INTERNATIONAL STANDARD

ISO/IEC 23000-19

First edition
2018-01

AMENDMENT 1
2018-10

Information technology — Multimedia application format (MPEG-A) —

Part 19: Common media application format (CMAF) for segmented media

AMENDMENT 1: SHVC media profile and additional audio media profiles

Technologies de l'information — Format pour application multimédia (MPEG-A) —

Partie 19: Format CMAF (Common Media Application Format) pour médias segmentés

AMENDEMENT 1: Profil média SHVC et profils médias audio supplémentaires

Reference number
ISO/IEC 23000-19:2018/Amd.1:2018(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

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 document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see 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 and IEC 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 (see 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 of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

A list of all parts in the ISO/IEC 23000 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Information technology — Multimedia application format (MPEG-A) —

Part 19: Common media application format (CMAF) for segmented media

AMENDMENT 1: SHVC media profile and additional audio media profiles

Clause 2

Add the following normative references:

ISO/IEC 23008-3:—[1)], *Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 3: 3D audio*

ISO/IEC 23091-3, *Information technology — Coding-independent code points — Part 3: Audio*

ISO/IEC 23003-4:2015, *MPEG audio technologies — Part 4: Dynamic range control*

Clause 3

Add the following terms and definitions to subclause 3.2:

3.2.10
audio programme
complete collection of all audio programme components and, if present, a set of accompanying presets

3.2.11
audio programme component
smallest addressable unit of an audio programme

Clause 5

Add the following new paragraphs at the end of Clause 5:

Annex H specifies the CMAF media profile for scalable HEVC (SHVC).

Annex I specifies the CMAF media profile for multichannel AAC.

Annex J specifies the CMAF media profile for MPEG-H audio.

1) Under preparation. Stage at the time of publication: ISO/FDIS 23008-3:2018.

Subclause 9.1

Before the last paragraph in subclause 9.1, add a new paragraph:

An additional CMAF media profile for scalable HEVC (SHVC) is defined in Annex H.

Subclause 10.1

Replace the first sentence with:

This clause specifies CMAF audio tracks derived from the CMAF track format, with additional constraints specific to CMAF audio tracks and audio CMAF media profiles.

Additional CMAF media profiles are defined in

— Annex I for multichannel AAC, and

— Annex J for MPEG-H 3D audio.

Annex G

Following Annex G, add new Annexes H, I and J:

Annex H
(normative)

Scalable HEVC media profile and track format

H.1 Dependent CMAF tracks

Dependent CMAF tracks can only be processed with the availability of one or more 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.

In a hypothetical CMAF receiver, the corresponding CMAF fragments of CMAF tracks depended on, are expected to be available prior to the CMAF fragments of the dependent CMAF track.

A manifest offering CMAF presentations with dependent CMAF tracks is expected to provide signalling of dependent and required CMAF tracks.

H.2 Scalable HEVC CMAF tracks

Scalable HEVC CMAF tracks shall conform to Clauses 7, 8, 9, and 12, and shall additionally conform to the constraints specified in this Annex.

Each scalable HEVC CMAF track that does not contain VCL NAL units with `nuh_layer_id` and `TemporalID` both equal to 0 is a dependent CMAF track and the constraints specified for dependent CMAF tracks in Clause H.1 shall apply. It is expected that the manifest provides signalling to express the dependency of a dependent CMAF track on another CMAF track, for example, using the `@dependencyID` in a DASH MPD.

H.3 CMAF switching set constraints for scalable HEVC CMAF tracks and media profiles

H.3.1 General

Subclause 9.2.3 applies, and additionally the following constraints apply:

— Each CMAF track with sample entry '`hev1`' shall conform to the HEVC media profile and track format as specified in Annex B.

— Each CMAF track containing a scalable HEVC bitstream or part thereof shall contain exactly one ISO BMFF track.

— When multiple CMAF tracks are present for carrying a scalable HEVC bitstream, the corresponding ISO BMFF tracks shall use distinct track IDs.

— CMAF switching sets containing a media profile listed in Clause H.6 with sample entry '`hev2`' and '`lhe1`' shall conform to single initialization CMAF switching set constraints.

Each coded video sequence in a scalable HEVC bitstream shall contain the necessary sequence parameter set and picture parameter set NAL units to signal decoding parameters changes allowed between CMAF tracks in the same switching set.

H.3.2 Sample Description Box ('stsd')

Subclauses 9.2.4 and 9.3.2.2 shall apply with the following additional restrictions:

A decoder configuration record:

— shall signal other sequence parameter set and picture parameter set fields used by the video track as specified in ISO/IEC 14496-15:2017, subclause 8.3.3.1,

— 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.

— may contain additional SEI NAL units to signal colour encoding and rendering, such as `mastering_display_colour_volume`, SEI `payloadType=137` (ISO/IEC 23008-2), or `alternative_transfer_characteristics`, SEI `payloadType =147` (ISO/IEC 23008-2).

H.3.3 Track Header Box ('tkhd')

The requirements of 7.5.4 apply.

NOTE Normalized `width` and `height` can be derived from a sequence parameter set NAL in each segment and coded video sequence for 'hev1', 'hev2', and 'lhe1' video samples. See subclause 9.2.3.3 for the storage and semantics of video sequence parameter sets.

H.3.4 Access units

Clause 9.2.6 applies.

Access units shall conform to the requirements of a sample of the indicated description ('hev1', 'hev2', or 'lhe1') as specified in ISO/IEC 14496-15.

CMAF fragments containing samples identified by the 'hev1', 'hev2', or 'lhe1' `type` shall contain all SPS and PPS NALs referenced by a coded video sequence in the first access unit of that sequence, immediately following its first access unit delimiter NAL (if any).

NOTE Access units of type 'hev1', 'hev2', and 'lhe1' can retain filler data (NAL units or SEI messages) and SEI messages that would change hypothetical reference decoder bitstream conformance if such conformance is necessary, such as the case where bitstreams are to be repackaged and conformance tested in MPEG-2 transport streams.

H.3.5 Decoding of adaptively switched scalable HEVC CMAF tracks

Subclause 6.6.6 applies to switching between single-layer bitstreams, wherein a “conceptual” track is generated by concatenating segments from different tracks among which the switching occurs. The same applies for each of the multiple tracks carrying a scalable bitstream, with the only difference being that, in this case multiple “conceptual” tracks are generated by concatenation of CMAF fragments. Once the multiple “conceptual” tracks are generated, the process specified in ISO/IEC 14496-15:2017, Clause 9 is applied to construct the bitstream to be decoded by the video decoder.

H.4 Sample and CMAF fragment constraints

H.4.1 Storage of scalable high efficiency video coding elementary streams

H.4.1.1 Conformance

Scalable high efficiency video coding (scalable HEVC) video tracks shall comply with ISO/IEC 14496-15:2017, Clause 9, with the following constraints:

— Each track shall carry only one layer or only a subset of one layer.

— Aggregators (as defined in ISO/IEC 14496-15:2017, Annex A) shall not be included in any track.

— The external base layer sample group shall not be included in any track.

The base layer (coded according to HEVC specification) shall be stored as described in subclause B.3.1.

H.4.1.2 Visual sample entry

The base track's syntax and values for a visual sample entry shall conform to sample entry 'hev1' as defined in ISO/IEC 14496-15.

— For each track that carries a subset of the base layer where the subset contains VCL NAL units with TemporalId greater than 0 only, the sample entry type shall be 'hev2'.

— For each track that carries a layer for which the VCL NAL units have nuh_layer_id greater than 0 or a subset of such a layer, the sample entry type shall be 'lhe1'.

H.4.1.3 HEVCDecoderConfigurationRecord and LHEVCDecoderConfigurationRecord

The HEVCDecoderConfigurationRecord and the HEVC compatible base layer shall conform to subclause B.3.1.3.

The LHEVCDEcoderConfigurationRecord and the enhancement layers shall conform to subclause B.3.1.3. inclusion, use and passing of SEI messages.

H.4.2 Constraints on scalable HEVC elementary streams

H.4.2.1 General

The following constraints apply to all CMAF scalable HEVC elementary streams. See Clause H.6 for media profile constraints on tier, profile, level, and frame rates.

H.4.2.2 General constraints

— 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.

H.4.2.3 Picture rate related constraints

— An enhancement layer shall only be present, if the base layer has constant picture rate.

— The enhancement layer, if present, shall have the same number of sub-layers as the base layer.

— When a layer has two sub-layers, the sub-layer representation with TemporalId equal to 0 shall have a constant picture rate (as indicated by the presence of elemental_duration_in_tc_minus1[0]), and the picture rate shall be exactly half of that of the entire layer's picture rate

(i.e., elemental_duration_in_tc_minus1[0] is equal to 2 * elemental_duration_in_tc_minus1[1]).

— The enhancement layer, when present, shall have the same picture rate as the base layer.

— The vps_vui_present_flag in each VPS shall be set equal to 1, pic_rate_present_vps_flag shall be set equal to 1, pic_rate_present_flag[i][j] shall be set equal to 1 and constant_pic_rate_idc[i][j] shall be set equal to 1 for all i, for all j.

— The vui_parameters_present_flag in each SPS shall be set equal to 1, vui_timing_info_present_flag in each SPS shall be set equal to 1, vui_hrd_parameters_present_flag in each SPS shall be set equal to 1, and fixed_pic_rate_general_flag[i] shall be set equal to 1 or fixed_pic_rate_within_cvs_flag[i] shall be set equal to 1 for all values of i in the range of 0 to maxNumSubLayersMinus1, inclusive.

— If vps_num_hrd_parameters is greater than 0, the following shall apply:

  — For each hrd_parameters() syntax structure in the VPS: fixed_pic_rate_general_flag[i] shall be set equal to 1 or fixed_pic_rate_within_cvs_flag[i] shall be set equal to 1 for all values of i in the range 0 to maxNumSubLayersMinus1, inclusive.

  — For each value of i in the range 0 to maxNumSubLayersMinus1, inclusive, the value of the syntax element elemental_duration_in_tc_minus1[i] in the hrd_parameters() syntax structure applicable to the enhancement layer shall be equal to the value of the syntax element elemental_duration_in_tc_minus1[i] in the hrd_parameters() syntax structure applicable to the base layer.

  NOTE This profile is restricted to equal picture rates on both layers. Implementations can, however, be capable of switching between different picture rates.

H.4.2.4 Picture type

All pictures shall be encoded as coded frames, and shall not be encoded as coded fields.

H.4.2.5 Video parameter sets (VPS)

H.4.2.5.1 VPS fields

Each scalable 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 scalable HEVC track shall conform to ISO/IEC 23008-2:2017, Annexes F and H with the following additional constraints:

— 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 0.

  — vps_extension_flag shall be set to 1.

  — vps_vui_present_flag shall be set to 1.

— The condition of the following fields for each profile_tier_level() structure in VPS shall not change throughout an scalable HEVC elementary stream:

  — general_profile_space

  - general_profile_idc
  - general_tier_flag
  - general_level_idc
- The value of vps_max_layers_minus1 of each VPS shall be set equal to 1.
- The value of sub_layer_level_present_flag[0] shall be equal to 1 only when the value of sub_layer_level_idc[0] is different than the value of general_level_idc.

H.4.2.5.2 VPS visual usability information (VPS VUI) fields

VPS VUI parameters that occur within a CMAF scalable HEVC track shall conform to ISO/IEC 23008-2:2017, Annexes F and H with the following additional constraints:

- pic_rate_present_vps_flag shall be set equal to 1.
- pic_rate_present_flag[i][j] shall be set equal to 1.
- constant_pic_rate_idc[i][j] shall be set equal to 1 for all i, for j equal to MaxSubLayersInLayerSetMinus1[i].

The values of the following fields in each video_signal_info() in VPS VUI shall not change throughout a CMAF track and switching set:

- vps_video_format
- video_full_range_vps_flag
- colour_primaries_vps
- transfer_characteristics_vps
- matrix_coeffs_vps

H.4.2.6 Sequence parameter sets (SPS)

H.4.2.6.1 SPS fields

Sequence parameter set NAL units that occur within a CMAF scalable HEVC track shall conform to ISO/IEC 23008-2:2017, Annexes F and H with the following additional constraints:

- The following fields shall have pre-determined values as follows:
  - 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 0.
  - vui_parameters_present_flag shall be set to 1.
  - vui_timing_info_present_flag shall be set to 1, vui_hrd_parameters_present_flag in shall be set to 1, and fixed_pic_rate_general_flag[i] shall be set equal to 1 or fixed_pic_rate_within_cvs_flag [maxNumSubLayersMinus1] shall be set equal to 1.