

DRAFT AMENDMENT

ISO/IEC 14496-12:2020/DAM 1

ISO/IEC JTC 1/SC 29

Secretariat: JISC

Voting begins on:
2021-01-20

Voting terminates on:
2021-04-14

Information technology — Coding of audio-visual objects —

Part 12: ISO base media file format

AMENDMENT 1: Support for new media types (haptics, volumetric visual) and other improvements

Partie 12: Format ISO de base pour les fichiers médias

AMENDEMENT 1: Titre manqué

ITeH STANDARD PREVIEW
(standards.iteh.ai)

ICS: 35.040.40

[ISO/IEC 14496-12/DAMd 1](#)

<https://standards.iteh.ai/catalog/standards/sist/e58781ae-49d9-400c-b526-0398b3ed598f/iso-iec-14496-12-damd-1>

THIS DOCUMENT IS A DRAFT CIRCULATED FOR COMMENT AND APPROVAL. IT IS THEREFORE SUBJECT TO CHANGE AND MAY NOT BE REFERRED TO AS AN INTERNATIONAL STANDARD UNTIL PUBLISHED AS SUCH.

IN ADDITION TO THEIR EVALUATION AS BEING ACCEPTABLE FOR INDUSTRIAL, TECHNOLOGICAL, COMMERCIAL AND USER PURPOSES, DRAFT INTERNATIONAL STANDARDS MAY ON OCCASION HAVE TO BE CONSIDERED IN THE LIGHT OF THEIR POTENTIAL TO BECOME STANDARDS TO WHICH REFERENCE MAY BE MADE IN NATIONAL REGULATIONS.

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.

This document is circulated as received from the committee secretariat.



Reference number
ISO/IEC 14496-12:2020/DAM 1:2021(E)

© ISO/IEC 2021

iTeh STANDARD PREVIEW (standards.iteh.ai)

[ISO/IEC 14496-12/DAMd 1](https://standards.iteh.ai/catalog/standards/sist/e58781ae-49d9-400c-b526-0398b3ed598f/iso-iec-14496-12-damd-1)
<https://standards.iteh.ai/catalog/standards/sist/e58781ae-49d9-400c-b526-0398b3ed598f/iso-iec-14496-12-damd-1>



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2021

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

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

Published in Switzerland

Foreword

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

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents) or the IEC list of patent declarations received (see patents.iec.ch).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

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

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.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

[ISO/IEC 14496-12/DAmD 1](https://standards.iteh.ai/catalog/standards/sist/e58781ae-49d9-400c-b526-0398b3ed598f/iso-iec-14496-12-damd-1)

<https://standards.iteh.ai/catalog/standards/sist/e58781ae-49d9-400c-b526-0398b3ed598f/iso-iec-14496-12-damd-1>

Information technology — Coding of audio-visual objects —

Part 12: ISO base media file format

AMENDMENT 1: Support for new media types (haptics, volumetric visual) and other improvements

In clause 3, add the following definitions, renumbering as needed

3.1.7

haptic media

timed tactile signals to be presented as part of the media presentation

3.1.31

volumetric visual media

timed visual media defining a visual coding in a three dimensional space

Note 1 to entry: In contrast to video media, which defines a planar coding.

[ISO/IEC 14496-12/DAmD 1](https://standards.iteh.ai/catalog/standards/sist/e58781ae-49d9-400c-b526-05963cd99878/iso-iec-14496-12-damd-1)

<https://standards.iteh.ai/catalog/standards/sist/e58781ae-49d9-400c-b526-05963cd99878/iso-iec-14496-12-damd-1>
in 8.3.3 insert after the bullet for `cdsc`, and number the notes

- `'cdtg'` describes the referenced media tracks and track groups collectively; the `'cdtg'` track reference shall only be present in timed metadata tracks.

NOTE A timed metadata track containing `'cdsc'` track reference to a `track_group_id` value describes each track in the track group individually.

in 8.3.3 add at the end

- `'shsc'` links a shadow sync track to a main track; see clause 9.5

in 8.5.2.3 replace

`maxBitrate` gives the maximum rate in bits/second over any window of one second.

`avgBitrate` gives the average rate in bits/second over the entire presentation.

with

`maxBitrate` gives the maximum rate in bits/second over any window of one second; this is a measured value for stored content, or a value that a stream is configured not to exceed; the stream shall not exceed this bitrate.

`avgBitrate` gives the average rate in bits/second of the stream; this is a measured value (cumulative over the entire presentation) for stored content, or the configured target average bitrate for a stream.

in 8.6.14.3 change

`range_size` indicates the size of the partial subsegment.

to

`range_size` indicates the size of the subsegment; the value 0 may be used in the last entry to indicate the remaining bytes of the segment, to the end of the segment

in 8.7.8.1 replace

If `aux_info_type` and `aux_info_type_parameter` are omitted then the implied value of `aux_info_type` is either (a) in the case of transformed content, such as protected content, the `scheme_type` included in the `ProtectionSchemeInfoBox` or otherwise (b) the sample entry type.

with

If `aux_info_type` and `aux_info_type_parameter` are omitted then the implied value of `aux_info_type` is either (a) in the case of transformed content, such as protected content, the `scheme_type` included in the `ProtectionSchemeInfoBox` or `ScrambleSchemeInfoBox`, or otherwise (b) the sample entry type. In the case of tracks containing multiple transformations, `aux_info_type` and `aux_info_type_parameter` shall not be omitted.

iTeh STANDARD PREVIEW (standards.iteh.ai)

in 8.8.13.1 replace

For example, `padding_flag` can be set equal to 1 when the following conditions are true:

- Each fraction contains two or more AVC, SVC, or MVC [ISO/IEC 14496-15^[16]] tracks of the same video bitstream.
- The samples for each track of a fraction are contiguous and in decoding order in a `MediaDataBox`.
- The samples of the first AVC, SVC, or MVC level contain extractor NAL units for including the video coding NAL units from the other levels of the same fraction.

with

The use of `padding_flag` is deprecated.

In 8.9.3.1, after the paragraph

There may be multiple instances of this box if there is more than one sample grouping for the samples in a track. Each instance of the `SampleGroupDescriptionBox` has a `type` that distinguishes different sample groupings. There shall be at most one instance of this box with a particular `grouping_type` in a track (i.e. defined in a `SampleTableBox` or `TrackFragmentBox`).

insert

The `flags` field of the `SampleGroupDescriptionBox` shall be zero when the box is in a `TrackFragmentBox`. When the box is in a `SampleTableBox`, either or both of the two lowest bits may be set:

- `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.

Note 1 the `static_mapping` flag is only useful when `default_group_description_index` is non-zero, since the default value of `default_group_description_index` is 0, indicating no mapping.

These flags may be used in combination with the version of the `SampleGroupDescriptionBox` to signal various possibilities.

- `static_group_description` without `static_mapping`:
the sample group definitions are only in the `MovieBox`, but samples can map to any of them.
- `static_mapping` without `static_group_description`:
everything in a fragment maps to at most one group; there may be new `SampleGroupDescriptionBoxes` of this type in fragments; depending on their version, the `SampleGroupDescriptionBoxes` can identify a default sample group, or that samples are unmapped.
- both `static_group_description` and `static_mapping`:
every sample maps to the default indicated in the `SampleGroupDescriptionBox` in the `MovieBox`; that `SampleGroupDescriptionBox` can indicate a default sample group or indicate that all samples are unmapped, depending on its version.

and delete the sentence

ITeh STANDARD PREVIEW

For video tracks, an abstract `VisualSampleGroupEntry` is used with similar types for audio and hint tracks.

and re-name the existing Note as Note 2.

[ISO/IEC 14496-12/DAMd 1](https://standards.iteh.ai/catalog/standards/sist/e58781ae-49d9-400c-b526-0398b3ed598f/iso-iec-14496-12-damd-1)

<https://standards.iteh.ai/catalog/standards/sist/e58781ae-49d9-400c-b526-0398b3ed598f/iso-iec-14496-12-damd-1>

In 8.9.3.2 add

```
abstract class HapticSampleGroupEntry (unsigned int(32) grouping_type)
extends SampleGroupDescriptionEntry (grouping_type)
{
}
abstract class VolumetricVisualSampleGroupEntry (unsigned int(32)
grouping_type) extends SampleGroupDescriptionEntry (grouping_type)
{
}
```

and change the definition of `SampleGroupDescriptionBox` to:

```
aligned(8) class SampleGroupDescriptionBox (unsigned int(32) handler_type)
extends FullBox('sgpd', version, flags){
unsigned int(32) grouping_type;
if (version>=1) { unsigned int(32) default_length; }
if (version>=2) {
unsigned int(32) default_group_description_index;
}
unsigned int(32) entry_count;
int i;
for (i = 1 ; i <= entry_count ; i++){
if (version>=1) {
if (default_length==0) {
unsigned int(32) description_length;
}
}
SampleGroupDescriptionEntry (grouping_type);
// an instance of a class derived from SampleGroupDescriptionEntry
// that is appropriate and permitted for the media type
}
```

}

in 8.12.1 add the following rows to Table 5

Haptics	encp	HapticSampleEntry
Volumetric visual	enc3	VolumetricVisualSampleEntry

in 8.17.1, add the following rows to the table

Haptics	icpp
Volumetric visual	icp3

add the following as clause 9.5

9.5 Shadow sync support

There are two forms of support for shadow sync; the sample table box in clause 8.6.3, and the use of the shadow sync track reference type in 8.3.3.

A track containing an 'shsc' track reference is called a shadow sync sample track, and the tracks pointed to by the 'shsc' track reference are called main tracks.

The shadow sync sample track provides an optional set of sync samples that can be used when seeking to a position or for similar operations performed to any of the associated main tracks.

When an 'shsc' track reference is present, the following constraints shall be obeyed:

- All samples of the shadow sync sample track shall be sync samples.
- Each main track shall have a sample that is aligned in decoding time with each sample of the shadow sync sample track.
- A concatenation of the following samples in the following order shall conform to the sample entry of the main track:
 - o Any selected sample of the shadow sync sample track, with the sample duration of the sample of the main track that is aligned in decoding time with the selected sample of the shadow sync sample track.
 - o Samples of the main track following the sample of the main track that is aligned in decoding time with the selected sample of the shadow sync sample track.

An 'shsc' track reference indicates that the decoded samples resulting from the concatenation specified above have acceptable quality for playback.

NOTE 1 The samples in the main track that are aligned in decoding time with the samples in the shadow sync sample track are "switchable" samples that are constrained so that no samples preceding a "switchable" sample in decoding order are used as a prediction reference for any sample following the "switchable" sample in decoding order.

Add the following as clause 10.9 and 10.10

10.9 Pixel Aspect Ratio Sample Grouping

10.9.1 Definition

The Pixel Aspect Ratio sample group ('pasr') may be used to signal the pixel aspect ratio of samples in a video track, when the pixel aspect ratio of the samples within a track change dynamically and a single value in a `PixelAspectRatioBox` in a sample entry, specified in 12.1.4 cannot therefore be used.

When the Pixel Aspect Ratio sample group is used in a track, the `PixelAspectRatioBox` shall not be present in any sample entry of that track.

10.9.2 Syntax

```
class PixelAspectRatioEntry() extends VisualSampleGroupEntry ('pasr'){
    unsigned int(32) hSpacing;
    unsigned int(32) vSpacing;
}
```

10.9.3 Semantics

`hSpacing`, `vSpacing`: define the relative width and height of a pixel as defined for the `PixelAspectRatioBox` in 12.1.4

10.10 Clean Aperture Sample Grouping

10.10.1 Definition iTeh STANDARD PREVIEW

The Clean Aperture sample group ('casg') may be used to signal the clean aperture of samples in a video track, when the clean aperture of the samples within a track change dynamically and a single value in a `CleanApertureBox` in a sample entry, specified in 12.1.4 cannot therefore be used.

When the Clean Aperture sample group is used in a track, the `CleanApertureBox` shall not be present in any sample entry of that track.

10.10.2 Syntax

```
class CleanApertureEntry() extends VisualSampleGroupEntry ('casg'){
    unsigned int(32) cleanApertureWidthN;
    unsigned int(32) cleanApertureWidthD;

    unsigned int(32) cleanApertureHeightN;
    unsigned int(32) cleanApertureHeightD;

    unsigned int(32) horizOffN;
    unsigned int(32) horizOffD;

    unsigned int(32) vertOffN;
    unsigned int(32) vertOffD;
}
```

10.10.3 Semantics

`cleanApertureWidthN`, `cleanApertureWidthD`, `cleanApertureHeightN`, `cleanApertureHeightD`, `horizOffN`, `horizOffD`, `vertOffN` and `vertOffD` define the clean aperture width, height and horizontal and vertical offsets of the clean aperture center as defined for the `CleanApertureBox` in 12.1.4