

ISO/IEC 23000-19:2020-03 (E)

Information technology - Multimedia application format (MPEG-A) - Part 19: Common media application format (CMAF) for segmented media

| 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 | CMAFmediaprofilesandCMAFpresentationprofiles | 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) | CMAFpresentationprofiles,mediaprofilesandsupplementaldata | 79 |
| Annex B (normative) | HEVCvideoCMAFtrackformatandCMAFmediaprofiles | 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-DUSAC track format and CMAF media profile | 120 |
| Annex L (normative) IMSC 1.1 media profiles | 122 |
| Bibliography | 124 |