**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

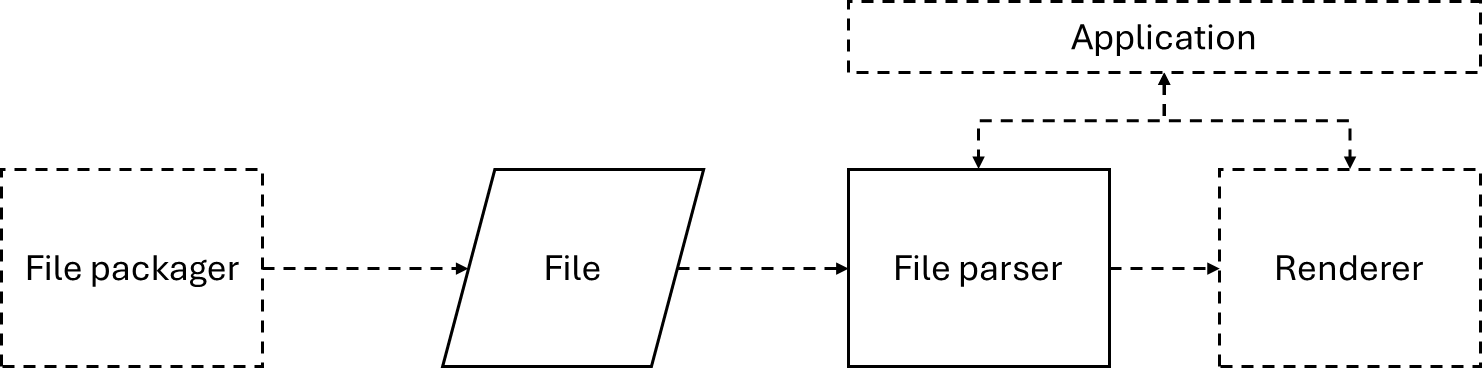
term

text of the definition

# Example application scenarios

## File-based consumption (informative)

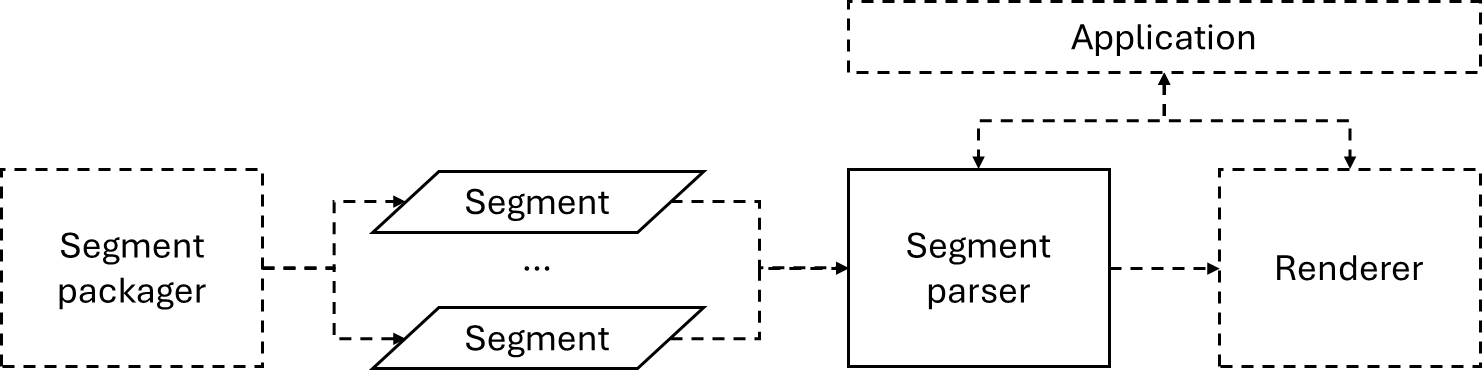
An application may consume a file compliant to this document.



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: 'depx'   
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 ('depx', version = 0, flags){  
 if (!is\_unit\_interval\_mode){  
 float(32) near\_plane;  
 float(32) far\_plane; unsigned int(6) units;  
 unsigned int(2) reserved;  
 }

}

### 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\_unit\_interval\_mode: Flag mask is 0x000001. The value 1 indicates that the near\_plane value is equal to 0, the far\_plane value is equal to 1, and the units value is equal to 0.

depth\_mapping\_type: Flag mask is 0x000006. The value indicates the type of mapping between decoded sample values and depth values, as specified in Table 1.

Table 1 – Definition of depth\_mapping\_type.

|  |  |
| --- | --- |
| depth\_mapping\_type | Interpretation |
| 0 | When depth\_mapping\_type is equal to 0, a linear relationship is defined between the decoded sample values and the depth values, where the decoded sample value equal to 0 corresponds to , and the decoded sample value equal to maxVal corresponds to . |
| 1 | When depth\_mapping\_type is equal to 1, an inverse relationship is defined between the decoded sample values and the depth values, where the decoded sample value equal to 0 corresponds to the inverse of , and the decoded sample value equal to maxVal corresponds to the inverse of . |
| 2 | When depth\_mapping\_type is equal to 2, a linear relationship is defined between the decoded sample values and the depth values, where the decoded sample value equal to 0 corresponds to , and the decoded sample value equal to maxVal corresponds to . |
| 3 | When depth\_mapping\_type is equal to 3, an inverse relationship is defined between the decoded sample values and the depth values, where the decoded sample value equal to 0 corresponds to the inverse of , and the decoded sample value equal to maxVal corresponds to the inverse of . |

The variable of Table 1 is specified by Equation 1:

(1)

The variable of Table 1 is specified by Equation 2:

(2)

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

near\_plane and far\_plane specify the nearest and the farthest depth values, respectively. The near\_plane value can be less than the far\_plane value or it can be greater than the far\_plane value. The near\_plane value and the far\_plane value shall have the same sign that can be either positive or negative.

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;  
}

### 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.

## Track definitions

### Depth video track definition

**7.4.1.1 General**

Depth video tracks use the 'depv' handler type in the HandlerBox of the MediaBox, as defined in 8.4.3 of ISOBMFF [1].

A depth video track is coded the same as a video track, but uses this different handler type, and is not intended to be visually displayed. Depth video tracks may be linked to a video track by a 'cdsc' track reference as described in 7.5.

Depth video tracks use VisualSampleEntry.

### Alpha video track definition

**7.4.2.1 General**

Alpha video tracks use the 'alpv' handler type in the HandlerBox of the MediaBox, as defined in 8.4.3 of ISOBMFF [1].

An alpha video track is coded the same as a video track, but uses this different handler type, and is not intended to be visually displayed. Alpha video tracks are linked to a video track by a 'cdsc' track reference as described in 7.5.

Alpha video tracks use VisualSampleEntry.

## Track reference

When a depth video track is linked to a video track, the track reference 'cdsc' shall be used as described in 8.3.3.3 of ISOBMFF[1].

When an alpha video track is linked to a video track, the track reference 'cdsc' shall be used as described in 8.3.3.3 of ISOBMFF[1].

# Brands

TDB

# Integration with delivery format

## Carriage of Depth and Alpha using HEVC

### Compatibility with single-layer HEVC

There is no provision for the carriage of Depth representation information SEI message in single-layer HEVC and therefore cannot be present in the file.

There is no provision for the carriage of Alpha channel information SEI message in single-layer HEVC therefore cannot be present in the file.

### Compatibility with multi-layer HEVC and MV-HEVC

If the Depth representation information SEI message is present in the sample entry or in the media sample of the track, the values of near\_plane and far\_plane of the DepthInformationBox should be equal to the values of ZNear and ZFar as defined in Section G.14.3.3 of HEVC. In case both the Depth representation information SEI message and DepthInformationBox are present, the information of DepthInformationBox shall be used.

If the Alpha channel information SEI message is present in the sample entry or in the media sample of the track, the values of opaque\_value and transparent\_value of the AlphaInformationBox should be equal to the values of alpha\_opaque\_value and alpha\_transparent\_value as defined in Section F.14.3.8. In case both the Alpha channel information SEI message and AlphaInformationBox are present, the information of AlphaInformationBox shall be used.

[Editor’s Note: Further constraints specific to the box fields with respect to the SEI messages parameters are to be defined as the box definitions become stable.]

# Annex I (informative)

## Handling of parameters in SEI messages and container level

NOTE Some of the following rules are the consequences of the file format constraints defined for the carriage of depth and alpha track (see Section 9).

It is recommended:

1. If a Depth representation information SEI message, as specified in AVC [6], HEVC [7] or VVC/VSEI [4], with depth\_representation\_type equal to 1 or 3 is present, the DepthInformationBox should not be present. If present, the parameters of DepthInformationBox shall be ignored.
2. If a Depth representation information SEI message, as specified in AVC [6], HEVC [7] or VVC/VSEI [4], with depth\_representation\_type equal to 0 or 2, and a DepthInformationBox are both present, the parameters of the DepthInformationBox should be set as follows:
   1. If depth\_representation\_type = 0, then depth\_mapping\_type = 3.
   2. If depth\_representation\_type = 2, then depth\_mapping\_type = 2.
   3. If z\_far\_flag = 1, then far\_plane = ZFar and units = unspecified
   4. If z\_near\_flag = 1, then near\_plane = ZNear and units = unspecified
3. If an Alpha channel information SEI message, as specified in HEVC [7] or VVC/VSEI [4], and an AlphaInformationBox are both present, the parameters of the AlphaInformationBox should be set as follows:
   1. If alpha\_channel\_use\_idc = 0, then is\_premultiplied = 0
   2. If alpha\_channel\_use\_idc = 1, then is\_premultiplied = 1
   3. transparent\_value = alpha\_transparent\_value
   4. opaque\_value = alpha\_opaque\_value

## Lossless conversion between SEI-based and IEEE 754 32-bit floating-point formats

### SEI-based floating point format

The SEI-based refers to the floating-point format specified by the Depth representation information element syntax, provided in clause G.14.2.4.2 of HEVC [1] and given in Table 2.

Table 2 – Depth representation information element syntax, as defined in HEVC, clause G.14.2.4.2 [1].

A screenshot of a computer

AI-generated content may be incorrect.

The three main fields of this floating-point format are defined as follows:

* **Sign**:1 bit
* **Exponent**: 7 bits **| Bias** = 31 **| Unbiased exponent values** = [-30, 95]
* **Mantissa:** 1-32 bits

The length of the mantissa field (i.e., 1 to 32 bits) is determined by the da\_mantissa\_len\_minus1 parameter.

### IEEE 754 32-bit floating-point format

The IEEE 754 32-bit refers to the single precision floating-point format specified in the IEEE 754 standard [2].

The three main fields of this floating-point format are defined as follows:

* **Sign**:1 bit
* **Exponent**: 8 bits **| Bias** = 127 **| Unbiased exponent values** = [-126, 127]
* **Mantissa:** 23 bits

### Lossless conversion rules

There are two rules that should be met for lossless conversion of the SEI-based floating point format () to the IEEE 754 32-bit floating-point format (), within the common range of real (representable) numbers ():

1. The mantissa length of must be set equal to 23, and
2. The unbiased exponent values of and must be equal, according to Equation 3

(3)

where represents the unsigned integer value of the exponent field of and takes a value in the range 0 to 127, represents the bias specified for and is equal to , represents the unsigned integer value of the exponent field of and takes a value in the range 0 to 255, represents the bias specified for and is equal to , and is determined by the smallest negative and the largest positive real (representable) number of X.

NOTE 1 Special cases, such as the is equal to 0 and 127, require special handling.

NOTE 2 To respect all aforementioned definitions, should take a value in the range 97 to 222.

NOTE 3 The smallest negative number represented by X is , and the largest positive number represented by X is .

1. ISO/IEC 23008-2 “Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 2: High efficiency video coding (HEVC)”
2. IEEE Computer Society (2019-07-22). IEEE Standard for Floating-Point Arithmetic. IEEE STD 754-2019. IEEE. pp. 1–84.

## Interpretation of depth decoded sample values

Decoded sample values are interpreted as depth values that are computed based on the depth\_mapping\_type value, as specified in Table 1, and the variables and , as specified by Equations 1 and 2, respectively, according to Equation 4:

(4)

Indicative graphs representing Equation 4 are shown in Figure 3.

A blue dotted line in a black background

AI-generated content may be incorrect.A blue dotted line on a black background

AI-generated content may be incorrect.

Figure 3 – The left plot corresponds to graphs obtained for depth\_mapping\_type value equal to 0 and 1. The right plot corresponds to graphs obtained for depth\_mapping\_type value equal to 2 and 3. A solid line is used to represent a linear relationship, specified with depth\_mapping\_type value equal to 0 and 2, and a dashed line is used to represent an inverse relationship, specified with depth\_mapping\_type value equal to 1 and 3, between the decoded sample values and the depth values .

## Recommended implementation of alpha blending composition

A recommended implementation of the alpha blending composition is provided in clause 8.23.2 of VSEI [4] specification. An identical implementation is obtained using the parameters specified in the AlphaInformationBox by applying the following changes in the corresponding VSEI [4] text:

* “alpha\_opaque\_value” is replaced by “opaque\_value”.
* “alpha\_transparent\_value” is replaced by “transparent\_value”.
* “alpha\_channel\_use\_idc equal to 0” is replaced by “is\_premultiplied equal to 0”.
* “alpha\_channel\_use\_idc equal to 1” is replaced by “is\_premultiplied equal to 1”.

Bibliography

1. ISO 14496‑12, *Information technology — Coding of audio-visual objects — Part 12: ISO base media file format*
2. ISO/IEC 23002-7*, Information technology — MPEG video technologies — Part 7: Versatile supplemental enhancement information messages for coded video bitstreams*
3. ISO/IEC 14496-10, *Information technology — Coding of audio-visual objects — Part 10: Advanced Video Coding (AVC)*
4. ISO/IEC 23008-2, *Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 2: High efficiency video coding (HEVC)*
5. ISO #####‑##:20##, *General title — Part ##: Title of part*