**ISO/IEC 23008-12:2024/AMD 2:2023(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: Support for tone map derived items and other improvements**

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.

Contents

[Foreword v](#_Toc166227934)

[1 Coding Constraints box related changes 1](#_Toc166227935)

[2 New tone-map derivation item 2](#_Toc166227936)

[6.6.2.4 Tone-map derivation 2](#_Toc166227939)

[6.6.2.4.1 Definition 2](#_Toc166227940)

[6.6.2.4.2 Syntax 4](#_Toc166227941)

[6.6.2.4.3 Semantics 4](#_Toc166227942)

[10.2.6 'tmap' brand 4](#_Toc166227943)

[J.6 Tone-map derivation 4](#_Toc166227944)

[3 New colour format enhancement derived item 7](#_Toc166227964)

[6.6.2.5 Colour format enhancement derivation 8](#_Toc166227965)

[6.6.2.5.1 Definition 8](#_Toc166227966)

[6.6.2.5.2 Syntax 9](#_Toc166227967)

[6.6.2.5.3 Semantics 9](#_Toc166227968)

[4 New constrained extents grid property 10](#_Toc166227969)

[6.5.37 Constrained Extents Grid Property 10](#_Toc166227970)

[6.5.37.1 Definition 10](#_Toc166227971)

[6.5.37.2 Syntax 11](#_Toc166227972)

[6.5.37.3 Semantics 11](#_Toc166227973)

[5 New stereo related signaling 11](#_Toc166227974)

[6.5.38 Disparity adjustment information 11](#_Toc166227975)

[6.5.38.1 Definition 11](#_Toc166227976)

[6.5.38.2 Syntax 12](#_Toc166227977)

[6.5.38.3 Semantics 12](#_Toc166227978)

[6.5.39 Stereo aggressors item property 12](#_Toc166227979)

[6.5.39.1 Definition 12](#_Toc166227980)

[6.5.39.2 Syntax 13](#_Toc166227981)

[6.5.39.3 Semantics 13](#_Toc166227982)

[6.8.10 'stem' entity grouping 14](#_Toc166227983)

[7 New HDR signaling aligning with ISO 22028-5 15](#_Toc166227984)

[6.5.40 Reference viewing environment 15](#_Toc166227985)

[6.5.40.1 Definition 15](#_Toc166227986)

[6.5.40.2 Syntax 15](#_Toc166227987)

[6.5.40.3 Semantics 15](#_Toc166227988)

[6.5.41 Nominal Diffuse White 16](#_Toc166227989)

[6.5.41.1 Definition 16](#_Toc166227990)

[6.5.41.2 Syntax 16](#_Toc166227991)

[6.5.41.3 Semantics 16](#_Toc166227992)

[7 Unified identifier handling clarifications 16](#_Toc166227993)

[8 Overview images 17](#_Toc166227994)

[6.4.10 Overview images 17](#_Toc166227995)

[6.8.11 Image Pyramid Entity Group 17](#_Toc166227996)

[6.8.11.1 Definition 17](#_Toc166227997)

[6.8.11.2 Syntax 18](#_Toc166227998)

[6.8.11.3 Semantics 18](#_Toc166227999)

[9 Region partition group 19](#_Toc166228000)

[6.8.12 Region Partition Group 19](#_Toc166228001)

[6.8.12.1 Definition 19](#_Toc166228002)

[6.8.12.2 Syntax 20](#_Toc166228003)

[6.8.12.3 Semantics 21](#_Toc166228004)

[10 Change all mention of SingleItemTypeReferenceBox to include SingleItemTypeReferenceBoxLarge 21](#_Toc166228005)

[11 Add T.35 metadata to Annex A 22](#_Toc166228006)

[A.5 ITU-T T.35 metadata 23](#_Toc166228007)

[12 Clarify that items may share auxiliary images, thumbnails and metadata 23](#_Toc166228008)

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 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: Support for tone map derived items and other improvements

# Coding Constraints box related changes

In clause 7.2.3.4, renumber NOTE2 to NOTE3.

In clause 7.2.3.4, add the following text after NOTE1:

NOTE 2: When a track contains inter-predicted images and the value of all\_ref\_pics\_intra is equal to 0, it is possible for inter-predicted images to be derived from non-intra coded images. In such cases, derived specifications can suggest guidelines for the frequency of sync samples.

Add the following text as a new subclause after subclause 7.2.3.4:

7.2.3.5 Recommendations for CodingConstraintsBox

Encoding image sequences complying with the constraint that either all samples are sync samples or the all\_ref\_pics\_intra field in the CodingConstraintsBox can be set to one is suggested in the following cases:

1. For ensuring compatibility with players implementing codec brands in annexes B, E, L and M specified in an earlier edition of this document, which required all samples to be sync samples or to have the all\_ref\_pics\_intra field in the CodingConstraintsBox to be equal to 1.
2. In applications and usages where fast random access is essential, the constraint ensures random access to any image in the image sequence by decoding at most up to two images.
3. In image sequences that have an edit list causing backward playback, the constraint makes backward playback possible with an approach that each displayed image is obtained by decoding a compliant bitstream of up to two images.

In clause B.3.2, replace the following paragraph

For a track containing an HEVC image sequence, either all samples shall be sync samples or the all\_ref\_pics\_intra field in the CodingConstraintsBox specified in 7.2.3 shall be set to one.

with the following text

For a track containing an HEVC image sequence, either all samples should be sync samples or the all\_ref\_pics\_intra field in the CodingConstraintsBox specified in 7.2.3 should be set to one.

NOTE: Clause 7.2.3.5 contains recommendations that are important for backwards compatibility.

In clause E.3.2, replace the following paragraph

For a track containing an AVC image sequence, either all samples shall be sync samples or the all\_ref\_pics\_intra field in the CodingConstraintsBox specified in 7.2.3 shall be set to one.

with the following text

For a track containing an AVC image sequence, either all samples should be sync samples or the all\_ref\_pics\_intra field in the CodingConstraintsBox specified in 7.2.3 should be set to one.

NOTE: Clause 7.2.3.5 contains recommendations that are important for backwards compatibility.

In clause L.3.2, replace the following paragraph

For a track containing an VVC image sequence, either all samples shall be sync samples or the all\_ref\_pics\_intra field in the CodingConstraintsBox specified in ‎7.2.3 shall be set to one.

with the following text

For a track containing an VVC image sequence, either all samples should be sync samples or the all\_ref\_pics\_intra field in the CodingConstraintsBox specified in 7.2.3 should be set to one.

NOTE: Clause 7.2.3.5 contains recommendations that are important for backwards compatibility.

In clause M.3.2, replace the following paragraph

For a track containing an EVC image sequence, either all samples shall be sync samples or the all\_ref\_pics\_intra field in the CodingConstraintsBox specified in ‎7.2.3 shall be set to one.

with the following text

For a track containing an EVC image sequence, either all samples should be sync samples or the all\_ref\_pics\_intra field in the CodingConstraintsBox specified in 7.2.3 should be set to one.

NOTE: Clause 7.2.3.5 contains recommendations that are important for backwards compatibility.

# New tone-map derivation item

Add the following new subclause after subclause 6.6.2.3:

## Tone-map derivation

### 6.6.2.4.1 Definition

An item with an item\_type value of 'tmap' defines a derived image item whose reconstructed image is formed from one base input image and a secondary input image that will be referred to as a gain map input image.

The input images are given by the SingleItemTypeReferenceBox/ SingleItemTypeReferenceBoxLarge of type 'dimg' for this derived image item within the ItemReferenceBox. In the SingleItemTypeReferenceBox/ SingleItemTypeReferenceBoxLarge of type 'dimg', the value of from\_item\_ID identifies the derived image item of type 'tmap', the value of reference\_count shall be equal to 2, and the values of to\_item\_ID identify the input images, of which the first shall be the base input image and the second shall be the gain map input image. The gain map input image may have different dimensions than the base as documented in ISO 21496-1 section 4.2.

Reconstruction is done by applying the gain map to the base image according to ISO 21496-1 section 6. As described in ISO 21496-1 section 6.3, the gain map may be scaled by a weight during application to adjust for local viewing conditions.

The base input image shall be associated with a 'colr' item property. This corresponds to the colorimetry metadata described in ISO 21496-1 section 5.3.1.

The gain map input image shall be associated with a 'colr' item property of type 'nclx' which indicates any transformations that the encoder has done to improve compression. In this item property, colour\_primaries and transfer\_characteristics shall be set to 2. The full\_range\_flag may be set to either 1 or 0. If the full\_range\_flag is 0, the reader shall clip the gain map pixel values to the logical range 0.0 to 1.0 after applying matrix\_coefficients and full\_range\_flag.

A 'tmap' derived item shall be associated with a 'colr' item property. This corresponds to the colorimetry metadata described in ISO 21496-1 section 5.3.2, which describes the colour properties of the reconstructed image if the gain map input item is fully applied according to ISO 21496-1 section 6.3.

When a 'tmap' derived item is the input to another derived item, that derived item shall treat the reconstructed image of the 'tmap' derived item as if the gain map has been fully applied and has the colour properties of the 'colr' item property associated with the 'tmap' derived item.

The base input image and the 'tmap' derived item should be associated with 'clli' item properties as appropriate to further document the optimal viewing conditions of each representation.

A 'tmap' derived item should be associated with a PixelInformationProperty item property. This property provides a hint to decoders on the approximate amount of colour resolution available after fully applying the gain map.

The number of channels in the gain map input item may be different than the number of channels in the channels field in the body of the 'tmap' derived item. If the gain map input item is single channel while the channels field is multi-channel, the gain map input item shall be treated as if it has three identical colour channels. If the gain map input item is multi-channel while the channels field is single channel, the channels field shall be treated as if it is multi-channel with identical values for all channels.

The gain map input image should be marked as hidden by setting (flags & 1) equal to 1 in its infe entry.

The requirements of clause 10.2.6 apply to any file containing a tone-map derived item.

NOTE 1: Backwards compatibility with parsers that do not support the tone-map derivation can be achieved by placing the base input image item and the 'tmap' derived item in an 'altr' entity group.

NOTE 2: The ToneMapImage has a version field. This is separate from the GainMapVersion field in the GainMapMetadata structure in ISO 21496-1 clause C.2. The ToneMapImage version field allows for versioning of how the tone-map derivation is done in this specification.

### Syntax

aligned(8) class ToneMapImage {  
 unsigned int(8) version = 0;  
 if(version == 0) {  
 GainMapMetadata;  
 }  
}

### Semantics

version shall be equal to 0. Readers shall not process a ToneMapImage with an unrecognized version number.

GainMapMetadata: Gain map metadata as defined in ISO 21496-1 clause C.2.

Add the following new subclause after subclause 10.2.5:

## 'tmap' brand

#### Definition

This brand enables file players to identify and decode HEIF files containing tone-map derived items. When present, this brand shall be among the brands included in the compatible\_brands array of the FileTypeBox.

#### Requirements on files

A file containing the 'tmap' brand in the compatible\_brands array of the FileTypeBox shall contain one or more tone-map derived items.

#### Requirements on readers

Readers shall support the following:

* Support for tone-map derived items
* Within the entity groups, support for EntityToGroupBox with grouping\_type equal to 'altr'

Add the following new subclause after subclause J.6:

## Tone-map derivation

This example illustrates how a tone-map derived item can be stored in a file in a backwards compatible manner using an 'altr' entity group. Both the base image item and the gain map image items are tiled using grid-derived items. The base image and gain map are explicitly marked as having zero rotation.

FileTypeBox 'ftyp': major-brand='heic', compatible-brands='tmap, mif1, heic'  
MetaBox 'meta': (container)  
 HandlerBox 'hdlr': 'pict'  
 PrimaryItemBox 'pitm': item\_ID=1;  
 ItemInfoBox 'iinf': entry\_count=11

// base image

1) 'infe': item\_ID=1, item\_type='grid';

2) 'infe': item\_ID=2(Hidden), item\_type='hvc1’;

3) 'infe': item\_ID=3(Hidden), item\_type='hvc1';

4) 'infe': item\_ID=4(Hidden), item\_type='hvc1';

5) 'infe': item\_ID=5(Hidden), item\_type='hvc1';

// gain map

6) 'infe': item\_ID=6(Hidden), item\_type='grid';

7) 'infe': item\_ID=7(Hidden), item\_type='hvc1';

8) 'infe': item\_ID=8(Hidden), item\_type='hvc1';

9) 'infe': item\_ID=9(Hidden), item\_type='hvc1';

10)'infe': item\_ID=10(Hidden), item\_type='hvc1';

// tmap

11)'infe': item\_ID=11, item\_type='tmap';

ItemLocationBox 'iloc': item\_count=11  
 item\_ID=1, construction\_method=1, extent\_count=1, extent\_offset=P1, extent\_length=Q1;

item\_ID=2, extent\_count=1, extent\_offset=P2, extent\_length=Q2;

item\_ID=3, extent\_count=1, extent\_offset=P3, extent\_length=Q3;

item\_ID=4, extent\_count=1, extent\_offset=P4, extent\_length=Q4;

item\_ID=5, extent\_count=1, extent\_offset=P5, extent\_length=Q5;

item\_ID=6, construction\_method=1, extent\_count=1, extent\_offset=P6, extent\_length=Q6;

item\_ID=7, extent\_count=1, extent\_offset=P7, extent\_length=Q7;

item\_ID=8, extent\_count=1, extent\_offset=P8, extent\_length=Q8;

item\_ID=9, extent\_count=1, extent\_offset=P9, extent\_length=Q9;

item\_ID=10, extent\_count=1, extent\_offset=P10, extent\_length=Q10;

item\_ID=11, extent\_count=1, extent\_offset=P11, extent\_length=Q11;

ItemReferenceBox 'iref':

// base image grid of 4 tiles

referenceType='dimg', from\_item\_ID=1, reference\_count=4,

to\_item\_ID=2,

to\_item\_ID=3,

to\_item\_ID=4;

to\_item\_ID=5;

// A (hidden) gain map is composed of 4 tiles in the grid

referenceType='dimg', from\_item\_ID=6, reference\_count=4,

to\_item\_ID=7,

to\_item\_ID=8,

to\_item\_ID=9;

to\_item\_ID=10;

// tmap  
 referenceType='dimg', from\_item\_ID=11, ref\_count=2,

to\_item\_ID=1, // base

to\_item\_ID=6; // gain map

ItemPropertiesBox 'iprp':  
 ItemPropertyContainerBox 'ipco':

// Base image, item 1, grid of tiles  
 'ispe'  
 'pixi'

'colr'  
 'clli'

// tiles, items 2-5

'ispe'

'hvcC'

// Gain map, item 6, grid of tiles

'ispe'  
 'pixi'

'colr'

// tiles, items 7-10

'ispe'

'hvcC'

// tmap, item 11  
 'colr'

'clli'

'pixi'

// Explicitly mark orientation of base and gain map  
 'irot': angle=0

ItemPropertyAssociation 'ipma': entry\_count=11  
 1) item\_ID=1, association\_count=5 // base  
 essential=0, property\_index=1;  
 essential=0, property\_index=2;  
 essential=1, property\_index=3;

essential=0, property\_index=4;

essential=1, property\_index=15;

2) item\_ID=2, association\_count=4 // base - tile 1

essential=1, property\_index=6;  
 essential=0, property\_index=5;  
 essential=0, property\_index=2;  
 essential=1, property\_index=3;

3) item\_ID=3, association\_count=4 // base - tile 2

essential=1, property\_index=6;  
 essential=0, property\_index=5;  
 essential=0, property\_index=2;  
 essential=1, property\_index=3;

4) item\_ID=4, association\_count=4 // base - tile 3

essential=1, property\_index=6;  
 essential=0, property\_index=5;  
 essential=0, property\_index=2;  
 essential=1, property\_index=3;

5) item\_ID=5, association\_count=4 // base - tile 4

essential=1, property\_index=6;  
 essential=0, property\_index=5;  
 essential=0, property\_index=2;  
 essential=1, property\_index=3;

6) item\_ID=6, association\_count=4 // gain map  
 essential=0, property\_index=7;  
 essential=0, property\_index=8;  
 essential=1, property\_index=9;  
 essential=1, property\_index=15;

7) item\_ID=7, association\_count=4 // gain map - tile\_1

essential=1, property\_index=11;

essential=0, property\_index=10;  
 essential=0, property\_index=8;

essential=1, property\_index=9;

8) item\_ID=8, association\_count=4 // gain map - tile\_2

essential=1, property\_index=11;

essential=0, property\_index=10;  
 essential=0, property\_index=8;

essential=1, property\_index=9;

9) item\_ID=9, association\_count=4 // gain map - tile\_3

essential=1, property\_index=11;

essential=0, property\_index=10;  
 essential=0, property\_index=8;

essential=1, property\_index=9;

10) item\_ID=10, association\_count=4 // gain map - tile\_4

essential=1, property\_index=11;

essential=0, property\_index=10;  
 essential=0, property\_index=8;

essential=1, property\_index=9;

11) item\_ID=11, association\_count=5 // tmap

essential=0, property\_index=1;

essential=1, property\_index=12;

essential=0, property\_index=13;

essential=0, property\_index=14;

essential=1, property\_index=15;

ItemDataBox 'idat':

Grid derivation data block (at idat\_offset P1, with length Q1)

Grid derivation data block (at idat\_offset P6, with length Q6)

Tone-map derivation data block (at idat\_offset P11, with length Q11)

GroupsListBox 'grpl':

EntityToGroupBox 'altr': group\_id=12, num\_entities\_in\_group=2

entity\_id=11;

entity\_id=1;

MediaDataBox 'mdat':  
 HEVC Image (at file offset P2, with length Q2)

HEVC Image (at file offset P3, with length Q3)

HEVC Image (at file offset P4, with length Q4)

HEVC Image (at file offset P5, with length Q5)

HEVC Image (at file offset P7, with length Q7)

HEVC Image (at file offset P8, with length Q8)

HEVC Image (at file offset P9, with length Q9)

HEVC Image (at file offset P10, with length Q10)

# New colour format enhancement derived item

Add the following new subclause after subclause 6.6.2.4:

## Colour format enhancement derivation

### Definition

An item with an item\_type value of 'cfen' is a colour format enhancement derived image item whose reconstructed image is formed from one or more input images that carry components in the luma plane that are used to reconstruct a picture with an enhanced colour format.

The input images are ordered using the reference type 'dimg' for this colour format enhancement derived image item within the ItemReferenceBox, where the value of from\_item\_ID identifies the colour format enhancement derived image item, and the values of to\_item\_ID identify the input images. The reference\_count gives the number of input image items and shall be greater or equal to one.

The colour format enhancement derived image item shall:

* have a Pixel Information property, an Image Spatial Extents property, and a Colour Information property with colour\_type set to 'nclx'.

EDITORS NOTE: Version 1 of 'pixi' property could be mandated for the use with 'cfen'.

The input image item(s):

* shall not signal duplicate entries for channel\_id greater than 1.
* shall each have a Pixel Information property and an Image Spatial Extents property.
* can have a Colour Information property that shall match the signaling from the Colour Information property of the colour formant enhancement derived image item.

EDITORS NOTE: It should be investigated if this restriction should only apply for version 0 of 'cfen'

The colour format enhancement derived image item should be signaled as a displayable item. The first entry signaled by the reference\_count array may be signaled as a displayable item. All remaining items signaled by the reference\_count array shall be hidden.

NOTE: Both the colour format enhancement derived image item and the first entry signaled by the reference\_count array can be incorporated into an 'altr' alternative group to signify that both image items are interlinked alternatives of each other. The colour format enhancement derived image item can be signaled as the first item in the 'altr' alternative group, while the second item in the group can store a 4:2:0 coded image to allow backwards compatibility.

### Syntax

aligned(8) class ColourFormatEnhancement {  
 unsigned int(8) version = 0;  
 for (i=0; i<reference\_count; i++) {  
 bit(7) reserved = 0;  
 unsigned int(1) is\_packed\_flag;  
 if(is\_packed\_flag == 1) {  
 // packed replacement of components  
 unsigned int(3) num\_cols\_minus1;  
 unsigned int(3) num\_rows\_minus1;  
 unsigned int(5) hor\_guard\_band\_log2;  
 unsigned int(5) ver\_guard\_band\_log2;  
 for(i=0; i<num\_cols\_minus1+num\_rows\_minus1+2; i++) {  
 unsigned int(8) channel\_id;  
 }  
 }  
 else {  
 unsigned int(8) channel\_id;  
 }  
 }  
}

### Semantics

version shall be equal to 0.

is\_packed\_flag this flag determines if the picture channels are packed in the referenced image item. When set to 1 the picture channels are packed into the first channel, i.e. the luma channel, using a packing method based on a matrix. When set to 0 the picture channels are stored in separate image items.

num\_cols\_minus1 plus 1 specifies the number of columns partitioning the picture.

num\_rows\_minus1 plus 1 specifies the number of rows partitioning the picture.

hor\_guard\_band\_log2 specifies if a horizontal guard band region is present between two picture channels and, if present, its size. If hor\_guard\_band\_log2 is equal to 0, then no horizontal guard band is present. If hor\_guard\_band\_log2 is larger than 0, then the horizontal guard band between two picture channels is equal to (1 << hor\_guard\_band\_log2).

ver\_guard\_band\_log2 specifies if a vertical guard band region is present between two picture channels and, if present, its size. If ver\_guard\_band\_log2 is equal to 0, then no vertical guard band is present. If ver\_guard\_band\_log2 is larger than 0, then the vertical guard band between two picture channels is equal to (1 << ver\_guard\_band\_log2).

channel\_id provides the identifier for the channel for the referenced items.

|  |  |  |
| --- | --- | --- |
| channel\_id | Mapping (depending on the 'colr' box) | |
| 0 | Unused | |
| 1 | Unspecified | |
| 2 | Y | R |
| 3 | Cb | G |
| 4 | Cr | B |
| 5 | Alpha | |
| 6 | Depth | |
| 7-255 | Reserved for future use. | |

# New constrained extents grid property

Add the following new subclause after subclause 6.5.36:

## Constrained Extents Grid Property

### Definition

Box type: 'cexg'

Property type: Descriptive item property

Container: ItemPropertyContainerBox

Mandatory (per item): No

Quantity (per item): At most one

The ConstrainedExtentsGridProperty descriptive item property indicates that each extent of the associated image item in the itemLocationBox is constrained to enclose data units of the item that are extractable as a contiguous byte range and are independently decodable and renderable as image tiles.

All data units or properties required to configure the decoder and decode an image tile shall be declared in the decoder configuration and initialization properties associated with the image item.

The reconstructed image of the associated image item is formed from one or more image tiles in a given grid order within a larger canvas.

The image tiles corresponding to the extents are inserted in row-major order, top-row first, left to right, in the order of the extents for the associated image item within the ItemLocationBox. The value of extent\_count within the ItemLocationBox shall be equal to (1+rows\_minus\_one)\*(1+columns\_minus\_one). All image tiles shall have exactly the same width and height, image\_tile\_width and image\_tile\_height. The reconstructed image is formed by tiling the image tiles into a grid with a column width equal to image\_tile\_width and a row height equal to image\_tile\_height, without gap or overlap. The grid of image tiles shall completely “cover” the reconstructed image of the associated image item, where image\_tile\_width\*columns is greater than or equal to image\_width and image\_tile\_height\*rows is greater than or equal to image\_height, where image\_width and image\_height are signalled in the ImageSpatialExtentsProperty associated with the image item.

### Syntax

aligned(8) class ConstrainedExtentsGridProperty  
extends ItemFullProperty('cexg', version = 0, flags) {  
 // this is a temporary,non-parsable variable  
 unsigned int FieldLength = ((flags & 1) + 1) \* 16;  
 unsigned int(16) rows\_minus\_one;  
 unsigned int(16) columns\_minus\_one;  
 unsigned int(FieldLength ) image\_tile\_width;  
 unsigned int(FieldLength ) image\_tile\_height;  
}

### Semantics

(flags & 1) equals to 0 specifies that the length of the fields image\_tile\_width and image\_tile\_height is 16 bits. (flags & 1) equals to 1 specifies that the length of the fields image\_tile\_width and image\_tile\_height is 32 bits. The values of flags greater than 1 are reserved.

image\_tile\_width, image\_tile\_height: specify respectively the width and height in pixels of the image tiles.

rows\_minus\_one, columns\_minus\_one: specify the number of rows of image tiles, and the number of image tiles per row. The value is one less than the number of rows or columns respectively. Image tiles enclosed in extents populate the top row first, followed by the second row and following rows, in the order of extents.

# New stereo related signaling

Add the following new subclause after subclause 6.5.37:

## Disparity adjustment information

### Definition

|  |  |
| --- | --- |
| Box type: | 'dadj' |
| Property type: | Descriptive item property |
| Container: | ItemPropertyContainerBox |
| Mandatory (per associated identifier value): | No |
| Quantity (per associated identifier value): | At most one |
|  |  |

The disparity adjustment descriptive item property defines the suggested global disparity adjustment amount for a stereo pair.

This item property should only be associated with an item or entity group that describes a stereo pair. If associated with a stereo pair entity group, the FileTypeBox or the ExtendedTypeBox associated with the FileTypeBox shall indicate that the requirements of the 'unif' brand apply in the file.

EDITORS NOTE: See TuC on frame-packed stereo items

### Syntax

aligned(8) class DisparityAdjustmentProperty  
extends ItemFullProperty('dadj', version = 0, flags = 0) {  
 signed int(32) disparity\_adjustment;  
}

### Semantics

disparity\_adjustment in units of 1/10 000 image widths. Positive values denote increased disparity with respect to parallel view direction.

Add the following new subclause after subclause 6.5.38:

## Stereo aggressors item property

### Definition

|  |  |
| --- | --- |
| Box type: | 'stag' |
| Property type: | Descriptive item property |
| Container: | ItemPropertyContainerBox |
| Mandatory (per associated identifier value): | No |
| Quantity (per associated identifier value): | Zero or more |

The stereo aggressors descriptive item property specifies the presence and characteristics of stereo aggressors detected within a stereo pair. Stereo aggressors are identified as elements that potentially cause discomfort when viewing the stereo pair on a stereoscopic display. This item property allows to identify and characterize these aggressors in detail.

This item property, when present, shall be used with a stereo pair. It can also be associated with other items only if also associated with a stereo pair so it is clear in which stereo context it applies. If associated with a stereo pair entity group ('ster' or 'stem'), the 'unif' brand shall be present in the compatible brands.

If this item property is associated with another item in addition to the stereo pair, it serves as a hint that the aggressors described by the property are localized to that specific item. For example, this can be used to indicate that a "Lens occlusion" is present in only the left or right image item, or using a region item, in a specific region of an image item. Multiple stereo aggressor properties may be associated with the same stereo pair since not all aggressors may be localized to the same area.

NOTE 1: An image item can be associated with multiple stereo pair entity groups; however a stereo aggressor can only apply to a specific group, which is why there is a requirement that the stereo aggressor has to be associated with a stereo group.

NOTE 2: As localization of the stereo aggressors is indicated with item property associations, all stereo aggressors within a single 'stag' box need to share the same localization. Specifying multiple unique localizations requires using multiple 'stag' boxes.

The aggressor\_severity shall be interpreted as follows:

* A value of 0 indicates an unknown severity. The file creator has no recommendation for the reader.
* A value of 1 to 42 (inclusive) indicates a mild severity. The file creator recommends that sensitive viewers are warned about potential discomfort when viewing in stereo.
* A value of 43 to 84 (inclusive) indicates a medium severity. The file creator recommends that all viewers are warned about potential discomfort when viewing in stereo.
* A value of 85 to 127 (inclusive) indicates a high severity. The file creator recommends that all viewers are warned about potential discomfort when viewed in stereo and that viewing should default to monoscopic rather than stereoscopic display unless overridden by the viewer.

If sub\_type\_uri is not specified for a specific aggressor, the aggressor\_severity is a relative measure only used to provide the viewer recommendations above and to rank the aggressors in the file relative to each other. If sub\_type\_uri is specified for an aggressor, aggressor\_severity shall be an absolute measure that can be compared between files.

### Syntax

aligned(8) class StereoAggressorsProperty  
extends ItemFullProperty('stag', version = 0, flags = 0) {  
 unsigned int(8) aggressor\_count\_minus\_one;  
 for(int i = 1; i <= aggressor\_count\_minus\_one + 1; i++) {  
 unsigned int(8) aggressor\_type;  
 unsigned int(1) sub\_type\_present;  
 unsigned int(7) aggressor\_severity;  
 if (sub\_type\_present) {  
 utf8string sub\_type\_uri;  
 }  
 }  
}

### Semantics

version shall be equal to 0.

aggressor\_count\_minus\_one is the number of aggressors minus one.

aggressor\_type has the following defined values:

|  |  |
| --- | --- |
| 0 | Unspecified |
| 1 | Lens occlusion |
| 2 | Image condition mismatch (lighting conditions, color, glare, sharpness, processing differences) |
| 3 | Stereo window violation |
| 4 | Object(s) too close |
| 5 | Stereo calibration error(s) |
| 6 | Temporal alignment mismatch |
| 7 | Poor image quality in one or both views (e.g. due to specific noise patterns) |
| 8 - 255 | Reserved |

aggressor\_severity provides an indication of how severe the file creator considers the aggressor to be. If sub\_type\_uri is not defined, the severity is file-relative and can not be used to compare aggressor severities between files.

sub\_type\_uri is an optional null-terminated UTF-8 character string of the Uniform Resource Identifier (URI) used to more explicitly identify the type of the stereo aggressor as well as how to interpret the aggressor\_severity. If specified, the aggressor\_severity shall be an absolute metric that can be compared between files. If not present, it defaults to the empty string.

Add the following text to the end of subclause 6.8.5:

NOTE: The 'stem' entity group can be used to indicate which item should be used for monoscopic fallback if the primary item is not part of the 'ster' group.

Add the following new subclause after 6.8.9:

## 'stem' entity grouping

The stereo pair entity grouping with monoscopic fallback ('stem'), is functionally equivalent to the 'ster' entity grouping with the following changes:

* The entity group shall contain exactly three entity\_id values that point to image items. The first two of these values shall be interpreted in the same way as the two entity\_id values of a 'ster' entity group.
* The third entity\_id value (with i equal to 2) indicates the image to display if monoscopic viewing of the stereo pair is desired. This may be the same as either the first or second entity\_id, but may also point to another image item more suitable for monoscopic display.
* If the primary item is part of the 'stem' group, the third entity\_id shall designate the primary item for monoscopic fallback.
* The value of (flags & 3) indicates how the third entity\_id relates to the stereo pair. The following values are defined:  
  0: Unspecified  
  1: Co-located with left view  
  2: Co-located with right view  
  3: Centered between left and right views

NOTE: To be backwards compatible with parsers that do not understand 'stem' entity groups, file creators may add both a 'stem' and a 'ster' entity group to the file. An 'altr' group can then be added containing the 'stem' and 'ster' group id's in that order to indicate that only one should be used. In order to use group id's in an 'altr' group, the 'unif' brand is required.

# New HDR signaling aligning with ISO 22028-5

Add the following new subclauses after subclause 6.5.39:

EDITORS NOTE: Do we need these properties for video as well?

## Reference viewing environment

### Definition

|  |  |
| --- | --- |
| Box type: | 'reve' |
| Property type: | Descriptive item property |
| Container: | ItemPropertyContainerBox |
| Mandatory (per item): | No |
| Quantity (per item): | At most one |

The reference viewing environment applies to display-viewing colorimetry, not to scene-referred colorimetry. It specifies the luminance and chromaticity parameters for the “surround” and “periphery” of the display. The “surround” is the area surrounding a display that can affect the adaptation of the eye, typically the wall or curtain behind the display, while “periphery” is the remaining environment outside of the surround.

### Syntax

class ReferenceViewingEnvironmentBox extends ItemFullProperty('reve', 0, 0){  
 unsigned int(32) surround\_luminance;  
 unsigned int(16) surround\_light\_x;  
 unsigned int(16) surround\_light\_y;  
 unsigned int(32) periphery\_luminance;  
 unsigned int(16) periphery\_light\_x;  
 unsigned int(16) periphery\_light\_y;  
}

### Semantics

surround\_luminance specifies the luminance of the surround in units of 0.0001 candelas per square metre.

[Ed. note]: disallowing 0 could be considered.

surround\_light\_x and surround\_light\_y specify the normalized x and y chromaticity coordinates, respectively, of the environmental reference surround light in the nominal viewing environment. These parameters are according to the CIE 1931 definition of x and y as specified in ISO 11664-1 (see also ISO 11664-3 and CIE 15) and are in normalized increments of 0.0001. The values of surround\_light\_x and surround\_light\_y shall be in the range of 0 to 10 000, inclusive.

periphery\_luminance specifies the luminance of the periphery in units of 0.0001 candelas per square metre.

[Ed. note]: disallowing 0 could be considered.

periphery\_light\_x and periphery\_light\_y specify the normalized x and y chromaticity coordinates, respectively, of the environmental reference periphery light in the nominal viewing environment. These parameters are according to the CIE 1931 definition of x and y as specified in ISO 11664-1 (see also ISO 11664-3 and CIE 15) and are in normalized increments of 0.0001. The values of periphery\_light\_x and periphery\_light\_y shall be in the range of 0 to 10 000, inclusive.

## Nominal Diffuse White

### Definition

|  |  |
| --- | --- |
| Box type: | 'ndwt' |
| Property type: | Descriptive item property |
| Container: | ItemPropertyContainerBox |
| Mandatory (per item): | No |
| Quantity (per item): | At most one |

### Syntax

class NominalDiffuseWhiteBox extends ItemFullProperty('ndwt', 0, 0){  
 unsigned int(32) diffuse\_white\_luminance;  
 unsigned int(16) diffuse\_white\_light\_x;  
 unsigned int(16) diffuse\_white\_light\_y;  
}

### Semantics

diffuse\_white\_luminance indicates the default nominal diffuse white luminance in units of 0.0001 candelas per square metre.

[Ed. note]: disallowing 0 could be considered.

diffuse\_white\_light\_x and diffuse\_white\_light\_y specify the normalized x and y chromaticity coordinates, respectively, of the environmental reference periphery light in the nominal viewing environment. These parameters are according to the CIE 1931 definition of x and y as specified in ISO 11664-1 (see also ISO 11664-3 and CIE 15) and are in normalized increments of 0.0001. The values of periphery\_light\_x and periphery\_light\_y shall be in the range of 0 to 10 000, inclusive and should be the same as the chromaticity coordinates of the white point of the content.

[Ed. note]: We could also add a note mentioning that values diffuse\_white\_light\_x = 3 127 and diffuse\_white\_light\_y = 3 290 could be used to signal D65.

# Unified identifier handling clarifications

Replace the word "item\_ID" in clauses 6.5.27.1, 6.5.18.1, 6.5.19.1, 6.5.35.1 with the phrase "identifier value"

Append the following text as a new paragraph in clauses 6.5.27.1, 6.5.18.1, 6.5.19.1, 6.5.35.1:

If this property is associated with an entity group, the FileTypeBox or the ExtendedTypeBox associated with the FileTypeBox shall indicate that the requirements of the 'unif' brand apply in the file.

# Overview images

Add the following new subclause after subclause 6.4.9:

## Overview images

An overview image is a grid derived image item or a tiled pre-derived coded image item whose reconstructed image is formed from generating a lower resolution, ‘binned’ version of a base image item. The base image item is also a tiled image item. The tiling may be implemented using a feature of a specific codec, or by using a grid derived image item. When a grid derived image item is used, the input items to the grid define the tiles. Derived image items shall not be used as inputs to the image grid, due to the need for in place byte range accessing of content. Individual tiles shall be written contiguously in memory, thereby allowing access with a single read or write action.

A pre-defined coded image representing an overview image or an image item representing the base image that are tiled using a feature of a specific codec shall be stored in such a way that each extent identifies that data range corresponding to a tile, and shall be associated with a ConstrainedExtentsGridProperty indicating the constraint on the extents and describing the tiling grid.

An overview image shall be tiled using the same tiling scheme as the base image, i.e. if tiles in the base image are X by Y pixels, they are X by Y pixels in the overview image. In cases where the binned resolution results in a fractional, or incomplete tile at the end of a row (column), the last tile in a row (column) of tiles shall be padded with the value zero at the end of the row (column) to complete the last tile in the row (column). If necessary, the clean aperture transformative property ('clap') may be applied to crop padded rows and/or columns. The number of tiles in a row (column) of tiles is determined by dividing the width (height) of the overview image by the tile size in X (tile size in Y) and rounding up.

An image pyramid is generated by stacking a series of progressively binned overview images and creating an ImagePyramidEntityGroup. Each overview image is associated with the original full resolution base image, using a reference of type 'base'. The amount of binning of each overview image is indicated in the ImagePyramidEntityGroup. The image format of the overviews is the same as the base image item. i.e. number of bands, bit depth, color format, etc.

Region items associated with the base image may be replicated for individual overviews using an appropriate scaling associated with the level of binning for a particular overview and referenced to the specific overview.

NOTE 1: In this version of the document, the exact derivation process (approaches such as the sum, average, median, minimum, or maximum value of a binned region) used to produce an overview from the base image is left to the implementer.

NOTE 2: When removing (or modifying) an item that is marked as the base image of an overview image, the content of associated image overview items might need to be removed (or rewritten).

Add the following new subclause after subclause 6.8.10:

## Image Pyramid Entity Group

### Definition

Box Type: 'pymd'

Container: GroupsListBox in a MetaBox at file level

Mandatory: No

Quantity: Zero or more

The ImagePyramidEntityGroup indicates a set of image items, formed as a base image item and a series of progressively binned overview image items, which together form an image pyramid.

Each overview image item has a reference to the original full resolution base image item, using a reference of type 'base'.

The ImagePyramidEntityGroup also provides overall information for the individual tiles inside the overview image items and base image item of the image pyramid.

The image format of the overview images shall be the same as the base image (i.e. number of bands, bit depth, color format, etc).

This entity group shall contain entity\_id values that point to a base image item and a set of overview image items and shall contain no entity\_id values that point to tracks. The entities shall be listed in the order of lowest resolution overview image item to the highest resolution overview image item, followed finally by the base image item of the image pyramid.

There may be multiple ImagePyramidEntityGroups in the same file with different group\_id values.

NOTE: All the entities of a same ImagePyramidEntityGroup, or only some of them, can also be members of a same entity group of type 'prgr' if they are stored in the file for allowing a progressive refinement. They can also be members of a same entity group of type 'altr' if they are proposed by the content creator as alternatives to be displayed for players not supporting the ImagePyramidEntityGroup.

### Syntax

aligned(8) class ImagePyramidEntityGroup  
extends EntityToGroupBox ('pymd', version = 0, flags = 0) {  
 unsigned int(16) tile\_size\_x;  
 unsigned int(16) tile\_size\_y;  
 for(i=0; i<num\_entities\_in\_group;i++) {  
 unsigned int(16) layer\_binning;  
 unsigned int(16) tiles\_in\_layer\_row\_minus1;  
 unsigned int(16) tiles\_in\_layer\_column\_minus1;  
 }  
}

### Semantics

num\_entities\_in\_group: as defined for EntityToGroupBox. In addition, it also specifies the number of layers of the image pyramid.

tile\_size\_x, tile\_size\_y: indicate the size in pixels of a tile in the width and height dimension, respectively, for all layers of the image pyramid.

layer\_binning: Indicates for each layer of the pyramid the level of binning between the base image and the overview image. A 2x2 binning is defined to be a layer\_binning of 2, a 4x4 binning is defined to be 4, etc. The width and height for an overview image with layer\_binning of 2 is half the width and half the height of the base image, etc. A base image has a layer\_binning of 1.

tiles\_in\_layer\_row\_minus1, tiles\_in\_layer\_column\_minus1: Indicate the number of tiles minus one in a row and a column, respectively, of a specific layer. If the layer is represented by a grid derived image item, tiles\_in\_layer\_row\_minus1 is equal to rows\_minus\_one and tiles\_in\_layer\_column\_minus1 is equal to columns\_minus\_one. If the layer is represented by a tiled pre-derived coded image item with a ConstrainedExtentsGridProperty, then tiles\_in\_layer\_row\_minus1 is equal to rows\_minus\_one and tiles\_in\_layer\_column\_minus1 is equal to columns\_minus\_one.

# Region partition group

Add the following new subclause after subclause 6.8.11:

EDITORS NOTE: How do partitions and pyramids interact?

## Region Partition Group

### Definition

Box Type: 'rgpa'

Container: GroupsListBox in a MetaBox at file level

Mandatory: No  
Quantity: Zero or more

A region partition group lists region items that are contained inside an area of an image. This area is the area covered by the region partition group and is defined as a rectangle in the image.

The region partition group is associated with the image item containing the region items it groups through an item reference of type 'rpds' from the image item to the region partition group.

The area covered by a region partition group is either defined in the region partition group structure itself or is the area of the image item the region partition group is associated with through an item reference of type 'rpds'.

A diagram of a diagram

Description automatically generated with medium confidence

Figure 1: Example of two region partition groups ('rgpa') associated with an image, the area of the top group is the whole image, while the area of the bottom group is the bottom-center part of the image

A region partition group associated with an image item should only contain region items associated with the image item or with another image item that this image item is a part of.

A diagram of a grid

Description automatically generated

Figure 2: Example of two region partition groups ('rgpa') related to a grid, the top group is associated with the grid and its area is the whole image, the bottom group is associated with an input item of the grid and its area is the whole input item

When a region item is contained in a region partition group, at least one of its regions shall intersects the area covered by the region partition group.

A region partition group shall list only region items it contains. A region partition group may list all the region items it contains. A region partition group may list only the region items it contains and that are adapted to being rendered when the rendered size of the area covered by the region partition group is greater than or equal to the display area.

When rendering a part of an image item, a renderer may use region partition groups to select which region items to parse and render. The renderer may use region partition groups matching the area to display. When several region partition groups correspond to the area to render, the renderer may use those matching the scale of the rendered area. The renderer may use region partition groups associated with the image item to render or associated with items corresponding to parts of the image item to render.

When using region partition groups jointly with an image pyramid, the area covered by a region partition group should correspond to the area of a tile of the image pyramid.

### Syntax

aligned(8) class RegionPartitionGroupBox  
extends EntityToGroupBox('rgpa', version, flags) {  
 unsigned int field\_size = ((flags & 1) + 1) \* 16;  
 if (flags & 2 == 2) {  
 unsigned int(field\_size) reference\_width;  
 unsigned int(field\_size) reference\_height;  
 unsigned int(field\_size) top;  
 unsigned int(field\_size) left;  
 unsigned int(field\_size) width;  
 unsigned int(field\_size) height;  
 }  
}

### Semantics

The following flags values are defined to describe the area covered by a group of region annotation items contained in a region partition group:

* Value 0x000001, when set specifies that the length of the fields reference\_width, reference\_height, top, left, width, and height is 32 bits. When not set, it specifies that the length of the fields reference\_width, reference\_height, top, left, width, and height is 16 bits.
* Value 0x000002, when set, specifies that the fields reference\_width, reference\_height, top, left, width, and height are present. When not set, it specifies that the fields reference\_width, reference\_height, top, left, width, and height are not present. In this case, the area covered by the region partition group is the area of the image it is associated with.

reference\_width, reference\_height specify, in pixel units, the width and height, respectively, of the reference space in which the region partition group is specified.

top, left specify the coordinates of the top-left corner of the area covered by the region partition group relatively to the reference space.

width, height specify the coordinates of the width and the height of the area covered by the region partition group relatively to the reference space.

# Change all mention of SingleItemTypeReferenceBox to include SingleItemTypeReferenceBoxLarge

EDITORS NOTE: Once we have the text for HEIF 3ed, make sure that we have not missed any mention of SingleItemTypeReferenceBox.

In subclause 6.6.1 replace:

The number of SingleItemTypeReferenceBoxes with the box type 'dimg' and with the same value of from\_item\_ID shall not be greater than 1.

*With:*

The number of SingleItemTypeReferenceBox or SingleItemTypeReferenceBoxLarge entries with the box type 'dimg' and with the same value of from\_item\_ID shall not be greater than 1.

In subclause 6.6.2.2.1 replace:

The input images are listed in the order they are layered, i.e. the bottom-most input image first and the top-most input image last, in the SingleItemTypeReferenceBox of type 'dimg' for this derived image item within the ItemReferenceBox.

With:

The input images are listed in the order they are layered, i.e. the bottom-most input image first and the top-most input image last, in the SingleItemTypeReferenceBox or SingleItemTypeReferenceBoxLarge of type 'dimg' for this derived image item within the ItemReferenceBox.

In subclause 6.6.2.2.3 replace:

reference\_count is obtained from the SingleItemTypeReferenceBox of type 'dimg' where this item is identified by the from\_item\_ID field.

With:

reference\_count is obtained from the SingleItemTypeReferenceBox or SingleItemTypeReferenceBoxLarge of type 'dimg' where this item is identified by the from\_item\_ID field.

In subclause 6.6.2.3.1 replace:

The input images are inserted in row-major order, top-row first, left to right, in the order of SingleItemTypeReferenceBox of type 'dimg' for this derived image item within the ItemReferenceBox. In the SingleItemTypeReferenceBox of type 'dimg', the value of from\_item\_ID identifies the derived image item of type 'grid', the value of reference\_count shall be equal to rows\*columns, and the values of to\_item\_ID identify the input images.

With:

The input images are inserted in row-major order, top-row first, left to right, in the order of SingleItemTypeReferenceBox or SingleItemTypeReferenceBoxLarge of type 'dimg' for this derived image item within the ItemReferenceBox. In the SingleItemTypeReferenceBox or SingleItemTypeReferenceBoxLarge of type 'dimg', the value of from\_item\_ID identifies the derived image item of type 'grid', the value of reference\_count shall be equal to rows\*columns, and the values of to\_item\_ID identify the input images.

In subclause 6.11.1 replace:

The number of SingleItemTypeReferenceBoxes with the box type 'drgn' and with the same value of from\_item\_ID shall not be greater than 1.

With:

The number of SingleItemTypeReferenceBox or SingleItemTypeReferenceBoxLarge entries with the box type 'drgn' and with the same value of from\_item\_ID shall not be greater than 1.

# Add T.35 metadata to Annex A

In subclause A.1 replace:

This annex specifies the format to store metadata complying with Exif (JEITA CP-3451E), XMP (ISO 16684-1), or MPEG-7 (ISO/IEC 15938-3) in files conforming to the Image File Format. When Exif, XMP, or MPEG-7 metadata is associated with items or tracks conforming to the Image File Format, the metadata shall follow the specifications of this annex. However, it is not required for a reader conforming to this document to understand Exif, XMP, or MPEG-7 metadata.

With:

This annex specifies the format to store metadata complying with Exif (JEITA CP-3451E), XMP (ISO 16684-1), ITU-T T.35, or MPEG-7 (ISO/IEC 15938-3) in files conforming to the Image File Format. When Exif, XMP, ITU-T T.35, or MPEG-7 metadata is associated with items or tracks conforming to the Image File Format, the metadata shall follow the specifications of this annex. However, it is not required for a reader conforming to this document to understand Exif, XMP, ITU-T T.35, or MPEG-7 metadata.

Add the following after subclause A.4:

## A.5 ITU-T T.35 metadata

ITU-T T.35 metadata as specified may be associated with image items using an item type of 'it35' as defined in ISO/IEC 14496-12.

For image sequences, ITU-T T.35 metadata may be embedded within samples and signalled using an ITU-T T.35 sample group or may be stored in an ITU-T T.35 sample group as defined in ISO/IEC 14496-12.

Add the following in the Bibliography:

*[xx]* Recommendation ITU-T T.35*, Procedure for the allocation of ITU-T defined codes for non-standard facilities*

# Clarify that items may share auxiliary images, thumbnails and metadata

In subclause 6.4.4 replace:

The thumbnail image and the master image are linked using a reference type 'thmb' from the thumbnail image to the master image.

With:

The thumbnail image and master image(s) are linked using a reference type 'thmb' from the thumbnail image to the master image.

In subclause 6.4.5 replace:

The auxiliary image and the master image are linked using an item reference of 'auxl' from the auxiliary image to the master image.

With:

The auxiliary image and master image(s) are linked using an item reference of 'auxl' from the auxiliary image to the master image.