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

ISO/IEC JTC1/SC29/WG11 MPEG2014/M34354
July 2014, Sapporo, Japan

Source **Telecom ParisTech, Canon Research Centre France**
Status **For consideration at the 109th MPEG Meeting**
Title **On HEVC Still Image File Format**
Author Jean Le Feuvre, Cyril Concolato, Franck Denoual, Frédéric Mazé, Eric Nassor

1 Introduction

During its 108th meeting, the WG11 Systems subgroup has issued the CD of the HEVC still image file format. The CD features a tool to describe an HEVC image as a collection of other HEVC tiles, or potentially of other HEVC independently coded image, as presented in contribution m33221.
This contribution addresses remaining issues in the current text and proposes to simplify the design and to extend it to include other image metadata descriptions as suggested in m33273.

2 Problematic

In the current CD text, the relationship between the HEVC tile and its location in the grid description is established by an index carried in a binary blob on 32 bits at the beginning of the item data (one dedicated item extend for this).
This has several drawbacks:

- it breaks definition of the HEVC tile item: media format shall be a set of NALUs as defined in 14496-15, according to the CD text
- it makes the editing painful (add/remove the blob)
- it breaks describing a tiling grid made of independent tiles (HEVC still), since the blob cannot be inserted without breaking the HEVC compliance.

This blob mechanism was introduced to avoid any change in the base specification 14496-12, but obviously the resulting text is not really practical.
One approach would be to carry the index of the item in the grid in a dedicated item (ItemIndexDataBlock was suggested), and have one item relation from the tile image to the index item, and one item relation between the tile index and the grid description. Although possible, this increases the number of items and item relation entries, and the benefit of the double indirection in terms of byte saved is not obvious.
We therefore suggest removing the complete grid description item and have only the spatial location of the image described (x/y/w/h) in the description item, with an item relation between the image (tile) and the spatial description. This makes the solution compatible with the original design (being able to describe an image as a collection of potentially independently coded HEVC images).

The technical tool is still the same (addressing tiles and items collection), only the syntax needs to be reworked and is proposed in section 4 below.

3 Image enhancements

14496-12 provides generic tools to indicate image cropping, sample aspect ratio modification and color adjustments through CleanApertureBox (clap), PixelAspectRatioBox (pasp) and ColourInformationBox (colr). These properties can obviously be applied to a single image, however their storage is currently only possible for media tracks. We therefore suggest the possibility to indicate these in the same item as the spatial description, as described by the syntax below.
Finally, we have identified use cases in images collection where a spatial location is not enough to describe the image, typically when images overlap. We suggest introducing a layer identification, as done in the media track header as described below.

4 Proposed solution

We suggest replacing section 7.9 by the following

"7.9 Simple Metadata support

This section specifies how to describe image items using box structures as defined in 14496-12, such as clean aperture, color information or sample aspect ratio, or boxes defined in this section, such as spatial relation.

7.9.1 Image Spatial Relation

The ImageSpatialRelationBox 'isre' gives the ability to describe the relative position of an image with other images in an image collection. It provides a subset of the functionalities of the matrix in the track header of a media. Coordinates in the ImageSpatialRelationBox are expressed on a square grid giving the author's intended display size of the collection; these units may or may not match the coded size of the image. The intended display size is defined by:
- Horizontally: the maximum value of (horizontal_display_offset + display_width) for all 'isre' boxes
- Vertically: the maximum value of (vertical_display_offset + display_height) for all 'isre' boxes

When some images do not have any 'isre' associated while other images in the file have 'isre' associated, the default images without any 'isre' shall be treated as if their horizontal and vertical offsets are 0, their display size is the intended display size and their layer is 0.

The ImageSpatialRelationBox indicates the relative spatial position of images after any cropping or sample aspect ratio has been applied to the images.

Definition:

Box Type: 'isre'
Container: Simple image meta-data item ('simd')

Mandatory: No
Quantity: Zero or one per item

Syntax:

```
aligned(8) class ImageSpatialRelationBox
extends FullBox('isre, version = 0, 0) {
  unsigned int(32) horizontal_display_offset;
  unsigned int(32) vertical_display_offset;
  unsigned int(32) display_width;
  unsigned int(32) display_height;
  int(16) layer;
}
```

Semantics:

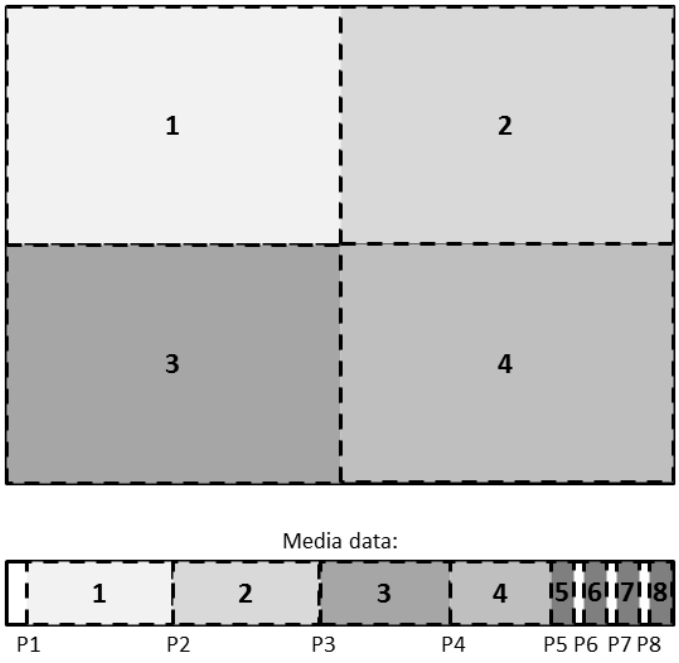
`horizontal_display_offset` specifies the horizontal offset of the image.
`vertical_display_offset` specifies the vertical offset of the image.
`display_width` specifies the width of the image.
`display_height` specifies the height of the image.
`layer` specifies the front-to-back ordering of the image; images with lower numbers are closer to the viewer. 0 is the normal value, and -1 would be in front of layer 0, and so on.

7.9.2 Simple Meta-data Description

An image item, whether a tile or a complete HEVC image, can have spatial relation, color or cropping information associated to it. The information are described through boxes contained in an item of type 'simd':

```
aligned(8) class SimpleImageMetaData {
 CleanApertureBox clap; // optional
 PixelAspectRatioBox pasp; // optional
 ColourInformationBox colour; // optional
 ImageSpatialRelationBox location; // optional
 Box other_boxes[]; // optional
}
```

The image item is associated to a SimpleImageMetaData item through an item reference of type 'simr' (see figure below, note that for clarity reason the hvcC item is omitted.). Depending on use cases, it is possible to have several items sharing the same metadata, for example when the same cropping is to be applied to all images.
It is possible for an image item to have multiple 'simr' references to different SimpleImageMetaData, for example when cropping is shared among images but not spatial information.
"



ftyp box: major-brand = 'hevc',
compatible-brands = 'hevc'
meta box: (container)
handler box: hdlr = 'hvc1'

Item information:
// No full image coded
Item type = 'hvc1', itemID=1 (sub image 1)
Item type = 'hvc1', itemID=2 (sub image 2)
Item type = 'hvc1', itemID=3 (sub image 3)
Item type = 'hvc1', itemID=4 (sub image 4)
Item type = '**simd**' itemID=5 (sub image descriptor)
Item type = '**simd**' itemID=6 (sub image descriptor)
Item type = '**simd**' itemID=7 (sub image descriptor)
Item type = '**simd**' itemID=8 (sub image descriptor)

Item Reference:
type='**simr**', fromID=1, toID=5;
type='**simr**', fromID=2, toID=6;
type='**simr**', fromID=3, toID=7;
type='**simr**', fromID=4, toID=8;

Item Location:
itemID = 1, extent_count = 1, extent_offset = P1, extent_length = L1;
itemID = 2, extent_count = 1, extent_offset = P2, extent_length = L2;
itemID = 3, extent_count = 1, extent_offset = P3, extent_length = L3;
itemID = 4, extent_count = 1, extent_offset = P4, extent_length = L4;
itemID = 5, extent_count = 1, extent_offset = P5, extent_length = L5;
itemID = 6, extent_count = 1, extent_offset = P6, extent_length = L6;
itemID = 7, extent_count = 1, extent_offset = P7, extent_length = L7;
itemID = 8 extent_count = 1, extent_offset = P8, extent_length = L8;

5 Editorial fixes

In 5.8 replace:
"HEVC tiles are stored as meta-data items of type 'hvc1'"

with

"HEVC tiles are stored as meta-data items of type 'hvt1'"

Replace
"this reference implicitly indicates that the tile item shares the same HEVCDecoderConfigurationRecord as the 'hbc1' item, and no 'init' reference shall be set on an 'hvc1' item"

with

"this reference implicitly indicates that the tile item shares the same HEVCDecoderConfigurationRecord as the 'hvc1' item, and no 'init' reference shall be set on an 'hvt1' item"

Replace
"Tile spatial information is given through 'tile' item references, as defined in section 7.9"
with
"Tile spatial information may be expressed using the tools defined in section 7.9".

Additionally, the use of the CodingConstraintsBox is unclear. The specification currently says: “The HEVCSampleEntry shall be used as specified in 14496-15. The coding constraints box shall be present in the HEVCSampleEntry”. Given that Part 15 does not define the CodingConstraintsBox, the HEVCSampleEntry cannot be used as specified in Part 15. Additionally, this box should be optional.

We propose to clarify as follows:

“The HEVCSampleEntry shall be used as specified in 14496-15, with an optional CodingConstraintsBox present in the HEVCSampleEntry extra_boxes array:

Assuming that the syntax of the HEVCSampleEntry in Part 15 is fixed as follows:

```
class HEVCSampleEntry() extends VisualSampleEntry (‘hvc1’ or 'hev1'){
   HEVCConfigurationBox config;
   MPEG4BitRateBox ();              // optional
   MPEG4ExtensionDescriptorsBox (); // optional
   Box   extra_boxes[];           // optional
}
”
```

6 Conclusion

We suggest the File Format experts to include the proposed modification in the DIS text of the HEVC still image file format.