|  |  |
| --- | --- |
| 3GPP TR 26.955 V0.3.0 (2020-08) | |
| Technical Report | |
| 3rd Generation Partnership Project;  Technical Specification Group TSG SA;  5G Video Codec Characteristics  (Release 17) | |
|  | |
|  |  |
|  | |
| The present document has been developed within the 3rd Generation Partnership Project (3GPP TM) and may be further elaborated for the purposes of 3GPP. The present document has not been subject to any approval process by the 3GPPOrganizational Partners and shall not be implemented. This Specification is provided for future development work within 3GPPonly. The Organizational Partners accept no liability for any use of this Specification. Specifications and Reports for implementation of the 3GPP TM system should be obtained via the 3GPP Organizational Partners' Publications Offices. | |

|  |
| --- |
|  |
| ***3GPP***  Postal address  3GPP support office address  650 Route des Lucioles - Sophia Antipolis  Valbonne - FRANCE  Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16  Internet  http://www.3gpp.org |
| ***Copyright Notification***  No part may be reproduced except as authorized by written permission. The copyright and the foregoing restriction extend to reproduction in all media.  © 2020, 3GPP Organizational Partners (ARIB, ATIS, CCSA, ETSI, TSDSI, TTA, TTC).  All rights reserved.  UMTS™ is a Trade Mark of ETSI registered for the benefit of its members  3GPP™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners LTE™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners  GSM® and the GSM logo are registered and owned by the GSM Association |

Contents

Foreword 6

Introduction 7

1 Scope 8

2 References 8

3 Definitions of terms, symbols and abbreviations 10

3.1 Terms 10

3.2 Symbols 11

3.3 Abbreviations 11

4 Overview of video codec capabilities in 3GPP services in Release 16 12

4.1 Introduction 12

4.2 TV Video Profiles 12

[4.3 VR Video Profiles – To be kept only if interest raised on VR 360 video codec evaluation 14

4.4 5G Media Streaming 14

4.4.1 Introduction 14

4.4.2 5GMS Downlink Streaming default profile 14

4.4.2.1 H.264 (AVC) 14

4.4.2.2 H.265 (HEVC) 15

4.4.2.3 Encapsulation format 15

4.4.3 5GMS Uplink Streaming default profile 15

4.4.3.1 H.265 (HEVC) 15

4.4.3.2 Encapsulation format 15

4.4.4 5GMS Television (TV) profile 15

4.4.4.1 H.264 (AVC) 15

4.4.4.2 H.265 (HEVC) 15

4.4.4.3 Encapsulation format 15

[4.4.5 5GMS Downlink 360 Virtual Reality (VR) profile – to be removed in case no interested in evaluating 360 VR 16

4.4.5.1 H.264 (AVC) 16

4.4.5.2 H.265 (HEVC) 16

4.4.5.3 Encapsulation format 16

4.5 Multimedia Telephony Services over IMS 16

4.5.1 Introduction 16

4.5.2 H.264 (AVC) 16

4.5.3 H.265 (HEVC) 16

4.5.4 Encapsulation format 17

4.6 Messaging Services 17

4.7 Screen Content Coding 17

5 Common Test Conditions and Parameters 17

5.1 Introduction 17

5.2 Video Test Sequences 18

5.3 Key Performance Indicators and Metrics 18

5.4 Reference Software Tools 18

6 Relevant Scenarios 18

6.1 Introduction 18

6.2 Scenario 1: Full HD Streaming 18

6.2.1 Motivation 18

6.2.2 Description of the Anticipated Application 18

6.2.3 Source Format Properties 20

6.2.4 Encoding and Decoding Constraints 21

6.2.5 Performance Metrics 21

6.2.6 Interoperability Considerations 21

6.2.7 Test Sequences 22

6.2.7.1 Standard Dynamic Range 22

6.2.7.2 High Dynamic Range 22

6.2.8 Detailed Test Conditions 22

6.2.8.1 Overview 22

6.2.8.2 Reference Software AVC 1 22

6.2.8.3 Reference Software HEVC 1: HM16.20 22

6.2.9 External Performance Data 23

6.2.10 Additional Information 23

6.3 Scenario 2: 4K-TV 23

6.3.1 Motivation 23

6.3.2 Description of the Anticipated Application 23

6.3.3 Source Format Properties 25

6.3.4 Encoding and Decoding Constraints 26

6.3.5 Performance Metrics 26

6.3.6 Interoperability Considerations 26

6.3.7 Test Sequences 26

6.3.8 Detailed Test Conditions 26

6.3.8.1 Overview 26

6.3.8.2 Reference Software HEVC 1: HM16.20s 27

6.3.8.2 HEVC open-source implementation libx265 27

6.3.9 External Performance Data 27

6.4 Scenario 3: Screen Content Scenario 27

6.4.1 Motivation 27

6.4.2 Description of the Anticipated Application 28

6.4.3 Source Format Properties 28

6.4.4 Screen content codec tools 28

6.4.5 Encoding and Decoding Constraints 29

6.4.6 Performance Metrics 30

6.4.7 Interoperability Considerations 30

6.4.8 Test Sequences 30

6.4.9 Detailed Test Conditions 30

6.4.10 External Performance Data 30

6.4.11 Additional Information 30

6.5 Scenario 4: Messaging and Social Sharing 30

6.5.1 Motivation 30

6.5.2 Description of the Anticipated Application 31

6.5.3 Source Format Properties 33

6.5.4 Encoding and Decoding Constraints 33

6.5.5 Performance Metrics 34

6.5.6 Interoperability Considerations 34

6.5.7 Test Sequences 34

6.5.8 Detailed Test Conditions 34

6.5.9 External Performance Data 34

6.5.10 Additional Information 34

6.6 Scenario 5: Online Gaming 35

6.6.1 Motivation 35

6.6.2 Description of the Anticipated Application 36

6.6.3 Source Format Properties 36

6.6.3.1 Introduction 36

6.6.3.2 Category A: Low/medium dynamicity with low/medium complexity. 36

6.6.3.3 Category B: games with high dynamicity and low/medium complexity. 37

6.6.3.4 Category C: games with high dynamicity and high complexity. 37

6.6.3.5 Category D: photo-realistic games or games based on natural images/video. 37

6.6.3.6 Category E: XR game content 37

6.6.3.7 Summary 37

6.6.4 Codec and compression tools for games 38

6.6.5 Encoding and Decoding Constraints 38

6.6.6 Performance Metrics 39

6.6.7 Interoperability Considerations 39

6.6.8 Test Sequences 39

6.6.9 Detailed Test Conditions 39

6.6.10 External Performance Data 39

6.6.11 Additional Information 39

7 Characterization of Existing Codecs 39

8 Gaps and Optimization Potential 40

8.1 Identified Gaps and Deficiencies with Existing Codecs 40

8.2 Potential Requirements for New Codecs 40

9 Initial Information on new Codecs 40

9.1 Introduction 40

9.2 Versatile Video Coding (VVC) 40

9.2.1 Overview 40

9.3 Essential Video Coding (EVC) 41

10 Conclusions and Proposed Next Steps 41

Annex A Scenario Template 41

A.1 Introduction 41

A.2 Template 41

Annex B Details on Performance Metrics 43

Annex C Candidate Test Sequences 44

C.1 Introduction 44

C.2 Gaming Test Sequences 44

C.2.1 Overview 44

C.2.2 AOV 44

C.2.2.1 Introduction 44

C.2.2.2 Source sequence properties 44

C.2.2.3 Copyright information 45

C.2.3 Baolei-Man 45

C.2.3.1 Introduction 45

C.2.3.2 Source sequence properties 45

C.2.3.3 Copyright information 45

C.2.4 Baolei-Woman 46

C.2.4.1 Introduction 46

C.2.4.2 Source sequence properties 46

C.2.4.3 Copyright information 46

C.2.5 Jianling-Temple 47

C.2.5.1 Introduction 47

C.2.5.2 Source sequence properties 47

C.2.5.3 Copyright information 47

C.2.6 Jiangling-Beach 48

C.2.6.1 Introduction 48

C.2.6.2 Source sequence properties 48

C.2.6.3 Copyright information 48

Annex <X> (informative): Change history 49

For definitive guidance on drafting 3GPP TSs and TRs, see [3GPP TS 21.801](http://www.3gpp.org/DynaReport/21801.htm) supplemented by the 3GPP web page <http://www.3gpp.org/specifications-groups/delegates-corner/writing-a-new-spec>.

Ensure all blue guidance text is removed before submitting the TS/TR to the TSG for approval.

# 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.

In the present document, modal verbs have the following meanings:

**shall** indicates a mandatory requirement to do something

**shall not** indicates an interdiction (prohibition) to do something

The constructions "shall" and "shall not" are confined to the context of normative provisions, and do not appear in Technical Reports.

The constructions "must" and "must not" are not used as substitutes for "shall" and "shall not". Their use is avoided insofar as possible, and they are not used in a normative context except in a direct citation from an external, referenced, non-3GPP document, or so as to maintain continuity of style when extending or modifying the provisions of such a referenced document.

**should** indicates a recommendation to do something

**should not** indicates a recommendation not to do something

**may** indicates permission to do something

**need not** indicates permission not to do something

The construction "may not" is ambiguous and is not used in normative elements. The unambiguous constructions "might not" or "shall not" are used instead, depending upon the meaning intended.

**can** indicates that something is possible

**cannot** indicates that something is impossible

The constructions "can" and "cannot" are not substitutes for "may" and "need not".

**will** indicates that something is certain or expected to happen as a result of action taken by an agency the behaviour of which is outside the scope of the present document

**will not** indicates that something is certain or expected not to happen as a result of action taken by an agency the behaviour of which is outside the scope of the present document

**might** indicates a likelihood that something will happen as a result of action taken by some agency the behaviour of which is outside the scope of the present document

**might not** indicates a likelihood that something will not happen as a result of action taken by some agency the behaviour of which is outside the scope of the present document

In addition:

**is** (or any other verb in the indicative mood) indicates a statement of fact

**is not** (or any other negative verb in the indicative mood) indicates a statement of fact

The constructions "is" and "is not" do not indicate requirements.

# Introduction

Predictions on mobile video consumption are ever increasing. Different studies point to dominance of video traffic in 5G networks reaching from 65% total traffic in the short-term all the way to 90% by the end of the decade. Video is expected to be integral for services such as enhanced mobile media (such as mobile streaming services), home broadband and TV (for example in the context of 5G fixed wireless access services), immersive and interactive media in the context of eXtended Realities (XR) and cloud gaming as well as new media services from new verticals. This indicates that the user experience and efficiency of 5G networks will be heavily impacted by the quality of video compression technologies that are used with 5G services. Efficient video compression and decompression technologies required dedicated hardware for power and resource efficient real-time execution but are at the same time complex and costly in terms of implementation on integrated platforms. Hence, typically state-of-the-art video compression technologies last for several years and are used as generic service enablers for different applications and services, including traditional streaming and conversational services, but also new media services. This document analyzes the currently defined 3GPP-defined video compression technologies for their suitability for existing and emerging services in the context of 5G and identifies gaps and optimization potentials that would warrant the introduction of new video compression technologies.

# 1 Scope

The present document relevant interoperability requirements, performance characteristics and implementation constraints of video codecs in 5G services, and to characterize existing 3GPP video codecs, in particular H.264/AVC and H.265/HEVC in order to have a benchmark for the addition of potential future video codecs. For this purpose, the document:

* Collects a summary of the video coding capabilities in 3GPP services.
* Collects a subset of relevant scenarios for video codecs in 5G-based services and applications, including video formats (resolution, frame rates, color space, etc.), encoding and decoding requirements, adaptive streaming requirements.
* Collects relevant and exemplary test conditions and material for such scenarios, including test sequences.
* Defines performance metrics for such scenarios with focus on objective performance metrics.
* Collects relevant interoperability functionalities and enabling elements for video codecs in different 5G services such as MTSI and Telepresence (i.e. RTP based conversational communications), or 5G media streaming (e.g. based on DASH/CMAF) supporting the identified scenarios.
* Collects relevant criteria and key performance indicators for the integration of video codecs in 5G processing platforms, taking into account factors such as encoding and decoding complexity in the context of the defined scenarios.
* Characterizes the existing codecs H.264/AVC and H.265/HEVC in the context of the above scenarios and document the findings in a consistent manner.
* Identifies gaps and deficiencies of existing codecs in such use cases and derive requirements for potential new codecs.
* Collects initial information on how new codecs under development in ISO/IEC SC29 WG11 (MPEG)/JVET (in particular including VVC and EVC) may meet the above criteria.

# 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] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[2] 3GPP TS 26.114: "IP Multimedia Subsystem (IMS); Multimedia telephony; Media handling and interaction".

[3] 3GPP TS 26.116: "Television (TV) over 3GPP services; Video profiles".

[4] 3GPP TS 26.118: "3GPP Virtual reality profiles for streaming applications".

[5] 3GPP TR 26.925: "Typical traffic characteristics of media services on 3GPP networks".

[6] 3GPP TR 26.928: "Extended Reality over 5G".

[7] Recommendation ITU-T H.264 (06/2017): "Advanced video coding for generic audiovisual services" | ISO/IEC 14496-10:2014/Amd.3 2016: "Information technology — Coding of audio-visual objects — Part 10: Advanced Video Coding".

[8] Recommendation ITU-T H.265 (11/2019): "High efficiency video coding" | ISO/IEC 23008-2:2020: "High Efficiency Coding and Media Delivery in Heterogeneous Environments - Part 2: High Efficiency Video Coding".

[9] "2020 Mobile Internet Phenomena Report", February 2020, accessible here:  
<https://www.sandvine.com/download-report-mobile-internet-phenomena-report-2020-sandvine>.

[10] "2019 Ericsson Mobility Report", November 2019, accessible here:  
https://www.ericsson.com/4acd7e/assets/local/mobility-report/documents/2019/emr-november-2019.pdf.

[11] T. Fautier, "New Codecs for 5G", DASH-IF Workshop on "Media Streaming meets 5G", December 2019, accessible here:  
<https://dashif.org/docs/workshop-2019/04-thierry%20fautier%20-%20Harmonic%20Codec%20Comparison%205G%20Media%20Workshop_Final%20v3.pdf>

[12] "Bitmovin Video Developer Report" , September 2019, accessible here:  
<https://go.bitmovin.com/video-developer-report-2019>.

[13] 3GPP TS 26.511: "5G Media Streaming (5GMS); Profiles, codecs and formats".

[14] Recommendation ITU-R BT.709-6 (06/2015): "Parameter values for the HDTV standards for production and international programme exchange".

[15] Recommendation ITU-R BT.2020-2 (10/2015): "Parameter values for ultra-high definition television systems for production and international programme exchange".

[16] Recommendation ITU‑R BT.2100-1 (06/2017): "Image parameter values for high dynamic range television for use in production and international programme exchange".

[17] “You can watch Netflix on any screen you want, but you’re probably watching it on a TV”, https://www.vox.com/2018/3/7/17094610/netflix-70-percent-tv-viewing-statistics

[18] YouTube Revenue and Usage Statistics (2020), https://www.businessofapps.com/data/youtube-statistics/

[19] Ultra HD Forum service tracker B2C, https://ultrahdforum.org/uhd-service-tracker/“

[20] HIS Market "4K-TV and UHD: the whole picture",  
<https://cdn.ihs.com/www/pdf/4ktv-uhd-ebook.pdf>

[21] A. Mercat, M. Viitanen, and J. Vanne, “UVG dataset: 50/120fps 4K sequences for video codec analysis and development,” Accepted to ACM Multimedia Syst. Conf., Istanbul, Turkey, June 2020.

[21] R. Jullian, Y. Chen, F. Galpin, E. François, M. Kerdranvat (InterDigital), "Extra results to JVET-N605 “Comparative study of video coding solutions VVC, AV1 and EVC versus HEVC", document JVET-O0898, 15th Meeting: Gothenburg, SE, 3–12 July 2019.

[22] R. Jullian, Y. Chen, F. Galpin, E. François, M. Kerdranvat (InterDigital), "Comparative study of video coding solutions VVC, AV1 and EVC versus HEVC", document JVET-N0605, 14th Meeting: Geneva, CH, 19–27 March 2019.

[23] S. Iwamura, S. Nemoto, A. Ichigaya, "[AHG13] Compression performance analysis for 4K and 8K HLG test sequences", document JVET-N0828, 14th Meeting: Geneva, CH, 19–27 March 2019.

[24] A. Segall, E. François, W. Husak, S. Iwamura, D. Rusanovskyy, "JVET common test conditions and evaluation procedures for HDR/WCG video," document JVET-P2011, 16th JVET meeting: Geneva, CH, 1–11 Oct. 2019.

[25] Nick Zarzycki, "The Complete Guide to Social Media Video Specs in 2018", July 23, 2018, <https://blog.hootsuite.com/social-media-video-specs/>

[26] YouTube Help, "Recommended upload encoding settings", <https://support.google.com/youtube/answer/1722171?hl=en>

[27] Facebook Help Center, "What are the video format guidelines for live streaming on Facebook?", <https://www.facebook.com/help/1534561009906955>

[28] Deepak Kumar, "All about TikTok video size [ Full Guide]-2020", February 2020, <https://tiktoktip.com/tiktok-size/>

[29] Snapchat business center, "Single Image or Video Specifications", <https://businesshelp.snapchat.com/en-US/article/top-snap-specs>

[30] ISO/IEC 23000-19: "Information Technology Multimedia Application Format (MPEG-A) – Part 19: Common Media Application Format (CMAF) for segmented media".

[31] GSMA White Paper, "Chinese operators make major RCS commitment: Whitepaper", April 9, 2020, <https://www.gsma.com/futurenetworks/latest-news/china-operators-make-major-rcs-commitment-whitepaper/>

[32] 3GPP TS 26.140, "Multimedia Messaging Service (MMS); Media formats and codecs".

[33] GSMA RCC.71, "RCS Universal Profile Service Definition Document", version 2.4, October 2019.

[34] GSMA NG.114, "IMS Profile for Voice, Video and Messaging over 5GS", February 2020.

[35] 3GPP TS 26.223, "Telepresence using the IP Multimedia Subsystem (IMS); Media handling and interaction"

[36] "NVIDIA VIDEO CODEC SDK", <https://developer.nvidia.com/nvidia-video-codec-sdk>

[37] Microsoft Online Documentation, "Real-time media calls and meetings with Microsoft Teams", <https://docs.microsoft.com/en-us/microsoftteams/platform/bots/calls-and-meetings/real-time-media-concepts>

[38] Recommendation ITU-T H.266 (07/2020): "Versatile Video Coding" | ISO/IEC 23090-3:2020: "Information technology — Coded representation of immersive media — Part 3: Versatile video coding".

[39] Recommendation ITU-T H.274 (07/2020): "Versatile supplemental enhancement information for coded video bitstreams" | ISO/IEC 23002-7:2020 "Information technology — MPEG video technologies — Part 7: Supplemental enhancement information messages for coded video bitstreams".

# 3 Definitions of terms, symbols and abbreviations

This clause and its three subclauses are mandatory. The contents shall be shown as "void" if the TS/TR does not define any terms, symbols, or abbreviations.

## 3.1 Terms

For the purposes of the present document, the terms given in 3GPP TR 21.905 [1] and the following apply. A term defined in the present document takes precedence over the definition of the same term, if any, in 3GPP TR 21.905 [1].

Definition format (Normal)

**<defined term>:** <definition>.

**example:** text used to clarify abstract rules by applying them literally.

## 3.2 Symbols

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

Symbol format (EW)

<symbol> <Explanation>

## 3.3 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905 [1] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in 3GPP TR 21.905 [1].

AAC

ABR

ACM

AOV

ARIB

ATSC

AVC Advanced Video Coding

AVCHD

AVI

BMFF

CABAC

CAE

CBP

CBR

CGI

CHP

CMAF

CMP

CRA

CTA

CTU

DASH

DVB

ERP

EVC Essential Video Coding

FPS

GOP

HDMI

HDR

HDTV

HEVC High-Efficiency Video Coding

HFR

HLG

HLS

HMD

HRD

HTML

IDR

JVET

MKV

MMO

MMORPG

MMS

MOBA

MOV

MPD

MPEG

MSE

MTSI

OBS

PSNR

RAP

RCC

RCS

RDPCM

RGB

RPG

RTMP

RTP

RTS

RWP

SDK

SDR

SEI

SSIM

SVOD

UDP

UHD

UMTS

UVG

VBR

VVC Versatile Video Coding

# 4 Overview of video codec capabilities in 3GPP services in Release 16

## 4.1 Introduction

This clause summarizes the video coding capabilities in relevant existing 5G services.

As of today, two codecs are prominently referenced and available, namely H.264 (AVC) [7] and H.265 (HEVC) [8].

Both codecs are defined as part of the TV Video Profiles in 3GPP TS 26.116 [3] and are also the foundation of the VR Video Profiles in 3GPP TS 26.118 [4]. The highest defined profile/level combinations are:

* H.264 (AVC) Progressive High Profile Level 5.1 [7] with the following additional restrictions and requirements:
  + the maximum VCL Bit Rate is constrained to be 120Mbps with cpbBrVclFactor and cpbBrNalFactor being fixed to be 1250 and 1500, respectively.
  + the bitstream does not contain more than 10 slices per picture
* H.265 (HEVC) Main-10 Profile Main Tier Level 5.1 [8] without any restrictions

More details on the codec capabilities and the the necessary interoperability requirements for different services are collected in the remainder of this clause.

## 4.2 TV Video Profiles

The TV Video Profiles in TS 26.116 [3] address coded representations of TV distribution signals up to UHD-1 phase 2. Table 4.2-1 provides an overview of the TV relevant formats considered in the context of 3GPP TV Video Profiles.

In the context of TV Video Profiles, the following aspect are defined:

* **Bitstream:** A media bitstream that conforms to a video encoding format and certain Operation Point.
* **Operation Point:** A collection of discrete combinations of different content formats including spatial and temporal resolutions, colour mapping, transfer functions, etc. and the encoding format.
* **Receiver:** A receiver that can decode and render any bitstream that is conforming to a certain Operation Point.

Table 4.2-1: TV over 3GPP services Video Profile Operation Points (TS 26.116 [3])

| Operation Point name | Resolution format | Picture aspect ratio | Scan | Max. frame rate | Chroma format | Chroma sub-sampling | Bit depth | Colour space format | Transfer  Characteristics |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| H.264/AVC 720p HD | 1280 × 720 | 16:9 | Progressive | 30 | Y'CbCr | 4:2:0 | 8 | BT.709 | BT.709 |
| H.265/HEVC 720p HD | 1280 × 720 | 16:9 | Progressive | 30 | Y'CbCr | 4:2:0 | 8 | BT.709 | BT.709 |
| H.264/AVC Full HD | 1920 × 1080 | 16:9 | Progressive | 60 | Y'CbCr | 4:2:0 | 8 | BT.709 | BT.709 |
| H.265/HEVC Full HD | 1920 × 1080 | 16:9 | Progressive | 60 | Y'CbCr | 4:2:0 | 8; 10 | BT.709; BT.2020 | BT.709; BT.2020 |
| H.265/HEVC UHD | 3840 × 2160 | 16:9 | Progressive | 60 | Y'CbCr | 4:2:0 | 10 | BT.2020 | BT.2020 |
| H.265/HEVC Full HD HDR | 1920 x 1080 | 16:9 | Progressive | 60 | Y'CbCr | 4:2:0 | 10 | BT.2020 | BT.2100 PQ |
| H.265/HEVC UHD HDR | 3840 x 2160 | 16:9 | Progressive | 60 | Y'CbCr | 4:2:0 | 10 | BT.2020 | BT.2100 PQ |
| H.265/HEVC Full HD HDR HLG | 1920 x 1080 | 16:9 | Progressive | 60 | Y'CbCr | 4:2:0 | 10 | BT.2020 | BT.2100 HLG |
| H.265/HEVC UHD HDR HLG | 3840 x 2160 | 16:9 | Progressive | 60 | Y'CbCr | 4:2:0 | 10 | BT.2020 | BT.2100 HLG |

Each of the Operation Point is associated to the support of a video codec and its Profile, Level and Tier (for HEVC) combined with additional bitstream constraints defined in TS 26.116 [3], clause 4. The Table 4.2-2 summarizes the video codec profile, tier and level associated to each operation point.

Table 4.2-2: Operation point video codec Profile/Tier/Level

| Operation Point name | Video Codec | Profile | Tier | Level |
| --- | --- | --- | --- | --- |
| H.264/AVC 720p HD | AVC | High | - | 3.1 |
| H.265/HEVC 720p HD | HEVC | Main | Main | 3.1 |
| H.264/AVC Full HD | AVC | High | - | 4.2 |
| H.265/HEVC Full HD | HEVC | Main-10 | Main | 4.1 |
| H.265/HEVC UHD | HEVC | Main-10 | Main | 5.1 |
| H.265/HEVC Full HD HDR | HEVC | Main-10 | Main | 4.1 |
| H.265/HEVC UHD HDR | HEVC | Main-10 | Main | 5.1 |
| H.265/HEVC Full HD HDR HLG | HEVC | Main-10 | Main | 4.1 |
| H.265/HEVC UHD HDR HLG | HEVC | Main-10 | Main | 5.1 |

For TV Video profiles, interoperability with ISO BMFF based systems and the DASH Streaming is of most relevance. Hence, for a codec to be used in the context of TV Video Profiles, the following is defined in terms of interoperability:

1. The receiver requirements on elementary stream level
2. The encapsulation of an elementary stream into an ISO Base Media File Format track
3. The provisioning of the media as part of DASH Adaptation Set to support seamless switching
4. All MPD-level signalling for the codec to support capability discovery

For details, refer to TS 26.116 [3], clause 4 and clause 5.

## [4.3 VR Video Profiles – To be kept only if interest raised on VR 360 video codec evaluation

The VR profiles for streaming applications defined in TS 26.118 [4] address the coded representation of 360 VR distribution signals. Table 4.X-1 provides an overview of the 360 VR relevant formats considered in the context of 3GPP VR profiles. The VR profiles follow the same logic as the TV Video profiles, they represent a list of interoperability points that are amended by the 3GPP services such as 5GMS.

Table 4.3-1: High-level Summary of Operation Points

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Operation Point name | Decoder | Bit depth | Typical  Original Spatial Resolution | Frame Rate | Colour space format | Transfer  Characteristics | Projection | Rotation | RWP | Stereo |
| Basic H.264/AVC | H.264/AVC HP@L5.1 | 8 | Up to 4k | Up to 60 Hz | BT.709 | BT.709 | ERP w/o padding | No | No | No |
| Main H.265/HEVC | H.265/HEVC MP10@L5.1 | 8, 10 | Up to 6k in mono and 3k in stereo | Up to 60 Hz | BT.709  BT.2020 | BT.709 | ERP w/o padding | No | Yes | Yes |
| Flexible H.265/HEVC | H.265/HEVC MP10@L5.1 | 8, 10 | Up to 8k in mono and 3k in stereo | Up to 120 Hz | BT.709  BT.2020 | BT.709,  BT.2100 PQ | ERP w/o padding CMP | No | Yes | Yes |

]

## 4.4 5G Media Streaming

### 4.4.1 Introduction

5G Media Streaming (5GMS) services are associated with a series of 5GMS profiles each of which containing a set of capability requirements associated to a service scenario. Default profiles (for downlink and uplink streaming) are also defined in case no other profile is claimed to be supported. The detailed requirements for 5G Media Streaming profiles and their associated codec requirements are defined in 3GPP TS 26.511 [13]. The following clauses focus on video codec requirements and recommendations in 5GMS specification as well as the encapsulation format.

NOTE: In some profiles the HD-HDR capability is used as the ability for a UE to present video signals with all the following features: at least Full-HD resolution, bit depth of at least 10, at least 60 frames per second, Wide Colour Gamut and High Dynamic Range.

### 4.4.2 5GMS Downlink Streaming default profile

#### 4.4.2.1 H.264 (AVC)

TS 26.511 [13] requires the support of H.264 (AVC) Progressive High Profile Level 3.1 [7] decoding, with the maximum VCL Bit Rate constrained to be 14 Mbps.

If the 5GMS client supports the reception of video and HD-HDR capabilities, then TS 26.511 [13] requires the support of H.264 (AVC) Progressive High Profile Level 4.0 [7] decoding.

#### 4.4.2.2 H.265 (HEVC)

TS 26.511 [13] recommends the support of H.265 (HEVC) Main Profile, Main Tier, Level 3.1 [8] decoding with no interlace support.

If the 5GMS client supports the reception of video and HD-HDR capabilities, then TS 26.511 [13] requires the support of H.265 (HEVC) Main10 Profile, Main Tier, Level 4.1 [8] decoding with no interlace support.

#### 4.4.2.3 Encapsulation format

TS 26.511 [13] requires the support of Common Media Application Format (CMAF) encapsulation as defined in ISO/IEC 23000-19 [30].

### 4.4.3 5GMS Uplink Streaming default profile

#### 4.4.3.1 H.265 (HEVC)

TS 26.511 [13] requires the support of H.265 (HEVC) Main10 Profile, Main Tier, Level 4.1[8] encoding with no interlace support.

#### 4.4.3.2 Encapsulation format

TS 26.511 [13] requires the support of Common Media Application Format (CMAF) as defined in ISO/IEC 23000-19 [30].

### 4.4.4 5GMS Television (TV) profile

#### 4.4.4.1 H.264 (AVC)

TS 26.511 [13] requires the support of *H.264/AVC 720p HD* as specified in TS 26.116 [3], clause 4.4.2.6.

TS 26.511 [13] recommends the support of *H.264/AVC Full HD* as specified in TS 26.116 [3], clause 4.4.3.6.

NOTE: Associated codec profiles and levels are listed in Table 4.2-2 of the present document.

#### 4.4.4.2 H.265 (HEVC)

TS 26.511 [13] recommends the support of *H.265/HEVC 720p HD*, *H.265/HEVC Full HD*, *H.265/HEVC, H.265/HEVC Full HD, H.265/HEVC UHD HDR*, *H.265/HEVC Full HD HDR HLG* and *H.265/HEVC UHD HDR HLG* as specified in TS 26.116 [3].

If the 5GMS client supports the Television (TV) profile and HD-HDR capabilities, TS 26.511 [13] requires the support of *H.265/HEVC Full HD* as specified in TS 26.116 [3].

NOTE: Associated codec profiles and levels are listed in Table 4.2-2 of the present document.

#### 4.4.4.3 Encapsulation format

TS 26.511 [13] requires the support of DASH as defined in 3GPP TS 26.116 [3], clause 5.

NOTE: Full alignment between DASH support in TS 26.116 [3] and CMAF support in TS 26.511 is still to be investigated.

### [4.4.5 5GMS Downlink 360 Virtual Reality (VR) profile – to be removed in case no interested in evaluating 360 VR

#### 4.4.5.1 H.264 (AVC)

TS 26.511 [13] requires the support of the *Basic H.264/AVC* Operation Point as defined in 3GPP TS 26.118 [4], clause 5.1.4.

NOTE: Associated codec profiles and levels are listed in Table 4.3-1 of the present document.

#### 4.4.5.2 H.265 (HEVC)

TS 26.511 [13] recommends the support of the *Main* *H.265/HEVC* Operation Point as defined in TS 26.118 [4], clause 5.1.5.

TS 26.511 [13] indicates that the *Flexible* *H.265/HEVC* Operation Point as defined in TS 26.118 [4], clause 5.1.6, may be supported.

NOTE: Associated codec profiles and levels are listed in Table 4.3-1 of the present document.

#### 4.4.5.3 Encapsulation format

TS 26.511 [13] requires the support of DASH as defined in 3GPP TS 26.118 [4], clause 5.2.

NOTE: Full alignment between DASH support in TS 26.116 [3] and CMAF support in TS 26.511 is still to be investigated.

]

## 4.5 Multimedia Telephony Services over IMS

### 4.5.1 Introduction

The Multimedia Telephony Services over IMS (MTSI) specification 3GPP TS 26.114 [2] covers the media aspects of 3GPP conversational services. For MTSI services the UE is associated with both encoder and decoder requirements. Therefore, the support for a given video codec with its profile and level is interpreted as the ability to:

- encode any video input which characteristics fit in the level limits and produce a bitstream conformant with the codec profile.

- decode any bitstream in the codec/profile/level configuration.

NOTE: Additional constraints related to conversational services such as decoding process, location of the parameter sets are defined in TS 26.114 [2], clause 5.2.2.

### 4.5.2 H.264 (AVC)

TS 26.114 [2], in clause 5.2.2 requires the support of H.264 (AVC) [7] Constrained Baseline Profile (CBP) Level 1.2.

TS 26.114 [2], in clause 5.2.2 also recommends H.264 (AVC) [7] Constrained High Profile (CHP) Level 3.1.

### 4.5.3 H.265 (HEVC)

TS 26.114 [2] requires the support of H.265 (HEVC) [8] Main Profile, Main Tier, Level 3.1 except for constrained terminals for which it is only recommended.

### 4.5.4 Encapsulation format

MTSI is built upon the Real Time Protocol (RTP) for which the payload formats are specified in TS 26.114 [2], clause 7.4.3.

## 4.6 Messaging Services

3GPP TS 26.140 [32] specifies the media types, formats and codecs for the MMS within the 3GPP system. The document extends to codecs for speech, audio, video, still images, bitmap graphics, and other media in general, as well as scene description, multimedia integration and synchronization schemes.

Specifically, for video, the following capabilities are defined:

- TS 26.140 [32] requires the support of H.264 (AVC) [7] Constrained Baseline Profile (CBP) Level 1.3.

- TS 26.140 [32] recommends the support of H.264 (AVC) [7] High Profile Level 3.1 with frame\_mbs\_only\_flag=1 by MMS clients supporting HDTV video content at a resolution of 1280x720 (720p) with progressive scan at 30 frames per second.

- TS 26.140 [32] also recommends the support of H.265 (HEVC) [8] Main Profile, Main Tier, Level 3.1.

NOTE: The specification has not been updated since 2014.

## 4.7 Screen Content Coding

3GPP TS 26.223 [35] specifies a client for the IMS-based telepresence service supporting conversational speech, video and text transported over RTP. Telepresence is defined as a conference with interactive audio-visual communications experience between remote locations, where the users enjoy a strong sense of realism and presence between all participants (i.e. as if they are in same location) by optimizing a variety of attributes such as audio and video quality, eye contact, body language, spatial audio, coordinated environments and natural image size.

As specified in 3GPP TS 26.223 [35] clause 5.2, telepresence UEs are required to support:

- H.264 (AVC) [7] Constrained High Profile (CHP), Level 3.1

- H.264 (AVC) [7] Constrained Baseline Profile (CBP), Level 1.2 (for interworking with MTSI clients)

- H.265 (HEVC) [8] Main Profile, Main Tier, Level 4.1

As specified in 3GPP TS 26.223 [35], Telepresence UEs are also recommended to support dedicated HEVC extensions defined specifically for such types of content:

- H.265 (HEVC) [8] Screen-Extended Main, Main Tier, Level 4.1

- H.265 (HEVC) [8] Screen-Extended Main 4:4:4, Main Tier, Level 4.1

# 5 Common Test Conditions and Parameters

## 5.1 Introduction

This clause defines common test conditions and parameters.

## 5.2 Video Test Sequences

## 5.3 Key Performance Indicators and Metrics

## 5.4 Reference Software Tools

# 6 Relevant Scenarios

## 6.1 Introduction

This clause collects relevant scenarios based on the template defined in Annex A.

## 6.2 Scenario 1: Full HD Streaming

### 6.2.1 Motivation

The 2020 Mobile Internet Phenomena Report from Sandvine [9] shows that mobile video downstream traffic accounts for more than 65% of the global application category traffic share.

According to Ericsson mobility report [10], video traffic in mobile networks is forecast to grow by around 30 percent annually through 2025 to account for three-quarters of mobile data traffic, from slightly more than 60 percent in 2019. The video traffic growth is driven by the increase of embedded video in many online applications, growth of video-on-demand (VoD) streaming services in terms of both subscribers and viewing time per subscriber, and the evolution toward higher screen resolutions on smart devices. All of these factors have been influenced by the increasing penetration of video-capable smart devices.

Furthermore, while UHD and 4K are trendy formats, the main application for mobile streaming is Full HD with 1080p at 50 or 60 frames per second is expected to be the format of choice for mobile streaming at scale. The distribution version may be downsampled to support adaptive bitrate streaming, possibly with High Dynamic Range (HDR) support. For detailed discussion please refer to the presentation at the DASH-IF Workshop Dec 2019 [11].

In terms of distribution, while in the past, streaming video was delivered primarily via RTMP or RTP, fewer and fewer devices support these aging protocols each year. Instead, the latest web standards support built-in video playback and HTML5 is now by far the preferred method for video playback. And adaptive bitrate protocols dominate the distribution. According to the developer report [12], adaptive bitrate streaming through HLS/DASH, using the CMAF/DASH based segment formats, provide vast majority for streaming video. The distribution is used for On-Demand and Live Streaming.

### 6.2.2 Description of the Anticipated Application

In the context of 3GPP services, 5G Media Streaming [13] as well as the TV Video Profiles [3] are specifications addressing this streaming scenario. Both, 5G Media Streaming [13] and TV Video Profiles [3] builds on CMAF-based Segment formats and DASH distribution. From TS 26.116, the following operation points may be considered in scope of the Full HD Streaming Scenario (pending availability of appropriate test content):

- H.265/HEVC Full HD HDR, see TS26.116 [3] clause 4.5.3.

- H.264/AVC Full HD, see TS26.116 [3] clause 4.4.3.

- H.265/HEVC Full HD, see TS26.116 [3] clause 4.5.5.

- H.265/HEVC Full HD HDR HLG, see TS26.116 [3] clause 4.5.7.

These operation points are further informed by relevant operational experience with commercially available encoders and decoders.

The considered scenario is the distribution of content through DASH/CMAF based streaming. Important aspects that are expected to be considered when evaluating a codec in the context of this:

- Quality and Coding Efficiency:

- High and uninterrupted visual quality, taking into account the service constraints.

- Any savings can provide significant benefits due to the expected large volume of the traffic either in quality or network utilization.

- Adaptive Bitrate streaming:

- Multiple bit rates are provided, typically with a ladder of 30–50% to permit bandwidth adaptation. The use of constant bit rate (CBR) encoding maximises reuse of a common ladder of encoded representations across multiple distribution networks. The use of capped variable bit rate (VBR) encoding allows the bit rate to be varied according to the difficulty of the source material while maintaining the ability to distribute the encoded representations through distribution networks with fixed capacity. This also maximises reuse of a common ladder across multiple distribution networks.

- CMAF Fragments of size typically in the range of 1–6s to permit seamless switching for bit rate adaptation.

- Regular Random Access, typically every 1–2 seconds according to TS 26.116 [3]. To achieve clean switching in both sound and picture when moving between different encoded representations in the ladder, 3.84 seconds enables video segment boundaries to be aligned with an integer number of audio Access Units, if a 50fps video signal and 48kHz audio signal is used.

- Encoding in this scenario is typically done as

- Live and On-Demand distribution and encoding

- Server and Cloud-based Encoding

- No specific encoding latency constraints

### 6.2.3 Source Format Properties

Table 6.2-1 provides an overview of the different source signal properties following the information from TS26.116 [3]. This information is used to select proper test sequences.

Table 6.2-1 Source Format Properties for different operation point

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Source Format Properties | H.264/AVC Full HD | H.265/HEVC Full HD | H.265/HEVC Full HD HDR | H.265/HEVC Full HD HLG |
| Spatial resolutions | 1920 × 1080 (Permitted Distribution formats: 1920 × 1080, 1600 × 900, 1280 × 720, 960 × 540, 854 × 480, 640 × 360,426 × 240) | | | |
| Chroma Format | Y'CbCr | | | |
| Chroma Subsampling | 4:2:0 | | | |
| Picture Aspect ratios | 16:9 | | | |
| Frame rates | 24; 25; 30; 50; 60; 24/1.001; 30/1.001; 60/1.001 Hz | | | |
| Bit Depth | 8 | 8, 10 | 10 | 10 |
| Colour space formats | BT.709 [14] | BT.709 [14]; BT.2020 [15] | BT.2020 [15] | BT.2020 [15] |
| Transfer Characteristics | BT.709 [14] | BT.709 [14]; BT.2020 [15] | BT.2100 [16] PQ | BT.2100 [16] HLG |

### 6.2.4 Encoding and Decoding Constraints

Table 6.2-2 provides an overview of encoding and decoding constraints for H.264/AVC Full HD and H.265/HEVC Full HD Profiles. This will support the definition of detailed test conditions.

Table 6.2-2 Encoding and Decoding Configurations

|  |  |  |
| --- | --- | --- |
| Encoding and Decoding Constraints | H.264/AVC Full HD | H.265/HEVC Full HD |
| Relevant Codec and Codec Profile/Levels according to TS26.116 and TS26.511. | H.264/AVC Progressive High Profile Level 4.2 [7] | HEVC/H.265 Main-10 Profile  Main Tier Level 4.1 [8] |
| Random access frequency | 1 second, 3.84 seconds [other numbers tbd] | 1 second, 3.84 seconds [other numbers tbd] |
| Error resiliency requirements | None | None |
| Bit rates and quality configuration | QP = [20, 23, 26, 29]  others | QP = [23[[1]](#footnote-2), 25, 28, 31, 34]  others |
| Bit rate parameters (CBR, VBR, CAE, HRD parameters) | Fixed QP  CBR 8–12 Mbit/s  VBR capped at 12 Mbit/s  others | Fixed QP  CBR 5–8 Mbit/s  VBR capped at 12 Mbit/s  others |
| ABR encoding requirements (switching frequency, etc.) | 1 second [other numbers tbd]  ABR through multiple QPs | 1 second [other numbers tbd]  ABR through multiple QPs |
| Latency requirements and specific encoding settings | No latency requirements beyond RAP so picture reordering allowed | No latency requirements beyond RAP so picture reordering allowed |
| Encoding complexity context | real-time encoding, cloud-based encoding, offline encoding, etc.  detailed parameters tbd | real-time encoding, cloud-based encoding, offline encoding, etc.  detailed parameters tbd |
| Required decoding capabilities | H.264/AVC Progressive High Profile Level 4.2 [7] | HEVC/H.265 Main-10 Profile  Main Tier Level 4.1 [8] |

### 6.2.5 Performance Metrics

The following performance metrics are considered for this scenario:

tbd

### 6.2.6 Interoperability Considerations

In order to use a codec in the context of 5G Media Streaming services in TS 26.511 and for TV Video profiles in TS 26.116, the following list provides a set of potentially relevant interoperability aspects for Full HD Streaming:

1. The receiver requirements on elementary stream level, in particular the profile/level and additional considerations.
2. The encapsulation of an elementary stream into an ISO Base Media File Format track
3. The definition of a CMAF media profile.
4. The static mapping of parameters to a DASH MPD, in particular to the MPD parameters, such as @mimeType, @codecs, etc
5. The dynamic mapping of parameters to a DASH MPD from a CMAF Master Header, in particular to the MPD parameters, such as @width, @height, etc.
6. All MPD-level signalling for the codec to support capability discovery
7. Encryption requirements and recommendations.
8. Capability discovery options, for example mapping to HTML-5, MSE and media capability APIs.
9. Source Buffer Initialization Requirements.
10. Playback Requirements, for example by referencing CTA WAVE Specifications
11. Relation to other specifications, such as in DVB, ATSC, MPEG, ETSI, etc.

For additional details, please refer to TS 26.116 and TS 26.511.

### 6.2.7 Test Sequences

#### 6.2.7.1 Standard Dynamic Range

Example sequences are here https://media.xiph.org/video/derf/

#### 6.2.7.2 High Dynamic Range

tbd

### 6.2.8 Detailed Test Conditions

#### 6.2.8.1 Overview

tbd

#### 6.2.8.2 Reference Software AVC 1

tbd

#### 6.2.8.3 Reference Software HEVC 1: HM16.20

As reference software for HEVC, the following was used:

- <https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.20/>

Example settings: <https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.20/cfg/encoder_randomaccess_main10.cfg> with following proposed changes

- IntraPeriod: Intra Period such that 1 second is achieved

- DecodingRefreshType: 1 (CRA) 🡺 2 (IDR)

- GOPSize: adjusted to Intra

- QP: [25, 28, 31, 34]

### 6.2.9 External Performance Data

tbd

### 6.2.10 Additional Information

Tbd

## 6.3 Scenario 2: 4K-TV

### 6.3.1 Motivation

Streaming towards mobile devices is undoubtly the first natural use-case expected for 5G-media streaming. However, consumption of video services on fixed receivers (e.g. TV sets) remains a preferable way of experiencing the high-quality content, whether it is for on-demand (e.g., blockbuster movies) or live services (e.g., sport events). Recent reports from largest VOD platforms such as Netflix confirm that assumption and show that the primary way of watching content remains fixed TV screens, covering 70% of devices 6 months after subscription [17]. In the same way, YouTube indicates that service usage on fixed TV set remains an inevitable way of accessing the content, with 250M of hours viewed per day on TV screens [18]. As 5G media streaming targets a wide range of connected devices and should be able to deliver video streams to many compatible high-resolution receivers, (e.g. 5G-HDMI-sticks, 5G-StB/5G-MediaGateway or even 5G-TV sets) the inclusion of 4K TV Scenarios for 5G Video codec evaluation is important.

First, the 4K-TV set is currently the most established way of displaying premium quality services using latest technology improvements for video content, including High-Dynamic-Range (HDR) and Wide-Colour-Gamut (WCG). Latest statistics from the Ultra-HD forum indicate that 148 UHD services are currently on-air, 74% being linear, 45% of those using HDR [19]. In addition, a large number of SVOD operators propose 4K access in their subscription packages (e.g. Netflix™ and Amazon Prime™). All these services may eventually take advantage of 5G-network capabilities to increase the device reach and enlarge audiences. This scenario is also endorsed by strong shipment forecasts, as indicated in the latest IHS 4K-TV UHD bluebook [20].

### 6.3.2 Description of the Anticipated Application

In the context of 3GPP services, 5G Media Streaming [13] as well as the TV Video Profiles [3] are specifications addressing this 4K-TV scenario. Both, 5G Media Streaming [13] and TV Video Profiles [3] build on CMAF-based Segment formats and DASH distribution. From 3GPP TS 26.116, the following operation points may be considered in scope of the 4K-TV Streaming Scenario (pending availability of appropriate test content):

- H.265/HEVC UHD, see 3GPP TS 26.116 [3] clause 4.5.4.

- H.265/HEVC UHD HDR, see 3GPP TS 26.116 [3] clause 4.5.6.

- H.265/HEVC UHD HDR HLG, see 3GPP TS 26.116 [3] clause 4.5.8.

This scenario is based on CMAF (including LL-DASH and HLS-LL) distribution of UHD-TV video services over 5G networks to 5G/non-5G capable devices. This includes 5G-equipped devices (e.g. smartphone, tablets, …) but also other devices gateway (e.g. TV sets, HDMI-Stick…) accessing services through a “5G-gateway” which can be a mobile phone or a home gateway. As multiple linear services will be delivered in parallel (news, sport, talk show…) in a similar manner as traditional TV services in a multiplex (potentially using multicast/broadcast delivery over 5G). In certain environments, High Frame Rate (HFR) beyond 60 fps is considered, e.g. in DVB and ATSC broadcast specifications. 3GPP does not have any HFR TV video profiles yet.

Important aspects that are expected to be considered when evaluating a codec in the context of this 4K-TV scenario are:

- Quality and Coding Efficiency:

- High and uninterrupted visual quality, considering the service constraints.

- Any savings can provide significant benefits due to the expected large volume of the traffic either in quality or network utilization.

- Adaptive Bitrate streaming:

- Multiple bit rates are provided, typically with a ladder of 30–50% to permit bandwidth adaptation. The use of constant bit rate (CBR) encoding maximises reuse of a common ladder of encoded representations across multiple distribution networks. The use of capped variable bit rate (VBR) encoding allows the bit rate to be varied according to the difficulty of the source material while maintaining the ability to distribute the encoded representations through distribution networks with fixed capacity. This also maximises the usage of a common ladder across multiple distribution networks.

- CMAF Fragments of size typically in the range of 1–6s to permit seamless switching for bit rate adaptation.

- Regular Random Access, typically every 1–2 seconds according to 3GPP TS 26.116 [3]. To achieve clean switching in both sound and picture when moving between different encoded representations in the ladder, 3.84 seconds enables video segment boundaries to be aligned with an integer number of audio Access Units, if a 50fps video signal and 48kHz audio signal is used.

- Encoding in this scenario is typically done as

- Live and On-Demand distribution and encoding

- Server and Cloud-based Encoding

- Capable of encoding multiple services at variable bitrate, inside a fixed dedicated resource (statistical multiplexing).

### 6.3.3 Source Format Properties

Table 6.3-1 provides an overview of the different source signal properties for 4K-TV. This information is used to select proper test sequences.

Table 6.3-1 4K-TV source format properties

|  |  |
| --- | --- |
| Source format properties | 4K-TV |
| Spatial resolution | 3840 x 2160  (Permitted distribution formats: 2560 × 1440, 1920 × 1080, 1600 × 900, 1280 × 720) |
| Chroma format | Y’CbCr |
| Chroma subsampling | 4:2:0 |
| Picture aspec ratio | 16:9 |
| Frame rates | 24; 50; 60; 24/1.001; 60/1.001; [100; 120] Hz |
| Bit depth | 10 |
| Colour space formats | BT.2020 [15] |
| Transfer characteristics | BT.2020 [15], BT.2100 [16] (PQ & HLG) |

NOTE: High Frame Rate (HFR) is not supported by 3GPP TV Video profiles defined in 3GPP TS 26.116 [3] in release 16. However, HFR is introduced in this clause for consideration on the video codec performances.

### 6.3.4 Encoding and Decoding Constraints

Table 6.3-2 provides an overview of encoding and decoding constraints for 4K-TV category using legacy codec HEVC. This will support the definition of detailed test conditions. It is noted that no relevant profiles exist in TS26.116 and TS26.511 for HFR 4K-TV content.

Table 6.3-2 Encoding and Decoding Configurations for 4K-TV with legacy HEVC codec

|  |  |  |
| --- | --- | --- |
| Encoding and Decoding Constraints | H.265/HEVC 4K-TV | H.265/HEVC 4K-TV HFR |
| Relevant Codec and Codec Profile/Levels according to TS26.116 and TS26.511. | H.265/HEVC Main-10 Profile  Level 5.1 [8] | No relevant 3GPP profiles, should be aligned with H.265/HEVC Main-10 Profile Level 5.2 [8] |
| RAP period | 3.84sec, 1sec | 3.84sec, 1sec |
| Bit rate parameters (CBR, VBR, CAE, HRD parameters) | B = {10,15,20,25} Mbps  CBR and capped-VBR | B = {10,15,20,25} Mbps  CBR and capped-VBR |
| Latency requirements and specific encoding settings | No latency requirements beyond RAP so picture reordering allowed | No latency requirements beyond RAP so picture reordering allowed |
| Encoding complexity context | real-time encoding, cloud-based encoding, offline encoding, etc. | real-time encoding, cloud-based encoding, offline encoding, etc. |
| Required decoding capabilities | H.265/HEVC Main-10 Profile  Level 5.1 [8] | H.265/HEVC Main-10 Profile  Level 5.2 [8] |

### 6.3.5 Performance Metrics

Performance is assessed using BD-Rate computation, with PSNR, SSIM and VMAF metrics as objective quality criterion. Regarding complexity considerations, encoding/decoding runtime is provided.

### 6.3.6 Interoperability Considerations

In order to use a codec in the context of 5G Media Streaming services in 3GPP TS 26.511 [13] and for TV Video profiles in 3GPP TS 26.116 [3], the same considerations for interoperability as for FullHD according to clause 6.2.6 apply.

For additional details, please refer to 3GPP TS 26.116 [3] and 3GPP TS 26.511 [13].

### 6.3.7 Test Sequences

Following sources exist and may be used:

- 4K sequences from Ultra Video Group [21],

- Derf test material available at https://media.xiph.org/video/derf/

### 6.3.8 Detailed Test Conditions

#### 6.3.8.1 Overview

First, the legacy codec HEVC is tested to assess the relevance of what’s already in the 3GPP specification for this particular 4K-TV scenario. For this first test, two HEVC implementations are compared (x265 and reference software HM), according to encoding constraints derived from Table 6.3-2. For x265, encoding presets are selected to cover the desired encoding complexity contexts, for live and offline encoding. The tested rate-control modes are CBR and capped-VBR to fit with the possible delivery methods (single or multiple services inside a resource). In this first test, the HEVC reference implementation HM is evaluated with CBR rate-control on, at coding-tree-unit (CTU) granularity. In addition, fixed QP encoding is also carried out to provide additional anchor points aligned with formal MPEG/JVET common test conditions (potentially for future comparison with other codecs, if needed).

#### 6.3.8.2 Reference Software HEVC 1: HM16.20s

As reference software for HEVC, the following is used

- <https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.20/>

Example setting: <https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.20/cfg/encoder_randomaccess_main10.cfg> with following proposed changes

- IntraPeriod: Intra Period such that 1 second is achieved

- DecodingRefreshType: 1 (CRA) 🡺 2 (IDR)

- GOPSize: adjusted to Intra

- QP: [25, 28, 31, 34]

Tbd

#### 6.3.8.2 HEVC open-source implementation libx265

Tbd

### 6.3.9 External Performance Data

Tbd (with [21] [22] [23] as references ).

## 6.4 Scenario 3: Screen Content Scenario

### 6.4.1 Motivation

This scenario mostly motivates cases for which content goes beyond videographic content, an in particular includes computer generated imagery (CGI). Several application spaces rely on screen content types such as videoconferencing and webinars with presentation slides displayed, or remote desktop applications for demos or remote assistance. These services include massive online videoconferencing systems and telepresence systems both being particularly popular with remote collaboration.

For telepresence with screen-sharing applications, some information related to video is collected in the following:

* MS Teams™ [26] as of end of 2019.
  + There are several formats supported for video. Two key properties of a video format are its frame size and color format. Supported frame sizes include 640x360 ("360p"), 1280x720 ("720p"), and 1920x1080 ("1080p"). Supported color formats include NV12 (12 bits per pixel) and RGB24 (24 bits per pixel).
  + A "720p" video frame contains 921,600 pixels (1280 times 720). In the RGB24 color format, each pixel is represented as 3 bytes (24-bits) comprised of one byte each of red, green, and blue color components. Therefore, a single 720p RGB24 video frame requires 2,764,800 bytes of data (921,600 pixels times 3 bytes/pixel). At a frame rate of 30fps, sending 720p RGB24 video frames means processing approximately 80 MB/s of content (which is substantially compressed by the H.264 video codec before network transmission).
* Other tools are for further study.

### 6.4.2 Description of the Anticipated Application

3GPP until now has very restricted set of services but based on the considerations in clause 6.4.1, the following encoding benchmark capabilities are considered for decoding:

- H.264 (AVCHD) YUV 4:2:0, YUV 4:4:4, 8 bit, Max Resolution 1920x1080

- H.265 (HEVC) YUV 4:2:0, YUV 4:4:4, 10 bit, Max Resolution 4096 x 2048

The considered scenario is low-latency streaming or conversational. Important aspects that are expected to be considered when evaluating a codec in the context of this:

- Quality and Coding Efficiency:

- The ability to compress computer-generated content. Typically, it means the ability to have non perceptible intra refreshes and the ability to maintain stability on low frequency areas (such as uniform backgrounds) as well as maintaining details on high frequencies (particularly for text)

- The ability compress YUV 4:2:0 and 4:4:4 content.

- Considered settings for encoding:

- Low-latency settings

- No specific error resilience mechanisms

- Encoding in this scenario is typically done as

- Real-time encoding

### 6.4.3 Source Format Properties

Table 6.4-1 provides an overview of the different source signal properties for Screen Content Sharing. This information is used to select proper test sequences.

Table 6.4-1 Screen Content Sharing source properties

|  |  |  |  |
| --- | --- | --- | --- |
| Source format properties | | Screen Content |  |
| Spatial resolution | | 1920 x 1080 |  |
| Chroma format | | Y’CbCr, RGB |  |
| Chroma subsampling | | 4:2:0, 4:4:4 |  |
| Picture aspect ratio | | 16:9 |  |
| Frame rates | | 25, 30, 50, 60 Hz |  |
| Bit depth | | 8 |  |
| Colour space formats |  | BT.709 | |
| Transfer characteristics | | BT.709 |  |

### 6.4.4 Screen content codec tools

Table 6.4-2 provides an overview of the different codec tools per profile that may be suitable for coding screen content sequences.

Table 6.4-1 Screen Content Tools per Profile

|  |  |  |
| --- | --- | --- |
| **Screen content tools** | **AVC** | **HEVC** |
| main profile | Not applicable | Transform skip |
| range extension profile | not applicable | Residual Differential Pulse Code Modulation (RDPCM) (implicit intra/explicit inter), |
| screen content profile | Not applicable | Intra Block Copy (full frame or less), Palette, Adaptive Color Transform |

NOTE 1: HEVC specifies nested profiles. The screen content profile contains all the screen content tools from the main and range extension profiles.

NOTE 2: Tools specific results could be gathered from JCT-VC.

### 6.4.5 Encoding and Decoding Constraints

Table 6.4-3 provides an overview of encoding and decoding constraints for Screen Content scenario using AVC and HEVC codecs operating points. This will support the definition of detailed test conditions.

Table 6.4-3 Encoding and Decoding Configurations for Screen Content

|  |  |  |
| --- | --- | --- |
| **Encoding and Decoding Constraints** | **AVC** | **HEVC** |
| **Relevant Codec and Codec Profile/Levels according to TS26.116 and TS26.511.** | H.264/AVC Main Profile  Level 4.0 [X] | H.265/HEVC Main-10 Profile  H.265/HEVC Screen Content profile  Level 4.1, Level 5.1, Level 6.1 |
| **RAP period** | Gradual refresh: 1 second, no fix interval intra | Gradual refresh: 1 second, no fix interval intra |
| **Bit rate parameters (CBR, VBR, CAE, HRD parameters)** | Fixed QP | Fixed QP |
| **Latency requirements and specific encoding settings** | Low-latency requirements, Low-delay configuration | Low-latency requirements,  Low-delay B, low-delay-P |
| **Encoding complexity context** | real-time encoding. | real-time encoding. |
| **Required decoding capabilities** | H.264/AVC Main Profile  Level 4.0 [X] | H.265/HEVC Main-10 Profile  H.265/HEVC Screen Content profile  Level 4.1, 5.1, 6.1 [8] |

### 6.4.6 Performance Metrics

tbd

### 6.4.7 Interoperability Considerations

tbd

### 6.4.8 Test Sequences

Tests sequences illustrating the screen content scenario should ideally contain either snyhtetic content from a presentation such as a slide deck with text and graphics or natural content presenting participants to a video conference call, or a mix of the two]

### 6.4.

### 9 Detailed Test Conditions

For screen content, would like to at least test HEVC, 720p and 1080p, Main and Screen Extended Main profiles]

### 6.4.10 External Performance Data

tbd

### 6.4.11 Additional Information

tbd

## 6.5 Scenario 4: Messaging and Social Sharing

### 6.5.1 Motivation

According to the 2020 Mobile Internet Phenomena Report from Sandvine [9] Video traffic continues to grow worldwide, and the increasing popularity of mobile consumers sharing video has not only caused growth in downstream traffic, but also in upstream traffic as well. Instagram ™ grew in the upstream as more consumers share images and videos. TikTok ™, Snapchat ™ (video), FaceTime ™, and even Facebook ™ Live were all in the top 50 applications worldwide on the upstream that are video-sharing-centric. Messaging applications, especially on the upstream, continue to become a critical part of the mobile experience, replacing old style text messaging, and increasingly are video-based. Four of the top 20 applications on the upstream are messaging apps.

Some typical examples and restrictions in April 2020 are provided in the following:

1. WhatsApp™ [25]
   1. The maximum size of the video that you can share is 16 MB.
   2. Various container formats that are supported by include MP4, MKV, AVI, 3GP, and MOV. H.264/AVC video codec and AAC audio codec are needed today.
2. YouTube™ [26]
   1. H.264/AVC is the recommended codec with the following settings
      1. Progressive scan
      2. High Profile
      3. 2 consecutive B frames
      4. Closed GOP. GOP of half the frame rate.
      5. CABAC
      6. Variable bitrate. No bitrate limit required, though we offer recommended bit rates below for reference
      7. Chroma subsampling: 4:2:0
   2. Resolution Formats: 360p, 480p, 720p, 1080p, 1440p, 2160p
   3. Both SDR and HDR are possible
   4. The standard aspect ratio is 16:9
3. Facebook Live™ [27]. To live stream on Facebook™, these video format guidelines are provided:
   1. Recommended max bit rate is 4000 Kbps (4 mbps).
   2. Max: 1080p (1920x1080) resolution, at 60 frames per second.
   3. An I-frame (keyframe) must be sent at least every 2 seconds throughout the stream.
   4. H264 encoded video.
4. TikTok™ [28], some video restrictions
   1. Upload from Android, up to 72 MB at most. upload videos from iOS up to 287.6 MB.
   2. Video dimensions should be 1080 x 1920
   3. MP4 or MOV file format. Video should be H.264/AVC encoded
5. Snapchat™ [25][2)], The latest information from 2018
   1. Recommended size: 1080 by 1920 pixels (9:16 aspect ratio)
   2. Recommended specs: .MP4 or MOV, H.264 encoded, maximum file size 1GB

According to Sandvine's report [9], sharing and uploading content as part of social sharing is predominantly pictures and videos that uploaded directly into a cloud and uploaded to one or many social networks, and then discussed (or shared again) over messaging networks. The relevant quality-of-experience factors include the quality of shared content, the time it takes to upload, the costs associated with the upload and also the processing and battery consumption requirements on the device to prepare the content for upload.

In another activity, three large operators released a 5G messaging white paper [31] to promote enhanced messaging services based on Rich Communication Services (RCS). This also shows the relevance of operator-based messaging services. Generally, uplink resources are even more precious and costly in 5G network operation and hence efficient technologies are vital for mass-scale services. This aspect is also considered by GSMA RCS Universal Profile specification, promoted as the industry standard for RCS Business Messaging, ensuring the telecoms industry remains at the centre of digital communications [34].

### 6.5.2 Description of the Anticipated Application

In the context of 3GPP services, 5G Media Streaming [13] provides the following encoding benchmark capabilities:

- **HEVC-FullHD-Enc**: the capability to encode a video signal with

- up to 133,693,440 luma samples per second, and

- up to a luma picture size of 2,228,224 samples, and

- up to 240 frames per second, and

- the Chroma format being 4:2:0, and

- the bit depth being either 8 or 10 bit,

to a bitstream that is decodable by a decoder that is **HEVC-FullHD-Dec** capable as defined in clause 4.2.2.1 of TS26.511 and defined as the capability to decode H.265 (HEVC) Main10 Profile, Main Tier, Level 4.1 [3] bitstreams that have general\_progressive\_source\_flag equal to 1, general interlaced\_source\_flag equal to 0, general\_non\_packed\_constraint\_flag equal to 1, and general\_frame\_only\_constraint\_flag equal to 1.

Based on the considerations in clause 6.X.1, it is also recommened to take into account the AVC-FullHD-Enc capabilities as defined in TS26.511 [13]:

**AVC-FullHD-Enc**: the capability to encode a video signal with

- up to 245,760 macroblocks per second, and

- up to a frame size of 8,192 macroblocks, and

- up to 240 frames per second, and

- the Chroma format being 4:2:0, and

- the bit depth being 8 bit,

to a bitstream that is decodable by a decoder that is **AVC-HD-Dec** capable as defined in clause 4.2.1.1 of TS26.511 and defined as the capability to decode H.264 (AVC) Progressive High Profile Level 4.0 [2] bitstreams.

Based on future expectations of higher quality uploads, it is also recommened to take into account the HEVC-UHD-Enc capabilities as defined in TS26.511 [13]:

**HEVC-UHD-Enc**: the capability to encode a video signal with

- up to 534,773,760 luma samples per second, and

- up to a luma picture size of 8,912,896 samples, and

- up to 480 frames per second, and

- the Chroma format being 4:2:0, and

- the bit depth being either 8 or 10 bit,

to a bitstream that is decodable by a decoder that is **HEVC-UHD-Dec** capable as defined in clause 4.2.2.1 of TS26.511 and defined as the capability the capability to decode H.265 (HEVC) Main10 Profile, Main Tier, Level 5.1[3] bitstreams that have general\_progressive\_source\_flag equal to 1, general interlaced\_source\_flag equal to 0, general\_non\_packed\_constraint\_flag equal to 1, and general\_frame\_only\_constraint\_flag equal to 1.

The considered scenario is the uploading and uplink streaming into the ISO/BMFF and CMAF container formats. Important aspects that are expected to be considered when evaluating a codec in the context of this:

- Quality and Coding Efficiency:

- The ability to compress a video sequence targeting the maximum file size and maintaining high quality.

- The ability to compress a video stream in real time to the available uplink streaming resources.

- Considered settings for encoding:

- Regular random access at least every 2 seconds, preferably more often

- No specific encoding latency constraints are applicable

- Encoding in this scenario is typically done as

- Real-time encoding for social sharing

- Offline encoding for messaging

- UE-based Encoding

### 6.5.3 Source Format Properties

Table 6.5-1 provides an overview of the different source signal properties for Social Sharing and Messaging. This information is used to select proper test sequences.

Table 6.5-1 Social sharing and source format properties

|  |  |
| --- | --- |
| Source format properties | Social Sharing |
| Spatial resolution | 1920 x 1080  (Permitted encoding formats: 1920 × 1080, 1280 × 720, 854 × 480) |
| Chroma format | Y’CbCr |
| Chroma subsampling | 4:2:0 |
| Picture aspec ratio | 16:9 |
| Frame rates | 25, 30, 50, 60 Hz |
| Bit depth | 8, 10 |
| Colour space formats | BT.709, BT.2020 |
| Transfer characteristics | BT.709, BT.2100 (HDR) |

### 6.5.4 Encoding and Decoding Constraints

Table 6.5-2 provides an overview of encoding and decoding constraints for Social sharing and messaging category using AVC and HEVC codecs. This will support the definition of detailed test conditions.

Table 6.5-2 Encoding and Decoding Configurations for Social sharing and messaging

|  |  |  |
| --- | --- | --- |
| Encoding and Decoding Constraints | AVC HD | HEVC HD |
| Relevant Codec and Codec Profile/Levels according to TS26.116 and TS26.511. | H.264/AVC Main Profile  Level 4.0 [X] | H.265/HEVC Main-10 Profile  Level 4.0 [8] |
| RAP period | 1 and 5 seconds | 1 and 5 seconds |
| Bit rate parameters (CBR, VBR, CAE, HRD parameters) | B = {0.5, 1, 2, 5} Mbps  Capped-VBR (social sharing) and VBR (messaging)  Fixed QP | B = {0.5, 1, 2, 5} Mbps  Capped-VBR (social sharing) and VBR (messaging)  Fixed QP |
| Latency requirements and specific encoding settings | No latency requirements beyond RAP so picture reordering allowed | No latency requirements beyond RAP so picture reordering allowed |
| Encoding complexity context | real-time encoding (social sharing), offline encoding (messaging) | real-time encoding (social sharing), offline encoding (messaging). |
| Required decoding capabilities | H.264/AVC Main Profile  Level 4.0 [X] | H.265/HEVC Main-10 Profile  Level 4.0 [8] |

### 6.5.5 Performance Metrics

tbd

### 6.5.6 Interoperability Considerations

tbd

### 6.5.7 Test Sequences

Tbd

https://photos.app.goo.gl/6QrmTTMoizVtxAoQA

### 6.5.8 Detailed Test Conditions

tbd

### 6.5.9 External Performance Data

tbd

### 6.5.10 Additional Information

Tbd

## 6.6 Scenario 5: Online Gaming

### 6.6.1 Motivation

Online gaming for HMD-based consumption was discussed and introduced in detail in TR 26.928 [6].

As an example, a comprehensive set of APIs including high-performance tools, samples and documentation for hardware accelerated video encode and decode on Windows and Linux for NVIDIA™ Video Codec SDK is available [N]. As another

For games rendered on 2D displays such as smart phones, typical characteristics of rasterized frames produced by the game engine are:

- Resolution of 720p, 1080p or 4K

- Frame rate of 30fps, 60fps or 120 fps

Typical color bit depth of 8 bits (RGB frames) but higher bit depth may be offered for HDR compatible games.

- High Quality: 4k at 60/120fps with an average throughput of 60/100 Mbps

- Main Quality: 1080p at 60/120fps with an average throughput of 30/40 Mbps

- Low Quality: 720p/1080p at 30fps with an average throughput of 10/12 Mbps

### 6.6.2 Description of the Anticipated Application

3GPP until now has very restricted set of services but based on the considerations in clause 6.Y.1, the following encoding benchmark capabilities are considered for decoding 2D displays:

The considered online gaming scenario is low-latency streaming, typically using UDP/IP based distribution to minimize protocol latencies. Important aspects that are expected to be considered when evaluating a codec in the context of this:

- Quality and Coding Efficiency:

- The ability to compress traditional computer-generated content.

- The ability to compress photorealistic computer-generated content.

- The ability to compress YUV 4:2:0 and 4:4:4 content

- Considered settings for encoding:

- Ultra low latency and Low latency settings

- No specific error resilience mechanisms

- Encoding in this scenario is typically done as

- Real-time encoding

### 6.6.3 Source Format Properties

#### 6.6.3.1 Introduction

Video games have different characteristics that are important to take into account when encoding the rasterized frames produced by the game engine. In TR 26.928 [6], clause 4.2.2, a few different types of games and their interaction delay tolerance are documented. However, TR 26.928 [6] does not differentiate the characteristics of the content. This aspect is addressed in the following.

In particular, the following characteristics are important:

- Dynamicity of content: how frequent rasterized frames change when compared to previous frame

- Complexity of content: how much content changes between frames and how complex such changes are

- Type of content: traditional CGI, photo-realistic CGI or natural images/video

Depending on these characteristics as well as the interaction delay tolerance, video games can be organized into different categories as document in the remainder of this clause.

#### 6.6.3.2 Category A: Low/medium dynamicity with low/medium complexity.

This category includes games such as board games, turn-by-turn strategy games, management/simulation games or non-realtime role-playing games (RPG) in which content may not change over several consecutive frames and changes are typically limited.

This category also includes games such as adventure games, casual games, or platform games in which although content may change at every single frame, changes are limited to animation of sprites or simple global movements of the content.

The common characteristics of the games in this category is that their playability can support longer interaction delay tolerance (500 – 1000ms according to TR 26.928 [6]) and their content is typically considered to video encode.

#### 6.6.3.3 Category B: games with high dynamicity and low/medium complexity.

This category includes games such as fighting games, racing games, real-time strategy (RTS) games or real-time RPGs in which content is very dynamic but changes are either limited or simple transforms.

The common characteristics of the games in this category is that their playability requires shorter interaction delay tolerances (100ms according to TR 26.928 [6]) while their content is still considered simple to video encode (with high benefits from prediction coding).

#### 6.6.3.4 Category C: games with high dynamicity and high complexity.

This category includes games such as first-person shooters (FPS), Massive Multiplayer Online (MMO) games and racing games in which content is very dynamic with possibly very significant changes regularly in the content.

The common characteristics of the games in this category is that their playability requires shorter interaction delay tolerances (100ms according to TR 26.928) and their content is typically considered as complex content to video encode.

#### 6.6.3.5 Category D: photo-realistic games or games based on natural images/video.

The main characteristics of the games in this category is that their content is typically considered as more complex content to video encode.

#### 6.6.3.6 Category E: XR game content

For a detailed analysis, please refer to clause 6.6.1 and TR26.928, clause 4.4.

#### 6.6.3.7 Summary

Table 6.6.3.7-1 provides an overview of the different source signal properties for Online Gaming. This information is used to select proper test sequences.

Table 6.6.3.7-1 Online Gaming source properties

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Source format properties | Category A  low/med dynamicity & low/med complexity games | Category B  high dynamicity & low/med complexity games | Category C  high dynamicity & high complexity games | Category D  photo-realistic or natural video games | Category E  XR games |
| Spatial resolution | 1280x720, 1920x1080, 3840x2160 | 1280x720, 1920x1080, 3840x2160 | 1280x720, 1920x1080, 3840x2160 | 1280x720, 1920x1080, 3840x2160 | 1920x1080, 2048x1024, 3840x2160, 4096x2048, 8192x4096 |
| Chroma format | Y’CbCr | Y’CbCr | Y’CbCr | Y’CbCr | Y’CbCr |
| Chroma subsampling | 4:2:0, 4:4:4 | 4:2:0, 4:4:4 | 4:2:0, 4:4:4 | 4:2:0, 4:4:4 | 4:2:0, 4:4:4 |
| Picture aspect ratio | 16:9 | 16:9 | 16:9 | 16:9 | 16:9; 2:1 |
| Frame Buffers | 1 | 1 | 1 | 1 | 2 (1 per eye) |
| Frame rates | 30, 50, 60 Hz | 30, 50, 60, 90, 120 Hz | 30, 50, 60, 90, 120 Hz | 30, 50, 60, 90, 120 Hz | 30, 50, 60, 90, 120 Hz |
| Bit depth | 8 | 8 | 8, 10 | 8, 10 | 8, 10 |
| Colour space formats | BT.709, BT.2020 | BT.709, BT.2020 | BT.709, BT.2020 | BT.709, BT.2020 | BT.709, BT.2020 |
| Transfer characteristics | N/A | N/A | BT.2100 (HDR) | BT.2100 (HDR) | BT.2100 (HDR) |

### 6.6.4 Codec and compression tools for games

Table 6.6.4-1 provides an overview of the different coding tools per profile that may be suitable for coding different types of gaming sequences.

Table 6.6.4-1 Gaming Codec Tools per Profile

|  |  |  |
| --- | --- | --- |
| **Profile** | **AVC** | **HEVC** |
| **main profile** | **NA** | TBD |
| **range extension profile** | NA | TBD |
| **screen content profile** | NA | Palette, TBD |

NOTE: Palette being heavily used in the design of games, this tool should be particularly tested for game content.

For photo-realistic games (Category D), certain "screen content" tools may not provide much benefits for this type of game content and would need to be tested.

### 6.6.5 Encoding and Decoding Constraints

Table 6.6.5-1 provides an overview of encoding and decoding configurations for the online gaming scenario using AVC and HEVC codecs as operation points. This will support the definition of detailed test conditions.

Table 6.6.5-1 Encoding and Decoding Configurations for Online Gaming

|  |  |  |
| --- | --- | --- |
| **Encoding and Decoding Constraints** | **AVC** | **HEVC** |
| **Relevant Codec and Codec Profile/Levels according to TS26.116 and TS26.511.** | H.264/AVC Main Profile  Level 4.0 [X] | H.265/HEVC Main-10 Profile  Level 4.1, Level 5.1, Level 6.1 |
| **RAP period** | 1 second, infinite | 1 second, infinite |
| **Bit rate parameters (CBR, VBR, CAE, HRD parameters)** | Fixed QP | Fixed QP |
| **Latency requirements and specific encoding settings** | Low-latency requirements  Low-delay configuration  With and without use of future reference frames | Low-latency requirements,  Low-delay-P  Low-delay-B  With (moderate latency to non-critical latency games) and without (ultra low and low latency game) use of future reference frames. |
| **Encoding complexity context** | real-time encoding | real-time encoding. |
| **Required decoding capabilities** | H.264/AVC Main Profile  Level 4.0 [X] | H.265/HEVC Main-10 Profile  Level 4.1, 5.1, 6.1 [8] |

NOTE: Depending on the type of game (First person, Massive Multi-Player Online Role Play Game, Strategy) the latency requirements differs, usually ranging from less than 50ms to more than 200ms, thus leading to different encoding constraints and configuration.

[Table 6.6.5-2 provides an overview of latency constraints for each game category and identifies which coding tools are applicable and potentially useful for each category.

**Table 6.6.5-2 Latency constraints tools for different game categories**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Latency Constraints** | **Category A** | **Category B** | **Category C** | **Category D** | **Category E** |
| **Low latency requirement** | optional | Yes | Yes | Yes/No | Yes |
| **Low delay P** | optional | required | required | optional | required |
| **Low delay B** | optional | required | required | optional | required |
| **Use of future reference frames** | optional | No | No | optional | No |

]

### 6.6.6 Performance Metrics

tbd

### 6.6.7 Interoperability Considerations

tbd

### 6.6.8 Test Sequences

This should include FPS, RPG and Strategy games.]

For candidate test sequences, please refer to Annex C.2.

### 6.6.9

### 6.6.10 External Performance Data

tbd

### 6.6.11 Additional Information

tbd

# 7 Characterization of Existing Codecs

# 8 Gaps and Optimization Potential

## 8.1 Identified Gaps and Deficiencies with Existing Codecs

ffs

## 8.2 Potential Requirements for New Codecs

Ffs

Reference to interop requirements in clause 4.6.

# 9 Initial Information on new Codecs

## 9.1 Introduction

This clause collects initial information on how new codecs under development in ISO/IEC SC29 WG11 (MPEG)/JVET (in particular including VVC and EVC) may meet the above criteria based on the characterization results provided for example by ISO/IEC SC29 WG11 (MPEG)/JVET.

## 9.2 Versatile Video Coding (VVC)

### 9.2.1 Overview

ITU-T SG 16 informed that the development of the new Versatile Video Coding (VVC) standard has been completed for Consent at the June-July 2020 meeting of ITU-T SG16 and is also thereafter undergoing final approval in ISO/IEC JTC 1/SC 29/WG 11 (MPEG). After final approval, the dual reference to the jointly developed VVC standard will be Rec. ITU-T H.266 | ISO/IEC 23090-3 [38].

VVC is the latest in a series of such jointly developed ITU-T Recommendations and International Standards for video coding, and is the direct successor to HEVC (Rec. ITU-T H.265 | ISO/IEC 23008-2) [8] and AVC (Rec. ITU-T H.264 | ISO/IEC 14496-10) [7]. VVC provides significant coding efficiency improvements over HEVC. Plans are underway to conduct a verification test with formal subjective testing to confirm that VVC achieves about a 50% bit rate reduction vs. HEVC for equal subjective video quality. Test results demonstrate that VVC provides about a 40% bit rate reduction for 4K/UHD test sequences using objective metrics. Application areas especially targeted for the use of VVC include ultra-high definition 4K and 8K video, video with a high dynamic range and wide colour gamut, and video for immersive media applications such as 360° omnidirectional video, as well as conventional standard-definition and high-definition video content. In addition to improving coding efficiency, VVC also provides highly flexible syntax supporting such use cases as subpicture bitstream extraction, bitstream merging, and layered coding scalability.

The VVC standard includes the specification of a "Main 10" profile that supports 8- and 10-bit 4:2:0 video, a "Main 10 4:4:4" profile with 4:4:4 and 4:2:2 format support, corresponding "Multilayer Main 10" and "Multilayer Main 10 4:4:4" profiles with support for layered coding, and "Main 10 Still Picture" and "Main 10 4:4:4 Still Picture" profiles for still image coding employing the same coding tools as in the corresponding video profiles.

ITU-T also announced AAP Consent for Rec. ITU-T H.274 “Versatile supplemental enhancement information for coded video bitstreams” (VSEI), developed jointly as twin text with ISO/IEC JTC 1/‌SC 29/‌WG 11 as ISO/IEC 23002-7 [39]. VSEI specifies the syntax and semantics of video usability information (VUI) parameters and supplemental enhancement information (SEI) messages for use with coded video bitstreams. VSEI is particularly intended for use with VVC, although it is drafted in a manner intended to be sufficiently generic that it may also be used with other types of coded video bitstreams.

## 9.3 Essential Video Coding (EVC)

# 10 Conclusions and Proposed Next Steps

ffs

Annex A  
Scenario Template

## A.1 Introduction

This annex provides a proposed template to introduce a Scenario for 5G Video. This template has been used to collect the scenarios in this report.

## A.2 Template

The following aspects are considered to be important for a scenario:

1. Scenario name <give the scenario a catchy name>
2. Motivation for the scenario: Why is the scenario relevant for 5G and video? What is the expected traffic?
3. Description of the scenario: This provides a description of the scenario addressing potentially the relation to a service *5G-based services and applications*, including video formats (resolution, frame rates, color space, etc.), encoding and decoding requirements, adaptive streaming requirements, predominantly based on scenarios defined for 5G media streaming as well as for TR 26.925 and TR 26.928.
4. Supporting companies and 3GPP members:
   1. This documents the 3GPP members that support this scenario in terms of providing the information, test material, test requirements and the characterization for the tests. For each of the identified necessities, a tick box is is created in the template.
   2. Preferably several 3GPP members are included in the support, and in addition a video service provider may be included (not necessarily a 3GPP member).
   3. Cross-verification is preferably done by the supporters of the scenario
5. Source format properties: This defines a clear range of the considered and relevant source formats, including the signal properties, but also the characteristics of the content. As an example, the source formats as defined in TS26.116 may be used which include:
   1. Spatial resolutions
   2. Chroma Format
   3. Chroma Subsampling
   4. Aspect ratios
   5. Frame rates
   6. Colour space formats
   7. Transfer Characteristics
   8. Bit depth
   9. Other signal properties
6. Encoding and decoding constraints and settings: Typical encoding constraints and settings such as
   1. Relevant Codec and Codec Profile/Levels according to TS26.116 and TS26.511.
   2. Random access frequency
   3. Error resiliency requirements
   4. Bitrates and quality requirements
   5. Bitrate parameters (CBR, VBR, CAE, HRD parameters)
   6. ABR encoding requirements (switching frequency, etc.)
   7. Latency requirements and specific encoding settings
   8. Encoding context: real-time encoding, on device encoding, cloud-based encoding, offline encoding, etc.
   9. Required decoding capabilities
7. Performance Metrics and Requirements:
   1. A clear definition on how the performance needs to be evaluated including metrics, etc addressing the main KPIs of the scenario.
   2. Objective measures such as PSNR, VMAF, etc, may be used.
   3. Subjective evaluation is not excluded and may be done, but needs commitment
8. Interoperability Considerations for the application
   1. Streaming with DASH/HLS/CMAF
   2. RTP based delivery
9. Test Sequences
   1. A set of selected test sequences that are provided by the proponents in order to do the evaluation. They should cover a set of source format properties
10. Detailed test conditions:
    1. Provides a proposal for detailed test conditions, for example based on a reference software together with the sequences and configuration parameters.
11. External Performance data
    1. References to external performance data that can be added, for example other SDOs, public documents and so on.
12. Additional Information

Annex B  
Details on Performance Metrics

ffs

Annex C  
Candidate Test Sequences

## C.1 Introduction

tbd

## C.2 Gaming Test Sequences

### C.2.1 Overview

tbd

### C.2.2 Arena Of Valor

#### C.2.2.1 Introduction

Arena of Valor (AOV) ©Tencent Games – is a 3D MOBA (multiplayer online battle arena) game. Within a limited size map 2 groups of 5 vs 5 or 3 vs 3 players take part of a battle for which each player chooses a hero with particular power and abilities. The figure C.2.2-1 illustrates the gameplay of AOV with a screenshot of the AOV test sequence.



Figure C.2.2-1: Screenshot of AOV test sequence

#### C.2.2.2 Source sequence properties

The AOV test sequence has the following properties:

- Resolution: 1920 x1080

- Scan: Progressive

- Frame rate: 60 fps

- Bit depth: 8-bit

- Length: 600 frames (10s)

- YUV format: YUV 4:2:0

- Colour Components: Y’CbCr

- Colour space: ITU-R BT.709

#### C.2.2.3 Copyright information

The AOV video sequence is made available under the following copyright disclaimer.

*Copyright © 1998 - 2018 Tencent. All Rights Reserved.*

*This sequence and all intellectual property rights therein remain the property of Shenzhen Tencent Computer System Co. Ltd. and Tencent America LLC. This sequence may only be used for the purpose of developing, testing and promulgating technology standards (the Purpose). Shenzhen Tencent Computer System Co. Ltd. and Tencent America LLC. make no warranties with respect to this sequence and expressly disclaim any warranties regarding its fitness for any purpose. This disclaimer shall remain attached to this sequence except when this sequence is used for the Purpose.*

### C.2.3 Baolei-Man

#### C.2.3.1 Introduction

The *Baolei-Man* test sequence is an extract from a famous online multiplayer game in which many players play in a battle royale mode meaning that in survival mode, the last standing one is the winner. The figure C.2.3-1 shows a screenshot of the *Baolei-Man* test sequence.

Une image contenant herbe, intérieur, homme, table

Description générée automatiquement

Figure C.2.3-1: Screenshot of Baolei-Man test sequence

#### C.2.3.2 Source sequence properties

The Baolei-Man test sequence is available in two original versions (fullHD and 4K) with the following properties specified hereafter in Table C.2.3-1:

Table C.2.3-1: Baolei-Man sequences properties

|  |  |  |
| --- | --- | --- |
|  | 1080p | 2160p |
| Resolution | 1920x1080 | 3840x2160 |
| Scan | Progressive | Progressive |
| Frame Rate | 60 fps | 60 fps |
| Bit Depth | 8 | 10 |
| Length | 600 frames | 600 frames |
| YUV format | YUV 4:2:0 | YUV 4:2:0 |
| Colour components | Y’CbCr | Y’CbCr |
| Colour space | BT.709 | BT.709 |

#### C.2.3.3 Copyright information

The Baolei-Man video sequences (in FullHD and 4K resolutions) are made available under the following copyright disclaimer.

*Copyright © 1998 - 2020 Tencent. All Rights Reserved.*

*This sequence and all intellectual property rights therein remain the property of Shenzhen Tencent Computer System Co. Ltd.. This sequence may only be used for the purpose of developing, testing and promulgating technology standards (the Purpose). Shenzhen Tencent Computer System Co. Ltd. make no warranties with respect to this sequence and expressly disclaim any warranties regarding its fitness for any purpose. This disclaimer shall remain attached to this sequence except when this sequence is used for the Purpose.*

### C.2.4 Baolei-Woman

#### C.2.4.1 Introduction

The *Baolei-Woman* test sequence is also an extract from the same game as for *Balei-Man*. The figure C.2.4-1 shows a screenshot of the *Baolei-Woman* test sequence.

Une image contenant télévision, vidéo, moniteur, pièce

Description générée automatiquement

Figure C.2.4-1: Screenshot of Baolei-Woman test sequence

#### C.2.4.2 Source sequence properties

Similarly to the *Baolei-Man* test sequence, *Baolei-Woman* is also available in two original versions (fullHD and 4K) with the following properties specified hereafter in Table C.2.4-1:

Table C.2.4-1: Baolei-Woman sequences properties

|  |  |  |
| --- | --- | --- |
|  | 1080p | 2160p |
| Resolution | 1920x1080 | 3840x2160 |
| Scan | Progressive | Progressive |
| Frame Rate | 60 fps | 60 fps |
| Bit Depth | 8 | 10 |
| Length | 600 frames | 600 frames |
| YUV format | YUV 4:2:0 | YUV 4:2:0 |
| Colour components | Y’CbCr | Y’CbCr |
| Colour space | BT.709 | BT.709 |

#### C.2.4.3 Copyright information

The Baolei-Woman video sequences (in FullHD and 4K resolutions) are made available under the following copyright disclaimer.

*Copyright © 1998 - 2020 Tencent. All Rights Reserved.*

*This sequence and all intellectual property rights therein remain the property of Shenzhen Tencent Computer System Co. Ltd.. This sequence may only be used for the purpose of developing, testing and promulgating technology standards (the Purpose). Shenzhen Tencent Computer System Co. Ltd. make no warranties with respect to this sequence and expressly disclaim any warranties regarding its fitness for any purpose. This disclaimer shall remain attached to this sequence except when this sequence is used for the Purpose.*

### C.2.5 Jianling-Temple

#### C.2.5.1 Introduction

The *Jianling-Temple* test sequence is an extract from another famous massive multiplayer online role-playing game (MMORPG) having two gameplay modes Player-versus-Environment (PvE) and Player-versus-player (PvP). The figure C.2.5-1 is a screenshot of the *Jianling-Temple* test sequence.

Une image contenant herbe, extérieur, équitation, homme

Description générée automatiquement

Figure C.2.5-1: Screenshot of Jiangli-Temple test sequence

#### C.2.5.2 Source sequence properties

The *Jiangli-Temple* test sequence has the following properties:

- Resolution: 1920 x1080

- Scan: Progressive

- Frame rate: 60 fps

- Bit depth: 8-bit

- Length: 600 frames (10s)

- YUV format: YUV 4:2:0

- Colour Components: Y’CbCr

- Colour space: ITU-R BT.709

#### C.2.5.3 Copyright information

The *Jiangli-Temple* video sequence is made available under the following copyright disclaimer.

*Copyright © 1998 - 2020 Tencent & NCSOFT. All Rights Reserved.*

*This sequence and all intellectual property rights therein remain the property of Shenzhen Tencent Computer System Co. Ltd. and NCSOFT Co.Ltd.. This sequence may only be used for the purpose of developing, testing and promulgating technology standards (the Purpose). Shenzhen Tencent Computer System Co. Ltd. and NCSOFT Co.Ltd. make no warranties with respect to this sequence and expressly disclaim any warranties regarding its fitness for any purpose. This disclaimer shall remain attached to this sequence except when this sequence is used for the Purpose.*

### C.2.6 Jiangling-Beach

#### C.2.6.1 Introduction

The *Jianling-Beach* test sequence is another extract from the same MMORPG game as the *Jiangli-Temple* sequence. The figure C.2.6-1 is a screenshot of the *Jianling-Beach* test sequence.

Une image contenant ordinateur, portable, différent, assis

Description générée automatiquement

Figure C.2.6-1: Screenshot of Jiangli-Beach test sequence

#### C.2.6.2 Source sequence properties

The *Jiangli-Beach* test sequence has the following properties:

- Resolution: 1920 x1080

- Scan: Progressive

- Frame rate: 60 fps

- Bit depth: 8-bit

- Length: 600 frames (10s)

- YUV format: YUV 4:2:0

- Colour Components: Y’CbCr

- Colour space: ITU-R BT.709

#### C.2.6.3 Copyright information

The *Jiangli-Beach* video sequence is made available under the following copyright disclaimer.

*Copyright © 1998 - 2020 Tencent & NCSOFT. All Rights Reserved.*

*This sequence and all intellectual property rights therein remain the property of Shenzhen Tencent Computer System Co. Ltd. and NCSOFT Co.Ltd.. This sequence may only be used for the purpose of developing, testing and promulgating technology standards (the Purpose). Shenzhen Tencent Computer System Co. Ltd. and NCSOFT Co.Ltd. make no warranties with respect to this sequence and expressly disclaim any warranties regarding its fitness for any purpose. This disclaimer shall remain attached to this sequence except when this sequence is used for the Purpose.*

Annex <X> (informative):  
Change history

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Change history** | | | | | | | |
| **Date** | **Meeting** | **TDoc** | **CR** | **Rev** | **Cat** | **Subject/Comment** | **New version** |
| 2020-04 | SA4#108 | S4-200661 |  |  |  | Initial Version | 0.0.1 |
| 2020-04 | SA4#108 | S4-200666 |  |  |  | Version agreed at SA4#108-e | 0.1.0 |
| 2020-05 | SA4#109 | S4-200891 |  |  |  | Version agreed at SA4#109-e | 0.2.0 |
| 2020-08 | SA4#110 | S4-201243 |  |  |  | Version agreed at SA4#110-e | 0.3.0 |

1. Achieves the objective of high quality for more difficult source material. [↑](#footnote-ref-2)