**ISO/IEC 14496-32:20xx(X)**

ISO/IEC JTC 1/SC 29/WG 3

Secretariat: JISC

**Information technology — Coding of audio-visual objects — Part 32 2nd edition: File format reference software and conformance**

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 2017, Published in Switzerland

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

Ch. de Blandonnet 8 • CP 401

CH-1214 Vernier, Geneva, Switzerland

Tel.  + 41 22 749 01 11

Fax  + 41 22 749 09 47

copyright@iso.org

www.iso.org

Contents

[Foreword iv](#_Toc517088186)

[Introduction v](#_Toc517088187)

[1 Scope 1](#_Toc517088188)

[2 Normative references 1](#_Toc517088189)

[3 Terms and definitions 1](#_Toc517088190)

[4 Reference Software 2](#_Toc517088191)

[5 Conformance 2](#_Toc517088192)

[5.1 Overview 2](#_Toc517088193)

[5.2 Standard specific considerations 2](#_Toc517088194)

[5.2.1 Considerations for ISO/IEC 14496-12, 14496-14, 14496-15 2](#_Toc517088195)

[5.2.2 Considerations for ISO/IEC 14496-30 2](#_Toc517088196)

[5.2.3 Considerations for ISO/IEC 23008-12 3](#_Toc517088197)

[5.3 File Documentation 4](#_Toc517088198)

[5.3.1 Files from related conformance programs 4](#_Toc517088199)

[5.3.2 Files unique to this conformance program 5](#_Toc517088200)

[Annex A (informative) Other tools 20](#_Toc517088201)

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 ISO documents 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](https://www.iso.org/directives-and-policies.html)).

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 (see [www.iso.org/patents](https://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](https://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*.

The second edition introduces new conformance files to the conformance program and documents them in clause 5.

A list of all parts in the ISO/IEC 14496 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](http://www.iso.org/members.html).

Information technology — Coding of audio-visual objects — Part 32 2nd edition: File format reference software and conformance

# Scope

This document describes the reference software and conformance suite for the file format documents in MPEG-4 and MPEG-H. Since these documents share a lot of technology, their reference software and conformance program are being handled together. These documents 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.

# 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*

# 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 terminological databases for use in standardization at the following addresses:

— IEC Electropedia: available at <http://www.electropedia.org/>

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

# Reference software

The reference software for ISO/IEC 14496-12, ISO/IEC 14496-14, ISO/IEC 14496-15, ISO/IEC 14496-30 andISO/IEC  23008-12 is provided at:

<http://standards.iso.org/iso-iec/14496/-32/ed-1/en/reference_software/>

Additional tools exercising features of the ISO base media file format (ISOBMFF) are described in Annex A.

# Conformance

## Overview

The conformance suite for ISO/IEC 14496-12, ISO/IEC 14496-14, ISO/IEC 14496-15, ISO/IEC 14496-30 andISO/IEC 23008-12 is provided at:

[http://standards.iso.org/iso-iec/14496/-32/ed-1/en/conformance/](http://standards.iso.org/iso-iec/14496/-32/ed-1/en/)

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 (see Annex A). 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 may have validation tools tailored to their areas.

An HTML report (conformance\_report.html) indicates which code points (e.g. box codes and track reference types) defined in these documents are used in which conformance file.

## Standard specific considerations

### (void)

For ISOBMFF conformance, an associated file (ff-conformance.xls) is used to list the features in groups, and to document which files exercise each feature. Features not currently covered by any file are marked with the sign “---“.

The file ff-conformance.xls has two spreadsheets. The first briefly lists the areas and features covered, and then has a column for each proposed file. The second sheet provides a brief description of each area and feature, by line.

[Editor's note: Editors will restructure this text to avoid empty sections.]

### 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.

### (void)

For HEIF conformance, an associated file (HEIF\_conformance.xlsx) is used to describe the current files, associated input bitstreams, and a specification coverage analysis of the conformance.

The spreadsheet "HEIF\_conformance.xlsx" contains information of the tested features and their mapping to the provided conformance files. It was created by picking all normative statements (shall, shall not, etc.) in the HEIF specification. These are collected in the "Areas tested" column of the sheet.

The column "Description" presents a brief description of the tested feature, exercising the associated normative statement in the specification.

The column "File ID" provides the file name.

The column "brands" indicates the HEIF brands that are included in FileTypeBox of the respective file (in the "Test Content" sheet) or concern the respective test case (in the "Test Cases" sheet. The aim of the column is to help implementers to pick files that concern the brand of their interest. For example, implementers of the 'heic' brand should only consider the files with that brand included.

Note that the spreadsheet enables filtering by content, e.g. by selecting only a particular File ID (at cell A1), one can see which features are exercised by a particular file.

[Editor's note: Editors will restructure this text to avoid empty sections.]

## File documentation

### Files from related conformance programs

#### 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.

#### 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.

#### 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 <http://uvvuwiki.com/cff/cff-test-files.html>.

#### 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 <http://testassets.dashif.org/>.

#### 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.

### Files unique to this conformance program

#### a1-foreman-QCIF.mp4

This file is about as simple as it gets. 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.

#### a2-foreman-QCIF-hinted.mp4

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

#### 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).

#### 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.

#### a5-foreman-AVC.mp4

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

#### 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.

#### 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 (also 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).

#### 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.

#### 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.

#### a10-foreman\_QCIF-raw.mp4

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

#### 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 4GB of zeroes. Therefore, the chunk offset table is also a co64, not an stco. The actual media data is a simple AAC tone.

#### 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.

#### f2.mp4

This file is a protected AVC + AAC file. The 128-bit key for the decryption process is 0x01020304050607080102030405060708 for both tracks. The salt (counter offset) is 0x0000000000000001 for the audio and 0x0000000000000002. Because of the usage of protected streams, “isom” was replace with “iso2” in the list of compatible brands.

#### 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 12.2kbps, no DTX

#### male\_amr122DTX.3gp

AMR 12.2kbps, DTX

#### female\_amr67\_hinted.3gp

AMR 6.7kbps, no DTX, hint track

#### female\_amr67DTX\_hinted.3gp

AMR 6.7kbps, DTX, hint track

#### 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.

#### 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.

#### 01-simple.mp4

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

#### 02-dref\_edts\_img.mp4

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

#### 03-hinted.mp4

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

#### 04-bifs\_video.mp4

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

#### 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 video I-frames are encrypted.

#### 06-bifs.mp4

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

#### 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.

#### 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.

#### 09-text.mp4

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

#### 10-fragments.mp4

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

#### 12-metas.mp4

File with a single image track, containing 3 metas (root, moov and track level). Meta at moov level has an item referencing the whole file.

#### 13-long.mp4

Long duration file, with MPEG-4 ASP track (only I-frames present). Total file duration is 5000000000 sec. (158 Years 81 Days 08:53:20).

#### 14\_large.mp4

Large file exercising 64 bits chunk offset.

#### 16\_vtt.mp4

The file 16\_vtt.mp4 contains the following boxes:

* wvtt
  + vttC
* vttc
  + ctim
  + iden
  + sttg
  + payl
* vtte
* vtta

#### 17\_negative\_ctso.mp4

This file contains a single AVC video track with a ctts box using version 1. The box declares negative composition offset such that the offset of the first frame is 0.

#### 18\_pssh\_v2.mp4

This file contains an encrypted video track. The file also contains 2 pssh boxes: one with only opaque data and the other one with KID information.

#### 19\_ttml.mp4

This file contains a simple TTML track.

#### 20\_stxt.mp4

This file contains a track of type 'stxt' whose content is SVG, the mime type is image/svg+xml. The SVG content is distributed over 4 samples. The configuration of the track contains the header of the SVG content.

#### 21\_segment.mp4

This file contains a self-initializing DASH media segment with the following boxes: 'sdix', 'prft', trun v1, sgpd.

#### 22\_tx3g.mp4

The file provides a simple 3GPP Timed Text stream stored in an mp4 file.

#### compact-no-code-fec-1.iso3

This file contains one JPEG file stored as an item. Compact No-Code FEC is used and the file is partitioned into one source block. The file contains also one hint track for FLUTE transmission.

#### compact-no-code-fec-2.iso3

This file contains one JPEG file stored as items. Compact No-Code FEC is used and the file is partitioned into three source blocks. The file contains also one hint track for FLUTE transmission.

#### mbms-fec.iso3

This file contains two JPEG files stored as items. MBMS-FEC is used and both files are partitioned into one source block and several sub-blocks. Parity symbols for the source block are stored as an FEC reservoir item. The file contains also three hint tracks for FLUTE transmission with FEC overheads 10%, 20% and 40%. Each hint track defines transmission of both JPEG files over one FLUTE channel.

#### fragment-random-access-1+AF8-rev1.mp4

This uses movie fragment random access boxes (movie fragment random access, track fragment random access, movie fragment random access offset in ISO/IEC 14496-12). These boxes help readers to search where random access points are.

There are random access points at 1-sec interval.

For the purpose of reference, the initial 1-second movie is followed by a 14-second movie fragment. The movie consists of audio and video tracks. Fragment aware readers should play 15.3 seconds of content, fragment-unaware readers only 1 second.

#### fragment-random-access-2.mp4

This uses movie fragment random access boxes as 1.4.

There are random access points at 5-sec interval.

For the purpose of reference, the initial 1-second movie is followed by a 10.1-second movie fragment. The movie consists of audio and video tracks. Fragment aware readers should play 15.1 seconds of content, fragment-unaware readers only 5 seconds.

#### timed-metadata.mp4

Simple audio file with a timed-metadata track.

#### restricted.mp4

|  |  |
| --- | --- |
| **Features** | * Post-decoder requirements on media   + ‘resv’ box   + ‘rinf’ box   + ‘stvi’ box   + ‘schi’ box |

This bitstream exercises the method of restricting access to video media. The media itself is in stereoscopic 3D format with side-by-side frame packing.

#### sg-tl-st.mp4

|  |  |
| --- | --- |
| **Features** | * Level assignment   + ‘leva’ box * Sub-tracks   + ‘strk’ box   + ‘stri’ box   + ‘strd’ box   + ‘stsg’ box * Segment index box   + ‘sidx’ box * Sub-segment index box   + ‘ssix’ box * Temporal level sample group   + Contained in ‘sgpd’ box * Sample groups   + ‘spgd’ box   + ‘sbgp’ box |

This bitstream exercises a method of splitting video media data such that it can be efficiently transmitted, decoded, and played back at full framerate or half framerate.

The first M samples in the ‘mdat’ box (sample group 1) give an independently decodable media segment, and the remaining samples (sample group 2) are not referenced by any samples in group 1.

Sample groups 1 and 2 are assigned to temporal levels 1 and 2 respectively. Each level is referenced as a byte range in the sub-segment index. Two sub-tracks are defined, the first of which applies just to sample group 1, and the second of which applies to the full range of samples.

#### subs\_tile\_hvc1.mp4

|  |  |
| --- | --- |
| **Features** | * sub-sample information   + ‘subs’ box * hvc1 sample entry   + ‘hvc1’ box |

This file contains HEVC video track with associated tile-based sub-sample (flags = 2). The sub-sample information box (‘subs’) is used to provide information such as sub-sample sizes in bytes, priority, etc. without having to decode HEVC data. The sub-samples are derived by using entry point offsets found from the HEVC slice header. The values of discardable and subsample\_priority fields in the ‘subs’ box are set to 0 and the field subsample\_size is set to the entry point offset value.

#### subs\_slice\_hvc1.mp4

|  |  |
| --- | --- |
| **Features** | * sub-sample information   + ‘subs’ box * hvc1 sample entry   + ‘hvc1’ box |

The file contains HEVC video track with associated slice-based sub-sample (flags = 4). Sub-sample is constructed by using the sizes of the slices. The values of discardable and subsample\_priority fields in the ‘subs’ box are set to 0 and the field subsample\_size is set to the slice size in bytes.

[Editor's note: There is a [problem](https://github.com/MPEGGroup/FileFormatConformance/issues/8) with this file. The sample\_delta of the first subsample entry is set to 0 while it shall always be 1]

#### aggr\_hvc1.mp4

|  |  |
| --- | --- |
| **Features** | * aggregators   + Aggregator container in the HEVC bit stream   + Aggregates SHVC enchancement layer * hvc1 sample entry   + ‘hvc1’ box |

This file includes an aggregator container around SHVC enhancement layer NAL units so that the video can be viewed and decoded as normal HEVC. Decoder supporting aggregators can increase video resolution by extracting the enhancement layer. Aggregator is a special type of NAL unit including or referring to a set of data. In this coded bit stream, the aggregator includes the aggregated data.

#### trgr\_hvc1.mp4

|  |  |
| --- | --- |
| **Features** | * track groups   + ‘trgr’ box * hvc1 sample entry   + ‘hvc1’ box |

This bit stream includes a ‘trgr’ box that groups related tracks together and a ‘hvc1’ box that contains the HEVC video configurations.

#### alst\_hvc1.mp4

|  |  |
| --- | --- |
| **Features** | * alternative startup (alst) sequences sample group   + ‘sgpd’ box   + ‘sbgp’ box * RAP sample group   + ‘sgpd’ box   + ‘sbgp’ box * hvc1 sample entry   + ‘hvc1’ box |

The alternative startup sequences sample group is indicated by two parameters: roll\_count and first\_output\_sample, which specify the number of samples in the group and the first sample in the sequence to be sent to the output, respectively. In addition, the random access point (RAP) sample group is used to point the location of the sync samples where alst sample group could be found in the stream. The RAP sample group is composed of CRA/BLA/IDR/IRAP pictures. A position in the stream is sought among the RAP samples preceding the position of interest. If the selected sample also belongs to the alst sample group, alternative startup sequence is activated. With roll\_count = 2 and first\_output\_sample = 1, *S*0 is sent to a decoder, *S*1...*S*7 are skipped, and *S*8 is sent to a decoder after which alternative startup sequence is complete and the decoding continues normally.

#### rtp\_rtcp\_reception\_hint\_tracks\_v2.mp4

|  |  |
| --- | --- |
| **Features** | * RTP Reception Hint Track sample entry   + ‘rrtp’ box * Timescale entry box   + ‘tims’ box * Time offset box   + ‘tsro’ box * Timestamp Synchrony box   + ‘tssy’ box * RTP Track SDP Hint Information box   + ‘sdp ’ box * RTCP Reception Hint Track sample entry   + ‘rtcp’ box * Track Reference Type Box: Content description   + ‘cdsc’ box * RTP movie information box   + ‘rtp ’ box * Hint media header (in RTP reception hint tracks)   + ‘hmhd’ box * Null media header box (in RTCP reception hint tracks)   + ‘nmhd’ box |

This file contains two media tracks, two RTP Reception Hint Tracks, and two RTCP Reception Hint Tracks. The file is composed of Received RTP Hint Sample Entry box (‘rrtp’) for the RTP Reception Hint Track description, timescale entry box (‘tims’) for timescale, time offset box (‘tsro’) for first sample RTP timestamp offset, and timestamp synchrony box (‘tssy’) for signaling whether a track has been synchronized during recording. timestamp\_sync field has been set to 1 indicating no synchronization was done during recording. Both RTP Reception Hint Tracks also have an RTP Track SDP Hint Information box (‘sdp ’) containing a media description part of an SDP message. The RTCP reception Hint Tracks are described by Received RTCP Hint Sample Entry box (‘rtcp’). In addition, RTCP tracks have a Content Description box (‘cdsc’) that refers to the associated RTP Reception Hint Track. The part of the SDP message before the Media description can be found in rtpmoviehintinformation box (‘rtp ’). The Hint Media Header box (‘hmhd’) has statistics on the hint track.

#### hvc2\_extractors.mp4

|  |  |
| --- | --- |
| **Features** | * 'hvc1' and 'hvc2' sample entries * Extractors with both inline constructors and sample constructors in samples of 'hvc2' track * 'scal' track reference type |

The file contains two tracks: an hvc1 base track, and an hvc2 extractor track. Each of these tracks carry their sample data in separate mdat boxes; that is, there is one mdat box with coded samples for the hvc1 track, and there is another mdat box that carries the extractor NAL units for the hvc2 track.

The hvc1 track carries a HEVC coded sequence containing 31 frames, where the first frame in decode order is an IDR frame and the rest are predicted B frames. Each frame of the coded sequence in this track has a spatial resolution of 512 x 256. Furthermore, every frame of this coded sequence is coded in a 2 x 2 tiling mode, where each tile has a spatial resolution of 256 x 128. Each tile is coded in a single slice and every slice contains only one tile. The tiles are motion-constrained, i.e., depend only on the collocated tiles in the reference pictures.

The hvc2 track is an extractor track, hence it contains a ‘scal’ type track reference in its tref box. Each sample of this track extracts one tile (the first tile in raster scan order) of the linked hvc1 track. The extracted NAL data is such that, resolving and constructing the samples of this track results in a valid HEVC bitstream.

#### hvc1\_only.mp4

This file contains a single, non-scalable ‘hvc1’ track

#### hev1\_only.mp4

This file contains a single, non-scalable ‘hev1’ track

#### hevc\_tiles\_single\_track\_nalm.mp4

This file contains:

* + independent 3x3 tiled HEVC single layer;
  + one ‘hvc2’ track;
  + “trif” and “nalm” with groupID=0 sample group descriptions;
  + sample to group box with grouping\_type=“nalm” and grouping\_type\_parameter=“trif”.

#### hevc\_tiles\_single\_track\_nalm\_rle.mp4

This file contains:

* + independent 3x3 tiled HEVC single layer;
  + one ‘hvc2’ track;
  + sample to group box with grouping\_type=“nalm” and grouping\_type\_parameter=“trif”;
  + “trif” and “nalm” run-length encoded with groupID=0 sample group descriptions.

#### hevc\_tiles\_multiple\_tracks\_v2.mp4

This file contains:

* independent 3x3 tiled HEVC single layer;
* one ‘hvc2’ and 9 ‘hvt1’ tracks each using “trif” with default sample group description (‘sgpd’ with version=2 and no ‘sbgp’);
* ‘tbas’ / ‘sabt’ track references for implicit reconstruction.

#### hevc\_tiles\_single\_track\_trif\_full\_picture\_v2.mp4

This file contains a single ‘hvc2’ track with ‘trif’ sample group description box (‘sgpd’ with version=2) using full\_picture flag and has no corresponding ‘sbgp’ box as allowed in 10.1 of ISO/IEC 14496-15.

#### hevc\_tiles\_single\_track\_nalm\_all\_intra.mp4

This file contains:

* + independent 3x3 tiled HEVC single layer;
  + one ‘hvc2’ track;
  + “trif” with independent\_idc=2 and “nalm” with groupID=0 sample group descriptions;
  + sample to group box with grouping\_type=“nalm” and grouping\_type\_parameter=“trif”.

#### shvc\_hvc1\_single\_track.mp4

See mhvc\_hvc1\_single\_track.mp4.

#### mhvc\_hvc1\_single\_track.mp4

These files contain SHVC or MV-HEVC with the following features:

* + Simple LHEVC bitstream in ‘hvc1’ track;
  + HEVC and LHEVC config;
  + ‘oinf’ and ‘linf’ sample groups.

#### shvc\_hev1\_single\_track.mp4

See mhvc\_hev1\_single\_track.mp4.

#### mhvc\_hev1\_single\_track.mp4

These files contain SHVC or MV-HEVC with the following features:

* + Simple LHEVC bitstream in ‘hev1’ track;
  + all parameter sets are stored in-band;
  + HEVC and LHEVC config;
  + ‘oinf’ and ‘linf’ sample groups.

#### shvc\_hvc2\_single\_track.mp4

See mhvc\_hvc2\_single\_track.mp4.

#### mhvc\_hvc2\_single\_track.mp4

These files contain SHVC or MV-HEVC with the following features:

* + Simple LHEVC bitstream in ‘hvc2’ track;
  + HEVC and LHEVC config;
  + ‘oinf’ and ‘linf’ sample groups.

#### shvc\_hev2\_single\_track.mp4

See mhvc\_hev2\_single\_track.mp4.

#### mhvc\_hev2\_single\_track.mp4

These files contain SHVC or MV-HEVC with the following features:

* + Simple LHEVC bitstream in ‘hev2’ track;
  + all parameter sets are stored in band;
  + HEVC and LHEVC config;
  + ‘oinf’ and ‘linf’ sample groups.

#### shvc\_hvc1\_hvc2\_multiple\_tracks\_extractors.mp4

See mhvc\_hvc1\_hvc2\_multiple\_tracks\_extractors.mp4.

#### mhvc\_hvc1\_hvc2\_multiple\_tracks\_extractors.mp4

These files contain SHVC or MV-HEVC with the following features:

* LHEVC bitstream in ‘hvc1’ track (base layer) and ‘hvc2’ track (enhancement layer);
* HEVC and LHEVC config;
* ‘oinf’ and ‘linf’ sample groups Extractors, ‘oref’, ‘sbas’ and ‘scal’ track references;
* ‘hvce’ brand.

#### shvc\_hev1\_hev2\_multiple\_tracks\_extractors.mp4

See mhvc\_hev1\_hev2\_multiple\_tracks\_extractors.mp4.

#### mhvc\_hev1\_hev2\_multiple\_tracks\_extractors.mp4

These files contain SHVC or MV-HEVC with the following features:

* LHEVC bitstream in ‘hev1’ track (base layer) and ‘hev2’ track (enhancement layer);
* HEVC and LHEVC config;
* ‘oinf’ and ‘linf’ sample groups;
* Extractors, ‘oref’, ‘sbas’ and ‘scal’ track references;
* all parameter sets are stored in-band;
* ‘hvce’ brand.

#### shvc\_hvc1\_lhv1\_multiple\_tracks\_implicit.mp4

See mhvc\_hvc1\_lhv1\_multiple\_tracks\_implicit.mp4.

#### mhvc\_hvc1\_lhv1\_multiple\_tracks\_implicit.mp4

These files contain SHVC or MV-HEVC with the following features:

* LHEVC bitstream in ‘hvc1’ track (base layer) and ‘lhv1’ track (enhancement layer);
* LHEVC config in ‘lhv1’ only;
* ‘oinf’ and ‘linf’ sample groups;
* No extractors, ‘oref’ and ‘sbas’ track references;
* ‘hvci’ brand.

#### shvc\_hev1\_lhe1\_multiple\_tracks\_implicit.mp4

See mhvc\_hev1\_lhe1\_multiple\_tracks\_implicit.mp4.

#### mhvc\_hev1\_lhe1\_multiple\_tracks\_implicit.mp4

These files contain SHVC or MV-HEVC with the following features:

* LHEVC bitstream in ‘hev1’ track (base layer) and ‘lhe1’ track (enhancement layer);
* LHEVC config in ‘lhe1’ only;
* ‘oinf’ and ‘linf’ sample groups No extractors, ‘oref’ and ‘sbas’ track references;
* all parameter sets are stored in-band;
* ‘hvci’ brand.

#### shvc\_hev1\_lhe1\_multiple\_tracks\_implicit.mp4

See mhvc\_hev1\_lhe1\_multiple\_tracks\_implicit.mp4.

#### mhvc\_hev1\_lhe1\_multiple\_tracks\_implicit.mp4

These files contain SHVC or MV-HEVC with the following features:

* LHEVC bitstream in ‘hev1’ track (base layer) and ‘lhe1’ track (enhancement layer);
* LHEVC config in ‘lhe1’ only;
* ‘oinf’ and ‘linf’ sample groups No extractors, ‘oref’ and ‘sbas’ track references;
* all parameter sets are stored in-band;
* ‘hvci’ brand.

#### lhevc\_avc1\_lhv1.mp4

This file contains an AVC base and an SHVC enhancement layer with the following features:

* one avc1 track;
* one lhv1 track with oinf and linf;
* one sbas track reference from SHEVC to AVC track;
* all parameter sets in config records.

#### lhevc\_avc1\_lhe1.mp4

This file contains an AVC base and an SHVC enhancement layer with the following features:

* one avc1 track;
* one lhv1 track with oinf and linf;
* one sbas track reference from SHEVC to AVC track;
* all parameter sets in config records for AVC, inband for SHVC.

#### lhevc\_avc3\_lhv1.mp4

This file contains an AVC base and an SHVC enhancement layer with the following features:

* one avc1 track;
* one lhv1 track with oinf and linf;
* one sbas track reference from SHEVC to AVC track;
* all parameter sets in config records for SHVC, inband for AVC.

#### lhevc\_avc3\_lhe1.mp4

This file contains an AVC base and an SHVC enhancement layer with the following features:

* one avc1 track;
* one lhv1 track with oinf and linf;
* one sbas track reference from SHEVC to AVC track;
* all parameter sets in-band AVC and SHVC.

#### hevc\_hvc1\_hvc2\_implicit\_v2.mp4

This file contains an HEVC base and an HEVC temporal sublayer with the following features:

* one hvc1 track with temporalID 0;
* one hvc2 track with temporalID 1;
* linf default sample group;
* sbas track references from hvc2 to hvc1 track;
* all parameter sets in config records;
* no extractors;
* ‘hvci’ brand.

#### hevc\_hev1\_hev2\_implicit\_v2.mp4

This file contains an HEVC base and an HEVC temporal sublayer with the following features:

* one hev1 track with temporalID 0;
* one hev2 track with temporalID 1;
* linf (default) sample group;
* sbas track references from hev2 to hev1 track;
* all parameter sets in-band;
* no extractors;
* ‘hvci’ brand.

#### hevc\_hvc1\_hvc2\_extractors\_v2.mp4

This file contains an HEVC base and an HEVC temporal sublayer with the following features:

* + one hvc1 track with temporalID 0;
  + one hvc2 track with temporalID 1;
  + linf (default) sample group;
  + sbas and scal track references from hvc2 to hvc1 track;
  + all parameter sets in config records;
  + extractors;
  + ‘hvce’ brand.

#### hevc\_hev1\_hev2\_extractors\_v2.mp4

This file contains an HEVC base and an HEVC temporal sublayer with the following features:

* one hev1 track with temporalID 0;
* one hev2 track with temporalID 1;
* linf (default) sample group;
* sbas and scal track reference from hev2 to hev1 track;
* all parameter sets in-band;
* extractors;
* ‘hvce’ brand.

#### C001.heic – C041.heic

See HEIF\_conformance.xlsx.

#### iff\_hevc\_single\_item.heic

Simple image item describing an HEVC image.

#### iff\_hevc\_single\_item\_main10.heic

Simple image item describing an HEVC image conforming to the Main 10 profile.

#### iff\_hevc\_tile\_multiple\_items\_tbas.heic

The file contains:

* 4 tile items of type 'hvt1' coding entirely an image, with a 2x2 tiling. Each item has 2 essential properties: the 'ispe' property indicates tile width and height and the 'hvcC' property contains the configuration for the complete image (i.e. not for decoding the tile as small image). The 'hvcC' property is the same for all tile items and is shared with the same property index. The 'rloc' property is present in each item and a 'tbas' item reference is used.
* An 'hvc1' item corresponding to the complete image. The NAL units are shared between the 'hvc1' and 'hvt1' items by using construction\_method = 2.

The brands contain 'mif1' and 'heic'.

#### multilayer001.heic - multilayer004.heic

multilayer001: A base quality and an enhanced quality presentations of same image.

multilayer002: Two image grid derived items, 2x2 each. Image items in the second grid have enhanced SNR.

multilayer003: A multi-view file with 'ster' grouping.

multilayer004: A base quality and an enhanced quality presentations of same image. The baselayer is an AVC-coded item.

[Editor's note: There is a [problem](https://github.com/MPEGGroup/FileFormatConformance/issues/9) with the above files. Undefined version 2 of the ItemInfoBox is used.]

#### multilayer005.heic

This file contains a single MV-HEVC coded two-layer image. The HTM (version 15.0) reference software was used to encode this image. The left view is coded as the base layer, and the right view is coded as the enhancement layer. The heic file contains two image items. The first is an ‘hvc1’ image item for the left view. The other is an ‘lhv1’ image item, associated with an ‘lsel’ property, for the right-view. The two image items are also grouped using the ‘ster’ grouping to indicate that they form a stereo-pair.

[Editor's note: There is a [problem](https://github.com/MPEGGroup/FileFormatConformance/issues/9) with the above file. Undefined version 2 of the ItemInfoBox is used.]

#### c042.heic

This file contains a HEVC coded image item associated with an ‘imir’ property, which indicates that the reconstructed image is mirrored about the vertical axis.

#### FX-VY-9436R.3\_qhd-variant.mp4

The compliant sample file for ISO/IEC 23001-12:2015 consists of following three files;

* FX-VY-9436R.3\_qhd.mp4
* FX-VY-9436R.3\_qhd-variant.mp4
* FX-VY-9436R.3\_qhd.vinput.merged.xml

FX-VY-9436R.3\_qhd.mp4 is the ISOBMFF media track and FX-VY-9436R.3\_qhd-variant.mp4 is the associated variant track, ISOBMFF metadata track.

These files rely on the “external context” option for associating a variant track to a media track. As such, there is no Track Reference box within the media track to make that association.

The keys for decrypting the content are in the XML file, FX-VY-9436R.3\_qhd.vinput.merged.xml, in Keyset Delivery Format (KDF). The KDF file contains the collection of all keys used to create the files, including all media and metadata keys for the purpose of providing all the information as conformance file. Typically, an individual player/decoder would receive just the subset of keys that particular player needs.

#### hev1\_clg1\_header.mp4

This file is a CMAF Header / DASH Initialisation Segment containing two colr boxes in an HEVC track with hev1 sample entry. The associated file hev1\_clg1\_segment.m4s is a CMAF Segment / DASH Media Segment which can be used in conjunction with the above. The video contained in the segment is only black frames (with some text in the top right), to avoid licensing issues. The information in the colour box matches the VUI, but since there is a VUI and an SEI and HLG is in effect we see two colour boxes with identical colour primaries and matrix coefficients, but the first documents the HLG transfer function and the second the BT.2020 code-point for standard gamma (i.e. the same as 709).

#### Compressed boxes

All sequences show a 2s video counter at 25fps.

**comp\_moov\_isoc.mp4**

The file contains a compressed moov and ‘isoC’ brand.

**comp\_moov\_otyp.mp4**

The file contains a compressed moov, a ‘ftyp’ with major brand ‘comp’ and a ‘otyp’ wrapping the original ‘ftyp’ with various brand info.

**comp\_moof\_nobrand.mp4**

The file contains an empty ‘moov’, no changes in brand info and compressed ‘moof’. A player not understanding compressed boxes could see this file as an init segment (empty ‘moov’ only).

**comp\_moof\_otyp.mp4**

The file contains ‘ftyp’=comp, ‘otyp’, an uncompressed empty ‘moov’ and compressed ‘moof’.

**comp\_moof\_sidx\_otyp.mp4**

The file contains ‘ftyp’=comp, ‘otyp’, an uncompressed empty ‘moov’ and compressed ‘moof’ and ‘sidx’.

**comp\_moof\_sidx\_ssix\_otyp.mp4**

The file contains ‘ftyp’=comp, ‘otyp’, an uncompressed empty ‘moov’ and compressed ‘moof’, ‘sidx’ and ‘ssix’.

**comp\_all\_otyp.mp4**

The file contains ‘ftyp’=comp, ‘otyp’ and compressed empty ‘moov’, ‘moof’, ‘sidx’ and ‘ssix’.

NOTE Due to the sequence being short (2s), the ‘sidx’ and ‘ssix’ boxes are being forced to their compressed versions, although their original sizes are smaller than their compressed sizes.

#### CENC conformance files

All DRM configuration files are present in the folder ***./drm\_cfg/***, each file containing the KID and key value for each key used.

Each protected ISOBMF file also contains a PSSH box using GPAC test system ID, which contains the key values in the PSSH for simple decryption without KMS.

All video sequences show a 2s video counter at 25fps, 1280x720, 420 8 bit HEVC 3x3 motion constrained tile-set.

All audio sequences play a 2s audio bip/bop at 44100Hz, mono, AAC.

All image sequences show a single HEVC picture of size 1280x720, 420 8 bit, 3x3 tiled.

**\*frag1s.mp4**

‘seig’ sample to group mapping in movie fragments.

**image\_\* and video\_cenc\_mkey\_\***

Both CENC-128 and CBCS with constant IV.

**\*\_cenc\_mkey\_subs\* and \*\_cbcs\_mkey\_const\_iv\_subs\***

Partial encryption of the tiles in the source frames.

#### VVC conformance files

**vvc\_basic\_track.mp4**

Single layer coded bitstream in VVC track with sample entry 'vvc1'

**vvc\_mixed\_nal\_subpicture\_tracks.mp4**

Single layer coded bitstream with subpictures having mixed NAL unit type. One VVC base track with sample entry 'vvc1'. Four subpictures coded in 4 VVC subpicture tracks with sample entry 'vvs1'. VVC merge base track has 'subp' track reference to VVC subpicture tracks. VVC merge base track has 'mixn' track reference to VVC subpicture tracks. Subpicture order sample grouping 'spor' with num\_subpic\_ref\_idx equal to 0 (same order). Mixed NAL unit type pictures sample group 'minp'. Same NAL unit type track grouping 'snut'.

**vvc\_subpicture\_tracks.mp4**

Single layer coded bitstream with subpictures. One VVC base track with sample entry 'vvc1'. Four subpictures coded with two different QPs in 8 VVC subpicture tracks with sample entry 'vvs1'. VVC merge base track has 'subp' track reference to VVC subpicture tracks. VVC subpicture tracks are grouped by 'alte' track grouping. Subpicture order sample grouping 'spor' with num\_subpic\_ref\_idx equal to 0 (same order). Subpicture layout map entry 'sulm'.

**vvc\_subpicture\_tracks\_spor\_ordered.mp4**

Single layer coded bitstream with subpictures. One VVC base track with sample entry 'vvc1'. Four subpictures coded in 4 VVC subpicture tracks with sample entry 'vvs1'. VVC merge base track has 'subp' track reference to VVC subpicture tracks. Subpicture order sample grouping 'spor' with num\_subpic\_ref\_idx greater than 0 (different order).

**vvc\_substitute\_subpicture\_single\_sample\_track.mp4**

Single layer coded bitstream with subpictures. One VVC base track with sample entry 'vvc1'. Four subpictures coded in 4 VVC subpicture tracks with sample entry 'vvs1', 4 substitute subpicture track with sample entry 'vvs1' (only one sample in the track) (In VvcNALUConfigBox flags&1=1). VVC merge base track has 'subp' track reference to VVC subpicture tracks. Subpicture order sample grouping 'spor' with num\_subpic\_ref\_idx equal to 0 (same order) (11.4.16),

**vvc\_substitute\_subpictures\_all\_frames.mp4**

Single layer coded bitstream with subpictures. One VVC base track with sample entry 'vvc1'. Four subpictures coded in 4 VVC subpicture tracks with sample entry 'vvs1', 4 substitute subpicture track with sample entry 'vvs1' (time aligned samples) (In VvcNALUConfigBox flags&1=1). VVC base track has 'subp' track reference to VVC subpicture tracks. Subpicture order sample grouping 'spor' with num\_subpic\_ref\_idx equal to 0 (same order).

**vvc\_substitute\_subpictures\_single\_sample\_track.mp4**

[Editor's note: missing description]

**vvc\_substitute\_subpictures\_track.mp4**

[Editor's note: missing description]

**RA\_BodeMuseum\_dcfi\_1920x1080\_qp37.mp4**

Test Vector with two vvc1 sample entries with a dcfi sample group to which all samples in the track are mapped.

**RA\_BodeMuseum\_inband\_ps\_1280x720\_qp37.mp4**

Test Vector with inband parameter sets.

**RA\_BodeMuseum\_mixed\_res\_qp37.mp4**

Test Vector with inband parameter sets and resolution change.

**RA\_BodeMuseum\_outband\_ps\_1920x1080\_qp37.mp4**

Test Vector with a vvc1 sample entry.

**RA\_BodeMuseum\_pase\_unused\_1280x720\_qp37.mp4**

Test Vector with SAP samples mapped to 'pase' sample groups where the max\_widh and max\_height do not match the value in the grouping\_type\_parameter ad therefore the 'pase' sample groups are not used.

**RA\_BodeMuseum\_pase\_used\_1920x1080\_and\_1280x720\_qp37.mp4**

Test Vector with SAP samples mapped to 'pase' sample groups with a resolution change and with the max\_widh and max\_height matching the value in the grouping\_type\_parameter and therefore the 'pase' sample groups are used.

#### Event Message Track Sample Files

**IF1\_V2\_captures/data-2-Data3\_3.mp4**

Captured file from Ateme Titan Live of event message track.

**IF1\_V2\_captures/audio-und-128000.isma**

Associated audio track capure from titan live.

**video-720x480-500000.ismv**

Associated video track capure from titan live

**out\_avail\_track.cmfm**

Avail track generated with unified open-source implementation using reference without overlaps. An event message track of 600 seconds (10 minutes) with 2 second segments, and 30 second ad break avails using a SCTE-35 splice\_insert message every 180 seconds (3 minutes).

**out\_avail\_track.mpd**

Corresponding EventStream element.

**out\_overlap.cmfm**

Avail track generated with unified open- source implementation using reference with overlaps.

**avail\_overlap.mpd**

Corresponding edited EventStream element

#### Multiplexed timed metadata tracks (mebx)

**test\_mebx\_me4c.mp4**

uses key\_namespace=me4c for all metadata items from Table 1. The color sample format is defined as:

class ColorSample extends Box(local\_key\_id){  
 unsigned int(32) x;  
 unsigned int(32) y;  
 unsigned int(32) width;  
 unsigned int(32) height;  
}

key\_value from the MetadataKeyDeclarationBox is used as the local\_key\_id with the following FourCCs:

* redd Red color
* blue Blue color
* ylow Yellow color
* whte White color
* blck Black color

The mebx sample file uses the pattern NIbFD which repeats 6 times giving a total of 30 frames for the video track (1 second duration). The metadata associated with each frame is summarized in the table below:

Table 1: Metadata per frame (this pattern repeats 6 times)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Metadata | N  Netherlands | I  Indonesia | b  Blue | F  France | D  Germany |
| Red | 0,0,64,16 | 0,0,64,24 | NULL | 44,0,20,48 | 0,16,64,16 |
| Blue | 0,32,64,16 | NULL | 0,0,64,48 | 0,0,20,48 | NULL |
| Yellow | NULL | NULL | NULL | NULL | 0,32,64,16 |
| White | 0,16,64,16 | 0,24,64,24 | NULL | 20,0,24,48 | NULL |
| Black | NULL | NULL | NULL | NULL | 0,0,64,16 |

#### HEVC\_1440x960\_derived\_image\_SSTR.heic

HEIF Single Stream Item Property – A file with SingleStreamProperty associated with overlay derived image item.

#### HEVC\_8192x4096\_25fps\_NALM\_PPRR\_PIPM.mp4

Picture-in-Picture - A file with PicInPicRegionReplacementEntry when region\_id\_type is equal to 1.

#### HEIF extrinsic intrinsic properties

**ex\_in\_trinsics/intrinsics/intrinsics.heic**

A file with intrinsics properties with the following contents:

Item 1:

fx = 512

fy = 512

s = 0

cx = 256

cy = 96

Coded as:

flags: 6656

Values: [33554432, 16777216, 8388608]

Item 2:

fx = 512

fy = 501.5

s = -100.5

cx = 256

cy = 96

Coded as:

flags: 72193

Values: [33554432, 16777216, 8388608, 43821739, -201]

**ex\_in\_trinsics/extrinsic\_position/position.heic**

The file contains 6 image items with 5 different extrinsic positions:

Item 1:

Coordinate System ID: 0

Position: 0.0, 0.0, 0.0

Coded as flags: 0; No values

Item 2:

Coordinate System ID: 1

Position: 0.01, 0.0, 0.0

Coded as flags: 33; Values: [10000, 1]

Item 3:

Coordinate System ID: 0

Position: 0.0, 0.02, 0.0

Coded as flags: 2; Values: [20000]

Item 4:

Coordinate System ID: 1

Position: 0.0, 0.0, -0.04

Coded as flags: 36; Values: [-40000, 1]

Item 5:

Coordinate System ID: 1

Position: -0.03, 0.05, -0.08

Coded as flags: 39; Values: [-30000, 50000, -80000, 1]

Item 6:

Coordinate System ID: 0

Position: 0.0, 0.0, 0.0

Coded as flags: 0; No values

Coordinate System ID: 1

Position: -0.03, 0.05, -0.08

Coded as flags: 36; Values: [-30000, 50000, -80000, 1]

**ex\_in\_trinsics/extrinsic\_rotation/ rotation\_[00,...,15].heic**

The files with extrinsics properties with rotation signaling. All files have extrinsics position set to 0.1, 0.1, -0.1 (in meters).

rotation\_00.heic:

Quaternion: 0.987965 0.0940609 -0.0789265

Input angles: X:170, Y:10, Z:10

Corresponding input matrix:

0.9698, 0.2007, -0.1383,

0.1710, -0.9646, -0.2007,

-0.1736, 0.1710, -0.9698,

rotation\_01.heic:

Quaternion: 0.363338 0.665546 -0.522402

Input angles: X:-110, Y:64, Z:170

Corresponding input matrix:

-0.4317, 0.8912, 0.1396,

0.0761, 0.1902, -0.9788,

-0.8988, -0.4119, -0.1499,

rotation\_02.heic:

Quaternion: -0.0469889 0.386794 0.913955

Input angles: X:45, Y:10, Z:170

Corresponding input matrix:

-0.9698, -0.2437, 0.0019,

0.1710, -0.6750, 0.7177,

-0.1736, 0.6964, 0.6964,

rotation\_03.heic:

Quaternion: 0.0789265 0.0940609 0.0789265

Input angles: X:10, Y:10, Z:10

Corresponding input matrix:

0.9698, -0.1413, 0.1986,

0.1710, 0.9751, -0.1413,

-0.1736, 0.1710, 0.9698,

// This creates a quaternion with w < 0, which should trigger a sign change

rotation\_04.heic:

Quaternion: 0.242726 0.241691 -0.789943

Input angles: X:170, Y:141, Z:62

Corresponding input matrix:

-0.3648, 0.9208, -0.1376,

-0.6862, -0.3659, -0.6287,

-0.6293, -0.1349, 0.7653,

// This is the worst case wrt precision when running the precision tests

rotation\_05.heic:

Quaternion: -0.99981 0.00872521 0.0174517

Input angles: X:0, Y:178, Z:179

Corresponding input matrix:

0.9992, -0.0175, -0.0349,

-0.0174, -0.9998, 0.0006,

-0.0349, -0.0000, -0.9994,

// Testing setting various angles to zero

rotation\_06.heic:

Quaternion: 0.707107 0 0

Input angles: X:90, Y:0, Z:0

Corresponding input matrix:

1.0000, 0.0000, 0.0000,

0.0000, 0.0000, -1.0000,

-0.0000, 1.0000, 0.0000,

rotation\_07.heic:

Quaternion: 0 0.707107 0

Input angles: X:0, Y:90, Z:0

Corresponding input matrix:

0.0000, 0.0000, 1.0000,

0.0000, 1.0000, 0.0000,

-1.0000, 0.0000, 0.0000,

rotation\_08.heic:

Quaternion: 0 0 0.707107

Input angles: X:0, Y:0, Z:90

Corresponding input matrix:

0.0000, -1.0000, 0.0000,

1.0000, 0.0000, 0.0000,

-0.0000, 0.0000, 1.0000,

rotation\_09.heic:

Quaternion: 0.130526 0 0

Input angles: X:15, Y:0, Z:0

Corresponding input matrix:

1.0000, 0.0000, 0.0000,

0.0000, 0.9659, -0.2588,

-0.0000, 0.2588, 0.9659,

rotation\_10.heic:

Quaternion: 0.12941 0.12941 -0.0170371

Input angles: X:15, Y:15, Z:0

Corresponding input matrix:

0.9659, 0.0670, 0.2500,

0.0000, 0.9659, -0.2588,

-0.2588, 0.2500, 0.9330,

rotation\_11.heic:

Quaternion: 0.111411 0.145194 0.111411

Input angles: X:15, Y:15, Z:15

Corresponding input matrix:

0.9330, -0.1853, 0.3085,

0.2500, 0.9504, -0.1853,

-0.2588, 0.2500, 0.9330,

rotation\_12.heic:

Quaternion: -0.0170371 0.12941 0.12941

Input angles: X:0, Y:15, Z:15

Corresponding input matrix:

0.9330, -0.2588, 0.2500,

0.2500, 0.9659, 0.0670,

-0.2588, 0.0000, 0.9659,

rotation\_13.heic:

Quaternion: 0 0 0.130526

Input angles: X:0, Y:0, Z:15

Corresponding input matrix:

0.9659, -0.2588, 0.0000,

0.2588, 0.9659, 0.0000,

-0.0000, 0.0000, 1.0000,

rotation\_14.heic:

Quaternion: 0 0.130526 0

Input angles: X:0, Y:15, Z:0

Corresponding input matrix:

0.9659, 0.0000, 0.2588,

0.0000, 1.0000, 0.0000,

-0.2588, 0.0000, 0.9659,

**camera\_parameters\_1\_v1.heic**

An image item with CameraIntrinsicMatrixProperty ('cmin') and CameraExtrinsicMatrixProperty ('cmex').

CameraExtrinsicMatrixProperty is with version is equal to 0 and flags field is equal 63

* pos\_x/y/z are present
* orientation is present using quaternions and in large format
* ID is present

CameraIntrinsicMatrixProperty is with version equal to 0 and flags field equal 65793

* denominatorShiftOperand is equal to 1
* skewDenominatorShiftOperand is equal to 1
* focal\_length\_y and skew\_factor are present

**camera\_parameters\_2\_v1.heic**

An image item with CameraIntrinsicMatrixProperty ('cmin') and CameraExtrinsicMatrixProperty ('cmex').

CameraExtrinsicMatrixProperty is with version is equal to 1 and flags field is equal 47

* pos\_x/y/z are present
* orientation is present using ViewpointGlobalCoordinateSysRotationStruct
* ID is present

CameraIntrinsicMatrixProperty is with version equal to 0 and flags field equal 65793

* denominatorShiftOperand is equal to 1
* skewDenominatorShiftOperand is equal to 1
* focal\_length\_y and skew\_factor are present

#### text\_item\_v1.heic

A text item with ImageSpatialExtentsProperty ('ispe') and TextLayoutProperty ('txlo'). Text item is associated with image item by an item reference of type 'text'

TextLayoutProperty has the following fileds

* reference\_width = 960
* reference\_height = 960
* x = 128
* y = 128

#### progressive-diagonal.heic

The progressive-diagonal.heic conformance sample is an example where a base quality image is progressively refined starting from the top-left and ending at the bottom-right. In this figure, for each cell, the large number indicates the item\_id of the corresponding input image, and the small number at the bottom left indicates the progressive rendering step for this cell.

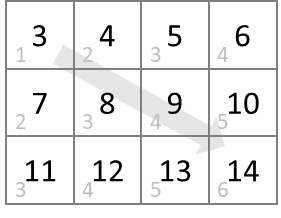


Figure 1: Progressive Diagonal Display Order

#### progressive-focus.heic

The progressive-focus.heic file, according to the ‘prgr’ entity group and the ‘prdi’ item properties stores the data with the thumbnail first, followed by the input images corresponding to the in-focus areas, and finishing with the remaining input images. The ‘prgr’ entity group and the ‘prdi’ item property indicate to a player that when realizing a progressive rendering of the grid, two intermediate reconstructions are of particular interest: a first preview of the primary item (the grid image item) using the thumbnail and an intermediate reconstruction (as shown on Figure 2) of the grid where the most interesting areas are displayed in high quality.

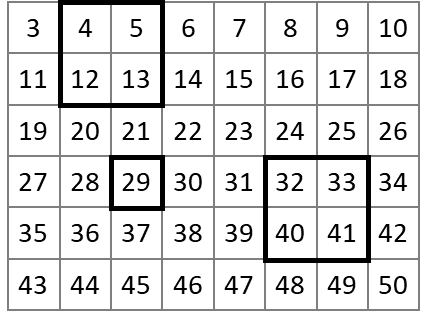


Figure 2: Progressive Focus Display Order (numbers correspond to item\_IDs)

#### Conformance Samples for Region Items

9 conformance samples for region items, are described in table below. There are 7 conformance samples for the different types of geometry that can be used in a region item. There are 2 conformance samples containing a list of regions inside a single region item. One is using 16 bits for defining the characteristics of the regions, the other is using 32 bits (i.e., in the first case, the value flags & 1 for the ‘rgan’ item is 0, in the second case it is 1).

|  |  |  |
| --- | --- | --- |
| Number | Filename | Feature |
| 1 | region-point.heic | Point region with ‘udes’ annotation. |
| 2 | region-rectangle.heic | Rectangle region with ‘udes’ annotation. |
| 3 | region-ellipse.heic | Ellipse region with ‘udes’ annotation. |
| 4 | region-polygon.heic | Polygon region with ‘udes’ annotation. |
| 5 | region-polyline.heic | Polyline region with ‘udes’ annotation. |
| 6 | region-mask.heic | Mask region with ‘udes’ annotation, where the mask is defined as an item. |
| 7 | region-inline-mask.heic | Mask region with ‘udes’ annotation, where the mask is defined inline in the region item. |
| 8 | region-list.heic | One rectangle and two points regions in a single region item, with ‘udes’ annotation. |
| 9 | region-list-large.heic | One rectangle and two points regions in a single region item, with ‘udes’ annotation, where the regions are defined using a field\_size of 32. |

3 conformance samples for derived region items are described in the table below. Each of these conformance samples contains an image item and a derived image item. A region item is associated with the image item inside which the regions are defined using an item reference of type ‘cdsc’ from the region item to the image item. A derived region item, includes a ‘drgn’ item reference to the region item. The derived region item is associated with the derived image item using an item reference of type ‘cdsc’ from the derived region item to the derived image item. Each region item contains a rectangle region and two point regions and is annotated with a ‘udes’ item property.

|  |  |  |
| --- | --- | --- |
| Number | Filename | Feature |
| 1 | region-derived-crop.heic | Derived region item using a ‘clap’ transformative item property. |
| 2 | region-derived-mirror.heic | Derived region item using an ‘imir’ transformative item property. |
| 3 | region-derived-rotation.heic | Derived region item using an ‘irot’ transformative item property. |

#### region-rectangle.mp4

This conformance sample contains a video track of 5 frames corresponding to a panning over a still image. This video track is annotated with a region track containing a rectangle region corresponding to the same area in the original still image. Therefore, the rectangle region is moving to follow the panning. This rectangle region is annotated with a ‘udes’ description stored in an ‘sbgp’ box.

#### essg\_sample1.mp4

A file with a ‘essg’ restricted track, with a sample group description with version 3, and an ‘esgh’ sample group listing the v3 sample group.

A file with the ‘esgh’ sample group with ‘stsd’ and no ‘cenc’ (a synthetic VVC track with a SAP sample group marked as essential).

A file with the ‘esgh’ sample group with num\_grouping > 1.

#### essg\_sample2.mp4

A file with a ‘essg’ restricted track, with a sample group description with version 3, and an ‘esgh’ sample group listing the v3 sample group.

A file with the ‘esgh’ sample group with ‘cenc’ (an encrypted synthetic VVC track).

#### Conforming files for picture-in-picture

**NALFF\_PiP\_FourPeople\_1280x720\_BQSquare\_416x240\_Offset\_0x480.mp4**

**NALFF\_PiP\_FourPeople\_1280x720\_BQSquare\_416x240\_Offset\_864x480.mp4**

The two conforming files use the following file format features:

1. The supplementary track reference ('supm'), indicating the relationship between a main video track and a supplementary video track,
2. The matrix fields of the TrackHeaderBoxes of the supplementary video track and the main video track, indicating the position and size of the overlaying window in the main video for overlaying the supplementary video, and
3. The layer fields of the TrackHeaderBox of the supplementary video track and the main video track, indicating that the supplementary video is overlaid on top of the main video.

There is only one difference between the two conforming files, on the position of the overlaying window, one at the bottom-right corner (as illustrated by Fig. 1), the other at the bottom-left corner.

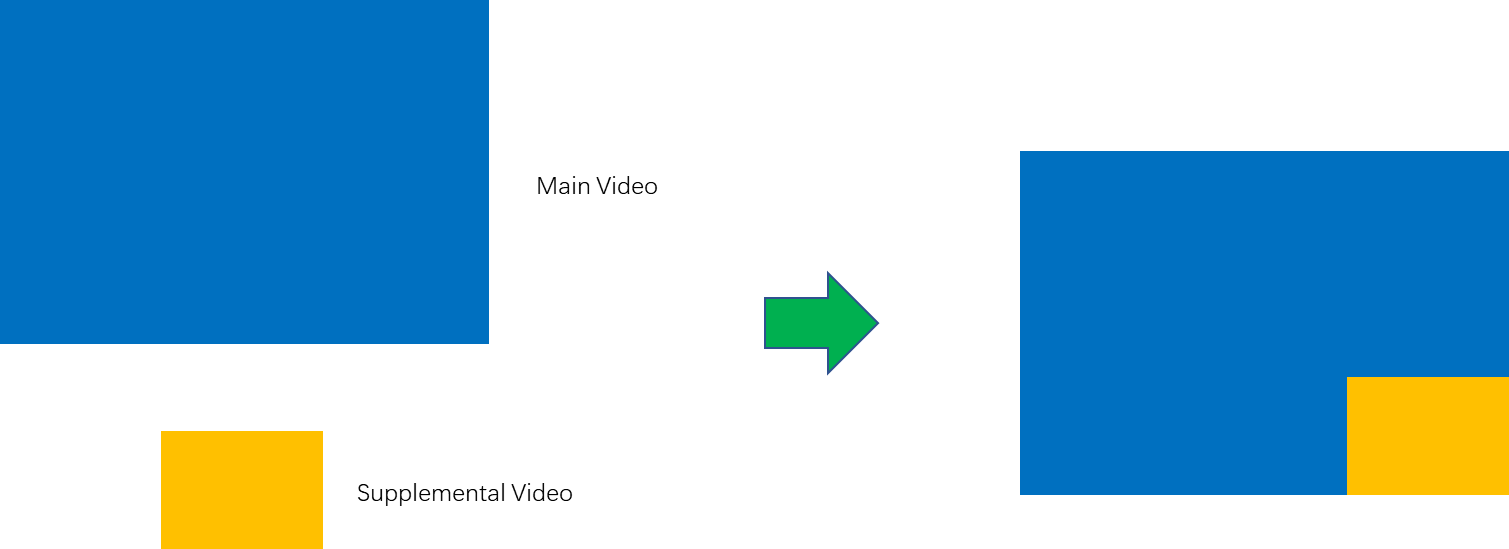


Figure 3: A picture-in-picture example by overlaying a supplemental video on top of a main video at the bottom-right corner

#### Static sample groups

This section provides conformance samples for the following flags values of the SampleGroupDescriptionBox:

* static\_group\_description: with value 1: when set to 1, this flag indicates that there are no SampleGroupDescriptionBoxes of this grouping\_type in any TrackFragmentBox of this track.
* static\_mapping: with value 2: when set to 1, this flag indicates that there are no SampleToGroupBoxes of this grouping\_type in this track (in neither the SampleTableBox nor any TrackFragmentBox of this track); all samples therefore map to the default

**sample\_group\_static\_group.mp4**

A ‘sgpd’ with version=1 and flags value set to 0x000001 for indicating static\_group\_description

**sample\_group\_static\_group\_frag.mp4**

A fragmented file with one ‘sgpd’ box with version=1 and flags value set to 0x00001 indicating static\_group\_description

**sample\_group\_static\_mapping.mp4**

A ‘sgpd’ with version=2 and flags value set to 0x000002 for indicating static\_mapping (i.e. no ‘sbgp’ box is present).

#### Default sample groups

This section provides conformance samples for the default sample group feature. This feature requires a SampleGroupDescriptionBox with version=2. It applies to both fragmented and non-fragmented files.

**sample\_group\_default\_desc.mp4**

This file provides an example with 50 samples mapped on 2 sample group description entries. The first 26 samples are explicitly mapped to a group\_description\_index in a ‘sbgp’ box and the remaining ones are implicitly mapped to the default\_group\_description\_index indication in the ‘sgpd’ box with version=2.

**sample\_group\_default\_desc\_frag.mp4**

This fragmented file contains 2 movie fragments, each with a ‘sgpd’ with version=2, declaring a default\_group\_description\_index with a value offset equal to 0x10000. In each ‘moof’, the 1st sample is explicitly mapped in a ‘sbgp’ to group\_description\_index =1 of the ‘sgpd’, the remaining ones to the default\_group\_description\_index

#### hevc\_tiles\_multiple\_tracks\_empty\_base\_v2.mp4

This file contains 1 tile base track ‘hvc2’ and 9 ‘hvt1’tile tracks (with ‘sabt’ and ‘tbas’ track references for implicit reconstruction). The tile base track has empty samples. Each tile track has a ‘trif’ sample group description box with version=2 and has no corresponding ‘sbgp’ box as allowed in 10.1 of ISO/IEC 14496-15.

#### EDRAP

**vvc\_edrap\_sap1.mp4**

This file contains two video tracks, a main track, containing a VVC single-layer bitstream with EDRAP samples, and an associated external stream track.

The main track has sample entry ‘vvc1’, and the EDRAP samples are grouped by the EDRAP sample group ‘edrp’ with edrap\_type equal to 1. The main track has an ‘aest’ track reference that references the associated external stream track.

The associated external stream track has the restricted sample entry ‘resv’ with the ‘rinf’ box, which contains the ‘schm’ box with scheme\_type equal to ‘spkt’.

**vvc\_edrap\_sap2.mp4**

Same as above, except that the value of edrap\_type is equal to 2.

**vvc\_edrap\_sap3.mp4**

Same as above, except that the value of edrap\_type is equal to 3.

#### vvc\_nnsei\_1.mp4

A Single layer coded bitstream in VVC track with sample entry 'vvc1'. Neural-network post-filter characteristics sample group ('nfcs') and neural-network post-filter activation sample group ('nfas').

#### Conformance for Efficient Sub-samples

The video sequence used in the conformance files from this clause is a synthetic sequence of 3 pictures, with size of 832x480 pixels, encoded in intra mode. Each picture contains 4 subpictures (SP1 to SP4) as illustrated in Figure 4:



Figure 4: Subpicture organisation

The VVC bitstream sequence is composed of 3 frames, as illustrated in Figure 5, with:

* The first frame comprising 1 PH (Picture Header) and 4 IDR VCL NAL units,
* All other frames comprising 1 PH and 4 CRA VCL NAL units.



Figure 5: VVC bitstream sequence

All the conformance samples of the VVC sequence use single track encapsulation, with ‘vvc1’ sample entry.

**efficientsubs\_subpictures\_vvc\_v1.mp4**

*‘subs’* version 2 syntax and *SubSampleReferenceTableBox* describing VVC subsamples corresponding to Subpicture-based sub-samples, as defined in section 11.3.6 of [4].

The values for the properties of the sub-samples of the first picture are not using references to the SubSampleReferenceTableBox (*has\_reference* = 0) and are provided in the subs box

The values for the properties of the sub-samples for following frames are identical to the properties in entries of the SubSampleReferenceTableBox (has\_reference=1).

**efficientsubs\_subsamples\_vvc\_v1.mp4**

*‘subs’* version 2 syntax and *SubSampleReferenceTableBox* describing VVC subsamples corresponding to NAL-unit-based sub-samples, as defined in section 11.3.6. of [4].

The values for the properties of all sub-samples are identical to the properties in entries of the SubSampleReferenceTableBox (*has\_reference* = 1)

#### mhm1\_preselection.mp4

This conformance file contains 2 preselections for MPEG-H 3D Audio (mhm1) signaled using the EntityToGroup box. The signaled audio scene contains a channel bed as well as 2 languages (English and Spanish) with the following 2 preselections:

* “Default”: Contains the channel bed and a language selection containing English and Spanish. The default prominence of the language track is set to 0.
* “Dialog+”: Contains the channel bed and a language selection containing English and Spanish. The default prominence of the language track is set to 9 for increased intelligibility.

Additionally, both preselections contain prominence as well as position interactivity.

#### mhm2\_preselection.mp4

This conformance file contains ISOBMFF preselections for multistream MPEG-H 3D Audio (mhm2). The described audio scene contains a channel bed as well as 3 languages (English, German and Korean) that result in 3 preselections:

* “English”: Contains channel bed and English commentary.
* “German”: Contains channel bed and German commentary.
* “Korean”: Contains channel bed and Korean commentary.

1. (informative)  
     
   Other tools

The status, utility, and correctness of these tools is unknown. This information is provided merely in case it is of assistance.

1. There is a set of open-source validation tools, distributed under the Apple Public Source Licence, in the open-source repository: <https://svn.macosforge.org/repository/ValidateMP4/>
2. The GPAC project has several tools dedicated to packaging and parsing of ISOBMFF files. More information on these tools can be found on the following website, <http://gpac.io>.

* MP4Box, a command-line packager, dumper and analyser of ISOBMFF files. It is distributed under the LGPL licence.
* mp4box.js, a JavaScript-based ISOBMFF parser, analyser and on-the-fly segmenter. It is distributed under the BSD-3 licence. The source code for mp4box.js is available at: <https://github.com/gpac/mp4box.js>. An analyser for inspecting ISOBMFF files based on this code is available at: <https://gpac.github.io/mp4box.js/test/filereader.html>

1. A website that contains example files and some background information of the High Efficiency Image file format (HEIF) is accessible at <https://nokiatech.github.io/heif>. The C++ source code and the JavaScript reader are provided through the GitHub Repository (<https://github.com/nokiatech/heif>). The webpages use JavaScript versions of the HEIF reader source code and a HEVC decoder (libde265.js, <http://www.libde265.org/>), i.e. parse and decode the example HEIF files in any JavaScript-capable browser implementation. The file writer configurations and input HEVC bitstreams to generate the HEIF-specific conformance files are provided in <https://github.com/nokiatech/heif_conformance>.
2. The FFmpeg project has support for reading and writing ISOBMFF files, included using movie fragments. More documentation on the project can be found at <http://ffmpeg.org>. FFmpeg is distributed under GPL or LGPL licence (depending on build settings).
3. Other tools include:

* <http://atomicparsley.sourceforge.net/>;
* <https://mediaarea.net/fr/MediaInfo>;
* <http://mp4creator.sourceforge.net/>: MP4 file creator, no longer maintained;
* <http://mp4parser.com/>, an online MP4 file parser.

**Bibliography**

[1] ISO/IEC 14496-26, *Information technology — Coding of audio-visual objects — Part 26: Audio conformance*

[2] ISO/IEC 23000-6, *Information technology — Multimedia application format (MPEG-A) — Part 6: Professional archival application format*

[3] ISO/IEC 23000-10, *Information technology — Multimedia application format (MPEG-A) — Part 10: Surveillance application format*

[4] ISO/IEC 23009 (all parts), *Information technology — Dynamic adaptive streaming over HTTP (DASH)*