Source	**Telecom ParisTech, Canon CRF**
Status	**For consideration at the 106th MPEG meeting**
Title	**Inputs to WD of 14496-15 AMD2**
Author	Jean Le Feuvre, Cyril Concolato, Franck Denoual, Frédéric Mazé, Eric Nassor

1   Introduction

During the Vienna meeting, a solution for describing HEVC tiling in the ISOBMFF was proposed as a working draft. This contribution investigates whether the tools in the WD also fit to describe scalable tiling as can be used in SHVC. It does not propose a complete SHVC support but rather updates the existing tools in the working draft to ensure their compatibility with hierarchically tiled descriptions.

2   Problems

2.1   Tiling and scalability

Using the existing tools and the ability to map each NALU to a given tile, this section shows what is missing in the current WD to map each NALU to a layer or a tile in a layer.

2.1.1   Example 1

Consider the setup of Figure 1 below:
- one layer of tiled HEVC in HD: High Definition
- one spatial layer of tiled SHVC at twice the resolution (UHD: Ultra High Definition),
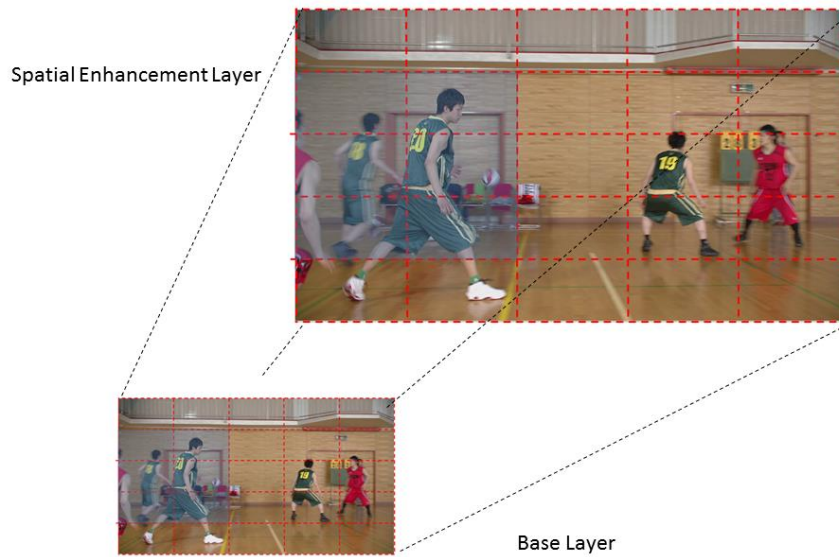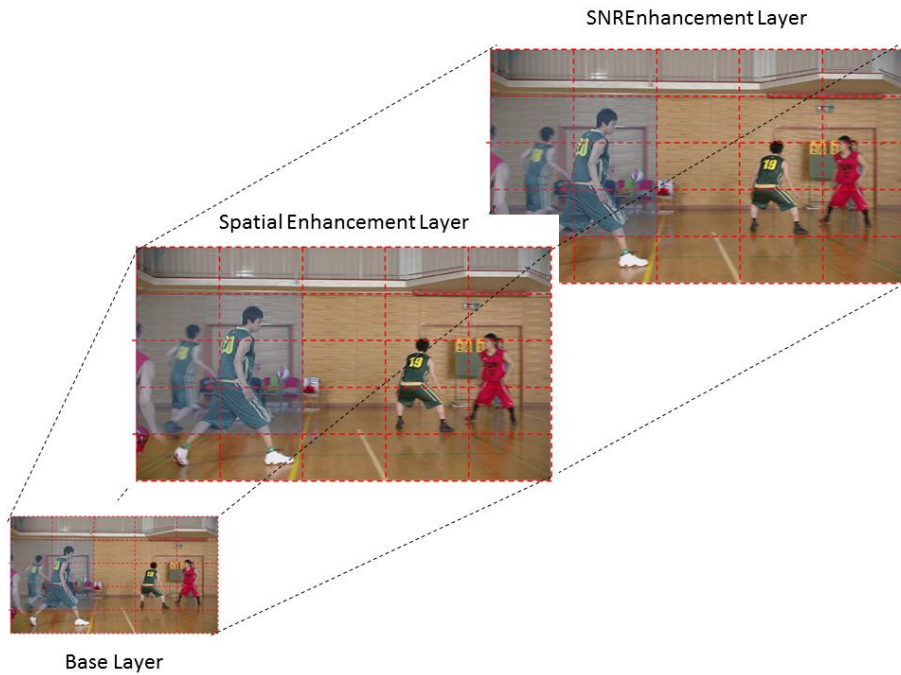
Figure 1: SHVC with 1 tile-based scalability layer

To describe this in a single track, we need:
- set of tiles group description for base and enhancement layers
  - if in the same track but in different luma spaces, we would need to add resolution info for each tile since 0x0x320x240@HD != 0x0x320x240@QFHD
  - a single set of tiles in unified luma space is then required, but which one should be picked ?

- declare dependencies between layers:
  - we need to add dependencyID to the base tile (TileRegionGroupEntry) for SHVC
  - If one tile from enhancement depends on several tiles from lower layer, tileSetRegion can be used

2.1.2 Example 2

Consider the setup of Figure 2:
- one layer of tiled HEVC
- one spatial layer of tiled SHVC at twice the resolution
- one SNR layer of tiled SHVC at twice the resolution on top of the spatial tiled SHVC layer.

Figure 2: SHVC with 2 tile-based scalability layers

To describe this configuration, we also need to indicate dependencies between tiled SNR and tiled spatial.

Once we have all tiling and dependencies, we need to map NALUs to this. If each tile@layer is given a unique groupID, the current draft solution works.

2.1.3 Example 3

Consider the setup of Figure 3 below:
- one layer of HEVC
- one spatial layer of SHVC at twice the resolution
- one SNR layer of SHVC at twice the resolution

Figure 3: SHVC with 2 tile-based scalability layers

In this case, no more tiling is used. We again need to indicate dependencies between SNR and spatial layers. If each layer is defined through a visual sample group description with its own groupID, the technologies in the draft also works.

2.1.4 Design

We suggest defining HEVC layers with groupIDs and unifying identifier namespace for tiles, tileSets and layers. By adding dependency IDs to the different structures, this enables:
- description of dependencies between tiled layers
- description of dependencies between non-tiled layers
- description of dependencies between a non-tiled enhancement to tiled base
- description of dependencies between a tiled enhancement to non-tiled base

If layers or tiles are stored in different tracks, as would be the case for some DASH-based solution, the proposed approach still works, if we define that the namespace for groupIDs is the base layer, e.g. that there shall not be the same value of groupID used in tracks referring to the same base (in most cases, groupID namespace's scope would be the moov, since only one base layer of video would be present in a typical file).

2.1.5 Final optimization

The last problem may come from variable NALU patterns from one sample to another, in which case the NALU Map Entry as defined in the WD is not sufficient. Indeed, it works fine only when NALU mapping to groups is constant between samples. One possibility could be to use Aggregators but they are SVC and MVC specific, are not available for HEVC and needs bitstream rewriting before extraction. We propose reusing the new semantics of the SubSampleInformation box introduced for HEVC file format: by assigning the reserved field to the groupID of the NAL unit, dynamic NALU maps can easily be described. Furthermore, by combining this new tool with the existing NALUMapEntry and the defaultSampleGroup

introduced during the 105th MPEG meeting, the solution provides a simple way to describe temporary modifications to a default "constant" (i.e. regular) NALU pattern.

3 Proposal

Change TileRegionGroupEntry() to

```
class TileRegionGroupEntry() extends VisualSampleGroupEntry
('trif')
{
 unsigned int(32) groupID;
 unsigned int(32) dependentGroupID;
 unsigned int(32) layerGroupID;
 unsigned int(2)  independent;
 unsigned int(6)  reserved=0;
 unsigned int(16) horizontal_offset;
 unsigned int(16) vertical_offset;
 unsigned int(16) region_width;
 unsigned int(16) region_height;
}
```

and add to semantics:

dependentGroupID: gives the identifier of a Tile (as defined by a TileRegionGroupEntry), of a Tile Set (as defined by a TileSetGroupEntry) or of an HEVC layer (as defined by HEVCLayerDefinitionBox) that this tile depends on. If set to 0, dependencies are derived from the track reference box.

layerGroupID: gives the identifier of the HEVC layer (as defined by HEVCLayerDefinitionBox) this tile belongs to. If set to 0, dependencies are derived from the track reference box.

region_width, region_height, vertical_offset and horizontal offset are given in luma samples of the layer identified by layerGroupID if not 0, or in luma samples of the frame as indicated in the visual sample entry of the stsd box.

Change the TileSetGroupEntry() semantics, replacing

```
class TileSetGroupEntry() extends VisualSampleGroupEntry
('tsif') {
  unsigned int(16) groupID;
```

by
```
class TileSetGroupEntry() extends VisualSampleGroupEntry
('tsif') {
  unsigned int(32) groupID;
```

and change the definition of dependencyTileGroupID to

dependencyTileGroupID gives the identifier of a Tile (as defined by a TileRegionGroupEntry) or a Tile Set (as defined by a TileSetGroupEntry) that this tile set depends on. If set to 0, dependencies are derived from the track reference box.

Note: dependencies from tile sets to layers are derived from the dependencies between the tile set and its tile regions, each tile region indicating its dependency layer or sub-layer tile region/set.

In definition of NALUMapEntry replace

```
  unsigned int(16) groupID;
```

by

```
  unsigned int(32) groupID;
```

and change the semantics of groupID as follows:

groupID indicates the corresponding scalable, multiview, tile, tile set or HEVC layer group entry, as indicated in the sample group descriptions. If 0, no group is associated to these identified NALUs.

Create a HEVCLayerDefinitionBox :

```
class HEVCLayerDefinitionBox() extends VisualSampleGroupEntry
('hvcl')
{
    unsigned int(32)  groupID;
    unsigned int(32)  dependentGroupID;
    unsigned int(16)  visualWidth;
    unsigned int(16)  visualHeight;
    //to be completed once SHVC high-level syntax is frozen
    // for example with profile_tier_level info, frame rate…
}
```

Semantics:

groupID: is a unique identifier for the layer described by this group. Value 0 is reserved for special use in the 'nalm' box.

dependentGroupID: indicates the groupID of an HEVC layer (as defined by a HEVCLayerDefinitionBox) this layer depends on. If 0, dependencies are derived from the track reference box. This is for example the case when an SHVC bitstream enhances an AVC|H264 track.

visualWidth gives the value of the width of the coded picture/view in luma samples
visualHeight gives the value of the height of the coded picture/view in luma samples

Add a new flag for the SubSampleInformationBox in HEVC (8.4.8):

Replace

```
        } else if (flags == 3 || flags == 4) {
```

```
        unsigned int(2) vcl_idc;
        unsigned int(30) reserved = 0;
    }
```

with

```
    } else if (flags == 3 || flags == 4) {
        unsigned int(2) vcl_idc;
        unsigned int(30) reserved = 0;


    } else if (flags == 5) {

        unsigned int(32) groupID;

    }
```

And add the following sentence:

"

5:    groupID based sub-samples. A sub-sample is mapped to a HEVC layer, tile or tileSet
      identified by its groupID. If a NALUMapEntry sample group description is present
      with a default `default_sample_description_index`, the default value is ignored
      (e.g. the SubSampleInformationBox overrides the definition present in the
      NALUMapEntry). If groupID is 0, no group is associated with this NALU or group
      of NALUs.

"

And change the sentence in 8.5.3:

"Each of these groups is assigned a unique identifier, which can be used to associate a NALU to
a group. Tile and Tile Sets share the same namespace for groupID (ie, there shall not be two tile
and tile sets with the same groupID)."

To

"Each of these groups is assigned a unique identifier, which can be used to associate a NALU to
a group. Tile, Tile Sets and HEVC layers share the same namespace for groupID, scoped by the
base HEVC layer; i.e. there shall not be two tile, tile sets or HEVC layer with the same groupID
in any tracks having the same base layer."

4   Conclusion

We suggest including the updates introduced in this contribution to the WD of 14496-15 AMD2.