FINAL
DRAFT

INTERNATIONAL
STANDARD

**ISO/IEC
FDIS
23001-16**

ISO/IEC JTC **1**/SC **29**

Secretariat: JISC

Voting begins on:
2021-09-03

Voting terminates on:
2021-10-29

Information technology — MPEG systems technologies —

Part 16: Derived visual tracks in the ISO base media file format

iTeh STANDARD PREVIEW
(standards.iteh.ai)

ISO/IEC FDIS 23001-16
https://standards.iteh.ai/catalog/standards/sist/12b76749-56ee-47e0-ac45-02bd57c6dc99/iso-iec-fdis-23001-16

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.

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.

Reference number
ISO/IEC FDIS 23001-16:2021(E)

© ISO/IEC 2021

iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/IEC FDIS 23001-16
https://standards.iteh.ai/catalog/standards/sist/12b76749-56ee-47e0-ac45-02bd57c6dc99/iso-iec-fdis-23001-16

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

© ISO/IEC 2021 – All rights reserved

Contents

Page

iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/IEC FDIS 23001-16
https://standards.iteh.ai/catalog/standards/sist/12b76749-56cc-47e0-ac45-02bd57c6dc99/iso-iec-fdis-23001-16

© ISO/IEC 2021 – All rights reserved

iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/IEC FDIS 23001-16
https://standards.iteh.ai/catalog/standards/sist/12b76749-56ee-47e0-ac45-02bd57c6dc99/iso-iec-fdis-23001-16

© ISO/IEC 2021 – All rights reserved

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).

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.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

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.

ISO/IEC FDIS 23001-16
https://standards.iteh.ai/catalog/standards/sist/12b76749-56ce-47e0-ae45-02bd57c6dc99/iso-iec-fdis-23001-16

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 23001 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.

© ISO/IEC 2021 – All rights reserved

Introduction

Derived visual tracks are designed to enable defining a timed sequence of visual transformation operations to be applied to input still images and/or samples of timed sequences of images in the same presentation. It is built using tools defined in the ISO base media file format (ISO/IEC 14496-12). This document specifies the core design and an initial base set of transformation operations.

The International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC) draw attention to the fact that it is claimed that compliance with this document may involve the use of a patent.

ISO and IEC take no position concerning the evidence, validity and scope of this patent right.

The holder of this patent right has assured ISO and IEC that he/she is willing to negotiate licences under reasonable and non-discriminatory terms and conditions with applicants throughout the world. In this respect, the statement of the holder of this patent right is registered with ISO and IEC. Information may be obtained from the patent database available at www.iso.org/patents.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights other than those in the patent database. ISO [and/or] IEC shall not be held responsible for identifying any or all such patent rights.

iTeh STANDARD PREVIEW
(standards.iteh.ai)
ISO/IEC FDIS 23001-16
https://standards.iteh.ai/catalog/standards/sist/12b76749-56ee-47e0-ac45-02bd57c6dc99/iso-iec-fdis-23001-16

© ISO/IEC 2021 – All rights reserved

Information technology — MPEG systems technologies —

Part 16:
Derived visual tracks in the ISO base media file format

1 Scope

This document defines a storage format for derived visual tracks and an initial base set of related transformation operations. The format defined in this document enables the interchange, editing, and display of timed sequences of images that result from transformation operations applied to input still images or samples of timed sequences of images in the same presentation.

This format defines normative structures used to contain the description of transformation operations, how to link that transformation operations to the inputs, and defines how to process those transformation operations to obtain a timed sequence of video frames.

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.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

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

ISO/IEC FDIS 23001-16
https://standards.iteh.ai/catalog/standards/sist/12b76749-56ee-47e0-ac45-

ISO/IEC 23001-10, *Information technology — MPEG systems technologies — Part 10: Carriage of timed metadata metrics of media 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 and the following apply.

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

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

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

3.1
derivation operation
container box representing an operation applying a *derivation transformation* (3.2) on an ordered list of *inputs* (3.5)

3.2
derivation transformation
visual transformation operation identified by a 32-bit value and a set of parameters that transforms *inputs* (3.5) into *visual outputs* (3.8)

Note 1 to entry: The 32-bit value is also known as a four-character code in ISO/IEC 14496-12.

© ISO/IEC 2021 – All rights reserved

3.3
derived sample
sample containing an ordered list of *derivation operations* (3.1)

3.4
derived visual track
video or picture track that contains a timed sequence of *derived samples* (3.3)

3.5
input
parameter input (3.6) or *visual input* (3.7)

3.6
parameter input
metadata from an input item or track that is used as input for a *derivation transformation* (3.2) of a *derivation operation* (3.1)

Note 1 to entry: The parameter input is either an input metadata item from file-level `MetaBox` or an interval of an input metadata track (possibly spanning multiple samples).

3.7
visual input
video or still image that is used as input for a *derivation transformation* (3.2) of a *derivation operation* (3.1)

Note 1 to entry: The visual input is either an input image item from file-level `MetaBox`, an interval of an input track (possibly spanning multiple samples), the visual output of a preceding *derivation operation* (3.1) or the default input fill picture signalled in the configuration record of the *derived visual track* (3.4).

iTeh STANDARD PREVIEW
(standards.iteh.ai)

3.8
visual output
one video frame or a sequence of video frames that is output from a *derivation transformation* (3.2) of a *derivation operation* (3.1)

ISO/IEC FDIS 23001-16
https://standards.iteh.ai/catalog/standards/sist/12678749-56ee-47e0-ac45-02bd57c6dc99/iso-iec-fdis-23001-16

4 Derived visual tracks, design principles

A derived visual track describes a timed sequence of derived samples composed of an ordered list of derivation operations, each derivation operation applying a derivation transformation for the duration of the derived sample on an ordered list of inputs represented in the same presentation.

A derived visual track shall be either a video track (with the `'vide'` handler type in the `HandlerBox` of the `MediaBox` as defined in ISO/IEC 14496-12) or a picture track (with the `'pict'` handler type in the `HandlerBox` of the `MediaBox` as defined in ISO/IEC 23008-12). A derived visual track is identified by its containing sample entry of type `'dtrk'` `DerivedVisualSampleEntry`. Each sample described by a `DerivedVisualSampleEntry` is a derived sample.

A derived visual track shall include a `TrackReferenceTypeBox` with `reference_type` equal to `'dtrk'` referring to all the inputs. Each reference shall be one of:

a) the `track_ID` of a track used by derived samples in the track, or, if unified IDs are in use as defined by ISO/IEC 14496-12, a `track_group_id`;

b) the `item_ID` of an image item, in the file-level `MetaBox`, used by derived samples in the track.

An ID value in the track references is resolved to a `track_ID` whenever the file contains a track with such ID, is resolved to a `track_group_id` whenever unified IDs are in use and the file contains a track group with such ID, and is resolved to an `item_ID` otherwise.

NOTE 1 A `track_ID` can be an ID of a derived visual track.

 © ISO/IEC 2021 – All rights reserved

If a referenced track is a member of an alternate group or switch group, or if the reference is to a track group, then the reader should pick a track from the group as the input to the derived visual track.

NOTE 2 The `TrackSelectionBox` can be used to provide guidance on the selection between members of an alternate group or switch group.

Similarly, if a referenced image item is a member of an alternate group (which may contain both tracks and images), then the reader should pick one member of the group as the input to the derived visual track.

A derived sample contains an ordered list of the derivation operations to be performed, each derivation operation applying a derivation transformation on an ordered list of inputs. The `layer` syntax element in `TrackHeaderBox` has no impact on ordering the inputs for derived samples.

The four-character codes of derivation transformation from all derivation operations used by the derived samples in the track are listed in the `DerivedVisualSampleEntry`, and also default inputs and parameter values can be supplied there. A derived sample in the track may use all or some of the derivation operations listed in the linked `DerivedVisualSampleEntry`, but derived samples shall not use a derivation operation not listed in the sample entry.

The derived sample durations document the time over which the derivation represented by the ordered list of derivation operations is active. Therefore, the number of samples defined in a derived visual track does not necessarily match 1:1 with the number of input image items or samples of input tracks that are being transformed. A single derivation duration may span multiple samples in the source track(s), and also derivation transformations in derived samples may have 'internal time structure' (e.g. a cross-fade) so the picture may change during the sample duration. This is in contrast to 'classic video'.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

Derived visual tracks do not respect edit lists on inputs. They operate on the composition timeline (i.e. before the application of edit lists) of their input tracks (including on derived visual tracks when used as visual inputs). However, the input tracks shall not have edit lists. Any edit lists of the input tracks shall be ignored if present.

ISO/IEC FDIS 23001-16
https://standards.iteh.ai/catalog/standards/sist/12b76749-56ee-47e0-ac45-[illegible]

NOTE 3 When time-alignment adjustment between input tracks is needed, signed composition offsets in input tracks can be used.

NOTE 4 A derived visual track can have an edit list; thus, a derived visual track using the identity transform, and with an edit-list, can provide a visual output that is a temporal re-mapping of the input track.

The inputs for a derivation operation in a derived sample can be either input image items from file-level `MetaBox` or intervals (possibly spanning multiple samples) of input video tracks, image sequence tracks, metadata items or metadata tracks, the visual output of a preceding derivation operation or a default input fill picture.

Transformative item properties or transformations (e.g. clean aperture, track matrix etc...) associated with input image items or samples of input tracks are always applied before performing the derivation operation.

NOTE 5 If a derived sample needs to refer to one explicit sample value in a referred track (other than the time-aligned sample value), an item can be created and referred to that has the same data as the desired sample value.

The visual inputs in a derived sample shall have consistent pixel aspect ratio and bit depth. The input image items, samples of input tracks or derived samples may have various width and height. When differences in width and height result in pixels that never get 'painted' by a derivation operation, those empty pixels are filled according to the value of `default_derivation_input` parameter signalled in `DerivedVisualTrackConfigRecord` (black, white or grey pixels). When differences in width and height result in pixels that end up outside the visual output size by a derivation operation, those pixels are cropped. This default behaviour may be overridden by derivation operation specifications.

A derived sample is reconstructed by performing the specified derivation operations in sequence. Some derivation operations can be marked as non-essential which indicates that the derivation operation may be skipped by the reader. However, the operations marked as essential shall be used in order to obtain a valid derived sample.

© ISO/IEC 2021 – All rights reserved

When more than one derivation operation is listed in a derived sample, the derivation operation that is not first in the list may include the output result (e.g. the visual output) of any of the previous derivation operations, only new inputs, or a combination of both.

In many cases the source tracks pointed to by the `'dtrk'` track reference are not intended for display. When a track is not intended for display, `track_in_movie` shall be equal to 0 for that track.

The visual output of a derived sample is the output from the last derivation operation in the sample. If there is no derivation operation, an empty derived sample (i.e. sample size of 0) is equivalent to an empty edit, i.e., there is no visual output from the derived visual track at that time.

Using derived visual tracks, it is possible to build either a chain of derivation operations on one single derived visual track or a hierarchy of multiple derived visual tracks when they are used as a visual input to another derived visual track. The latter should only be used when each derived visual track in the hierarchy is also needed on its own.

5 Derivation operation

5.1 Definition

Box Type: `'dimg'`
Container: derived sample or `DerivedVisualTrackConfigRecord` in a `DerivedVisualSampleEntry`
Mandatory: Yes, in a `DerivedVisualTrackConfigRecord`, and No in a derived sample
Quantity: At least one in a `DerivedVisualTrackConfigRecord`, and Zero or more in a derived sample

iTeh STANDARD PREVIEW
(standards.iteh.ai)

A derivation operation in either a derived sample entry or derived sample is represented by a container box of type `'dimg'` that always carries a derivation transformation box inherited from VisualDerivationBase, and can carry a `VisualDerivationInputs` providing the inputs for the derivation transformation.

ISO/IEC FDIS 23001-16
https://standards.iteh.ai/catalog/standards/sist/12b7b749-56cd-47c0-ad45-02bd57c6dc99/iso-iec-fdis-23001-16

A derivation transformation in a derivation operation is identified by a 32-bit value, also known as a four-character code in ISO/IEC 14496-12, unless that code is `'uuid'`, whereupon a UUID identifies a vendor-specific derivation transformation.

A derivation transformation's parameters shall

a) be single, countable

b) have defined default values in the specification

For both inputs and parameters, there is a bit-mask in the sample entry and a bit-mask in the derived sample that uses the derivation operation, defining whether the parameter value or input is supplied there.

Each parameter takes the value defined in that derived sample, if any, or else the value defined in the sample entry, if any, or else the default value for that parameter defined in the derivation transformation specification.

Each input takes, in precedence order:

a) the input listed in the derived sample if present;

b) the default input listed in the corresponding derivation operation in the configuration record in the sample entry, if present;

c) or the default input fill picture defined in the configuration record for the derived visual track.

 © ISO/IEC 2021 – All rights reserved

Each input can be:

a) an index into the track reference box for the derived visual track (itself indicating either an input image item from the file-level `MetaBox`, an input track, or, if unified IDs are in use as defined by ISO/IEC 14496-12, a track group);

b) a relative index of a previous derivation operation in the same derived sample containing the derivation operation that uses it;

NOTE 1 A relative index declared in the configuration record is pointing to a derivation operation within the derived sample that uses it.

c) or the default input fill picture defined in the configuration record for this derived visual track.

The default input fill picture is signalled in `VisualDerivationDefault` in the `DerivedVisualTrackConfigRecord` in the derived visual track sample entry (e.g. either a full black, mid grey or full white picture).

NOTE 2 The size of the default input fill picture is given by `width` and `height` in the sample entry.

If no inputs are defined in a derivation operation listed in the sample entry, `VisualDerivationInputs` may be absent there; if none are defined in a derivation operation listed in the derived sample (presumably at least some are defined in the sample entry) then `VisualDerivationInputs` may be absent there.

Bits in the masks are assigned from least-significant (first input or parameter) upwards. Their default value is 0 (i.e. if the operation has 9 inputs and only 8 bits are supplied, the 9th bit is assumed to be zero).

iTeh STANDARD PREVIEW
(standards.iteh.ai)

The version of `VisualDerivationBase` is currently constrained to be 0; only one flag is defined, the low-order bit. When set to 1, it indicates that the operation is essential. If any derived sample sets this but for a given derivation transformation type (code), then the corresponding derivation operation listed in the sample entry shall set this bit. A parser shall not process a derived visual track that contains a derivation operation marked in the sample entry as essential that is not recognized or not supported by the parser. If a non-essential derivation operation is not supported, the derived sample containing it may be processed as if this non-essential derivation operation was a null derivation operation, i.e. the visual output of a non-essential and unsupported derivation operation is the visual output from the previous derivation operation in the sample. If there is no previous derivation operation, the visual output of the unsupported derivation operation is the default input fill picture defined in the configuration record for the derived visual track.

ISO/IEC FDIS 23001-16
https://standards.iteh.ai/catalog/standards/sist/12b76749-5bee-47e0-ac45-02bd57e0de99/iso-iec-fdis-23001-16

5.2 Syntax

```
aligned(8) class VisualDerivationBase
extends FullBox (code, version = 0, flags, optional unsigned int(8)[16] uuid_code){
   // the box may be empty and terminate before the following field;
   unsigned int(16) highest_param_idx;
   unsigned int(floor((highest_param_idx+7)/8))*8) parameter_defined_flags;
   // the remaining bytes are the values of parameters signalled by parameter_defined_
flags;
}

aligned(8) class VisualDerivationInputs
extends FullBox ('dinp', version = 0, flags = 0){
   unsigned int(16) highest_input_idx;
   unsigned int(floor((highest_input_idx+7)/8))*8) input_present_flags;
   unsigned int(16) reference_index[];
}

aligned(8) class VisualDerivation extends Box ('dimg'){
   VisualDerivationBase()   derivation_transformation;    // actually an instance of a
derived class
   VisualDerivationInputs   inputs;    // optional
}
```

© ISO/IEC 2021 – All rights reserved