ISO/IEC JTC 1/SC 29/WG 03 N1567

**ISO/IEC JTC 1/SC 29/WG 03  
MPEG Systems   
Convenorship: KATS (Korea, Republic of)**

**Document type:** Output Document

**Title:** Potential improvement of ISO/IEC 23000-19 3rd edition DAM 2 New Structural CMAF Brand Profile

**Status:** Approved

**Date of document:** 2025-07-11

**Source:** ISO/IEC JTC 1/SC 29/WG 03

**No. of pages:** 25 (with cover page)

**Email of Convenor:** young.L@samsung.com

**Committee URL:** <https://isotc.iso.org/livelink/livelink/open/jtc1sc29wg3>

**INTERNATIONAL ORGANIZATION FOR STANDARDIZATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC 1/SC 29/WG 03 MPEG SYSTEMS**

**ISO/IEC JTC 1/SC 29/WG 03 N1567**

**July 2025 – Daejeon, KR**

|  |  |
| --- | --- |
| **Title** | **Potential improvement of ISO/IEC 23000-19 3rd edition DAM 2 New Structural CMAF Brand Profile** |
| **Source** | **WG 03, MPEG Systems** |
| **Status** | **Approved** |
| **Serial Number** | **25314** |

**ISO 23000-19:2025(X)**

ISO/IEC JTC1 /SC 29 /WG 03 /N0XXX

Secretariat: XXXX

Information technology — Multimedia application format (MPEG-A) — Part 19: Common media application format (CMAF) for segmented media, AMENDMENT 2: Additional structural CMAF brand profile

DIS stage

**Warning for WDs and CDs**

This document is not an ISO International Standard. It is distributed for review and comment. It is subject to change without notice and may not be referred to as an International Standard.

Recipients of this draft are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation.

© ISO 2020

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office

Case postale 56 • CH-1211 Geneva 20

Tel.  + 41 22 749 01 11

Fax  + 41 22 749 09 47

E-mail  copyright@iso.org

Web  www.iso.org

Published in Switzerland.

# Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2. [www.iso.org/directives](http://www.iso.org/directives)

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received. [www.iso.org/patents](http://www.iso.org/patents)

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the WTO principles in the Technical Barriers to Trade (TBT) see the following URL: [Foreword - Supplementary information](http://www.iso.org/iso/home/standards_development/resources-for-technical-work/foreword.htm)

The committee responsible for this document is ISO/IEC JTC1 SC29.

The main changes compared to the previous edition are as follows:

* definition of a set of MV-HEVC media profiles
* carriage of metadata tracks in a single CMAF track
* Definition of the cmf1 brand

# Clause 3, Terms and definitions

Add the following note after the definition of the CMAF track (3.2.1)

Note: In addition to a Primary Media Track (3.2.14) the CMAF track may also contain Supplementary Metadata Tracks (3.2.15).

Add 2 new definitions to Clause 3.2

**3.2.14**

**Primary Media Track**

The primary media track refers to the main content track that carries the audio, video, or audiovisual data. This track is essential for the core playback experience and is always present in the MovieBox.

**3.2.15**

**Supplementary Metadata Track**

A Supplementary Metadata Track refers to a metadata track that is directly related to the media data of the Primary Media Track (3.2.14) within the same CMAF track (3.2.1). This track uses a handler type equal to 'meta' and provides metadata that enhances or describes the associated primary media content.

Add 3 new definitions to Clause 3.3

**3.3.19**

**Texture picture**

A picture that carries pixel data corresponding to information that can be rendered on a screen or some other device, either directly or indirectly after the consideration of additional information, such as alpha, depth, display, colour format, colour processing, and/or other information that can be associated with such pixel data.

**3.3.20**

**Alpha picture**

A picture that contains a video signal representing information indicating the opacity level of each pixel in an associated foreground picture and how that associated picture should be blended with other pictures for purposes of alpha blending.

# Clause 4, Abbreviated terms

Add the following abbreviated terms:

|  |  |
| --- | --- |
| MSE | media source extension |
| MV | multi-view |
| NAL | network abstraction layer |

# Clause 7, CMAF track format

## Clause 7.1, Overview

Replace the following text:

The CMAF track format is derived from the ISO base media file format in this clause and structural brands are specified. At this point, the 'cmfc' and the 'cmf2' CMAF structural brands are defined. The 'cmf2' brand further restricts the 'cmfc' brand.

Several CMAF media objects are derived from the CMAF track format.

with:

The CMAF track format is derived from the ISO base media file format in this clause and structural brands are specified. At this point, the 'cmfc', the 'cmf1', and the 'cmf2' CMAF structural brands are defined. The 'cmf2' brand further restricts the 'cmfc' brand. The 'cmf1' brand both extends and restricts the 'cmfc' brand.

Several CMAF media objects are derived from the CMAF track format.

## Clause 7.2, CMAF brands

Replace Table 1 with:

**Table 1 — CMAF brands**

|  |  |  |
| --- | --- | --- |
| **Brand** | **Location** | **Conformance requirements** |
| cmfc | FileTypeBox and SegmentTypeBox | 7.6 |
| cmf2 | FileTypeBox and SegmentTypeBox | 7.7 |
| cmf1 | FileTypeBox and SegmentTypeBox | 7.8 |
| cmfs | SegmentTypeBox | 7.3.3.1 |
| cmfl | SegmentTypeBox | 7.3.3.2 |
| cmff | SegmentTypeBox | 7.3.2.3 |
| cmfr | SegmentTypeBox | 7.3.2.4 |

Replace Table 2 with:

Table 2 — Common media application format file extensions

|  |  |  |
| --- | --- | --- |
| **Track type** | **File extension** | **Internet media type (MIME type)** |
| Video | .cmfv | video/mp4 |
| Audio | .cmfa | audio/mp4 |
| Text (subtitle) | .cmft | application/mp4 |
| Metadata | .cmfm | application/mp4 |

## Clause 7.3, CMAF media objects

### Clause 7.3.1, CMAF boxes

Update the following entry of Table 3:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | trak |  |  |  |  | 1 | ISO/IEC 14496-12 |  | Container for each track |

with:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | trak |  |  |  |  | + | ISO/IEC 14496-12 | CMAF 7.5.21 | Container for each track |

Update the following entry of Table 3:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | trex |  |  |  | 1 | ISO/IEC 14496-12 | CMAF 7.5.14 | Track extends box |

with:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | trex |  |  |  | + | ISO/IEC 14496-12 | CMAF 7.5.14 | Track extends box |

Update the following entry of Table 5:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | traf |  |  |  |  | 1 | ISO/IEC 14496-12 |  | Track fragment |

with:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | traf |  |  |  |  | + | ISO/IEC 14496-12 | CMAF 7.5.21 | Track fragment |

### Clause 7.3.2.1, CMAF header

Replace the bullet point c.4:

1. The MovieBox shall contain exactly one track containing media data as specified in subclause 7.3.2.2.

NOTE Timed metadata tracks can be provided as separate CMAF tracks in a separate selection set.

with:

1. The MovieBox shall contain exactly one Primary Media Track containing media data as specified in subclause 7.3.2.2.

NOTE Timed metadata tracks can be provided as separate CMAF Timed Metadata Tracks in a separate selection set. Alternatively, Supplementary Metadata Tracks may be included in the same CMAF selection set.

### Clause 7.3.2.2, CMAF track

Replace the bullet point d:

d) Each CMAF track contains a single ISO BMFF track and TrackBox, as determined by CMAF header constraints specified in subclause 7.3.2.1.

with:

d) Each CMAF track contains a single Primary Media Track and TrackBox for the media data, as determined by the CMAF header constraints specified in subclause 7.3.2.1. A CMAF track can contain additional TrackBoxes if it carries Supplementary Metadata Tracks.

### Clause 7.3.2.3, CMAF chunk

Replace the following bullet points:

b) The MovieFragmentBox shall conform to the constraints of the structural CMAF brand constraints specified in Table 5, such as containing only one TrackFragmentBox that contains only one TrackRunBox.

NOTE 1 Since there is only one TrackRunBox per MovieFragmentBox, all media samples of a CMAF chunk are located in a single track run in one MediaDataBox.

c) The MediaDataBox shall contain all media samples referenced by the TrackRunBox and should immediately follow the TrackRunBox in byte order.

f) A CMAF chunk shall contain a TrackFragmentDecodeTimeBox containing the baseMediaDecodeTime of the first media sample.

with:

b) The MovieFragmentBox shall conform to the constraints of the structural CMAF brand constraints specified in Table 5 and clause 7.5.21. The MovieFragmentBox shall contain only one TrackFragmentBox of the Primary Media Track and may contain TrackFragmentBoxes of Supplementary Metadata Tracks.

c) The MediaDataBox shall contain all media samples referenced by the TrackRunBox. The MediaDataBox should follow the TrackRunBox in byte order unless another TrackFragmentBox follows within the same MovieFragmentBox. If multiple TrackFragmentBoxes are present within a MovieFragmentBox, the MediaDataBox shall follow the last TrackFragmentBox in that MovieFragmentBox, ensuring all samples referenced by the TrackRunBoxes are contained in the corresponding MediaDataBox.

f) A CMAF chunk shall contain a TrackFragmentBaseMediaDecodeTimeBox containing the baseMediaDecodeTime of the first media sample.

### Clause 7.3.3.3, CMAF track file

Replace the bullet points d) and e)

d) A video CMAF track may use an offset edit list as specified in subclauses 7.5.13 and 10.2.6.

e) A video CMAF track may use v1 TrackRunBoxes using negative composition offsets to adjust the composition time of the earliest presented video media sample in each CMAF fragment to its baseMediaDecodeTime and the earliest video media sample in the CMAF track file to zero, without using an offset edit list.

with

d) A CMAF track may use an offset edit list as specified in subclauses 7.5.13, 10.2.6 and 9.2.7.

e) A video CMAF track may use version 1 TrackRunBoxes using negative composition offsets to adjust the composition time of the earliest presented video media sample in each CMAF fragment to its baseMediaDecodeTime and the earliest video media sample in the CMAF track file to zero, without using an offset edit list.

## Clause 7.5, Constraints on ISO Base Media File Format boxes

### Clause 7.5.14, Track Extends Box ('trex')

Replace the paragraph:

A TrackExtendsBox shall be present in a CMAF track since it is a fragmented file as defined in ISO/IEC 14496-12.

with:

At least one TrackExtendsBox shall be present in a CMAF track since it is a fragmented file as defined in ISO/IEC 14496-12.

### Clause 7.5.21

Add a new Clause 7.5.21

**7.5.21 Constraints on the number of tracks in a single CMAF track**

A single Primary Media Track shall be present in a CMAF Track.

Additional Supplementary Metadata Tracks may be present, and if present, shall use the track reference box with at least one track reference type box that describes the association of the Supplementary Metadata Track to the Primary Media Track. If a specific track reference type is not defined to link the metadata track to the Primary Media Track it describes, the 'cdsc' track reference type should be used.

## Clause 7.6, The structural CMAF Brand 'cmfc'

Replace:

A CMAF track conforming to the CMAF structural brand 'cmfc' shall conform to the CMAF track constraints defined in subclauses 7.1, 7.2, 7.3, 7.4, and 7.5..

with:

A CMAF track conforming to the CMAF structural brand 'cmfc' shall conform to the CMAF track constraints defined in subclauses 7.1, 7.2, 7.3, 7.4, and 7.5. In addition, no Supplementary Metadata Tracks shall be present.

## Clause 7.7, The structural CMAF Brand 'cmf2'

### Clause 7.7.2, Edit List Box ('elst')

Replace:

For video CMAF Tracks, the EditBox and in particular the EditListBox shall not be present.

For video CMAF Track files as well as any other media types, the EditListBox may be present following the constraints in subclause 7.5.13.

with:

For video CMAF Tracks that are not video CMAF track files, the EditBox and, in particular, the EditListBox shall not be present.

For other CMAF tracks and for video CMAF Track files as well as any other media types, the EditListBox may be present following the constraints in subclause 7.5.13.

## New Clause 7.8

Add a new Clause 7.8

**7.8 The Structural CMAF Brand 'cmf1'**

**7.8.1 General**

A CMAF track conforming to the CMAF structural brand 'cmf1' shall conform to constraints defined in subclauses 7.1, 7.2, 7.3, 7.4, and 7.5 and all remaining constraints and exceptions in this clause 7.8. These constraints are introduced to ensure that the CMAF tracks and CMAF switching set track headers are conforming, including the case where all CMAF Tracks of the presentation conforming to this brand would be included in a single ISO BMFF file, as well as use cases where supplementary metadata is carried alongside primary media data within the same CMAF Track.

**7.8.2 Movie Box (**'moov'**)**

The MovieBox shall contain exactly one TrackBox of the Primary Media Track and may contain one or more additional TrackBoxes of Supplementary Metadata Track(s) with closely associated metadata.

**7.8.3 Track Header Box (**'tkhd'**)**

CMAF TrackHeaderBoxes shall conform to subclause 7.5.4 with the following additional constraints.

— The MovieFragmentBox shall contain exactly one TrackFragmentBox of the Primary Media Track and may contain one or more additional TrackFragmentBox of Supplementary Metadata Track(s) with closely associated metadata.

— If all CMAF tracks of the CMAF presentation are combined into a single ISO BMFF container, then the track header shall be valid. This for example included that:

— The track\_ID is set to a unique identifier over within this CMAF presentation.

— All tracks within one CMAF switching set have alternate\_group shall be set to the same value. Each CMAF switching set in the CMAF presentation shall have a unique value for the alternate\_group.

— For a video track, every decoder output signal shall have its decoded and cropped image size in video spatial samples measured on a uniformly sampled square grid identical to the value of width and height defined in the Track Header.

**7.8.4 Track Reference Box (**'tref'**)**

The TrackReferenceBox shall only be present in Supplementary Metadata Tracks and shall contain a single TrackReferenceTypeBox specifying its relation to the corresponding Primary Media Track.

**7.8.5 Media Header Box (**'mdhd'**)**

CMAF MediaHeaderBoxes shall conform to subclause 7.5.5 with the following additional constraints.

— If one or more Supplementary Metadata Track(s) are present, the value of the timescale should be the same as set in the Primary Media Track. Differences in timescale values may require additional processing to ensure proper synchronization of metadata with media samples.

**7.8.6 Track fragment decode time box (**'tfdt'**)**

The value of the baseMediaDecodeTime in the TrackFragmentBaseMediaDecodeTimeBox of a Supplementary Metadata Track should be set to the same value as the baseMediaDecodeTime in the TrackFragmentBaseMediaDecodeTimeBox of the corresponding Primary Media Track.

**7.8.7 Duration Consistency**

The duration of a Supplementary Metadata Track should match the duration of the corresponding Primary Media Track within a CMAF Track. The sum of sample durations in the TrackRunBox elements of a Supplementary Metadata Track should be equal to the sum of sample durations in the TrackRunBox elements of the corresponding Primary Media Track.

**7.8.8 Metadata Track Switching and Presence in CMAF Fragments**

Supplementary Metadata Tracks, when present, shall also be present in each CMAF Fragment of a CMAF Track. The boundaries of Supplementary Metadata Tracks shall align with the boundaries of the Primary Media Track within a CMAF Track to allow seamless switching at CMAF Fragment boundaries. This ensures consistent availability of metadata across all fragments and enables adaptive switching without playback interruptions.

# Clause 9, Video CMAF tracks

## Clause 9.2.3, Track Header Box ('tkhd')

Replace the following bullet point in clause 9.2.3

— flags = 0x000007

with:

— flags = 0x000007. This value implies that track\_enabled, track\_in\_movie, and track\_in\_preview are set to 1; the track\_in\_preview flag was deprecated in ISOBMFF.

## Clause 9.2.4, Sample Description Box ('stsd')

Add the following bullet point at the end of the clause 9.2.4 (before the note):

— one or more ColourInformationBoxes with sub-type 'nclx' as specified in ISO/IEC 14496-12.

## Clause 9.3.2, CMAF track format constraints for NAL structured video

Remove the following bullet point from clause 9.3.2.1:

— The CleanApertureBox should not be present, since the cropped aperture is defined to be active image (“clean”) in a video CMAF track.

In clause 9.3.2.2 replace:

NOTE 1 The size of the NAL header length field defined in video tracks conforming to ISO/IEC 14496-15 is stored in the field LengthSizeMinusOne in the corresponding decoder configuration record, e.g. for AVC video in the AVCDecoderConfigurationRecord.

— shall contain one or more ColorInformationBoxes with sub-type 'nclx' and a PixelAspectRatioBox 'pasp', as documented in ISO/IEC 14496-12, if the first sample entry contains no SPS NAL with VUI in the decoder configuration record.

NOTE 2 A decoder configuration record without a parameter set is valid for a sample description such as 'avc3' or 'hev1' that can store the parameter set NALs necessary for decoding and display in each CMAF fragment.

with:

NOTE 1 The size of the NAL header length field defined in video tracks conforming to ISO/IEC 14496-15 is stored in the field LengthSizeMinusOne in the corresponding decoder configuration record, e.g. for AVC video in the AVCDecoderConfigurationRecord.

NOTE 2 A decoder configuration record without a parameter set is valid for a sample description such as 'avc3' or 'hev1' that can store the parameter set NALs necessary for decoding and display in each CMAF fragment.

— shall contain one or more ColourInformationBoxes with sub-type 'nclx' and a PixelAspectRatioBox 'pasp', as documented in ISO/IEC 14496-12, if the first sample entry contains no SPS NAL unit with a VUI in the decoder configuration record.

— should not contain a CleanApertureBox, since the cropped aperture is defined to be active image ("clean") in a video CMAF track.

## Clause 9.3.7, Single initialization CMAF switching set constraints for NAL structured video tracks and media profiles

Fix the spelling of the ColorInformationBox to ColourInformationBox in bullet point c):

2) can contain a sample entry without NAL units that shall include one or more ColourInformationBoxes with sub-type 'nclx' and a PixelAspectRatioBox, as specified in ISO/IEC 14496-12.

## Clause 9.6, Video Media Profiles

Add the following bullet points at the end of the clause:

* Annex O specifies packaging and codec constraints for some CMAF media profiles using the LCEVC video codec. Systems claiming conformance to CMAF using LCEVC shall conform to the provisions of Annex O.
* Annex P specifies packaging and codec constraints for some CMAF media profiles using the MV-HEVC video codec. Systems claiming conformance to CMAF using MV-HEVC shall conform to the provisions of Annex P.

# Clause 13, Timed metadata tracks

Change the title and the contents of Clause 13 from:

**13 Timed metadata tracks**

Timed metadata tracks carry information relating to the CMAF presentation, such as programme metadata or geo-location information.

CMAF timed metadata tracks shall conform to timed metatracks as defined in ISO/IEC 14496-12 and shall adhere to the CMAF track structure defined in Clause 7.

CMAF timed metadata tracks use the null media header ('nmhd') and MetaSampleEntry is used to declare the type of the metadata track. The URIMetaSampleEntry can be used to signal a specific timed metadata scheme by a URI in URIBox.

Each ISO BMFF sample in a timed metadata track encloses metadata relevant to the presentation interval of that sample as defined in ISO/IEC 14496-12.

Metadata applying to other intervals can be present if this is supported by a metadata scheme, however such metadata should also be carried in an ISO BMFF sample with an overlapping presentation time interval (if such sample is present). In other words, metadata should be repeated in all samples during which it applies; otherwise players can miss metadata in case of random access to the metadata track.

All samples are sync samples. In addition, no knowledge of a prior sample should be required to correctly interpret the metadata carried in an ISO BMFF sample.

Timed metadata track files can use the filename extension 'cmfm'.

to:

**13 CMAF timed metadata tracks**

CMAF timed metadata tracks carry information relating to the CMAF presentation, such as programme metadata or geo-location information. It is important to clarify that the CMAF Timed Metadata Tracks are distinct from Supplementary Metadata Tracks, as defined in clause 3.2.15 of this document. While Supplementary Metadata Tracks are carried within the same CMAF track as the Primary Media Track and maintain a one-to-one dependency with the media data, CMAF timed metadata tracks are standalone metadata tracks. These tracks contain a single TrackBox in a MovieBox and are intended to carry general timed metadata, typically independent of the media tracks.

CMAF timed metadata tracks shall conform to timed metadata tracks as defined in Clause 12.3 of ISO/IEC 14496-12 and shall adhere to the CMAF track structure defined in Clause 7 of this document. MetaDataSampleEntry is used to declare the type of metadata in the track. The URIMetaSampleEntry is one example of a metadata sample entry that can be used to signal a specific timed metadata scheme via a URI in URIBox.

Each ISO BMFF sample in a CMAF timed metadata track encloses metadata relevant to the presentation interval of that sample as defined in ISO/IEC 14496-12.

Metadata applying to other intervals can be present if this is supported by a metadata scheme, however such metadata should also be carried in an ISO BMFF sample with an overlapping presentation time interval (if such sample is present). In other words, metadata should be repeated in all samples during which it applies; otherwise, players can miss metadata in case of random access to the metadata track.

All samples are sync samples. In addition, no knowledge of a prior sample should be required to correctly interpret the metadata carried in an ISO BMFF sample.

CMAF timed metadata track files can use the filename extension 'cmfm'.

# Annex B, HEVC video CMAF track format and CMAF media profiles

## Annex B.1, HEVC video CMAF tracks

Replace the text:

This annex defines HEVC video tracks and specific CMAF media profiles with HEVC elementary stream constraint sets. Applications that do not conform to the HEVC video track or any of these CMAF media profiles can either specify their own HEVC video track definition or CMAF media profile or both. Applications can also signal brand conformance to just a CMAF structural brand defined in this document (e.g. 'cmfc' or 'cmf2').

HEVC tracks shall conform to subclause 9.3, as additionally constrained in this annex.

with:

This annex defines HEVC video tracks and specific CMAF media profiles with HEVC elementary stream constraint sets. Applications that do not conform to the HEVC video track or any of these CMAF media profiles can either specify their own HEVC video track definition or CMAF media profile, or both. Applications can also signal brand conformance to just a CMAF structural brand defined in this document (e.g. 'cmfc', 'cmf1', or 'cmf2').

HEVC tracks shall conform to subclause 9.3, as additionally constrained in this annex.

# Annex H, Scalable HEVC media profile and track format

## Annex H.1, Dependent CMAF tracks

Replace the first paragraph:

Dependent CMAF tracks can only be processed with the availability of one or more other CMAF tracks. Specifically, CMAF fragments of dependent CMAF track can only be decoded and/or rendered if the corresponding CMAF fragment of all of the CMAF tracks it depends on are available. CMAF tracks that are not dependent CMAF tracks are regular CMAF tracks.

with:

Dependent CMAF tracks can only be processed with the availability of one or more other CMAF tracks. Specifically, CMAF fragments of a dependent CMAF track can only be decoded and/or rendered if the corresponding CMAF fragment of all CMAF tracks it depends on are available. CMAF tracks that are not dependent CMAF tracks are regular CMAF tracks.

## Annex H.3.2, Sample Description Box ('stsd')

Replace the second bullet point:

* For a visual sample entry with codingname 'hev1', 'hev2', and 'lhe1' shall contain one or more decoding parameter sets (containing VPS, SPS, and PPS NALs for HEVC video). Each video sample in the CMAF track shall reference a parameter set in the sample entry.

with:

* A visual sample entry with codingname 'hev1', 'hev2', and 'lhe1' shall contain one or more parameter sets (containing VPS, SPS, and PPS NAL units for HEVC video). Each video sample in the CMAF track shall reference a parameter set in the sample entry.

## Annex H.4.1.3, H.4.1.3 HEVCDecoderConfigurationRecord and LHEVCDecoderConfigurationRecord

Replace the text from:

The HEVCDecoderConfigurationRecord and the HEVC compatible base layer shall conform to subclause B.2.4.

The LHEVCDEcoderConfigurationRecord and the enhancement layers shall conform to subclause B.2.4. inclusion, use and passing of SEI messages.

with:

The HEVCDecoderConfigurationRecord, the HEVC compatible base layer, the LHEVCDEcoderConfigurationRecord, and the HEVC enhancement layers shall conform to subclause B.2.4.

## Annex H.4.2.2, General constraints

Remove the 4th bullet point on scalable resolution restriction in H.4.2.2:

— The bitstream shall contain at most two layers, a base layer and possibly an enhancement layer.

— The base layer shall conform to the HEVC Main 10 profile and main tier.

— The enhancement layer, when present, shall conform to the HEVC Scalable Main 10 profile and main tier.

— ~~The spatial resolution of the enhancement layer shall be equal to X times that of the base layer both horizontally and vertically. The value of X shall be 1.5, 2 or 3.~~

— Each layer shall contain at most two sub-layers, with TemporalId equal to 0 and 1 when there are two sub-layers, and the value of sps\_max\_sub\_layers\_minus1 of each SPS shall be set equal to 0 or 1. If there is only one sub-layer, the TemporalId shall be 0.

— The value of sub\_layer\_level\_present\_flag[ 0 ] shall be equal to 1. This constraint requires the signalling of the level of the sub-layer representation with TemporalId equal to 0.

## Annex H.4.2.5.2, VPS visual usability information (VPS VUI) fields

In clause H.4.2.5.2 replace:

— vps\_video\_format

with:

— video\_vps\_format

# Annex O, LCEVC media profile and track format

## Annex O.5.1, LCEVC signalling of "codecs" parameters

Replace:

Presentation applications should signal video codec profile and levels of each Base Layer and LCEVC Enhancement Layer track and switching set using parameters conforming to IETF RFC 6381 and ISO/IEC 14496-15:2022, Clause E.4.

with:

Presentation applications should signal the video coding specification profile and levels of each Base Layer and LCEVC Enhancement Layer track and switching set using parameters conforming to IETF RFC 6381 and ISO/IEC 14496-15:2022, Clause E.4. Table O.1 specifies the video codecs parameters.

clv1.vprf<prof>.vlev<level>

**Table O.1 — Video codecs parameters for the LCEVC media profiles**

|  |  |  |  |
| --- | --- | --- | --- |
| **codecs parameter** | Description | CMAF LCEVC Main Profile | CMAF LCEVC Main 4:4:4 Profile |
| <prof> | profile\_idc | 0 | 1 |
| <level> | level\_idc | <level>  Examples  For level 4.1: 41  For level 5.1: 51  For level 6.1: 61 | |

## Annex O.6, LCEVC media profile and track brands

Replace:

LCEVC media profiles and track brands shall conform to Table O.1.

with:

LCEVC media profiles and track brands shall conform to Table O.2.

Replace:

Table O.1 - LCEVC video media profiles

with:

Table O.2 - LCEVC video media profiles

# New Annex P on MV-HEVC

Add a new Annex P

**Annex P**(normative)

**CMAF Track and media profiles for MV-HEVC**

* 1. Introduction

This Annex defines media profiles for MV-HEVC and specifies constraints on CMAF Tracks and CMAF Switching Sets. It introduces three media profiles: the *MV-HEVC Stereo Media Profile* (Clause P.2), the *HEVC* *with Alpha* *Media Profile* (Clause P.3), and the *MV-HEVC Stereo with Alpha Media Profile* (Clause P.4). An overview of these profiles is provided in Table P.1. The *MV-HEVC Stereo* Media Profile can be used for the transport of stereoscopic video. The *HEVC with Alpha* Media Profile targets monoscopic video with a single alpha channel. The *MV-HEVC Stereo with Alpha* Media Profile supports stereo video with an auxiliary alpha channel.

Applications that do not conform to the MV-HEVC video track or any of these CMAF media profiles can either specify their own MV-HEVC video track definition or CMAF media profile, or both. Applications can also signal brand conformance to just a CMAF structural brand defined in this document.

Table P.1: Overview of MV-HEVC profiles in CMAF

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Profile Name and Brand** | **Number of Layers** | **Aux video allowed** | **Number of Texture Layers** | **layer\_id = 0** | | | **layer\_id > 0** | | |
| **Profile** | **BitDepth** | **Chroma format** | **Profile** | **BitDepth** | **Chroma format** |
| MV-HEVC Stereo  mvst | 2 | No | 2 | Main | 8 | 4:2:0 | Multiview Main  Multiview Extended | 8 | 4:2:0 |
| Main 10 | 8 or 10 | 4:2:0 | Multiview Main 10  Multiview Extended 10 | 8 or 10  (same as layer 0) | 4:2:0 |
| HEVC with Alpha  hval | 2 | Yes | 1 | Main | 8 | 4:2:0 | Multiview Main  Multiview Extended  Multiview Monochrome | 8 | 4:0:0  4:2:0 |
| Main 10 | 8 or 10 | 4:2:0 | Multiview Main 10  Multiview Extended  Multiview Monochrome  Multiview Monochrome 10 | 8 or 10 | 4:0:0  4:2:0 |
| MV-HEVC Stereo with Alpha  mvsa | 4 | Yes | 2 | Main | 8 | 4:2:0 | Multiview Main  Multiview Extended  Multiview Monochrome | 8 | 4:0:0  4:2:0 |
| Main 10 | 8 or 10 | 4:2:0 | Multiview Main 10  Multiview Extended  Multiview Monochrome  Multiview Monochrome 10 | 8 or 10  (texture layers shall match)  (alpha layers shall match) | 4:0:0  4:2:0 |

* 1. MV-HEVC Stereo Media Profile
     1. General

The Clause P.2 defines the MV-HEVC Stereo Media Profile for CMAF. If a CMAF track conforms to the requirements in this clause, it can use the brand 'mvst'.

A CMAF track compliant to this profile carries two video layers in the single Primary Media Track: one representing the left view and one representing the right view of a stereo pair. This clause specifies constraints on the CMAF track, the CMAF switching set, and the MV-HEVC elementary stream.

* + 1. CMAF track constraints

The Primary Media Track of a CMAF track conforming to the MV-HEVC Stereo Media Profile:

* shall conform to the single track layered HEVC elementary stream and sample definitions from Clause 9 of ISO/IEC 14496-15
* shall be a single video track
* shall contain exactly 2 layers in the single video track
* shall not contain extractors
* shall not contain aggregators
* can contain the Operating Points Information ('oinf') sample group
* can contain the Layer Information ('linf') sample group

NOTE Sample groups 'oinf' and 'linf' are not mandatory for a single L-HEVC track.

* shall not contain the external base layer sample group ('lbli')
* shall conform to the requirements for the Track Header Box ('tkhd') as defined in clause 9.2.3.
* shall conform to the requirements for the Sample Description Box ('stsd') as defined in clause 9.2.4 and shall use the same sample entry type for all sample entries. Constraints on the visual sample entry are defined in clause P.2.4.
* shall conform to the requirements of NAL structured video CMAF tracks as defined in subclause 9.3 with the following exceptions:
  + The sample entry shall contain a single VPS as well as one SPS and one PPS per layer. Since exactly two layers are allowed in this profile, the parameter sets should be ordered as follows: the VPS, followed by the SPS and PPS of layer 0, followed by the SPS and PPS of layer greater than 0.
  + The sample entry should contain one hvcC and one lhvC configuration boxes and the lengthSizeMinusOne field should be set to 3 for both configuration boxes.
  + The sample entry boxes such as 'colr', 'clli', 'mdcv', 'cclv', and 'amve' apply to all texture layers but not auxiliary layers. When equivalent metadata is carried in SEI messages, the container-level signalling takes precedence, and the metadata values should be identical.
  + 'clap' should not be present, but if present it applies to all layers.
    1. CMAF switching set constraints

For a CMAF switching set conforming to the MV-HEVC Stereo Media Profile the following applies:

* Each CMAF track in the CMAF switching set shall conform to CMAF track constraints as defined in clause P.2.2.
* The CMAF switching set shall conform to constraints for NAL structured video CMAF switching sets as specified in subclause 9.3.6 or 9.3.7.
* the additional constraints defined in clause P.2 for CMAF switching sets.
  + 1. Visual sample entry

The syntax and values of a visual sample entry shall conform to HEVCSampleEntry

or HEVCLHVCSampleEntry as defined in ISO/IEC 14496-15. The sample entry type shall be either 'hvc1' or 'hev1'.

The width and height fields specified in the VisualSampleEntry shall be set according to clause 4.5 of ISO/IEC 14496-15 using only the base layer information.

* + - 1. HEVC Configuration Box ('hvcC')

The HEVCDecoderConfigurationRecord and the HEVC compatible base layer shall conform to subclause B.2.4.

Furthermore, it shall include the three\_dimensional\_reference\_displays\_info SEI message as specified in ISO/IEC 23008-2.

* + - 1. Layered HEVC Configuration Box ('lhvC')

The LHEVCDecoderConfigurationRecord and the enhancement layers shall conform to subclause B.2.4.

* + 1. Constraints on MV-HEVC elementary streams

The following constraints apply to CMAF MV-HEVC elementary streams.

* Access units and media samples shall conform to subclause 9.3 and to the requirements of a media sample of the indicated description ('hvc1' or 'hev1') as specified in ISO/IEC 14496-15.
* The bitstream shall contain exactly two layers.
* The base layer with nuh\_layer\_id equal to 0 shall carry the texture picture.
* The first enhancement layer with nuh\_layer\_id greater than 0 shall carry the texture picture.
* The base layer with nuh\_layer\_id equal to 0 shall conform to HEVC Main or HEVC Main 10 profile.
* The enhancement layer with nuh\_layer\_id greater than 0 shall conform to one of these profiles:
  + Multiview Main or Multiview Extended when the base layer conforms to HEVC Main profile.
  + Multiview Main 10 or Multiview Extended 10 when the base layer conforms to HEVC Main 10 profile. Furthermore, the bitdepth of the enhancement layer shall match the bitdepth of the base layer.
* All pictures shall be encoded as coded frames and shall not be encoded as coded fields.

CMAF fragments containing access units identified by the 'hev1' sample description shall contain a VPS and SPS and PPS NAL units for both layers referenced from a coded video sequence in the first access unit of that sequence, immediately following its first access unit delimiter NAL unit, if an access unit delimiter NAL unit is present.

Access units identified by the 'hev1' sample description can retain filler data (in NAL units or SEI messages) and SEI messages that would change the hypothetical reference decoder bitstream conformance if removed.

Access units of type 'hvc1' shall reference a video parameter set in the sample entry of the CMAF header associated with the containing CMAF track.

**Video parameter sets (VPS)**

Each MV-HEVC video sample in the CMAF track shall reference the VPS in the CMAF header sample entry according to ISO/IEC 14496-15. The VPS shall not change within CMAF tracks or between CMAF tracks in a switching set. A CMAF MV-HEVC track shall conform to the multi-layer extensions and multiview high efficiency video coding specification requirements specified in Annex G of ISO/IEC 23008-2 with the following additional constraints:

* vps\_extension\_flag shall be set to 1
* vps\_extension( ) shall be present as specified in Annex F of ISO/IEC 23008-2.
* The value of vps\_max\_layers\_minus1 shall be set to 1.
* The value of vps\_num\_rep\_formats\_minus1 shall be set to 0.

The following fields shall have values set as follows for each profile\_tier\_level() structure in the VPS:

* general\_progressive\_source\_flag shall be set to 1.
* general\_frame\_only\_constraint\_flag shall be set to 1.
* general\_interlaced\_source\_flag shall be set to 0.
* general\_non\_packed\_constraint\_flag shall be set to 1.

The condition of the following fields for each profile\_tier\_level() structure in the VPS shall not change throughout the multiview HEVC elementary stream:

* general\_profile\_space
* general\_profile\_idc
* general\_tier\_flag
* general\_level\_idc

**Sequence parameter sets (SPS)**

Sequence parameter set NAL units that occur within a CMAF MV-HEVC track shall conform to the multi-layer extensions and multiview high efficiency video coding specified in ISO/IEC 23008-2 with the following additional constraints:

* general\_progressive\_source\_flag shall be set to 1.
* general\_frame\_only\_constraint\_flag shall be set to 1.
* general\_interlaced\_source\_flag shall be set to 0.
* general\_non\_packed\_constraint\_flag shall be set to 1.
* vui\_parameters\_present\_flag shall be set to 1.
* If present, chroma\_format\_idc shall be equal to 1.
* If present, update\_rep\_format\_flag shall be equal to 0.

The value for each of the following fields in the active SPS shall not change from one coded video sequence to another throughout a CMAF MV-HEVC track:

* general\_profile\_space
* general\_profile\_idc
* general\_tier\_flag
* general\_level\_idc
  + 1. Encryption

The encryption of a CMAF track conforming to the MV-HEVC mixed media profile shall be compliant with Clause 8.

* 1. HEVC with Alpha Media Profile
     1. General

The Clause P.3 defines the HEVC with Alpha Media Profile for CMAF, which targets the transport of monoscopic video carrying *texture pictures (3.3.19)* together with a single auxiliary video layer containing *alpha pictures (3.3.20)*. If a CMAF track conforms to the requirements in this clause, it can use the brand 'hval'.

A CMAF track compliant to this profile carries two video layers in the single Primary Media Track: one carrying the texture pictures and one carrying the associated alpha pictures. This clause specifies constraints on the CMAF track, the CMAF switching set, and the MV-HEVC elementary stream.

* + 1. CMAF track constraints

The Primary Media Track of a CMAF Track conforming to the HEVC with Alpha Media Profile shall comply with the constraints defined in Clause P.2.2.

* + 1. CMAF switching set constraints

The constraints from Clause P.2.3 apply.

* + 1. Visual sample entry

The constrains from Clause P.2.4 apply with additional constraints defined in this clause:

* The three\_dimensional\_reference\_displays\_info SEI message as specified in ISO/IEC 23008-2 should not be present.
* The alpha\_channel\_information SEI message shall be present and shall comply with the following constraints:
  + alpha\_channel\_use\_idc shall be set to 1 if alpha is premultiplied, or 0 if alpha is straight.
  + alpha\_channel\_cancel\_flag, alpha\_channel\_incr\_flag, alpha\_channel\_clip\_flag, and alpha\_transparent\_value should be set to 0.
* The depth field specified in the VisualSampleEntry shall be set to 0x0020 indicating that the track referencing this sample entry carries alpha pictures as auxiliary video.
  + 1. Constraints on MV-HEVC elementary streams

The constrains from Clause P.2.5 apply with additional constraints defined in this clause:

* The first enhancement layer with nuh\_layer\_id greater than 0 shall carry an auxiliary picture with AuxId equal to AUX\_ALPHA.
* The enhancement layer with nuh\_layer\_id greater than 0 shall conform to one of these profiles:
  + Multiview Main, Multiview Extended, or Multiview Monochrome when the base layer conforms to HEVC Main profile.
  + Multiview Main 10, Multiview Extended, Multiview Monochrome, or Multiview Monochrome 10 when the base layer conforms to HEVC Main 10 profile.

**Video parameter sets (VPS) additional constraints**

* The value of vps\_num\_rep\_formats\_minus1 can be set to either 0 or 1.

**Sequence parameter sets (SPS) additional constraints**

* If present, update\_rep\_format\_flag can be set to either 0 or 1.
  + 1. Encryption

The constraints from Clause P.2.6 apply.

* 1. MV-HEVC Stereo with Alpha Media Profile
     1. General

The Clause P.4 defines the MV-HEVC Stereo with Alpha Media Profile for CMAF that can be used for the transport of stereoscopic video carrying 2 *texture pictures (3.3.19)* together with 2 auxiliary video layers containing *alpha pictures (3.3.20)*. If a CMAF track conforms to the requirements in this clause, it can use the brand 'mvsa'.

A CMAF track compliant to this profile carries four video layers in the single Primary Media Track. This clause specifies constraints on the CMAF track, the CMAF switching set, and the MV-HEVC elementary stream.

* + 1. CMAF track constraints

The Primary Media Track of a CMAF Track conforming to the MV-HEVC Auxiliary Media Profile shall comply with the constraints from Clause P.2.2, with the following additional constraints:

* shall contain exactly 4 layers in the single video track
  + 1. CMAF switching set constraints

The constraints from Clause P.2.3 apply.

* + 1. Visual sample entry

The constrains from Clause P.2.4 apply with additional constraints defined in this clause.

* The alpha\_channel\_information SEI message shall be present and the signaled information shall be the same for both auxiliary video layers carrying alpha pictures:
  + alpha\_channel\_use\_idc shall be set to 1 if alpha is premultiplied, or 0 if alpha is straight.
  + alpha\_channel\_cancel\_flag, alpha\_channel\_incr\_flag, alpha\_channel\_clip\_flag, and alpha\_transparent\_value should be set to 0.
* The depth field specified in the VisualSampleEntry shall be set to 0x0020 indicating that the track referencing this sample entry carries alpha pictures as auxiliary video.
  + 1. Constraints on MV-HEVC elementary streams

The constrains from Clause P.2.5 apply with additional constraints defined in this clause.

* The bitstream shall contain exactly four layers.
* Two enhancement layers with nuh\_layer\_id greater than 0 shall carry auxiliary picture with AuxId equal to AUX\_ALPHA.
* In addition to the base layer, another enhancement layer with nuh\_layer\_id greater than 0 shall carry the texture picture. The profile of this enhancement layer shall match that of the base layer.
* Both enhancement layers with nuh\_layer\_id greater than 0 that carry auxiliary pictures shall conform to one of these profiles:
  + Multiview Main, Multiview Extended, or Multiview Monochrome when the base layer conforms to HEVC Main profile.
  + Multiview Main 10, Multiview Extended, Multiview Monochrome, or Multiview Monochrome 10 when the base layer conforms to HEVC Main 10 profile.

**Video parameter sets (VPS) additional constraints**

* The value of vps\_max\_layers\_minus1 shall be set equal to 3.
* The value of vps\_num\_rep\_formats\_minus1 can be set to either 0 or 1.

**Sequence parameter sets (SPS) additional constraints**

* If present, update\_rep\_format\_flag can be set to either 0 or 1.
  + 1. Encryption

The constraints from Clause P.2.6 apply.

* 1. Video codecs string parameter

Presentation applications should signal video codec profile and levels of each HEVC track and CMAF switching set using parameters conforming to IETF RFC 6381 and ISO/IEC 14496-15. In addition to the requirements above, Clause K.2.3 of ISO/IEC 14496-12 defines an extension to the codecs parameter syntax allowing for descriptive tags via the desc component. This extension should be used to signal the CMAF MV-HEVC profile employed in the presentation. In particular, the usecase tag should be used to distinguish MV-HEVC profiles. Examples include:

MV-HEVC Stereo Media Profile:

codecs="desc.usecase=vstereo+codec=hvc1.1.6.L93.B0"

HEVC with Alpha Media Profile:

codecs="desc.usecase=valpha+codec=hvc1.1.6.L93.B0"

MV-HEVC Stereo with Alpha Media Profile:

codecs="desc.usecase=valphastereo+codec=hvc1.1.6.L93.B0"