ISO/IEC JTC 1/SC 29/WG 03 N1349

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

**Document type:** Output Document

**Title:** CD of ISO/IEC 23000-19 3rd edition AMD 2 New Structural CMAF Brand Profile

**Status:** Approved

**Date of document:** 2024-11-08

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

**No. of pages:** 23 (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 N1349**

**November 2024, Sapporo, JP**

|  |  |
| --- | --- |
| **Title** | **CD of ISO/IEC 23000-19 3rd edition AMD 2 New Structural CMAF Brand Profile** |
| **Source** | **WG 03, MPEG Systems** |
| **Status** | **Approved** |
| **Serial Number** | **24412** |

**ISO 23000-19:2024(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: MV-HEVC, metadata tracks and other improvements

CD 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.

[Ed. Note (DP): Can we make a 4th edition with integrated AMD1 to 3rd edition and the text of this AMD? Then we would list the changes in the foreword:

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

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

]

[Ed. Note on switching group box in ISOBMFF: Contribution m67527 at the joint Exploration project of ISOBMFF/CMAF/DASH on that topic was accepted in the next Amendment of part 12. For CMAF the questions are:

* Are the semantics for this box corresponding to switching set semantics for CMAF?
* Should CMAF has an informative Annex/Example on how to use it?

We need to resolve this before the DIS stage of Amd. 2.]

# Clause 3 changes

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 can 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 primary 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 within the same CMAF track (3.2.1). It is designed to enhance or provide necessary information for the playback of the media content and shall use the 'cdsc' track reference type to link the metadata to the Primary Media Track (3.2.14). This metadata maintains a one-to-one dependency with the media data, ensuring synchronization and relevance. Examples include, but are not limited to, metadata for spatial media (e.g., stereoscopic video parallax data), or other timed information that directly impacts the media experience. Such metadata is not intended for late binding (3.3.3) or separate selection during playback.

[Ed. Note (DP): I feel like we also need a definition for the CMAF Timed Metadata Track in Clause 3.3 as specified in Clause 13]

# Clause 4 changes

Add new abbreviation for MV - multi-view

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

# Clause 7 changes

## Clause 7.1

Replace 7.1 overview

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 extends and restricts the 'cmfc' brand.

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

## Clause 7.2

Add the 'cmf1' entry to Table 1 — CMAF brands

Table 1 — CMAF brands

|  |  |  |
| --- | --- | --- |
| **Brand** | **Location** | **Conformance requirements** |
| cmfc | FileTypeBox and SegmentTypeBox | 7.6 |
| cmf1 | FileTypeBox and SegmentTypeBox | 7.8 |
| cmf2 | FileTypeBox and SegmentTypeBox | 7.7 |
| ... | ... | ... |

## Clause 7.3

### Clause 7.3.1

Update the following entry of Table 3 with

| NL 0 | NL 1 | NL 2 | NL 3 | NL 4 | NL 5 | **Format Req.** | **Specification** | **Constraints** | **Description** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | trak |  |  |  |  | **+** | I**SO/IEC 14496-12** | **CMAF 7.5.21** | **Container for each track** |

Update the following entry of Table 5 with

| NL 0 | NL 1 | NL 2 | NL 3 | NL 4 | NL 5 | **Format Req.** | **Specification** | **Constraints** | **Description** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | traf |  |  |  |  | **+** | ISO/IEC14496**-**12 | **CMAF 7.5.21** | **Track fragment** |

### Clause 7.3.2.1

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

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 CMAF header constraints specified in subclause 7.3.2.1. A CMAF track may contain additional TrackBoxes if it carries Supplementary Metadata Tracks.

### Clause 7.3.3.3

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 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.

## Clause 7.5

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 that carries the primary media shall always be present in a CMAF Track. Additional Supplementary Metadata Tracks are allowed only under the 'cmfm' brand. Otherwise, only a single Primary Media Track is allowed. Supplementary Metadata Tracks shall use the 'cdsc' track reference type to link the metadata track to the Primary Media Track it describes. If multiple Supplementary Metadata Tracks are present, all of them shall have a 'cdsc' track reference type to the same Primary Media Track.

[Editor’s note: Comments are welcome for the presence or not of the Track Reference Box in the CMAF track. And if present, what values in the Track\_IDs array should be used. This note applies generically to any track using a Track Reference box when they are all stored in a single file, for instance in the case of scalable HEVC.]

## Clause 7.7

Replace Clause 7.7.2

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.

[Ed. Note (DP) I move this statement from the discussion at MPEG#147 to this note.

It is not clear what is the benefit of permitting the use of edit list only to non-video tracks in ‘cmf2’. Would it be clearer for implementers to say: 'cmfc' allows edit lists, while 'cmf2' does not. Would it match implementations? It seems that more and more deployments use negative offsets for video and avoid audio priming? Feedback from implementers is needed, in order to decide if we should implement this in CMAF.]

## 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 of the CMAF structural brand 'cmfc' and all remaining constraints and exceptions in this clause 7.8.

These constraints introduced to signal that the CMAF tracks and CMAF switching set track headers are conforming as if all CMAF Tracks of the presentation also conforming to this brand would be included in a single ISO BMFF file.

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

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

— 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 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.

## New Clause 7.9

Add a new clause 7.9.

**7.9 The structural CMAF Brand 'cmfm'**

**7.9.1 General**

A CMAF track conforming to the CMAF structural brand 'cmfm' shall conform to constraints of the CMAF structural brand 'cmfc' and all remaining constraints in subclause 7.8.

**7.9.2 Movie Box (**'moov'**)**

Under the 'cmfm' brand, the MovieBox shall contain exactly one TrackBox of the Primary Media Track and one or more additional TrackBox of Supplementary Metadata Track(s) with closely associated metadata.

**7.9.3 Movie Fragment Box (**'moof'**)**

Under the 'cmfm' brand, the MovieFragmentBox shall contain exactly one TrackFragmentBox of the Primary Media Track and one or more additional TrackFragmentBox of Supplementary Metadata Track(s) with closely associated metadata.

# Clause 9 changes

In Clause 9.2.4 add the following bullet point at the end of the clause (before the note)

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

In Clause 9.3.2.2 fix the spelling of the ColorInformationBox to ColourInformationBox

— 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 with VUI in the decoder configuration record.

In Clause 9.3.7 fix the spelling of the ColorInformationBox to ColourInformationBox

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

[Ed. Note (DP): With the generalized requirement in clause 9.2.4 other shall requirements for the nclx colr box could be removed in derived clauses.]

In Clause 9.6 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 changes

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 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 metadata the 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 as defined in ISO/IEC 14496-12.

Metadata applying to other intervals may 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 may 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 may 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, which are carried within the same CMAF track alongside the primary media. Supplementary Metadata Tracks are designed to enhance or provide necessary information for the playback of the primary media, maintaining a one-to-one dependency with the media data. The CMAF Timed Metadata Tracks defined in this clause are standalone tracks with a single TrackBox in a MovieBox and are intended to carry general timed metadata not directly embedded with the Primary Media Track.

CMAF timed metadata tracks shall conform 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. The URIMetaSampleEntry can be used to signal a specific timed metadata scheme by a URI in URIBox.

[Ed. Note (DP): I changed this paragraph, please verify. It is already a requirement from ISOBMFF that nmhd is used, I refer to clause 12.3 in ISOBMFF for that. There is also no such type defined for MetaSampleEntry.]

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 may 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 may 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 may use the filename extension 'cmfm'.

# Annex B changes

Replace Clause B.1 with the following:

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 specifiy 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

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 changes

## Annex H.1 changes

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 of the CMAF tracks it depends on are available. CMAF tracks that are not dependent CMAF tracks are regular CMAF tracks.

## Annex H.3.2 changes

Replace the second bullet point of H.3.2:

* 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 NALs for HEVC video). Each video sample in the CMAF track shall reference a parameter set in the sample entry.

## Annex H.4.1.3 changes

Replace the text from clause H.4.1.3

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 enhancement layers shall conform to subclause B.2.4.

## Annex H.4.2.2 changes

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 HEVC Main 10 profile and main tier.

— The enhancement layer, when present, shall conform to 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.

[Ed. Note (DP): NB comments are welcome on this change as it removes an existing restriction to allow other important resolution ratios such as 1.0, 1.25, 4.0, etc.

(YK): How to address the backward compatibility issue caused by this change?

(DP): IIRC there is no such restriction on the elementary stream level and we are not aware of any deployments of S-HEVC to cause troubles in terms of backward compatibility.]

# Annex O changes

Replace Clause O.6

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

CMAF file with brand 'clv1' shall contain only LCEVC samples with sample entry 'lvc1'.

**Table O.1 — LCEVC video media profiles**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Media profile** | **Codec** | **Profile** | **Level** | **Max frame height** | **Max frame width** | **Max frame rate** | **CMAF file brand** |
| **LVC**  **Main** | LCEVC | Main | 4.1 | 4 320 | 7 680 | 120 | 'clv1' |

with

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

CMAF file with brand 'clv1' shall contain only LCEVC samples with sample entry 'lvc1'.

clv1.vprf<prof>.vlev<level>

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

|  |  |  |  |
| --- | --- | --- | --- |
| **codec 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 | |

[Ed. Note (DP): To me it looks like this should be another table and not replace the existing one from FDAM1, this new table seems to be for the codecs string and most likely belong to O.5 and not O.6. Please verify, this change. This looks like a bug to me.]

# 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 CMAF tracks and media profiles for MV-HEVC. It also defines the CMAF Switching Set constraints for MV-HEVC in CMAF. This Annex defines three media profiles for MV-HEVC, the *MV-HEVC Mixed Media Profile* in Clause P.2, the *MV-HEVC* *Auxiliary* *Media Profile* in Clause P.3, and the *MV-HEVC Stereo Media Profile* in Clause P.4. These profiles are specified in an order of increasing constraints. The overview of these profiles is depicted in Table P.1. The codecs parameter for CMAF Tracks conforming to MV-HEVC media profiles is provided in clause P.5.

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

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Media Profile Name and Brand** | **Max Layers** | **BitDepth** | **Profile layer\_id=0** | **Profile(s) layer\_id>0** | **Auxiliary video allowed** | **Max Number of Texture Layers** |
| MV-HEVC Stereo  'mvst' | 2 | 8 or 10 | Main or Main 10 | Multiview Main  Multiview Main Extended  Multiview Main Extended 10  Multiview Main 10 | No | 2 |
| MV-HEVC Auxiliary  'mvax' | 2 | 8 or 10 | Main or Main 10 | Multiview Main  Multiview Extended  Multiview Extended 10  Multiview Monochrome  Multiview Monochrome 10  Multiview Main 10 | Yes | 1 |
| MV-HEVC Mixed  'mvmx' | 3 or 4 | 8 or 10 | Main or Main 10  Multiview | Multiview Main  Multiview Extended  Multiview Extended 10  Multiview Monochrome  Multiview Monochrome 10  Multiview Main 10 | Yes | 2 |

[Ed. Note (DP): revisit the order of profiles being listed in the table to align with the profiles documented in the Annex]

[Ed. Note (DP) make sure to remove capitalized words]

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

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

A CMAF track compliant to this profile carries all video layers in the single Primary Media Track. The definition includes the definition of a CMAF track, CMAF switching set, and MV-HEVC elementary stream constraints.

This profile can be used for the transport of stereo video in association of auxiliary video or for a monoscopic video with multiple auxiliary videos. For a CMAF application with a monoscopic video and a single auxiliary video, the MV-HEVC Auxiliary Media Profile as defined in clause P.3 should be used. For a CMAF application that only intends to use a stereo video without auxiliary video, the MV-HEVC Stereo Media Profile as defined in Clause P.4 should be used.

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.

* + 1. CMAF track constraints

The Primary Media Track of a CMAF Track conforming to the MV-HEVC mixed 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 2 or more layers in the single video track
* shall not contain extractors
* shall not contain aggregators
* may contain the Operating Points Information ('oinf') sample group
* may contain the Layer Information ('linf') sample group

[Ed. Note (DP): Be aware that we have an amendment to part 15 where we clarify that sample groups 'oinf' and 'linf' are not mandatory for a single L-HEVC track. We should also mention these here because the readers of the old version of Part 15 could think that these are mandatory.]

* 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.

[Ed. Note (DP): we probably need to say more about the track header width/height as we now deal with multiple videos in layers.]

* 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
* shall conform to the additional constraints defined in clause P.2.
  + 1. CMAF switching set constraints

For a CMAF Switching Set conforming to the MV-HEVC mixed 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.

The depth field specified in the VisualSampleEntry shall be set to 0x0020 if samples of the track referencing this sample entry carry alpha auxiliary video in any of the layers.

[Ed. Note (DP) the FileFormat group probably needs to consider other values in this field. For example for signaling of depth or other auxiliary video types in the stream]

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

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

Question: I assume all the MDCV, CLLV, AMVE etc. SEIs are the same in both config records (for layer 0 and layer 1 for example) or can they differ? If yes, we need some additional bullet points here.

[Ed. Note (DP): the above was taken from Annex H. However, I believe we need to think a little bit more about colour signaling aspects here. Probably this signaling only applies to texture layers? SEIs could be layer dependent while in ISOBMF we don't have a layer dependent signaling for this. We need to clarify how to use colr, cllv and so on when multiple layers are present.]

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

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

[Ed. Note (DP): Similar concern as for hvcC. Furthermore, we can say something like: If alpha auxiliary picture is present in the track, the alpha channel information SEI message should be present.]

* + 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 three or four layers.
* The base layer with nuh\_layer\_id equal to 0 shall carry the primary picture.
* The first enhancement layer with nuh\_layer\_id greater than 0 shall carry the primary picture if stereoscopic video is carried by the CMAF track.
* The base layer with nuh\_layer\_id equal to 0 shall conform to HEVC Main or HEVC Main 10 profile. [Ed. Note (DP): probably need to say something about the tier too?]
* The enhancement layers with nuh\_layer\_id greater than 0 shall conform to one of these profiles:
  + Multiview Main
  + Multiview Main Extended
  + Multiview Main Extended 10
  + Multiview Mono
  + Multiview Mono 10
* 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 all SPS and PPS NALs referenced from a coded video sequence in the first access unit of that sequence, immediately following its first access unit delimiter NAL, if an access unit delimiter NAL is present.

Access units identified by the 'hev1' sample description may retain filler data (in NAL units or SEI messages) and SEI messages that would change 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. 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 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 of each VPS shall be set to 2 or 3.
* vps\_vui\_present\_flag shall be set to TBD.

The following fields shall have values set as follows for each profile\_tier\_level() structure in 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 VPS shall not change throughout the multiview HEVC elementary stream:

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

[Ed. Note (DP): VUI suff in case vps\_vui\_present\_flag = 1 TBD]

**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.

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

[Ed. Note (DP): verify these

VUI parameters that occur within a CMAF MV-HEVC track shall conform to ISO/IEC 23008-2 with the following additional constraints:

— The following fields shall have pre-determined values as follows.

— aspect\_ratio\_info\_present\_flag shall be set to 1.

— video\_full\_range\_flag shall be set to 0.

— The following fields have the following values.

— colour\_description\_present\_flag should be set to 1.

NOTE As defined in ISO/IEC 23008-2, if the colour\_description\_present\_flag is set to 1, the colour\_primaries, transfer\_characteristics and matrix\_coefficients fields are present in the VUI.

— If colour\_description\_present\_flag is set to 1, then colour\_primaries, transfer\_characteristics and matrix\_coefficients shall be set to one of the values permitted for the media profile (see Table B.1).

— If colour\_description\_present\_flag is set to 0, this shall indicate the following values are to be assumed:

— colour\_primaries = 1;

— transfer\_characteristics = 1;

— matrix\_coefficients = 1.

— overscan\_info\_present\_flag shall be set to 0, therefore overscan\_appropriate shall not be present,

— aspect\_ratio\_idc shall be set to 1.

— The values of the following fields shall not change throughout a CMAF track and CMAF switching set.

— low\_delay\_hrd\_flag

— colour\_description\_present\_flag

— colour\_primaries, when present

— transfer\_characteristics, when present

— matrix\_coeffs, when present

— The values of the following fields should not change throughout a CMAF track.

— vui\_time\_scale

— vui\_num\_units\_in\_tick

]

* + 1. Encryption

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

[Ed. Note (DP): Input contributions are welcome on this topic]

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

The Clause P.3 defines the MV-HEVC Auxiliary Media Profile for CMAF that can be used for the transport of monoscopic video in combination with a single auxiliary video such as depth or alpha. If a CMAF track conforms to the requirements in this clause, it may use the brand 'mvax'.

A CMAF track compliant to this profile carries two video layers in the single Primary Media Track. The definition includes the definition of a CMAF track, CMAF switching set, and MV-HEVC elementary stream constraints.

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.

* + 1. CMAF track constraints

The constraints from Clause P.2.2 apply.

* + 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.

[Ed. Note (DP): here we can say that only 2 SPSs PPSs are allowed. layer count shall be 2. layer with nuh\_layer\_id = 1 shall carry 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 two layers.
* The base layer with nuh\_layer\_id equal to 0 shall carry the primary picture.
* The first enhancement layer with nuh\_layer\_id greater than 0 shall carry an auxiliary picture.

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

* The value of vps\_max\_layers\_minus1 of each VPS shall be set equal to 1.
  + 1. Encryption

The constraints from Clause P.2.6 apply.

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

The Clause P.4 defines the MV-HEVC Stereo Media Profile for CMAF that can be used for the transport of stereoscopic video. If a CMAF track conforms to the requirements in this clause, it may use the brand 'mvst'.

A CMAF track compliant to this profile carries two video layers in the single Primary Media Track. The definition includes the definition of a CMAF track, CMAF switching set, and MV-HEVC elementary stream constraints.

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.

* + 1. CMAF track constraints

The constraints from Clause P.2.2 apply.

* + 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.

[Ed. Note (DP): here we can say that only 2 SPSs PPSs are allowed. layer count shall be 2. layer with nuh\_layer\_id = 1 shall carry another texture video. VisualSampleEntry.depth = 0x0018.]

* + 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 two layers.
* The base layer with nuh\_layer\_id equal to 0 shall carry the primary picture.
* The first enhancement layer with nuh\_layer\_id greater than 0 shall carry the primary picture.
* The first enhancement layer with nuh\_layer\_id greater than 0 shall conform to one of these profiles:
  + Multiview Main
  + Multiview Main Extended
  + Multiview Main Extended 10

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

* The value of vps\_max\_layers\_minus1 of each VPS shall be set equal to 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.

[Ed. Note (DP): This section depends on the outcome of the ISOBMFF work. Current work in the file format group in the ISOBMFF amendment working draft is useful for this section. This section should also clarify how lhevcptl MIME type parameter should be used in combination to what will be defined in ISOBMFF.]

# Definitions (add in the definitions section)

**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, color format, color processing, and/or other information that may be associated with such pixel data.

**Alpha picture**

A picture that contains a video signal representing information indicating either the opacity level of each pixel or the corresponding picture segment each pixel in the video belongs to.

**Depth picture**

A picture that contains a video signal representing depth information indicating the distance of objects from the camera at each pixel location.

[Ed. note: (DP) consider moving these definitions to ISOBMFF so that other specs can inherit them]

[Ed. note (DP) make sure to check that the defined terms are actually used in text]