**ISO #####-#:####(X)**

ISO/TC ###/SC ##/WG #

Date: YYYY-MM-DD

**Carriage of depth and alpha** (Introductory element — Main element — Part #: Part title)

At this point this document is a placeholder WD and parts of its content may be later moved to different amendments.

WD 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.

*A model document of an International Standard (the Model International Standard) is available at:*[*https://www.iso.org/drafting-standards.html*](https://www.iso.org/drafting-standards.html)

© ISO 20XX

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

Contents

*This template allows you to work with default MS Word functions and styles. You can use these if you want to maintain the Table of Contents automatically and apply auto-numbering.*

*To update the Table of Contents please select it and press "F9".*

[Foreword iv](#_Toc181692558)

[Introduction v](#_Toc181692559)

[1 Scope 1](#_Toc181692560)

[2 Normative references *(mandatory)* 1](#_Toc181692561)

[3 Terms and definitions *(mandatory)* 1](#_Toc181692562)

[4 Example application scenarios 2](#_Toc181692563)

[4.1 File-based consumption (informative) 2](#_Toc181692564)

[4.2 Segment-based delivery (informative) 2](#_Toc181692565)

[5 Concepts 2](#_Toc181692566)

[5.1 Depth elementary stream 2](#_Toc181692567)

[5.2 Alpha elementary stream 2](#_Toc181692568)

[6 Elementary stream constraints 3](#_Toc181692569)

[7 Encapsulation in ISOBMFF 3](#_Toc181692570)

[7.1 Sample Definitions 3](#_Toc181692571)

[7.2 Depth Sample entry 3](#_Toc181692572)

[7.2.1 General 3](#_Toc181692573)

[7.2.2 Depth information box 3](#_Toc181692574)

[7.2.3 Syntax 3](#_Toc181692575)

[7.2.4 Semantics 3](#_Toc181692576)

[7.3 Alpha Sample entry 4](#_Toc181692577)

[7.3.1 General 4](#_Toc181692578)

[7.3.2 Alpha information box 4](#_Toc181692579)

[7.3.3 Syntax 4](#_Toc181692580)

[7.3.4 Semantics 4](#_Toc181692581)

[7.4 Track definitions 5](#_Toc181692582)

[7.4.1 Depth track definition 5](#_Toc181692583)

[7.5 Auxiliary video track definition 5](#_Toc181692584)

[7.5.1 General 5](#_Toc181692585)

[7.5.2 Auxiliary video type box 5](#_Toc181692586)

[7.6 Track reference 6](#_Toc181692587)

[8 Brands 6](#_Toc181692588)

[9 Integration with delivery format 6](#_Toc181692589)

[Bibliography 7](#_Toc181692590)

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

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)).

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

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

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

This document was prepared by Technical Committee *[or Project Committee]* ISO/TC *[or ISO/PC]* ###, *[name of committee]*, Subcommittee SC ##, *[name of subcommittee]*.

This second/third/… edition cancels and replaces the first/second/… edition (ISO #####:####), which has been technically revised.

The main changes are as follows:

— xxx xxxxxxx xxx xxxx

A list of all parts in the ISO ##### 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](https://www.iso.org/members.html).

Introduction

This document specifies the carriage of depth and alpha sequences based on the ISO Base Media File Format (ISBOMFF) to enable interoperability points for applications producing and consuming content with a depth and alpha component.

To define those interoperability points, this document defines specific constraints on the ISOBMFF structure as well as on the elementary streams used to represent the depth and alpha sequences. In addition, the metadata required to interpret those depth and alpha sequences are also specified by this document.

Carriage of depth and alpha (Introductory element — Main element — Part #: Part title)

# Scope

[Editor’s note: At this point this document is a placeholder WD and parts of its content may be later moved to different amendments.]

This document specifies the carriage of depth and alpha sequences along with their metadata in an ISOBMFF structure. The rendering of the content is considered to be application-specific and thus out-of-scope of this document.

This document defines the following:

* TBD

# Normative references *(mandatory)*

*Two options of text (remove the inappropriate option).*

*1) The normative references shall be introduced by the following wording.*

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 #####‑#, *General title — Part #: Title of part*

ISO #####‑##:20##, *General title — Part ##: Title of part*

*2) If no references exist, include the following phrase below the clause title:*

There are no normative references in this document.

# Terms and definitions *(mandatory)*

*Four options of text (remove the inappropriate options).*

*1) If all the specific terms and definitions are provided in Clause 3, use the following introductory text:*

For the purposes of this document, the following terms and definitions apply.

*2) If reference is given to an external document, use the following introductory text:*

For the purposes of this document, the terms and definitions given in [external document reference xxx] apply.

*3) If terms and definitions are provided in Clause 3, in addition to a reference to an external document, use the following introductory text:*

For the purposes of this document, the terms and definitions given in [external document reference xxx] and the following apply.

*4) If there are no terms and definitions provided, use the following introductory text:*

No terms and definitions are listed in this document.

*The text below is always included after each option:*

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

* ISO Online browsing platform: available at <https://www.iso.org/obp>
* IEC Electropedia: available at <https://www.electropedia.org/>

3.1

term

text of the definition

Note 1 to entry: Text of the note.

[SOURCE: …]

3.2

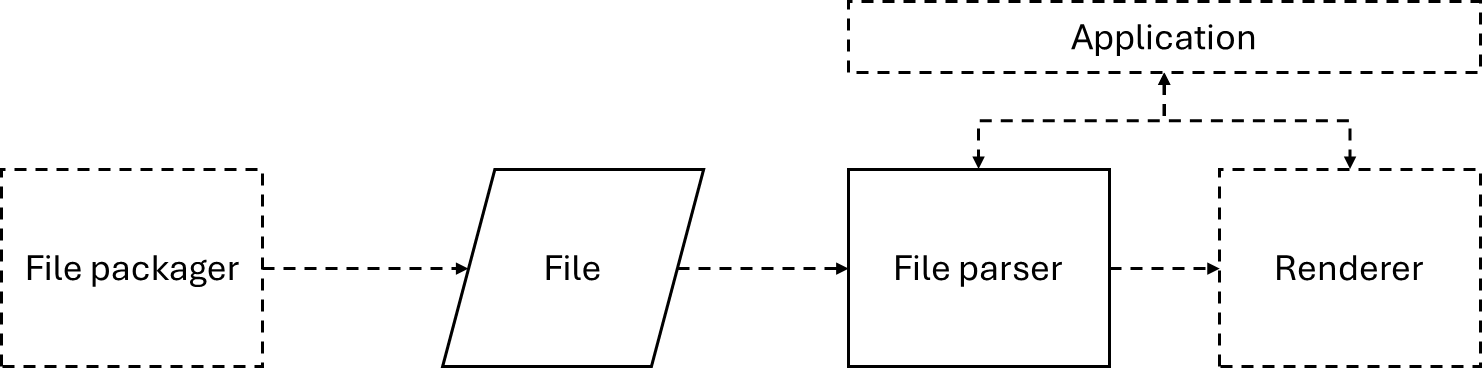
term

text of the definition

# Example application scenarios

## File-based consumption (informative)

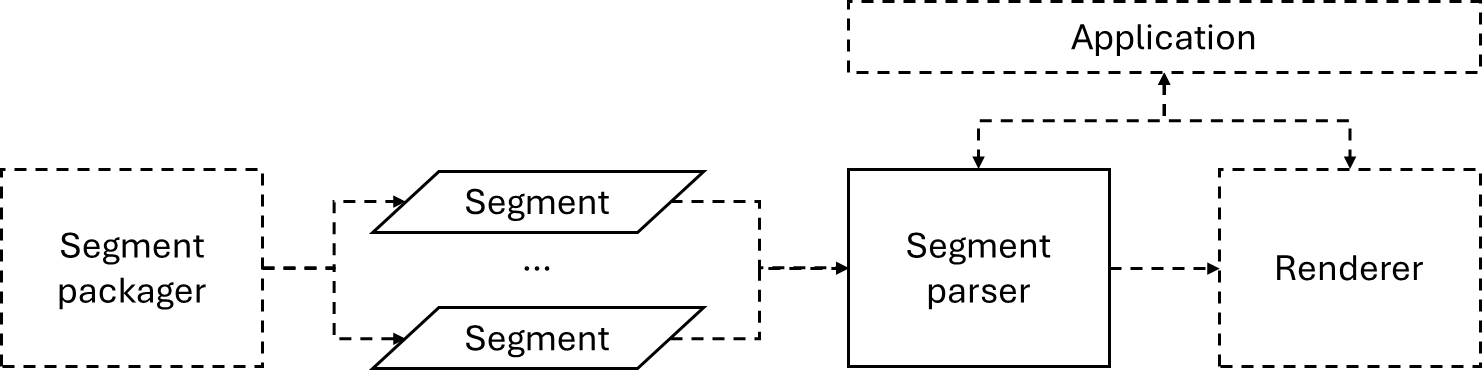
An application may consume a file compliant to this document.



1. File-based application scenario

## Segment-based delivery (informative)

An application may consume segment containing content compliant to this document.



1. Segment-based application scenario

# Concepts

## Depth elementary stream

An depth elementary stream is:

* TODO

## Alpha elementary stream

An alpha elementary stream is:

* TODO

# Elementary stream constraints

TDB

# Encapsulation in ISOBMFF

## Sample definitions

[Editor’s note: This document is a placeholder for this clause but some of this clause may be converted later into an amendments of corresponding specifications and replaced with mere references to them.]

## Depth samples

### Depth sample entry

[Editor’s note: The definition of a depth media handler is to be further studied]

### Depth information box

Box Type: 'depi'   
Container: VisualSampleEntry  
Mandatory: No  
Quantity: One

The DepthInformationBox may be used to provide information independent of the coding, to interpret the depth data.

When the DepthInformationBox is present, the decoded sample values are uniformly quantized into the range [0, maxVal], with maxVal = 2^bit\_depth – 1.

### Syntax

class DepthInformationBox extends FullBox ('depi', version = 0, flags){  
 float(32) near\_plane;  
 float(32) far\_plane;  
 unsigned int(16) focal\_plane;  
 unsigned int(6) units;  
}

[Editor’s Note: The optionality of some of those parameters is to be studied]

### Semantics

version is an integer that specifies the version of this box.

flags is a 24-bit integer with flags; the following values are defined:

is\_normalised: Flag mask is 0x000001. Specifies whether the depth values are normalized in the range 0 to 1.

is\_reverse: Flag mask is 0x000002. Specifies whether the mapping of the minimum and the maximum decoded sample values to the near\_plane and the far\_plane is reversed. When is\_reverse is equal to 0, the decoded sample values equal to 0 correspond to the near\_plane value and the decoded sample values equal to maxVal correspond to the far\_plane value. When is\_reverse is equal to 1, the decoded sample values equal to 0 correspond to the far\_plane value and the decoded sample values equal to maxVal correspond to the near\_plane value.

is\_inverse: Flag mask is 0x000003. Specifies whether there is an inverse relationship between the decoded sample values and the depth values.   
When is\_inverse is equal to 0 and is\_reverse equals to 0, the decoded sample value equal to 0 correspond to the far\_plane value and the decoded sample values equal to maxVal correspond to the near\_plane value.  
When is\_inverse is equal to 0 and is\_reverse equals to 1, the decoded sample value equal to 0 correspond to the near\_plane value and the decoded sample values equal to maxVal correspond to the far\_plane value.  
When is\_inverse is equal to 1 and is\_reverse equals to 0, the decoded sample values equal to 0 correspond to the 1/far\_plane value and the decoded sample values equal to maxVal correspond to the 1/near\_plane value.  
When is\_inverse is equal to 1 and is\_reverse equals to 1, the decoded sample values equal to 0 correspond to the value 1/near\_plane and the decoded sample values equal to maxVal correspond to the 1/far\_plane value.

[Editor’s Note: The parameter bit\_depth is to be properly specified]

[Editor’s Note: Editorial alignment is required between the definitions of is\_inverse and is\_reverse.]

[Editor’s Note: An informative section containing the equations to retrieve the depth values from decoded sample values and related figures, could be added.]

near\_plane and far\_plane specify the nearest and the farthest depth values, respectively.

focal\_plane specifies the focal plane of the depth values. units specifies the units of the depth values, as follows:

0: unspecified

1: the values are in meters

2: the values are in millimetres

3-63: reserved.

## Alpha samples

### General

### Alpha information box

Box Type: 'alpi'   
Container: VisualSampleEntry  
Mandatory: No  
Quantity: One

The AlphaInformationBox may be used to provide information independent of the coding, to interpret the alpha data.

The AlphaInformationBox is optional and if it is absent it is assumed that a value of 0 indicates full transparency and a value equal to 2^bit\_depth - 1 indicates full opacity.

[Editor’s Note: The parameter bit\_depth is to be properly specified]

### Syntax

class AlphaInformationBox extends FullBox ('alpi', version = 0, flags){  
  
 unsigned int(16) opaque\_value;  
 unsigned int(16) transparent\_value;  
 unsigned int(6) reserved;  
}

[Editor’s Note: possible alignment of the semantics with other definitions of alpha parameters (e.g. in SEI) is for further study]

### Semantics

version is an integer that specifies the version of this box.

flags is a 24-bit integer with flags; the following values are defined:

is\_ premultiplied: Flag mask is 0x000001. Specifies if the frame values of the primary video stream comprised in the referenced video track alpha values are premultiplied by the alpha values. The value of 0 specifies that the frame values of the primary video stream are not premultiplied by the alpha values. The value of 1 specifies that the frame values of the primary video stream are premultiplied by the alpha values.

opaque\_value specifies the alpha value for which the referenced video track values are considered opaque for the purposes of alpha blending.

transparent\_value specifies the alpha value for which the referenced video track values are considered transparent for the purposes of alpha blending.

TDB

## Track definitions

### Depth track definition

[Editor’s note: the introduction of 'dept' as handler\_type is to be further studied. This section currently lists the two high-level approaches agreed by the group for consideration.]

Currently, in ISOBMFF both depth and alpha are carried as auxiliary video by using the 'auxv' handler type. Auxiliary video tracks are typically associated to a video track and are not intended to be visually displayed. The type of reference indicates the specific relation to the associated track. Specifically for alpha, the reference\_type can be set to 'auxl', while for depth to 'vdep'.

This seems to be sufficient for carriage of alpha, since it is not visually displayed and it is always associated with a video stream (upon which the alpha map is based). However, carriage of depth streams is more complicated for several reasons:

* A depth map might not be associated with a visual scene (e.g. for privacy reasons, for scene analysis, IR night capture etc.).
* A depth map might be coded with a non-video based codec (e.g. RVL).
* Dimensions and projection of a depth map might not be aligned with the dimensions and projection of the video stream (requiring additional signalling/processing).
* The range of the depth map (near, far) might vary over time.

To address the aforementioned issues the following two approaches for possible solutions are studied by the group.

***1. Introducing a depth handler***

With this approach a media handler for depth can be used (e.g. 'dept'). It will be used to describe tracks containing depth map data.

Pros:

* Depth will not be rendered by incompatible players.
* Can carry depth coded in any scheme, including with depth-specific codecs.
* Can be easily ignored on legacy platforms.
* Can carry depth without being associated with a video stream.
* Supports tracks auxiliary to the depth track.

Cons:

* Requires signalling for the depth codec used.

***2. Using Depth Info Box***

With this approach a compatible player will look for a box containing information on the depth map (e.g. 'depi'). If the box is found it is assumed to be a track carrying depth data.

Pros:

* Can carry depth without being associated with a video stream.
* All information on the stream is in the same place.

Cons:

* Identifying the type of content (happening during track selection) is mixed with parsing parameters to interpret (happening after track selection).
* Does not support non video-based codecs.
* Legacy players will ignore the box therefore processing the content as regular (or auxiliary) video.
* Less robust to file conversion, box can be easily dropped.
* The question of which handler to use that will support all the requirements still remains.

TDB

### Auxiliary video track definition

[Editor’s note: This document is a placeholder for this clause but some of this clause may be converted later into an amendments of corresponding specifications and replaced with mere references to them.]

#### General

[Editor’s note: the usage of handler type is to be further studied – especially with regards to requirements and use cases concerning backwards compatibility and unintended visual playback of content].

Auxiliary videos carried in an auxiliary video track use the 'auxv' handler type in the HandlerBox of the MediaBox, as defined in 8.4.3 of ISOBMFF [1].

When the 'auxv' handler type is used for an auxiliary video track, the AuxiliaryVideoTypeBox shall be present as defined in clause 7.5.2.

[Editor’s note: This is assumed that there will be a brand defined for this constraint which will prevent exiting files to become invalid.]

When the 'auxv' is used to carry depth, the aux\_video\_type in AuxiliaryVideoTypeBox shall be equal to 0. When aux\_video\_type is equal to 0 the DepthInformationBox, as defined in clause 7.2.2, shall be present in the VisualSampleEntry.

When the 'auxv' is used to carry alpha, the aux\_video\_type in AuxiliaryVideoTypeBox shall be equal to 1. When aux\_video\_type is equal to 1 the AlphaInformationBox, as defined in clause 7.3.2, shall be present in the VisualSampleEntry.

#### Auxiliary video type box

##### Definition

Box Type: 'auxt'   
Container: Visual sample entry  
Mandatory: see below  
Quantity: Zero or one per sample entry

The AuxiliaryVideoTypeBox is used to indicate the types of auxiliary videos present in track samples associated with this box through the sample entry.

**Table 1 – Auxiliary video types**

|  |  |
| --- | --- |
| Value | Description |
| 0 | Depth video |
| 1 | Alpha video |

#### Syntax

aligned(8) class AuxiliaryVideoTypeBox extends FullBox('auxt') {  
 unsigned int(16) aux\_video\_type;  
}

#### Semantics

aux\_video\_type indicates the type of the auxiliary video, as defined in Table 1

[Editor’s Note: Using the CICP code point for alpha and depth is to be considered for further study.]

## Track reference

The track reference 'auxl' shall be used as described in 8.3.3.3 of ISOBMFF[1].

# Brands

TDB

# Integration with delivery format

TDB

Bibliography

[1] ISO 14496‑12, *Information technology — Coding of audio-visual objects — Part 12: ISO base media file format*

[2] ISO #####‑##:20##, *General title — Part ##: Title of part*