

**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/M36558
June 2015, Warsaw, Poland**

Source **Telecom ParisTech, Canon Research Centre France**
Status **For consideration at the 112th MPEG Meeting**
Title **Fixing the oinf box in the Layered HEVC file format**
Author Cyril Concolato, Jean Le Feuvre, Franck Denoual, Frédéric Mazé, Naël Ouedraogo,
Eric Nassor

1 Introduction

The current text requires only one VPS per set of related base and layer tracks. This is unnecessarily restrictive as there are several use cases where this might be useful. This contribution reviews some use cases and proposes a modification of the existing draft text.

2 Analysis

Some use cases are:

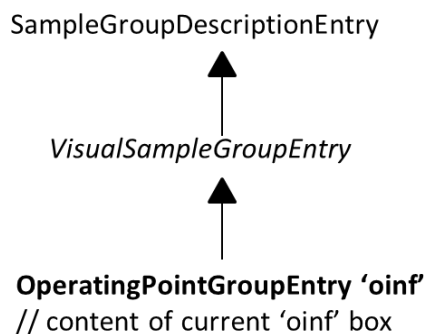
- Storage of spliced HEVC bitstreams with different VPS.
- Concatenating files obeying to the same profile constraints. Such concatenation could follow different strategies:
 - o Samples in tracks containing the base layers from both files could simply be concatenated, leading to multiple VPS/SPS/PPS in different sample entries or in larger hvcC NALU arrays.
 - o Sample from non-base layers could be concatenated by inspecting tracks one by one and concatenating them with sample from tracks corresponding to layers with similar constraints, if any, possibly leading to a change in time of the (single) layer id associated to the track;
 - o Sample from the non-base layer tracks of the second file could be added to new sets of tracks, shifted in time to maintain sync with the concatenated base track. The latter approach is complex and might not be preferable. In such scenario, it might be useful to allow for track-layer configuration changes.
- Streaming a long-running fragmented MP4 file with changes in layer configuration. In such case, a content producer defined two scalable layers for two classes of devices (e.g. SD, HD). After a period of time, a new class of device (e.g. UHD) is available. It should be possible to reuse the two tracks and to add an extra layer. In a configuration where movie fragments are used, the operation of removing fragments should not lead to information loss. If the layer configuration changed during the streaming this should be captured.

Allowing information to change over time in the ISOBMFF can mean: using a dedicated track; using sample entries; or using sample groups. Using dedicated tracks allows sharing information across tracks. However, the precedent of AVC Parameter Sets tracks (signaled with 'avcp') does not vote for this option. Using sample entries in regular tracks (ie. for HEVC the 'hvcC' box) does not allow for modification after the "moov" has been produced. In a context of streaming layered HEVC using movie fragments, this does not seem acceptable. The chosen approach is therefore to use sample groups.

3 Proposal

The layer configuration is currently provided by the 'oinf' and 'tcon' mandatory boxes.

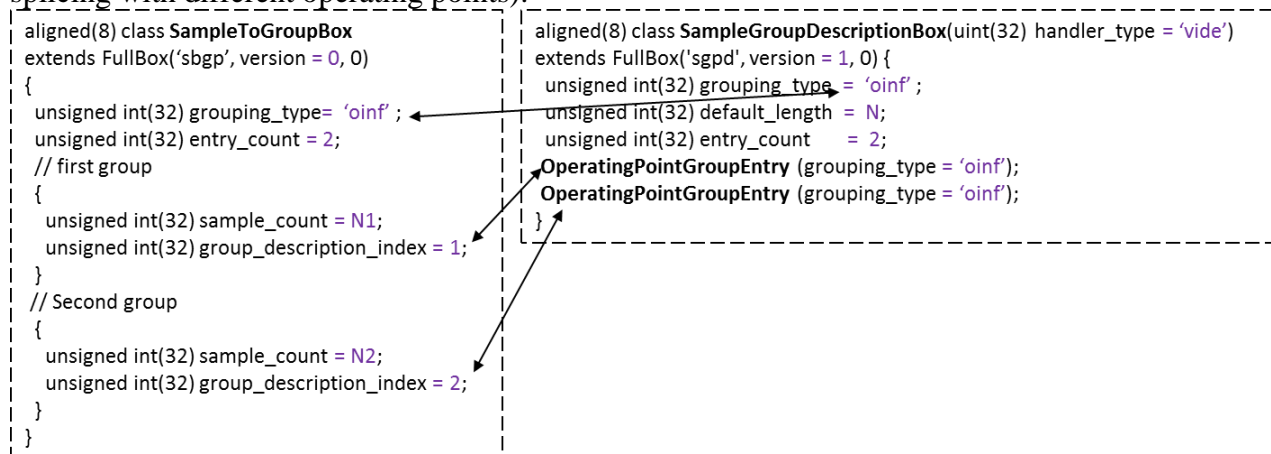
We propose to move those boxes out of the 'minf' box and to make them inherit from the VisualSampleGroupEntry class as below.



```

class OperatingPointGroupEntry extends VisualSampleGroupEntry ('oinf') {
    // content of 'oinf' box
}
  
```

Usage of this new kind of VisualSampleGroupEntry is the same as for any other sample group description as illustrated below (example shows two sample groups resulting from stream splicing with different operating points):



The default sample grouping mechanism can also be applied as shown below.

```

aligned(8) class SampleGroupDescriptionBox(uint(32) handler_type = 'vide')
extends FullBox('sgpd', version = 2, 0) {
    unsigned int(32) grouping_type = 'oinf';
    unsigned int(32) default_length = N;
    unsigned int(32) default_sample_description_index = 1;
    unsigned int(32) entry_count = 1;
    OperatingPointGroupEntry (grouping_type = 'oinf');
}

```

As well, since layer organization inside a track can vary along time, the 'tcon' box should follow the same transformation, i.e. moving from 'minf' box to a new grouping_type and new VisualSampleGroupEntry, for example LayerInfoGroupEntry 'tcon' as follows:

```

class LayerInfoGroupEntry extends VisualSampleGroupEntry ('tcon') {
    unsigned int (2) reserved;
    unsigned int (6) num_layers_in_track;
    for (i=0; i<num_layers_in_track; i++) {
        unsigned int (4) reserved;
        unsigned int (6) layer_id;
        unsigned int (3) min_sub_layer_id;
        unsigned int (3) max_sub_layer_id;
    }
}

```

For a track with constant layer organization, the default sample grouping could be used as follows (definition of a sample group description box with 'tcon' grouping_type):

```

aligned(8) class SampleGroupDescriptionBox(uint(32) handler_type = 'vide')
extends FullBox('sgpd', version = 2, 0) {
    unsigned int(32) grouping_type = 'tcon';
    unsigned int(32) default_length = N;
    unsigned int(32) default_sample_description_index = 1;
    unsigned int(32) entry_count = 1;
    LayerInfoGroupEntry (grouping_type = 'tcon');
}

```

In case the layer organization varies along time, a sample group box can be defined to map samples to their appropriate layer to track organization:

<pre> aligned(8) class SampleToGroupBox extends FullBox('sbgp', version = 0, 0) { unsigned int(32) grouping_type= 'tcon'; unsigned int(32) entry_count = 2; // first group { unsigned int(32) sample_count = N1; unsigned int(32) group_description_index = 1; } // Second group { unsigned int(32) sample_count = N2; unsigned int(32) group_description_index = 2; } } </pre>	<pre> aligned(8) class SampleGroupDescriptionBox(uint(32) handler_type = 'vide') extends FullBox('sgpd', version = 1, 0) { unsigned int(32) grouping_type = 'tcon'; unsigned int(32) default_length = N; unsigned int(32) entry_count = 2; LayerInfoGroupEntry(grouping_type = 'tcon'); LayerInfoGroupEntry(grouping_type = 'tcon'); } </pre>
---	---

The current approach uses the 'oref' reference track type to signal which track contains the 'oinf' information. There are 2 options here:

- discard the 'oref' type and mandate that the 'oinf' information be in the base track, as indicated by the 'sbas' track reference. In case of hybrid coding (AVC base track + HEVC enhancement track), this means that the sample group 'oinf' would be allowed on an AVC track.
- if the first option is not possible, keep the 'oref' to signal which track contains the sample group and sample group description boxes indicating the 'oinf' information.

4 Recommendation

We propose to adopt the proposal in the FDIS of 14496-15 Amendment on Carriage of Layered HEVC.