.1202.2: "Parametric non-intrusive bitstream assessment of video media streaming quality - Higher resolution application area".

[i.8] Recommendation ITU-T P.1203.1: "Parametric bitstream-based quality assessment of progressive download and adaptive audiovisual streaming services over reliable transport - video quality estimation module".

[i.9] Recommendation ITU-T P.1203.2: "Parametric bitstream-based quality assessment of progressive download and adaptive audiovisual streaming services over reliable transport - audio quality estimation module".

[i.10] Recommendation ITU-T P.1203.3: "Parametric bitstream-based quality assessment of progressive download and adaptive audiovisual streaming services over reliable transport - Quality integration module".

[i.11] Recommendation ITU-T J.343.1: "Hybrid-NRe objective perceptual video quality measurement for HDTV and multimedia IP-based video services in the presence of encrypted bitstream data".

[i.12] Recommendation ITU-T J.343.2: "Hybrid-NR objective perceptual video quality measurement for HDTV and multimedia IP-based video services in the presence of non-encrypted bitstream data".

[i.13] Recommendation ITU-T J.343.3: "Hybrid-RRe objective perceptual video quality measurement for HDTV and multimedia IP-based video services in the presence of a reduced reference signal and encrypted bitstream data".

[i.14] 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".

[i.15] Recommendation ITU-T J.343.5: "Hybrid-FRe objective perceptual video quality measurement for HDTV and multimedia IP-based video services in the presence of a full reference signal and encrypted bitstream data".

[i.16] Recommendation ITU-T J.343.6: "Hybrid-FR objective perceptual video quality measurement for HDTV and multimedia IP-based video services in the presence of a full reference signal and non-encrypted bitstream data".

[i.17] Recommendation ITU-T J.246 (2008): "Perceptual visual quality measurement techniques for multimedia services over digital cable television networks in the presence of a reduced bandwidth reference".

[i.18] Recommendation ITU-T J.247 (2008): "Objective perceptual multimedia video quality measurement in the presence of a full reference".

[i.19] Recommendation ITU-T J.341 (2016): "Objective perceptual multimedia video quality measurement of HDTV for digital cable television in the presence of a full reference".

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the following terms and definitions apply:

bitstream model: computational model that predicts the subjectively perceived quality of video, audio or multimedia, based on analysis of the payload and transport headers

hybrid model: computational model that predicts the subjectively perceived quality of video, audio, or multimedia, based on the media signal and the payload and transport headers

live Streaming: streaming of live content e.g. web cam, TV programs, etc.

parametric model: computational algorithm that predicts the subjectively perceived quality of video, based on transport layer and client parameters

perceptual model: computational algorithm that aims to predict the subjectively perceived quality of video, based on the media signal

streaming on demand: streaming of stored content e.g. movies

3.2 Abbreviations

For the purposes of the present document, the following abbreviations apply:

AVC	Advanced Video Coding
BLER	BLock Error Rates
CPU	Central Processing Unit
DCT	Discrete Cosine Transform
FR	Full Reference Algorithm

HD	High Definition

NOTE: 1 280 x 720 pixels, fullHD 1 920 x 1 080 pixels.

HEVC	High Efficiency Video Coding
IMS	IP Multimedia Subsystem
IP	Internet Protocol
ITU	International Telecom standardization Union
JPEG	Joint Photographic Expert Group (Standard)
MM	Multimedia
MOS	Mean Opinion Score
MPEG	Moving Picture Expert Group (Standard)
MPEG-TS	MPEG Transport Stream
MTSI	Multimedia Service for IMS
NR	No Reference Algorithm
PLR	Packet Loss Rates
PSNR	Peak Signal Noise Ratio
RR	Reduced Reference
RTP	Real Time Protocol
SD	Standard Definition
TCP	Transport Control Protocol
TV	Television
UDP	User Datagram Procotol
VGA	Video Graphics Adapter
VHS	Video Home System

4 General

Video quality assessment has become a central issue with the increasing use of digital video compression systems and their delivery over mobile networks. Due to the nature of the coding standards and delivery networks the provided quality will differ in time and space. Thus, methods for video quality assessment represent important tools to compare the performance of end-to-end applications.

The present document sets the guidelines of video quality algorithms applicable for mobile applications and the scenarios of their application. Any eligible algorithm needs to predict the quality perceived by the user using mobile terminal equipment. The goal is to have one or more objective video quality measurement algorithms, which predict the video quality as perceived by a human viewer, which is in conformance with the minimum requirements list given in the present document.

ITU-T has approved many different algorithms for objective prediction of visual quality in the last years. They can be differentiated in algorithms based on image analysis (the actual image is input to the algorithm) and bitstream-based measures (the IP stream is input of the model), and so-called hybrid models which combine image analysis with meta-information from the bitstream. These models have different scopes and limitations, they require different input information and result in different predictions accuracy.

For image based analysis ITU recommends a reduced reference algorithm in Recommendation ITU-T J.246 [i.17] up to VGA resolution, multiple full-reference algorithms as in Recommendation ITU-T J.247 [i.18] for all video resolutions up to VGA and the algorithm in Recommendation ITU-T J.341 [i.19] for HD resolutions.

Bitstream based models are described in Recommendations ITU-T P.1201 and P.1202 series ([i.4] to [i.7]); there are differentiations for individual resolutions and for encrypted and non-encrypted bitstreams.

In 2015 ITU approved a set of hybrid models in the Recommendation ITU-T J.343 series ([i.11] to [i.16]), where J.343.1 [i.11] and J.343.2 [i.12] are no-reference hybrid models, J.343.3 [i.13] and J.343.4 [i.14] are reduced reference hybrid models and J.343.5 [i.15] and J.343.6 [i.16] are full-reference hybrid models. The odd suffix stands for models to be applicable for encrypted and non-encrypted bitstreams, while the models having an even suffix are only applicable to non-encrypted bitstreams.

It is common to all services treated in the present document that quality as seen from the user's perspective depends on the server and client applications used. For example, is has to be expected that under the same network conditions, two different video streaming clients will exhibit different video quality due to differences in the way these clients use available bandwidth. Therefore, for full validation of tools type and version of clients used has to be fully documented and are seen as part of the information needed to reproduce and calibrate measurements.

NOTE: The present document focuses on those visual continuous media reproductions where the source and the player are connected via a (mobile) telecommunication network rather than the replay of a clip that has been completely stored on the same device as the player and is replayed from there.

5 Services

5.0 Introduction

The aspect of video quality is of interest wherever there are services where the transfer of moving pictures or still images is involved. Three major fields of transferring video content can be identified that make use of packet switched and circuit switched services.

Table 1: Requirement profiles of the services

Application	Symmetry	Data rates	One Way Delay	Lip-sync
Video telephony	Two-way	32 kbps to 2 Mbps	< 150 ms preferred < 400 ms limit	< 80 ms
Streaming	One-way	32 kbps to 10 Mbps	< 10 s	
Conversational Multimedia	Two-way		< 150 ms	Mutual service dependency, echo

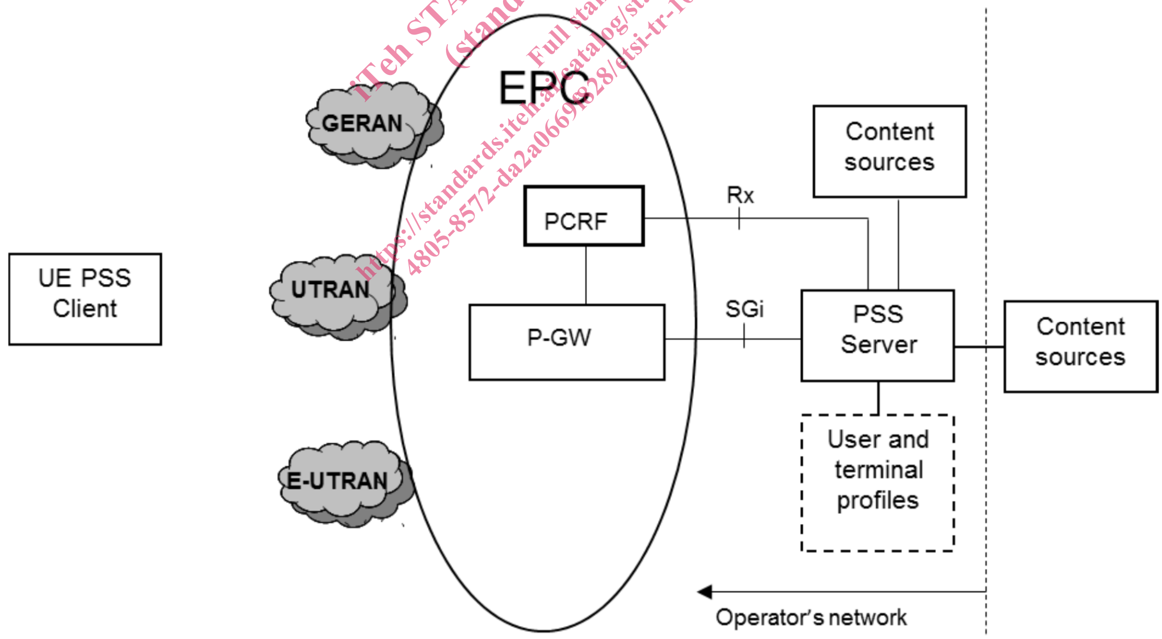

Figure 1: Streaming (ETSI TS 126 233 [i.1])

5.1 Streaming

Streaming refers to replay of media streams like audio and video in a continuous way while those streams are being received by the client over a data network. The client plays the incoming multimedia stream from a buffer in which the packets are stored after arrival.

Streaming accounts for a large percentage of the data network traffic. Typical applications can be classified into on-demand and live information delivery applications. Examples of the first group are video-on-demand applications like YouTube™. Live delivery of radio and television programs is an example of the second category.

NOTE: YouTube™ is the trade name of a product supplied by Google. This information is given for the convenience of users of the present document and does not constitute an endorsement by ETSI of the product named.

5.2 Conversational Multimedia

Multimedia services combine two or more media components within a call. The service where two or more parties exchange video, audio and text and maybe even share documents is a multimedia service. This is a peer-to-peer set up in which one party acts as the source (server) and the other as client(s) and vice versa in real time. Another example of a new multimedia conversational service is the 3GPP standardized MTSI service [i.2].

5.3 Video Telephony

Video telephony is a full-duplex system, carrying both video and audio and intended for use in a conversational environment. In principle the same delay requirements as for conversational voice will apply, i.e. no echo and minimal effect on conversational dynamics, with the added requirement that the audio and video have to be synchronized within certain limits to provide "lip-synch".

6 QoS Scenarios

6.0 General

The different services that are making use of video can be delivered in a variety of ways and situations. To obtain the full picture of the quality of these services they need to be tested accordingly. However for practical purposes and general feasibility, key scenarios need to be identified to facilitate video quality measurements.

6.1 Measurement Scenarios

The key scenarios are live streaming, streaming on demand, video telephony and conversational multimedia. These services can be tested by drive test or in a static fashion.

The algorithms for estimating video and audiovisual quality can be classified depending on:

- Type of input:
  - Perceptual (access to the video signal).
  - Hybrid (access to both the video signal and either the transport layer payload or the transport header information).
  - Bitstream (access to the transport layer payload, but not the video signal).
  - Parametric (access to transport header, client information, and knowledge about used codecs).

NOTE: The accessibility of the needed information depends on presence of encryption and its depth.

- Access to reference video: The algorithm models that are used are:
  - Full reference model (FR).
  - No reference model (NR).