|  |  |
| --- | --- |
| 3GPP TS 26.265 V0.4.0(2024-11) | |
| Technical Specification | |
| 3rd Generation Partnership Project;  Technical Specification Group Services and System Aspects;  Media Delivery: Video Capabilities and Operation Points (Release 19) | |
|  | |
|  |  |
|  | |
| 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  https://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.  © 2024, 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 4

Introduction 5

1 Scope 6

2 References 6

3 Definitions of terms, symbols and abbreviations 6

3.1 Terms 7

3.2 Symbols 7

3.3 Abbreviations 7

4 Context and Definitions 7

4.1 Motivation 8

4.2 Reference architectures and definitions 9

4.3 Specification 10

4.4 Video representation formats 10

4.4.1 Overview 10

4.4.2 Video signal parameters 10

4.4.3 3GPP Video Formats 13

4.4.3.1 Introduction 13

4.4.3.2 High-Definition TV 13

4.4.3.3 High Dynamic Range TV 14

4.4.3.4 3GPP Stereoscopic Cinema Format 15

4.5 Reference API parameters 16

4.5.1 Introduction 16

4.5.2 Video Decoder API Parameters 16

4.5.3 Video Encoder API Parameters 17

5 Video Coding Capabilities 17

5.1 Overview 17

5.4 Single-Instance Encoding Capabilities 19

5.5 Multi-Instance Decoding Capabilities 20

5.6 Multi-Instance Encoding Capabilities 20

6 Video Operation Points 20

6.1 Introduction 21

7 Video Media Profiles and System Capabilities 21

Annex <A> (normative): Registration Information 21

Annex <B> (informative): Mapping of Reference Architecture to Implementations 22

B.1 Introduction 22

B.2 WebCodecs API 22

Annex <X> (informative): Change history 23

For definitive guidance on drafting 3GPP TSs and TRs, see [3GPP TS 21.801](https://www.3gpp.org/DynaReport/21801.htm).

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

# Foreword

This Technical Specification 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

This clause is optional. If it exists, it shall be the second unnumbered clause.

# 1 Scope

This clause shall start on a new page.

The present document …

# 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".

[bt709] Recommendation ITU-R BT.709-6 (06/2015): "Parameter values for the HDTV standards for production and international programme exchange"

[bt2100] Recommendation ITU-R BT.2100-2 (07/2018): "Image parameter values for high dynamic range television for use in production and international programme exchange"

[h264] Recommendation ITU-T H.264 (08/2021): "Advanced video coding for generic audiovisual services".

[h265] Recommendation ITU-T H.265 (09/2023): "High efficiency video coding".

[h273] Recommendation ITU-T H.273 (09/2023): "Coding-independent code points for video signal type identification".

[h274] Recommendation ITU-T H.274 (09/2023): "Versatile supplemental enhancement information messages for coded video bitstreams".

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

[CENC] ISO/IEC 23001-7: "MPEG systems technologies - Part 7: Common encryption in ISO base media file format files".

[DPC] CTA-5003-B: "Web Application Video Ecosystem (WAVE): Device Playback Capabilities Specification", available at https://shop.cta.tech/products/web-application-video-ecosystem-device-playback-capabilities-cta-5003-b .

[6381] IETF RFC 6381: The 'Codecs' and 'Profiles' Parameters for "Bucket" Media Types.

[MSE] 3GPP TR 26.857, "5G Media Service Enablers"

[3dtv] A. Quested and B. Zegel, "3D-TV production standards - first report of the ITU-R Rapporteurs", EBU Technical Review, 2011 Q2, https://tech.ebu.ch/publications/trev\_2011-Q2\_3dtv\_quested

# 3 Definitions of terms, symbols and abbreviations

This clause and its three (sub) clauses 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 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 TR 21.905 [1].

**Bitstream:** A sequence of bits that conforms to a specific video encoding format and aligns with a certain Operation Point.

**Chroma:** a sample array or single sample representing one of the two colour difference signals related to the primary colours, represented by the symbols *Cb* and *Cr*.

**Hero Eye**: The default eye in a stereo (stereoscopic) video pair, often determined by tags set by the cameras used to capture the video.

**Luma:** a sample array or single sample representing the monochrome signal related to the primary colours (denoted with the symbol *Y*),

**Operation Point:** A collection of discrete combinations of different content formats, including spatial and temporal resolutions, colour mapping, transfer functions, and the encoding format.

**Receiver:** A device capable of decoding and rendering any bitstream that is conforming to a certain Operation Point.

## 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 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 TR 21.905 [1].

Abbreviation format (EW)

<ABBREVIATION> <Expansion>

# 4 Context and Definitions

Editor’s Note from 619, clause 5.1

The principles of existing video capabilities are built around the following principles:

**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.

Decoding capabilities are defined which are a combination of

The capability to decode a bitstream conforming to a certain profile and level

The bitstream being restricted in terms of flags and settings

An illustration of an operation points and decoding capabilities is provided below.

A diagram of a diagram with Crust in the background

Description automatically generated

*Receivers* are a combination of decoding capabilities and the ability to rendering the formats included in an operation point.

The timing and the properties of the format may be signaled in the bitstream, or may be signaled by external means, for example on packaging level, i.e. on ISO BMFF or RTP level.

At the receiving end conformance always refers to real-time decoding and rendering.

Bitstreams can either conform to any of the above “circles”

Codec & Profile

Level

Decoding capabilites

Operation Point

Concurrent decoding capabilities are defined as the ability to decode several bitstreams in parallel.

Encoding capabilities are defined by the ability to encode a *video signal* with certain boundary parameters to a bitstream that is decodable (and possibly can be rendered). Typically, specifications would require real-time encoding.

Decoding capabilities can be shared across many different applications.

Operation Points are more specific towards applications and may not or only partially be defined in a new spec.

## 4.1 Motivation

Video codecs, encoders and decoders are core components of 3GPP services. At the same time, video encoders and decoders residing on 3GPP UEs and defined in 3GPP specifications also provide interoperability points for third-party services. Video capabilities are predominantly independent of the service in use. This specification addresses the definition of video capabilities and operating points such that 3GPP service specifications as well as third-party service providers can refer to the interoperability points defined in this specification.

The present specification makes use some of the concepts recommended in TR 26.857 [2], i.e. the concept of Media Service Enablers.

## 4.2 Reference architectures and definitions

In order to define the normative aspects of this specification, reference architectures are defined. The core architecture is provided in Figure 4.2-1. The workflow addresses the generation of a *video bitstream* from a video signal using a *video encoder* as well as the decoding of a video bitstream by a *video decoder* and providing the resulting decoded video as well as associated metadata to a rendering and display process. The video encoder as well as the video decoder may be configured to certain operations indicated by APIs in Figure 4.2-1. These APIs are not normatively specified but serve as an example reference to configure encoders and decoders as documented in Annex [A].



Figure 4.2-1 Reference architecture for video operating points and capabilities

A more system-centric architecture is provided in Figure 4.2-2. The workflow addresses the generation of a *transport stream* from a video signal using a *video encoder* and a *packager*. The package may include for example timing and metadata information. The de-packaging and decoding of the *transport stream* by a de-packager and a *video decoder*, respectively, allows for providing the resulting video signal as well as associated metadata to a rendering and display process. Again, the packager/encoder as well as the de-packager/decoder may be configured to certain operations indicated by APIs in Figure 4.2-2.



Figure 4.2-2 Reference architecture for system operating points and capabilities

Editor’s Note: A reference architecture for multiple decoders still needs to be defined.

Based on this introduction, the following terms are defined

**Operating Point:** A collection of different possible video formats including spatial and temporal resolutions, colour mapping, transfer functions, etc. and a video encoding format.

**Bitstream**: A compressed media representation presented as a sequence of bits that conforms to a particular video coding specification/format and one or more Operating Points.

**Receiver**: A device that can ingest and decode any bitstream that is conforming to a particular video coding specification and Operating Point, and optionally render it.

In addition, on system level the following terms are defined:

**System Operating Point:** A collection of different possible video formats including spatial and temporal resolutions, colour mapping, transfer functions, etc., a video encoding and a packaging format.

**Transport Stream:** A packaged media bitstream that conforms to a particular video coding and packaging specification/format and one or more Operating Points.

**System Receiver:** A receiver that can de-package and decode any system bitstream that is conforming to a particular System Operating Point, and optionally render it.

## 4.3 Specification

This specification defines the following capabilities:

- Video Decoding capability: The capability to decode any video bitstream that conforms to an operating point and provides a conforming output video signal and possibly associated metadata.

- System Receiver capability: The capability to un-package and decode any transport stream that conforms to a system operating point and provides a conforming output video signal and possibly associated metadata

- Video Encoding capability: The capability to encode any video signal included in the operating point to a bitstream that is decodable by decoder that conforms to the same operating point.

- System Transmitter capability: The capability to encode and package any video signal included in the operating point to a system bitstream that can be unpacked and decoded by a system receiver that conforms to the same operating point.

While not explicitly stated in the capabilities, it is a requirement for decoders and receivers to process the data in real-time. For encoder, real-time encoding is typically also a requirement.

## 4.4 Video representation formats

### 4.4.1 Overview

This clause defines video representation formats in the context of media delivery in 3GPP. For this purpose, a set of video signal parameters are defined in clause 4.4.2, with the restriction on what is defined in 3GPP media delivery. Based on the defined video signal parameters, clause 4.4.3 defines a set of video representation formats.

NOTE: These clause does not specify whether these parameters and formats are required, recommended or suggested to be supported. This aspect is left to specific service specifications or external specifications to refer to the parameters and formats defined in this clause.

### 4.4.2 Video signal parameters

Video signals considered in this specification are represented by a sequence of pictures, where a *picture* can represent either an array of *luma* samples in a monochrome format or an array of luma samples and two corresponding arrays of *chroma* samples in a 4:2:0, 4:2:2, or 4:4:4 colour format. Only *progressive* signals are considered. A component refers to an array or single sample from one of the three arrays (luma and two chroma) that compose a picture. The Luma component represents a sample array or single sample representing the monochrome signal related to the primary colours (denoted with the symbol *Y*), and a chroma component represents a sample array or single sample representing one of the two colour difference signals related to the primary colours, represented by the symbols *Cb* and *Cr*.

Video signals are typically described by a set of parameters that are required for the proper rendering of the decoded signal. Table 4.4.2-1 documents typical video signal parameters and provides a definition and/or reference.

Table 4.4.2-1 Video Signal Parameters

|  |  |  |
| --- | --- | --- |
| Parameter | Definition | 3GPP restrictions |
| Spatial Resolution width | the number of active samples per line for the luma component.  Example values are 1280 or 1920 for HD, and 3840 for UHD.  NOTE: The width does not restrict the encoding resolution to fixed values. Cropping parameters can be indicated that prescribe decoders the need to remove spatial video samples in a partially filled coding block that are not intended for presentation. | no direct restrictions, but services may provide subsets. |
| Spatial Resolution height | the number of active lines per picture for the luma component.  Typical values are 720 or 1080.  NOTE: The height does not restrict the encoding resolution to fixed values. Cropping parameters can be indicated that prescribe decoders the need to remove spatial video samples in a partially filled coding block that are not intended for presentation. | no direct restrictions, but services may provide subsets. |
| Scan Type | indicates the source scan type of the pictures as defined in clause 7.3 of Rec. ITU-T H.273.  Typical value is progressive | progressive only |
| Chroma format indicator | indicates whether the picture has only a luma component or that the picture has three colour components that consist of a luma component and two associated chroma components, such that the width and height of each chroma component are the width and height of the luma component divided by a factor defined by the chroma format as defined in Rec. ITU-T H.274, clause 7.3. | 4:2:0 |
| Bit depth | Indicates the bit depth for the samples of the luma component and the samples of the two associated chroma components.  Note that in general, the bit depth of the luma component and of the two associated chroma components may differ.  Typical values are 8 or 10 bits. | 8 or 10 bits |
| Colour primaries | indicates the chromaticity coordinates of the source colour primaries as specified in clause 8.1 of Rec. ITU-T H.273.  Typical values are 1 to refer to Rec. ITU-R BT.709-6 [bt709] or 9 to refer to Rec. ITU-R BT.2020-2 and Rec. ITU-R BT.2100-2. | BT.709 or BT.2020/BT.2100 |
| Transfer Characteristics | either indicates the reference opto-electronic transfer characteristic function of the source picture as a function of a source input linear optical intensity input or indicates the inverse of the reference electro-optical transfer characteristic function as a function of an output linear optical intensity as defined in clause 8.2 of Rec. ITU-T H.273.  Typical values are 1 to refer to Rec. ITU-R BT.709-6, 14 to refer to Rec. ITU-R BT.2020-2 (10 bit), 16 to refer to Rec. ITU-R BT.2100-2 perceptual quantization (PQ) system, or 18 to refer to Rec. ITU-R BT.2100-2 hybrid log-gamma (HLG) system | BT.709 SDR, BT.2020 SDR, BT.2100 PQ, or BT.2100 HLG |
| Matrix Coefficients | describes the matrix coefficients used in deriving luma and chroma signals from the green, blue, and red primaries. A video full range flag may be supplied with this parameter specifying the scaling and offset values applied in association with the Matrix coefficients. For detailed definition refer to clause 8.2 of Rec. ITU-T H.273.  Typical values are 1 to refer to the non constant luminance YCbCr representation Rec. ITU-R BT.709-6 or 9 to refer to the non constant luminance YCbCr representations in Rec. ITU-R BT.2020-2 and Rec. ITU-R BT.2100-2. | BT.709 or BT.2020/BT.2100 |
| Frame rate | Typical values, using frames per second, are: 120, 120/1.001, 100, 60, 60/1.001, 50, 30, 30/1.001, 25, 24, 24/1.001 | no direct restrictions, but services may only permit a restricted subset. |
| Frame packing | indicates a frame packing arrangement, if present, as defined in clause 8.4 of Rec. ITU-T H.273. | Typically restricted to no frame packing, but applications may use frame packing. |
| Projection | indicates a projection, if present, as defined in Rec. ITU-T H.274, clause 7.3, and typically refers to packing arrangements in clause 8.6 of Rec. ITU-T H.274. | Typically restricted to no projection, but applications may use projections. |
| Sample aspect ratio | indicates width-to-height aspect ratio of the luma samples of the associated pictures as defined in clause 7.3 of Rec. ITU-T H.273.  Typical value is 1 | No specific restrictions, but 1 is expected. |
| Chroma sample location type | specifies the location of the chroma samples relative to the luma samples for frames as defined in Rec. ITU-T H.273, clause 8.7.  Typical values are 0 (chroma samples are horizontally co-sited with and vertically centered between the first luma sample at the top-left corner and the first two luma samples at the top-left corner, respectively) or 2 (chroma samples are co-sited with the luma sample at the top-left corner). Note that 1 is common for still images. | No specific restrictions, but 0 is expected if not present. For HDR the value is typically set to 2. |
| Range | Specifies how luma and chroma samples are represented in digital video as defined in Rec. ITU-T H.273, clause 8.3 using the VideoFullRangeFlag.  Only the value set to 0 is used, i.e. the video range or restricted range is applied where the luma values range from 16 to 235 in an 8-bit system, and chroma values range from 16 to 240. For 10-bit systems, the values are multiplied by 4. Note that for still images full range is commonly used. | No specific restrictions, but 0 is expected if not present. |
| Stereoscopic Video | Visual media may be stereoscopic, in which a view is available to be presented to the left eye and another view is available to be presented simultaneously to the right eye. The presentation of both the left and right views allows for an effect known as stereopsis, which can be defined as "the perception of depth produced by the reception in the brain of visual stimuli from both eyes in combination; binocular vision." For signal representations, [3dtv] recommends that the Left and Right eyes comply to regular image formats such as Rec. ITU-R BT.709 and any necessary 3D-specific metadata is incorporated with the data. Hence, for stereoscopic video, two synchronized video signals are available, each with identical format parameters (such as the ones defined in this table).  Additional metadata that may be added with stereoscopic video:  - hero eye is the default eye in a stereo (stereoscopic) video pair, often determined by tags set by the cameras used to capture the video. If so signaled, this indicates the other stereo eye view is derived from the specified stereo eye and may be useful when choosing which eye to use in a monoscopic viewing environment. There is no requirement that either of the two eyes (or views) is tagged as the hero eye in which case no hero eye tagging may be present. |  |

### 4.4.3 3GPP Video Formats

#### 4.4.3.1 Introduction

While a variety of formats may be used based on the video signal parameters defined in clause 4.4.2, for consistent programs and signals, several video formats are defined by a set of restrictions using the video signal parameters in clause 4.4.2. These signals are primarily used to distribute TV and movie content.

New 3GPP Video formats may be defined.

#### 4.4.3.2 High-Definition TV

3GPP High-Definition TV (HDTV) formats are defined based on Rec. ITU-R BT-709-6 [bt709]. 3GPP HDTV formats shall conform to Rec. ITU-R BT-709-6 [bt709] with the following restrictions:

- Only the following formats are included 24/P, 25/P, 30P, 50/P and 60/P. Interlace and progressive segmented frame signals are excluded.

- Only the Non-Constant Luminance Y'C'BC'R signal format is considered.

An informative summary of the parameters of a 3GPP HDTV format based on the parameters defined in Table 4.4.2-1 is provided in Table 4.4.3.2-1.

Table 4.4.3.2-1 Video Signal Parameters for 3GPP HDTV format

|  |  |
| --- | --- |
| Parameter | Restrictions |
| Spatial Resolution width | the number of active samples per line is 1920. |
| Spatial Resolution height | the number of active lines per picture for the luma component is 1080.  NOTE: Typically the encoded signal has 1088 lines and cropping is applied to remove spatial samples that are not presented. |
| Scan Type | the source scan type of the pictures as defined in clause 7.3 of Rec. ITU-T H.273 is progressive |
| Chroma format indicator | The chroma format indicator is 4:2:0. |
| Bit depth | The permitted values are 8 or 10 bit. |
| Colour primaries | Only the value 1 as defined in clause 8.2 of Rec. ITU-T H.273 is permitted. |
| Transfer Characteristics | Only the value 1 as defined in clause 8.2 of Rec. ITU-T H.273 is permitted. |
| Matrix Coefficients | Only the value 1 as defined in clause 8.2 of Rec. ITU-T H.273 is permitted. |
| Frame rates | The permitted values are 60, 60/1.001, 50, 30, 30/1.001, 25, 24, 24/1.001 fps. |
| Frame packing | No frame packing is applied. |
| Projection | No projection is used. |
| Sample aspect ratio | The pixel aspect ratio is 1 (square pixel), i.e. only the value 1 as defined in clause 7.3 of Rec. ITU-T H.273 is permitted. |
| Chroma sample location type | the location of chroma samples relative to the luma samples for progressive frames as defined in Rec. ITU-T H.273, clause 8.7 is set to 0 (chroma samples are horizontally co-sited with and vertically centered between the first luma sample at the top-left corner and the first two luma samples at the top-left corner, respectively). |
| Range | The restricted video range is used. |

#### 4.4.3.3 High Dynamic Range TV

3GPP High Dynamic Range (HDR) TV formats are defined based on Rec. ITU-R BT-2100-2 [bt2100]. 3GPP HDR TV formats shall conform to ITU-R BT-2100-2 [bt2100] with the following restrictions:

- Only 4:2:0 colour subsampling is considered

- Only the Non-Constant Luminance Y'C'BC'R signal format is considered

- Only 10-bit representations are considered

An informative summary of the parameters of a 3GPP HDR TV format based on the parameters defined in Table 4.4.2-1 is provided in Table 4.4.3.3-1.

Table 4.4.3.3-1 Video Signal Parameters for 3GPP HDR TV format

|  |  |
| --- | --- |
| Parameter | Restrictions |
| Picture aspect ratio | 16:9 |
| Spatial Resolution width x height | 7680 × 4320, 3840 × 2160, 1920 × 1080  NOTE: For 1080, typically the encoded signal has 1088 lines and cropping is applied to remove spatial samples that are not presented. |
| Scan Type | the source scan type of the pictures as defined in clause 7.3 of Rec. ITU-T H.273 is progressive |
| Chroma format indicator | The chroma format indicator is 4:2:0. |
| Bit depth | The permitted value is 10 bit. |
| Colour primaries | Only the value 9 as defined in clause 8.2 of Rec. ITU-T H.273 is permitted. |
| Transfer Characteristics | Only the values 16 (for PQ) and 18 (for HLG) as defined in clause 8.2 of Rec. ITU-T H.273 are permitted.  Editor’s Note: How about BT.2020 SDR signals? They should be added as well, proper integration needs to be done. |
| Matrix Coefficients | Only the value 9 as defined in clause 8.2 of Rec. ITU-T H.273 is permitted. |
| Frame rates | The permitted values are 120, 120/1.001,100, 60, 60/1.001, 50, 30, 30/1.001, 25, 24, 24/1.001 fps. |
| Frame packing | No frame packing is applied. |
| Projection | No projection is used. |
| Sample aspect ratio | The pixel aspect ratio is 1 (square pixel), i.e. only the value 1 as defined in clause 7.3 of Rec. ITU-T H.273 is permitted. |
| Chroma sample location type | the location of chroma samples relative to the luma samples for progressive frames as defined in Rec. ITU-T H.273, clause 8.7 is set to 2 (chroma samples are co-sited with the luma samples at the top-left corner). |
| Range | The restricted video range is used. |

#### 4.4.3.4 3GPP Stereoscopic Cinema Format

The stereoscopic 3D TV format uses two signals, one for the left eye and another view for the right eye as defined in Table 4.4.2-1. The components for each eye closely follow the specifications of the 3GPP HDR signals, but there are some restrictions and extensions, namely:

- Only 4:2:0 colour subsampling is considered

- Frame rates include high frame rate for movies, namely 48 fps.

- the spatial resolution is restricted to a maximum value of 4K

- Only the Non-Constant Luminance Y'C'BC'R signal format is considered

An informative summary of the parameters of a 3GPP Stereoscopic 3D TV format based on the parameters defined in Table 4.4.2-1 is provided in Table 4.4.3.4-1.

Table 4.4.3.4-1 Video Signal Parameters for 3GPP Stereoscopic 3D Cinema format

|  |  |
| --- | --- |
| Parameter | Restrictions |
| Picture aspect ratio | 16:9 |
| Spatial Resolution width x height | 3840 × 2160, 1920 × 1080  NOTE: For 1080, typically the encoded signal has 1088 lines and cropping is applied to remove spatial samples that are not presented. |
| Scan Type | the source scan type of the pictures as defined in clause 7.3 of Rec. ITU-T H.273 is progressive |
| Chroma format indicator | The chroma format indicator is 4:2:0. |
| Bit depth | The permitted values are 8 or 10 bit. 8 bit is only permitted for SDR. |

|  |  |
| --- | --- |
| Colour primaries  Transfer Characteristics  Matrix Coefficients | Only the following value combinations are permitted: (1, 1, 1), (9,14, 9), (9, 16, 9), and (9, 18, 9) for SDR HD, SDR UHD, HDR PQ, and HDR HLG, respectively. |
| Frame rates | The permitted values are 60, 60/1.001, 48, 48/1.001, 50, 30, 30/1.001, 25, 24, 24/1.001 fps. |
| Frame packing | No frame packing is applied. |
| Projection | No projection is used. |
| Sample aspect ratio | The pixel aspect ratio is 1 (square pixel), i.e. only the value 1 as defined in clause 7.3 of Rec. ITU-T H.273 is permitted. |
| Chroma sample location type | For SDR HD, the location of chroma samples relative to the luma samples for progressive frames as defined in Rec. ITU-T H.273, clause 8.7 is set to 0.  For SDR UHD, HDR PQ, and HLG, the location of chroma samples relative to the luma samples for progressive frames as defined in Rec. ITU-T H.273, clause 8.7 is set to 2. |
| Range | The restricted video range is used. |
| Stereoscopic Video | A signal for the Left and for the Right Eye is provided whereby the signals have the identical parameters as above and are timely synchronized. |

## 4.5 Reference API parameters

### 4.5.1 Introduction

When media is played back, the decoder and the playback pipeline need to be initialized. For this purpose, certain parameters are required. In CTA-5003 [DPC], a media playback model is described that is aligned with with HTML 5.1 and the <video> element, as well as the Media Source Extensions.

### 4.5.2 Video Decoder API Parameters

Based on CTA-5003 [DPC], Table 4.5.2-1 provide relevant parameters that need to be attached to the content, in order to establish media playback properly, and serve as an API. The parameters are used for the following purposes:

- to identify the capability of the device in order to check of the signal can be played back

- to initialize the decoding and playback platform to allocate the resources for decoding and rendering

Table 4.4.3.4-1 Video Signal Parameters for 3GPP Stereoscopic 3D TV format

|  |  |  |
| --- | --- | --- |
| Parameter | Restrictions | Status |
| width | specifies the width of a video player, in pixels | required |
| height | specifies the width of a video player, in pixels. | required |
| media type | specifies the media type of the component, in this case video | required |
| format | specifies the format of the media, for example mp4 | required |
| profiles | specifies the profile of the format, for example 'cmfc' | optional |
| codecs | specifies through a well-defined string the codec used for the signal | required |
| Video format parameters | specifies additional video format parameters as defined in Table 4.4.2.1 to describe the signal and to initialize the encoder. | optional |

### 4.5.3 Video Encoder API Parameters

Video encoder API parameters are for further study.

# 5 Video Coding Capabilities

## 5.1 Overview

This clause defines video decoding capabilities and video encoding capabilities for 3GPP media delivery.

NOTE: These clause does not specify whether these capabilities are required, recommended or suggested to be supported. This aspect is left specific service specifications or external specifications to refer to the capabilities defined in this clause.

5.2 Codecs, Profiles and Levels

5.2.1 Codec & profile

This specification defines capabilities based on the following video codecs and video codec profiles:

- AVC/H.264 Progressive High Profile [h264],

- HEVC/H.265 Main Profile Main Tier [h265],

- HEVC/H.265 Main-10 Profile Main Tier [h265].

- HEVC/H.265 Multiview Main 10 Main Tier [h265].

- HEVC/H.265 Multiview Extended 10 Tier [h265].

5.2.2 Codec & profile & Levels

This specification defines capabilities based on the following video codec profile and levels:

- AVC/H.264 Progressive High Profile Level 3.1,

- AVC/H.264 Progressive High Profile Level 4.0,

- AVC/H.264 Progressive High Profile Level 4.2,

- AVC/H.264 Progressive High Profile Level 5.1,

- AVC/H.264 Progressive High Profile Level 6.1,

- HEVC/H.265 Main Profile Main Tier Level 3.1,

- HEVC/H.265 Main-10 Profile Main Tier Level 4.1,

- HEVC/H.265 Main-10 Profile Main Tier Level 5.0,

- HEVC/H.265 Main-10 Profile Main Tier Level 5.1,

- HEVC/H.265 Main-10 Profile Main Tier Level 6.1,

- HEVC/H.265 Multiview Main 10 Profile Main Tier Level 5.1,

- HEVC/H.265 Multiview Extended 10 Profile Main Tier Level 5.1.

5.3 Single-Instance Decoding Capabilities

Editor’s Note: This is copy and paste from S4-240619, clause 5.2.3. More edits are needed.

5.3.1 AVC Decoding Capabilities

The following decoding capabilities are defined:

**- AVC-FullHD-Dec**: the capability to decode AVC/ITU-T H.264 Progressive High Profile Level 4.0 [h264] bitstreams.

**- AVC-UHD-Dec:** the capability to decode AVC/ITU-T H.264 Progressive High Profile Level 5.1 [h264] bitstreams with the following additional requirements:

- the maximum VCL Bit Rate is constrained to be 120 Mbps with cpbBrVclFactor and cpbBrNalFactor being fixed to be 1250 and 1500, respectively; and,

- the bitstream does not contain more than 10 slices per picture.

**- AVC-8K-Dec:** the capability to decode AVC/ITU-T H.264 Progressive High Profile Level 6.1 [h264] bitstreams with the following requirements:

- the maximum VCL Bit Rate is constrained to be 120 Mbps with cpbBrVclFactor and cpbBrNalFactor being fixed to be 1250 and 1500, respectively; and,

- the bitstream does not contain more than 16 slices per picture.

- the bitstream shall not include horizontal motion vector component values that exceed the range from −2048 to 2047, inclusive, or that have vertical motion vector component values that exceed the range from −512 to 511, inclusive, in units of ¼ luma sample displacement. This constraint should be indicated by using values of log2\_max\_mv\_length\_horizontal less than or equal to 11 and values of log2\_max\_mv\_length\_vertical less than or equal to 9.

5.3.2 HEVC Decoding Capabilities

The following decoding capabilities are defined:

- **HEVC-HD-Dec**: the capability to decode HEVC/ITU-T H.265 Main Profile, Main Tier, Level 3.1 [h265] 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.

- **HEVC-FullHD-Dec**: the capability to decode HEVC/ITU-T H.265 Main 10 Profile, Main Tier, Level 4.1 [h265] 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.

- **HEVC-UHD-Dec**: the capability to decode HEVC/ITU-T H.265 Main 10 Profile, Main Tier, Level 5.1 [h265] 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.

- **HEVC-8K-Dec**: the capability to decode HEVC/ITU-T H.265 Main10 Profile, Main Tier, Level 6.1 [h265] 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 with the following further limitations:

- **MV-HEVC-UHD-Dec**: the capability to decode bitstreams with an HEVC/ITU-T H.265 Main 10 Profile base layer (layer\_id=0), and a single HEVC/ITU-T H.265 Multiview Main 10 [or Multiview Extended 10] layer (layer\_id=1) [h265]. Each layer shall conform to Main Tier, Level 5.1, while the device should be capable of supporting single layer decoding of HEVC/ITU-T H.265 Main 10 Profile bitstreams at Main Tier, Level 5.2. All layers shall 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.

Editor’s Note: Adding operating point(s) for 8k stereoscopic is FFS.

Editor’s Note: Adding operating point(s) for frame packed stereoscopic video is FFS.

- the bitstream does not exceed the maximum luma picture size in samples of 33,554,432,

- the maximum VCL Bit Rate is constrained to be 80 Mbps with CpbVclFactor and CpbNalFactor being fixed to be 1000 and 1100, respectively.

## 5.4 Single-Instance Encoding Capabilities

Editor’s Note: This is copy and paste from S4-240619, clause 5.2.4. More edits are needed.

The following encoding capabilities are defined:

**- AVC-FullHD-Enc:** the capability to encode a video signal to a bitstream that is decodable by a decoder that is *AVC-FullHD-Dec* capable as defined in clause 5.3 with the following additional constraints:

- up to 245,760 macroblocks per second;

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

- up to 240 frames per second;

- the chroma format being 4:2:0; and

- the bit depth being 8 bit;

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

- up to 33,177,600 luma samples per second;

- up to a luma picture size of 983,040 samples;

- up to 120 frames per second;

- 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 **HEVC-HD-Dec** capable as defined in clause 5.3.

**- HEVC-FullHD-Enc:** the capability to encode a video signal to a bitstream that is decodable by a decoder that is *HEVC-FullHD-Dec* capable as defined in clause 5.3 with the following additional constraints:

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

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

- up to 240 frames per second;

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

- the bit depth being either 8 or 10 bit;

**- HEVC-UHD-Enc:** the capability to encode a video signal to a bitstream that is decodable by a decoder that is *HEVC-UHD-Dec* capable as defined in clause 5.3 with the following additional constraints:

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

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

- up to 480 frames per second;

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

- the bit depth being either 8 or 10 bit;

## 5.5 Multi-Instance Decoding Capabilities

Editor’s Note: This is copy and paste from S4-240619, clause 5.2.5. More edits are needed.

The following multi-instance decoding capabilities are defined:

**- AVC-FullHD-Dec-2**: The capability of supporting up to two (*N*=2) concurrent decoder instances with the aggregate capabilities of *AVC-FullHD-Dec* as defined in clause 5.4.

**- AVC-UHD-Dec-4**: The capability of supporting up to four (*N*=4) concurrent decoder instances with the aggregate capabilities of *AVC-UHD-Dec* as defined in clause 5.4.

**- HEVC-UHD-Dec-4:** The capability of supporting up to four (*N*=4) concurrent decoder instances with the aggregate capabilities of *HEVC-UHD-Dec* as defined in clause 5.4.

**- UHD-Dec-4**: The capability supporting up to four (*N*=4) concurrent decoder instances with either:

- the aggregate capabilities of *AVC-UHD-Dec-4* as defined in this clause,

- the aggregate capabilities of *HEVC-UHD-Dec-4* as defined in this clause, or,

- the capability of decoding up to 4 bitstreams for which each bitstream does not exceed the capability of being decodable either with *AVC-FullHD-Dec* or *HEVC-FullHD-Dec* as defined in clause 5.4.

**- AVC-8K-Dec-8:** The capability of supporting up to eight (*N*=8)concurrent decoder instances with the aggregate capabilities of *AVC-8K-Dec* as defined in clause 5.4.

**- HEVC-8K-Dec-8:** The capability of supporting up to eight (*N*=8)concurrent decoder instances with the aggregate capabilities of *HEVC-8K-Dec* as defined in clause 5.4.

**- 8K-Dec-8**: The capability supporting up to eight (*N*=8)concurrent decoder instances with either:

- the aggregate capabilities of *AVC-8K-Dec-8* as defined in this clause,

- the aggregate capabilities of *HEVC-8K-Dec-8* as defined in this clause, or,

- the capability of decoding up to:

- eight bitstreams for which each bitstream does not exceed the capability of being decodable either with *AVC-FullHD-Dec* or *HEVC-FullHD-Dec* as defined in clause 5.4; or,

- four bitstreams for which each bitstream does not exceed the capability of being decodable either with *AVC-UHD-Dec* or *HEVC-UHD-Dec* as defined in clause 5.4.

## 5.6 Multi-Instance Encoding Capabilities

This specification does not define multi-instance encoding capabilities.

# 6 Video Operation Points

Editor’s Note: A collection of different possible video formats including spatial and temporal resolutions, colour mapping, transfer functions, etc. and a video encoding format.

* See again S4-240619 for exising ones

## 6.1 Introduction

Video operation points define a restricted subset of representation signals and media capabilities.

Table 6.1-1 provides an overview of defined video operation points.

Table 6.1-1 Video Operation Points

|  |  |  |  |
| --- | --- | --- | --- |
| Name | Video Format | Decoding Capabilities | Definition |
| 3GPP-AVC-HDTV | 3GPP-HDTV (see clause 4.4.3.2) | AVC-FullHD-Dec (see clause 5.4) | 6.2 |
| 3GPP-HEVC-HDTV | 3GPP-HDTV (see clause 4.4.3.2) | HEVC-FullHD-Dec (see clause 5.4) | 6.3 |
| 3GPP-HEVC-HD-HDR | 3GPP-HDR (see clause 4.4.3.3) | HEVC-FullHD-Dec (see clause 5.4) | 6.4 |
| 3GPP-HEVC-UHD-HDR | 3GPP-HDR (see clause 4.4.3.3) | HEVC-UHD-Dec (see clause 5.4) | 6.5 |
| 3GPP-HEVC-3DTV | 3GPP-3DTV (see clause 4.4.3.4) | HEVC-UHD-Dec-2 (see clause 5.5) | 6.6 |
| 3GPP-MVHEVC-3DTV | 3GPP-3DTV (see clause 4.4.3.4) | MVHEVC-UHD-2 (see clause 5.5) | 6.7 |

# 7 Video Media Profiles and System Capabilities

7.1 Introduction

System operation points define a restricted subset of video operation points mapped to delivery options.

Table 7.1-1 provides an overview of defined system operation points. Note that the operation points need to conform to both requirements, so they are an intersection.

Editor’s Note: This table is considered as a starting point and alignment with above clauses needs to happen in terms of terminology.

[

**Table 7.1-1 System Operation Points**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Video Operation Point** | **CMAF Media Profile** | **Definition** |
| 3GPP-AVC-HDTV-CMAF | 3GPP-AVC-HDTV | 'cfhd' (see CMAF) | 7.2 |
| 3GPP-HEVC-HDTV-CMAF | 3GPP-HEVC-HDTV | 'chh1’ | 7.3 |
| 3GPP-HEVC-HD-HDR-CMAF | 3GPP-HEVC-HD-HDR | 'chd1' or 'clg1' | 7.4 |
| 3GPP-HEVC-UHD-HDR-CMAF | 3GPP-HEVC-UHD-HDR | 'chd1' or 'clg1' | 7.5 |
| 3GPP-HEVC-3DTV-CMAF | 3GPP-HEVC-3DTV | Not defined yet | 7.6 |
| 3GPP-MVHEVC-3DTV-CMAF | 3GPP-MVHEVC-3DTV | Not defined yet | 7.7 |

]

Editor’s Note: In the remainder of the clause, mapping to DASH delivery needs to be done in alignment with TS 26.116.

Annex <A> (normative):  
Registration Information

Editor’s Note: Will collect and registration information such as URNs.

Annex <B> (informative):  
Mapping of Reference Architecture to Implementations

# B.1 Introduction

This annex provides some background on how to map the reference architectures defined in clause 4 into concrete implementations. The mapping of the capabilities, the configuration of the encoders and decoders through APIs as well as some workflow aspects are provided.

The Annex is not considered to prescribe any implementation but is expected to support implementors to integrate the capabilities and operating points defined in this specification into their workflows.

The Annex also serves as an analyis on what functionalities are available in existing implementations and where there are potential gaps that may be addressed by the owners of the implementation to fully support all features.

# B.2 WebCodecs API

Editor’s Note: Analyze the configuration information with the APIs defined in WebCodecs. More work on this is needed.

The configuration of the codec is here

dictionary ***VideoDecoderConfig*** {

required [DOMString](https://webidl.spec.whatwg.org/#idl-DOMString) [codec](https://www.w3.org/TR/webcodecs/#dom-videodecoderconfig-codec);

[AllowSharedBufferSource](https://webidl.spec.whatwg.org/#AllowSharedBufferSource) [description](https://www.w3.org/TR/webcodecs/#dom-videodecoderconfig-description);

[[EnforceRange](https://webidl.spec.whatwg.org/#EnforceRange)] [unsigned long](https://webidl.spec.whatwg.org/#idl-unsigned-long) [codedWidth](https://www.w3.org/TR/webcodecs/#dom-videodecoderconfig-codedwidth);

[[EnforceRange](https://webidl.spec.whatwg.org/#EnforceRange)] [unsigned long](https://webidl.spec.whatwg.org/#idl-unsigned-long) [codedHeight](https://www.w3.org/TR/webcodecs/#dom-videodecoderconfig-codedheight);

[[EnforceRange](https://webidl.spec.whatwg.org/#EnforceRange)] [unsigned long](https://webidl.spec.whatwg.org/#idl-unsigned-long) [displayAspectWidth](https://www.w3.org/TR/webcodecs/#dom-videodecoderconfig-displayaspectwidth);

[[EnforceRange](https://webidl.spec.whatwg.org/#EnforceRange)] [unsigned long](https://webidl.spec.whatwg.org/#idl-unsigned-long) [displayAspectHeight](https://www.w3.org/TR/webcodecs/#dom-videodecoderconfig-displayaspectheight);

[VideoColorSpaceInit](https://www.w3.org/TR/webcodecs/#dictdef-videocolorspaceinit) [colorSpace](https://www.w3.org/TR/webcodecs/#dom-videodecoderconfig-colorspace);

[HardwareAcceleration](https://www.w3.org/TR/webcodecs/#enumdef-hardwareacceleration) [hardwareAcceleration](https://www.w3.org/TR/webcodecs/#dom-videodecoderconfig-hardwareacceleration) = "no-preference";

[boolean](https://webidl.spec.whatwg.org/#idl-boolean) [optimizeForLatency](https://www.w3.org/TR/webcodecs/#dom-videodecoderconfig-optimizeforlatency);

};

dictionary ***VideoEncoderConfig*** {

required [DOMString](https://webidl.spec.whatwg.org/#idl-DOMString) [codec](https://www.w3.org/TR/webcodecs/#dom-videoencoderconfig-codec);

[[EnforceRange](https://webidl.spec.whatwg.org/#EnforceRange)] required [unsigned long](https://webidl.spec.whatwg.org/#idl-unsigned-long) [width](https://www.w3.org/TR/webcodecs/#dom-videoencoderconfig-width);

[[EnforceRange](https://webidl.spec.whatwg.org/#EnforceRange)] required [unsigned long](https://webidl.spec.whatwg.org/#idl-unsigned-long) [height](https://www.w3.org/TR/webcodecs/#dom-videoencoderconfig-height);

[[EnforceRange](https://webidl.spec.whatwg.org/#EnforceRange)] [unsigned long](https://webidl.spec.whatwg.org/#idl-unsigned-long) [displayWidth](https://www.w3.org/TR/webcodecs/#dom-videoencoderconfig-displaywidth);

[[EnforceRange](https://webidl.spec.whatwg.org/#EnforceRange)] [unsigned long](https://webidl.spec.whatwg.org/#idl-unsigned-long) [displayHeight](https://www.w3.org/TR/webcodecs/#dom-videoencoderconfig-displayheight);

[[EnforceRange](https://webidl.spec.whatwg.org/#EnforceRange)] [unsigned long long](https://webidl.spec.whatwg.org/#idl-unsigned-long-long) [bitrate](https://www.w3.org/TR/webcodecs/#dom-videoencoderconfig-bitrate);

[double](https://webidl.spec.whatwg.org/#idl-double) [framerate](https://www.w3.org/TR/webcodecs/#dom-videoencoderconfig-framerate);

[HardwareAcceleration](https://www.w3.org/TR/webcodecs/#enumdef-hardwareacceleration) [hardwareAcceleration](https://www.w3.org/TR/webcodecs/#dom-videoencoderconfig-hardwareacceleration) = "no-preference";

[AlphaOption](https://www.w3.org/TR/webcodecs/#enumdef-alphaoption) [alpha](https://www.w3.org/TR/webcodecs/#dom-videoencoderconfig-alpha) = "discard";

[DOMString](https://webidl.spec.whatwg.org/#idl-DOMString) [scalabilityMode](https://www.w3.org/TR/webcodecs/#dom-videoencoderconfig-scalabilitymode);

[VideoEncoderBitrateMode](https://www.w3.org/TR/webcodecs/#enumdef-videoencoderbitratemode) [bitrateMode](https://www.w3.org/TR/webcodecs/#dom-videoencoderconfig-bitratemode) = "variable";

[LatencyMode](https://www.w3.org/TR/webcodecs/#enumdef-latencymode) [latencyMode](https://www.w3.org/TR/webcodecs/#dom-videoencoderconfig-latencymode) = "quality";

[DOMString](https://webidl.spec.whatwg.org/#idl-DOMString) [contentHint](https://www.w3.org/TR/webcodecs/#dom-videoencoderconfig-contenthint);

};

For video codec registry, see here: <https://www.w3.org/TR/webcodecs-codec-registry/#video-codec-registry>

For HEVC codec registrations, please go here: <https://www.w3.org/TR/webcodecs-hevc-codec-registration/>

Annex <X> (informative):  
Change history

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Change history | | | | | | | |
| Date | Meeting | TDoc | CR | Rev | Cat | Subject/Comment | New version |
| 2024-04 | SA4#127bis-e | S4-240616 |  |  |  | Initial version | 0.0.0 |
| 2024-04 | SA4#127bis-e | S4-240758 |  |  |  | Version agreed at SA4#127bis-e | 0.1.0 |
| 2024-05 | SA4#128 | S4-241369 |  |  |  | Version agreed at SA4#128 including S4-240911, S4-241296, S4-241298 | 0.2.2 |
| 2024-08 | SA4#129-e | S4-241669 |  |  |  | Version agreed at SA4#129-e including S4-241479, S4-241705 | 0.3.0 |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 2024-10 | Post SA4#129-e Video SWG AHG | S4aV240073 |  |  |  | Version agreed during Post SA4#129-e Video SWG AHG October 29, 2024 adding S4aV240060 | 0.3.1 |
| 2024-11 | SA4#130 | S4-241892 |  |  |  | Version submitted for SA4#130 adding agreed S4aV240073 | 0.3.2 |
| 2024-11 | SA4#130 | S4-242064 |  |  |  | Version agreed at SA4#130 including S4-241894, S4-242174, S4-242209, S4-242211 | 0.4.0 |

,