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

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

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

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

This document is a preview generated by ERS



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2020

All rights reserved. Unless otherwise specified, or required in the context of its implementation, 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
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

Contents

Page

Foreword	vii
Introduction	viii
1 Scope	1
2 Normative references	1
3 Terms and definitions	2
3.1 Media objects	3
3.2 Logical structure	3
3.3 Application model	4
4 Abbreviated terms	6
5 Document organization	8
6 CMAF hypothetical application model, media object model and profiles	10
6.1 Overview of the hypothetical application model and media object model	10
6.2 CMAF content processing model	11
6.3 Late binding CMAF track synchronization	12
6.4 Adaptive switching of CMAF tracks in CMAF switching sets	13
6.5 CMAF specified objects and profiles	14
6.5.1 Object derivation and interoperability code points	14
6.5.2 Encoded media objects	14
6.5.3 Logical media object sets	14
6.5.4 Addressable media objects	14
6.5.5 CMAF profiles, brand and identifiers	15
6.6 CMAF media object model	16
6.6.1 CMAF fragments	16
6.6.2 CMAF tracks	17
6.6.3 CMAF track files	17
6.6.4 CMAF segments	18
6.6.5 CMAF chunks	18
6.6.6 CMAF switching sets and adaptive switching	19
6.6.7 CMAF selection sets and late binding	22
6.6.8 CMAF presentation timing model	23
6.6.9 Manifest information	26
6.6.10 CMAF addressable media objects, resources, and resource identifiers	26
7 CMAF track format	27
7.1 Overview	27
7.2 CMAF brands	27
7.3 CMAF media objects	28
7.3.1 CMAF boxes	28
7.3.2 CMAF track media objects	30
7.3.3 CMAF addressable media objects	34
7.3.4 CMAF switching sets	36
7.3.5 CMAF selection sets	39
7.3.6 CMAF presentations	39
7.4 Additional boxes, not defined in the ISO Base Media File Format	40
7.4.1 Track Encryption Box ('tenc')	40
7.4.2 Sample Encryption Box ('senc')	40
7.4.3 Protection System Specific Header Box ('pssh')	40
7.4.4 Media profile specific boxes	40
7.4.5 Event Message Box ('emsg')	40
7.5 Constraints on ISO Base Media File Format boxes	41
7.5.1 Movie Header Box ('mvhd')	41
7.5.2 Metadata Boxes	41
7.5.3 Kind Box ('kind')	41

7.5.4	Track Header Box ('tkhd')	42
7.5.5	Media Header Box ('mdhd')	42
7.5.6	Video Media Header Box ('vmhd')	42
7.5.7	Sound Media Header Box ('smhd')	43
7.5.8	Subtitle Media Header Box ('sthd')	43
7.5.9	Data Reference Box ('dref')	43
7.5.10	Sample Description Box ('stsd')	43
7.5.11	Protection Scheme Information Box ('sinf')	43
7.5.12	Track contained media sample information boxes	43
7.5.13	Edit List Box ('elst')	44
7.5.14	Track Extends Box ('trex')	44
7.5.15	Movie Fragment Header Box ('mfhd')	44
7.5.16	Track Fragment Header Box ('tfhd')	44
7.5.17	Track Run Box ('trun')	45
7.5.18	Sample Group Description Box ('sgpd')	45
7.5.19	Media Data Box ('mdat')	45
7.5.20	Sub-sample Information Box ('subs')	46
7.6	The Structural CMAF Brand 'cmfc'	46
7.7	The structural CMAF Brand 'cmf2'	46
7.7.1	General	46
7.7.2	Edit List Box ('elst')	46
7.7.3	Track Run Box ('trun')	46
8	Common encryption of CMAF tracks	46
8.1	Multiple DRM system support	46
8.2	Track encryption	47
8.2.1	General requirements	47
8.2.2	CMAF track constraints	48
8.2.3	Encryption constraints	49
8.2.4	CMAF presentation encryption	50
9	Video CMAF tracks	50
9.1	Overview	50
9.2	General video CMAF track format	51
9.2.1	General video CMAF track structure and constraints	51
9.2.2	Video Media Header ('vmhd')	51
9.2.3	Track Header Box ('tkhd')	52
9.2.4	Sample Description Box ('stsd')	52
9.2.5	Video CMAF fragment presentation time	53
9.2.6	Video media sample dependencies	53
9.2.7	Video edit lists	53
9.2.8	General video CMAF fragment random access constraints	53
9.2.9	Additional random access pictures within CMAF video fragments	53
9.2.10	Image framing and encoding constraints	54
9.2.11	General video CMAF switching set constraints	54
9.3	NAL structured video CMAF tracks	55
9.3.1	Overview	55
9.3.2	CMAF track format constraints for NAL structured video	56
9.3.3	NAL structured video access units contained in media samples	57
9.3.4	NAL structured video coding sequences corresponding to CMAF fragments	57
9.3.5	Elementary stream constraints	58
9.3.6	General CMAF switching set constraints for NAL structured video	58
9.3.7	Single initialization CMAF switching set constraints for NAL structured video tracks and media profiles	58
9.4	AVC video CMAF tracks	59
9.4.1	Storage of AVC elementary streams	59
9.4.2	Constraints on AVC elementary streams	60
9.5	AVC video Internet Media Type parameters	61
9.5.1	AVC signalling of "codecs" parameters	61

10	Audio CMAF tracks	62
10.1	Overview	62
10.2	General audio CMAF track format	62
10.2.1	Derivation	62
10.2.2	Track Header Box ('tkhd')	62
10.2.3	Sound Media Header Box ('smhd')	63
10.2.4	Sample Description Box ('stsd')	63
10.2.5	AudioSampleEntry	63
10.2.6	Audio offset edit list	63
10.3	AAC audio CMAF tracks	63
10.3.1	Overview	63
10.3.2	"codecs" parameter signalling	63
10.3.3	Considerations for AAC audio encoding	64
10.3.4	AAC track constraints	65
10.3.5	AAC elementary stream constraints	66
10.4	AAC core audio CMAF media profile	67
10.5	AAC adaptive switching audio CMAF media profile	67
10.5.1	General constraints	67
10.5.2	CMAF fragment encoding constraints	68
10.5.3	General considerations and requirements	68
10.5.4	Constraints for AAC-LC	68
10.5.5	Constraints for HE-AAC	69
10.5.6	Constraints for HE-AACv2	70
11	Subtitles and captions	71
11.1	Overview	71
11.2	WebVTT	71
11.3	IMSC text and image tracks	72
11.3.1	General	72
11.3.2	Common constraints	72
11.3.3	IMSC1 text track constraints	73
11.3.4	IMSC1 image track constraints	73
11.4	CTA-608 and CTA-708	73
11.5	Metadata for subtitles	74
12	CMAF media profiles and CMAF presentation profiles	74
12.1	CMAF media profiles	74
12.1.1	General guidelines for specifying CMAF media profiles	74
12.1.2	Guidelines for audio CMAF media profiles	75
12.1.3	Guidelines for video CMAF media profiles	75
12.2	CMAF presentation profiles	76
12.2.1	General	76
12.2.2	CMAF profile conformance	76
	Annex A (normative) CMAF presentation profiles, media profiles and supplemental data	79
	Annex B (normative) HEVC video CMAF track format and CMAF media profiles	83
	Annex C (informative) Subsampling of NAL structured video tracks in CMAF switching sets	88
	Annex D (informative) Hypothetical player model	98
	Annex E (informative) Event messages	101
	Annex F (informative) Error handling for missing media	102
	Annex G (informative) Recommendations for AAC CMAF switching set encoding	103
	Annex H (normative) Scalable HEVC media profile and track format	106
	Annex I (normative) AAC multichannel CMAF media profiles and track format	112
	Annex J (normative) MPEG-H 3D audio track format and CMAF media profile	115
	Annex K (normative) MPEG-D USAC track format and CMAF media profile	120

Annex L (normative) IMSC 1.1 media profiles **122**

Bibliography **124**

This document is a preview generated by EVS

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.

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) or the IEC list of patent declarations received (see <http://patents.iec.ch>).

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 Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

This second edition cancels and replaces the first edition (ISO/IEC 23000-19:2018), which has been technically revised. It also incorporates the Amendments ISO/IEC 23000-19:2018/Amd.1:2018 and ISO/IEC 23000-19:2018/Amd.2:2019.

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

- addition of supplemental data brands;
- modification to the structural brand cmfc for compatibility with DASH segments;
- definition of a stricter brand 'cmf2' for legacy devices;
- refinements and updates to HEVC media profiles for SDR and HDR;
- definition of the scalable HEVC media profile;
- definition of AAC multichannel media profiles;
- definition of MPEG-H 3D audio track format and CMAF media profile;
- definition of MPEG-D USAC track format and CMAF media profile;
- definition of IMSC1.1 media profile.

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.

Introduction

Common media application format (CMAF) combines and constrains several MPEG specifications to define a multimedia format that is optimized for delivery of a single adaptive multimedia presentation to a variety of devices, using a variety of adaptive streaming, broadcast, download and storage methods.

Several MPEG specifications have been adopted for much of the video delivered over the internet and other IP networks (cellular, cable, broadcast, etc.). Various organizations have taken MPEG's core coding, file format and system standards and combined them into their own specifications for their specific application. While these specifications are similar, their differences result in unnecessary duplication of engineering effort and duplication of identical content in slightly different formats, which results in increased storage and delivery costs.

CMAF provides a common media specification that application specifications, such as MPEG dynamic adaptive streaming over HTTP (DASH), can reference and a common media format that allows a single encoded multimedia presentation to be used by many applications.

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

Part 19:

Common media application format (CMAF) for segmented media

1 Scope

This document specifies the CMAF multimedia format, which contains segmented media objects optimized for streaming delivery and decoding on end user devices in adaptive multimedia presentations.

CMAF specifies a track format derived from the ISO base media file format, then derives addressable media objects from CMAF tracks that can be used for storage and delivery.

CMAF specifies sets of tracks that share encoding and packaging constraints that enable the selection of multiple tracks to form a multimedia presentation and allow seamless switching of alternative encodings of the same content at different bit rates, frame rates, resolution, etc.

CMAF specifies a hypothetical application model that determines how tracks in a CMAF presentation are intended to be combined and synchronized to form a multimedia presentation. The model abstracts delivery to allow any delivery method. The hypothetical application model assumes a manifest and player, but CMAF does not specify a manifest, player, or delivery protocol, with the intent that any that support the hypothetical application model can be used.

CMAF specifies media profiles and brands that constrain media encoding and packaging of CMAF tracks to enable seamless adaptive switching of tracks and allow devices to identify compatible content by its brand.

CMAF specifies presentation profiles that conditionally require sets of CMAF tracks conforming to specified media profiles and allow content creators and devices to identify compatible multimedia presentations.

CMAF enables extensibility by specifying how new media profiles and presentation profiles can be specified and identified and includes guidelines for those specifications.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 14496-1, *Information technology — Coding of audio-visual objects — Part 1: Systems*

ISO/IEC 14496-3, *Information technology — Coding of audio-visual objects — Part 3: Audio*

ISO/IEC 14496-10, *Information technology — Coding of audio-visual objects — Part 10: Advanced video coding*

ISO/IEC 14496-12, *Information technology — Coding of audio-visual objects — Part 12: ISO base media file format*

ISO/IEC 14496-14, *Information technology — Coding of audio-visual objects — Part 14: MP4 file format*

ISO/IEC 14496-15, *Information technology — Coding of audio-visual objects — Part 15: Carriage of network abstraction layer (NAL) unit structured video in the ISO base media file format*

ISO/IEC 14496-30, *Information technology — Coding of audio-visual objects — Part 30: Timed text and other visual overlays in ISO base media file format*

ISO/IEC 23001-7, *Information technology — MPEG systems technologies — Part 7: Common encryption in ISO base media file format files*

ISO/IEC 23008-2, *Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 2: High efficiency video coding*

ISO/IEC 23009-1, *Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 1: Media presentation description and segment formats*

ISO/IEC 23008-3:2019, *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*

ISO/IEC 23003-3:2012, *Information technology — MPEG audio technologies — Part 3: Unified speech and audio coding*

IETF RFC 5234, *Augmented BNF for Syntax Specifications: ABNF*, <https://tools.ietf.org/html/rfc5234>

IETF RFC 6381:2011, *The 'Codecs' and 'Profiles' Parameters for "Bucket" Media Types*, <https://tools.ietf.org/html/rfc6381>

ITU-R Recommendation BT.709, *Parameter values for the HDTV standards for production and international programme exchange*

ITU-R Recommendation BT.1886, *Reference electro-optical transfer function for flat panel displays used in HDTV studio production*

ITU-R Recommendation BT.2035, *A reference viewing environment for evaluation of HDTV program material or completed programmes*

ITU-T Recommendation X.667:2014, *Information technology — Open Systems Interconnection — Procedures for the operation of OSI Registration Authorities: Generation and registration of Universally Unique Identifiers (UUIDs) and their use as ASN.1 object identifier components*, <https://www.itu.int/rec/T-REC-X.667>

ANSI/CTA-608-E R-2014, *Line 21 Data Services*, http://www.techstreet.com/standards/cta-608-e-r2014?product_id=1815447

ANSI/CTA-708-E, *Digital Television (DTV) Closed Captioning*, http://www.techstreet.com/standards/cta-708-e?product_id=1860354

W3C IMSC1, *TTML Profiles for Internet Media Subtitles and Captions 1.0*, <http://www.w3.org/TR/ttml-imsc1>

W3C IMSC1.1, *TTML Profiles for Internet Media Subtitles and Captions 1.1*, <http://www.w3.org/TR/ttml-imsc1.1>

W3C, *TTML Media Type Definition and Profile Registry*, *W3C Working Group Note*, <https://www.w3.org/TR/ttml-profile-registry>

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.