ETSI TR 126 902 V13.0.0 (2016-01)

TECHNICAL REPORT

Digital cellular telecommunications system (Phase 2+);
Universal Mobile Telecommunications System (UMTS);
LTE;
Video codec performance
(3GPP TR 26.902 version 13.0.0 Release 13)

Reference

RTR/TSGS-0426902vd00

Keywords

GSM,LTE,UMTS

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
http://portal.etsi.org/tb/status/status.asp

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

Intellectual Property Rights

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.

Foreword

This Technical Report (TR) has been produced by ETSI 3rd Generation Partnership Project (3GPP).

The present document may refer to technical specifications or reports using their 3GPP identities, UMTS identities or GSM identities. These should be interpreted as being references to the corresponding ETSI deliverables.

The cross reference between GSM, UMTS, 3GPP and ETSI identities can be found under http://webapp.etsi.org/key/queryform.asp.

Modal verbs terminology

In the present document "**shall**", "**shall not**", "**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.

Contents

Foreword

This Technical Report has been produced by the 3rd Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

- x the first digit:
  - 1 presented to TSG for information;
  - 2 presented to TSG for approval;
  - 3 or greater indicates TSG approved document under change control.
- y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.
- z the third digit is incremented when editorial only changes have been incorporated in the document.

1 Scope

The present document comprises a technical report on Video Codec Performance, for packet-switched video-capable multimedia services standardized by 3GPP.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] IETF RFC 2429: "RTP Payload Format for the 1998 Version of ITU-T Rec. H.263 Video (H.263+)".

[2] IETF RFC 3550: "RTP: A Transport Protocol for Real-Time Applications", Schulzrinne H. et al, July 2003.

[3] ITU-T Recommendation H.263 (1998): "Video coding for low bit rate communication".

[4] 3GPP TS 26.110: "Codec for Circuit Switched Multimedia Telephony Service; General Description".

[5] 3GPP TS 26.111: "Codec for Circuit Switched Multimedia Telephony Service; Modifications to H.324".

[6] ITU-T Recommendation H.263 - Annex X (2004): "Annex X: Profiles and levels definition".

[7] ITU-T Recommendation H.264 (2003): "Advanced video coding for generic audiovisual services" | ISO/IEC 14496-10:2003: "Information technology - Coding of audio-visual objects - Part 10: Advanced Video Coding".

[8] ISO/IEC 14496-10/FDAM1: "AVC Fidelity Range Extensions".

[9] IETF RFC 3984: "RTP payload Format for H.264 Video".

[10] 3GPP TS 26.141: "IP Multimedia System (IMS) Messaging and Presence; Media formats and codecs".

[11] 3GPP TS 26.234: "Transparent end-to-end Packet-switched Streaming Service (PSS); Protocols and codecs".

[12] 3GPP TS 26.346: "Multimedia Broadcast/Multicast Service (MBMS); Protocols and codecs".

[13] 3GPP TS 26.235: "Packet switched conversational multimedia applications; Default codecs".

[14] 3GPP TS 26.236: "Packet switched conversational multimedia applications; Transport protocols".

[15] 3GPP TS 26.114: "IP Multimedia Subsystem (IMS); Multimedia telephony; Media handling and interaction".

[16] 3GPP TR 26.936: "Performance characterization of 3GPP audio codecs".

[17] 3GPP TR 25.101: "User Equipment (UE) radio transmission and reception (FDD)".

3 Abbreviations

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

APSNR	Average PSNR
AVC	Advanced Video Codec
DCCH	Dedicated Control CHannel
DPDCH	Dedicated Physical Data CHannel
DTCH	Dedicated Traffic CHannel
HSPA	High-Speed Packet Access
IETF	Internet Engineering Task Force
IMS	Internet protocol Multimedia Subsystem
IP	Internet Protocol
MAC	Medium Access Control
MBMS	Multimedia Broadcast/Multicast Service
MSE	Mean Square Error
MTSI	Multimedia Telephony over IMS
NAL	Network Abstraction Layer
NSD	Normalized Square Difference
PANSD	PSNR of Average Normalized Square Difference
PDU	Protocol Data Unit
PDVD	Percentage of Degraded Video Duration
PSC	Packet-Switched Conversational
PSNR	Peak Signal-to-Noise Ratio
PSS	Packet-switched Streaming Service
RFC	IETF Request For Comments
RLC	Radio Link Control
RTCP	RTP Control Protocol
RTP	Real-time Transport Protocol
SDP	Session Description Protocol
TTI	Transmission Time Interval
UDP	User Datagram Protocol
UE	User Equipment
UTRAN	UMTS Terrestrial Radio Access Network

4 Document organization

The present document is organized as discussed below.

- Clause 5 introduces the service scenarios, including their relationship with 3GPP services. Furthermore, it discusses the performance measurement metrics used in the present document.
- Clause 6 (performance figures) defines representative test cases and contains a listing, in the form of tables, performance of video codecs for each of the test cases.
- Clause 7 (supplementary information on figure generation) contains pointers to accompanying files containing video sequences, anchor bit streams, and error prone test bit streams. It also describes the mechanisms used to generate the anchor compressed video data, compressed video data exposed to typical error masks, and descriptions on the creation of error masks.
- Annex A sketches one possible environment that could be used by interested parties as a starting point for defining a process to assess the performances of a particular video codec against the performance figures.
- Annex B introduces details on the H.263 encoder and decoder configurations.
- Annex C introduces details of the H.264 encoder and decoder configurations
- Annex D introduces details on the usage of 3G file format in the present document.
- Annex E introduces details on the usage of RTPdump format in the present document.

- Annex F introduces details on the simulator, bearers, and dump files.
- Annex G introduces the details on the Quality Metric Evaluation.
- Annex H introduces the details on the Video Test Sequences.
- Annex I provides information on verification of appropriate use of the tools provided in this document.

5 Service scenarios and metrics

Video transmission in a 3GPP packet-switched environment conceptually consists of an Encoder, one or more Channels, and a Decoder. The Encoder, as defined here, comprises the steps of the source coding and, when required by the service, the packetization into RTP packets, according to the relevant 3GPP Technical Specification for the service and media codec in question. The Channel, as defined here comprises all steps of conveying the information created by the Encoder to the Decoder. Note that the Channel, in some environments, may be prone to packet erasures, and in others it may be error free. In an erasure prone environment, it is not guarantied that all information created by the Encoder can be processed by the Decoder; implying that the Decoder needs to cope to some extent with compressed video data not compliant with the video codec standard. The Decoder, finally, de-packetizes and reconstructs the - potentially erasure prone and perhaps non-compliant - packet stream to a reconstructed video sequence. The only type of error considered at the depacketizer/decoder is RTP packet erasures.

5.1 Service scenarios

3GPP includes video in different services, e.g. PSS [11], MBMS [12], PSC [13], [14], and MTSI [15]. This report lists the performance figures only one service scenario focusing on an RTP-based conversational service such as PSC or MTSI.

- **Service scenario A (PSC/MTSI-like)** relates to conversational services involving compressed video data (an erasure prone transport, low latency requirements, application layer transport quality feedback, etc.). In this scenario, UE-based video encoding and decoding are assumed. The foremost examples for this service scenario are PSC or MTSI. Within the this service scenario, the performance of an encoder and a decoder is of importance for the service quality. Service scenario A refers to the performance of a decoder to consume a possibly non-compliant (due to transmission errors) compressed video data generated by an encoder that fulfils the provision of sufficient quality in this scenario.

5.2 Performance metrics

This clause defines performance metrics as used in section 6, to numerically and objectively express a Decoder's reaction to compressed video data that is possibly modified due to erasures. Only objective metrics are considered which can be computed from sequences being available in a 3G format as described in Annex D by using the method detailed in annex G.

The following section provides a general description of the quality metrics. For the exact computation with the availability of sequences in 3G format please refer to annex G.

The following acronyms are utilized throughout the remainder of this subclause:

- *OrigSeq*: The original video sequence that has been used as input for the video encoder.
- *ReconSeq*: The reconstructed video sequence, the output of a standard compliant decoder that operates on the output of the video encoder without channel simulation, i.e. without any errors. Timing alignment between the OrigSeq and ReconSeq are assumed.

- *ReceivedSeq*: The video sequence that has been reconstructed and error-concealed by an error-tolerant video decoder, after a) the video encoder operated on the OrigSeq and produced an error free packet stream file as output, b) the channel simulator used the error free packet stream file and applied errors and delays to it so to produce an error-prone packet file which is used as the input of the error-tolerant video decoder. For comparison purpose, a constant delay between OrigSeq and the ReceivedSeq is assumed, whereby this constant delay is removed before comparison.

Each of the following metrics generates a single value when run for a complete video sequence.

5.2.1 Average Peak Signal-to-Noise Ratio (APSNR)

The average Peak Signal-to-Noise Ratio (APSNR) calculated between all pictures of the *OrigSeq* and the *ReconSeq* or the *ReceivedSeq*, respectively. First, the Peak Signal-to-Noise Ratio (PSNR) of each picture is calculated with a precision sufficient to prevent rounding errors in the future steps. Thereafter, the PSNR values of all pictures are averaged. The result is reported with a precision of two digits.

NOTE: This is the traditional metric referred to as PSNR in the academic literature and in the context of video compression research.

Only the luminance component of the video signal is used.

In case that results from several *ReceivedSeq* are to be combined, the average of all PSNR values for all *ReceivedSeq* is computed as the final result.

5.2.2 PSNR of Average Normalized Square Difference (PANSD)

The PSNR of Average Normalized Square Difference (PANSD) is calculated between all pictures of the *OrigSeq* and the *ReceivedSeq*, respectively. First, the normalized square difference, also know as Mean Square Error (MSE) of each picture is calculated with a precision sufficient to prevent rounding errors in the future steps. Thereafter, the NSD values of all pictures are averaged. The result is reported with a precision of two digits. Then, a conversion of this value into a PSNR value is carried out.

Only the luminance component of the video signal is used.

In case that results from several *ReceivedSeq* are to be combined, the average of all NSD values for all *ReceivedSeq* is computed and the final result is the PSNR over this averaged NSD.

5.2.3 Percentage of Degraded Video Duration (PDVD)

The Percentage of Degraded Video Duration (PDVD) is defined as the percentage of time of the entire display time for which the PSNR of the erroneous video frames are more than x dB worse than PSNR of the reconstructed frames whereby x is set to 2 dB. This metric computation requires three sequences, the *OrigSeq*, the *ReconSeq*, and the *ReceivedSeq*.

Only the luminance component of the video signal is used.

In case that results from several *ReceivedSeq* are to be combined, the average of all PDVD values for all *ReceivedSeq* is computed as the final result.

6 Test case definition and performance figures

6.1 Preliminary remarks

The test cases defined in this clause represent typical or worst-case transmission scenarios for the evaluated 3GPP video services. In all cases the respective constraints of the service have been taken into account, especially in the selection of video encoding tools. Purposely, some well-known and recognized optimization features, for example for error resilience, have been excluded. It is obvious that these optimizations are possible and, at least to some extent, also expected in real-world implementations.

It is also *not* the intention of this document to provide a comparison of different video codecs.

6.2 Definitions

6.2.1 Video test sequences

The input to the encoding process is a source sequence in 3G file format. All applied test sequences are attached to the present document in zip-folder *TestSequences.zip*. Three different test sequences are provided.

Table 1: Video test sequences

Sequence	Name	File Name	Frame Rate
1	**Stunt**	**stunt_qcif.3gp**	**15 fps**
2	**Bar**	**bar-30s.3gp**	**12 fps**
3	**Party**	**lt-party-30s.3gp**	**12 fps**

Details on the test sequences can be found in annex H.

6.2.2 Encoding parameters

The video encoding process, the encoder is made aware of the service it is operating in and the expected bitrate. The only varying parameter in the considered scenario is the expected transmission bitrate in kbit/s. The encoder is not made aware of the expected error conditions.

The two considered bitrates are 64 kbit/s and 128 kbit/s. However, the encoder should take into account packetization overhead due to RTP/UDP/IP headers such that the actual encoding bitrate should be lower.

More detailed encoding parameters for each H.263 and H.264 are described in annexes B and C, respectively.

6.2.3 Transport parameters

The transport parameters define the bearer settings, the applied loss masks, and the number of statistical experiments. The definition of the bearer parameters is strongly related to the applied transport software.

The applied bearers and radio dump files are numbered from 1 to 8 (Bearer ID). In all cases, for statistical significance, 128 independent trials are applied.

For details of transport parameter definition we refer to clause 7.1.3 and annex F.

6.3 Service scenario A (PSC/MTSI-like service)

6.3.1 Encoding anchors

Encoding parameter combinations are defined by the following parameters:

- Anchor: Anchor number.
- Sequence: Sequence number and name.
- Bitrate: 64 kbit/s and 128 kbit/s.
- Codec: H.263 and H.264

The codec refers to the application of an encoder, which generates for:

- H.263 an RTP packet stream according to H.263 baseline (Profile 0) Level 45 [6] and RTP payload format for the ITU-T Recommendation H.263 video codec [3] in RFC 2429 [1].