

**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/M35083
October 2014, Strasbourg, France**

Source Canon Research Centre France , Telecom ParisTech
Status For consideration at the 110th MPEG Meeting
Title Image composition with HEVC Still Image
Author Franck Denoual, Frédéric Mazé, Eric Nassor (Canon), Jean Le Feuvre and Cyril Concolato (Telecom Paris Tech)

1 Introduction

During its 109th meeting, the WG11 Systems subgroup has issued the DIS of the HEVC Still Image file format. This file format features a descriptor for image composition and cropping. This contribution addresses remaining issues in the current text to expose resulting composition image as the primary item as suggested in m34256.

2 Problematic

In the current DIS text, the ISOBMFFMetaData item ('simd') describes spatial relation, color or cropping information associated to an image, image being a tile or a complete image. In particular, the ImageSpatialRelationBox ('isre') provides position, size and layering information that is useful to organize sub-images into a composite bigger image.

While these tools work for tiled image description (example 1) or composition description without associating an item to the resulting composite image (example 2), we found an issue when an author would like to declare the composite image as the primary item in the file (example 3).

2.1 Composition of HEVC tiles (example 1)

The example below illustrates the description of 4 tiles from an HEVC coded image. The HEVC image contains 4 tiles covering the whole picture as shown on the figure below. The full image is the main item and each tile is described as an image item with position and size respectively to the full image.

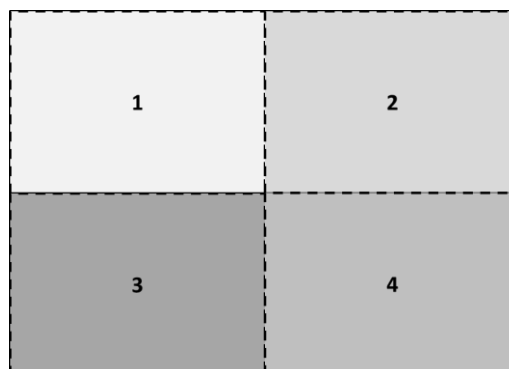


Figure 1: Tiled HEVC picture.

Below is provided the description of such tile-based composition with current Still Image descriptors.

```
ftyp box: major-brand = 'mif1', compatible-brands = 'heic'
meta box: (container)
handler box: hdlr = 'hvc1' primary item: itemID = 1;
Item Information Entries:
item_type = 'hvc1', itemID=1, // composite image, the HEVC encoded one
item_type = 'hvt1', itemID=2, item_protection_index = 0... // sub-image
item_type = 'hvt1', itemID=3, item_protection_index = 0... // sub-image
item_type = 'hvt1', itemID=4, item_protection_index = 0... // sub-image
item_type = 'hvt1', itemID=5, item_protection_index = 0... // sub-image
item_type = 'simd' itemID=6 (sub-image descriptor)... (isre: 0, 0, W, H, 0)
item_type = 'simd' itemID=7 (sub-image descriptor)... (isre: W, 0, W, H, 0)
item_type = 'simd' itemID=8 (sub-image descriptor)... (isre: 0, H, W, H, 0)
item_type = 'simd' itemID=9 (sub-image descriptor)... (isre: W, H, W, H, 0)
item_type = 'hvcC', itemID=10 (decoder config record)
item_type = 'simd' itemID=11 (composition descriptor)... (isre: 0, 0, 2W, 2H, 0)
Item Reference Entries:
type= 'init', from_item_ID=10, to_item_ID=1... (assuming reverse order)
type='base', from_item_ID=1, ref_count=4, to_item_ID=2, 3, 4, 5 (to each tile)
// declare sub-images as tiles of the full image
type= 'tbas', from_item_ID=2, to_item_ID=1...
type= 'tbas', from_item_ID=3, to_item_ID=1...
type= 'tbas', from_item_ID=4, to_item_ID=1...
type= 'tbas', from_item_ID=5, to_item_ID=1...
// providing positions and sizes
type= 'simr', from_item_ID=6, to_item_ID=2
type= 'simr', from_item_ID=7, to_item_ID=3
type= 'simr', from_item_ID=8, to_item_ID=4
type= 'simr', from_item_ID=9, to_item_ID=5
type= 'simr', from_item_ID=11, to_item_ID=1
...
Item Location:
itemID = 1, extent_count = 4, // full image is composed of 4 tiles
extent_offset = P2, extent_length = L2... // data for tile 1
extent_offset = P3, extent_length = L3... // data for tile 2
extent_offset = P4, extent_length = L4... // data for tile 3
extent_offset = P5, extent_length = L5... // data for tile 4
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;
itemID = 9, extent_count = 1, extent_offset = P9, extent_length = L9;
itemID = 10, extent_count = 1, extent_offset = P0, extent_length = L0;
itemID = 11, extent_count = 1, extent_offset = P11, extent_length = L11;
Media data box:
1 HEVC Decoder Configuration Record ('hvcC' at offset P0)
5 1 HEVC Image (with 4 tiles at file offsets P2, P3, P4, P5)
5 simple image metadata (at file offsets P6, P7, P8, P9, P11)
```

2.2 Composition of independent images (example 2)

The figure below illustrates a composition of independent HEVC images into a bigger image. This bigger image is not exposed as an item in the file format. Only described are the 4 sub images with each associated with spatial relation information (isre).

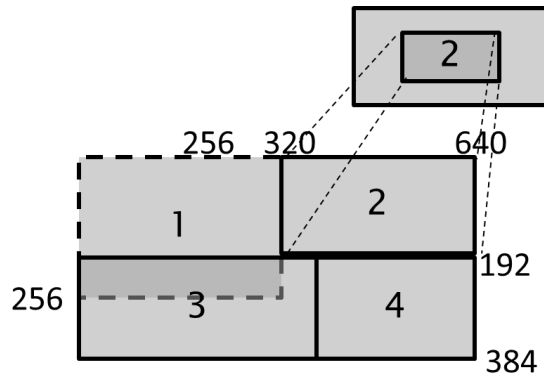


Figure 2: Tiled HEVC picture.

```

ftyp box: major-brand = 'mif1', compatible-brands = 'heic'
meta box: (container)
handler box: hdlr = 'hvc1' // no primary item provided
Item Information Entries:
item_type = 'hvc1', itemID=1 (sub-image)
item_type = 'hvc1', itemID=2 (sub-image)
item_type = 'hvc1', itemID=3 (sub-image)
item_type = 'hvc1', itemID=4 (sub-image)
item_type='simd' itemID=5 (sub-image 1 descriptor)
item_type='simd' itemID=6 (sub-image 2 descriptor)
item_type='simd' itemID=7 (sub-image 3 descriptor)
item_type='simd' itemID=8 (sub-image 4 descriptor)
item_type = 'hvcC', item_ID=9 (decoder config, shared among all sub-images)

```

Item Reference:

```

type='simr' from_item_ID=5, to_item_ID=1
type='simr' from_item_ID=6, to_item_ID=2
type='simr' from_item_ID=7, to_item_ID=3
type='simr' from_item_ID=8, to_item_ID=4
type='init', from_item_ID=9, reference_count=4, to_item_ID=1,2,3,4 (reverse order)

```

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;
itemID = 9, extent_count = 1, extent_offset = P0, extent_length = L0;

```

Media data box:

```

1 HEVC Decoder Configuration Record ('hvcC' at offset P0)
4 HEVC Images (at file offsets P1, P2, P3, P4)
4 Simple Image Metadata items (at file offsets P5, P6, P7, P8)

```

```

'simd' { // for item 5, subimg 1
  'isre': 0, 0, 320, 256, 0 (back)
}

```

```

'simd' { // for item 6, subimg 2
  'clap': 320, 1, 192, 1, 0,1 0,1, (center of img2)
  'isre': 320, 0, 320, 192, -1 (front)
}

```

```

'simd' { // for item 7, subimg 3
  'isre': 0, 192, 384, 192, -1
}

```

```

'simd' { // for item 8, subimg 4
  'isre': 384, 192, 256, 192, -1
}

```

Example description in Still Image File Format

From the above description, in order to determine the size of the composite image, the parser has to interpret all 'isre' boxes and deduce the size of the composite image as the maximum value of (isre::display_offsets + isre::display_sizes). Moreover there is no primary item.

2.3 Exposing the composition of sub images (example 3)

Considering the previous example, when the author would like to expose the composite image as the primary item, it could be useful, to avoid parsing work, to describe the size of the composite image and for player to directly render the composition image.

The differences, compared to example 2 are highlighted in yellow.

```
ftyp box: major-brand = 'hevc', compatible-brands = 'mif1'
meta box: (container)
handler box: hdlr = 'hevc' primary item box: item_ID = 10
Item Information Entries:
item_type = 'hvc1', itemID=1 (sub-image)
item_type = 'hvc1', itemID=2 (sub-image)
item_type = 'hvc1', itemID=3 (sub-image)
item_type = 'hvc1', itemID=4 (sub-image)
item_type='simd' itemID=5 (sub-image 1 descriptor, same isre as in example 2)
item_type='simd' itemID=6 (sub-image 2 descriptor, same isre as in example 2)
item_type='simd' itemID=7 (sub-image 3 descriptor, same isre as in example 2)
item_type='simd' itemID=8 (sub-image 4 descriptor, same isre as in example 2)
item_type = 'hvcC', item_ID=9 (decoder config, shared among all sub-images)
item_type = '????', itemID=10 (the composite image) // Issue with the type here
item_type='simd' itemID=11 (full image, isre = {0, 0, 640, 384, 0}) // the referential for composition
```

Item Reference:

```
type='simr' from_item_ID=5, to_item_ID=1
type='simr' from_item_ID=6, to_item_ID=2
type='simr' from_item_ID=7, to_item_ID=3
type='simr' from_item_ID=8, to_item_ID=4
type='init', from_item_ID=9, reference_count=4, to_item_ID=1, 2, 3, 4 (reverse order)
type='simr' from_item_ID=11, to_item_ID=10
```

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;
itemID = 9, extent_count = 1, extent_offset = P0, extent_length = L0;
itemID = 10, extent_count = 4, // full composed from 4 sub-images extents P1, P2, P3, P4
```

Media data box:

```
1 HEVC Decoder Configuration Record ('hvcC' at offset P0)
4 HEVC Images (at file offsets P1, P2, P3, P4)
4 Simple Image Metadata items (at file offsets P5, P6, P7, P8)
```

For that, a 'simd' can also be associated to the composite image, optionally with reference to each sub-image. However a problem remains regarding the type of the corresponding image item. Declaring it as an 'hvc1' item requires the presence of an 'hvcC' item which is not available for this composite image.

We then propose a new item type to be able to describe composite image (see proposal in next section).

3 Proposals

3.1 New item type for composite images

We propose the creation of a new item type for pictures resulting from composition of HEVC images or tiles.

For example: ‘hvco’ for **HEVC Composition** image. Such image, in opposition to ‘hvc1’, does not require to be linked to an ‘hvcC’ item because not always available, for example when composing from different HEVC sub-images.

To describe this new kind of image item, we propose a new sub-section in section 6.8, for example 6.8.6:

In section 6.8 add a sub-section:

Section 6.8.6: Composite images

“Composite still images that are related to a primary still image shall also be stored as items. An HEVC composite image uses the item type ‘hvco’ for HEVC Composition image.

Optionally, the image item resulting from composition is linked to a ‘simd’ (spatial relation and other metadata descriptor) so that a parser can easily have the display sizes of the image (without having to parse all the simd of the sub-images involved in the composition).

Optionally, when composition is done with sub-images sharing the same ‘hvcC’ information, the item for the composite image can also be linked to this ‘hvcC’ item. The link to ‘hvcC’ item is not mandatory for images with ‘hvco’ image items.

Optionally, a composite image can be optionally linked to the sub-images through a ‘tile’ or ‘base’ reference image item, ‘tile’ indicating the referenced sub-images are spatial parts of the composite image and ‘base’ indicating that the composite image is derived from the referenced sub-images.”

3.2 Declaration of initialization data for HEVC images (section 6.8.2.1)

In order to improve readability and compactness of initialization data declarations, we propose to Replace:

“Images must be linked to an item of type ‘hvcC’ by an item reference of type ‘init’. Items of type ‘hvcC’ must contain the HEVC decoder configuration record defined in the following sections. It is structurally identical to the HEVC Decoder Configuration Record defined in ISO/IEC 14496-15.”

With:

“Image items must be linked to an item of type ‘hvcC’ by an item reference of type ‘init’ **from the ‘hvcC’ item to the list of image items using it**. Items of type ‘hvcC’ must contain the HEVC decoder configuration record defined in the following sections. It is structurally identical to the HEVC Decoder Configuration Record defined in ISO/IEC 14496-15”

Indeed, reversing the order of ‘init’ information declaration as suggested above would have the following benefits:

- Better readability by reading the declaration of the reference as : item X “from_item_ID” is ‘init’ data for the list of provided “to_item_ID”.

- Exploit the `reference_count` parameter to declare only one `ItemReference` for all `hvcC` items shared among images or sub-images (as done in examples 1 and 2).

4 Conclusion

We suggest the File Format experts to include the proposed modification (new item type for composed images, compaction of initialization data declaration) in the DIS text of the HEVC still image file format.