**ISO/IEC 23008-12:2024/CDAM 2:2024(E)**

ISO/IEC JTC1/SC 29

Secretariat: JISC

**Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 12: Image File Format — Amendment 2: Low-overhead image file format**

CD stage

**Copyright notice**

This ISO document is a working draft or committee draft and is copyright-protected by ISO. While the reproduction of working drafts or committee drafts in any form for use by participants in the ISO standards development process is permitted without prior permission from ISO, neither this document nor any extract from it may be reproduced, stored or transmitted in any form for any other purpose without prior written permission from ISO.

Requests for permission to reproduce this document for the purpose of selling it should be addressed as shown below or to ISO's member body in the country of the requester:

ISO copyright office

Case postale 56 • CH-1211 Geneva 20

Tel. + 41 22 749 01 11

Fax + 41 22 749 09 47

E-mail copyright@iso.org

Web www.iso.org

Reproduction for sales purposes may be subject to royalty payments or a licensing agreement.

Violators may be prosecuted.

**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 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](http://www.iso.org/directives) or [www.iec.ch/members\_experts/refdocs](http://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](http://www.iso.org/patents)) or the IEC list of patent declarations received (see [patents.iec.ch](https://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 the following URL:  [www.iso.org/iso/foreword.html](http://www.iso.org/iso/foreword.html). In the IEC, see [www.iec.ch/understanding-standards](https://www.iec.ch/understanding-standards)

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

A list of all parts in the ISO/IEC 14496 and ISO/IEC 23008 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](https://www.iso.org/members.html) and [www.iec.ch/national-committees](http://www.iec.ch/national-committees).

**Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 12: Image File Format — Amendment 2: Low-overhead image file format**

1. Update the table of contents to reflect the status of current amendment.
2. Update Introduction

Add the following text at the end:

Annex O specifies the low-overhead image file format suitable for small and simple images used in specific use cases. In this format, the top-level MetaBox is replaced by a MinimizedImageBox to reduce the overhead relative to the size of the image or metadata payloads, while logically representing the contents of the MetaBox. Annex O also specifies the requirements and brands for this format, as well as the procedure for expanding it to the regular Image File Format.

1. Add a Note to Clause 6.1

Add the following NOTE after the paragraph: "Images can be stored as items using the support for untimed data storage, called the MetaBox for historical reasons, in the ISO base media file format. A file may contain any number of image items."

NOTE For small and simple files, a pre-processed version of the MetaBox called the MinimizedImageBox, as defined in Annex O, can be used. A file containing a MinimizedImageBox has the same compliance and rendering requirements as an Image File Format file because the MinimizedImageBox is expanded into the equivalent MetaBox as described in detail in Annex O.4. Once expanded, the rest of Clause 6 applies.

1. Update Clause 6.5.6

Change the following text in clause 6.5.6.1:

The PixelInformationProperty descriptive item property indicates the number and bit  
depth of colour components in the reconstructed image of the associated image item.

To:

The PixelInformationProperty descriptive item property indicates the number and bit depth of colour and alpha/depth components, if present, in the reconstructed image of the associated image item.

Change the syntax in clause 6.5.6.2 to:

aligned(8) class PixelInformationProperty

extends ItemFullProperty('pixi', version, px\_flags){

if(version == 0 || version == 1) {

unsigned int(8) num\_channels;

for (i=0; i<num\_channels; i++) {

unsigned int(8) bits\_per\_channel;

}

}

if(version == 1) {

unsigned int alpha\_flag = (px\_flags & 1) != 0;

unsigned int alpha\_premultiplied\_flag = (px\_flags & 2) != 0;

unsigned int subsampling\_flag = (px\_flags & 4) != 0;

for (i=0; i<num\_channels; i++) {

unsigned int(3) channel\_idc;

unsigned int(2) channel\_format;

unsigned int(1) channel\_label\_flag;

unsigned int(2) reserved = 0;

if(subsampling\_flag) {

unsigned int(4) subsampling\_type;

unsigned int(4) subsampling\_location;

}

if(channel\_label\_flag) {

utf8string channel\_label;

}

}

}  
}

Add the following text to the end of clause 6.5.6.3:

alpha\_flag: If true, indicates that the coded image contains alpha. If true, at least one of the channels shall have channel\_idc set to 0 and one shall have channel\_idc set to 1. Only applicable when version is 1.

alpha\_premultiplied\_flag: If true, the channels with channel\_idc set to 0 have been premultiplied by the channel with channel\_idc set to 1. If true, alpha\_flag shall also be true. Only applicable when version is 1.

subsampling\_flag: If true, indicates that subsampling information is present. Only applicable when version is 1.

channel\_idc: This field indicates the contents of the channel. A value of 0 indicates colour/grayscale. A value of 1 indicates alpha. A value of 2 indicates depth. Values 3-7 are reserved for future use. At most one channel shall have a channel\_idc of 1.

channel\_format: This field indicates the data type of the channel as defined by the component\_format values in ISO/IEC 23001-17. component\_bit\_depth is considered to be equal to bits\_per\_channel.

channel\_label\_flag: This flag indicates the presence of the channel\_label.

subsampling\_type: This field indicates the subsampling type as specified by GenericSubsamplingType in Rec. ITU-T H.273 | ISO/IEC 23091-2.

subsampling\_location: This field indicates the subsampling sample location as specified by GenericSubsamplingSampleLocType in Rec. ITU-T H.273 | ISO/IEC 23091-2.

channel\_label: The human readable description of the channel.

EDITORS NOTE:  
ITU-T H.273 | ISO/IEC 23091-2 (CICP for video) is currently being updated to include the definitions of the new subsampling types that we would like to use in pixi v1. We need to monitor the standardization progress of ISO/IEC 23091-2 with the WG5.

EDITORS NOTE:  
Anything other than unsigned integers may need clarifications with regards to colr-nclx as noted here: <https://git.mpeg.expert/MPEG/Systems/FileFormat/HEIF/-/issues/95#note_81705>

1. Add the following new subclauses after subclause 11.3.3:

#### 11.3.4 Groups of regions

##### 11.3.4.1 Union of regions entity group

A union of regions entity group ('unrg') indicates the union of all the regions represented by one or more entities describing regions of an image.

Each entity\_id value in the entity group shall refer to a region item.

All the region items in the union of regions entity group shall be associated with the same image item, inside which the regions are defined, using an item reference of type 'cdsc' from each region item to the same image item.

If unique IDs are used:

* the union of regions entity group may also be associated with the image item inside which the regions are defined using an item reference of type 'cdsc' from the union of regions entity group to the image item,
* an annotation may be associated with the union of regions entity group by associating:
  + an item property, using the ItemPropertyAssociationBox;
  + a metadata item, using an item reference of type 'cdsc' from the metadata item to the union of regions entity group;
  + an image item or another entity group, using an item reference of type 'eroi' from the union of regions entity group to the image item or the other entity group.

##### 11.3.4.2 Compound region entity group

A compound region entity group ('corg') associates one main region item with one or more region items. It indicates an inclusion relationship between a main object covered by regions of a main entity and other objects covered by regions described by one or more other entities, the main object logically including the other objects.

NOTE 1 For example, a compound region entity group can be used to associate a main region corresponding to a body with regions corresponding to body parts (e.g., the head, legs or arms of the body) to indicate that the body is logically including the body parts.

The entities in a compound region entity group shall be region items.

The number of entities in a compound region entity group shall be at least 2. The first entity\_id value shall indicate the main region item. It indicates the region covering the main object that is logically including the objects covered by the regions described by the second and following entity\_ids.

This inclusion relationship does not convey information at the geometry level. A main region signaled as including others regions by a compound region entity group may or may not geometrically include the other regions.

NOTE 2 For example, the main region item corresponding to the first entity\_id value can represent a bounding box or a region encompassing partially the regions described by the region items corresponding to the second and following entity\_id values.

All the region items in the compound region entity group shall be associated with the same image item, inside which the regions are defined, using an item reference of type 'cdsc' from each region item to the same image item.

If unique IDs are used:

* the compound region entity group may also be associated with the image item, inside which the regions are defined, using an item reference of type 'cdsc' from the compound region entity group to the image item,
* an annotation may be associated with the compound region entity group by associating:
  + an item property, using the ItemPropertyAssociationBox;
  + a metadata item, using an item reference of type 'cdsc' from the metadata item to the compound region entity group;
  + an image item or another entity group, using an item reference of type 'eroi' from the union of regions entity group to the image item or the other entity group.

1. Add a new subclause A.2.3

#### A.2.3 Untimed compressed Exif metadata

When an ExifDataBlock, as described in A.2.1, is compressed using the deflate() algorithm defined in IETF RFC 1951, and

the resulting untimed compressed Exif metadata is stored as a metadata item, the item\_type value shall be set to 'dExf'.

1. Add a new subclause J.7

### J.7 Single image in a low-overhead image file

A file with a single coded image item could be structured as follows:

FileTypeBox 'ftyp': major\_brand='mif3', minor\_version='heic', compatible\_brands=''

MinimizedImageBox 'mini': (container)

width\_minus1=W

height\_minus1=H

main\_item\_codec\_config\_size=C,

main\_item\_data\_size\_minus1=D,

main\_item\_codec\_config=HEVCDecoderConfigurationRecord (with length C)

main\_item\_data=HEVC Image Data (with length D+1)

1. Rename clause L.4.1

### L.4.1 'mif1'-compliant VVC image and image collection brands

1. Add new subclause L.4.4

### L.4.4 'mif3'-compliant VVC image and image collection brand

#### L.4.4.1 Requirements on files

Files shall include 'mif3' as the major\_brand and the brand 'vvi3' as the minor\_version in the FileTypeBox and conform to the specifications in ‎O.2.1 of this document ('mif3' structural brand).

The 'vvi3' brand defines the coded image item type to be 'vvc1' as defined in L.2.2.1.2 and the alpha\_item\_codec\_config (if present), gainmap\_item\_codec\_config (if present) and main\_item\_codec\_config to have the format defined by CompactVvcDecoderConfigurationRecord specified in L.4.4.3 of this document.

Files that include 'mif3' as the major\_brand and the brand 'vvi3' as the minor\_version in the FileTypeBox shall comply with the specifications in ‎L.2 and shall additionally be constrained as follows:

* explicit\_codec\_types\_flag in the MinimizedImageBox shall be equal to 0,
* float\_flag in the MinimizedImageBox shall be 0,
* If high\_bit\_depth\_flag in the MinimizedImageBox is not 0, bit\_depth\_minus9 in the MinimizedImageBox shall be between 0 and 7 inclusive,
* If gainmap\_item\_data\_size in the MinimizedImageBox is not 0, gainmap\_float\_flag in the MinimizedImageBox shall be 0,
* If gainmap\_item\_data\_size in the MinimizedImageBox is not 0 and high\_bit\_depth\_flag in the MinimizedImageBox is not 0, gainmap\_bit\_depth\_minus9 in the MinimizedImageBox shall be between 0 and 7 inclusive,
* If alpha\_flag in the MinimizedImageBox is equal to 1, alpha\_item\_data\_size in the MinimizedImageBox shall not be equal to 0,
* The data of each codec configuration property shall use the syntax defined in subclause L.4.4.3.2 of this document,
* Each coded image item shall contain a single IDR picture per layer,
* The data of each coded image item shall consist of only one NAL unit excluding the length and NAL unit header fields, which are inferred by the reader as specified in subclause L.4.4.2 of this document,
* The equivalent file as specified in L.4.4.1.2 shall conform to the 'vvic' brand.

#### L.4.4.2 Requirements on readers

The requirements on readers specified in ‎L.4.1.3 and ‎O.2.1 of this document ('mif3' structural brand) shall be supported.

As a response to 'mif3' as the major\_brand and the brand 'vvi3' as the minor\_version in the FileTypeBox, the readers shall treat the file as if an equivalent file were created containing FileTypeBox with 'vvic' as the equivalent major\_brand and the equivalent MetaBox derived with the following additional operations in addition to those specified in subclause O.4:

* infer infe\_type to be equal to 'vvc1',
* infer codec\_config\_type to be equal to 'vvcC',
* expand the data of each codec configuration property from the compact syntax to the full structure as defined in section L.4.4.3.4 of this document,
* when multi\_layer\_flag is equal to 0, expand the data of the respective image item so that the NAL unit length is inferred from the item length, and so that the NAL unit header fields nal\_unit\_type, nuh\_temporal\_id\_plus1 and nuh\_layer\_id are set to 8 (IDR\_N\_LP), 1 and 0, respectively.

#### L.4.4.3 Compact VVC decoder configuration

##### L.4.4.3.1 Definition

This subclause specifies the compact decoder configuration information for ISO/IEC 23090-3 video content. The compact decoder configuration provides essential parameters that are relevant for still images.

##### L.4.4.3.2 Syntax

aligned(8) class CompactVvcDecoderConfigurationRecord {  
 unsigned int(1) multi\_layer\_flag;  
 unsigned int(2) lengthSizeMinusOne;  
 if (multi\_layer\_flag) {  
 unsigned int(1) ptl\_present\_flag;  
 if (ptl\_present\_flag) {  
 VvcPTLRecord(1) native\_ptl;  
 if (native\_ptl.ptl\_multilayer\_enabled\_flag == 1)  
 unsigned int(3) ols\_idx;  
 }  
 unsigned int(1) nal\_units\_present\_flag;  
 }  
 if (multi\_layer\_flag == 0 || nal\_units\_present\_flag) {  
 unsigned int(1) additional\_nal\_unit\_flag;  
 if (additional\_nal\_unit\_flag) {  
 unsigned int(3) num\_aps\_nal\_unit;  
 unsigned int(3) num\_sei\_nal\_unit;  
 }  
 if (multi\_layer\_flag)  
 unsigned int(1) array\_completeness;  
 trailing\_bits();  
 if (multi\_layer\_flag) {  
 unsigned int(8) vps\_nal\_unit\_length;  
 bit(8\*vps\_nal\_unit\_length) vps\_nal\_unit;  
 }  
 unsigned int(8) sps\_nal\_unit\_length;  
 bit(8\*sps\_nal\_unit\_length) sps\_nal\_unit;  
 unsigned int(8) pps\_nal\_unit\_length;  
 bit(8\*pps\_nal\_unit\_length) pps\_nal\_unit;  
 if (additional\_nal\_unit\_flag) {  
 for (i=0; i< num\_aps\_nalus; i++) {  
 unsigned int(8) aps\_nal\_unit\_length;  
 bit(8\*aps\_nal\_unit\_length) aps\_nal\_unit;  
 }  
 for (i=0; i< num\_sei\_nalus; i++) {  
 unsigned int(8) sei\_nal\_unit\_length;  
 bit(8\*sei\_nal\_unit\_length) sei\_nal\_unit;  
 }  
 }  
 }  
}

**L.4.4.3.3 Semantics**

multi\_layer\_flag: When equal to 0, the data of the coded image item shall consist of only one NAL unit that excludes the length and NAL unit header fields and has nal\_unit\_type, nuh\_temporal\_id\_plus1 and nuh\_layer\_id equal to to 8 (IDR\_N\_LP), 1 and 0, respectively. When equal to 1, the data of the coded image item may include any number of NAL units, which include the length and NAL unit header fields.

nal\_units\_present\_flag: indicates that NAL units are present in the decoder configuration.

additional\_nal\_unit\_flag: equal to 1 indicates the presence of additional NAL units in the decoder configuration record. additional\_nal\_unit\_flag equal to 0 indicates the absence of additional NAL units in the decoder configuration record.

num\_aps\_nal\_unit: indicates the number of APS NAL units included in the configuration record for the referenced CVS.

num\_sei\_nal\_unit: indicates the number of SEI NAL units included in the configuration record for the referenced CVS.

all

vps\_nal\_unit\_length: indicates the length in bytes of the NAL unit. When equal to 0, the VPS NAL unit is not present.

vps\_nal\_unit: contains the VPS NAL unit as specified in ISO/IEC 23090-3.

sps\_nal\_unit\_length: indicates the length in bytes of the NAL unit. When equal to 0, the SPS NAL unit is not present.

sps\_nal\_unit: contains the SPS NAL unit as specified in ISO/IEC 23090-3.

pps\_nal\_unit\_length: indicates the length in bytes of the NAL unit. When equal to 0, the PPS NAL unit is not present.

pps\_nal\_unit: contains the PPS NAL unit as specified in ISO/IEC 23090-3.

aps\_nal\_unit\_length: indicates the length in bytes of the APS NAL unit.

aps\_nal\_unit: contains the APS NAL unit as specified in ISO/IEC 23090-3.

sei\_nal\_unit\_length: indicates the length in bytes of the SEI NAL unit.

sei\_nal\_unit: contains the SEI NAL unit as specified in ISO/IEC 23090-3.

The semantics of the other parameters are the same as for VvcDecoderConfigurationRecord as defined in ISO/IEC 14496-15.

##### L.4.4.3.4 Equivalence with the VVC decoder configuration

CompactVvcDecoderConfigurationRecord shall be considered equivalent to VvcDecoderConfigurationRecord as defined in ISO/IEC 14496-15 with the following fields:

* if ptl\_present\_flag is present and set to 1:
  + num\_sublayers is set to 1,
  + constant\_frame\_rate is set to 1,
  + if the codec configuration property is associated with the main image item:
    - chroma\_format\_idc is set to the value of the chroma\_subsampling field from the MinimizedImageBox,
  + if the codec configuration property is associated with the alpha auxiliary image item:
    - chroma\_format\_idc is set to 0,
  + if the codec configuration property is associated with the main image item or with the alpha auxiliary image item:
    - bit\_depth\_minus8 is set to 0 if the value of the high\_bit\_depth\_flag field from the MinimizedImageBox is 0, or to the value of the bit\_depth\_minus9 field from the MinimizedImageBox plus 1,
    - max\_picture\_width is set to the value plus 1 of the width\_minus1 field from the MinimizedImageBox,
    - max\_picture\_height is set to the value plus 1 of the height\_minus1 field from the MinimizedImageBox,
  + if the codec configuration property is associated with the gain map image item:
    - bit\_depth\_minus8 is set to 0 if the value of the gainmap\_high\_bit\_depth\_flag field from the MinimizedImageBox is 0, or to the value of the gainmap\_bit\_depth\_minus9 field from the MinimizedImageBox plus 1,
    - max\_picture\_width is set to the value plus 1 of the gainmap\_width\_minus1 field from the MinimizedImageBox,
    - max\_picture\_height is set to the value plus 1 of the gainmap\_height\_minus1 field from the MinimizedImageBox,
  + avg\_frame\_rate is set to 0,
* if nal\_units\_present\_flag is set to 1:
  + num\_of\_arrays is set to the number of entries below:
    - if vps\_nal\_unit\_length is present and not 0, there is a VPS NAL unit array with:
      * NAL\_unit\_type set to 14 (VPS\_NUT as defined in ISO/IEC 23090-3),
      * num\_nalus set to 1,
      * nal\_unit\_length set to vps\_nal\_unit\_length,
      * nal\_unit set to vps\_nal\_unit.
    - if sps\_nal\_unit\_length is not 0, there is a SPS NAL unit array with:
      * NAL\_unit\_type set to 15 (SPS\_NUT as defined in ISO/IEC 23090-3),
      * num\_nalus set to 1,
      * nal\_unit\_length set to sps\_nal\_unit\_length,
      * nal\_unit set to sps\_nal\_unit.
    - if pps\_nal\_unit\_length is not 0, there is a PPS NAL unit array with:
      * NAL\_unit\_type set to 16 (PPS\_NUT as defined in ISO/IEC 23090-3),
      * num\_nalus set to 1,
      * nal\_unit\_length set to pps\_nal\_unit\_length,
      * nal\_unit set to pps\_nal\_unit.
    - if additional\_nal\_unit\_flag is set to 1 and num\_aps\_nal\_unit is not 0, there is a prefix APS NAL unit array with:
      * NAL\_unit\_type set to 17 (PREFIX\_APS\_NUT as defined in ISO/IEC 23008-2),
      * num\_nalus set to num\_aps\_nal\_unit, and for each prefix APS NAL unit:
        + nal\_unit\_length is set to aps\_nal\_unit\_length,
        + nal\_unit is set to aps\_nal\_unit.
    - if additional\_nal\_unit\_flag is set to 1 and num\_sei\_nal\_unit is not 0, there is a prefix SEI NAL unit array with:
      * NAL\_unit\_type set to 23 (PREFIX\_SEI\_NUT as defined in ISO/IEC 23008-2),
      * num\_nalus set to num\_sei\_nal\_unit, and for each prefix SEI NAL unit:
        + nal\_unit\_length is set to sei\_nal\_unit\_length,
        + nal\_unit is set to sei\_nal\_unit.
* if array\_completeness is not present in CompactVvcDecoderConfigurationRecord, it is set equal to 1 in VvcDecoderConfigurationRecord.
* the other parameters are carried over as is, and repeated if needed.

1. Add new Annex O

**Annex O**(normative)

**Low-overhead image file format**

### O.1 General

The low-overhead image file format provides a more compact representation of the image file format for specific use-cases. This format is designed for small and simple files where the traditional use of the MetaBox would result in significant overhead relative to the size of the image and/or metadata payloads. In this format, the top-level MetaBox is replaced by a MinimizedImageBox, which logically maintains the presence of the MetaBox by representing its contents.

When a parser encounters a MinimizedImageBox, it shall expand it to a MetaBox as described in Clause O.4, at which point the specifications of Clause 6 apply.

When a brand specified in Clause O.2 is the major brand or among the compatible brands of a file, either explicitly or implicitly, the requirements specified in this annex shall be applied.

### O.2 Brands

#### O.2.1 'mif3' structural brand

##### O.2.1.1 Requirements on files

Files containing the brand 'mif3' as the major brand or in the compatible brands array of the FileTypeBox shall conform to the constraints defined in this subclause.

When the 'mif3' brand is the major brand or present among the compatible brands of the FileTypeBox, the file may be identified by the MIME type defined in Annex P. When the 'mif3' brand is the major brand, the defined file extension and MIME type defined in Annex P should be used.

The boxes listed in Table XX1 are required in a file under the 'mif3' brand. The Version column in the following table lists the versions of the boxes allowed by this brand. Other versions of the boxes shall not be present. Other file-level boxes shall not be present.

**Table XX1 — Required boxes in a file under the** 'mif3' **brand**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Hierarchy of boxes** | | | **Version** | **Box description** |
| ftyp |  |  | - | file type and compatibility |
| mini |  |  | 0 | metadata and image data |

When the 'mif3' brand is present as the major brand of the FileTypeBox, the minor version of the FileTypeBox shall be 0 or a brand.

##### O.2.1.2 Requirements on readers

When the 'mif3' brand is present as the major\_brand of the FileTypeBox, the minor\_version of the FileTypeBox shall be 0 or a brand that is either structurally compatible with the 'mif3' brand, such as a codec brand complying with the 'mif3' structural brand, or a brand to which the file conforms after the equivalent MetaBox has been transformed from MinimizedImageBox as specified in Clause O.4.

**Table XX2 — Boxes to be supported under the 'mif3' brand  
in addition to those required for readers of 'mif1' brand**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Hierarchy of boxes** | | | **Version** | **Box description** |
| mini |  |  | 0 | metadata and image data |

A file containing a MinimizedImageBox shall be treated as if:

* it contained the 'mif1' brand in the compatible\_brands array in the FileTypeBox,
* it contained the equivalent MetaBox as described in Clause O.4,
* the major\_brand of the FileTypeBox was the equivalent brand specified by the brand provided in the minor\_version, if the major\_brand is 'mif3' and the minor\_version is not 0.

### O.3 Minimized Image Box

#### O.3.1 Definition

|  |  |
| --- | --- |
| Box type: | 'mini' |
| Container: | File |
| Mandatory: | No |
| Quantity: | Zero or one |

The minimized image box format provides a more compact representation of the MetaBox for a subset of use cases. It is meant to be used for small and simple files where the full MetaBox would result in considerable overhead compared to the image data payload.

When MinimizedImageBox is present in a file, the 'mif3' brand or a derived brand that implies the 'mif3' brand shall be the major brand or present among the compatible brands in the FileTypeBox.

#### O.3.2 Syntax

aligned(8) class MinimizedImageBox extends Box('mini') {

bit(2) version = 0;

// flags

explicit\_codec\_types\_flag

float\_flag

full\_range\_flag

alpha\_flag

explicit\_cicp\_flag

bit(1) hdr\_flag;

bit(1) icc\_flag;

bit(1) exif\_flag;

bit(1) xmp\_flag;

bit(3) orientation\_minus1;

// Spatial extents

bit(1) small\_dimensions\_flag;

unsigned int(small\_dimensions\_flag ? 7 : 15) width\_minus1;

unsigned int(small\_dimensions\_flag ? 7 : 15) height\_minus1;

// Pixel information

if (chroma\_subsampling == 1 || chroma\_subsampling == 2)

bit(1) chroma\_is\_horizontally\_centered;

if (chroma\_subsampling == 1)

bit(1) chroma\_is\_vertically\_centered;

if (float\_flag)

bit(2) bit\_depth\_log2\_minus4;

else {

bit(1) high\_bit\_depth\_flag;

if (high\_bit\_depth\_flag)

bit(3) bit\_depth\_minus9;

}

if (alpha\_flag)

bit(1) alpha\_is\_premultiplied;

// Colour properties

if (explicit\_cicp\_flag) {

bit(8) colour\_primaries;

if (chroma\_subsampling != 0)

else

}

else {

colour\_primaries = icc\_flag ? 2 : 1;

icc\_flag ? 2 : 13

chroma\_subsampling == 0 ? 2 : 6

}

// Optional unless minor\_version of FileTypeBox is a brand defining these

if (explicit\_codec\_types\_flag) {

bit(32) infe\_type;

bit(32) codec\_config\_type;

}

// High Dynamic Range properties

if (hdr\_flag) {

bit(1) gainmap\_flag;

if (gainmap\_flag) {

unsigned int(small\_dimensions\_flag ? 7 : 15) gainmap\_width\_minus1;

unsigned int(small\_dimensions\_flag ? 7 : 15) gainmap\_height\_minus1;

bit(8) gainmap\_matrix\_coefficients;

bit(1) gainmap\_full\_range\_flag;

bit(2) gainmap\_chroma\_subsampling;

if (gainmap\_chroma\_subsampling == 1 || gainmap\_chroma\_subsampling == 2)

bit(1) gainmap\_chroma\_is\_horizontally\_centered;

if (gainmap\_chroma\_subsampling == 1)

bit(1) gainmap\_chroma\_is\_vertically\_centered;

bit(1) gainmap\_float\_flag;

if (gainmap\_float\_flag)

bit(2) gainmap\_bit\_depth\_log2\_minus4;

else {

bit(1) gainmap\_high\_bit\_depth\_flag;

if (gainmap\_high\_bit\_depth\_flag)

bit(3) gainmap\_bit\_depth\_minus9;

}

bit(1) tmap\_icc\_flag;

bit(1) tmap\_explicit\_cicp\_flag;

if (tmap\_explicit\_cicp\_flag) {

bit(8) tmap\_colour\_primaries;

bit(8) tmap\_transfer\_characteristics;

bit(8) tmap\_matrix\_coefficients;

bit(1) tmap\_full\_range\_flag;

}

else {

tmap\_colour\_primaries = 1;

tmap\_transfer\_characteristics = 13;

tmap\_matrix\_coefficients = 6;

tmap\_full\_range\_flag = 1;

}

}

bit(1) clli\_flag;

bit(1) mdcv\_flag;

bit(1) cclv\_flag;

bit(1) amve\_flag;

bit(1) reve\_flag;

bit(1) ndwt\_flag;

if (clli\_flag)

ContentLightLevel clli;

if (mdcv\_flag)

MasteringDisplayColourVolume mdcv;

if (cclv\_flag)

ContentColourVolume cclv;

if (amve\_flag)

AmbientViewingEnvironment amve;

if (reve\_flag)

ReferenceViewingEnvironment reve;

if (ndwt\_flag)

NominalDiffuseWhite ndwt;

if (gainmap\_flag) {

bit(1) tmap\_clli\_flag;

bit(1) tmap\_mdcv\_flag;

bit(1) tmap\_cclv\_flag;

bit(1) tmap\_amve\_flag;

bit(1) tmap\_reve\_flag;

bit(1) tmap\_ndwt\_flag;

if (tmap\_clli\_flag)

ContentLightLevel tmap\_clli;

if (tmap\_mdcv\_flag)

MasteringDisplayColourVolume tmap\_mdcv;

if (tmap\_cclv\_flag)

ContentColourVolume tmap\_cclv;

if (tmap\_amve\_flag)

AmbientViewingEnvironment tmap\_amve;

if (tmap\_reve\_flag)

ReferenceViewingEnvironment tmap\_reve;

if (tmap\_ndwt\_flag)

NominalDiffuseWhite tmap\_ndwt;

}

}

// Chunk sizes

if (icc\_flag || exif\_flag || xmp\_flag || (hdr\_flag && gainmap\_flag))

bit(1) few\_metadata\_bytes\_flag;

bit(1) few\_codec\_config\_bytes\_flag;

bit(1) few\_item\_data\_bytes\_flag;

if (icc\_flag)

unsigned int(few\_metadata\_bytes\_flag ? 10 : 20) icc\_data\_size\_minus1;

if (hdr\_flag && gainmap\_flag && tmap\_icc\_flag)

unsigned int(few\_metadata\_bytes\_flag ? 10 : 20) tmap\_icc\_data\_size\_minus1;

if (hdr\_flag && gainmap\_flag)

unsigned int(few\_metadata\_bytes\_flag ? 10 : 20) gainmap\_metadata\_size;

if (hdr\_flag && gainmap\_flag)

unsigned int(few\_item\_data\_bytes\_flag ? 15 : 28) gainmap\_item\_data\_size;

if (hdr\_flag && gainmap\_flag && gainmap\_item\_data\_size > 0)

unsigned int(few\_codec\_config\_bytes\_flag ? 3 : 12) gainmap\_item\_codec\_config\_size;

unsigned int(few\_codec\_config\_bytes\_flag ? 3 : 12) main\_item\_codec\_config\_size;

unsigned int(few\_item\_data\_bytes\_flag ? 15 : 28) main\_item\_data\_size\_minus1;

if (alpha\_flag)

unsigned int(few\_item\_data\_bytes\_flag ? 15 : 28) alpha\_item\_data\_size;

if (alpha\_flag && alpha\_item\_data\_size > 0)

unsigned int(few\_codec\_config\_bytes\_flag ? 3 : 12) alpha\_item\_codec\_config\_size;

if (exif\_flag)

unsigned int(few\_metadata\_bytes\_flag ? 10 : 20) exif\_data\_size\_minus1;

if (xmp\_flag)

unsigned int(few\_metadata\_bytes\_flag ? 10 : 20) xmp\_data\_size\_minus1;

trailing\_bits(); // bit padding till byte alignment

// Chunks

if (alpha\_flag && alpha\_item\_data\_size > 0 && alpha\_item\_codec\_config\_size > 0)

unsigned int(8) alpha\_item\_codec\_config[alpha\_item\_codec\_config\_size];

if (hdr\_flag && gainmap\_flag && gainmap\_item\_codec\_config\_size > 0)

unsigned int(8) gainmap\_item\_codec\_config[gainmap\_item\_codec\_config\_size];

if (main\_item\_codec\_config\_size > 0)

unsigned int(8) main\_item\_codec\_config[main\_item\_codec\_config\_size];

if (icc\_flag)

unsigned int(8) icc\_data[icc\_data\_size\_minus1 + 1];

if (hdr\_flag && gainmap\_flag && tmap\_icc\_flag)

unsigned int(8) tmap\_icc\_data[tmap\_icc\_data\_size\_minus1 + 1];

if (hdr\_flag && gainmap\_flag && gainmap\_metadata\_size > 0)

unsigned int(8) gainmap\_metadata[gainmap\_metadata\_size];

if (alpha\_flag && alpha\_item\_data\_size > 0)

unsigned int(8) alpha\_item\_data[alpha\_item\_data\_size];

if (hdr\_flag && gainmap\_flag && gainmap\_item\_data\_size > 0)

unsigned int(8) gainmap\_item\_data[gainmap\_item\_data\_size];

unsigned int(8) main\_item\_data[main\_item\_data\_size\_minus1 + 1];

if (exif\_flag)

unsigned int(8) exif\_data[exif\_data\_size\_minus1 + 1];

if (xmp\_flag)

unsigned int(8) xmp\_data[xmp\_data\_size\_minus1 + 1];

}

EDITORS NOTE:  
HDR-related specifications such as ISO 21496-1 were not finalized at the time of writing this document. The changes in this document should be kept as close to the up-to-date HDR-related specifications as possible.

#### O.3.3 Semantics

version: specifies the version of the MinimizedImageBox. The version shall be set to 0 in this version of this document.

small\_dimensions\_flag: if set to 0, the length of the fields signaled among width\_minus1, height\_minus1, gainmap\_width\_minus1 and gainmap\_height\_minus1 is 7 bits; otherwise, it is 15 bits.

width\_minus1: plus 1 specifies the width of the reconstructed image in pixels.

height\_minus1: plus 1 specifies the height of the reconstructed image in pixels.

orientation\_minus1: plus 1 specifies the Exif orientation value as defined in JEITA CP-3451E section 4.6.4.A "Orientation".

icc\_flag: equal to 1 indicates that the main image is associated with an ICC profile as defined in ISO 15076-1 or ICC.1[13].

exif\_flag: equal to 1 indicates the presence of Exif metadata.

xmp\_flag: equal to 1 indicates the presence of XMP metadata.

full\_range\_flag: is a binary value representing the VideoFullRangeFlag as defined in Rec. ITU-T H.273 | ISO/IEC 23091-2. This signaling applies exclusively to the main image.

chroma\_subsampling: 0 specifies that there is exactly one channel of coded color samples (monochrome), otherwise there are exactly three channels of coded color samples (one luma channel and two chroma channels). A value of 1 specifies that these chroma channels are subsampled both horizontally and vertically by a factor 2 (i.e. 4:2:0). A value of 2 specifies that these chroma channels are subsampled by a factor 2 horizontally (i.e. 4:2:2). A value of 3 specifies that there is no subsampling of these chroma channels (i.e. 4:4:4).

chroma\_is\_horizontally\_centered: 0 specifies that the chroma samples of the main image are co-located horizontally with the luma samples of the main image, otherwise they are horizontally centered between the luma samples of the main image. 0 unless chroma\_subsampling is 1 or 2.

chroma\_is\_vertically\_centered: 0 specifies that the chroma samples of the main image are co-located vertically with the luma samples of the main image, otherwise they are vertically centered between the luma samples of the main image. 0 unless chroma\_subsampling is 1.

float\_flag: specifies the format of the pixel values of the reconstructed main and alpha image items as the channel\_format values, as specified in PixelInformationProperty with version 1 in clause 6.5.6.

bit\_depth\_log2\_minus4: specifies the format of floating-point numbers used for the pixel values of the reconstructed main and alpha image items. The values 0, 1, and 2 respectively correspond to the bits\_per\_channel values 16, 32 and 64, as specified in PixelInformationProperty with version 1 in clause 6.5.6. Other values are reserved. When float\_flag is set to 0, the value is undefined.

high\_bit\_depth\_flag: 0 specifies that the number of bits per channel for the pixel values of the reconstructed main and alpha image items, as specified in PixelInformationProperty with version 1 in clause 6.5.6, is 8. Otherwise bit\_depth\_minus9 is signaled. When float\_flag is set to 1, the value is undefined.

bit\_depth\_minus9: specifies the number of bits, minus nine, per channel for the pixel values of the reconstructed main and alpha image items, as specified in PixelInformationProperty with version 1 in clause 6.5.6. When high\_bit\_depth\_flag is set to 0 or float\_flag is set to 1, the value is undefined.

alpha\_flag: 0 specifies that the image is opaque. Otherwise the image has an alpha layer, whether the codec has native translucency support or an alpha auxiliary image item is used.

alpha\_is\_premultiplied: when set to 1 specifies that the color channels are pre-multiplied by the alpha channel, otherwise the color channels are not pre-multiplied. Ignored if alpha\_flag is 0.

explicit\_cicp\_flag: equal to 0 indicates the sRGB on-screen colors as the values of ColourPrimaries and TransferCharacteristics, as defined in Rec. ITU-T H.273 | ISO/IEC 23091-2, respectively set to 1 and 13 if icc\_flag is 0, and to 2 and 2 otherwise. 0 specifies sRGB on-screen colors as the value of MatrixCoefficients, as defined in Rec. ITU-T H.273 | ISO/IEC 23091-2, set to 2 if chroma\_subsampling is 0, and to 6 otherwise. When the value is equal to 1 it indicates that these values are signaled explicitly.

colour\_primaries: carries a ColourPrimaries value as defined in Rec. ITU-T H.273 | ISO/IEC 23091-2 for the main image.

transfer\_characteristics: carries a TransferCharacteristics value as defined in Rec. ITU-T H.273 | ISO/IEC 23091-2 for the main image.

matrix\_coefficients: carries a MatrixCoefficients value as defined in Rec. ITU-T H.273 | ISO/IEC 23091-2 for the main image.

explicit\_codec\_types\_flag: 0 specifies that the minor\_version of the FileTypeBox carries a brand defining a single coded image item type and a single codec configuration property box type. Shall be set to 1 otherwise.

infe\_type: carries the coded image item type. Corresponds to the item\_type field of the version 2 of the ItemInfoEntry box. Defined by the brand carried by the minor\_version of the FileTypeBox if explicit\_codec\_types\_flag is 0.

codec\_config\_type: carries the codec configuration property box type. Defined by the brand carried by the minor\_version of the FileTypeBox if explicit\_codec\_types\_flag is 0.

hdr\_flag: 0 specifies that the image is SDR and has no associated HDR-related signaling. Otherwise the image is either SDR with a SDR-to-HDR gain map, or HDR with an optional HDR-to-SDR gain map.

gainmap\_flag: 0 specifies that the file has no tone-mapped image and no associated HDR-related ISO 21496-1 gain map. Otherwise the file contains a tone-mapped image and is associated with a gain map, whether the codec has native gain map support or a separate **gain map image item** is used. 0 if hdr\_flag is 0.

gainmap\_width\_minus1: carries the width minus one of the gain map image in pixels.

gainmap\_height\_minus1: carries the height minus one of the gain map image in pixels.

gainmap\_matrix\_coefficients: carries a MatrixCoefficients value as defined in Rec. ITU-T H.273 | ISO/IEC 23091-2 for the gain map image.

gainmap\_full\_range\_flag: carries a VideoFullRangeFlag as defined in Rec. ITU-T H.273 | ISO/IEC 23091-2 for the gain map image.

gainmap\_chroma\_subsampling: 0 specifies that there is exactly one channel of coded gain map samples (monochrome), otherwise there are exactly three channels of coded gain map samples (one luma channel and two chroma channels). A value of 1 specifies that these chroma channels are subsampled both horizontally and vertically by a factor 2 (i.e. 4:2:0). A value of 2 specifies that these chroma channels are subsampled by a factor 2 horizontally (i.e. 4:2:2). A value of 3 specifies that there is no subsampling of these chroma channels (i.e. 4:4:4).

gainmap\_chroma\_is\_horizontally\_centered: 0 specifies that the chroma samples of the gain map image are co-located horizontally with the luma samples of the gain map image, otherwise they are horizontally centered between the luma samples of the gain map image. Ignored unless gainmap\_chroma\_subsampling is 1 or 2.

gainmap\_chroma\_is\_vertically\_centered: 0 specifies that the chroma samples of the gain map image are co-located vertically with the luma samples of the gain map image, otherwise they are vertically centered between the luma samples of the gain map image. Ignored unless gainmap\_chroma\_subsampling is 1.

gainmap\_float\_flag: specifies the format of the pixel values of the reconstructed gain map image item as the channel\_format values, as specified in PixelInformationProperty with version 1 in clause 6.5.6.

gainmap\_bit\_depth\_log2\_minus4: specifies the format of floating-point numbers used for the pixel values of the reconstructed gain map image item. The values 0, 1, and 2 respectively correspond to the bits\_per\_channel values 16, 32 and 64, as specified in PixelInformationProperty with version 1 in clause 6.5.6. Other values are reserved. When float\_flag is set to 0, the value is undefined.

gainmap\_high\_bit\_depth\_flag: 0 specifies that the number of bits per channel for the pixel values of the reconstructed gain map image item, as specified in PixelInformationProperty with version 1 in clause 6.5.6, is 8. Otherwise gainmap\_bit\_depth\_minus9 is signaled. When gainmap\_float\_flag is set to 1, the value is undefined.

gainmap\_bit\_depth\_minus9: specifies the number of bits, minus nine, per channel for the pixel values of the reconstructed gain map image item, as specified in PixelInformationProperty with version 1 in clause 6.5.6. When high\_bit\_depth\_flag is set to 0 or float\_flag is set to 1, the value is undefined.

tmap\_icc\_flag: if 1, specifies that the tone-mapped image is associated with an **ICC profile** as defined in ISO 15076-1 or ICC.1[23]. 0 if gainmap\_flag is 0.

tmap\_explicit\_cicp\_flag: 0 specifies sRGB on-screen colors as the values of ColourPrimaries, TransferCharacteristics and MatrixCoefficients, as defined in Rec. ITU-T H.273 | ISO/IEC 23091-2, associated with the tone-mapped image, set to 1, 13 and 6, respectively. Otherwise these values are signaled explicitly.

tmap\_colour\_primaries: carries a ColourPrimaries value as defined in Rec. ITU-T H.273 | ISO/IEC 23091-2 for the tone-mapped image.

tmap\_transfer\_characteristics: carries a TransferCharacteristics value as defined in Rec. ITU-T H.273 | ISO/IEC 23091-2 for the tone-mapped image.

tmap\_matrix\_coefficients: carries a MatrixCoefficients value as defined in Rec. ITU-T H.273 | ISO/IEC 23091-2 for the tone-mapped image.

tmap\_full\_range\_flag: carries a VideoFullRangeFlag as defined in Rec. ITU-T H.273 | ISO/IEC 23091-2 for the tone-mapped image. Set to 1 if tmap\_explicit\_cicp\_flag is 0.clli\_flaggainmap\_flag

clli\_flag: 1 specifies that there is signaling for ContentLightLevel attached to the main image. Otherwise no such signaling is present. 0 if hdr\_flag is 0.

mdcv\_flag: 1 specifies that there is signaling for MasteringDisplayColourVolume attached to the main image. Otherwise no such signaling is present. 0 if hdr\_flag is 0.

cclv\_flag: 1 specifies that there is signaling for ContentColourVolume attached to the main image. Otherwise no such signaling is present. 0 if hdr\_flag is 0.

amve\_flag: 1 specifies that there is signaling for AmbientViewingEnvironment attached to the main image. Otherwise no such signaling is present. 0 if hdr\_flag is 0.

reve\_flag: 1 specifies that there is signaling for ReferenceViewingEnvironment attached to the main image. Otherwise no such signaling is present. 0 if hdr\_flag is 0.

ndwt\_flag: 1 specifies that there is signaling for NominalDiffuseWhite attached to the main image. Otherwise no such signaling is present. 0 if hdr\_flag is 0.

tmap\_clli\_flag: 1 specifies that there is signaling for ContentLightLevel attached to the tone-mapped image. Otherwise no such signaling is present. 0 if gainmap\_flag is 0.

tmap\_mdcv\_flag: 1 specifies that there is signaling for MasteringDisplayColourVolume attached to the tone-mapped image. Otherwise no such signaling is present. 0 if gainmap\_flag is 0.

tmap\_cclv\_flag: 1 specifies that there is signaling for ContentColourVolume attached to the tone-mapped image. Otherwise no such signaling is present. 0 if gainmap\_flag is 0.

tmap\_amve\_flag: 1 specifies that there is signaling for AmbientViewingEnvironment attached to the tone-mapped image. Otherwise no such signaling is present. 0 if gainmap\_flag is 0.

tmap\_reve\_flag: 1 specifies that there is signaling for ReferenceViewingEnvironment attached to the tone-mapped image. Otherwise no such signaling is present. 0 if gainmap\_flag is 0.

tmap\_ndwt\_flag: 1 specifies that there is signaling for NominalDiffuseWhite attached to the tone-mapped image. Otherwise no such signaling is present. 0 if gainmap\_flag is 0.

clli: The box body of the ContentLightLevelBox as defined in ISO/IEC 14496-12 attached to the main image. Only present if clli\_flag is 1.

mdcv: The box body of the MasteringDisplayColourVolumeBox as defined in ISO/IEC 14496-12 attached to the main image. Only present if clli\_flag\_mdcv is 1.

cclv: The box body of the ContentColourVolumeBox as defined in ISO/IEC 14496-12 attached to the main image. Only present if clli\_flag\_cclv is 1.

amve: The box body of the AmbientViewingEnvironmentBox as defined in ISO/IEC 14496-12 attached to the main image. Only present if clli\_flag\_amve is 1.

reve: The box body of the ReferenceViewingEnvironmentBox attached to the main image. Only present if clli\_flag\_reve is 1.

ndwt: The box body of the NominalDiffuseWhiteBox attached to the main image. Only present if clli\_flag\_ndwt is 1.

tmap\_clli: The box body of the ContentLightLevelBox as defined in ISO/IEC 14496-12 attached to the tone-mapped image. Only present gainmap\_flagif tmap\_clli\_flag is 1.

tmap\_mdcv: The box body of the MasteringDisplayColourVolumeBox as defined in ISO/IEC 14496-12 attached to the tone-mapped image. Only present gainmap\_flagif tmap\_clli\_flag\_mdcv is 1.

tmap\_cclv: The box body of the ContentColourVolumeBox as defined in ISO/IEC 14496-12 attached to the tone-mapped image. Only present gainmap\_flagif tmap\_clli\_flag\_cclv is 1.

tmap\_amve: The box body of the AmbientViewingEnvironmentBox as defined in ISO/IEC 14496-12 attached to the tone-mapped image. Only present gainmap\_flagif tmap\_clli\_flag\_amve is 1.

tmap\_reve: The box body of the ReferenceViewingEnvironmentBox attached to the tone-mapped image. Only present gainmap\_flagif tmap\_clli\_flag\_reve is 1.

tmap\_ndwt: The box body of the NominalDiffuseWhiteBox attached to the tone-mapped image. Only present gainmap\_flagif tmap\_clli\_flag\_ndwt is 1.

few\_metadata\_bytes\_flag: 0 specifies that the length of the signaled fields among icc\_data\_size\_minus1, tmap\_icc\_data\_size\_minus1, gainmap\_metadata\_size, exif\_data\_size\_minus1 and xmp\_data\_size\_minus1 is 10 bits, otherwise 20 bits. Undefined unless one of these fields is signaled.

few\_codec\_config\_bytes\_flag: 0 specifies that the length of the signaled fields among gainmap\_item\_codec\_config\_size, main\_item\_codec\_config\_size and alpha\_item\_codec\_config\_size is 3 bits, otherwise 12 bits.

few\_item\_data\_bytes\_flag: 0 specifies that the length of the signaled fields among gainmap\_item\_data\_size, main\_item\_data\_size\_minus1 and alpha\_item\_data\_size is 15 bits, otherwise 28 bits.

icc\_data\_size\_minus1: carries the size minus one in bytes of the **ICC profile** as defined in ISO 15076-1 or ICC.1[23], associated with the main image. Undefined if icc\_flag is 0.

tmap\_icc\_data\_size\_minus1: carries the size minus one in bytes of the **ICC profile** as defined in ISO 15076-1 or ICC.1[23], associated with the tone-mapped image. Undefined if tmap\_icc\_flag is 0.

gainmap\_metadata\_size: carries the size of the gain map metadata. 0 if gainmap\_flag is 0.

gainmap\_item\_data\_size: carries the size of the coded sample data for the HDR-related gain map image item in bytes. If gainmap\_flag is set to 1, a size of 0 is reserved for future use. 0 if gainmap\_flag is 0.

gainmap\_item\_codec\_config\_size: carries the size of the codec configuration for the gain map auxiliary image item in bytes. The value 0 specifies that the codec does not need any configuration data for the gain map. 0 if gainmap\_item\_data\_size is 0.

main\_item\_codec\_config\_size: carries the size of the codec configuration for the main image item in bytes.

main\_item\_data\_size\_minus1: carries the size minus one of the coded sample data for the main image item in bytes.

alpha\_item\_data\_size: carries the size of the coded sample data for the alpha auxiliary image item in bytes. If alpha\_flag is set to 1, the value 0 specifies that the codec has native translucency support and that the alpha samples are coded alongside the color samples in the main\_item\_data chunk. 0 if alpha\_flag is 0.

alpha\_item\_codec\_config\_size: carries the size of the codec configuration for the alpha auxiliary image item in bytes. The value 0 specifies that the codec does not need any configuration data for alpha. 0 if alpha\_item\_data\_size is 0.

exif\_data\_size\_minus1: specifies the size minus one of the **Exif** metadata in bytes. -1 if exif\_flag is 0.

xmp\_data\_size\_minus1: specifies the size minus one of the **XMP** metadata in bytes. -1 if xmp\_flag is 0.

trailing\_bits: padding bits to ensure payloads are 8-bit aligned. Shall be 0.

alpha\_item\_codec\_config: carries the optional alpha image codec configuration data. When alpha\_item\_codec\_config\_size is 0, alpha\_item\_codec\_config is not present.

gainmap\_item\_codec\_config: carries the HDR-related gain map image item codec configuration data. When gainmap\_item\_codec\_config\_size is 0, gainmap\_item\_codec\_config is not present.

main\_item\_codec\_config: carries the main image item codec configuration data. When main\_item\_codec\_config\_size is 0, main\_item\_codec\_config is not present.

icc\_data: carries the **ICC profile** data of the main image as defined in ISO 15076-1 or ICC.1[23]. When icc\_flag is 0, icc\_data is not present.

tmap\_icc\_data: carries the **ICC profile** data of the optional HDR-related tone-mapped image as defined in ISO 15076-1 or ICC.1[23]. When tmap\_icc\_flag is 0, tmap\_icc\_data is not present.

gainmap\_metadata: Gain map metadata as defined by the GainMapMetadata struct in ISO 21496-1. Not present if gainmap\_metadata\_size is 0.

alpha\_item\_data: carries the coded sample data of the optional alpha image. When alpha\_item\_data\_size is 0, alpha\_item\_data is not present.

gainmap\_item\_data: carries the coded sample data of the optional gain map image. When gainmap\_item\_data\_size is 0, gainmap\_item\_data is not present.

main\_item\_data: carries the coded sample data of the main image.

exif\_data: specifies the optional **Exif** metadata. When exif\_flag is set to 0, exif\_data is not present.

xmp\_data: specifies the optional **XMP** metadata. When xmp\_flag is set to 0, xmp\_data is not present.

### O.4 Equivalence with MetaBox

A MinimizedImageBox has a one-to-one mapping to a MetaBox. Readers shall treat a MinimizedImageBox as if it were the equivalent MetaBox that is transformed from MinimizedImageBox as specified in this subclause. When a reader encounters the MinimizedImageBox, it will create the equivalent MetaBox in memory and populate its contents based on the parsed contents of the MinimizedImageBox.

File writers can choose between two formats: they can either write a traditional image file based on the MetaBox or opt for a low-overhead image file format based on MinimizedImageBox.

The equivalent MetaBox has its version and flags set to 0 and the sub-boxes described by the following subclauses.

#### O.4.1 HandlerBox

The equivalent MetaBox shall have a HandlerBox with handler\_type equal to 'pict'.

#### O.4.2 PrimaryItemBox

The equivalent MetaBox shall have a PrimaryItemBox with item\_ID set to 1.

#### O.4.3 ItemInfoBox

The equivalent MetaBox shall have an ItemInfoBox containing the following entries:

* ItemInfoEntry of version 2 with flags set to 0, item\_ID set to 1 and item\_type set to infe\_type. All other fields are set to null or 0 as appropriate.
* If alpha\_item\_data\_size is not 0, ItemInfoEntry of version 2 with flags set to 1, item\_ID set to 2 and item\_type set to infe\_type. All other fields are set to null or 0 as appropriate.
* If gainmap\_flag is 1, ItemInfoEntry of version 2 with flags set to 0, item\_ID set to 3 and item\_type set to 'tmap'. All other fields are set to null or 0 as appropriate.
* If gainmap\_item\_data\_size is not 0, ItemInfoEntry of version 2 with flags set to 1, item\_ID set to 4 and item\_type set to infe\_type. All other fields are set to null or 0 as appropriate.
* If exif\_flag is 1, ItemInfoEntry of version 2 with flags set to 1, item\_ID set to 6 and item\_type set to 'Exif'. All other fields are set to null or 0 as appropriate.
* If xmp\_flag is 1, ItemInfoEntry of version 2 with flags set to 1, item\_ID set to 7 and item\_type set to 'mime' and content\_type set to 'application/rdf+xml'. All other fields are set to null or 0 as appropriate.

#### O.4.4 ItemReferenceBox

The ItemReferenceBox is populated with the following entries:

* If ItemInfoBox has an entry for item\_ID 2: Item type reference with referenceType set to 'auxl', from\_item\_ID set to 2, reference\_count set to 1 and to\_item\_ID set to 1.
* If ItemInfoBox has an entry for item\_ID 2 and alpha\_is\_premultiplied is set to 1: Item type reference with referenceType set to 'prem', from\_item\_ID set to 1, reference\_count set to 1 and to\_item\_ID set to 2.
* If gainmap\_flag is 1 and ItemInfoBox has an entry for item\_ID 4: item type reference with referenceType set to 'dimg', from\_item\_ID set to 3, reference\_count set to 2 and the following to\_item\_ID entries: 1, 4.
* If gainmap\_flag is 1 and ItemInfoBox has zero entry for item\_ID 4: item type reference with referenceType set to 'dimg', from\_item\_ID set to 3, reference\_count set to 1 and to\_item\_ID set to 1.

EDITORS NOTE:  
‘tmap’ items currently require 2 inputs. In a future where a codec supports coding the gain map interleaved with the base image in the same coded payload, a ‘tmap’ item would presumably only have one input. We’re leaving the option open here for now for that future.

* If ItemInfoBox has an entry for item\_ID 6: item type reference with referenceType set to 'cdsc', from\_item\_ID set to 6, reference\_count set to 1 and to\_item\_ID set to 1.
* If ItemInfoBox has an entry for item\_ID 7: item type reference with referenceType set to 'cdsc', from\_item\_ID set to 7, reference\_count set to 1 and to\_item\_ID set to 1.

If the resulting ItemReferenceBox contains at least one entry, it shall be added to the equivalent MetaBox. An empty ItemReferenceBox shall be ignored.

#### O.4.5 EntityToGroupBox

If gainmap\_flag is 1, the equivalent MetaBox shall have a GroupsListBox containing a single sub-box. That sub-box is an EntityToGroupBox with grouping\_type set to 'altr', version set 0, flags set to 0, group\_id set to 5, num\_entities\_in\_group set to 2 and the following entity\_id entries: 3, 1.

#### O.4.6 ItemPropertiesBox

The equivalent MetaBox shall have an ItemPropertiesBox containing an ItemPropertyContainerBox and an ItemPropertyAssociationBox.

The ItemPropertyContainerBox shall have 32 entries as listed below. Any entry for which the condition is not true is replaced with a FreeSpaceBox.

|  |  |  |
| --- | --- | --- |
| **Entry** | **Condition** | **Contents** |
| 1 | main\_item\_codec\_config\_size is not 0 | Item property with the type set to codec\_config\_type and with contents from main\_item\_codec\_config. |
| 2 | true | ImageSpatialExtentsProperty with image\_width set to width\_minus1+1 and image\_height set to height\_minus1+1. |
| 3 | true | PixelInformationProperty version 1 with field values defined in section O.4.7 of this document. |
| 4 | true | ColourInformationBox with colour\_type set to 'nclx' and colour\_primaries, transfer\_characteristics, matrix\_coefficients and full\_range\_flag set to the values from the MinimizedImageBox. |
| 5 | icc\_flag is 1 | ColourInformationBox with the colour\_type set to 'prof' and with ICC\_profile contents being icc\_data. |
| 6 | alpha\_item\_codec\_config\_size is not 0 | Item property with the type set to codec\_config\_type and with contents from alpha\_item\_codec\_config. |
| 7 | alpha\_item\_data\_size is not 0 | AuxiliaryTypeProperty with aux\_type set to 'urn:mpeg:mpegB:cicp:systems:auxiliary:alpha'. |
| 8 | alpha\_item\_data\_size is not 0 | PixelInformationProperty version 1 with field values defined in section O.4.7 of this document. |
| 9 | orientation is 2, 4, 5, 6 or 7 | ImageRotation property with angle set to 2, 1, 3, 1, 1, respectively |
| 10 | orientation is 1, 3, 4 or 6 | ImageMirror property with axis set to 1, 0, 0, 1, respectively |
| 11 | clli\_flag is 1 | ContentLightLevelBox with body set to clli |
| 12 | clli\_flag\_mdcv is 1 | MasteringDisplayColourVolumeBox with body set to mdcv |
| 13 | clli\_flag\_cclv is 1 | ContentColourVolumeBox with body set to cclv |
| 14 | clli\_flag\_amve is 1 | AmbientViewingEnvironmentBox with body set to amve |
| 15 | clli\_flag\_reve is 1 | ReferenceViewingEnvironmentBox with body set to reve and version and flags set to 0 |
| 16 | clli\_flag\_ndwt is 1 | NominalDiffuseWhiteBox with body set to ndwt and version and flags set to 0 |
| 17 | gainmap\_item\_codec\_config is not 0 | Item property with the type set to codec\_config\_type and with contents from gainmap\_item\_codec\_config. |
| 18 | gainmap\_item\_data\_size is not 0 | ImageSpatialExtentsProperty with image\_width set to gainmap\_width\_minus1 + 1 and image\_height set to gainmap\_height\_minus1 + 1. |
| 19 | gainmap\_item\_data\_size is not 0 | PixelInformationProperty version 1 with field values defined in section O.4.7 of this document. |
| 20 | gainmap\_item\_data\_size is not 0 | ColourInformationBox with colour\_type set to 'nclx' and colour\_primaries and transfer\_characteristics set to 2, matrix\_coefficients set to gainmap\_matrix\_coefficients and full\_range\_flag set to gainmap\_full\_range\_flag. |
| 21 | gainmap\_flag is 1 | ImageSpatialExtentsProperty equal to entry 2 if orientation is 0, 1, 2, 3 and with image\_width set to height\_minus1 + 1 and image\_height set to width\_minus1 + 1 for any other orientation. |
| 22 | gainmap\_flag is 1 and (tmap\_explicit\_cicp\_flag is 1 or tmap\_icc\_flag is 0) | ColourInformationBox with colour\_type set to 'nclx' and colour\_primaries, transfer\_characteristics, matrix\_coefficients and full\_range\_flag set to tmap\_colour\_primaries, tmap\_transfer\_characteristics, tmap\_matrix\_coefficients and tmap\_full\_range\_flag, respectively. |
| 23 | gainmap\_flag is 1 and tmap\_icc\_flag is 1 | ColourInformationBox with the colour\_type set to 'prof' and with ICC\_profile contents being tmap\_icc\_data. |
| 24 | gainmap\_flagtmap\_clli\_flag is 1 | ContentLightLevelBox with body set to tmap\_clli |
| 25 | gainmap\_flagtmap\_clli\_flag\_mdcv is 1 | MasteringDisplayColourVolumeBox with body set to tmap\_mdcv |
| 26 | gainmap\_flagtmap\_clli\_flag\_cclv is 1 | ContentColourVolumeBox with body set to tmap\_cclv |
| 27 | gainmap\_flagtmap\_clli\_flag\_amve is 1 | AmbientViewingEnvironmentBox with body set to tmap\_amve |
| 28 | gainmap\_flagtmap\_clli\_flag\_reve is 1 | ReferenceViewingEnvironmentBox with body set to tmap\_reve and version and flags set to 0 |
| 29 | gainmap\_flagtmap\_clli\_flag\_ndwt is 1 | NominalDiffuseWhiteBox with body set to tmap\_ndwt and version and flags set to 0 |
| 30 | false | Reserved |
| 31 | false | Reserved |
| 32 | false | Reserved |

The ItemPropertyAssociationBox shall have the entries below. Any association to a FreeSpaceBox shall be dropped.

* Item 1 shall be associated with ItemPropertyContainerBox entries:
  + 1, essential
  + 2, non-essential
  + 3, non-essential
  + 4, essential
  + 5, essential
  + 9, essential
  + 10, essential
* If hdr\_flag is 1, item 1 shall be associated with ItemPropertyContainerBox entries:
  + 11, non-essential
  + 12, non-essential
  + 13, non-essential
  + 14, non-essential
  + 15, non-essential
  + 16, non-essential
* If alpha\_item\_data\_size is not 0, item 2 shall be associated with ItemPropertyContainerBox entries:
  + 6, essential
  + 2, non-essential
  + 7, essential
  + 8, non-essential
  + 9, essential
  + 10, essential
* If gainmap\_flag is 1, item 3 shall be associated with ItemPropertyContainerBox entries:
  + 21, non-essential
  + 22, essential
  + 23, essential
* If hdr\_flag is 1 and gainmap\_flag is 1, item 3 shall be associated with ItemPropertyContainerBox entries:
  + 24, non-essential
  + 25, non-essential
  + 26, non-essential
  + 27, non-essential
  + 28, non-essential
  + 29, non-essential
* If gainmap\_item\_data\_size is not 0, item 4 shall be associated with ItemPropertyContainerBox entries:
  + 17, essential

18, non-essential

* + 19, non-essential
  + 20, non-essential
  + 9, essential
  + 10, essential

#### O.4.7 PixelInformationProperty

##### O.4.7.1 Reconstruction

The various PixelInformationProperty boxes associated with the image items in the file are reconstructed given the arguments

* main\_components
* alpha\_flag
* alpha\_is\_premultiplied
* subsampling
* chroma\_is\_horizontally\_centered
* chroma\_is\_vertically\_centered
* float\_flag
* bit\_depth\_log2\_minus4
* high\_bit\_depth\_flag
* bit\_depth\_minus9

Reconstruction happens as follows:

* version set to 1.
* px\_flags set to:
  + 7 if alpha\_flag and alpha\_is\_premultiplied are both true
  + 5 if alpha\_flag is true and alpha\_is\_premultiplied is false
  + 4 otherwise
* num\_channels set to:
  + main\_components if alpha\_flag is false
  + main\_components + 1 if alpha\_flag is true
* each bits\_per\_channel entry set to 2bit\_depth\_log2\_minus4+4 if float\_flag is 1, or to 8 if high\_bit\_depth\_flag is 0, or to bit\_depth\_minus9+9 otherwise.
* channel\_label\_flag is set to 0 for all entries
* the following entries for channel subsampling and formats
  + the following entry if alpha\_flag is true (skipped if false)
    - channel\_idc set to 1
    - channel\_format set to float\_flag,
    - subsampling\_type set to 0
    - subsampling\_location set to 0
  + the following entry if main\_components > 0 (skipped otherwise)
    - channel\_idc set to 0
    - channel\_format set to float\_flag,
    - subsampling\_type set to 0
    - subsampling\_location set to 0
  + the following entry, repeated twice, if main\_components > 1 (skipped otherwise)
    - channel\_idc set to 0
    - channel\_format set to float\_flag,
    - subsampling\_type set to
      * 2 if subsampling is 1
      * 1 if subsampling is 2
      * 0 if subsampling is 3
    - subsampling\_location set to
      * 1 if chroma\_is\_horizontally\_centered and chroma\_is\_vertically\_centered are both true
      * 3 if chroma\_is\_horizontally\_centered is true and chroma\_is\_vertically\_centered is false
      * 0 if chroma\_is\_horizontally\_centered is false and chroma\_is\_vertically\_centered is true
      * 2 if chroma\_is\_horizontally\_centered and chroma\_is\_vertically\_centered are both false

##### O.4.7.2 Main image PixelInformationProperty

The PixelInformationProperty associated with the main image item is reconstructed as described in section O.4.7.1 with the following arguments:

* main\_components = 1 if chroma\_subsampling is 0, else 3
* alpha\_flag = alpha\_flag is 1 and alpha\_item\_data\_size is 0
* alpha\_premultiplied\_flag = alpha\_is\_premultiplied
* subsampling = chroma\_subsampling
* chroma\_is\_horizontally\_centered = chroma\_is\_horizontally\_centered
* chroma\_is\_vertically\_centered = chroma\_is\_vertically\_centered
* float\_flag = float\_flag
* bit\_depth\_log2\_minus4 = bit\_depth\_log2\_minus4
* high\_bit\_depth\_flag = high\_bit\_depth\_flag
* bit\_depth\_minus9 = bit\_depth\_minus9

##### O.4.7.3 Alpha auxiliary image PixelInformationProperty

If alpha\_item\_data\_size is not 0, the PixelInformationProperty associated with the alpha auxiliary image item is reconstructed as described in section O.4.7.1 with the following arguments:

* main\_components = 0
* alpha\_flag = 1
* alpha\_premultiplied\_flag = 0
* subsampling = 0
* chroma\_is\_horizontally\_centered = 0
* chroma\_is\_vertically\_centered = 0
* float\_flag = float\_flag
* bit\_depth\_log2\_minus4 = bit\_depth\_log2\_minus4
* high\_bit\_depth\_flag = high\_bit\_depth\_flag
* bit\_depth\_minus9 = bit\_depth\_minus9

##### O.4.7.4 Gain map image PixelInformationProperty

If gainmap\_flag is not 0, the PixelInformationProperty associated with the gain map image item is reconstructed as described in section O.4.7.1 with the following arguments:

* main\_components = 1 if gainmap\_chroma\_subsampling is 0, else 3,
* alpha\_flag = 0
* alpha\_premultiplied\_flag set to 0,
* subsampling = gainmap\_chroma\_subsampling
* chroma\_is\_horizontally\_centered = gainmap\_chroma\_is\_horizontally\_centered
* chroma\_is\_vertically\_centered = gainmap\_chroma\_is\_vertically\_centered
* float\_flag = gainmap\_float\_flag
* bit\_depth\_log2\_minus4 = gainmap\_bit\_depth\_log2\_minus4
* high\_bit\_depth\_flag = gainmap\_high\_bit\_depth\_flag
* bit\_depth\_minus9 = gainmap\_bit\_depth\_minus9

EDITORS NOTE:  
The text for the ‘tmap’ item recommends it being associated with ‘pixi’ property that hints at the approximate bit-depth achieved after combining the base and gain map. We have these choices:

1. Simply ignore this recommendation
2. Code the number of extra bits achieved after reconstruction (i.e. if base is 8 bits, code value “2” to say you get 10 bits out). How does this work with floating point though?
3. Code the full set of fields required to create a v1 ‘pixi’.

#### O.4.8 ToneMapImage metadata

If gainmap\_flag is 1, tmap\_item\_data is defined as a data chunk of tmap\_item\_data\_size bytes, being gainmap\_metadata\_size+1 bytes, containing the ToneMapImage metadata as defined in section 6.6.2.4.2, with the following:

* version set to 0
* GainMapMetadata set to gainmap\_metadata

If gainmap\_flag is 0, tmap\_item\_data is defined as an empty chunk and tmap\_item\_data\_size as 0 byte.

#### O.4.9 ItemLocationBox

The equivalent MetaBox shall have an ItemLocationBox of version 1 or version 2 containing the following entries:

* item\_ID 1, with construction\_method set to 1, offset set to alpha\_item\_data\_size+tmap\_item\_data\_size+gainmap\_item\_data\_size and length set to main\_item\_data\_size\_minus1+1
* Optional item\_ID 2, with construction\_method set to 1, offset set to 0 and length set to alpha\_item\_data\_size
* Optional item\_ID 3, with construction\_method set to 1, offset set to alpha\_item\_data\_size, and length set to tmap\_item\_data\_size
* Optional item\_ID 4, with construction\_method set to 1, offset set to alpha\_item\_data\_size+tmap\_item\_data\_size, and length set to gainmap\_item\_data\_size
* Optional item\_ID 6, with construction\_method set to 1, offset set to main\_item\_data\_size\_minus1+1+alpha\_item\_data\_size+tmap\_item\_data\_size+gainmap\_item\_data\_size, and length set to exif\_data\_size\_minus1+1
* Optional item\_ID 7, with construction\_method set to 1, offset set to main\_item\_data\_size\_minus1+1+alpha\_item\_data\_size+tmap\_item\_data\_size+gainmap\_item\_data\_size+exif\_data\_size\_minus1+1, and length set to xmp\_data\_size\_minus1+1

#### O.4.10 ItemDataBox

The equivalent MetaBox shall have an ItemDataBox containing the non-empty chunks among alpha\_item\_data, tmap\_item\_data, gainmap\_item\_data, main\_item\_data, exif\_data and xmp\_data, concatenated in that order.

1. Add new Annex P

**Annex P**

(normative)

**Low-overhead image file format MIME type registration**

### P.1 General

The file extension and MIME type of a file deriving from the ISO base media file format usually reflect the major brand in the FileTypeBox. When the major brand indicates a brand related to Annex O (low-overhead image file format), the MIME type defined in this annex should be used. When such a brand is a major or compatible brand, this MIME type may also be used.

The registration below is the formal MIME type registration as recorded at IANA.

### P.2 Registration

MIME media type name: image

MIME subtype name: hif2

The semantics of the subtype are as follows:

hif2: High efficiency image file containing one image item  
 possibly with an alpha plane in any coding format. This subtype  
 name may be 'hif2' only if the file conforms to the  
 requirements of the 'mif3' brand.

Required parameters: none

Optional parameters:

profiles: Specified by RFC 6381 and its successors

codecs: Specified by RFC 6381 and its successors for files  
 conforming to specifications derived from ISO/IEC  
 14496-12. Note that for ISO-defined (MPEG) video  
 codecs, the format of a list item included in the  
 value of the codecs parameter is specified in  
 ISO/IEC 14496-15.

Encoding considerations: as for video/mp4

Security considerations: See section 4 of RFC 4337 and section 7 of  
 RFC 6381. This format does not supply integrity or  
 confidentiality protection and so they are applied externally  
 when needed. The security considerations of URLs are discussed  
 in RFC 3986.

Interoperability considerations: TBD

Published specification: ISO/IEC 23008-12

Applications: Multimedia, Imaging, Pictures

Fragment identifier considerations: Fragment identifiers are specified in Annex C of ISO/IEC 14496-12:2022

Additional information:

Magic number(s): none

File extension(s): hmg

Intended usage: Common

### P.3 Examples (informative)

Content-Type:

image/hif2; codecs="vvc1.1.L51.CQA"; profiles="mif3"

An image file with low-overhead HEIF container containing one VVC-coded image using Main 10 profile, Main Tier, Level 3.1.

Content-Type:

image/hif2; codecs="hvc1.A1.80.L93.B0"; profiles="mif3"

An image file with low-overhead HEIF container containing one HEVC-coded image.

1. Update Annex C

Replace the following text:

Fragment identifier considerations: Fragment identifiers are specified in Annex L of ISO/IEC 14496-12, available as a Publicly Available Standard at http://standards.iso.org/ittf/PubliclyAvailableStandards/index.html

with

Fragment identifier considerations: Fragment identifiers are specified in Annex C of ISO/IEC 14496-12:2022.

1. Add the following entries to Bibliography

[13] ICC.1:2001-04, *File format for color profiles*, International Color Consortium

1. Add the following entries to Clause 2 Normative References

ISO 15076-1, *Image technology colour management — Architecture, profile format and data structure — Part 1: Based on ICC.1:2010*