International Standard

ISO/IEC 14496-32

Second edition
2025-04

Information technology — Coding of audio-visual objects —

Part 32: File format reference software and conformance

PROOF/ÉPREUVE

Reference number
ISO/IEC 14496-32:2025(en)

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

Published in Switzerland

Contents

Page

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 or www.iec.ch/members_experts/refdocs).

ISO and IEC draw attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO and IEC take no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO and IEC had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents and https://patents.iec.ch. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

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. In the IEC, see www.iec.ch/understanding-standards.

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 14496-32:2021), which has been technically revised.

The main changes are as follows:

— new conformance files have been added to the conformance suite with additional documentation in Clause 5;

— obsolete spreadsheets and html files have been removed.

A list of all parts in the ISO/IEC 14496 series can be found on the ISO and IEC websites.

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 and www.iec.ch/national-committees.

Information technology — Coding of audio-visual objects —

Part 32:
File format reference software and conformance

1 Scope

This document describes the reference software and conformance suite for the file format documents in multiple standards. Since these standards share a lot of technology, their reference software and conformance program are being handled together. These standards are: ISO/IEC 14496-12, ISO/IEC 14496-14, ISO/IEC 14496-15, ISO/IEC 14496-30 and ISO/IEC 23008-12.

The purpose of the conformance suite is to cover the set of valid features that can be exercised in the file format. Media conformance is not covered, though of course to exercise the file format features, media will be stored.

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 14496-12, *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 23008-12, *Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 12: Image File Format*

3 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 14496-12, ISO/IEC 14496-14, ISO/IEC 14496-15, ISO/IEC 14496-30 and ISO/IEC 23008-12 apply.

ISO and IEC maintain terminology databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at https://www.electropedia.org/

4 Reference software

The reference software for ISO/IEC 14496-12, ISO/IEC 14496-14, ISO/IEC 14496-15, ISO/IEC 14496-30, and ISO/IEC 23008-12 is available at: http://standards.iso.org/iso-iec/14496/-32/ed-2/en/reference_software/

This link contains version v.0.3.0 of the software with the changelog included in the software package, which documents new features compared to older versions.

5 Conformance

5.1 Overview

The conformance suite for ISO/IEC 14496-12, ISO/IEC 14496-14, ISO/IEC 14496-15, ISO/IEC 14496-30 and ISO/IEC 23008-12 is provided at: http://standards.iso.org/iso-iec/14496/-32/ed-2/en/conformance/

The suite of conformance tests does not currently cover deliberately errored files. However, such files do occur in practice and implementations should be written to be resilient.

There is no official reference tool provided to check the conformance of files. However, such tools do exist. The reference software can be used to open files in 'debug' mode and provide a listing of what it finds, and other trade associations and standards bodies might have validation tools tailored to their areas.

5.2 Standard specific considerations

5.2.1 Considerations for ISO/IEC 14496-30

Two files are proposed regarding:

— WebVTT;

— TTML.

There are several general features of ISO/IEC 14496-30 which are not exercised in the above files, specifically:

— a track with `'mul'` language;

— a track sharing resources (images, ...) between samples through the use of a `MetaBox`;

— file with a subtitle and a font track.

There are several features specific to the carriage of TTML in MP4 which are not exercised:

— timing:

  — 'empty' sample;

  — sample with a document containing a larger time range than the sample presentation range;

  — sample with a document whose time range is smaller than the sample range;

— sample entry format:

  — with multiple namespace values;

  — with schema location;

  — with auxiliary mime types;

— sample format:

  — with additional resources.

5.3 File documentation

5.3.1 Files from related conformance programs

5.3.1.1 Overview

Some files provided in this suite originate from external conformance programs, from other ISO documents or external specifications. The latest version should be obtained from their original source as indicated in the following sections.

5.3.1.2 Audio-based conformance files

ISO/IEC 14496-26 provides several MP4 files containing MPEG-4 AAC, MPEG-D MPS, and MPEG-D USAC streams and exercising features of the ISOBMFF. Within these files, only some are necessary to cover all the ISOBMFF features. These files are from the "AAC-Conformance" part. The features specifically exercised by these files compared to existing files are: `'padb'` and `'m4ae'`.

The following files are also included for ISOBMFF conformance:

— ./mpeg-audio-conformance/aac-conformance/ac01.mp4 (`'padb'`)

— ./mpeg-audio-conformance/aac-conformance/sls2100_aot02_048_16.mp4 (`'m4ae'`)

Original files can be obtained from ISO/IEC 14496-26.

5.3.1.3 Timed text conformance files

DECE produced a set of Timed Text test vector files with video and audio tracks. Some of these files exercise features of ISO/IEC 14496-12 and ISO/IEC 14496-30. The following files are added as part of ISOBMFF conformance:

— Solekai002_1280_23_1x1_v7clear.uvu

— Solekai007_1920_29_1x1_v7clear.uvu

Original files can be obtained from https://www.uvcentral.com/cff/cff-test-files.html.

5.3.1.4 DASH-based conformance files

The ISO/IEC 23009 series specifies the use of ISOBMFF files for adaptive streaming. ISO/IEC 23009-2 contains the following files that exercise specific features of the ISOBMFF:

— ./green/video_2500000bps_0.mp4

— ./green/meta_2500000bps_0.mp4m

— ./nalu/svc/mp4-onDemand-LastTime-depRep.mp4

— ./nalu/mvc/DDF_10s_25fps.mp4

Original files can be obtained from https://testassets.dashif.org/.

5.3.1.5 MAF-based conformance files

The ISO/IEC 23000 series uses the ISOBMFF standard at its core and contains some files (related to the PAAF and VSAF standards). None of these files exercise new features, but they exercise some features differently. The following files are copied in this ISOBMFF conformance:

— ./maf/vsaf/1.mp4

— ./maf/paaf/01_output_a.paf

Original files can be obtained from ISO/IEC 23000-6 and ISO/IEC 23000-10.

5.3.2 Files unique to this conformance program

5.3.2.1 a1-foreman-QCIF.mp4

This file is very simple. It has an MPEG-4 video part 2 visual track, and an AAC track, interleaved; and a minimal BIFS scene and OD track, with an IOD.

5.3.2.2 a2-foreman-QCIF-hinted.mp4

This file is basically the same as a1-foreman-QCIF.mp4 but hinted for RTP transmission.

5.3.2.3 a3-tone-protected.mp4

This file uses the protected stream structures. The keys are also supplied (a3c-prot-keys.txt), and the result of de-protection (a3b-tone-deprot.mp4, for comparison).

5.3.2.4 a4-tone-fragmented.mp4

This file uses movie fragments. The initial 1-second movie is followed by a 1-second movie fragment. Fragment-aware readers should play 2 second of content, fragment-unaware readers only 1 second.

5.3.2.5 a5-foreman-AVC.mp4

This is a very simple video-only main profile AVC file. Since it is main profile, composition offsets are used.

5.3.2.6 a6_tone_multifile.mp4

This is the same tone as used in test a4-tone-fragmented.mp4, but the actual access units are stored in a separate file, referenced by a relative URL "./myData.dat" from the main file.

5.3.2.7 a7-tone-oddities.mp4

This file also uses the 1-second tone. However, it has:

— a UUID,

— a non-standard `'junk'` atom in it (which should be ignored),

— a free space box (which should also be ignored),

— the compact sample size table,

— a padding bits table (though the padding bits are all set to zero).

The `'mdat'` atom has an implied length (the length in the file is zero, meaning to end of file).

5.3.2.8 a8-foreman_QCIF_edit.mp4

This file has the "foreman" 10 second of video, with 5 seconds before and after, of the "container" video. However, the edit list should select only foreman. The container ship should not appear. Note that the I-frames do not land on the edit boundaries. A player will have to pre-roll the video from an I-frame to work correctly.

5.3.2.9 a9-aac-samplegroups-edit.mp4

This file demonstrates the suggested way of handling AAC: it has a pre-roll sample group, and a track edit that is not aligned at either start or end with an AAC sample boundary.

5.3.2.10 a10-foreman_QCIF-raw.mp4

This file contains 'raw' (YUV420) video. Since this is an unregistered codec type (it is supported in QuickTime movie files) this is an unrecognized codec type from an ISOBMFF reader's point of view.

5.3.2.11 LargerThan4GB.mp4

This file tests handling of very large (> 4GB) files.

The `'mdat'` atom has a large (64-bit) size, and all the samples are at the end, preceded by 4 GB of zeroes. Therefore, the chunk offset table is also a `'co64'`, not a `'stco'`. The actual media data is a simple AAC tone.

5.3.2.12 f1.mp4

This file is a simple AVC + AAC file. It has an MPEG-4 AVC Baseline visual track (including the optional `BitrateBox`), and an AAC track.

5.3.2.13 f2.mp4

This file is a protected AVC + AAC file. The 128-bit key for the decryption process is `0x010203040506070801020304050607 08` for both tracks. The salt (counter offset) is `0x0000000000000001` for the audio and `0x0000000000000002`. Because of the usage of protected streams, `'isom'` was replaced with `'iso2'` in the list of compatible brands.

5.3.2.14 male_amr122.3gp

This file and the following 3GP files contain AMR speech at 12.2 or 6.7 kbps, with or without DTX (silence frames). In addition, 3GP files with hint tracks are provided.

AMR 12.2 kbps, no DTX

5.3.2.15 male_amr122DTX.3gp

AMR 12.2 kbps, DTX

5.3.2.16 female_amr67_hinted.3gp

AMR 6.7 kbps, no DTX, hint track

5.3.2.17 female_amr67DTX_hinted.3gp

AMR 6.7 kbps, DTX, hint track

5.3.2.18 pdin_example.3gp

This file contains one video track with AVC and a progressive download information box specifying required initial delays for six different download rates. The download rates 5106, 7659, 10213, 12766, 15319 and 20426 bytes per second require initial delays of 20808, 7206, 1089, 652, 396 and 200 seconds, respectively.

5.3.2.19 rs_example_r1.3gp

This file contains three video tracks with AVC at different bitrates, three audio tracks with HE-AACv2 at different bitrates, track selection box, and rate share information with two operation points. For the first operation point (100 kilobits per second), the target rate shares are 20 % for audio and 80 % for video. For the second operation point (160 kilobits per second), the target rate share weights are 20 and 140 for audio and video, respectively. As the sum is not 100 for the second case, the numbers correspond to weights that need to be normalized by the server/player. Depending on the available bitrate, the server selects which tracks to stream/play.

5.3.2.20 01_simple.mp4

Simple AV file (MPEG-4 ASP video, AAC audio), BIFS + OD scene, 2 timelines (BIFS/OD and A/V), interleaved.

5.3.2.21 02_dref_edts_img.mp4

Image track, audio track with edit list, with media data located outside the file.

5.3.2.22 03_hinted.mp4

Simple video file with MPEG-4 ASP visual, hinted for RTP (RFC 3640 payload).

5.3.2.23 04_bifs_video.mp4

Video (MPEG-4 ASP visual) + BIFS text (reading 'unprotected video'), with a single timeline.

5.3.2.24 05_bifs_video_protected_v2.mp4

Protected video (MPEG-4 ASP visual) + BIFS text (reading 'protected video'), with a single timeline. Keys are described in an item located in a `'meta'` box at the file root level, ISMA KMS URI referring to this item. Keys are:

— key `0x2b7e151628aed2a6abf7158809cf4f3c`

— salt `0xf8f9fafbfcfdfeff`

Only I-frames of the video track are encrypted.

5.3.2.25 06_bifs.mp4

Simple animation with a single BIFS track. File `'moov'` box is located after `'mdat'` box.

5.3.2.26 07_bifs_sprite.mp4

Simple looping animation with two BIFS tracks, exercising decoding dependency and synchronization track references.

Animation track uses `ShadowSync` samples in-between regular samples.

5.3.2.27 08_bifs_carousel_v2.mp4

Simple animation with a single BIFS track. Random access samples are inserted in-between the samples for the BIFS carrousel and signalled with a sample dependency type box.

5.3.2.28 09_text.mp4

Sample MPEG-4 streaming text file, stored in 3GPP text track format, with 2 sample descriptions.

5.3.2.29 10_fragments.mp4

Simple AV file (MPEG-4 ASP video, AAC audio), BIFS + OD scene, stored as a sequence of 500 ms fragments.

5.3.2.30 12_metas.mp4

File with a single image track, containing 3 `'meta'` boxes (file, movie and track level). Meta at movie level has an item referencing the entire file.