INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC1/SC29/WG11 MPEG2015/M36556
June 2015, Warsaw, Poland

Source **Telecom ParisTech, Canon Research Centre France**
Status **For consideration at the 112th MPEG Meeting**
Title **Editorial changes to the Image File Format**
Author Cyril Concolato, Jean Le Feuvre, Franck Denoual, Frédéric Mazé, Naël Ouedraogo, Eric Nassor

1 Introduction

This contribution proposes various editorial changes to the Image File Format produced during the 111th MPEG meeting.

1.1 Item definition

14496-12 misses a definition of the term 'item'. We propose to define it as follows

3.1.X item
data which does not require timed processing, as opposed to sample data.

1.2 Cover image

The current text contains several texts for the notion of "cover image":

"3.1.4 cover image
image that is representative of an image collection or an image sequence and should be displayed when no other information is available on the preferred displaying method of the image collection or the image sequence"

> "NOTE: The primary item is typically a cover image as specified in **Erreur ! Source du renvoi introuvable.**. It is recommended not to have a thumbnail image or an auxiliary image as a primary item."

"6.3.3 Cover image

For a collection of still images carried as items in a MetaBox, the cover image, if available, is the primary item of the MetaBox. The cover image should be displayed when no other information is available on the preference to display a collection of still images."

These texts do not specify any syntax for a cover image. A cover image is the primary item. The specification could be simplified by removing all the above texts.

1.3 Auxiliary Images

In section 6.3.3: the definition of (generic) auxiliary images uses text that is specific to HEVC auxiliary images:

> `aux_type`: A null-terminated UTF-8 character string of the Uniform Resource Name (URN) used to identify the type of auxiliary image. The URN identifies the resource and the index in the resource which identifies the type of auxiliary image.

URNs (in the generic sense) do not generally identify resources and indexes. The last sentence should be deleted. Also it should be clarified if any generic URN is defined by MPEG or reserved, and how URN should be formed.

1.3.1 HEVC Auxiliary Images

In section B.2.5.1, there is a problem with the following note (missing "for"?).

> NOTE: While Table F.2 of ISO/IEC 23008-2 specifies auxiliary picture types for auxiliary pictures within auxiliary picture layers, urn:mpeg:hevc:2015:auxid:xxx may be used auxiliary pictures stored as items with an item reference of type `'auxl'` to another image

Additionally, the specification should be implementable without notes. The following note seems to contain normative text as to where the AuxConfigBox is allowed and where the HEVCConfigurationBox can be placed.

> NOTE: The auxiliary image is treated like a master image. Among other things, this means that the initialization data for HEVC coded auxiliary images is provided by an image entry item. The same image entry item may contain both AuxConfigBox and HEVCConfigurationBox. It is possible that a master image and an auxiliary image share the same image entry item, in which case AuxConfigBox is ignored for the initialization of the master image. When several image items share the same content of AuxConfigBox but different content for HEVCConfigurationBox, it may be preferable to have an image entry item containing AuxConfigBox separately from the image entry items containing HEVCConfigurationBox, and provide two `'init'` item references for HEVC auxiliary images.

Remove the normative parts out of the note and insert them in the specification text.

1.3.2 Integer representation

The following sentence is not precise enough:
"The xxx in the URN string is the ASCII representation of an integer"
Strings in ISOBMFF are UTF-8 encoded, so using ASCII does not add any value. More importantly the string representation of the integer is not defined: 'decimal', 'hexadecimal' …

1.4 Image Entry item

The draft text introduced the concept of "Image Entry Item". This concept is hard to understand. In particular, the analogy to "sample entry" is only understandable to few people. It should be renamed from "image entry item" to "image initialization item". Additionally, the text needs clarifications.

It is very hard to understand that the item_type defines the syntactical structure associated to it and that a complex hierarchy of structures is defined. For example, there is no explicit place where one can find the syntax associated to a given item_type 4CC. A table would help. For instance, the "ient" value is not clearly associated with the `ImageItemEntry` class. Another example is the "`AuxImageItemEntry`" for which the item type is hard to identify.

1.5 Reference samples

In 7.4.1.1, the text introduces s, r, and q but indicates in a note that the samples should be in the order q, r and s. This is confusing.

1.6 MIME Types

1.6.1 sub-type vs. codecs vs. Profiles (Annex D.3)

The proposed codecs sub-parameter mixes coding and storage aspects. It should be there to identify the media codecs used in the file not the way the codecs data is packed in the mp4 file. The MIME sub-type 'heic' already hints on that and the profiles sub-parameter could list brands for that as well. If files are declared with video/mp4, the profiles sub-parameter should indicate brands that hint that the file has IFF structures. We suggest removing the part defining 'metaitem' and 'irot' (and alike). The example below:

```
Content-Type: image/heic; codecs=metaitem.irot+hvc1.A1.80.L93.B0
```

would become:

```
Content-Type: image/heic; codecs=hvc1.A1.80.L93.B0;
profiles=mif1
```

1.6.2 order of tracks

In Annex E.2, the current MIME registration annex forces (for the first time in the ISOBMFF) to process the file tracks in a specific order in order to produce the codecs value, with the following sentence:

```
"the first list item shall represent a track having the handler
type 'pict'"
```

This means that one cannot simply process all items and tracks in the given file order. This introduces unnecessary complexity.

1.6.3 missing implementation

In Annex D.2, the MIME Type registration annexes indicate:

```
Interoperability considerations: A number of interoperating
          implementations exist within the ISO/IEC 14496
          community, and that community has reference software
          for reading and writing the file format.
```

We believe the annex should be delayed until such interoperable implementations are produced.

1.7 extra "the"

"3.1.13 Image File Format
the file format specified by this specification"
Extra "the"

1.8 too many "indicated"

"3.1.5 derived image
image that is represented in a file by an indicated operation to indicated images and can be obtained by performing the indicated operation to the indicated images"

1.9 too many "meta"

Consider renaming `ISOBMFFMetaData` to `ISOBMFFImageMetaData` to avoid confusion with the ISOBMFF "meta" box.

1.10 Clean aperture

In section 7.2.1, Note on CleanApertureBox:

> NOTE: When a CleanApertureBox is present in a sample entry, the clipping specified by the CleanApertureBox takes place before applying the rotation specified by the matrix syntax element.

This note defines a normative behavior. It is only a note because this behavior is already normatively defined in 14496-12, add "As specified in 14496-12 XXX"

1.11 No need for justification

In section 7.2.3.4, the following text is not proper standard text, no need to justify the usefulness of some feature:
"This is useful to a decoder because with this knowledge, as well as the profile-level-tier flags signalled in the sample entry, it can assign the appropriate amount of buffer space to store the reference images."
Make the text a note.

1.12 Reserved value

In 7.2.3.4
" Readers shall ignore reserved regardless of its value."
should be
"Readers compliant to this version shall ignore reserved regardless of its value."
or the sentence be removed.

1.13 ISOBMFFMetadata

1.13.1 Confusing text

In section 6.6.1, the following text is confusing:
" may also contain any boxes of `ISOBMFFMetaData` describing the linked image items "
should be:
" may also contain any of the possible boxes in `ISOBMFFMetaData` describing the linked image items "

1.13.2 Missing 'rloc'

In section 8.5.2, the 'rloc' box can be present in the ISOBMFFMetadata box like 'clap' … It should be added explicitely and not implicitely allowed in extra boxes.

1.13.3 Semantics of Image spatial extents

In section 8.5.3.3: definition of ImageSpatialExtentsBox mentions the term “associated image”

We suggest using “associated image” in the semantics section:
Replace :
display_width specifies the width of the image in pixels.
display_height specifies the height of the image in pixels

With:
display_width specifies the width of the associated image in pixels.
display_height specifies the height of the associated image in pixels

There should also be a note to indicate that the term "associated image" could mean a tile.

1.13.4 Terms and definitions misplaced

Section 7.4.1.1: The terms "direct reference sample" (and others) should be defined in the terms and definitions section.

1.13.5 Thumbnail and image sequence tracks

Section 7.5.2, the following sentence is unclear:
"the thumbnail track and the image sequence track are signaled to be a part of the same alternate group"
The verb "are" should be replaced by "should be".

1.13.6 Tile

In Section B.2.6.1., "An HEVC tile item is stored as items of type `'hvt1'",` "is" should be replaced by "shall be"

1.13.7 Tile

In B.2.6.2.3, the tile location semantics should be more explicit and say:
“indicates the position of the HEVC tile item within the HEVC image item (when only one ‘hvc1’ item is present) or within the HEVC image item referenced by the ‘tbas’ reference type (when several ‘hvc1’ items are present)”

1.14 Hidden images

In section 9.3.3.1: "An entity group of type `'altr'` that includes image items shall include either all hidden or all non-hidden items."
This is confusing. What happens if 2 non-hidden images are not alternate and there is a third hidden image?